BERGISCHE
UNIVERSITÄT
WUPPERTAL

Technische Hochschule
Ingolstadt

Accelerated Virtual Evaluation of Restraint System Performance for Vehicle Occupants with Varying Body Shapes

Dissertation to obtain a Doctoral Degree
in the
School of Mechanical Engineering and Safety Engineering
of the
University of Wuppertal

Submitted by
Franz Plaschkies

Wuppertal 2025
Oral Exam on 7th October 2025

Accelerated Virtual Evaluation of Restraint System Performance for Vehicle Occupants with Varying Body Shapes

Franz Plaschkies

Dissertation, University of Wuppertal, School of Mechanical Engineering and Safety Engineering, Chair for Optimization of Mechanical Structures, January 2025

Abstract

This thesis describes the application of machine learning in the context of development and assessment methods in the domain of safety systems for occupant safety. Supervised learning methods are applied to generate rapid predictions for the results of finite element simulations in the domain of the assessment of occupant safety systems. The key concept revolves around the question of how a metamodel could be designed without explicit characterization of the vehicle. The base machine learning architecture uses the result of finite element simulation as a reference to predict the result of a finite element simulation in the same vehicle configuration but different occupant anthropometrics. The input for the metamodel is of the type multivariate time-series. Different strategies to deal with this data type are compared. Starting from a detailed finite element model of the Honda Accord 2014, using an incremental and structured approach, a simplified version of the finite element model is derived. Each simplification step is reasoned and evaluated regarding the loss of the model quality in relation to the gained improvements in computation time. Databases featuring different designs of experiments and the anthropomorphic test device Hybrid III and the human body model VIRTHUMAN are generated. The influence of different database sizes and an adaptive querying strategy is studied. Transfer learning approaches as a method to deal with limited data are discussed as well. The homogeneous transfer between Hybrid III and VIRTHUMAN and the direct cross-domain prediction between them is investigated.

Keywords

Passive Safety, Occupant Protection, Finite Element Analysis, Machine Learning, Accelerated Data Generation

Accelerated Virtual Evaluation of Restraint System Performance for Vehicle Occupants with Varying Body Shapes

Franz Plaschkies

Dissertation, Bergische Universität Wuppertal, Fakultät für Maschinenbau und Sicherheitstechnik, Lehrstuhl für Optimierung mechanischer Strukturen, Januar 2025

Kurzfassung

Diese Dissertation beschreibt die Anwendung von maschinellem Lernen im Rahmen von Entwicklungs- und Bewertungsmethoden im Bereich der Sicherheitssysteme für den Insassenschutz. Methoden des überwachten Lernens werden angewandt, um schnelle Vorhersagen für die Ergebnisse von Finite-Elemente-Simulationen im Bereich der Bewertung von Insassenschutzsystemen zu generieren. Das Konzept dreht sich um die Frage, wie ein Metamodell ohne explizite Charakterisierung des Fahrzeugs trainiert werden kann. Die zugrundeliegende Architektur des maschinellen Lernens verwendet das Ergebnis einer Referenz-Finite-Elemente-Simulation, um das Ergebnis einer Finite-Elemente-Simulation mit derselben Fahrzeugkonfiguration, aber anderen anthropometrischen Merkmalen der Insassen vorherzusagen. Die Eingabe für das Metamodell ist vom Typ der multivariaten Zeitreihe. Es werden verschiedene Strategien zum Umgang mit dieser Datenart verglichen. Ausgehend von einem detaillierten Finite-Elemente-Modell des Honda Accord 2014 wird mittels eines inkrementellen und strukturierten Ansatzes eine vereinfachte Version des Finite-Elemente-Modells abgeleitet. Jeder Vereinfachungsschritt wird begründet und hinsichtlich des Verlusts an Modellqualität im Verhältnis zu den gewonnenen Verbesserungen bei der Rechenzeit bewertet. Es werden Datenbanken mit verschiedenen Versuchsplänen und dem Crashtest-Dummy Hybrid III sowie dem Menschmodell VIRTHUMAN erstellt. Der Einfluss verschiedener Datenbankgrößen und eine adaptive Datenstrategie werden untersucht. Transfer-Learning-Ansätze als Methode zum Umgang mit begrenzten Daten werden ebenfalls beleuchtet. Der homogene Transfer zwischen Hybrid III und VIRTHUMAN und die direkte bereichsübergreifende Vorhersage zwischen beiden Modellen wird untersucht.

Keywords

Passive Sicherheit, Insassenschutz, Finite-Elemente-Analyse, maschinelles Lernen, beschleunigte Datengenerierung

Acknowledgements

The content of this dissertation was created as part of my position at the CARISSMA Institute C-ISAFE of Technische Hochschule Ingolstadt. First, I thank my supervisor, Prof. Ondřej Vaculín, Ph. D., at Technische Hochschule Ingolstadt for all the encouraging discussions, generous support, and his trust. Second, I thank my supervisor, Prof. Dr.-Ing. Axel Schumacher, at the University of Wuppertal, for his invaluable guidance and enabling the inspiring exchange with other doctorate candidates. Third, I thank Prof. Dr.-Ing. Matthias Kröger from Technische Universität Bergakademie Freiberg for his teachings during my studies, his supervision of my master thesis and being finally third examiner for this thesis. Furthermore, I thank Univ.-Prof. Dr.-Ing. habil. Uwe Janoske from the the University of Wuppertal for being chairman of the examination board. Finally, I thank Univ.-Prof. Dr.-Ing. Fabian Brännström from the the University of Wuppertal for completing the board as reviewer.

I would like to thank all my colleagues at Technische Hochschule Ingolstadt and ARRK Engineering GmbH who made the presented research possible. Thanks to the students who supported this project. The resource provided by Technische Hochschule Ingolstadt made this project possible.

I would like to express my gratitude to the Free State of Bavaria for funding project arTIco – Artificial Intelligence and Correlations, a research project in the funding line Digitalization of the Bavarian Cooperative Research Programme (BayVFP). The project was funded under the grant number DIK-2110-0025//DIK0404/02. In this context, I would like to thank our project partner at IDIADA Fahrzeugtechnik GmbH.

I also appreciate the EU Horizon Europe programme for funding the AWARE2ALL project under the grant agreement number 101076868. Together with colleagues from the ESI Group, the database with VIRTHUMAN was created. In this context, I would especially like to thank Sebastian Müller.

I would like to thank the team from Humanetics Innovative Solutions Inc. The insights from our cooperation and discussions were invaluable inspiration for this research.

Finally, I especially thank my wife, Vishakha, for her tireless proofreading, relentless effort on keeping me focussed, and her continuous cheering. Furthermore, a special thanks goes to my parents, Katharina and Olaf, my family, and friends for their great support throughout my life.

Nuremberg, January 2026 Franz Plaschkies

Contents

Glossary

Activation Function

Maps in an Artificial Neural Network (ANN) the neuron weighted input to an output.

Airbag

An airbag is part of the advanced restrained system. It is a fabric hull, inflated by a gas mixture. Typical examples are the Passenger Airbag (PAB), the Driver Airbag (DAB), and the Curtain Airbag (CAB).

Attribute

Attributes are the observed characteristics of a system. An attribute is a particular characteristic observed for multiple instances.

Biofidelity

The degree to which a human surrogate model fits to a real human.

Bulkhead

The bulkhead (also known as firewall) is the part of the Body In White (BIW), which separates the engine compartment from the passenger compartment.

Channel

The output of a particular sensor over time.

Computer-hardware

The physical components of a computer. Major parts determining the computational capabilities are the Central Processing Unit (CPU), the Graphics Processing Unit (GPU), the persistent and temporary storage.

Convolutional Kernel

The tensor within the convolutional layer which is applied by a convolution to the input.

Convolutional Layer

A layer type of an Artificial Neural Network (ANN), where the input tensor is processed with a kernel by the convolution operation.

Cost-function

The cost-function quantifies how much a metamodel prediction deviates from the ground truth for a group (batch) of instances.

Dense Layer

A layer type of an Artificial Neural Network (ANN), where the input of one neuron is the combination of all neurons of the previous layer.

Development-set

A set of instances meant for the training and validation of a metamodel.

Dropout Layer

A layer type of an Artificial Neural Network (ANN), which sets during the training the output of random neurons of previous layer to zero. The remaining outputs are scaled to maintain the sum over all neurons.

Epoch

An epoch describes the pass of the whole training-set during the training of a metamodel.

Experiment

A particular combination of factors characterizing a system whose response is observed. In this work, one experiment is usually repeated as FE-simulations for different percentiles of human surrogates.

F1-score

Metric-function as Harmonic Mean of Precision and Recall.

Factor

The system's input varied in the scope of a Design of Experiments (DoE).

FE-model

A mathematical physical model based on the Finite Element analysis.

FE-simulation

A computational simulation with a mathematical physical model based on the Finite Element analysis.

Feature

The term feature indicate the input to a metamodel

Ground Truth

A ground truth is the recorded response of the observed system to a given input (features). Hence it is the reference for the targets.

Hardware-model

A physical model based on in hardware. It is labelled as a model, because it resembles a real world scenario but simplified and under laboratory conditions.

Human Surrogate

A representation of a human. In context of this work, a human surrogate can be either an Anthropomorphic Testing Device (ATD) or an Human Body Model (HBM).

Hyperparameter

The variables of a metamodel. Hyperparameters characterize the learning algorithm's architecture and learning strategy.

Inflator

Gas generator of an airbag.

Injury-criterion

An injury criterion is a physical quantity relating the measured values in a crash test to an injury severity.

Input Layer

The first layer of an Artificial Neural Network (ANN). Typically, this layer represents the interface and passes the presented data directly to the next layer.

Instance

One particular observation in general. Specifically in this work, an instance is a pair of results from a FE-simulation from the same experiment. One FE-simulation with a particular human surrogate is used as features of the metamodel. The other FE-simulation with another human surrogate is used as targets of the metamodel

Integration Interval

The distance of two neighboured timesteps of the numerical integration.

Intrusion

In general the deformation of the BIW. For test scenarios of frontal crashes and occupants in a vehicle's first seat row, intrusion is the deformation of the bulkhead.

ISO-code

Sensor channel naming convention in orientation on ISO 13499 (see page 31).

ISO-grade

Label related to the objective rating metric for non-ambiguous signals as defined in ISO/TS 18571:2014(E) (see page 30).

ISO-rating

Numerical value related to the objective rating metric for non-ambiguous signals as defined in ISO/TS 18571:2014(E) (see page 30).

Kernel Function

A function which maps an input into another space.

L1 Regularization Term

A term added to the cost-function to penalize increasing absolute value of the parameters.

L2 Regularization Term

A term added to the cost-function to penalize increasing squared value of the parameters.

Layer

An organizational part of an Artificial Neural Network (ANN) containing neurons.

Level

The possible values a factor of a DoE can take.

Load Case

A vehicle test scenario which is characterized by the impacting barrier, the impact velocity, and the impact direction. In this thesis are standard velocities implied. The Full Width Barrier (FWB) related to $56\,\mathrm{km\,h^{-1}}$, Offset Deformable Barrier (ODB) to $64\,\mathrm{km\,h^{-1}}$, and Oblique Movable Deformable Barrier (OMDB) to $90\,\mathrm{km\,h^{-1}}$.

Loss-function

The loss-function quantifies how much a metamodel prediction deviates from the ground truth for a single instance.

Metamodel

A metamodel refers to an algorithm, in which inner parameters can be fitted regarding a given dataset.

Metric-function

Quantifies how much a metamodel prediction deviates from the ground truth for a whole set of instances.

Multivariate Time-series

Multiple time-series of a particular instance.

Neuron

The smallest part of an Artificial Neural Network (ANN).

Objective-function

The objective-function quantifies how much a metamodel prediction deviates from the ground truth as a sum of the cost-function and the regularization term.

Output Layer

The last layer of an Artificial Neural Network (ANN). Typically, this layer is a dense layer with a task specific activation function.

Overfitting

Overfitting takes place, if a metamodel fits too close to the training-set. Overfitting is indicated by a significant better performance for predictions on the training-set than on unseen data like from the validation-set. This condition is also referred to as high variance of the metamodel

Panel

Samples of a particular channel for multiple instances.

Parameter

The variables within a metamodel. Parameters are adjusted by an optimization process so that the metamodel output fits to a reference.

Pooling Layer

A layer type of an Artificial Neural Network (ANN), which is usually combined with a convolutional layer. The pooling operation extracts from a defined window single scalars with a defined strategy.

Pull-out

The amount of the seatbelt webbing pulled from the spool of the retractor.

Pull-through

The amount of the seatbelt webbing pulled through the seatbelt-buckle from the lap part to the shoulder part or vice versa.

R2-score

Metric-function as Coefficient of Determination. In this thesis capped for values lower than 0.

Reference Report

Validation reports of Honda Accord 2024 FE-model with structural FE-simulations by Singh et al. (2018b) and occupant FE-simulations by Singh et al. (2018a).

Regularization Term

A term added to the cost-function to penalize increasing values of the parameters.

Retractor

The retractor houses the seat-belt on a spool. For buckling of the occupant and during the forward motion of the occupant in the event of a crash, the belt is pulled out. Additional functions can be pre-tensioning (pulling belt into the device) and the load limited release of webbing.

Sample

A sample is a particular value of a particular time-series at a particular timestep.

Sampling Interval

The distance of two neighboured timesteps.

Sampling Rate

The number of samples per time window (typically 1 s).

Seatbelt

The seatbelt (or safety belt) is the primary restraint system. It consists, as a very basic version, only of a part wrapped around the lap (pelvis). The extended version features an additional seatbelt part wrapped around the thorax (shoulder). The mounts of a seatbelt system are called anchor points. The seatbelt can technically be tensioned and load-limited at all anchor points. The lap- and thorax-seatbelt parts are connected at the seatbelt buckle; either as one continuous seatbelt which can be pulled through the buckle or an individual segments in case of a fully locked buckle.

Seatbelt-buckle

A seatbelt-buckle is a connection to one of the seatbelt-systems anchor points. It connects the lap- and thorax-seatbelt parts; either as one continuous seatbelt which can be pulled through the buckle or a individual segments in case of a fully locked buckle.

Sill

A sill is a part of the BIW. It describes the area under the doors.

Submarining

The slipping of the seatbelt over the pelvic iliac wings.

Target

The target is the output of a metamodel.

Testing-set

A set of instances neither used during the fitting process of a metamodel parameters nor during the metamodel hyperparameter optimization.

Time-series

Samples of a particular instance and particular channel.

Time-to-fire

Moment of activation of an active restraint system component like the airbag inflator and the seatbelt pre-tensioner.

Timestep

A particular point in time.

Training-set

A set of instances used during the fitting process of the metamodel parameters.

Transfer-set

A set of instances with a completely different FE-model than the metamodel was trained on.

Trial

A trial in this work is one training and evaluation run of the machine learning framework with a particular learning algorithm, database, and hyperparameters.

Underfitting

Underfitting takes place, if a metamodel does not fit the training-set sufficiently. Underfitting is indicated by low performance for predictions on the training-set and all other sets. This condition is also referred as a high bias of the metamodel

Validation

In the context of virtual modelling, validation is a process to fit a FE-model response to the responses of its precedent within the evaluated scenario. In the context of machine learning, validation is the performance assessment of the metamodel using the unseen validation-set during the training.

Validation-set

A set of instances not used during the fitting process of a metamodel parameters.

Vehicle-pulse

A vehicle's pulse is the rigid body motion of the body in white during the crash. The pulse is characterized by the x, y, and z translation and rotation over time. The kinematics can be expressed as displacement, velocity, or acceleration.

Acronyms

ANN Artificial Neural Network

ATD Anthropomorphic Testing Device

BIW Body In White

BMI Body Mass Index

CAB Curtain Airbag

CAC_3 Chest Acceleration over 3 ms Criterion

CDC Chest Deflection Criterion

CFC Channel Frequency Class

CISS Crash Investigation Sampling System

CNN Convolutional Neural Network

CoG Centre of Gravity

CORA CORrelation and Analysis

CPU Central Processing Unit

CT Computed Tomography

CVC Chest Viscous Criterion

DAB Driver Airbag

DoE Design of Experiments

DoF Degree of Freedom

Euro NCAP European New Car Assessment Programme

FCC Femur Compression Force Criterion

FEA Finite Element Analysis

FMVSS Federal Motor Vehicle Standards

FWB Full Width Barrier

GHBMC Global Human Body Models Consortium

GPR Gaussian Process Regression

HAC_3 Head Acceleration over 3 ms Criterion

HBM Human Body Model

HIC Head Injury Criterion

HIC_{15} Head Injury Criterion for up to 15 ms

HIII Hybrid III

HIII-05F Hybrid III 5^{th} Percentile Female Dummy

HIII-50M Hybrid III 50^{th} Percentile Male Dummy

HIII-95M Hybrid III 95^{th} Percentile Male Dummy

IQR Interquartile Range

ISO International Organization of Standardization

LS-DYNA Livermore Software technology corporation DYNA

LSTM Long Short-Term Memory

MADYMO MAthematical DYnamic MOdels

MAE Mean Absolute Error

MBS Multi Body Simulation

MCDM Multi-Criteria Decision Making

NASS National Automotive Sampling System

NEC Neck Extension Moment Criterion

NHTSA National Highway Traffic Safety Administration

Nij Neck Injury Criterion

NSC Neck Shear Force Criterion

NTC Neck Tension Force Criterion

ODB Offset Deformable Barrier

ODB-L Offset Deformable Barrier Left

ODB-R Offset Deformable Barrier Right

OLC Occupant Load Criterion

OMDB Oblique Movable Deformable Barrier

OMDB-L Oblique Movable Deformable Barrier Left

OMDB-R Oblique Movable Deformable Barrier Right

PAB Passenger Airbag

PC Principal Component

PCA Principal Component Analysis

PHMS Post Mortem Human Subject

ReLU Rectified Linear Unit

RNN Recurrent Neural Network

SAE Society of Automotive Engineers

SLL Switchable Load Limiter

SOLB Small OverLap Barrier

SOLB-L Small OverLap Barrier Left

SOLB-R Small OverLap Barrier Right

SVM Support Vector Machine

THOR Test Device for Human Occupant Restraint

THOR-50M Test device for Human Occupant Restraint, 50^{th} Percentile Male

THUMS Total Human Model for Safety

VHUM VIRTHUMAN

VHUM-05F VIRTHUMAN 5^{th} Percentile Female Human Body Model adapted to HIII-05F

VHUM-50M VIRTHUMAN 50^{th} Percentile Male Human Body Model adapted to HIII-50M

VHUM-95M VIRTHUMAN 95th Percentile Male Human Body Model adapted to HIII-95M

ViVA Virtual Vehicle Safety Assessment

VPS Virtual Performance Solution (ESI)

Nomenclature

The following notation was used within this work. The same letter keeps its principal meaning, irrespective of its annotations and font weight. Only the font type, like A and $\mathcal{A}$ and alphabet like E and Σ changes a symbol's meaning.

- Matrices are denoted by uppercase letters in a bold font in Latin $\boldsymbol{ABC}$, calligraphic Latin $\boldsymbol{\mathcal{ABC}}$, or Greek $\boldsymbol{\Phi\Psi\Omega}$.
- Vectors are denoted by lowercase letters in a bold font, like Latin $\boldsymbol{abc}$ or Greek $\boldsymbol{\phi\psi\omega}$.
- Scalars are denoted by a normal weight font, like Latin $ABCabc$, calligraphic Latin $\mathcal{ABC}$, or Greek $\Phi\Psi\Omega\phi\psi\omega$.
- Sets are denoted using the font blackboard bold $\mathbb{ABC}$.
- Value intervals are denoted by (or) for excluded border values and [or] for included border values.

Total derivatives are written either in the $^{\mathrm{d}u}/_{\mathrm{d}t}$ convention, or denoted for the special case of the derivatives by time, by a dot over the symbol, e. g. $\dot{u}$ for the first, and $\ddot{u}$ for the second derivative. Partial derivatives are denoted in the $^{\partial u}/_{\partial t}$ convention.

Some annotations are used with a fixed meaning. An estimation is denoted by $\hat{u}$, while the related ground truth is denoted without a hat. The arithmetic mean of a vector is denoted with a horizontal bar, like $\bar{u}$. and the median by a tilde, e. g. $\tilde{u}$. The result of a signal CFC transformation of u is denoted by $\breve{u}$. Results from other transformation are denoted by $\breve{u}$. Ultimately, the annotations around a symbol follow the convention:

$$\overset{\text{Identifier}}{}u\overset{\text{Operation}}{}_{\text{Spatial Direction / Matrix / Vector Element Index}}$$

The font weight in combination with the lower- and uppercase denotes the status of matrix, vector, and scalar for the complete annotated symbol. A symbol directly followed by round brackets indicates a dependency, such as e. g. $U(t)$.

Constants

$\mathcal{R}$	Molar Gas Constant	8.314 462 618 153 24 J mol^{-1} K^{-1}
e	Euler's Number, denoted as "exp"	$\approx$2.718 28

π	Mathematical Constant pi	$\approx$3.141 59
g	Physical Constant of Standard Acceleration of Gravity	9.806 65 m s^{-2}

Operators

∇	Nabla Operator
∂	Operation of Partial Derivative
d	Total Derivative Operator

Sets

$\mathbb{E}$	Set of Equations
$\mathbb{F}$	Set of Features with Size $\mathcal{F}$ and Index ξ
$\mathbb{N}$	Natural Numbers
$\mathbb{P}$	All Parameters of a Metamodel
$\mathbb{R}$	Real Numbers
$\mathbb{S}$	Set of Instances with Size $\mathcal{S}$ and Index ς
$\mathbb{T}$	Set of Target Attributes with Size $\mathcal{T}$ and Index τ
$\mathbb{Z}$	Integers

Physics

β	Angle Between the Seatbelt-sections before and after the D-ring in [°]
δ	Angle of Cantilever Measuring Chest Deflection of ATD HIII
ϵ	Lagrangian Multiplier
Γ	Physical Quantity of Temperature
γ	Wrapping Angle of a Capstan in [rad]
$\mathcal{U}$	Number of Generalized Coordinates in a Multibody System
$\mathcal{V}$	Number of Constraining Equations in a Multibody System

$\mathcal{Z}$	Lagrangian as the Difference of the Total Kinetic and Potential Energy of a Multibody System
μ	Friction Coefficient
ω	Angular Frequency, with ω_a and ω_d being the Angular Cut-off Frequencies for SAE J211 and ISO 6487 (CFC)
ϕ	Constraining Equations of a Multibody System
Π	Physical Quantity of Molar Mass
ρ	Physical Quantity of Density
a	Filter Coefficients a_0, a_1, and a_2 for SAE J211 and ISO 6487 (CFC)
b	Filter Coefficients b_1 and b_2 for SAE J211 and ISO 6487 (CFC)
D	Physical Quantity of Translational Displacement
E	Physical Quantity of Elastic Modulus
F	Physical Quantity of Force
K	Stiffness of a Finite Element System
L	Length
M	Physical Quantity of Moment of Force (Torque)
m	Physical Quantity of Mass
N	Nominal Filter Class for a Filter According to SAE J211 and ISO 6487 (CFC), Typical Values: 60, 180, 600 and 1000
P	Physical Quantity of Pressure
p	Generalized Forces at the Degrees of Freedom of a Multibody System and External Forces of a Finite Element System
q	Generalized Coordinate of a Multibody System
R	Resultant Direction as of Vector with X-, Y-, Z-component as Euclidean Norm
r	Radius
S	Damping of a Finite Element System

U Physical Quantity of Displacement

u Displacement Vector

V Physical Quantity of Volume

W Mass of a Multibody or Finite Element System

X Spatial Direction 1 of a System

Y Spatial Direction 2 of a System

Z Spatial Direction 3 of a System

z Discrete Signal, with $\breve{z}$ being the Filtered Signal According SAE J211 and ISO 6487 (CFC)

Statistics and Machine Learning

α Learning Rate

χ A Single Feature Attribute from $\boldsymbol{\mathcal{X}}$

η Neuron Identifier of ANN

$\imath$ Index of Gradient Descent Iteration

$\jmath$ Objective-function

κ Kernel

λ Weight of Regularization Term as Hyperparameter

$\mathcal{A}$ Length of Convolutional Kernel

$\mathcal{F}$ Number of Features in $\mathbb{F}$

$\mathcal{H}$ Absolute Frequency of Correctly of Wrongly Classified Instances

$\mathcal{I}$ A Particular Injury-criterion

$\mathcal{K}$ Covariance Matrix

$\mathcal{L}$ Loss-function

$\mathcal{M}$ Metric-function

$\mathcal{N}$ Normal Distribution

$\mathcal{P}$ Probability

$\mathcal{S}$ Number of Instances in a Dataset $\mathbb{S}$

$\mathcal{T}$ Number of Targets (Labels) in $\mathbb{T}$

$\mathcal{X}$ Features / Input to a Metamodel

$\mathcal{Y}$ Targets / Output of a Metamodel

ψ Layer Identifier of ANN

σ^2 Statistical Measure of Variance and Hyperparameter in Gaussian Radial Basis Function

τ Target's Attribute Index with $\tau = \{1, 2, \cdots, \mathcal{T}\} \in \mathbb{N}$

Θ Parameters of a Metamodel

θ Weights of Transfer Function of ANN

Υ Transfer-function of a Neuron of an ANN

υ Single Attribute of the Target

ε Bias

ϱ Output of a Layer of a Neural Network

ς Sample Index with $\varsigma = \{1, 2, \cdots, \mathcal{S}\} \in \mathbb{N}$

ξ Feature Attribute Index with $\xi = \{1, 2, \cdots, \mathcal{F}\} \in \mathbb{N}$

ζ Index Denoting a Particular Time-series or Channel

c Cost Function

k Index of Kernel of CNN

Others

$\mathcal{C}$ Number of Sensor Signals

$\mathcal{D}$ Number of Samples in the Time-series

$\mathcal{E}$ Number of Experiments

f	A Generic Function
s	Exponent of Sobol Algorithm
T	Time Signal Sampling Interval; Step Size in Numerical Time Integration
t	Time Index; Additional Index like t_1 Indicating a Particular Time Stamp

1 Introduction

1.1 Motivation

The field of vehicle safety has existed almost since the beginning of automobiles. Important milestones of passive safety, briefly outlined by Köllner (2023), were the patents of the crumble zone in 1952 by Béla Barényi, the seatbelt in 1959 by Nils Ivar Bohlin, and the airbag in 1971 by Walter Linderer. Alongside the technical innovations and different global stakeholders, the German regulatory bodies promoted passive safety. In 1974, seatbelts were made mandatory for new vehicles and since 1976, occupants were legally required to use them. Following the analysis of Evers (2021), the usage quota of seatbelts was almost immediately over 90 % for drivers. In contrast, for passengers sitting in the rear, it took about 20 years to reach this level.

In parallel with the technical development and regulations, the assessment methods also progressed. From the beginning with rocket propelled, full vehicle crash tests in the late 1950s, over the introduction of crash test dummies in the 1970s, there is now a broad portfolio of assessment methods available. As outlined in figure 1.1 on the following page, assessment usually includes the modelling of the reality. Derived from accident research and statistics, scenarios and criteria are defined, and the modelling methods are developed. The physical testing covers full-scale crash tests, reduced testing on a sled, and single component tests using the drop tower or specialized methods. In synergy with the experimental simulations are the numerical simulations. Multi Body Simulation (MBS) is primarily used to model kinematic processes like vehicle dynamics. Finite Element Analysis (FEA) adds the deformability of objects and is the major tool for analysing crash mechanics. The Finite Volume Analysis (FVA) is well suited for fluid simulation like aerodynamics and detailed airbag modelling.

Schöneburg (2023) described that the utilization of physical and virtual assessment methods in their interplay during the vehicle development process. Virtual prototyping is dominant in the concept phase. Nearer the development process progresses towards the vehicle's approval, more testing in hardware is used. Once prototypes are available, the continuous validation of the virtual models is central.

Lately, the industry has undergone substantial changes. As Schöneburg (2023) outlined, the increased competition alongside the market and regulatory requirements enforced the development cycles to shorten from seven to nine years to three to seven years while the product complexity drastically increased. To succeed in this environment, manufactures and

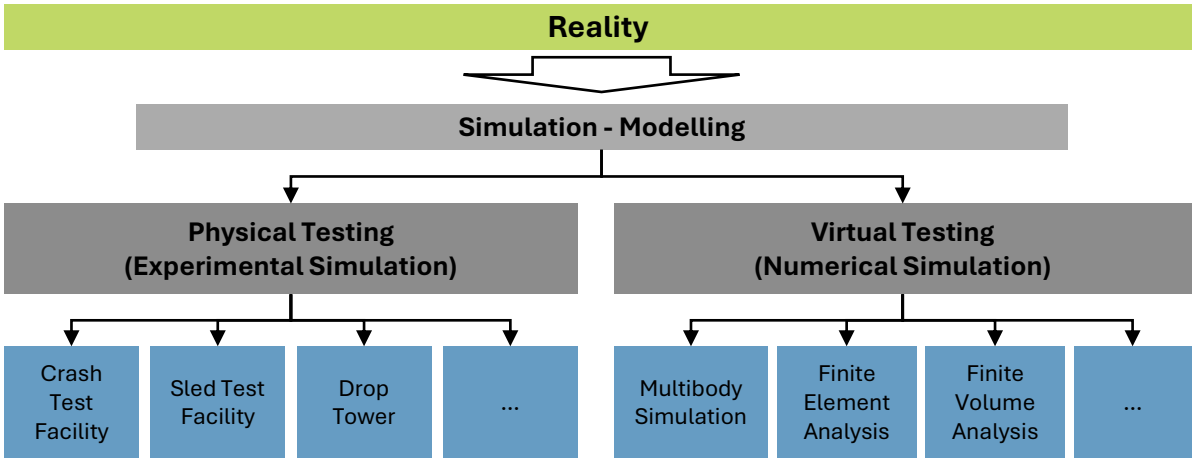


Figure 1.1: Overview on Simulation Methods

engineering service providers must push the limits of the current assessment methods.

Not only the market requirements itself, but also the worldwide commitment to vehicle safety under the Vision Zero, for example described by Elvebakk (2007), will increase the space of considered scenarios in which safety systems must operate effectively. The awareness of the complexity of road traffic scenarios and the diversity of road traffic users is derived from this philosophical concept. The vast variety challenges the current concepts of singular testing and local optimizations, while technical innovations are catalysts for new ways of thinking.

The machine-learning-powered metamodels, which is an algorithm where stored parameters contain statistical information on the modelled subject, can be the next significant step based on hardware and virtual physical models. According to the annually updated Gartner's Hype-cycle model, displayed in figure 1.2 on the next page, Artificial Intelligence (AI) powered simulations are in a phase of triggering innovation and are expected to reach the plateau of productivity within the next five to ten years. Impactful breakthroughs inspire efforts in the field in almost all domains. The over centuries grown know-how around vehicle safety and the increased computational power are the enablers for the promising technology.

This presented research roots in the domain of virtual assessment in the field of passive vehicle safety. Existing development processes should be accelerated and augmented. In addition to increasing efficiency, the acceleration of the processes should enable the consideration of a broader range of requirements. Specifically, the variance in human dimensions such as body size or types should be better reflected during the design process.

For the fast and cost-efficient estimation of the effect of different characteristics of the human body in a vehicle crash, complex FE-simulations with computation times in the range of hours should be extended by metamodels enabling estimations with computation times in the range of seconds.

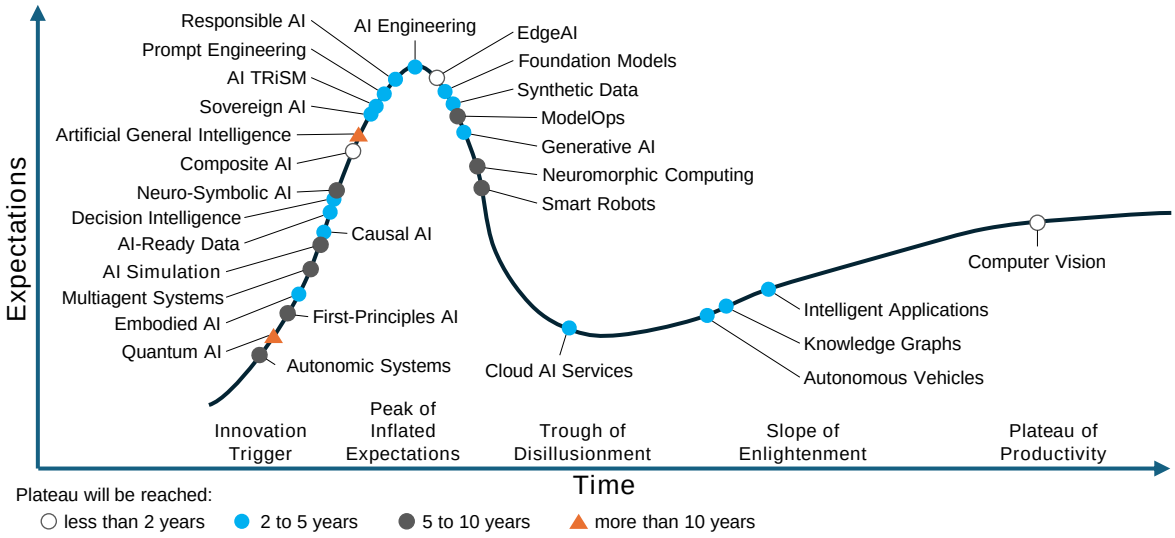


Figure 1.2: Gartner's Hype-cycle for Artificial Intelligence (adapted from Gartner, 2024)

1.2 Problem Statement

The motivation above implies two main objectives for a metamodel to be successful. The system must generate estimates faster, while the whole life cycle cost must be lower than for virtual simulations. Naturally, this includes a sufficient accuracy of the estimates.

The very nature of metamodels is the rapid prediction. To train a metamodel, a vast amount of data is necessary. If the distribution of the data changes over time, the metamodel must be trained again. For the presented research, it is assumed that the data will always contain the input and output of the metamodel. Hence, supervised learning, a subclass of machine learning, seems apt.

In supervised learning, the algorithm attempts to fit a function which maps an input, so-called features, to an output, the targets. To generate data, typically a Design of Experiments (DoE) is performed. The different settings of the factors in DoE are commonly used as features and the resulting system's response as target attributes. However, this approach has the major limitation that the relevant characteristics of the system must be known. In domains like vehicle safety, this cannot be always guaranteed. For example, the change of the airbag supplier can lead to similar but not parametrizable behaviour of this restraint component.

To overcome this limitation, the implicit information from reference simulations appears promising. One occupant simulation in a particular vehicle configuration would be necessary to capture implicitly the characteristics of this specific configuration. The learning algorithm should be able to capture the relationship of this reference simulation to different anthropometrical configuration of occupants. Typically, occupants are modelled using human surrogates like Anthropomorphic Testing Devices (ATDs) or Human Body Model (HBM) with the anthropometrical configuration of a population's 5^{th}, 50^{th}, and 95^{th} percentiles. Consequently, the database must contain variations in the vehicle's configuration as well as in the anthropometrics.

The basic idea forming the DoE is to vary the vehicle's factors by an appropriate scheme. The occupant's anthropometrical configuration is then repeated for each experiment of the DoE. The data, which is generated by simulation of those configurations, can be streamed through a general machine learning architecture, as displayed in figure 1.3. The database contains reference simulation of one anthropometrical configuration is used as features. The simulation with another anthropometrics but with the same vehicle configuration is used as targets.

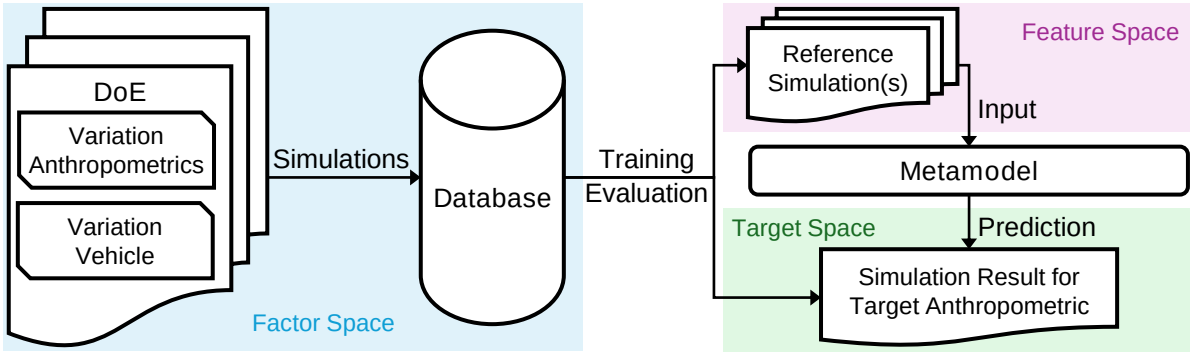


Figure 1.3: General Machine Learning Architecture and Data Flow

The output of the simulation can be recorded data as a signal in the time domain like an acceleration signal, or a three-dimensional representation of the geometry over time. Storing the geometry representation in a high resolution of the time domain would require tremendously higher storage capacities than using the sensor output. Because of this reason, the data should be limited to sensor signal within the presented research. Hence, the input and output of the metamodel are multivariate time-series. Feature attributes and targets attributes can be extracted from this data type.

The naming conventions are defined by figure 1.4. The tabular data can be expressed as a table of attributes and the temporal data as a cube of time-series. A dataset is formed by instances. Data from the metamodel input is denoted as feature and from the output as target.

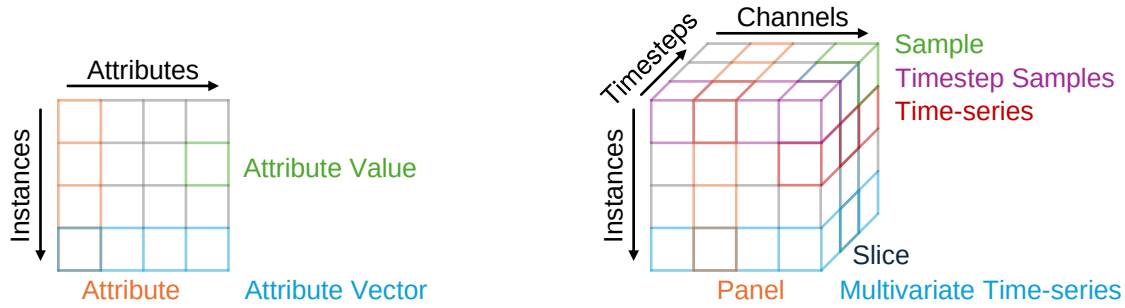


Figure 1.4: Naming Conventions for Data Tabular Data (left) and Temporal Data (right)

In terms of efficient data generation, the transition of the factor space to the feature space becomes relevant. The factors forming the DoE span the first space – the factor space. The

responses of the reference simulation for each experiment of the DoE form the second space – the feature space. The first space has fewer dimensions compared to the second space.

Data can be generated in advance or directly during the training of the metamodel. For the queried data generation, some directives seem necessary. Naturally, a form of confidence of the metamodel in areas of the factor space is used. Since the evaluation takes place in the second parameter space, new data cannot directly be generated. It is required to perform the transition into the first space before the data generation.

1.3 Research Questions and Outline

The following research questions were formulated from the problem statement explained above. The questions are aligned with the bigger picture of how a metamodel could be designed without explicit characterization of the vehicle.

- Can a human surrogate response sufficiently characterize a vehicle?
- How can the required data be generated and used efficiently?
- To which extent the non-linearity of the system limits the capabilities of the metamodel?
- Can domain knowledge be utilized to reduce the required amount of data?
- What are the related works?

To answer the questions, this report was structured in four chapters. After the state-of-the-art was described, a FE-model was selected and simplified. Based on this model, databases were generated and utilized for machine learning applications. For research purpose, the premise was to generate whole databases in advance. In later steps, the sequential data generation approaches were studied. The relevant software packages are outlined in appendix A.

The **chapter 2**, aims to provide an overview of this research's context and display the relevant theoretical fundamentals. The overviews cover the topics of machine learning, occupant safety, and human body diversity. The chapter starts with an overview of passive safety and the common human surrogates. Additionally, the typical assessment methods and criteria are introduced. Next, the relevant theoretical background for training and evaluation of supervised machine learning models is presented. Finally, the concepts of efficient data generation and utilization are collected. Both topics are brought together by collecting research of covering the human body diversity in passive occupant safety. A digression on Multi-Criteria Decision Making (MCDM) enlightens the extended methods for the evaluation and selection of different alternatives for setting up the discussed system.

The topic of **chapter 3** is the efficient generation of databases. The criteria for an appropriate virtual model were defined, being the fundament for the selection of the base FE-model. In the chapter, the order of the individual simplification steps is motivated, for each step the applied measure described and its impact on the FE-model version's validity and computational cost

discussed. The final stage of the simplification process was used to generate 2 development-sets for the training and evaluation of the metamodel. To assess the combination of the simplifications and metamodel, a 3rd database, called testing-set, was generated using an intermediate step of the simplification chain. To investigate the possibilities of the transfer between FE-models of different complexity, a so-called transfer-set was generated. The Anthropomorphic Testing Device (ATD) representative Hybrid III (HIII) was used as a base and the Human Body Model (HBM) representative VIRTHUMAN (VHUM) as a representant of a complex model. The FE-models and the generated databases were made available to the public as digital supplements.

Finally, in **chapter 4**, metamodels were presented and discussed. The input for the metamodels is of the type of unstructured data, or more specifically, they are multivariate time-series. The learnable feature attributes extraction architectures using Convolutional Neural Network (CNN) and the Long Short-Term Memory (LSTM) based Recurrent Neural Networks (RNNs) are compared. As an alternative to the learnable feature attributes extractors, the data transformations are discussed. The vehicle characterizing human surrogate was a 50th percentile. The task was the target attributes regression of injury-criteria for 5th and 95th percentiles. The task was extended by target attributes classification for different number of classes for the criticality of a crash. Furthermore, the target multivariate time-series regression of the channels from the FE-simulations was considered. For the extended estimates, a method from the collected alternatives was chosen. At last, the influence of different number of instances for training and the effectivity of an exemplary querying algorithm is presented. As another way of efficient data utilization, the transfer-set was used for transfer learning; the homogeneous transfer of the learnable feature attributes extractor from one domain to another. Additionally, the domain of the relatively simple ATD was used to directly estimate results from FE-simulations with more complex HBM.

2 Background, Relevant Fundamentals, and Related Work

This chapter aims to provide an overview of the related work of this dissertation. The first section introduces the basic concepts of passive safety assessment methods. Next, the fundamental techniques of machine learning are discussed. This chapter is focussed on the relevance of human body diversity and approaches to achieve broad coverage. Ultimately, the concepts for the assessment of cost and use factors are discussed.

2.1 Overview Passive Safety

2.1.1 Passive Safety Systems

Schöneburg (2023) defines passive safety as measures to reduce the severity of an accident. Accident research is crucial to identify fields of actions and control the success of measures. The insights from the accident research and the understanding of the human body's biomechanics led to the development of passive safety systems.

Passive safety performance is primarily determined by size, shape, structural design, and restraint systems of the vehicle. For occupant safety, the typical strategy is a very stiff design of the passenger compartment and energy absorbing structural components like crash boxes or the longitudinal members outside that compartment.

The primary restraint system consists of passive design elements of the interior such as limited local stiffnesses, the absence of sharp edges, and simple seatbelts. The seatbelt and its main components and force measuring location are presented in figure 2.1 on the following page. The extended or supplemental restraint systems are airbags and the extension of seatbelt systems by load limiters and pre-tensioners. Standard airbags are the Driver Airbag (DAB) and the Passenger Airbag (PAB). In most vehicles, side airbags and Curtain Airbags (CABs) can also be found. Especially in the US, vehicles are additionally equipped with knee airbags due to the unbelted test cases in the Federal Motor Vehicle Standards (FMVSS) 208. However, since incentivised by e. g. European New Car Assessment Programme (Euro NCAP) in form of improvement of the rating, knee airbags are installed in other regions as well. Furthermore, airbags for the vehicle's centre, belt, and seat ramp were developed.

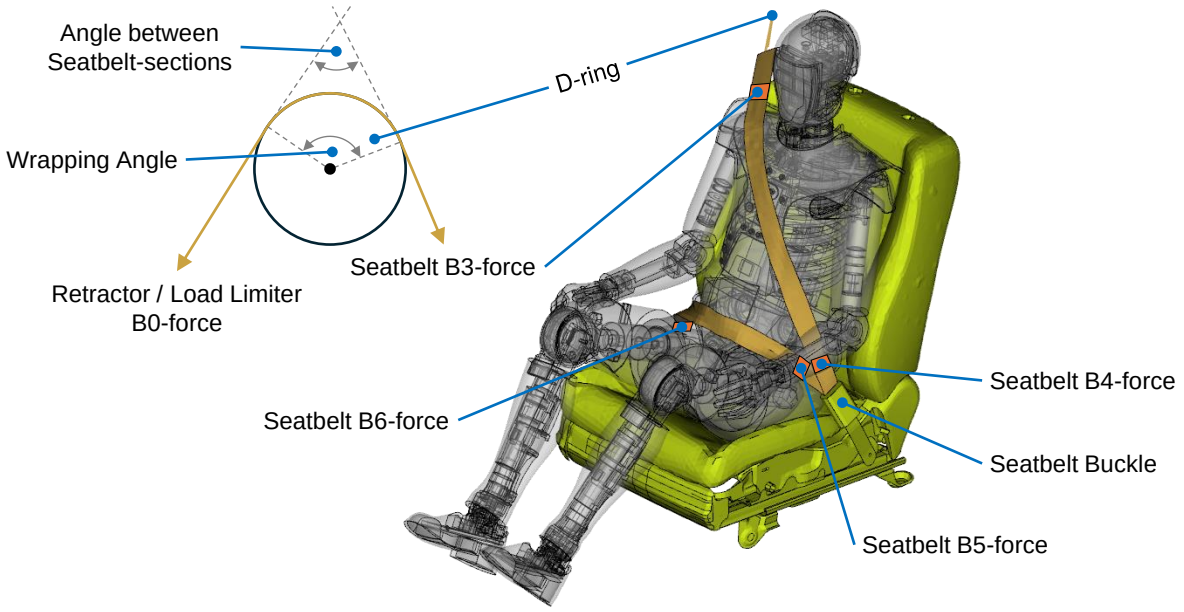


Figure 2.1: Seatbelt Force Sensor Locations and Naming Convention

Nowadays, the systems are activated a few milliseconds after the impact, depending on the measured accelerations. Newer development aims for earlier activation, even before the impact, by using the environment sensors, for example described by Lugner et al. (2020). In case of a frontal impact, firstly, the occupant is restraint by the seatbelt alone. In the further forward motion, once the frontal airbag is reached, both seatbelt and airbag are active to control the occupant's kinetic energy dissipation.

To take different body shapes and behavioural preferences of the occupant into account, the active restraint systems possess adaptivity to some extent. Typical are different timings to activate multiple loads limiting levels of the seatbelt system, together with the timing of opening and airbags adaptive went to control its inner pressure. The different levels of the adaptivity can be triggered, e. g. by the position of the seat, the occupant's weight measured in the seat, or the surveillance system.

Another motivation driving the development of adaptivities is the implications of autonomous driving. The absence of the driver's task can enable new activities during the ride. Jorlöv et al. (2017) and Östling and Larsson (2019) conducted studies in Sweden and China asking a general audience, which sitting positions and activities they could imagine during a ride with an autonomous vehicle. The authors identified the five major positions displayed in figure 2.2 in combination with different levels of how much the seat's backrest is reclined. In the case each position should offer a similar level of passive safety as in a conventional car, new restraint systems have to be developed. Next to airbags in new locations, extended belt systems and shape adaptive airbags are promising examples, as illustrated by Laakmann et al. (2019).

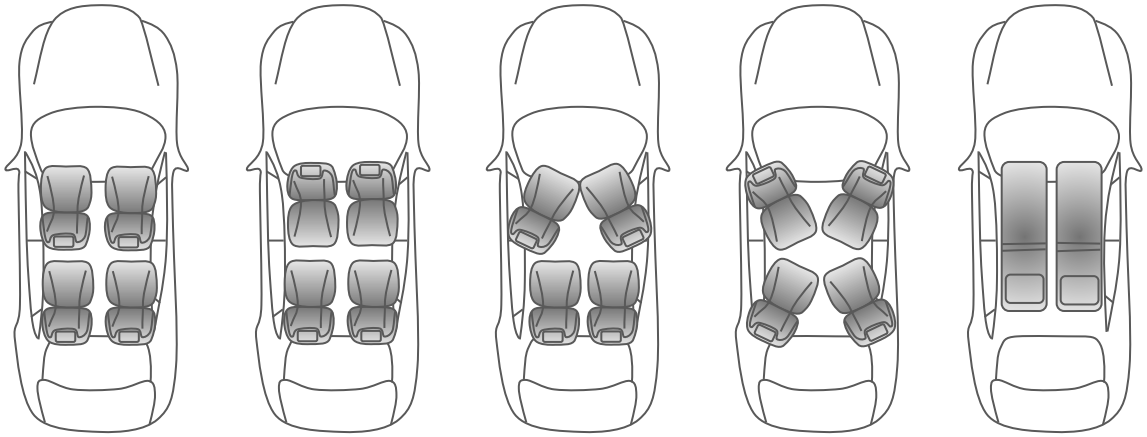


Figure 2.2: Potential New Sitting Positions (adapted from Östling and Larsson, 2019)

2.1.2 The Diversity of the Human Body

2.1.2.1 Anthropometrics

Anthropometric data exists as (a) general statistics about body height and weight conducted for several regions, (b) more detailed studies collecting numerous external anthropometrical measures, and as (c) external and internal geometric data conducted by studies based on Computed Tomography (CT) scans or cadaver analysis.

Robinette et al. (2002) introduced the CAESAR database, containing anthropometrical data for 18 to 65 years old in North America and Europe. North American data were collected in one Canadian city and the USA. The European data was taken from the population in the Netherlands and Italy.

The National Health and Nutrition Examination Survey (NHANES) collects data on height and weight of the US American population. The program named National Household Education Surveys Program (NHES) began in the early 1960s and is still ongoing, as documented in (U.S. Department of Health & Human Services, 2021).

An extensive collection of external measurements was carried out by Reed and Ebert (2013) and Reed and Rupp (2013) for sitting subjects with 310 soldiers and later an additional 200 elderly occupants. The data were used by numerous studies on the effect of human body variations on its crash behaviour. The database was also used by Park and Reed (2016) for BioHuman, an interactive model that predicts anthropometrics by body height, Body Mass Index (BMI), age, and the ratio of the upper to the lower body lengths (SHS).

Gordon et al. (1989) performed a survey on US army personnel, collecting detailed anthropometrical measures. The database was published under the name Anthropometric Survey of US Army Personnel (ANSUR). An update to ANSUR II was released by Gordon et al. (2014). The data were used, for example, by Östh et al. (2017) and Gayzik et al. (2011) as a reference for creating different Human Body Models (HBMs).

An example of an in-depth database for internal and external measures is the Reference Analytic Morphomic Population (RAMP) introduced by Wang et al. (2016). They collected data based on CT scans of approximately 5000 subjects.

Motivated by the risk of submarining, the slipping of the lap seatbelt over the pelvic iliac wings, the studies focused on the deviations in the pelvic region. For example, Izumiyama et al. (2018) and Izumiyama et al. (2020) used X-rays to determine the different pelvis shapes, angles, and spine alignment of seated occupants.

Reed and Ebert (2013) compared the data from the 1970s with the data collected by National Health and Nutrition Examination Survey (NHANES) in the year 2000. The North American population seemed not changed significantly in height. Nevertheless, the percentiles of the weight shifted. For example, the Hybrid III 50th Percentile Male Dummy (HIII-50M) matches with the 33rd percentile of the population in 2000.

Wang et al. (2016) compared the North American population with the East Asian population. The authors took data for America from the Reference Analytic Morphomic Population (RAMP) database of CT scan-based measures. They chose the Asian-RAMP, which was conducted in Taiwan, as a representative population for East Asia. Both populations showed significant differences in all properties.

The comparison of body height and weight is presented in table 2.1. It confirms the findings of Reed and Ebert (2013) and Wang et al. (2016). The North American data for the 1970s is taken from Schneider et al. (1983), for the 2000s from Reed and Ebert (2013). The Malaysian data from Rahman et al. (2018) were chosen as representative for East Asia, and the Belgium data from Motmans (2005) to represent Europe. The Malaysian population is significantly smaller.

Table 2.1: Height & Weight of Percentiles in Different Populations

	North America 1970s	North America 2000s	Malaysia 2000s	Belgium 2000s
5th percentile Female	1511 mm	1508 mm	1464 mm	1534 mm
	47.0 kg	50.3 kg	42.7 kg	/
50th percentile Female	1618 mm	1622 mm	1551 mm	1646 mm
	62.3 kg	71.0 kg	55.8 kg	70.0 kg
50th percentile Male	1753 mm	1761 mm	1685 mm	1766 mm
	77.0 kg	85.4 kg	66.0 kg	80.4 kg
95th percentile Male	1869 mm	1887 mm	1793 mm	1943 mm
	102.3 kg	123.4 kg	94.8 kg	/

2.1.2.2 Masses, Inertia, Tissue and Joints

The determination of masses of body sections can be performed either by CT scans combined with a density approximation or by cadaver analysis. Segmenting of the body (e. g. in arms,

thorax) was done, for example by Leva (1996). More detailed studies, as conducted by Yoganandan et al. (2009) determine all body parts' properties (e. g. organs, tissues).

Dumas et al. (2007) updated the historical tables of body segment inertia parameters. They distributed the total body mass over each segment, weighted by the individual volumes. The validity of the constant density assumption in the body segmenting is questioned. Ackland et al. (1988) analysed the CT scan of one leg. They concluded, with the limitation on the investigated body part, that the assumption of an equal density appears to be valid. A recent literature study collecting different approaches to determine the body segment inertia parameters was done by Diaz-Rodriguez et al. (2016). They added a method using rigid body dynamics for the determination of such parameters.

Assessment of material and joint properties suffers, as in all human-related data collections for passive safety, from the limited load applicable to living subjects. Furthermore, cadaver specimens can have properties different from those of in-vivo material. Ramachandra et al. (2019) reviewed different approaches to assess the properties of the rib cage. They concluded significant differences in stiffness response between tests performed with the pure rib cage and tests using the entire thorax.

Araújo (2008) collected data about human joint flexibility in sports. Engin and Chen (1987) analysed the kinematic and passive resistive properties of the human elbow.

2.1.2.3 Variations in Sitting Positions

As mentioned above, anthropometrical variations can lead to different sitting postures. In addition, the degrees of freedom provided by a car's seat and the occupants' various activities also lead to different sitting positions. Finally, the test protocols prescribe specific configurations.

Manary et al. (1998) developed a human surrogate position model based on 606 subjects. The resulting posture models were extended by Park et al. (2016) with parameters for different ages. Compared to the changes in spine-related inner posture by Izumiyama et al. (2018), those statistical models predict the posture as a whole.

Given an adaptivity which is dependent on the seat's position, a single setting on which all individuals plausibly and likely can sit had to be defined. Reed and Flannagan (2000) observed a trend for smaller occupants to sit near to the dashboard and taller further away. Despite this clear and, from an ergonomic perspective, absolutely plausible trend within the recorded 96 male and female drivers, the authors found an unsatisfying correlation of an occupant's statue and its exact position in the seat's length adjustment.

An extended insight into likely sitting positions can be provided by the crash databases. The open available data is from Crash Investigation Sampling System (CISS) containing cases from 2017 to 2021 in the US. The selection criteria are defined as follows:

> [...] the target population for the CISS is all police-reported motor vehicle crashes on trafficways involving passenger vehicles (automobiles, automobile derivatives, SUVs, van-based light trucks, light conventional trucks (pickup-style cab), and

other light trucks with gross vehicle weight ratings <4.5 kg) and in each of which a passenger vehicle is towed from the scene for any reason. (Radja et al., 2022)

The database contains information on the involved occupant's age, height, weight, and gender. Additionally, the occupant's seat track setting is recorded in five levels reaching from "forward most" to "rear most".

The following filter criteria were applied to extract relevant data:

- Only passenger vehicles bigger than lightweight vehicles but smaller than light trucks,
- Only adult occupants (age $\geq$18),
- Only length (track) adjustable seats in known positions,
- Only occupants on front left and front right seats (equal driver & passenger positions),
- Only normal postures (standard use of seat with upright backrest),
- Only non-pregnant occupants.

As the CISS database contains only accidents records, the data is potentially biased compared to the general US population. As reference, data from the National Health Interview Survey (NHIS) database was used. NHIS collects data within the US since 1957 by conducting household interviews. The selection criteria are defined as follows:

> The target population for the NHIS is the civilian noninstitutionalized population residing within the 50 states and the District of Columbia at the time of the interview. The NHIS universe includes residents of households and noninstitutional group quarters (e. g., homeless shelters, rooming houses, and group homes). Persons residing temporarily in student dormitories or temporary housing are sampled within the households that they reside in permanently. Persons excluded from the universe are those with no fixed household address (e. g., homeless and / or transient persons not residing in shelters), active duty military personnel and civilians living on military bases, persons in long-term care institutions (e. g., nursing homes for the elderly, hospitals for the chronically ill or physically or intellectually disabled, and wards for abused or neglected children), persons in correctional facilities (e. g., prisons or jails, juvenile detention centers, and halfway houses), and U.S. nationals living in foreign countries. While active-duty Armed Forces personnel are not sampled for inclusion in the survey, any civilians residing with Armed Forces personnel in non-military housing are eligible to be sampled. (*National Health Interview Survey (NHIS)* 2022)

Heights in the survey are reported in Inch with zero floating points. Are the values converted to cm, some semi-discrete values are the result.

Figure 2.3 on the facing page shows the relationship of anthropometrical attributes to the sitting position of an occupant. Additionally, to display the potential bias, the filtered CISS population is compared to the overall US population from the National Health Interview Survey (NHIS) in

density histograms with 30 bins. For obvious reasons, there are significantly more drivers than passengers in the database. The box plots indicate the 25th, 50th, and 75th percentiles; outliers were determined by the Interquartile Range (IQR) factor of 1.5 (see section 2.2, page 55).

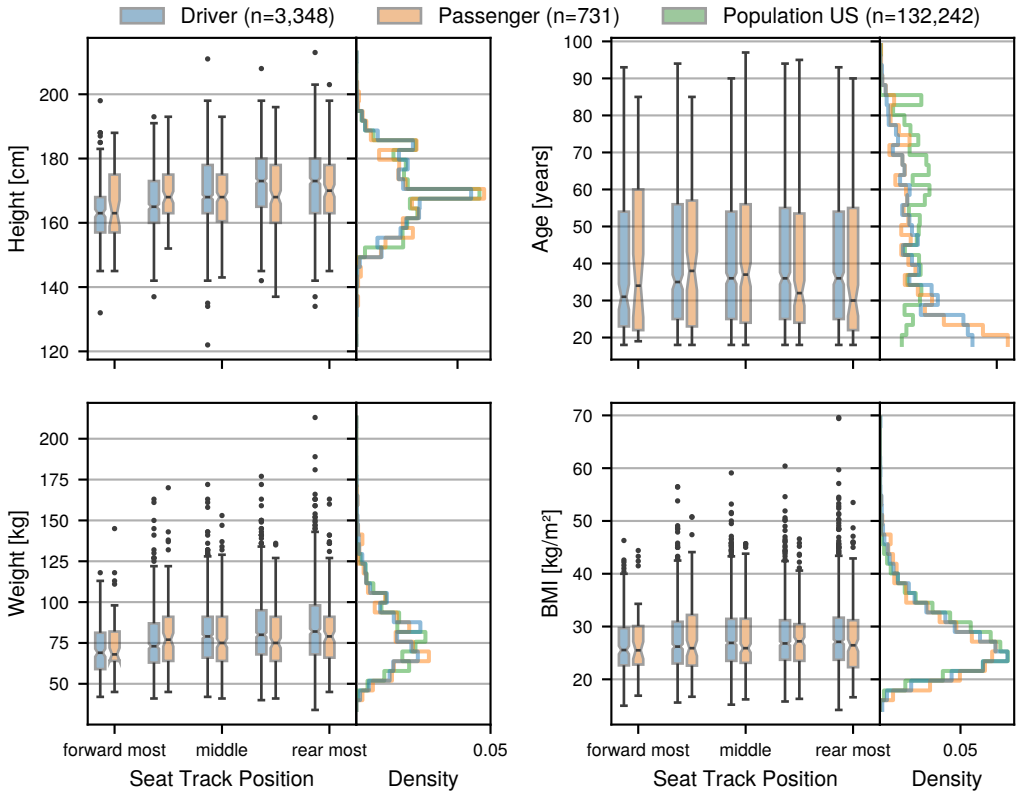


Figure 2.3: Seat Position to Age, Height, Weight, and BMI with Population Comparison based on CISS and NHIS Databases

For height, weight, and BMI, the crash-involved population seems to match the general population well. For age, the population which experienced a vehicle accident is seemingly much younger than the overall US population. Cox and Cicchino (2021) came to similar findings, associating the increased involvement with overall riskier behaviour. However, the wide spread of all ages in all seat track positions indicates a weak correlation between both factors.

In general, all anthropometrical parameters have a widespread on each seat track position. For drivers, there is a stronger correlation to height and weaker to weight. For passengers, the forward most position seems predominantly populated by smaller occupants. No trend is recognizable for the positions behind. Especially around the middle seat track position, all types of occupants can be found.

The aggregated data has one major limitation. The moment of assessment of the seat track position is not defined. There is a high chance that first responders had to change the position

to rescue the injured occupant. Consequently, the police recorded position can differ from the real one.

2.1.3 General Passive Safety Assessment Methods

2.1.3.1 Overview

In technical development, the final product is determined with legal, consumer, and internal requirements and validated by tests. However, often the validation is reduced to a couple of discrete example test cases, not necessarily representing the system's overall behaviour. This approach is also applied to the current passive safety assessment of vehicles. Unquestionably, the test cases are carefully selected to represent critical and statistically relevant scenarios. However, especially the point-wise testing as defined by the legislative entities can lead to a local optimization.

Testing, in general, can be understood as the simulation of real-world scenarios. The main assessment methods are either full vehicle crash tests, or sled tests in hardware. Both test types can be mimicked in simulations by virtual models. Virtual methods allow a high degree of modelling flexibility within the expertise of the engineers, especially during the development phase of vehicles. A systematic of testing methods is presented in figure 1.1 on page 2. In the following, the most important testing methods are briefly explained. A more in-depth explanation is provided for Finite Element Analysis (FEA) as this is the main method used in this work.

Based on Multi Body Simulation (MBS) or FEA methods, virtual testing generates faster, and less expensive results than hardware testing. However, the virtual models must be validated by physical testing. Due to this limitation, the potential of virtual methods cannot be fully utilized. The transition from anthropometric data to virtual models plays an essential role in the virtual assessment of the human body's diversity in passive safety systems.

2.1.3.2 Physical Testing

Crash Test Facility

Full vehicle hardware tests require a vehicle with the complete crash relevant structure and interior. The most important components of a test facility are sketched in figure 2.4 on the facing page. Usually, the vehicle under test is pulled towards an obstacle. Alternatively, a trolley is pulled towards the vehicle. For some tests, both crash partners are accelerated.

In the actual test zone, strong lights enable high-speed cameras to record the crash. From a pit, additional cameras capture the deformations of the vehicle's crumble zone. Next to cameras, typical sensors are acceleration sensors, electromechanical transducers, strain gauges, string potentiometers, and paint. The paint is used to identify the contact point of the Anthropomorphic Testing Device (ATD) with the interior and with the restraint system.

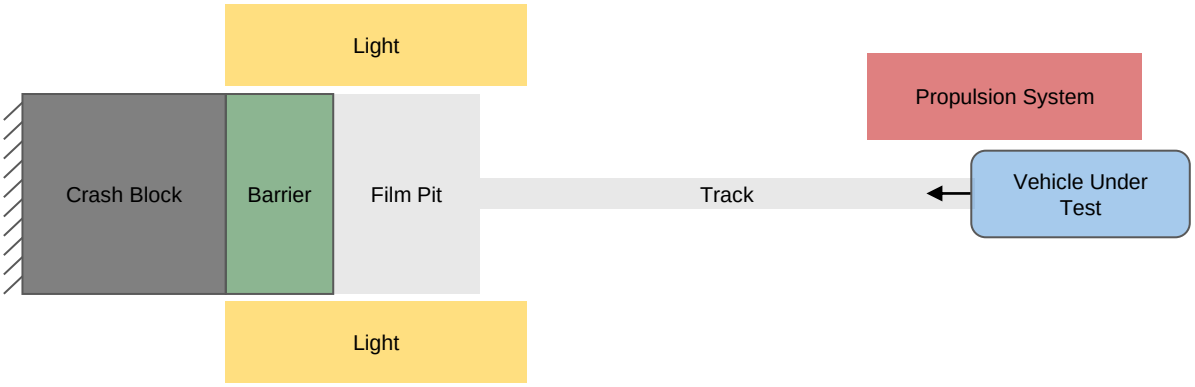


Figure 2.4: Schematic Representation of a Crash Test Facility

The vehicle motion is derived from the high-speed cameras using dynamic photogrammetry. Similar, but usually less accurate results, can be achieved by integration of the acceleration signals from the sensors mounted on the Body In White (BIW).

Sled Test Facility

As a simplification of the full vehicle testing, sled tests are used. Central is the sled, on which a reduced but reinforced BIW is mounted. Moreover, more specialised setups such as the seat alone or a seat with door are used. There are two principals of sled test devices – both are sketched in figure 2.5. The first one mimics a full vehicle crash: the sled is pulled towards an obstacle. The crash structure is replaced by a controlled energy absorption system like stacked metal wires, a servo hydraulic press, or an irreversible deformable element. The second principle is the acceleration sled. Here, the still-standing system is accelerated following the desired crash pulse. While enabling higher test accuracies and coming with lower space requirements, such systems are pricier than the deceleration sleds. Overall, the sensors and their evaluation are the same, as for the full vehicle crash.

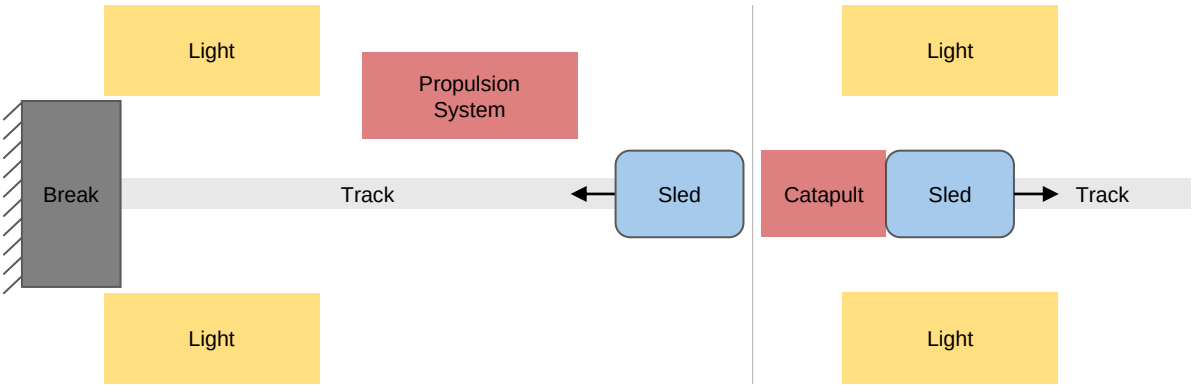


Figure 2.5: Schematic Representation of a Sled Test Facility (Left Acceleration Sled, Right Deceleration Sled)

2.1.3.3 Virtual Testing

Multi Body Simulation

Multi Body Simulation (MBS) model systems use rigid, hence undeformable, bodies with a defined mass and inertia, connected by joints or constraints. Each body in the system has six Degree of Freedoms (DoFs). Constraints deliver additional equations to the base system of equations of motion. In practise, the formulation of Lagrangian dynamics can be used.

For example, for the software MSC ADAMS, McGrath and McConville (1998) formulated the system of equations $\mathbb{E}$ based on the Lagrangian $\mathcal{Z} \in \mathbb{R}$ as

$$\frac{d}{dt}\frac{\partial \mathcal{Z}}{\partial \dot{\boldsymbol{q}}} - \frac{\partial \mathcal{Z}}{\partial \boldsymbol{q}} + \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{q}} \cdot \boldsymbol{\epsilon} = \boldsymbol{p}. \tag{2.1}$$

The Lagrangian is the difference of the total kinetic and potential energy of a multi body system. The generalized coordinates of the system are represented by the vector $\boldsymbol{q} \in \mathbb{E}^{\mathcal{U}}$. If no direct dependencies between some DoFs are considered, $\boldsymbol{q}$ contains all six degrees of freedom for each body in the system, leading to the $\mathcal{U}$ coordinates. The partial derivatives lead to vectors with $\mathbb{E}^{\mathcal{U}}$. The $\mathcal{V}$ equations of the constraints are coupled by $\boldsymbol{\phi} \in \mathbb{E}^{\mathcal{V}}$. Consequently, the partial derivative leads to the matrix of $\mathbb{E}^{\mathcal{V}\times\mathcal{U}}$, which is multiplied by the Lagrangian multipliers $\boldsymbol{\epsilon} \in \mathbb{E}^{\mathcal{V}}$. Ultimately, the generalized forces at the degrees of freedom are denoted by $\boldsymbol{p} \in \mathbb{E}^{\mathcal{U}}$. Additionally, the constraining equations are considered, leading to the complete system of second-order differential equations with $\mathcal{U} + \mathcal{V}$ equations and the unknown $\boldsymbol{q}$ and $\boldsymbol{\epsilon}$.

Those are solved using numerical time integration. Deformable bodies can be modelled using elastically connected small rigid bodies. However, the method does not scale well with an increasing number of bodies, hence limits DoFs of the whole system and constrains. Instead, the strength of the method is in the time domain. Consequently, MBS is used for vehicle dynamics, pre-crash MBS-simulations, or in crash MBS-simulations of highly limited degree of detail as simplified interior.

Finite Element Analysis – Basics

Finite Element Analysis (FEA) adds deformable bodies to the system. A simple example is shown in figure 2.6 on the facing page. Domains with a defined geometry are discretized by small, alias finite, elements. Each element is defined by its corner nodes shared with the neighbour elements. The method provides different level of abstractions, allowing the engineer to make the FE-model as simple as possible but as complex as necessary. For structural conditions, where the length dimension is much larger than the height and width dimensions, beams and truss formulations are used. If the planar size is significant bigger than the thickness, a shell formulation and else so-called solid elements are used.

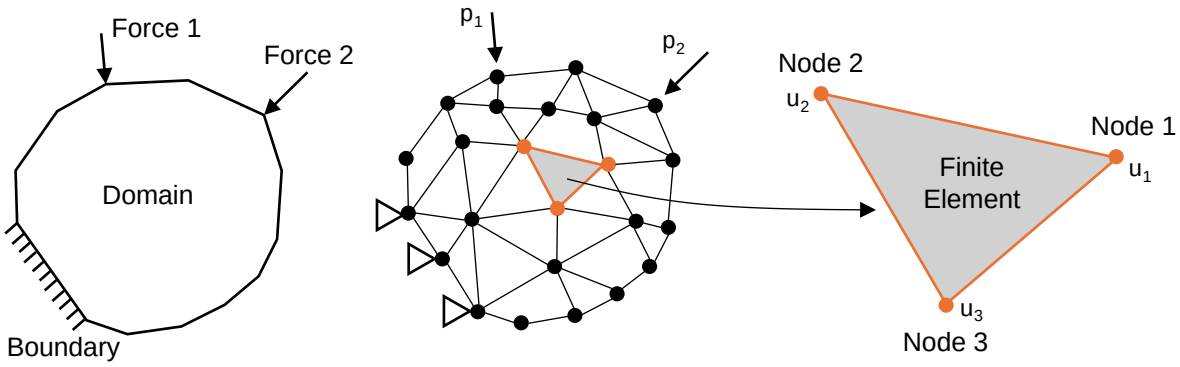


Figure 2.6: Schematic Representation of Finite Element Model in 2D (adapted from Tekkaya and Soyarslan, 2014)

The relation of the nodes of one element is governed by the material law, resulting in internal forces. The resulting mass matrix $\boldsymbol{W}$, the system's damping $\boldsymbol{S}$, the internal stiffness $\boldsymbol{K}$ and external forces $\boldsymbol{p}$ result in the equation of motion

$$\boldsymbol{W} \cdot \ddot{\boldsymbol{u}} + \boldsymbol{S} \cdot \dot{\boldsymbol{u}} + \boldsymbol{K} \cdot \boldsymbol{u} = \boldsymbol{p}, \tag{2.2}$$

being a system of second-order differential equations and connecting the nodal displacement vector $\boldsymbol{u}$ over the time.

The numerical time integration to solve the equation containing the nodal velocities $\dot{\boldsymbol{u}}$ and accelerations $\ddot{\boldsymbol{u}}$ is done either implicitly or explicitly. Using the implicit method, the solution of each timestep contains the information at the timestep itself. The methods try to achieve an equilibrium by iteration until a convergence criterion is met. In the explicit case, the solution of one timestep depends solely on the previous. To achieve accurate results, a rather small integration interval has to be used. On the contrary, no convergence needs to be achieved.

Consequently, implicit time integration usually used for static or processes of low dynamics. Explicit time integration is used for highly dynamic and non-linear systems like vehicle crashes. The popular software – the so-called solver – Livermore Software technology corporation DYNA (LS-DYNA) utilizes the central difference method for the time integration. With explicit time integration, the integration interval is crucial for the numerical stability and computational expense. The smaller the integration interval T, the higher the expense. The maximum integration interval is determined by the system's properties. For example, in LS-DYNA for beam elements T is defined as

$$T \leq L \cdot \sqrt{\frac{\rho}{E}}. \tag{2.3}$$

In the calculation of the maximum integration interval, L is the element's length, ρ its density, and E the elastic modulus. The equation is extended for higher dimensional elements, accordingly. Rigid bodies are not considered in the determination of the critical timestep. In general, the timestep of an element increases with its size and density and decreases with higher stiffness. Another important influence on the actual integration interval is the stability requirement for

the contact algorithm. The integration interval directly influences the computation time, since with a bigger integration intervals, fewer cycles have to be calculated. In summary, a coarser mesh can increase the integration interval and accelerate the FE-simulation. Another side effect of a coarser mesh is the existence of fewer nodes, and thus fewer computations to perform and less data to update and store.

Especially, the explicit time integration enables the FE-simulation of highly non-linear structural problems. Next to large and highly dynamic deformation, material properties such as plasticity, failure, and strain rate dependence can be modelled. The mass-scaling techniques utilizes the density dependency of the critical integration interval in equation 2.3 on the previous page to counter the integration interval reducing effect of compressed elements. Naturally, this method can only be applied, if the density increase does not lead to a mass increase which would significantly influence the overall dynamics.

Finite Element Analysis – Relevant Modelling Techniques for LS-DYNA

The theory of LS-DYNA is provided by ANSYS, Inc. (2024). In general, a FE-model is organized in card definitions – so-called keywords defined by Livermore Software Technology (2020a,b). The nodes are associated with elements. Elements are associated through the "part" definition with a material law and element formulation. Contacts and constraints are defined using "set" definitions that contain groups of nodes, elements, or parts. Finally, the "control" keywords allow the setting of parameters for the solver.

The software stores the **results** in so-called d3plot files. Since the full FE-model information at each timestep can consume large storage, the "database" definitions allow the recording of local information like the kinematics of a node of the force in an element. Such recordings are stored in so-called binout files. In the domain of crash testing, a typical sampling rate is 10 kHz as required, for example, by ISO 18571 (see page 29). The sampling rate of the full model output must be balanced with the available storage which is persistent for storing and temporary for post-processing and the required temporal resolution. In the author's experience, a sampling rate of 0.5 kHz to 1 kHz is a good choice for the given domain.

To compute a FE-simulation, the problem is viewed as temporal and spatial. The temporal progress is achieved by numerical time integration. For numerical time integration, the available number of floating points is relevant – 32 bit is called single **precision** and 64 bit double precision. The precision is associated with computation time, numerical stability, and storage consumption. The spatial decomposition enables the application of **parallelization techniques**. The system's physical problem is solved at each timestep for each decomposed region parallel. More Central Processing Units (CPUs) can lead to a higher degree of parallelization and, hence, to shorter computation time given a sufficient model complexity. The update of the full system equals the overhead and must be balanced with the system's complexity. The parallelization can be done using a shared memory (SMP) as is the case on a workstation, or by distributed CPUs with the respective memory with Message Passing Parallel (MPP) as on a cluster. For crash FE-simulations, the MPP implementation of LS-DYNA is optimized and should be used consequently.

Not to be forgotten should be the computational effort for **pre- and** especially **post-processing** of the FE-simulations. All models have to be generated and transferred from a local workstation to the cluster. After the FE-simulations, the results have to be transferred and stored. This process benefits from smaller file sizes achieved by the careful selection of the stored data and the overall model size. For post-processing, the access to the persistent storage is required. Rapid drives like Solid-State Disks (SSDs) can speed up the post-processing significantly. In addition to appropriate computer-hardware, optimized processing algorithms containing parallelization and vectorization are key.

The central **element types** are solid, shell, and beam. A solid element models a whole volume. The shell element assumes planar conditions with a thickness definition. Beam elements assume axial conditions with a definition of the beam's cross-section. Under the category elements, LS-DYNA also lists spring / damper and seatbelt related definitions.

For each element type, in the so-called "section" definition, various **element formulations** are available. For example, for beam elements, a less complex truss definition can be set if only axial loads are expected.

The **material law** governs the element response under load. Elastic definitions allow linear and non-linear material stiffness, and plastic definition add irreversible deformations. A rigid definition will lock all relative displacements of the nodes within a part to each other. The null material provides a massless element and can be used for contact improvements regarding stability and computational cost for the surfaces of the solid element. Finally, the fabric definition, used for airbags, enables modelling materials where the bending stiffness is of little significance.

In LS-DYNA, two major **separable contact** types exist. In general, the contact algorithm checks if a node is in contact with a surface formed by other nodes and applies some form of penalty related to the material's stiffness. The two-partners-contact allows the definition of two distinct surface groups. The self-contact enables the interaction of all surfaces within a single group. LS-DYNA provides multiple forms of contact treatment in terms of contact detection, contact force application, and initial penetration treatment. For self-contact, LS-DYNA additionally provides an edge detection algorithm, allowing robust but computationally expensive contacts at sharp edges. The initial penetrations are especially relevant for compact assemblies where due to modelling inaccuracies the distance between contact surfaces is lower than the shell thickness. The initial penetrations would lead to initial forces and can be handled by an algorithm adjusting the local thickness or moving nodes. Unlike the initial penetrations, an initial intersection cannot be handled. Since the solver has no information on the directions of the contact partners at the start of the FE-simulation, an intersecting node would be simply treated as being on the other surface side; the two contact surfaces will not separate.

Purposefully **non-separable** contacts are "tied" definitions. The algorithms defined a force element between a node from one contact partner to the surface of another partner. The force is a soft coupling between the two surfaces. Alternatively, a constraint-like formulation can be selected, too.

Similar to the MBS, in FEA, **constraints** can be defined by coupling the kinematics of nodes as well. If the relative motion is completely locked, the definition is called rigid. Joints leave some DoF open, like, for example, the revolute joint allows only relative rotation. Additionally, the kinematics can be prescribed for the whole time domain as a form of boundary condition.

Initial conditions can be modelled in multiple ways. One way would theoretically be the above-explained initial penetration. Other ways are "initial" definitions, such as velocity. Finally, the initial tension can be defined using the concept of "reference geometry". In this definition, the tension free geometry of an element is provided. From the difference between this tension free state and the one at the beginning of FE-simulation, the initial tension is calculated. The concept can be used to model initial tensions in a component from its production. Another application is in airbags. In its folded state, the fabric is heavily deformed and compressed. As explained above, would small elements lead to a small integration interval. In LS-DYNA, the reference geometry can be used instead of the actual element to determine the critical integration interval. To avoid unfolding without inflation, the inner tension can be deactivated too.

The technique of **sub-FE-models** is commonly used for FE-simulations with ATDs. Only for ATD interaction, relevant parts of the complete vehicle model are used. The actual crash is replaced by a prescribed vehicle-pulse and applied intrusions. Three techniques from LS-DYNA are typically used:

- *BOUNDARY_PRESCRIBED_MOTION
- *BOUNDARY_PRESCRIBED_ACCELEROMETER_RIGID and
- *INTERFACE_COMPONENT respective *INTERFACE_LINKING.

For the prescribed motion, a vehicle-pulse at one or multiple vehicle positions must be recorded. The prescribed motion algorithm is applied to a DoF of a node. By the combination of six acceleration signals, a node's complete kinematic over the entire FE-simulation time gets prescribed. If the node is connected to other nodes, or a modified algorithm for rigid bodies is used, whole components rigid body motion can be controlled in this way. The prescribed accelerometer algorithm applies accelerations to a submodel on multiple nodes. The deviation between the different sensor locations is treated by averaging. Nevertheless, the sensor locations should be affected by local deformations as little as possible. The interface functionality enables the recording of deformations in one FE-simulation and the application of it to another FE-simulation. It maps the displacements of single nodes from one FE-simulation to another. In the case of non-matching meshes, the displacement of the constrained nodes will be interpolated from the surrounding nodes of the reference FE-simulation.

The FE-simulation of **airbags** is almost a discipline by itself, as the problem is partially based on structural and partially fluid mechanics. An airbag in LS-DYNA is a volume enclosed by shell elements with a fabric material law. In the simplest case, the pressure P is assumed constant over the volume V and the conditions are governed by the ideal gas law

$$P \cdot V = \frac{m}{\Pi} \cdot \mathcal{R} \cdot \Gamma, \tag{2.4}$$

with Π being the chemical compound dependant molar mass, $\mathcal{R}$ the Molar Gas Constant, and Γ the temperature, and m the mass. A definition of the inflator adds mass to the system, while leakage and venting definitions remove the mass from one timestep to the next. The temperature is handled similar. The volume is coupled to the enclosing fabric and its response to the uniformly applied pressure. The jetting extension limits the area of applied pressure on the airbag fabric to a cone of sight from the inflator. Inhomogeneous conditions, relevant especially during the unfolding of an airbag, can be simulated using the mesh-free particle method. The mass flow can contain a single gas or a mixture (hybrid).

Other software like Virtual Performance Solution (ESI) (VPS) uses similar definitions and strategies to set up a FE-model. Naturally, the exact naming and algorithm implementation vary in detail. The similarities allow the existence of solver-agnostic pre- and post-processors like ANSA / META from BETA CAE Systems or Generator / Animator from GNS GmbH. However, due to differences in detail, the transfer between multiple solvers requires significant modelling and validation efforts.

2.1.4 Human Body Surrogates

2.1.4.1 Introduction

For the assessment of occupant safety, real-world crashes have to be simplified into crash tests. The most common approach is the combination of the tested vehicle with a fixed or movable barrier. Surrogate models like ATDs or purely virtual existing HBMs are used to assess the injury risk. With this, biofidelity, validity, and repeatability are key factors for the meaningful assessment of occupant safety.

Biofidelity in automotive safety describes the degree to which a model represents an actual human. The shape, structure, and biomechanical behaviour can be studied on living subjects. The biofidelity in trauma biomechanics is usually assessed by comparisons with the so-called Post Mortem Human Subject (PHMS). The utilization of PHMS is very limited for ethical reasons, but also due to the limited availability of subjects as well as the potentially non-representativity of the available subjects for a population. Iwamoto et al. (2015) and Kovar and Hlucha (2019) collected published tests relevant for the validation of human surrogate. Richardson et al. (2019) added a test with a reclined occupant.

Validity is understood here as a measure of how well a virtual model represents its physical origin. It is worth noting that validity is only given for the scenario in which the validation was carried out. Hence, the judgement if a model is after a change still valid relies mainly on the expertise of the developing engineers.

Repeatability describes the degree to which the model's estimations deviate if the same experiment is conducted several times. Certification procedures address repeatability. A reproducible test and measurement tool is crucial for an assessment. It seems obvious that neither the developing nor the assessing entities can progress if the repetition of a test results in divergent outcomes.

2.1.4.2 Anthropomorphic Testing Devices

Yang et al. (2006) reviewed the virtual representation of humans for automotive safety published until 2006. They structured the approaches into models using lumped masses, models using linked masses, and FE-models. Furthermore, they distinguished between models for only single regions (e. g. head, neck, thorax) and total body models.

The typical technical representation of the total human body is the Anthropomorphic Testing Device (ATD). The most common ones for frontal crash testing are the members of the Hybrid III (HIII) family. Recently, Euro NCAP introduced in Euro NCAP (2020) the Test Device for Human Occupant Restraint (THOR) 50^{th} percentile to their testing protocols. As ATDs are technical constructions of limited complexity, the validity of virtual twins of ATDs is usually high. Simplified real-world test setups can be mirrored into the virtual models, and separated effects can be compared.

Numerous sensors are located within the technical assembly, which is called an ATD. The HIII measures with around 41 sensors; the newer THOR with over 150 sensors. The output of the sensors are by so-called injury risk curves related to the human biomechanical limitations. Praxl (2011) reviewed the validity of such curves, identifying four factors affecting the reliability of injury risk curves: sampling interval, censoring, test severities, and the statistical model. He points out that a wrong choice of one of those factors, especially the statistical model, can lead to an unrealistic curve that impacts the safety assessment. Often, instead of risk values, absolute limits are used to report. Eppinger et al. (2000) defines injury-criteria for the sensor signals for each ATD.

Since an ATD is a technical assembly, it should be used only for the purpose it was designed for. Similarly, an injury risk curve is constructed in relation to a particular injury mechanism. For that reason, an ATD designed for a front crash should not be used for a lateral impact. Furthermore, an ATD for common upright sitting positions potentially cannot measure adequately with highly reclined backrests. Additionally, different injury mechanisms could be dominant that are maybe not covered by the injury risk curves.

For assessing the interaction of the vehicle with its occupants, few anthropometric configurations are usually tested. The 5^{th}, 50^{th}, and 95^{th} percentiles of ATDs are widely applied assuming a Gaussian distribution of body shapes in a population. However, the human body does not necessarily vary as a whole, but in some parameters (see page 9 ff). The reason for using only those few percentiles can be the limitation of resources, especially for physical testing. Furthermore, the occupants' sitting position can vary due to different anthropometrics, pre-crash manoeuvres, and possible activities while using a vehicle (see page 11 ff).

A statistical representation of a human population is limited by the time when and the geographical region where the data were collected. Restriction to selected percentiles of a population inflicts a limitation in the biofidelity of ATDs. The HIII-50M dates back to Hybrid II, whose anthropometric was based on a survey by Tennant et al. (1974) among the North American population. Schneider et al. (1983) reported that the survey was used for defining the outer dimensions of the HIII family. According to Martin and Shook (2007), the anthropometrics of

the THOR dates back to those studies, too. Interestingly, the Belgium 50th percentile male is near to the HIII anthropometrics (see table 2.1 on page 10 and figure 2.8 on page 28).

2.1.4.3 Update of an Anthropometrical Testing Device on the Example of HIII-50M and THOR-50M

The focus of this work is the frontal crash for vehicle passengers in the first row. Recently, multiple organizations, for example CATARC (2024) and Euro NCAP (2023a,c), replaced the widely used HIII-50M ATD with the newer Test device for Human Occupant Restraint, 50th Percentile Male (THOR-50M). After the initial efforts of the NHTSA (2015), the NHTSA (2024) announced the incorporation of THOR-50M to FMVSS 208 for 2025. Naturally, industry and research compared the responses of both surrogates in the crash scenarios.

An extensive comparison was published by Xu et al. (2000). The authors analysed the responses in pendulum tests as well as in full vehicle tests. As a major limitation of the study, the THOR-50M version should be mentioned. The used alpha version represents an early stage of the development.

In the pendulum tests, higher oscillations of the head acceleration of the THOR-50M were found. Furthermore, the peak chest acceleration was reported 25 % to 50 % higher than the acceleration of the HIII-50M. Moreover, the pelvis acceleration showed stronger oscillations and significantly higher peaks.

In the vehicle tests, approximately 10 g higher maximum head accelerations of the THOR-50M were measured. The maximum chest deflection was found to be similar. In the author's experience, the chest deflection of both ATDs is comparable only to a limited extent. While HIII-50M evaluates the deflection at a single point, for THOR-50M four points are used. Further findings of Xu et al. (2000) were higher loads in the lumbar spine of the HIII-50M. For this reason, the overall stiffer spine was identified. THOR-50M has deformable rubber blocks in multiple places along the spine. In comparison, the HIII family was designed with one single big rubber block. Regarding the lower extremities, the HIII-50M reacted with rougher signals to the vehicle crash. The authors provided the overall softer knee of the THOR-50M as a reason. However, no clear trend was identified between both signals regarding their magnitude.

The newer THOR-50M was compared with the HIII-50M under frontal test condition by Albert et al. (2023). They observed similar seatbelt B3-forces (see figure 2.1 on page 8) and reasoned them by the similar geometry and dominance of the retractor for both ATDs. However, the HIII-50M reported higher lap seatbelt forces. In contrast, they measured higher forces on the knee bolster for THOR-50M. The foot support forces were found to be similar. The study of Albert et al. (2023) was intended as a biofidelity comparison of both ATDs. Hence, most analysis compared the individual surrogate with the PHMS test. The ISO-rating was calculated for multiple responses. Unfortunately, since the scope was the biofidelity assessment, the scores for a direct ATD to-ATD comparison were not provided.

O'Connor et al. (2022) conducted a database-driven study. The authors compared data from crash tests with THOR-50M and HIII-50M with real word accidents from the CISS database

from the US. In addition, they compared their findings with those from the literature. Geometrical comparisons of both ATDs and comparisons of their responses in crashes were performed.

Given the same seat adjustment, the head of the THOR-50M was found to be ca. 50 mm rearward (away from the steering wheel) and ca. 30 mm higher than the head of the HIII-50M. At the same time, the H-Point (equals approximately the Centre of Gravity (CoG) of the pelvis) of the THOR-50M was approx. 16 mm more forward and 30 mm higher. The knees of the THOR-50M were measured in average ca. 50 mm closer to the instrument panel. The seatbelt B3-forces were found to be similar. The maximum chest accelerations were found to be comparable. The Head Acceleration over 3 ms Criterion (HAC_3) was reported 25 % higher in average for THOR-50M than for HIII-50M. The Neck Injury Criterion (Nij) showed an 73 % higher average and the Neck Tension Force Criterion (NTC) 24 %. The Femur Compression Force Criterion (FCC) were 93 % to 179 % higher for THOR-50M.

O'Connor et al. (2022) reasoned the observed differences and compared them with the literature. All trends were reported to be in agreement with the literature. The vertical difference of the head position was explained by the different sitting heights of both ATDs with 907 mm for the THOR-50M and 884 mm for the HIII-50M. The more forward head position of the HIII-50M seemed plausible, as its spine's overall stiffness is lower. Hence, the back can follow the seat backrest less. The different H-Point positions were not discussed. However, they can be expressed as distance to the cushion instead. Due to the overall thicker pelvis of the THOR-50M, its H-Point will be higher and more forward. The more forward knee positions of THOR-50M were explained by the combination of the pelvis position and the length of the upper leg. The head and neck loads of the THOR-50M were assumed to be higher due to higher initial distance of the head to the DAB. The higher femur loads were connected with the initial knee position.

In general, similar trends were found in the different studies. Although, the investigation of Xu et al. (2000) diverged because an older version of THOR-50M was used. Combining the reported results and the author's experience, the following set of expected differences between THOR-50M and HIII-50M were assembled given a matched seating procedure. The findings are summarized in table 2.2 on the next page. The arrow direction in the position columns indicates the direction of the difference. Pointing to the left is understood as towards the instrument panel, and pointing up as a vehicle's upward direction. The arrows in the load column are related to the maximum accelerations for the accelerometers. The arrow directions indicate the qualitative direction of the change in the observed values. Due to the different chest designs of both ATD, the chest deflection was left aside. Overall, the literature reports higher maximum chest deflections of the THOR-50M and reasons them by a softer thorax design and more sensitive sensors.

Table 2.2: Expectable Differences of an Update from HIII-50M to THOR-50M

	Position		Load
	X-direction	Z-direction	
Head Accelerometer	→ ca. 50 mm	↑ ca. 30 mm	↑
Neck			↑
Shoulder Belt			↔
Chest Accelerometer			↔
Lumbar Spine			↓
Lap Belt			↓
Pelvis	← ca. 15 mm	↑ ca. 30 mm	↑
Femur			↑
Knees	← ca. 50 mm	equal	↑

2.1.4.4 Human Body Models

A purely virtual approach to model occupants is via HBMs. An HBM is not limited by technical restrictions like ATDs. Of course, the degree of detail is limited by the targetted computational cost. The numerical model offers the possibility to define its behaviour precisely like in a human being. Extensive validation to PHMS has to be done to achieve this goal. In contrast, Schneider et al. (2017) stated that such validations are limited since every human is unique and so every PHMS is unique as well. Unlike the validation procedure for virtual ATDs, single component tests and complete assembly tests cannot be based on the same subject. Hence, a statistically significant number of tests must be carried out to get a well-validated model. This is contradicted by the limitations of PHMS testing stated above.

For human body modelling, MBS and FEA are used. Siemens Digital Industries Software (2024a), provides a MBS model focusing on pre-crash assessment and crash analysis. A hybrid approach to MBS modelling in FEA was used by Vychytil et al. (2014) to create VIRTHUMAN (VHUM). The objective of the developments was low computational cost and straightforward scalability. Kovar and Hlucha (2019) reviewed the status of VHUM and claimed it to be fully validated. Based on CT scans, the following FEA full-body models were created: (i) the HBM of the Global Human Body Models Consortium (GHBMC), introduced by Gayzik et al. (2011) and Schwartz et al. (2015), (ii) the Total Human Model for Safety (THUMS), described by Shigeta et al. (2009), and (iii) the HBM of Virtual Vehicle Safety Assessment (ViVA), introduced by Östh et al. (2017). In figure 2.7 on the following page, the MBS model of MAthematical DYnamic MOdels (MADYMO), the hybrid FE-model, and the highly detailed THUMS are compared.

The THUMS FE-models are widely used in research and industry. Shigeta et al. (2009) described the THUMS v4 with detailed organs and approx. 2 000 000 elements. The external anthropometric data were derived from the HIII ATD defined in Robbins (1983). THUMS FE-models exist for the whole range of the HIII family, ranging from 6-year-old to 95th percentile.

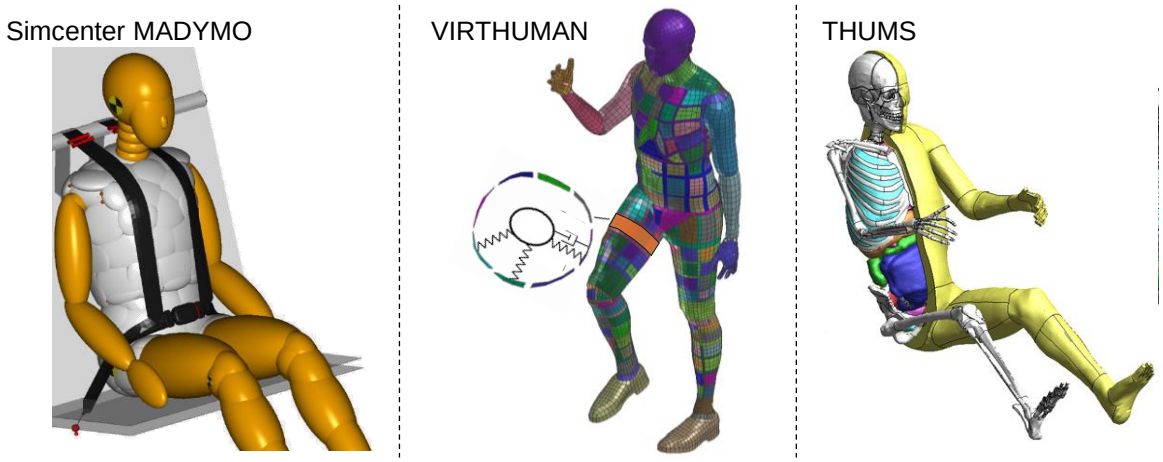


Figure 2.7: Examples of Human surrogate (from left, adapted from Siemens Digital Industries Software, 2024b, Kovar and Hlucha, 2019, Toyota Motor Corporation, 2011)

The GHBMC FE-models were created in 2006 and are widely used as an alternative to THUMS. Analogously, two different degrees of details were developed as well as the typical anthropometrical sizes.

The ViVA FE-model attempts open-source development. Unlike GHBMC and THUMS, ViVA represents an average female. In this context occurs the question for a reason of the selected gender of ATDs and the other HBMs 50th percentiles. The decision was reported by Schneider et al. (1983). In the report, four anthropometrical configurations originating from National Health and Nutrition Examination Survey (NHANES) data from the 1970s were suggested; the 5th percentile female, the 50th percentile male and female, and the 95th percentile male. Although the 5th and 95th percentiles covered the overall population's outer border, the 50th percentiles for the male and female populations were significantly different (table 2.1 on page 10). Schneider et al. (1983) stated that the funding was not enough for the development of ATDs in four configurations and, hence, the 50th percentile female was excluded.

HBMs are an active field of research. E. g. THUMS entered Open Access by Toyota (2020) and was recently released by Toyota (2023) as v7 for reclined postures. Other noteworthy HBMs are:

- The CHARM-70F and the CHARM-10 described by Jin et al. (2018) representing a 70-year-old female,
- The CAVEMAN focused on military application depicted in Butz et al. (2020),
- The Ford HBM in El-Jawahri et al. (2010),
- And the Human Model for Safety (HUMOS) and HUMOS 2 introduced by Robin (2001) and Vezin and Verriest (2005) respectively,
- The HANS FE-model developed in particular for industrial use from DYNAmore GmbH (2023), and the HBM Connect by Humanetics Innovative Solutions Inc (2024b).

Relevant publications with data on tests with PHMS can be found in the validation reports of various HBMs. Beillas and Berthet (2018) reviewed the development and validation of spleen, liver, and kidney in the GHBMC. They concluded that the selection of the specimens significantly influences the selection of material properties in the virtual model. Recent experiments address the upcoming new sitting positions in the context of automated driving.

Virtual models possess the ability to be geometrically scalable. For ATDs this is usually not done as they mirror their exact hardware twin. On the other hand, HBMs are commonly scaled. However, especially for comparability, HBMs are often published related to the established percentiles. Additionally, sensors similar to the ATD or FEA data could be used in HBMs. An example of more profound insights from the FEA data is the tension within the tissues.

In figure 2.8 on the following page for the assessment of frontal crashes, commonly used human surrogates are collected. Some models are only available with sitting pelvis, and thus, no exact standing height data was accessible. For the HIII family, the data were directly measured from the FEA model used in the thesis work. To measure standing height, the limits of the leg joint had to be violated. For the THORs, the approximations from National Highway Traffic Safety Administration (2022) were used. The ATDs weights were taken from Humanetics Innovative Solutions Inc (2024a). It must be considered that the specifications of an ATD are with tolerances. For example, the Dummy Testing and Equipment Committee (2022) defined up to ±15 mm tolerance for length measures. The data of THUMS was published by Kato et al. (2018), the data of VHUM by Kovar and Hlucha (2019), and the data of GHBMC by Park et al. (2013). The data for the ViVA and SAFER HBMs were summarized by Li et al. (2023) and used those base anthropometries for further studies, including scaling of the models. In figure 2.8 on the next page, additionally, the population data from table 2.1 on page 10 for North America (NA) and Malaysia were added.

In summary, the classic ATDs get increasingly accompanied by HBMs. Recent developments like the Biofidelic Dummy as described by Härtel (2024) have the potential to bridge both worlds. Despite a validated HBM can provide great insight into the injury mechanisms active during a car crash, their computational cost is significantly higher.

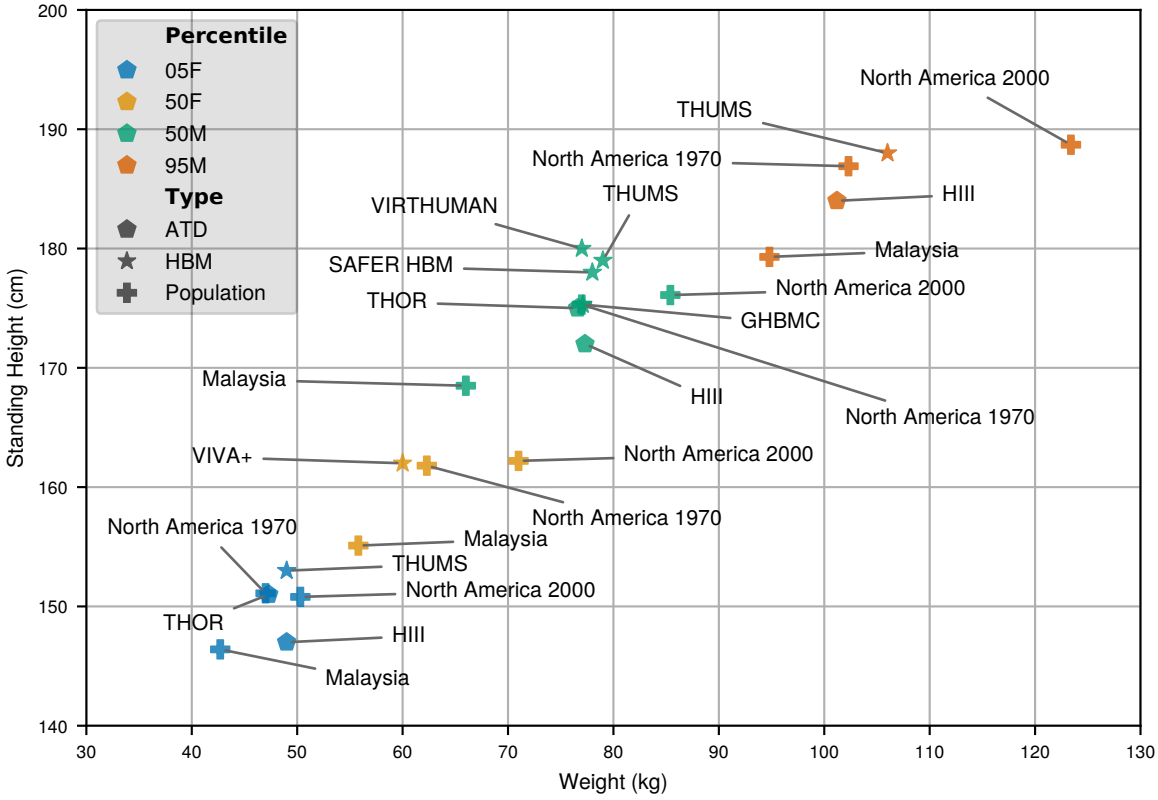


Figure 2.8: Standing Height and Weight of Selected Populations and Human surrogate

2.1.5 Typical Requirements on Passive Safety of Vehicles

2.1.5.1 Standards

Overview

Not obligatory by itself, but often referred in the requirements from legislatives and consumer organizations, are standards as they are defined, for example by International Organization of Standardization (ISO) and Society of Automotive Engineers (SAE). Moreover, standards reflect state-of-the-art and industry best practice. The following applications relevant for this work are illustrated: SAE J211 respectively ISO 6487 for signal filtering, ISO 18571 for signal comparison, and ISO 13499 for sensor channel naming conventions.

CFC Signal Filtering by SAE J211 and ISO 6487

Sensor signals in the passive safety domain are usually filtered according to SAE J211 from the Safety Test Instrumentation Standards Committee (2022). The Channel Frequency Class (CFC) filters are industry standard. It is defined as a digital, low-pass Butterworth filter of second order. The filter is passed twice, once forward and once backwards, resulting effectively in a

filter of fourth order and eliminating the phase shift from a single pass. The filter's differential equation

$$\breve{z}_t = a_0 \cdot z_t + a_1 \cdot z_{t-1} + a_2 \cdot z_{t-2} + b_1 \cdot \breve{z}_{t-1} + b_2 \cdot \breve{z}_{t-2}, \tag{2.5}$$

maps the padded discrete signal $\boldsymbol{z}$ over the time t to the filtered signal $\breve{\boldsymbol{z}}$ with T being the sampling interval. The filter coefficients $\boldsymbol{a}$ and $\boldsymbol{b}$ are defined by

$$a_0 = \frac{\omega_a^2}{1 + \sqrt{2} \cdot \omega_a + \omega_a^2}, \tag{2.6}$$

$$a_1 = 2 \cdot a_0, \tag{2.7}$$

$$a_2 = a_0, \tag{2.8}$$

$$b_1 = \frac{-2 \cdot (\omega_a^2 - 1)}{1 + \sqrt{2} \cdot \omega_a + \omega_a^2}, \text{ and} \tag{2.9}$$

$$b_2 = \frac{-1 + \sqrt{2} \cdot \omega_a + \omega_a^2}{1 + \sqrt{2} \cdot \omega_a + \omega_a^2}. \tag{2.10}$$

Within the coefficients, the angular cut-off-frequencies are being defined as

$$\omega_a = \tan\left(\omega_d \cdot \frac{T}{2}\right), \text{ and} \tag{2.11}$$

$$\omega_d = 2\pi \cdot N \cdot 2.0775, \tag{2.12}$$

with N indicating the filter class.

The standards SAE J211 and ISO 6487 from International Organization for Standardization (2015) both define the filter that only diverge in equation 2.12. While SAE uses the factor 2.0775, the ISO suggests $^{1.25}/_{0.6}$, where 0.6 transforms the 1 dB cut-off frequency to −3 dB. The 1.25 scales the single filter to 1.5 dB, resulting in 3 dB in total after the forward- and backward-passes.

Typical values for N are 60, 180, 600 and 1000. The filter class to be used is defined in the respective testing protocol. By convention, the filter class is denoted by indicating CFC in combination with N like CFC 60.

An implementation of the filter can be found in the code base of Plaschkies (2024a). The PYTHON script was obfuscated. Obfuscation allows the distribution of a script in a binary, not human-readable format.

Objective Signal Comparison by ISO 18571

The procedure defined by International Organization for Standardization (2014) as ISO 18571 describes a methodology for validation metrics and the calculation of the correlation of two bijective signals. The standard is the successor of the CORrelation and Analysis (CORA) and Enhanced Error Assessment of Response Time Histories (EEARTH) ratings. At the time of this thesis, an update of ISO 18571 was about to be released. However, the following explanations relate to the version of 2014.

A free usable implementation of the method is CORAplus by Thunert (2017). The distribution comes with detailed explanations of the methodology. Open source is implementation by Graz University of Technology (2023) on OpenVT. Proprietary implementations can be found in most test evaluation and post-processing tools like Meta from Beta and Animator from GNS.

The calculated ISO-rating reports on a sliding scale with defined meaning, as displayed in table 2.3. The ISO-rating take values from 0 to 1, where 1 is the best. Values smaller than 0.5 are considered not interpretable. One signal is called the reference, the other the test. The ISO-grade translates the ISO-rating into Excellent, Good, Fair and Poor.

Table 2.3: Sliding Scale of the Overall ISO-rating, adapted from ISO 18571 (2014)

ISO-rating Range	Grade	Description
$0.94 <$ ISO-rating ≤ 1.00	Excellent	Almost perfect characteristics of the reference signal is captured
$0.80 <$ ISO-rating ≤ 0.94	Good	Reasonably good characteristics of the reference signal is captured, but there are noticeable differences between both signals
$0.58 <$ ISO-rating ≤ 0.80	Fair	Basic characteristics of the reference signal are captured, but there are significant differences between the two signals
$0.00 <$ ISO-rating ≤ 0.58	Poor	Almost no correlation between the two signals

The **ISO-rating** is composed of four sub-ratings: corridor, phase, magnitude, and slope. The weighted arithmetic mean of the sub-rating form the ISO-rating. For processing, the starting time of both signals must be the same. The rating was calibrated for a 10 kHz sampling rate of the signals. To avoid artefacts, the signal filter strategy is important. The evaluated time interval should be limited to the relevant validation time window.

For the **corridor-sub-rating**, an inner and outer corridor is formed. The width of the corridors is constant and calculated from the maximum absolute value of the reference signal in the evaluated time interval. Each timestep of the test signal, is rated with 1 if within the inner corridor, 0 if outside the outer corridor, or interpolated between 0 and 1 if in the outer corridor. The arithmetic mean of all timestep ratings form the corridor-rating.

To calculate the **phase-sub-rating**, the test signal is shifted in time, until the maximum cross-correlation between both signals is reached. From the shifted time delta, the phase-rating is derived. The shifted, and truncated signal is used to calculate the last two sub-ratings.

Dynamic Time Warping (DTW) is used to calculate the **magnitude-sub-rating**. For dynamic time warping, the Euclidean distances of all timestep values of one signal to all timestep values of the other signal are calculated as the first step. In the resulting distance matrix, the upper-left corner corresponds to the start of both time signals and the lower-right corner corresponds to both ends. The objective is now to find the most cost-efficient path – the so-called warping path – from the upper-left to the lower-right corner, where a step's cost is the distance value of the respective field. The normalized length of the warping path is translated into the magnitude-sub-

rating. The dynamic time warping approach can be interpreted as a dynamic local time-scaling of signals to match their patterns.

Ultimately, the **slope-sub-rating** evaluates the differences in slopes between the two signals within 10 timesteps of the 10 kHz sampling rate. The average over the individual slope differences leads to the sub-rating.

For simple integration and fast computation for the presented research, the author implemented the procedure in PYTHON. In addition to NUMPY from Harris et al. (2020) and SCIPY by Virtanen et al. (2020), the package DTWALIGN from Suzuki (2022) was used. To validate the implementation and compare it to others, the digital attachment of ISO 18571 was used. In the ISO 18571, reference signals are provided for 4 accelerations, 2 angular displacements, 2 displacements, 3 forces, and 3 moments with 3 test cases each. For table 2.4 the ratings for the in total 42 cases were calculated and subtracted from the ratings given in the ISO 18571. The displayed values are the average over the cases of absolute differences, forming the Mean Absolute Error (MAE) (see section 2.2, equation 2.44 on page 44).

Table 2.4: Mean Absolute Error of Different Implementations to ISO 18571 Reference

Rating	CoraPlus	OpenVT	Own
ISO-rating	0.009	0.009	0.009
Corridor	0.000	0.001	0.001
Phase	0.019	0.018	0.021
Magnitude	0.006	0.008	0.006
Slope	0.023	0.023	0.019

For all evaluated implementations, there was an unusual deviation from the reference for the phase-sub-score of the signal called “Moment 2 – CAE1”. This signal is one of the provided examples for a torque signal. Cross-checking the value with the pre-released 2024 version of the standard, resulted in correcting the reference value of 2014 from 0.000 to 0.497. The new value was in accordance with the evaluated implementations. After corrections, all implementations are in good agreement with the reference. The adjustment is already considered in table 2.4. In tendency, the MAE of phase- and slope-ratings were observed higher while not contributing with 40 % to the total rating. Furthermore, higher MAEs were observed for reference rating lower than 0.5.

An implementation of the ISO-rating can be found in the code base of Plaschkies (2024a). The PYTHON script was obfuscated.

Naming Conventions by ISO 13499

ISO 13499 from International Organization for Standardization (2019) describes the conventions for data exchange in the context of vehicle crash testing. The standard features a folder structure to store meta- and sensor-data, file formats, and naming conventions, so-called ISO-codes.

In this work, the conventions of naming sensor signals, the so-called channel codes, are relevant. A channel code is 16 characters long and composed of a sequence of codes with fixed length and specific order, defining the test object, position, main location, fine locations, physical dimension, direction, and filter class. The fine location three was extended by two characters to specify the occupant's percentile as. The full convention as used in this work is presented in table 2.5.

Table 2.5: ISO MME Definitions as Used in this Study

Position	Example	Meaning	Options (Examples)
1	0	Test Object	0 – None
2	3	Position	1 – Driver 3 – Passenger
3	T	Main Location	HEAD – Head NECK – Neck CHST – Chest PELV – Pelvis FEMR – Femur TIBI – Tibia
4	I		
5	B		
6	I		
7	U	Fine Location 1	UP – Upper LO – Lower 00 – None
8	P		
9	L	Fine Location 2	LE – Left RI – Right 00 – None
10	E		
11	H	Fine Location 3	VH00 – Vehicle H350 – HIII-50M VH50 – VHUM-50M TH50 – THOR-50M OCCU – Occupant
12	3		
13	5		
14	0		
15	F	Physical Dimension	DS – Displacement / Deflection in [mm] VE – Velocity in [m s^{-1}] AC – Acceleration in [g] FO – Force in [kN] MO – Moment / Torque in [N m]
16	0		
17	Z	Direction	X – X-Direction Y – Y-Direction Z – Z-Direction R – Resultant-Direction
18	D	Filter Class	X – None A – CFC 1000 B – CFC 600 C – CFC 180 D – CFC 60

2.1.5.2 Requirements by Legislatives and Consumer Organizations

Overview

The development process in automotive passive safety is determined by fulfilling mandatory requirements in the vehicle approval procedures. For the European market, the type-approval and third-party certification procedure required by European and national legislatives, mostly refers to the regulations of the United Nations Economic Commission for Europe (UNECE). Although the UN regulations as such are not binding, they become mandatory once ratified on the national or EU level. In comparison, mainly for the US-American market, the requirements for self-certification are defined by the National Highway Traffic Safety Administration (NHTSA) in the FMVSS.

Consumer or insurance-oriented testing organizations such as the Euro NCAP significantly influence vehicle development requirements. Those organizations are defining test cases and rating criteria to motivate the manufacturers to develop safer vehicles. The results of those frameworks do not compromise the legal approval status of a vehicle. However, the market influence is high. Not being approval-relevant enables being the driver for new requirements aiming to increase the safety in mobility.

The passive safety assessment methods of the legislative entities and consumer testing organizations mentioned above are based on physical experiments, especially crash tests. An exception is the assessment of pedestrian safety by Euro NCAP through virtual testing defined by Euro NCAP (2023e). A HBM can, after its certification as described by Euro NCAP (2023d), be used to assess pedestrian safety. As by Euro NCAP (2022), it is planned to extend this procedure to other domains. However, as by the author's impression from discussions with industry experts, the real utilization of such approaches seems to remain limited.

Next to the potential cost reduction during the development, the new crash scenarios resulting from Advanced Driver Assistance Systems (ADAS) related actions are considered driver for virtual assessment activities. Larsson et al. (2019) analysed the impact of lane-change manoeuvre in the pre-crash phase. The manoeuvre can change the occupant's nominal sitting position. The assessment of passive safety with a manoeuvre-dependent initial position of the occupant in the crash phase leads to a drastically increased parameter space. Furthermore, the previously discussed potential new sitting positions and activities in the context of automated driving functionalities (see page 9) will add to the parameter space. The acceptance of virtual methods can help to make the testing within such a parameter space feasible.

Evaluated are sensor outputs and pre-defined behavioural characteristics before and during the crash. The evaluation of sensor output is summarized above for human surrogates. An overview of the criteria used is provided by *Safety Companion* (2024) and an in-depth discussion by Schmitt et al. (2019). Not directly sensor-related criteria are e. g. the airbag slap and submarining. The airbag slap evaluates whether the airbag front interacts with the occupant while unfolding. A very high relative velocity can cause injuries to the face. Submarining describes the loss of coupling between the lap seatbelt and the pelvis. This effect leads to less restraint and can influence the abdomen region by extensive compression.

Other criteria include the intrusions into the passenger compartment and the leakage of fuel or electrolytes. Furthermore, the accessibility of a crashed vehicle for first responders and the availability such as rescue sheets are incentivised.

Selected Injury Criteria

In the Technical Bulletin number 21 from Euro NCAP (2023b), the Euro NCAP relevant data formats and injury-criteria are defined. Then, the criteria for the adult occupants in the first seat row for the HIII ATD limited to Hybrid III 5$^{\text{th}}$ Percentile Female Dummy (HIII-05F) and HIII-50M percentiles shared criteria. The Technical Bulletin requires the naming and file structure conventions according to ISO 13499. For filtering the signals of each sensor, a specific CFC is prescribed. The CFC method implies a sampling rate of at least 10 kHz for the $\mathcal{D}$ timesteps. An implementation of the injury-criteria calculation can be found in the code base of Plaschkies (2024a).

For the head acceleration related criteria, the resultant acceleration

$$ {}^{\text{Head}}\breve{\ddot{U}}_{R,t} = \sqrt{{}^{\text{Head}}\breve{\ddot{U}}^2_{X,t} + {}^{\text{Head}}\breve{\ddot{U}}^2_{Y,t} + {}^{\text{Head}}\breve{\ddot{U}}^2_{Z,t}} \tag{2.13} $$

is used. For the directional acceleration signals ${}^{\text{Head}}\breve{\ddot{\boldsymbol{U}}} \in \mathbb{R}^{4\times\mathcal{D}}$, the filtering by CFC 1000 is prescribed to be applied on the X-, Y-, and Z-components first. The unit of acceleration is $[g]$. The **Head Acceleration over 3 ms Criterion (HAC_3)** in $[g]$ is defined by

$$ \text{HAC}_3 = \max_{t=[t_1,t_2]} \left\{ {}^{\text{Head}}\breve{\ddot{U}}_{R,t} \right\}, \text{ with } t_2 - t_1 = 3\,\text{ms}, \tag{2.14} $$

as an exceeded level of the resultant head acceleration within the time window of 3 ms. The **Head Injury Criterion for up to 15 ms (HIC_{15})** in the form of

$$ \text{HIC}_{15} = \max_{t_2-t_1\leq 15\,\text{ms}} \left\{ (t_2 - t_1) \cdot \left[\frac{T}{2\cdot(t_2-t_1)} \cdot \sum_{t=t_1}^{t_2-1} \left({}^{\text{Head}}\breve{\ddot{U}}_{R,t} + {}^{\text{Head}}\breve{\ddot{U}}_{R,t+1} \right) \right]^{2.5} \right\}, \tag{2.15} $$

usually treated as dimensionless, but actually has the unit $[\text{s}\,g^{2.5}]$. It is computed as the maximum value for all possible time intervals in [s] between t_1 and t_2, with any span up to 15 ms. The time-integral for the discrete signal is resolved using the trapezoidal-rule with a constant sampling interval T.

The **Neck Extension Moment Criterion (NEC)** in [N m] is calculated by

$$ \text{NEC} = \left| \min_t \left\{ {}^{\text{Upper Neck}}\breve{M}_{Y,t} \right\} \right|, \tag{2.16} $$

the **Neck Tension Force Criterion (NTC)** in [kN] by

$$ \text{NTC} = \max_t \left\{ {}^{\text{Upper Neck}}\breve{F}_{Z,t} \right\}, \tag{2.17} $$

and the **Neck Shear Force Criterion (NSC)** in [kN] by

$$\mathrm{NSC} = \max_t \left\{ {}^{\text{Upper Neck}}\breve{F}_{X,t} \right\}. \tag{2.18}$$

The raw signals of the torque ${}^{\text{Upper Neck}}\boldsymbol{M} \in \mathbb{R}^{4\times\mathcal{D}}$ and the force ${}^{\text{Upper Neck}}\boldsymbol{F} \in \mathbb{R}^{4\times\mathcal{D}}$ must be filtered by CFC 600 and measured at the ATD upper neck bracket. Positive shear force indicates a head-rearwards motion, and positive axial force indicates a tensile force in the neck.

The **Chest Deflection Criterion (CDC)** in [mm] is calculated by

$$\mathrm{CDC} = \max_t \left\{ {}^{\text{Chest}}\breve{D}_{X,t} \right\}, \tag{2.19}$$

using the CFC 180 chest deflection. For the HIII family, the chest deflection is not measured directly as a translation. Instead, a cantilever connects the spine box with the sternum plate. The relative motion of the sternum to the spine is recorded as an angle δ in [rad]. By applying a ATD specific linearization factor, the angle can be translated into the deflection. In the manual for the HIII FE-models from Guha et al. (2011), linearization factors

$${}^{\text{Chest}}D_{X,t} = \delta(t) \cdot \begin{cases} 96\,\mathrm{mm\,rad^{-1}} & \text{for HIII-05F} \\ 145\,\mathrm{mm\,rad^{-1}} & \text{for HIII-50M} \\ 158\,\mathrm{mm\,rad^{-1}} & \text{for HIII-95M} \end{cases} \tag{2.20}$$

were defined.

The **Chest Viscous Criterion (CVC)** in [m s^{-1}] is calculated by

$$\mathrm{CVC} = 1.3 \cdot \max_t \left\{ {}^{\text{Chest}}\breve{\dot{D}}_{X,t} \cdot {}^{\text{Chest}}\breve{\breve{D}}_{X,t} \right\}, \tag{2.21}$$

using an element-wise product where ${}^{\text{Chest}}\dot{\boldsymbol{D}}$ is the chest deflection velocity and ${}^{\text{Chest}}\breve{\boldsymbol{D}}$ the normalized chest deflection. The scaling factor is defined for the HIII family with 1.3.

The chest compression velocity is the derivative of the chest deflection and calculated from the padded discrete signal using a 5-point-stencil, defined as

$${}^{\text{Chest}}\breve{\dot{D}}_{X,t} = \frac{8\left[{}^{\text{Chest}}\breve{D}_{X,t+1} - {}^{\text{Chest}}\breve{D}_{X,t-1}\right] - \left[{}^{\text{Chest}}\breve{D}_{X,t+2} - {}^{\text{Chest}}\breve{D}_{X,t-2}\right]}{12 \cdot T}, \tag{2.22}$$

The normalized chest deflection is defined by

$${}^{\text{Chest}}\breve{\breve{D}}_{X,t} = {}^{\text{Chest}}\breve{D}_{X,t} \cdot \begin{cases} 0.187^{-1} & \text{for HIII-05F} \\ 0.229^{-1} & \text{for HIII-50M} \\ 0.254^{-1} & \text{for HIII-95M} \end{cases}. \tag{2.23}$$

The normalization value is specified for HIII-05F and HIII-50M by Euro NCAP. Additionally, the value for the Hybrid III 95$^{\text{th}}$ Percentile Male Dummy (HIII-95M) was provided by Mertz (2002).

Ultimately, the **Femur Compression Force Criterion (FCC)** in [kN] is calculated by

$$FCC = \max\left\{\left|\min_t\left\{{}^{\text{Left Femur}}F_{Z,t}\right\}\right|, \left|\min_t\left\{{}^{\text{Right Femur}}F_{Z,t}\right\}\right|\right\} \tag{2.24}$$

from the left and right femur axial forces by CFC 600 filtered.

In table 2.6, the key values for the above outlined injury-criteria are displayed. The thresholds are defined by Euro NCAP (2023a) in the assessment protocol for adult occupant protection. A criterion is graded with 4 points if its value is less than the given threshold. The grade 0 point is given if the value exceeds the respective threshold. For values in between both thresholds, the points are interpolated.

The 0 point threshold for HIII-95M in table 2.6 is an extrapolation from the HIII-05F and HIII-50M ATDs, assuming equal distance between them and the HIII-95M. This approach has no connection to biomechanical limits, and hence its purpose can be orientation alone.

Table 2.6: Injury Limits by Euro NCAP for Frontal Impact Crash Scenarios

			HIII-05F		HIII-50M		HIII-95M
		Unit	4 Points	0 Points	4 Points	0 Points	0 Points (Extrapolated)
Head	HIC_{15}		500	700	500	700	700
	HAC_3	g	72	80	72	80	80
Neck	NEC	N m	36	49	42	57	65
	NTC	kN	1.7	2.62	2.7	3.3	3.98
	NSC	kN	1.2	1.95	1.9	3.1	4.25
Chest	CDC	mm	18	34	22	42	50
	CVC	$m\,s^{-1}$	0.5	1.0	0.5	1.0	1.0
Femur	FCC	kN	2.6	6.2	3.8	9.07	11.9

Crash Severity Estimation

The crash severity is ultimately reflected as the occupant's injury severity, as defined above. However, the assessment using human surrogates can be costly and time-consuming. Furthermore, the complex interaction of surrogate, interior, and restraint system can obfuscate the relevant mechanism in e. g.the vehicle deformation. Consequently, other indicators for the crash severity were developed. A comprehensive overview is presented by Schöneburg (2023).

From a physical perspective, a crash is a collision in which the kinetic energy is partially transformed and dissipated by deformation and heat and partially conserved in the form of a rebound motion. Simple indicators are the change of velocity representing the total energy and the maximum or average acceleration at a sensor point in the vehicle during a defined time window.

Another approach – the Ride-Down-Effect (RDE) – evaluates the crumble zone of the vehicle by calculating the remaining deformation zone of the vehicle at the moment of a coupled

occupant in relation to the total possible deformation length. “Coupled” related to the moment, where the seatbelt slack is removed and the restraint system active.

The Occupant Load Criterion (OLC) in [g] idealizes the occupant's forward displacement during the crash in relation to the X-velocity ${}^{\text{Vehicle}}\dot{U}_{X,t}$ of the vehicle. Two phases are defined for the occupant: (1) the free motion phase between t_0 and t_1, described by

$$ {}^{\text{Occ. Rel., Free}}U_{X,t} = \frac{T}{2} \cdot \sum_{t=t_0}^{t_1-1} \left({}^{\text{Vehicle}}\dot{U}_{X,t} + {}^{\text{Vehicle}}\dot{U}_{X,t+1} \right) - {}^{\text{Vehicle}}\dot{U}_{X,t_0} \cdot [t_1 - t_0]. \tag{2.25} $$

And (2) the restrained phase from t_1 to t_2, as defined by

$$ {}^{\text{Occ. Rel., Restr.}}U_{X,t} = \frac{T}{2} \cdot \sum_{t=t_1}^{t_2-1} \left({}^{\text{Vehicle}}\dot{U}_{X,t} + {}^{\text{Vehicle}}\dot{U}_{X,t+1} \right) - {}^{\text{Vehicle}}\dot{U}_{X,t_2} \cdot [t_2 - t_1]. \tag{2.26} $$

Until the restraint system is in force, an occupant moves relative to the vehicle constantly with its initial X-velocity. The duration of this phase results mainly from the time passed until the crash was detected, leading to the activation of the restraint systems and on the initial seatbelt slack. The OLC, as defined by

$$ \text{OLC} = \frac{{}^{\text{Vehicle}}\dot{U}_{X,t_2} - {}^{\text{Vehicle}}\dot{U}_{X,t_1}}{t_2 - t_1}, \tag{2.27} $$

evaluates the average acceleration, an occupant will experience, if the remaining space is used to achieve the condition of equal X-velocity of vehicle and occupant at t_2.

In the literature and test procedures as defined e. g. by Euro NCAP (2023c), the free motion phase is fixed to the value of 65 mm and for restrained phase, a value of 235 mm was defined. The time t_0 equals the beginning of the crash-phase. The equations 2.25 and 2.26 must be solved iteratively to calculate t_1 and t_2. Typically, the trapezoidal rule is used to integrate the discrete signal of the X-velocity. An implementation of the algorithm for PYTHON can be found in the code base by Plaschkies (2024a).

Kübler et al. (2009) considered the OLC not being a sufficient predictor of the injuries of the occupant. Nevertheless, the criterion is commonly used as an estimate for the severity of a vehicle-pulse when, for example, comparing the configurations of the same vehicle. Another application is the compatibility rating by Euro NCAP (2023c).

Motivated to find better estimates for the injury severity, Kübler et al. (2009) proposed the OLC++ still using the vehicle-pulse alone, and Rabus et al. (2022) the Real Occupant Load Criterion for Prediction (ROLC_p) taking additional restraint system properties into account. Furthermore, Schneider (2023) developed a universal configurable spring-damper model for the approximation of vehicle-pulses with the motivation to approximate the airbags' deployment times.

2.2 Machine Learning

2.2.1 Machine Learning Algorithms – Supervised Learning

2.2.1.1 Introduction

Machine learning uses statistical data and trainable algorithms to generate predictions. The domain is often categorized in supervised, unsupervised, semi-supervised, and reinforced learning. The algorithms are parametrized, and the parameters are optimized to fulfil a particular objective.

In supervised learning, the parameters are optimized to fit an input to a known output as ground truth during the training phase. Unsupervised learning attempts to recognize patterns like clusters in the data. In semi-supervised learning, the known pairs of input and output are used to identify correlations, on which basis additional data is generated. In reinforcement learning, actions and outcomes are typically defined. By formulating reward functions, the algorithm is optimized to a maximum reward while performing actions.

As described in the introduction section, this work utilized supervised learning methods. In this domain, a function f maps an input $\boldsymbol{\mathcal{X}}$ to an output $\boldsymbol{\mathcal{Y}}$ as denoted by

$$\boldsymbol{\mathcal{Y}} = f(\boldsymbol{\mathcal{X}}). \tag{2.28}$$

On the input side of a metamodel, there can be structured or unstructured data. Structured data are also called tabular data or attributes, since the data can be represented in a plain table or 2D-matrix $\boldsymbol{\mathcal{X}}$ assembled from individual feature attributes $\boldsymbol{\chi}$. Each column represents an attribute ξ from the feature attribute space $\mathbb{F}$ with the number of instances $\mathcal{F}$ and each row an instance ς from the dataset $\mathbb{S}$ with the number of instances $\mathcal{S}$. Values in the raw data can be either numerical or a character string. Possible scales for numerical values are nominal, ordinal, interval, and ratio. The type of scale must be considered when selecting an appropriate learning algorithm. Since the algorithms are working in a numerical space, strings must be converted, leading to

$$\boldsymbol{\mathcal{X}} \in \mathbb{R}^{\mathcal{S}\times\mathcal{F}} \text{ and } \boldsymbol{\chi} \in \mathbb{R}^{\mathcal{S}} \text{ with } \mathcal{X}_{\varsigma,\xi} = {}^{\xi}\chi_{\varsigma}. \tag{2.29}$$

One-hot-encoding is one approach to map a string to multiple feature attributes using zeros and ones to indicate to which class it belongs, leading to the number space $\mathbb{Z}$ instead of $\mathbb{R}$ for those particular attributes.

In principle, the targets of a prediction model can be all type of data, like the input. Consequently, for targets attributes $\boldsymbol{\mathcal{Y}}$, assembled from individual attributes $\boldsymbol{\upsilon}$, the definition

$$\boldsymbol{\mathcal{Y}} \in \begin{cases} \mathbb{R}^{\mathcal{S}\times\mathcal{T}} & \text{Regression} \\ \mathbb{Z}^{\mathcal{S}\times\mathcal{T}} & \text{Classification} \end{cases}, \text{ and } \boldsymbol{\upsilon} \in \begin{cases} \mathbb{R}^{\mathcal{S}} & \text{Regression} \\ \mathbb{Z}^{\mathcal{S}} & \text{Classification} \end{cases}, \text{ with } \mathcal{Y}_{\varsigma,\tau} = {}^{\tau}\upsilon_{\varsigma}, \tag{2.30}$$

with attributes τ in target space $\mathbb{T}$ holds true.

There is a fundamental differentiation between regression and classification tasks. For classification, one or more classes are formed. The prediction task is to estimate which class a sample

belongs to. If only one class exists, the task is called binary classification, where all samples belong to a class. For a single attribute, the instances belonging to the class can be denoted as ${}^{+}\boldsymbol{\upsilon}$ and the ones not belonging to the class as ${}^{-}\boldsymbol{\upsilon}$. In the case of regression, the target values are continuous.

Examples of unstructured data are pictures, audio signals, and text. The unstructured data is mathematically represented by adding more axes to $\boldsymbol{\mathcal{X}}$. The common assumption of structured data is the independence of the attributes from each other. Hence, each of the attributes should contain a unique piece of information and the neighbourhood of two attributes carries no meaning. For unstructured data, this is not the case. For example, in pictures, the location of each pixel is normally relevant. Hence, for each data type, specific algorithms were developed to extract the information – feature attributes – from the unstructured state. Those algorithms are either fixed data transformers or can contain learnable parameters. Data like related audio signals are called multivariate time-series.

2.2.1.2 Basic Supervised Learning Strategy

The input matrix $\boldsymbol{\mathcal{X}}$ is composed of $\mathcal{F}$ feature attributes in $\mathbb{F}$ with index ξ for a set $\mathbb{S}$ of instances with index ς and the size $\mathcal{S}$. The prediction as output matrix $\boldsymbol{\mathcal{Y}}$ is composed of $\mathcal{T}$ target attributes in $\mathbb{T}$ with index τ for the same set of instances. As an example,

$$\upsilon_{\varsigma,\tau} = \sum_{\xi=1}^{\mathcal{F}} \Theta_{\xi,\tau} \cdot \mathcal{X}_{\varsigma,\xi} + \Theta_{\mathcal{F}+1,\tau}, \tag{2.31}$$

describes the prediction of a single instance of one target attribute by linear regression, with the weights and bias $\boldsymbol{\Theta}$ being the learnable parameters.

During the training and evaluation phase, the ground truth $\boldsymbol{\mathcal{Y}}$ must be available. The metamodel makes predictions $\hat{\boldsymbol{\mathcal{Y}}}$, which are compared by the loss-function $\boldsymbol{\mathcal{L}}$ with the ground truth. A typical loss-function for regression problems is the squared error

$${}^{\text{Squared Error}}\mathcal{L}_{\varsigma,\tau} = (\hat{\mathcal{Y}}_{\varsigma,\tau} - \mathcal{Y}_{\varsigma,\tau})^2. \tag{2.32}$$

The cost-function $\boldsymbol{c}$ is the statistical evaluation of the loss-function over all instances, like

$$c_\tau = \frac{1}{\mathcal{S}} \sum_{\varsigma=1}^{\mathcal{S}} \mathcal{L}_{\varsigma,\tau}, \tag{2.33}$$

by calculating of the average of the loss-functions from a set of instances with the size $\mathcal{S}$.

If the loss-function is additionally extended by regularization terms, the function is called the objective-function $\boldsymbol{\jmath}$. The regularization term, is added to the optimizing algorithm, punishing too adaptive behaviour. For example, in linear regression with regularization term – so-called ridge regression – the regularization term sums up the parameters $\boldsymbol{\Theta}$ leading to

$$\jmath_\tau = c_\tau + \lambda \sum_{\xi=1}^{\mathcal{F}} \Theta_{\xi,\tau}^2. \tag{2.34}$$

The hyperparameter λ is used to weight the regularization term in the objective-function. Alternatively, the objective-function in the form

$$\jmath_\tau = c_\tau + \lambda \sum_{\xi=1}^{\mathcal{F}} |\Theta_{\xi,\tau}| \tag{2.35}$$

is used, too. Linear regression with the absolute value of the parameters in the regularization term is called lasso regression. The regularization term of equation 2.35 is also referred to as L1 regularization term, and the term of equation 2.34 on the preceding page as L2 regularization term.

Since the optimizer attempts to minimize the objective-function, this additional term prevents single parameters values from becoming relatively high during the fitting process. An overly high single parameters values weights the related input higher; hence, overfitting (see page 42) can occur. The parameters of the algorithm are iteratively adjusted by the optimizer to optimize the objective-function. The direction of the optimization is determined by the chosen loss-function. Variables that are not adjusted by the optimizer are called hyperparameter.

2.2.2 Data-serialization Formats

When deciding how to store data persistently, a massive number of data-serialization formats are available; the Wikipedia Contributors (2024) alone lists 37 distinct formats. In this place, the relevant formats for this work should be discussed with a special focus on the specifics regarding the programming language PYTHON by the Python Software Foundation (2024).

To understand the persistent data-serialization formats, a short digression on temporary formats in PYTHON seems apt. First, everything in PYTHON are objects, where objects under the hood are referenced by identifiers. There are mutable and immutable objects. An immutable object will change its reference identifier when changed, while a mutable object allows changes under the same reference identifier. Examples of immutable objects are integers, floats, strings as a chain of characters, and tuples as a chain of mutable or immutable objects. Mutable objects are, for example, lists as a chain of objects with a specific order, sets as collections of immutable and unique objects with no specific order. Further, dictionaries as collections of key-value-pairs with no specific order made of an immutable and unique object as a key and an object as the corresponding value. The uniqueness of an object is determined by generating its hash. A hash is the output of a deterministic function that converts an arbitrary object into a fixed value of size.

Next to the above explained native object types, there are manifold libraries leveraging concepts from other programming languages, such as C and C++. The library NUMPY by Harris et al. (2020) introduces so-called arrays using the C style pre-allocation of temporary memory. The pre-allocation allows for rapid vectorized operation on the Central Processing Unit (CPU). Other libraries like TENSORFLOW from the TensorFlow Developers (2024) use a similar format. The widely used library PANDAS from the Pandas Development Team (2024) adds column and row properties of an array, enabling the filtering and manipulation of data. Finally,

PYARROW from the Apache Software Foundation (2024) provides access to the Apache Arrow Framework. In this framework, data is stored in blocks in combination with a central header, allowing distributed storage and parallelized access to the data. In contrast, to access an object within any iterable, either its exact location must be known or it must be compared to a condition. The comparing-iteration can be sped up by sorting the object, e. g. by their hash-values. Nevertheless, for stacked formats, each stack-level must be iterated over sequentially. For example, each column must be checked against a condition first, to allow subsequently the access of the row. The library POLARS by Vink et al. (2024) offers similar to PANDAS data handling strategies but leverages the advantages of the Apache Arrow Framework such as the parallel access of data blocks.

The data generated for this work originate from Livermore Software technology corporation DYNA (LS-DYNA) and Virtual Performance Solution (ESI) (VPS). As described in section 2.1 on page 18 ff, LS-DYNA generates files in a structured binary format, which can be parsed e. g. by the library LASSO-PYTHON. The library was released by Lasso GmbH (2022) and is maintained by the community. The data from VPS in section 3.3 on page 121 ff was provided using the ERF-HDF5, which is based on the Hierarchical Data Format (HDF) framework and can be parsed using H5PY from Collette et al. (2023). Although such formats are well suited to store the results of FE-simulations, they do not provide any database functionality. Consequently, they have to be parsed into one of the PYTHON internal formats for further use.

Each of the above-described temporary data formats has a more or less direct equivalent to a persistent format. Binary formats often consume less storage and are accessed faster from the persistent storage than human-readable formats. An example of human-readable formats is CSV (Comma Separated Values). This file type corresponds directly to the PANDAS tables. Another example is the JSON format, which contains programming language-agnostic formats of basic structures such as lists and dictionaries.

An example of binary files is the pickle format, which provides a method to serialize any kind of object. NUMPY arrays can be stored by a similar method. The Apache Arrow framework defines the PARQUET format. In table 2.7 on the following page, the data access time and storage size for the discussed formats are compared. A 1000×1000 matrix of random floats was generated and converted to the respective PYTHON internal structures. Next, each object was stored three times on the hard drive and read again. The presented values are the arithmetic mean and can only be evaluated relative to each other, since the measures highly depend on the hardware used. All three binary formates perform similarly. The higher overhead of the PARQUET format leads to a slightly higher cost. However, its scalability, parallelized filter operations, local read access, and compression functionalities would change the result drastically. It should be noted that internal data conversion in combination with some data can neglect the effect of a particular format in a way that even a CSV file processing is ultimately faster than using PARQUET.

Table 2.7: Read and Write Times Comparison for Data-serialization Formats for a 1000 × 1000 Data-matrix

Format	Time Read [s]	Time Write [s]	Size [GB]
NUMPY	0.27	0.39	0.75
PICKLE	0.28	0.42	0.75
PARQUET	0.60	7.62	0.91
CSV	17.16	84.46	1.79
JSON	22.91	14.25	1.84

As stated above, there are manifold data-serialization formats. For this work, the described ones were relevant. The JSON was selected for clear human readability, especially for dictionary-like objects. The CSV was preferred for numerical data in tabular form with a low number of rows and columns, again because of the clear human readability. The selection of binary formats was motivated by the ease of use within PYTHON and its performance. General objects were stored using PICKLE. The PARQUET format was used in combination with PYARROW or PANDAS due to the superior filtered reading capabilities and the extensive toolbox for data manipulation.

2.2.3 Performance Assessment

2.2.3.1 Generalization Assessment – Over- and Underfitting

Above, on page 39 ff, loss-functions, cost-functions, and objective-functions were introduced as global training metric-functions. For model selection and evaluation, an evaluation metric-function $\mathcal{M}$ is used. This metric-function can have the same definition as the cost-function. However, since this metric-function has no requirement regarding computational efficiency and convergence behaviour, other functions can be used too. A simple example for $\mathcal{M}$ is

$$\mathcal{M} = \frac{1}{\mathcal{T}} \cdot \sum_{\tau=1}^{\mathcal{T}} \jmath_\tau, \tag{2.36}$$

as the arithmetic mean of the normalized objective-functions.

Next to an as exact match as possible between $\boldsymbol{\mathcal{Y}}$ and $\hat{\boldsymbol{\mathcal{Y}}}$, the generalization performance of the algorithm is crucial. In this context, two basic extreme behaviours have to be balanced. In the first case, the model possesses a high variance; this condition is called overfitting. In the other case, the model has a high bias, also known as underfitting the data. High variance of a model is understood as the flexibility or degree of freedom the model has to adapt to the data. The biases are assumptions about the data and hence limited degrees of freedom impeding a precise modelling. The overfitting and underfitting can be evaluated by splitting the available data with existing ground truth. The first set ${}^{\blacktriangleright}\mathbb{S}$, so-called the training-set, is used to train the model. The second set of instances ${}^{\triangleright}\mathbb{S}$, the validation-set, is hold-out during the training and is passed as unseen data through the fitted algorithm. When the metric-functions ${}^{\blacktriangleright}\mathcal{M}$ and ${}^{\triangleright}\mathcal{M}$ are assessed,

and assuming a definition of a metric-function reporting in the interval, the definition

$$[0,1] = \{\mathcal{M} \in \mathbb{R} \,|\, 0 \leq \mathcal{M} \leq 1\} \text{ where 1 is best,} \tag{2.37}$$

would lead to the assessment of

$$^{\blacktriangleright}\mathcal{M} \ll 1 \rightarrow \text{Underfitting / High Bias, and} \tag{2.38}$$

$$^{\blacktriangleright}\mathcal{M} \ll {}^{\triangleright}\mathcal{M} \rightarrow \text{Overfitting / High Variance.} \tag{2.39}$$

Overfitting can be countered by adding more instances to the training-sets, the selection of less and better features, and higher weighted regularization term. Underfitting can be fought by choosing an algorithm of higher complexity, the addition of more features, and the reduction of regularization term weight.

2.2.3.2 Performance Metric-functions

The performance assessment of a metamodel is a statistical evaluation. Not the correctness of a single prediction is relevant, but the overall performance on a larger set of instances. The selection of an appropriate metric-function depends on the data type of the target and on the objectives of the project.

Powers (2011) discussed established metric-functions for the binary and general classification cases. By weighting classes, skew data distributions can be treated or preferences from the work's objectives implemented. Many metric-functions are based on a **Confusion Matrix** as defined in table 2.8, evaluating the absolute frequency $\mathcal{H}$ of correctly or wrongly classified instances in relation to the ground truth.

Table 2.8: Confusion Matrix for Single Attribute Binary Classification

		Prediction	
		Positives	Negatives
Ground Truth	Positives	True Positives $\mathcal{H}(^{+}\hat{\boldsymbol{v}} \cap {}^{+}\boldsymbol{v})$	False Negatives $\mathcal{H}(^{-}\hat{\boldsymbol{v}} \cap {}^{+}\boldsymbol{v})$
	Negatives	False Positives $\mathcal{H}(^{+}\hat{\boldsymbol{v}} \cap {}^{-}\boldsymbol{v})$	True Negatives $\mathcal{H}(^{-}\hat{\boldsymbol{v}} \cap {}^{-}\boldsymbol{v})$

Recall or sensitivity is defined by

$$\text{Recall} = \frac{\mathcal{H}(^{+}\hat{\boldsymbol{v}} \cap {}^{+}\boldsymbol{v})}{\mathcal{H}(^{+}\hat{\boldsymbol{v}})} \in \mathbb{R} = [0,1]. \tag{2.40}$$

Thus, recall evaluates only how many of the predicted class belongings are correct. In contrast, **Precision** or confidence evaluates according to

$$\text{Precision} = \frac{\mathcal{H}(^{+}\hat{\boldsymbol{v}} \cap {}^{+}\boldsymbol{v})}{\mathcal{H}(^{-}\hat{\boldsymbol{v}})} \in \mathbb{R} = [0,1], \tag{2.41}$$

how many of the class belongings in the dataset were classified correctly.

Finally,

$$\text{Specificity} = \frac{\mathcal{H}(^{-}\hat{\boldsymbol{v}} \cap {}^{-}\boldsymbol{v})}{\mathcal{H}(^{-}\hat{\boldsymbol{v}})} \in \mathbb{R} = [0, 1] \tag{2.42}$$

describes the **Specificity** of a model. Specificity evaluates only the correctness of samples, not belonging to a class.

The three metric-functions shown evaluate only a specific aspect of a binary classification. A combined metric-function like

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \in \mathbb{R} = [0, 1] \tag{2.43}$$

can be defined as the harmonic mean of recall and precision, and is called the **F1-score**. In the case of skewed data distributions, such as many more positive cases than negative cases, rates instead of counts are used. Usually, the metric-functions are then named differently. For example, the accuracy will be called the Cohen's Kappa.

For the multi-class case, two major strategies are used: one-vs-one and one-vs-rest. Both split the non-binary classes into a binary situation. In the one-vs-one strategy, the positive / negative situation is achieved by assigning a pair of classes. In the one-vs-rest strategy, one class is used as a positive, and all other classes form the negatives. The resulting scores are averaged per class. The macro average is defined as a simple average over all class scores. The micro average, on the other hand, uses the number of samples for the calculation. Hence, class imbalance is considered. Another way to achieve this would be to weigh each class by its occurrence and calculate the weighted average. Additionally, the metric-functions are defined for the single attribute case; condensing the individual values for each attribute can be done analogous equation 2.36 on page 42.

Typical examples for metric-functions used to evaluate regression tasks are **Mean Absolute Error (MAE)**

$$\text{MAE} = \frac{1}{\mathcal{S}} \cdot \sum_{\varsigma=1}^{\mathcal{S}} |\upsilon_\varsigma - \hat{\upsilon}_\varsigma| \in \mathbb{R} = [0, +\infty), \tag{2.44}$$

Mean Squared Error (MSE),

$$\text{MSE} = \frac{1}{\mathcal{S}} \cdot \sum_{\varsigma=1}^{\mathcal{S}} (\upsilon_\varsigma - \hat{\upsilon}_\varsigma)^2 \in \mathbb{R} = [0, +\infty), \text{ and} \tag{2.45}$$

Root Mean Squared Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{\mathcal{S}} \cdot \sum_{\varsigma=1}^{\mathcal{S}} (\upsilon_\varsigma - \hat{\upsilon}_\varsigma)^2} \in \mathbb{R} = [0, +\infty). \tag{2.46}$$

For all three metric-functions, 0 would be a perfect score. Again, the metric-functions were defined for the single attribute case; condensing can be done similar as above.

Another typical metric-function is the **R2-score**

$$\text{R2-score} = 1 - \frac{\sum_{\varsigma=1}^{\mathcal{S}} \left(v_\varsigma - \hat{v}_\varsigma\right)^2}{\sum_{\varsigma=1}^{\mathcal{S}} \left(v_\varsigma - \bar{v}\right)^2} \in \mathbb{R} = (-\infty, 1]. \tag{2.47}$$

The reference is the prediction of the arithmetic mean of the ground truth. This score ranges from minus infinity to plus one. One would be a perfect score, and zero indicates a random prediction near the arithmetic mean of all seen target values. Scores lower than zero can be interpreted as predictions that are worse than random.

All metric-functions evaluate the model's performance statistically. The averages in the regression metric-functions can be affected by outliers.

2.2.3.3 Standardized Evaluation – The Framework Approach

Performing a thorough evaluation is essential for reliable results. In terms of ensuring the quality during development, a standardized process seemed appropriate. The author developed a framework in the project arTIco, published by Plaschkies et al. (2024b), for a related task. There, solving a multi-class problem with multivariate time-series data as input was part of the project's content. To ensure the comparability between the developed solutions of the project partners, the author introduced the following framework, which was later adapted to be used specifically for this work. The framework was designed to remain agnostic about the actual machine learning library.

The framework was set up using PYTHON. The central packages, not built-in, are PANDAS and POLARS for data handling, NUMPY for calculations, and SCIKIT-LEARN by Pedregosa et al. (2011) for data splitting and metric-functions. For the variation of hyperparameters, the OPTUNA framework from Akiba et al. (2019) was used.

The heart of the framework is the standardized and parametrized pipeline. The main development section is the model module. Its inner structure is up to the researcher to design, the interfaces to the evaluation flow are fixed. The raw data is the input to the pipeline and is transformed to be used in the model module. The model module's output is evaluated. All variables in the pipeline and the model module are referred to as hyperparameters. A set of such hyperparameters together with the script version and execution date forms a trial.

Data splitting is the standard approach for assessing the overfitting tendency of a metamodel. The framework uses a 5-fold cross-validation strategy, where the instances are shuffled and split into five partitions. A metamodel is trained five times from scratch. Each time, one partition is hold-out as validation-set and the remaining instances are used as training-set.

The performance in the training-set is reported as the arithmetic mean over the predictions using the respective training-set. To report the generalization capabilities, all predictions on the hold-out sets of each fold were concatenated. Using a bootstrapping approach, next to the arithmetic mean, a confidence interval was calculated.

2.2.3.4 Multi Criterial Performance Evaluation

From the above explanations, the quality of a metamodel is assessed using statistical performance metric-functions that evaluates the prediction in comparison to a ground truth. In practise, more aspects such as the computational cost for the data generation and training should be addressed, too. Plaschkies et al. (2023) proposed an approach utilizing Multi-Criteria Decision Making (MCDM).

In MCDM, each set of hyperparameter settings (trial) is called an alternative. The alternatives are compared and ranked against each other using predefined criteria. Plaschkies et al. (2023) used the Preference Ranking Organization Method for Enrichment Evaluations (PROMETHEE), as first described by Brans and Vincke (1985). Plaschkies et al. (2023) discussed the computational complexity of the original algorithm and the by Keyser and Peeters (1996) described instability of the completely ordered list of alternatives. In the so-called reverse rank problem, a new alternative can affect the order of other alternatives to each other, making it necessary to evaluate with each new alternative the full set of alternatives. As a solution, the modification of the PROMETHEE algorithm proposed by Calders and Assche (2018) was presented.

Plaschkies et al. (2023) proposed a step-wise approach to define the relevant criteria and rank the alternatives. The authors proposed to consider aspects such as representativeness, correlation, transparency, and the level of scales during the definition of the criteria. The representativeness of the generated alternatives determines whether the criteria are correct and meaningful and represent various aspects of the problem. Correlated criteria should be merged to avoid unwanted higher weights on a specific aspect. Invariant criteria should be inspected if the invariance is only due to the generated alternatives or meaningful to the overall problem. In the first case, the criterion can remain and is dropped in the ladder. Criteria should be grouped into meaningful categories to support a transparent rating scheme. It depends on the actual use case to which category an aggregated criterium fits. Another aspect of transparency is the understandability of a criterium. A directly assessed criterium is more straightforward to process and understand than one resulting from complex calculations. Ultimately, the reporting scale of a criterion should be considered. Typically, nominal, ordinal, interval, and ratio scales are used, where nominal has the lowest level, and ratio has the highest level. The lowest-level scale of the ones used will determine which MCDM method can be utilised. Not all criteria can be assessed on ratio scales. However, if possible, the higher-level scale always seems preferable.

The overall method seemed promising if experts' knowledge on key performance indicators is available. In addition, it must be possible to translate such indicators into criteria and implement those criteria in the evaluation framework.

2.2.4 Examples of Supervised Learning Algorithms

2.2.4.1 Overview – A Systematics of Algorithms

The field of supervised learning is broad and ranges from a simple linear regression to complex Artificial Neural Network (ANN). In figure 2.9, examples are collected by Imam et al. (2024). There are multiple ways of categorizing supervised learning algorithms. One way is the differentiation between parametric and non-parametric models.

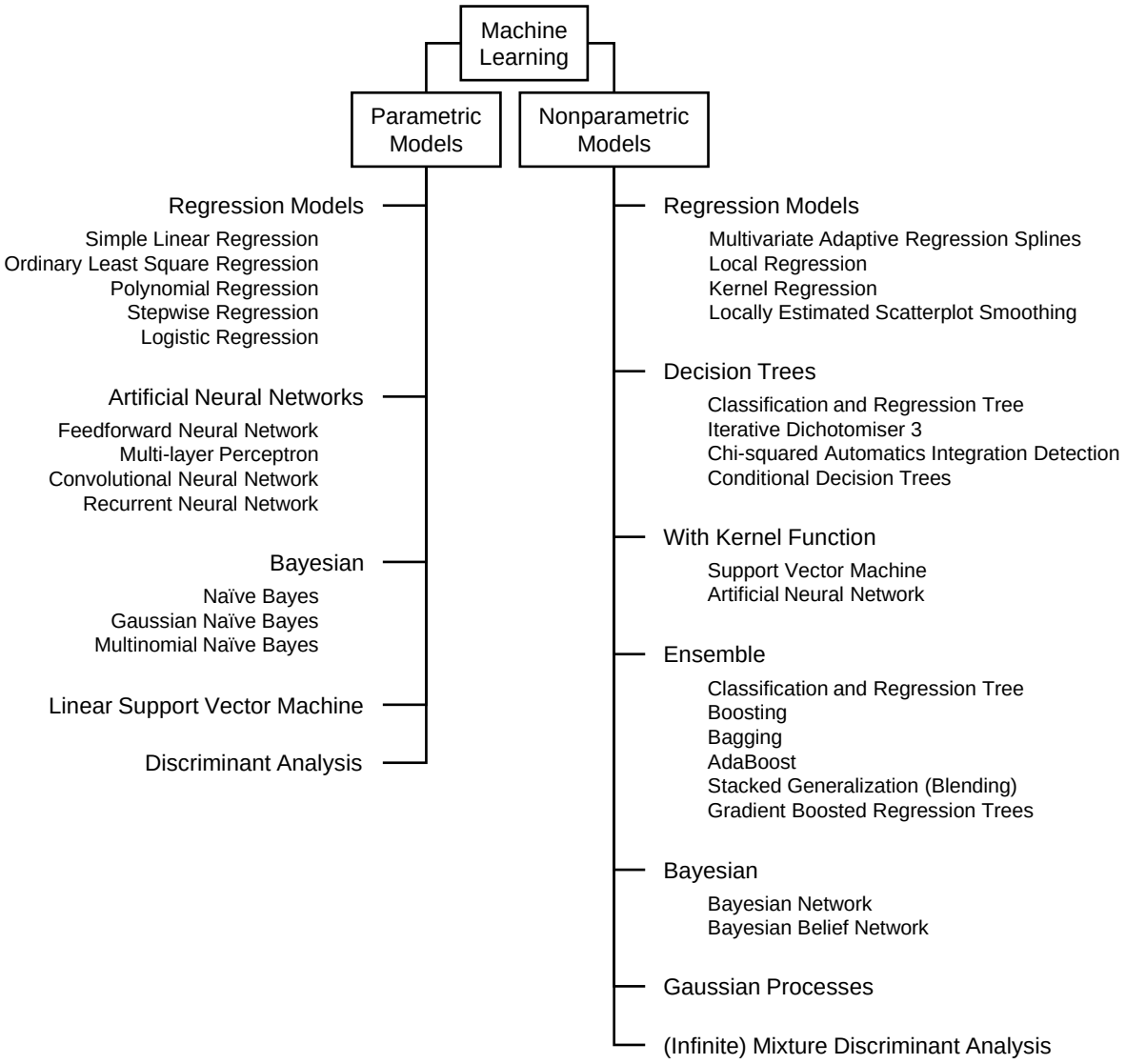


Figure 2.9: Examples of Supervised Learning Algorithms (adapted from Imam et al., 2024)

Imam et al. (2024) described parametric techniques make assumptions on the function $f(\boldsymbol{\mathcal{X}})$, which means that the "estimated function has a finite set of parameters that are not affected by new data. These parameters can be estimated by fitting the training data into the model"

(page 3). In contrast, non-parametric techniques make no assumption about $f(\boldsymbol{\mathcal{X}})$; "the number of parameters to estimate is not fixed and often increases with additional data" (page 3).

Some methods are developed to work for a specific task type. For example, the simple linear regression is for regression tasks and the logistic regression for classification task. Both algorithms require tabular data but have analytical determinable optima of their objective-function. Other algorithms provide robustness against outliers in the training-set. In this category belong the decision tree algorithms and the tree-based ensemble methods. Exceptional flexibility regarding task and data is offered by the group of Artificial Neural Network (ANN). By adding approaches from Bayesian statistics, the ANN can even be transformed from parametric to non-parametric algorithms. In practise, multiple algorithms can be equal candidates for a particular application. In the following, selected algorithms are briefly introduced to illustrate the core principles of machine learning.

The selection of the algorithm is often a trade-off between the required complexity and the computational cost. Naturally, a linear regression requires significantly less resources than a complex ANN but can only cover linear relationships. Additionally, the explainability can be a central factor. A decision tree can even be visualized, clearly documenting on which criteria a class was assigned. In contrast, especially ANNs are barely interpretable.

2.2.4.2 Linear Regression and Logistic Regression

One of the simplest algorithms, the linear regression and its adaption to lasso and ridge regression, was discussed above on page 39.

Logistic regression is used as one of the basic algorithms for classification tasks. Instead of the linear function (see equation 2.31 on page 39), the logistic function

$$\hat{\mathcal{Y}}_{\varsigma,\tau} = \left[1 + \exp\left(\Theta_{\mathcal{F}+1,\tau} - \sum_{\xi=1}^{\mathcal{F}} \Theta_{\xi,\tau} \cdot \mathcal{X}_{\varsigma,\xi}\right)\right]^{-1} \text{ with } \begin{cases} \leq 0.5 & \Rightarrow {}^{-}\upsilon_{\varsigma} \\ > 0.5 & \Rightarrow {}^{+}\upsilon_{\varsigma} \end{cases} \tag{2.48}$$

is used to model the class probability.

The logistic function can take values between 0 and 1. The usual threshold is 0.5 to assign a positive class assignment to the prediction.

2.2.4.3 Support Vector Machines

Another group of algorithms is the Support Vector Machines (SVMs). The basic idea is to divide the space defined by the feature attributes using hyperplanes. The algorithm evaluates the distance of the instances to the hyperplane. The smallest distance of all instances to the plane is called the margin. The optimizer tries to find a plane, maximizing this margin. Instances exactly on this margin are called support vectors, as the algorithm basically ignores all other points. The hyperplane is used as a decision boundary.

The transition of SVM from classification to regression can be achieved by evaluating subspaces. The location of the hyperplane is now the regression output. Hence, the margin is now minimized.

2.2.4.4 Gaussian Process Regression

The Gaussian Process Regression (GPR), also known as kriging, is a popular method in probabilistic machine learning. An intuitive introduction to the methodology is provided by Wang (2023). A problem is modelled using a multivariate normal distribution, defined by its mean and covariance. All features and target attributes are assumed to be individually normal distributed.

To generate single attribute estimates ${}^{\triangleright}\hat{\boldsymbol{\upsilon}} \in \mathbb{R}^{{}^{\triangleright}\mathcal{S}}$ from new data, the covariance matrices $\boldsymbol{\mathcal{K}}$ for combinations of features ${}^{\blacktriangleright}\boldsymbol{\mathcal{X}} \in \mathbb{R}^{{}^{\blacktriangleright}\mathcal{S}\times\mathcal{F}}$ from the training-set and features ${}^{\triangleright}\boldsymbol{\mathcal{X}} \in \mathbb{R}^{{}^{\triangleright}\mathcal{S}\times\mathcal{F}}$ from the prediction-set and the target attributes ${}^{\blacktriangleright}\boldsymbol{\upsilon} \in \mathbb{R}^{{}^{\blacktriangleright}\mathcal{S}}$ from the training-set are used as defined by

$$ {}^{\triangleright}\hat{\boldsymbol{\upsilon}} = \boldsymbol{\mathcal{K}}({}^{\blacktriangleright}\boldsymbol{\mathcal{X}}, {}^{\triangleright}\boldsymbol{\mathcal{X}}) \cdot \boldsymbol{\mathcal{K}}({}^{\blacktriangleright}\boldsymbol{\mathcal{X}}, {}^{\blacktriangleright}\boldsymbol{\mathcal{X}})^{-1} \cdot {}^{\blacktriangleright}\boldsymbol{\upsilon}. \tag{2.49} $$

Additionally, the related variance ${}^{\triangleright}\hat{\boldsymbol{\sigma}}^2 \in \mathbb{R}^{{}^{\triangleright}\mathcal{S}}$ by

$$ {}^{\triangleright}\hat{\boldsymbol{\sigma}}^2 = \operatorname{diag}\left\{\boldsymbol{\mathcal{K}}({}^{\triangleright}\boldsymbol{\mathcal{X}}, {}^{\triangleright}\boldsymbol{\mathcal{X}}) - \boldsymbol{\mathcal{K}}({}^{\blacktriangleright}\boldsymbol{\mathcal{X}}, {}^{\triangleright}\boldsymbol{\mathcal{X}}) \cdot \boldsymbol{\mathcal{K}}({}^{\blacktriangleright}\boldsymbol{\mathcal{X}}, {}^{\blacktriangleright}\boldsymbol{\mathcal{X}})^{-1} \cdot \boldsymbol{\mathcal{K}}({}^{\blacktriangleright}\boldsymbol{\mathcal{X}}, {}^{\triangleright}\boldsymbol{\mathcal{X}})^T\right\} \tag{2.50} $$

as the main diagonal of the covariance of the target normal distribution. The variances depend solely on the feature attribute vectors. The ${}^{\triangleright}\hat{\boldsymbol{\upsilon}}$ equals the means of the normal distribution $\mathcal{N}_\varsigma({}^{\triangleright}\hat{\upsilon}_\varsigma, {}^{\triangleright}\hat{\sigma}^2_\varsigma)$.

The variance collapses to ${}^{\blacktriangleright}\hat{\boldsymbol{\sigma}}^2 = \mathbf{0}$ for instances from the training-set because in the basic algorithm no noise is assumed.

The covariance matrix $\boldsymbol{\mathcal{K}}$ is formed as exemplary, shown here:

$$ \boldsymbol{\mathcal{K}}({}^{\blacktriangleright}\boldsymbol{\mathcal{X}}, {}^{\triangleright}\boldsymbol{\mathcal{X}}) = \begin{bmatrix} \kappa({}^{\blacktriangleright 1}\boldsymbol{\chi}, {}^{\triangleright 1}\boldsymbol{\chi}) & \cdots & \kappa({}^{\blacktriangleright 1}\boldsymbol{\chi}, {}^{\triangleright \mathcal{S}}\boldsymbol{\chi}) \\ \vdots & \ddots & \vdots \\ \kappa({}^{\blacktriangleright \mathcal{S}}\boldsymbol{\chi}, {}^{\triangleright 1}\boldsymbol{\chi}) & \cdots & \kappa({}^{\blacktriangleright \mathcal{S}}\boldsymbol{\chi}, {}^{\triangleright \mathcal{S}}\boldsymbol{\chi}) \end{bmatrix} \in \mathbb{R}^{{}^{\blacktriangleright}\mathcal{S}\times{}^{\triangleright}\mathcal{S}}. \tag{2.51} $$

Commonly used is the squared exponential kernel function, also known as Gaussian radial basis function, is defined by

$$ \kappa({}^{\blacktriangleright\varsigma}\boldsymbol{\chi}, {}^{\triangleright\varsigma}\boldsymbol{\chi}) = \Theta_{1,\tau} \exp\left(-\frac{||\,{}^{\blacktriangleright\varsigma}\boldsymbol{\chi} - {}^{\triangleright\varsigma}\boldsymbol{\chi}||^2}{2 \cdot \Theta_{2,\tau}}\right) \in \mathbb{R}. \tag{2.52} $$

The parameters $\boldsymbol{\Theta}$ being the output variance and a length scale for each target attribute, where a bigger $\Theta_{1,\tau}$ would allow the model a wider distribution around its mean, and $\Theta_{2,\tau}$ controls the variability of the functions.

The adjustment of parameters by the optimizer updates the prior assumption on the data by the posterior tuning to the actual data. The choice of the kernel function itself can be considered a

hyperparameter. The optimizer attempts to maximize the log marginal likelihood

$$^{\text{Optimal}}\boldsymbol{\Theta} = \arg\max_{\boldsymbol{\Theta}} \{\log \mathcal{P}(\boldsymbol{\mathcal{Y}}|\boldsymbol{\mathcal{X}}, \boldsymbol{\Theta})\} \tag{2.53}$$

as the probability $\mathcal{P}$ for $\boldsymbol{\mathcal{Y}}$ given $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\Theta}$.

2.2.4.5 Tree Algorithms

Classification and Regression Trees (CART) segment the dataset repeatedly to reach a classification. Each split achieves to maximize e. g. the class purity of a set. The extension of decision trees is the random forest and the boosted trees. Although, the bagging and boosting techniques are theoretically not limited to a decision tree algorithm.

To form a random forest, multiple trees are trained in parallel; hence, they are classified as the ensemble method with bootstrap aggregation (bagging). Each tree gets a bagged set of instances from the total dataset. Finally, the prediction from each tree will be averaged. In bagging, instances are drawn from the total dataset uniformly and with replacement.

Boosted trees are trained sequentially. In the initial step, a single tree is trained on the whole training-set. Next, the correctness of prediction for each instances in the training-set is evaluated. In the subsequent steps, the loss-function is modified in a way that the incorrectly predicted instances are higher weighted. As the last step, all trained trees are used and their predictions are averaged, weighted by their individual loss-function they achieved during training.

2.2.4.6 Artificial Neural Networks

Basics

The Artificial Neural Network (ANN) is a very flexible class of learning algorithms. A comprehensive collection of important techniques and algorithms was presented by Patterson and Gibson (2017). Below, the key concepts used in this work are briefly presented. Furthermore, the book “Deep Learning” by Goodfellow et al. (2016) can be considered reference literature.

An ANN is composed of neurons organized in layers. A neuron is defined as a function that maps the input values to an output. An input layer acts as the entry point for data, and an output layer delivers the actual prediction. Layers between the in- and output layers are called deep layers. The way of connecting the neurons and layers, the choice of the neurons’ formulation, and the structure and size of the layers define the architecture of an ANN. Furthermore, mathematical operations such convolution and pooling can be performed within a network. There are multiple basic topologies of ANNs like single layer or multilayer Feed Forward Networks (FFNs) and Recurrent Neural Networks (RNNs).

Layer Types and Activation Functions

One basic layer type is the **dense layer**, also called a fully connected layer. Here, one neuron η of the layer ψ receives the output $\boldsymbol{\varrho}$ of the previous layer. The neuron transfer-function Υ

$$^{\psi,\eta}\Upsilon = {}^{\psi-1}\boldsymbol{\varrho} \cdot {}^{\psi,\eta}\boldsymbol{\theta} + {}^{\psi,\eta}\varepsilon, \tag{2.54}$$

processes the input as linear combination with weights $\boldsymbol{\theta}$ and bias ε.

The result Υ is fed into an activation function. This function must be continuously differentiable and should introduce some non-linearity. One example is **Rectified Linear Unit (ReLU)**

$$^{\psi,\,\text{ReLU}}\varrho_\eta({}^{\psi,\eta}\Upsilon) = \max(0, {}^{\psi,\eta}\Upsilon) \in \mathbb{R} = [0, +\infty). \tag{2.55}$$

The evaluated activation function is the neuron output ϱ. Another common activation function is the **hyperbolic tangent function (tanh)**

$$^{\psi,\,\tanh}\varrho_\eta({}^{\psi,\eta}\Upsilon) = \frac{2}{1 + \exp\left(-2 \cdot {}^{\psi,\eta}\Upsilon\right)} - 1 \in \mathbb{R} = [-1, 1]. \tag{2.56}$$

The **output layer** for the single attribute case contains one single neuron. This output neuron contains a simple pass through as the **linear activation function**

$$^{\psi,\,\text{Linear}}\varrho_\eta({}^{\psi,\eta}\Upsilon) = {}^{\psi,\eta}\Upsilon \in \mathbb{R} = (-\infty, +\infty) \tag{2.57}$$

in the regression case. For binary classification, the **sigmoid activation function**

$$^{\psi,\,\text{Sigmoid}}\varrho_\eta({}^{\psi,\eta}\Upsilon) = \frac{1}{1 + \exp\left(-{}^{\psi,\eta}\Upsilon\right)} \in \mathbb{R} = [0, 1], \text{ and} \tag{2.58}$$

for multi-class classification, the **softmax activation function**

$$^{\psi,\,\text{SoftMax}}\varrho_\eta({}^{\psi,\eta}\Upsilon) = \frac{\exp\left({}^{\psi,\eta}\Upsilon\right)}{\sum_\eta \exp\left({}^{\psi,\eta}\Upsilon\right)} \in \mathbb{R} = [0, 1] \tag{2.59}$$

is used. Each class is represented by a neuron of the output layer.

For Feed Forward Networks, in a pass, the feature attributes of one instance are fed into the **input layer**. Layer by layer, the values of activation functions are calculated and passed as input into the next layer until finally the output layer is reached. The output is now compared with the ground truth using the loss-function.

The above described layers interpret the features as tabular data; a space $\mathbb{R}^{\mathcal{S}\times\mathcal{F}}$ is assumed. Convolutional Neural Networks (CNNs) with **convolutional layers** are architectures enabling the processing of unstructured data. For example, the signals of $\mathcal{C}$ multivariate time-series with $\mathcal{F}$ timesteps will be $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{\mathcal{S}\times\mathcal{F}\times\mathcal{C}}$. In

$$\mathcal{X}_{\varsigma,\xi,\zeta} * \boldsymbol{\kappa} = \sum_{k=1}^{\mathcal{A}} \left[\mathcal{X}_{\varsigma,\xi+k,\zeta} \cdot \kappa_k\right] \in \mathbb{R}, \tag{2.60}$$

the convolutional operation with a single one dimensional convolutional kernel $\kappa \in \mathbb{R}^{\mathcal{A}}$ with the index k for one timestep ξ of a single time-series ζ is shown. In equation 2.60, the convolution is applied left aligned to the evaluated timestep while often the centred application is used, too.

The time-series must be padded to enable the displayed operation for the whole time range. Furthermore, the shift from one evaluated timestep to another, the so-called stride, can be adjusted. Additionally, by dilation, timesteps can be skipped within the convolution itself. The convolutional kernel shape, the stride, and the dilation, the padding strategy, the number of convolutional kernels, and the multivariate time-series strategy are hyperparameters. For example, a convolutional kernel $\kappa \in \mathbb{R}^{\mathcal{A},\mathcal{C}}$ would apply the convolution directly to all multivariate time-series. The convolutional kernels values themselves are parameters and, hence, the trainable part of a convolutional layer. The unstructured data of a greyscale picture require a two-dimensional convolution and a coloured picture a three-dimensional convolution. As the convolutional kernel moves over the data, a feature map is generated.

Often, after a convolutional layer, a **pooling layer** is placed. The commonly used max-pooling operation picks from each result field the maximum value, resulting in a compressed feature map. The pooling strategy and window are hyperparameters. The convolutional layer and pooling layer act as a feature attribute extractor, transforming the unstructured data into a tabular representation. To do so, as a last step, the last array after convolution and pooling is simply flattened, which is called a **flatten layer**. The resulting flat array can now be passed in as regular layers. The flatten layer and pooling layer do not contain trainable parameters.

Closely related to the convolutional layer is the **transposed convolution**. When the convolutional operation results in a downsampling of a time-series, then generates the transposed operation an upsampled version of that time-series. The equation 2.60 on the previous page described the convolution as the sum of the products of each convolutional kernel element with the elements of the current window of the time-series. For the transposed operation, the convolutional kernel is iterated over the time-series, too. The current timestep is multiplied by the convolutional kernel, which results in a scaled version of that convolutional kernel. If this operation is performed on each element, the scaled convolutional kernels will overlap depending on the mode of iteration. The overlapping elements will be summed up.

So far, only Feed Forward Networks were presented. In contrast, a RNN has a backward connection between layers. Data are understood as some form of time-series. One instance has now additionally a hidden time dimension. One forward-pass means one timestep of an instance passing through the network. At all recurrent connections, the neuron output is copied into the so-called hidden states. The backwards-directed value is stored for the next passes. Important characteristics are the number of layers with recurrent connections and the length of those connections. A hidden state can be described as a short-term memory of a network, as it can store information about a limited time interval. As an extension, long-term memory cells with forget gates were introduced. ANNs using such an architecture are commonly known as **Long Short-Term Memory (LSTM)** models. The rather complex structure of a single LSTM unit is presented in figure 2.10 on the facing page.

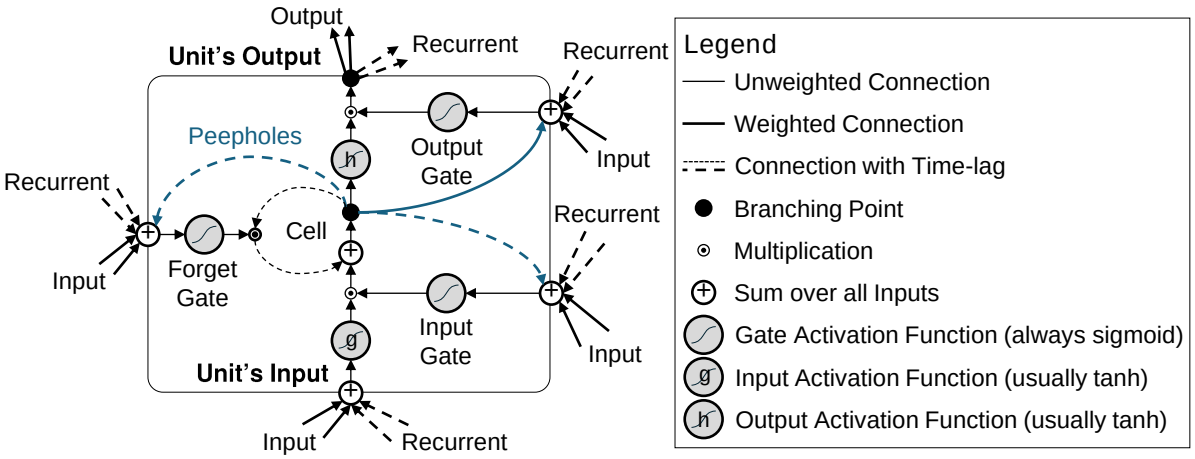


Figure 2.10: A single LSTM Unit (adapted from Patterson and Gibson, 2017)

Another type of layer is the **dropout layer**. During the training, it sets the output of random neurons of the previous layer to zero. The remaining outputs are scaled to maintain the sum over all neurons.

Training

During training an ANN, the parameters are adjusted to optimize the objective-function. In the PYTHON library TENSORFLOW from the TensorFlow Developers (2024) and the TENSORFLOW-wrapping library KERAS by Chollet et al. (2015), the regularization term is defined directly for each neuron and added internally to the cost-function. The available is the L1 regularization term and the L2 regularization term (see equation 2.34 and equation 2.35 on page 40). The parameters are grouped, and each group can be penalized individually.

Depending on the type of activation function of the output layer, an appropriate loss-function muss be selected and the reduction strategy for the cost-function respective objective-function be defined. The established practise was documented by Goodfellow et al. (2016) and summarized by Brownlee (2019). An established strategy to reduce the individual losses is the simple sum over all instances in a batch of instances. One target attribute τ of $\mathcal{Y}$ is predicted by one neuron of the output layer.

For regression problems with the linear activation function of the output layer (see equation 2.57 on page 51), absolute error or squared absolute error (see equation 2.32 on page 39) is typically used. For binary classification with the sigmoid activation function (see equation 2.58 on page 51) and for the multi-class classification with the softmax activation function (see equation 2.59 on page 51), it is the logarithmic loss which also called cross-entropy. The binary loss is defined by

$$^{\text{Binary}}\mathcal{L}_{\varsigma,\tau} = \mathcal{Y}_{\varsigma,\tau} \cdot \log\left(^{\text{Output, Sigmoid}}\varrho_{\tau}\right) + [1 - \mathcal{Y}_{\varsigma,\tau}] \cdot \log\left(^{\text{Output, Sigmoid}}\varrho_{\tau}\right), \tag{2.61}$$

with $\mathcal{Y} \in \mathbb{Z}^{\mathcal{S}\times\mathcal{T}} = [0, 1]$. In comparison, the cross-entropy for the multi-class case is defined

as categorial cross-entropy

$$^{\text{Multi-class}}\mathcal{L}_{\varsigma,\tau} = \mathcal{Y}_{\varsigma,\tau} \cdot \log\left[^{\text{Output, Softmax}}\varrho_{\tau}\right]. \tag{2.62}$$

The selection of the loss-function is driven by considerations of the convergence of the optimizer.

The most common method to adjust to the objective-function is the backpropagation method. In a forward-pass, the value of the objective-function $\jmath$ in the current iteration $\imath$ of the optimizer is determined. Next, in the backward-pass, the partial derivatives ∇ for all parameters $\mathbb{P}$ are calculated and the parameters according to

$$\mathbb{P}_{\imath+1} = \mathbb{P}_{\imath} - \alpha \cdot \nabla\jmath(\mathbb{P}_{\imath}) \tag{2.63}$$

are updated. The amount of change is controlled by the learning rate α.

After the backward-pass reaches the input layer, another forward-pass is started. There are three strategies of gradient descent. For stochastic gradient descent, the parameter update is performed after each forward-pass of a single randomly drawn instance without replacement from the training-set; evaluated is the loss-function. The other extreme is called batch gradient descent. Here, the whole training-set is passed and the resulting objective-function is used for the update. In between is the mini-batch gradient descent, where the training-set is split into batches on which the update will be based. Mini-batch attempts to balance the computational efficiency of using the full training-set and the execution speed of the stochastic approach. In general, a pass of all instances of the training-set is called an epoch. The epochs are repeated until an abort criterion is met.

The gradient descent base algorithm with the fixed learning rate can get stuck at a local optimum. The concept of the momentum, the gradient, is replaced by a form of a moving average of the gradients of multiple iteration. By this concept, the smaller "hills" in the optimizing surface should be overcome. Another approach is Root Mean Square Propagation (RMSprop) that introduces an adaptive learning rate for each parameter by taking into account the slope of gradient development. The combination of both is the widely used ADAptive Moment estimation (ADAM) algorithm, introduced by Kingma and Ba (2015).

The end of training an ANN can simply be determined by the number of past epochs. Additionally, early stopping can be defined. One criterion is to stop, once the cost-function of an epoch remains unchanged. Another is, to additionally evaluate the objective-function on unseen data, the validation-set. Usually, the training score improves during the iterations. Until the moment of overfitting, the same is true for the objective-function evaluated for the validation-set. Identifying this point can lead to a good epoch to stop the training.

2.2.5 Tabular Feature Processing – Feature Engineering

As described above, attributes are the column-wise representation of data. Each of the attributes is assumed to be independent. Furthermore, learning algorithms make assumptions about the distribution of the data and its relation to the target. Techniques processing such type of data, aim to clean the data, reduce the dimensionality, change the distribution, and the spread of the data.

In reality, data from a raw database is never clean. There may be missing information and outliers. Missing information in the form of an empty attribute value of one instance is not usable, since no mathematical operation can be applied. In addition to simply dropping the affected instances, less aggressive strategies such as imputing the missing values were developed. A simple imputing strategy is to fill the missing value with the average of all existing instances of that attribute. More advanced approaches apply, for example, a regression algorithm combining information from other features.

Outliers affect many learning algorithms, as those methods infer on the basis of statistics. Outliers must first be detected and then treated with an appropriate strategy. One common approach for detection is the application of the Interquartile Range (IQR). Each feature is evaluated as univariant. Two quartiles are defined and calculated as fences for the data. Frequent choices are the 25^{th} and 75^{th} quantiles. As outliers are now instances detected whose attribute values are outside the fence. Hereby, the fence is defined as the upper respective lower quartile plus respective minus the IQR. Typically, the IQR is multiplied by 1.5 for outliers and by 3 for extreme outliers. Once detected, the outlier affected instance can be treated as a instance with missing value.

Distributions can be shown using so-called box plots. A box plot indicates the median and the 25^{th} and 75^{th} quantiles as a box. The outlier threshold is represented by fences. Values outside those fences are drawn as points, so-called fliers. As an extension, the letter plots were introduced by Hofmann et al. (2017) for presenting a high number of data points. The non-parametric representation utilized boxes for more quantiles than in the box plot. An implementation in PYTHON can be found under the name boxenplot in the library SEABORN by Waskom (2021). The library is based on the fundamental plotting package MATPLOTLIB by Hunter (2007).

Often, feature attributes – but target attributes too – are on a continuous ratio scale. There can be discrete values or even categories. For example, by the above-mentioned one-hot-encoding, a attribute containing the belonging of an instance to one of multiple classes, is expanded to multiple attributes, each indicating, if the instance belongs to a particular class.

Many algorithms are sensitive to the magnitude of the data. The learning process is greatly improved if all attributes are in a similar range. In a first step, the data is often centred. This can be achieved simply by subtracting from each attribute value the arithmetic mean or median of the corresponding set of instances. The mean of a set of values is faster computed but sensitive to outlier. The computation of the median is more demanding computationally but robust against outliers. The scaling is achieved by the division of each attribute value by a statistical

measure of its corresponding instance set. This measure can be the difference of the minimum and maximum attribute value, a attribute standard deviation, or a quantile range. Again, the quantile range is a non-parametric measure and robust against outliers.

Combining attributes to reduce the dimensionality of a problem or transform their distribution and relationship is another basic set of techniques. As many learning algorithms assume linear relationships between the attributes, transformations aim to create this linearity. For example, in figure 2.11 on the left side, the two attribute vectors ${}^{1}\boldsymbol{\chi}$ and ${}^{2}\boldsymbol{\chi}$ from $\boldsymbol{\mathcal{X}}$ have a nearly quadratic relationship. The operation

$$\left\{ {}^{2}\boldsymbol{\chi} \propto {}^{1}\boldsymbol{\chi}^{2} \cap {}^{1}\check{\boldsymbol{\chi}} = {}^{1}\boldsymbol{\chi}^{2} \right\} \Rightarrow \left\{ {}^{2}\boldsymbol{\chi} \propto {}^{1}\check{\boldsymbol{\chi}} \right\} \tag{2.64}$$

for squaring the ${}^{1}\boldsymbol{\chi}$ would create a new attribute vector, ${}^{1}\check{\boldsymbol{\chi}}$ being direct proportional to ${}^{2}\boldsymbol{\chi}$.

Another approach is combining attributes which are not linear separable and mapping them to another attribute space. For example, in figure 2.11 on the upper-right side, the attribute vectors ${}^{3}\boldsymbol{\chi}$ and ${}^{4}\boldsymbol{\chi}$ have a circular relationship with radius r. The third attribute vector ${}^{5}\boldsymbol{\chi}$ is not correlated with ${}^{3}\boldsymbol{\chi}$. Given these three attribute vectors, a purely linear algorithm cannot separate the data into two classes. A new attribute vector ${}^{4}\check{\boldsymbol{\chi}}$ was generated according to

$$\left\{ {}^{4}\boldsymbol{\chi} \propto \sqrt{r^{2} - {}^{3}\boldsymbol{\chi}^{2}} \cap {}^{4}\check{\boldsymbol{\chi}} = {}^{3}\boldsymbol{\chi}^{2} + {}^{4}\boldsymbol{\chi}^{2} \right\} \Rightarrow \left\{ {}^{4}\check{\boldsymbol{\chi}} \propto r^{2} \right\}. \tag{2.65}$$

The new attribute space defined by ${}^{3}\boldsymbol{\chi}$, ${}^{4}\check{\boldsymbol{\chi}}$, and ${}^{5}\boldsymbol{\chi}$ is displayed in figure 2.11 on the lower-right side and has now linear separable properties.

For the visual, some random jitter was added to all ${}^{\varsigma}\boldsymbol{\chi}$ individually, representing noise in the data. For the sake of simplicity, the noise is not represented in equations 2.64 and 2.65

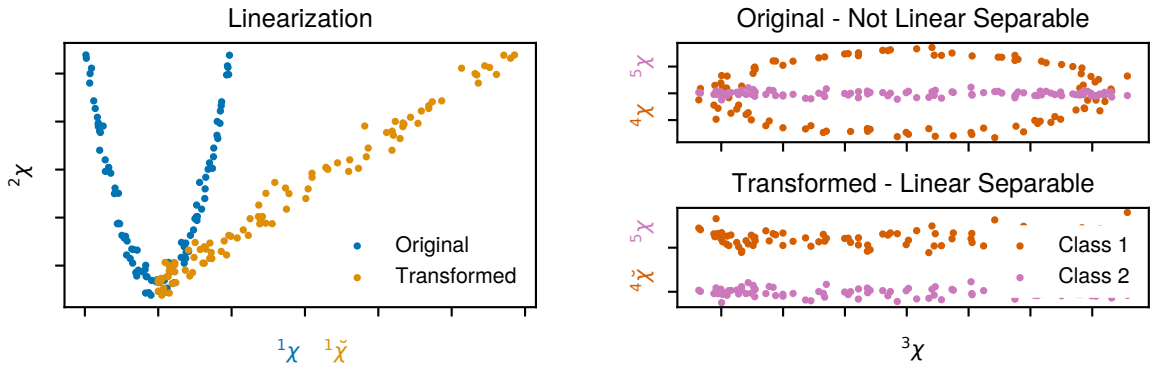


Figure 2.11: Attribute Vector Transformation (left) and Combination (right)

Mapping the attribute space of one data point into a higher dimensional space can be computationally expensive when it comes to the calculation of distances between points. With the so-called kernel function trick, a similarity between two points can be calculated in the original attribute space. The similarity is calculated as the dot product of two attribute vectors. With the

above described methods, both vectors must be transformed in the higher dimensional space and there the calculation is performed. Selecting an appropriate kernel function, enables the calculation in the original space. Typically, the kernel function trick is applied in SVMs to calculate the distance to the hyperplane. By applying the kernel function, non-linear decision boundaries can be achieved.

In contrast to the manual combination of two or more attributes to reduce dimensionality, Principal Component Analysis (PCA) leverages the Singular Value Decomposition (SVG) to combine correlated attributes into few ones. Data are first translated to a centre at zero. Next, the points are rotated to an axis preserving most information if all points would be projected on it. This axis is called the first Principal Component (PC). All higher components are orthogonal to the others and explain a decreasing part of the variance in the data.

All mentioned approaches are of a deterministic nature and can be inverted. However, like in the case of PCA, compressed data cannot be fully restored. Some transformations are independent of the data. Others contain parameters which must be fitted to the underlying data. Hence, such operations must be applied on training as well as on unseen data. Fitted parameters, however, must exclusively be determined based on the training data.

2.2.6 Multivariate Time-series Processing

All above described learning algorithms require tabular features. For ANN, the addition of the CNN and LSTM was presented. Both architectures are methods to extract feature attributes from time-series data. Consequently, they can be called a domain-agnostic, trainable feature attribute extractors. When the applied algorithm is independent of the actual data, it can be called a transformation. The calculation of injury-criteria is such a transformation for a specific domain. Additionally, there are domain-agnostic not trainable transformation strategies.

An overview of domain-agnostic transformations was provided by Bagnall et al. (2017). The authors named four categories: whole time-series processing, intervals, shapelets, and dictionary-based methods. For the whole time-series processing category, a time-series is simply considered a vector of independent attributes or is used for distance measures. The interval-based methods evaluate a time-series in time windows and calculate attributes like the time-series maximum within each time window. Shapelets are phase-independent patterns in a time-series. The CNN can be named in this category, too. Finally, dictionary-based methods count reoccurring patterns like the number of local maxima in a time-series.

The PYTHON library SKTIME by Löning et al. (2022) provides an extensive collection of domain-agnostic transformers. Bagnall et al. (2017) compared a selection for the application of time-series classification. However, for this thesis' research, regression and multivariate time-series regression should be studied, too. In this context, the shapelet-based algorithm ROCKET by Dempster et al. (2020) is of interest. The authors use the idea of convolution like in a CNN architecture but define the convolutional kernel weights as fixed. A high number like 1000 convolutional kernels of different size is randomly generated and applied on the time-series and processed by convolution and pooling to feature attributes.

Another approach is to combine the above approaches as a collection of attributes. Such collections were proposed by Lubba et al. (2019) as CATCH22 (22 CAnonical Time-series CHaracteristics) and Christ et al. (2018) as TSFRESH. Both libraries are also integrated in SKTIME.

If the metamodel output is one or multiple time-series, the computational cost must be considered. Many learning algorithms do not natively support a multi-output. For multi-attribute regression or classification, it can be acceptable to simply train a metamodel per attribute individually. To account for the temporal interdependency, the predictions can be chained, too; a predicted value of one timestep becomes a attribute for the next one. More efficient can be algorithms which natively support the multi-attribute output; examples of such algorithms are ANNs and some implementations of decision trees.

For machine learning, the ideal sampling rate of time-series does not necessarily equal the one required for time-series processing, like the assessment of injury-criteria. Hence, changing the time sampling interval is another relevant transformation. Plaschkies (2023) used averaging over fixed time windows. Additionally, for the time series prediction, the reversibility of PCA was used to predict only a limited number of PCs and restore the complete time-series. Instead of fixed time windows, a rolling average can be generated, too. An implementation can be found in the library PANDAS. To reduce the sampling interval, the given time-series can simply be interpolated. Another approach is the transformation of the time-series to its frequency-domain, dropping the not considered frequencies, and the transformation back to the time-domain. An implementation can be found in the PYTHON library, SCIPY.

2.3 Efficient Data Generation and Utilization

2.3.1 Efficient Data Generation with Finite Element Models

General approaches for efficient FE-simulations in Livermore Software technology corporation DYNA (LS-DYNA) were discussed above in section 2.1 on page 18 ff. As key concepts, the appropriate selection of element types and sizes can be highlighted. Furthermore, powerful computer-hardware and efficient pre- and post-processing algorithms are key.

Wielens (2022) presented a FE-model simplification approach which can be applied to a whole vehicle with the objective of an accelerated FE-simulation of a vehicle crash. The author presented two approaches for the simplification. The first approach defines a cutting plane through the vehicle. All elements on one side are considered relevant, and those on the other side will be dropped. The second approach requires one FE-simulation with the full vehicle. Next, e. g. the internal energy of each component is evaluated. Components whose energy is below a defined threshold during the crash can be neglected. To conduct now FE-simulations with the region of interest, again, two approaches were discussed. The first relates to an approach proposed by Schäffer et al. (2019a,b). The dropped components are replaced by in total two masses, representing those components' total mass and inertia. The masses were connected by elastic beams to the cut edges of the remaining FE-model. Wielens (2022)

proposed a completely rigid connection as one adaption of the method. Alternatively, the author suggested the utilization of a prescribed nodal displacement at the cut edges using the LS-DYNA interface functionality (see section 2.1, page 20).

The interface functionality is by the knowledge of this thesis' author commonly used for occupant-FE-simulations, too. The vehicle structure is merely considered a container for mounting the components of the interior and restraint system. When relevant, the Body In White (BIW) recorded deformation is applied by the interface functionality. The remaining structure is completely dropped or, if relevant for attaching any components, modelled rigid with a prescribed vehicle-pulse. Singh et al. (2018a) briefly presented a similar approach applied to the full vehicle model of a Honda Accord 2014, referred to as a "sled model". Similar, Sequeira (2022) applied this strategy to a FE-model of the Toyota Yaris 2010 as a base for crash scenario dependant crash severity estimations.

2.3.2 Design of Experiments

A reasonable Design of Experiments (DoE) is the heart of any data generation. Siebertz et al. (2017) provided a profound insight into established methods. The system under test is characterized by factors, where a combination of factor settings forms an experiment. Observed is the system's response. A Design of Experiments (DoE) is now the defined variation of the factor settings. Each factor should statistically uncorrelated of the other over the experiments to avoid unnecessary repetitions of an experiment. Furthermore, sufficient coverage of the factor space is essential.

Factors can be of a continuous or discrete nature. A scheme to vary discrete factor levels is the full factorial sampling, where every level of one factor is simply combined with all levels of another. The resulting number of experiments equals the number of levels to the power of the number of factors. Since not always the resulting high number of experiments is possible, screening methods were introduced. Such approaches assume, for example, a linear relationship between a factor and the observed response and additionally no combined effect of e. g. more than three factors. Since the description model requires a parameter per assumed influence of a factor and factor combination, removing a combination level reduces the number of parameters and hence the number of required experiments.

For continuous factors, it is assumed that all levels within a defined range per factor are possible and that the number of levels is infinite. Such a factor space can be transformed to a discrete space, or (semi) randomized methods can be used. The Monte-Carlo scheme, for example, defines the value of each factor for each experiment randomly. With the increasing number of experiments, the factors will be equally distributed, and hence the space-coverage will become homogeneous.

Another randomized approach is the Latin Hypercube Design. In a first step, the number of experiments is set. Second, the number of levels equals this number of experiments and is randomly permutated for each factor over the experiments. Next, random values are generated for each factor level value and subtracted from the respective value. The resulting difference is

normalized by the number of experiments. Finally, each factor is scaled to its assigned value range. There are multiple modifications of the base algorithm, improving the reliability of producing high quality test fields.

Instead of drawing from an equal distribution, sequences like the Halton- and Sobol-sequences are used to generate DoEs deterministically, while maintaining the qualities of randomized approaches. For this reason, such formalisms are also called pseudo-random. Intuitively, a Sobol-sequences split a unit-space incrementally in halves, resulting in a sequence, like

$$\left\{0, \frac{1}{2}, \frac{3}{4}, \frac{1}{4}, \frac{3}{8}, \frac{7}{8}, \frac{5}{8}, \frac{5}{8}, \dots\right\}.$$

This sequence is also called the binary Van der Corput Sequence. The construction takes for each element a natural number and converts it to its binary representation. Then, the binary is flipped and converted back to the decimal system. This new natural number is divided by 2 to the power of the number of bits required to represent the original number without leading zeros. The values of the other factors are permutations of the first one. The formalism for generating the sequence and its permutations can be expressed by a so-called primitive polynomial. To generate a Sobol-DoE different primitive polynomials can be used as well. Due to the formalism, the number of experiments $\mathcal{E}$ should be the function

$$\mathcal{E} = 2^s, \tag{2.66}$$

of the Sobol's s. An implementation of the sequence and permutation generating algorithm can be found in the PYTHON library SCIPY. The algorithm is deterministic; when the same primitive polynomial is selected, then the algorithm will always generate the same output. Furthermore, Siebertz et al. (2017) highlight the good uniform distribution of the resulting test-field.

2.3.3 Adaptive Design of Experiment

The approach of adaptive DoE is also referred to as active learning. Related to supervised learning, the terms summarize approaches, where the training process of the metamodel starts with an initial set of data with an existing ground truth, to which sequentially more data points are added by a query-algorithm. The additional instances can be generated by simply increasing the original DoE. In full factorial designs, more factor levels can be added and the quasi-random Sobol sequences can be continued. Both lead to a denser factor field. However, not all areas of the field account the same to the observed system's responses. Consequently, active learning approaches take these local differences into account with the objective of ultimately making less experiments necessary than when using a full DoE formalism. In other words, the potential informativeness of a data point without yet existing ground truth is the criterion to be selected for the subsequent data generation.

Settles (2010) reported multiple established approaches to active learning. The **uncertainty-based query-algorithms** were presented as the most common approach. Here, the uncertainty of a metamodel in the prediction for a particular instance without existing ground truth is

evaluated. The next candidate for data generation is the instance with the highest uncertainty. Another approach discussed is **community-based query-algorithms**, where multiple metamodels make predictions on instances without existing ground truth. The instance with the highest disagreement is queried next. Another presented group of algorithms – the **expected impact query-algorithms** – utilizes the expected impact of a new instance to a metamodel change of parameters, to the metric-function evaluated on the training-set, or to the metric-function evaluated on the validation-set (see section 2.2, page 42 ff). Finally, the authors discussed the risk of single-instance based algorithms, like the uncertainty-based ones, to query outliers or less relevant instances. As a solution, they presented **densities-weighted uncertainty-based algorithms**. Such an algorithm would reduce the chance of an unlikely feature attribute value combination to be queried.

An example of a recent application of community-based query-algorithms in combination with uncertainty-based query-algorithms was described by Gan and Gu (2019). The authors used the combination of an initial DoE and three metamodels. The uncertainties of each metamodel were individually assessed and regions of interest identified. The overlap of the three resulting regions of interest were used as the design space for further function calls. The authors proposed an iterative procedure with a convergence criterion.

Another approach was introduced by Schneider et al. (2022) on the example of generic functions and occupant safety assessment. Initially, a base dataset was generated by conducting FE-simulations with Total Human Model for Safety (THUMS) for a DoE. A metamodel was trained based on that data. The predicted values were understood as the design space. Upon discretizing this design space into a defined number of bins, the region of interest was defined as the bin with the highest error. Predictions instances without existing ground truth were generated and sorted into the existing bins. All instances, which came into the bin of interest, are the queried candidates.

Entirely on an uncertainty-based algorithm, based was the recent contribution of Ballal et al. (2023) for the example of a generic function and the application to pedestrian safety. The authors started with an initial DoE. A Gaussian Process Regression (GPR) model was fitted to the data. The confidence information of the model was leveraged to identify new samples to be generated. Similar approaches were presented by Azzimonti et al. (2021), Kaneko (2021), and Picheny et al. (2010).

2.3.4 Physics Informed Machine Learning

Physics informed machine learning holds great potential in generating efficiently more accurate metamodels. Karniadakis et al. (2021) presented a comprehensive review of such methods and defined the concept of bias as a key to infusing metamodels with physics. The authors categorized the observational bias, the inductive bias, and the learning bias. The **observational bias** can be understood as a weak mechanism for steering the learning process by presenting pre-selected data to the algorithm. For example, the symmetry of a vector field could be represented by duplicating instances with features for the symmetric location. The **inductive bias** is achieved by directly incorporating physics-related assumptions into the machine learning

architecture; the physical knowledge is enforced implicitly. The authors described Convolutional Neural Networks (CNNs) as one example for this approach, since the convolution and pooling operations lead to a representation of the extracted attributes regarding symmetry, for example. As further examples, Graph Neural Networks (GNNs) and methods like GPR where a kernel function can be selected were listed. Finally, the **learning bias** was defined as a modification of the loss-function penalizing non-physical behaviour. Additionally, the authors pointed out that all three bias-approaches can be used in combination, too. A special focus on the bias-approaches had the survey of Hao et al. (2022). The authors collected established approaches of incorporating the physics describing systems of differential equation into the loss-function.

A recent example of an inductive bias using architecture in the domain of vehicle safety was presented by Belaid et al. (2021). The authors utilized a CNN block to process a vehicle-pulse. In addition, pre-extracted characteristics of the vehicle-pulse, the information which Anthropomorphic Testing Device (ATD) was used, and its belting-status were processed as scalar feature attributes by some dense layers. Intuitively, the CNN block downsampled the vehicle-pulse, while the dense block mapped the scalar attributes to this downsampled space. The authors called this process the encoder part. The output of the metamodel was the driver's chest acceleration as a time-series. To generate the new time-series, a transposed convolution (see section 2.2, page 52) was performed in a so-called decoder block upsampling the encoded space. The result was passed through another dense layer to generate the final output. The authors used data from full crashes as physical testing. Additionally, they used FE-simulation to augment the database. The metamodel performance was found to be significantly better than the baseline in the form of the averaged chest deflection signal over the training-set.

Kayvantash (2018) proposed the decomposition of the result of a FE-simulation to reduce the complexity of the multi-attribute target. A FE-simulation result was decomposed into a spatial and a temporal operator in combination with the transfer matrix containing the singular values. The author claimed that for a small change in the factor space, it is sufficient to predict either the spatial or the temporal operator. By reversing the Singular Value Decomposition, the complete time-series can be restored, considering the change due to the factor shift. The principle was recently demonstrated by Diederich et al. (2023) predicting head kinematics of the MAthematical DYnamic MOdels (MADYMO) Active Human Body Model in driving manoeuvre scenarios. Both authors limit the methodology to small variations and relatively linear conditions.

Further relevant examples for dimensional reduction techniques as a form of observational bias of the metamodel were presented by Kracker et al. (2020) and Limousin et al. (2019). The first employed decomposition techniques to extract low-dimensional attributes from FE-simulation full geometry output on the example of b-pillars in a side-crash scenario. The latter projected finite elements into a voxel grid to extract meaningful characteristics of a crash FE-simulation while being mesh-independent. This new representation was compressed, e. g. by Principal Component Analysis (PCA), for further analysis. Bach et al. (2019) investigated the dimensional reduction of a FE-model to achieve shorter computation time. The solver was

equipped with an estimate of which parts of the model are relevant to the solution. By reducing the solved model size in such a way, a significant speed-up was achieved.

2.3.5 Transfer Learning

Another angle for efficient data utilization is offered by transfer learning. The basic idea of the category of techniques is to take advantage of the knowledge gained from one dataset to improve the predictions in other domains or tasks.

Zhao et al. (2024) provided a comprehensive literature survey on established approaches. For transfer learning, there is a source and a target domain on the one side, and a source and a target task of the metamodel on the other side. The actual transfer always originates from the trainable feature attribute extractor. The authors provided two approaches of categorization. The segmentation by the aspect of data and task led to the categories of unsupervised, transductive, and inductive transfer learning. The second "method examines whether the source and target domains share the same feature and target spaces" (Zhao et al. (2024), page 6), which led to the categories of homogeneous and heterogeneous transfer learning.

For **unsupervised transfer learning**, two databases with instances without an existing ground truth are available and two different but related tasks are to be solved. The feature extractor from the first data-task-pair is used to support the learning process for the second data-task-pair.

The category of **transductive transfer learning** deals with scenarios, where one database contains instances with an existing ground truth and another instances without an existing ground truth. Both databases originate in different domains, while the actual task is the same. As an example, Zhao et al. (2024) illustrated the recognition of tumours. The first database contained pictures of CT-scans and the ground truth on the existence of tumours. A CNN-based metamodel is trained on this dataset. The second dataset contained CT-scans of the lung. Again, the task was to recognize existing tumours. The trained metamodel was utilized to support the unsupervised learning with this second dataset.

Finally, **inductive transfer learning** assumes one database with instances with or without an existing ground truth. The second database contains instances with an existing ground truth, from the same domain as the first database. The first database is used to train a metamodel to perform a particular task. Next, the trained feature extractor is used to support the training for the second task.

Methods of multitask learning as part of the inductive transfer learning were collected by Vandenhende et al. (2022). The authors described that, first, a metamodel is trained on all available data to recognise feature attributes. Then, the data is split by the available target attributes with their ground truth and the related task trained.

Zhao et al. (2024) described **homogeneous transfer learning** to be used when data from two related domains of with the same metamodel input space is accessed. For **heterogeneous transfer learning**, only one database exists, but the used metamodel input space differs.

Specific for the PYTHON library KERAS and a CNN based metamodel, Chollet (2023) provided a compact introduction to homogeneous transfer learning. In the example, one pre-trained metamodel for picture classification was used for the estimation of new picture classes. The feature extractor architecture was replicated and the parameters from the trained metamodel was used to initialise the parameters of the new feature attribute extractor. In a first learning round, those parameters were held constant and the dense layers parameters trained. Next, the feature extractors' parameters were unfrozen, and the training continued with a smaller learning rate as so-called fine-tuning.

2.4 Approaches to Cover Human Body Diversity

2.4.1 Approaches using Physical or Virtual Testing

Depending on the selected model approach, the assessment of anthropometric variation leads to different levels of effort. VIRTHUMAN (VHUM) or models using Multi Body Simulation (MBS) can be scaled relatively easy in different dimensions, as demonstrated by Hynčík et al. (2013). Models like Total Human Model for Safety (THUMS) and Global Human Body Models Consortium (GHBMC) utilise finite elements, which means higher requirements on the algorithms.

Serre et al. (2006) proposed a control point interpolation based on the kriging method, also known as Gaussian Process Regression (GPR). Morphing based on radial basis functions is widely used – e. g. by Vavalle (2015) for target morphing. The different techniques were compared by Jolivet et al. (2015).

Due to ageing, humans change their average anthropometrics as well as their tolerance against loading. For example, Agnew et al. (2013) investigated the age-related effect on the stiffness of the ribs. They found a significantly higher risk of fracture in older subjects.

Cronin et al. (2012) studied the impact of the ageing effect on the validity of Human Body Models (HBMs) using the example of the neck of the GHBMC. They determined a higher failure load tension in the virtual model than in the reference Post Mortem Human Subject (PHMS). As a reason, they point out that the material properties of the GHBMC were developed for a 26 years old subject, whereas the reference PHMS was older. Schoell et al. (2015) compared the original 26 years old GHBMC responses with a 65 years old version. They found an increased probability for AIS 3+ injuries with the older model. Morris et al. (2003) and Brumbelow (2019) investigated the age effect using real-world statistics databases. Their conclusions seem to support the findings of the studies with the virtual models.

Motivated by the increasing obesity rate worldwide, numerous studies have analysed the effect of obesity on crash behaviour. An overview of the statistical evaluations and detailed virtual studies is provided by Parenteau et al. (2020). Cormier (2008) and Jehle et al. (2012) analysed the data in National Highway Traffic Safety Administration (NHTSA) Fatality Analysis Reporting System (FARS). Rupp et al. (2013) used data from the National Automotive Sampling System (NASS). Ryb and Dischinger (2008) based their analysis on data from the Crash Injury

Research Engineering Network (CIREN). For the virtual assessment, Kim et al. (2014) used the MAthematical DYnamic MOdels (MADYMO) model of the Hybrid III (HIII) Anthropomorphic Testing Device (ATD) and extended it with a fat vest. Shi et al. (2015) morphed the THUMS model based on an obese PHMS. Forman et al. (2009) investigated obesity by running three tests with obese PHMS and two subjects near the 50^{th} percentile, focussing on chest injuries. All studies point to an increase in mortality and the risk of severe injuries in obese occupants.

Klinich et al. (2000) collected case reports on the implications of the restraint system in case of pregnancy. They concluded that the seatbelt could have undesirable effects. It should still be used as the risk of adverse foetal outcome is higher with not properly applied restraint system. Moorcroft et al. (2003, 2004) and Duma et al. (2004, 2005) extended the 5^{th} percentile in MADYMO to model pregnancy. They found that the restraint system should be adapted to pregnancy and that new injury assessment thresholds should be developed. During their studies, they investigated how the seatbelt and airbag effect the uterus. They recommend using a 3-point-seatbelt, an as-low-as-possible worn lap seatbelt, and a sitting position with a high distance to the interior.

The effect of the occupant's gender is the content of active discussions. As stated above, the state-of-the-art safety assessment uses the male representation for the 50^{th} percentile and for few cases the females' 5^{th} percentile. Projects like Virtual Vehicle Safety Assessment (ViVA) introduced female representations. In terms of testing in context of gender, an under-representation of female mimicking surrogates can be and is stated. This could be a major issue, if the gender-specific outcomes of crashes are statistically significant while having high effect-size and the differences in the crash related biomechanics clear. Abrams and Bass (2020) analysed the NASS database for gender-specific differences. They could not determine any significant general differences. However, women appear to have an increased injury risk at a higher age than men in the same age group.

Carter et al. (2014) approached the impact of age, gender, and Body Mass Index (BMI) by analysing data in the NASS database. As above, they found for older occupants that females have a higher injury risk than males. A higher BMI was found to increase the injury risk in the thorax region. They concluded that the effects should be analysed by the usage of modified computational models.

Hu et al. (2016, 2017, 2019) performed a series of studies utilizing the GHBMC and the radial-basis-function-based morphing method. They performed sled and crash FE-simulations and used regression models to determine the sitting position previously published by Park et al. (2016). In the most recent study, Hu et al. (2019) analysed 100 morphed GHBMC in a sled model. The findings were a general higher risk of head injuries for taller occupants, increased injury risks for femur and pelvis injuries with higher BMI, and for older subjects, a higher risk of thorax and femur risk. Izumiyama et al. (2018) investigated the effect of different spine alignments using THUMS. They observed differences in the injury measures and an increased tendency to submarining when the pelvis is more tilted.

The effect of the presented variations on an occupant on the crashworthiness was the subject of interest in numerous studies. In general, it can be concluded that a deviation from the state-of-the-art development configuration usually leads to higher injury risks. Due to its flexibility, virtual modelling seems to be a promising approach to counter the various challenges. However, advanced HBMs are highly computationally expensive. A detailed model can, even with modern hardware, take hours to days for one FE-simulation. New and more efficient approaches should be found to explore the assessment within this new and drastically increased parameter space.

The development of surrogate models for detailed virtual models seems to be one possible way to overcome this limitation. Raja and Swain (2017) defined, for example, spring damper formulations to model the occupant. Another increasingly investigated approach is to utilise data analysis and machine learning.

2.4.2 Approaches using Machine Learning

2.4.2.1 From Real-World Data

Craig et al. (2011) investigated using data from the car's Event Data Recorder (EDR) if the occupant's injury can be predicted to provide fast information, e. g. to hospitals. The authors compared the results with cases from the NASS and Crash Injury Research Engineering Network (CIREN) databases. Meaningful parameters have been identified.

Hasija and Takhounts (2019) attempted to predict the Head Injury Criterion (HIC) criticality from crash-test information. They utilised the available data in the NHTSA database to train the logistic regression algorithm for a classification task. They achieved a F1-score of around 0.8.

Using NASS data, Lee et al. (2018) tried to estimate the occupant injury grade using the decision tree algorithm. Liu et al. (2020) continued the study and compared boosting algorithms with Support Vector Machine (SVM) and Artificial Neural Network (ANN). Both studies achieved a maximum accuracy of around 80 %. Chong et al. (2005) developed an ANN, a decision tree algorithm, and a combined estimator of both methods to predict the injury severity in three classes using NASS data. The best class accuracy reached from 60 % to 90 %. The authors concluded that better results can be achieved if the data contained the impact speed. It is known to be one factor with a high correlation with injury severity.

2.4.2.2 From Virtual Physical Models

The use of machine learning approaches with data from virtual simulation holds great potential. Not always is the human diversity the centre of research. However, successful approaches can, in general, accelerate the assessment and hence enable the coverage of the broad variety of the human body.

Zhao et al. (2010) performed 98 crash and occupant FE-simulations under the variation of the seatbelt and airbag factors. They used the varied factors to estimate a weighted injury-criterion in three classes. The accuracy for 11 instances from the testing-set was found to be 83 %.

Sequeira et al. (2020) published 131 car-to-car crash FE-simulations varying the impact angles. Vehicle level FE-simulations were carried out, followed by occupant (driver) level FE-simulations. They proposed using the data to train an occupant injury prediction metamodel. The metamodel should use available near-real-time information to decide on the activation of airbags.

Yang et al. (2020) formulated the proposal to approach the real-time prediction by using look-up tables. They used 9 FE-simulations with THUMS, varying stature, BMI, belt load limiter, and the backrest angle. The factors were used as feature attributes for the regression methods.

Bance and Nie (2019) proposed a multiple-step approach to achieve near-real-time prediction of the injury risk. First, they generated a limited dataset with a MADYMO sled MBS model. Then they used the data to configure a lumped mass model. This surrogate model was used to generate approx. 5690 instances, which were fed into an ANN. The estimation of the overall injury level was rated with an accuracy greater than 70 %. The prediction of peaks was found to require more research. The same model was used later by Bance et al. (2021) to achieve the near-real-time behaviour in an entirely deterministic manner in the form of two-dimensional lumped mass models. The study was extended by Chen et al. (2020). They proposed to chain a road scenario creator to define the initial impact conditions. With vehicle-barrier MBS simulations, they generated the crash pulse fed into a vehicle occupant model. The real-time features and the occupant MBS simulation responses were used to train a prediction algorithm. The metamodel accuracy of over 70 % seemed satisfactory, with the limitation that only 23 instances were used to determine this value.

Joodaki et al. (2021) conducted 450 FE-simulations with HBM. The purpose was the restraint system optimisation on a so-called 'Life Years Lost' criterion. Multiple learning algorithms were trained, tuned, and compared. The metamodel feature attributes were the Design of Experiments (DoE) factors and the target attribute, a so-called Life Years Lost criterion. The general problem was stated as a regression. The optimisation approach was understood as establishing response surfaces, hence accelerated function calls. However, the focus of the study was on the comparison of the learning algorithms and the effect of hyperparameter tuning.

Hasija and Takhounts (2022) developed a combined metamodel using Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) architectures to estimate head accelerations using only video recordings of a head during the accident. The authors used the GHBMC and the FE-model of the 2014 Honda Accord in front crash scenarios. In total, 1010 FE-simulations were conducted and the head animation extracted. For the predicted peak angular head velocities, the authors claim not nearer specified correlation coefficients from 0.73 to 0.92. A closely related approach using deep learning architectures designed for picture classification was proposed by Kunitomi and Konosu (2018). They used screenshots from different time frames of a car to pedestrian crash MBS simulation in MADYMO as a feature to predict three criticality classes of the HIC. The accuracy was greater than 80 % for pictures taken later than 60 ms after the collision.

2.4.2.3 Conclusions on the Role of Machine Learning in Occupant Safety

The prediction of injury severities using real-world data is mainly used to identify the most influential factors leading to injuries. Other authors tried to use the data to estimate the accurate injuries. It can be concluded that the databases may not be detailed enough for precise injury predictions.

Virtual models' usage as a deliverer of the data fed into machine learning algorithms has advantages compared to real-world data usage. First, the amount of information generated by each simulation can be set according to the needs. Additionally, it is possible to retrieve information that was initially not explicitly collected. Second, virtual simulations can be automated to a high degree. This automation enables generating the vast amount of data required for machine learning, given the appropriate computational power and efficient models.

Similar to occupant protection, in other passive safety domains, machine learning approaches are also used. There is a potential to transfer them to the domain of this research.

3 Data Generation via Finite Element Analysis

This chapter revolves around FE-simulations. The process of model selection and simplification, including conducting a DoE is displayed. The selection and simplification follows an initial requirement analysis. The central requirement is a computation time low enough to allow high volume DoEs. The stepwise simplified FE-models alongside the analysis from this chapter and four generated databases are supplemented by Plaschkies (2024c,d,e,f) and Plaschkies and Müller (2024). The related source code was published by Plaschkies (2024a).

The simplification process starts by defining the sitting position of an ATD. Each step follows the principles of efficient FE-simulations while ensuring compatibility by direct comparisons.

The DoE is prepared by a study of selecting meaningful factors. Afterwards, the DoE is defined, and the simulations are conducted accordingly. The results are displayed as summary statistics.

3.1 Selection of Vehicle Model

3.1.1 Requirements and Available Models

3.1.1.1 Requirements and Development Environment

According to the scope of this dissertation, there were multiple requirements for a physical virtual simulation model. The selected model should enable extensive data generation to fuel the subsequent studies. As the domain, passive occupant safety was pre-defined. Additional requirements were given through the context in which this research was conducted. Load cases, which are typical for frontal occupant protection and described by legislative bodies and consumer testing organizations, were considered. To take non-linearities into account, offset load cases like the Offset Deformable Barrier (ODB) were viewed as necessary.

The defined domain led to the critical components a model should contain. Firstly, a vehicle structure capable to generate vehicle-pulses is required. Second, the assessment of occupant responses necessitates a human surrogate like an Anthropomorphic Testing Device (ATD). Third, for realistic responses, the vehicle interior and restraint systems must exist. All, structure, interior, and restraints should orient on the state-of-the-art of passive vehicle safety measures. The dashboard, carpet, seat, Driver Airbag (DAB), Passenger Airbag (PAB), and seatbelt were identified as crucial systems. Extended systems, such as Curtain Airbag (CAB), knee airbag,

side airbag, far side airbag, or seat ramp airbag were considered optional, depending on the availability and necessity in the selected load cases.

To enable extensive simulations required to assemble a sufficiently sized database, a model must have a reasonable low computation time relative to the given computer-hardware capacity. At the author's institute, a cluster of Central Processing Units (CPUs) was available, consisting of 17 nodes, each equipped with two Intel Xeon E5-2687W v4 CPUs. Hence, in total, 408 physical cores could have been utilized. To enable multiple thousand simulations and some flexibility in studying the effects of the selected Design of Experiments (DoE), a computation time under 30 min on one node of the cluster seemed reasonable. Furthermore, the available computer-hardware required a software architecture optimized for CPU clusters.

The modelling of human surrogate responses in vehicle crashes is usually done using numerical methods. The stability of such models over repeated simulations and against changes in the computer-hardware are crucial for reliable results. The second base for reliability is the validation of a virtual model against its hardware-model counterpart.

To support the scientific community, the model should be free to use and results be publishable. Furthermore, for the planned DoE the parametrizability is another requirement.

3.1.1.2 Landscape of Available Models

The model selection was done at the end of 2019. At this time, various vehicle models for the Livermore Software technology corporation DYNA (LS-DYNA) finite element software were available. However, most of the models contained the structural components only, where the interior parts were not included. For other software packages, models were purchasable but not publicly available.

The LS-DYNA FE-models are widely used in the scientific world. For example, Sequeira (2022) adapted the Toyota Yaris model from George Mason University, Center for Collision Safety and Analysis (2016) for the use in the context of occupant safety. Since the restraint components were not available as data from a hardware-model, the FE-simulation results could have been checked only regarding their general plausibility.

Other candidates were the 'Gold Standard Sled Buck' created by THUMS User Community (2019), the LS-DYNA learning aid FE-model, the Virtual Vehicle Safety Assessment (ViVA) test buck as used by Östh et al. (2017), and the proprietary FE-models from TECOSIM Group GmbH and A2MAC1 Group (2022). Those FE-models were either not publicly available, or highly simplified without being directly related to a hardware-model test, or did not contain any restraint systems.

At the end of 2019, the only model, which includes structural and restraint components and is freely available, was the Honda Accord 2014 FE-model provided at National Highway Traffic Safety Administration (2019) as per the author's knowledge. In 2022, the Nissan Rogue FE-model was released by George Mason University, Center for Collision Safety and Analysis (2022). However, the schedule did not allow the switch to the more modern vehicle.

All above-mentioned alternatives were modelled using Finite Element Analysis (FEA). An approach based on Multi Body Simulation (MBS) was available within the Simcenter software by Siemens Digital Industries Software (2024a). Despite the expectable faster computation time in comparison to methods based on FEA, no open-source MBS models were available as per the author's knowledge.

Other options could have been 2D-models, as used by Raja and Swain (2017) or later Plaschkies and Vaculín (2020). Although, such approaches cannot take lateral movement of vehicle or human surrogate into account. This issue was tackled by Bance et al. (2021), who used in total four two-dimensional lumped mass models for the main load directions (front, rear, near- and far-side). The prediction accuracy ranged from moderate to good. However, as one of the focuses of this work is the assessment of lateral non-linearities, two-dimensional models were found to be not suitable in general.

Ultimately, the best choice seemed to be the Honda Accord 2014 model. The FE-model came with the ATD human surrogate Test device for Human Occupant Restraint, 50th Percentile Male (THOR-50M) sub-FE-models for driver and passenger positions and featured a validation reference report on both, crash and ATD responses. Newer versions of THOR-50M were made available by the University of Virginia School of Engineering and Applied Science, Center for Applied Biomechanics (2024) and various FE-models of the ATD human surrogate family Hybrid III (HIII) as published by Maurath et al. (2021) were available within the LS-DYNA licence package.

3.1.2 Model Analysis and Load Case Selection

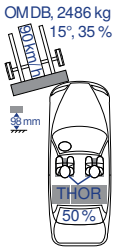
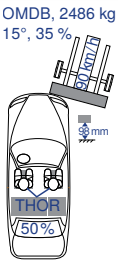
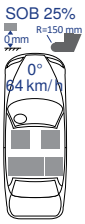
3.1.2.1 Description of the Available Honda Accord 2014 FE-model

The selected FE-model represents a Honda Accord 2014 midsized saloon. The FE-model development was done in two steps. First, Singh et al. (2018b) created a structural FE-model with the intention to assess the vehicle crash response. This FE-model was compared with crash tests of a hardware-model for the load cases, as displayed in table 3.1 on the next page.

In the table 3.1 on the following page, a checked mark ✓in the post-crash row indicates the availability of photos or snapshots taken either after the hardware-model test or at the end of a simulation. A checked mark in the vehicle-pulse row relates to the signals from one of the vehicle's accelerometers, and a checked mark in the ATD row marks the availability of signals from the ATD sensors. The barrier types are the Full Width Barrier (FWB) and the offset barriers Oblique Movable Deformable Barrier (OMDB), Offset Deformable Barrier (ODB), and the Small OverLap Barrier (SOLB), each for left and right side.

The accelerometers which were defined in the FE-model are displayed in figure 3.1 on the next page. In the reference report, there is no clear statement declaring which sensor location was used to generate the provided graphs. It seems likely the used sensor location is one of the sills or the Centre of Gravity (CoG) as those are often used in practise.

Table 3.1: Available Load cases and Data (Graphics adapted from Safety Companion *2024)*

Barrier	FWB	OMDB-L	OMDB-R	ODB-L	ODB-R	SOLB-L	SOLB-R
							
Post Crash	✓	✓	✓	✓		✓	
Vehicle-pulse	✓	✓	✓				
ATD		✓	✓				

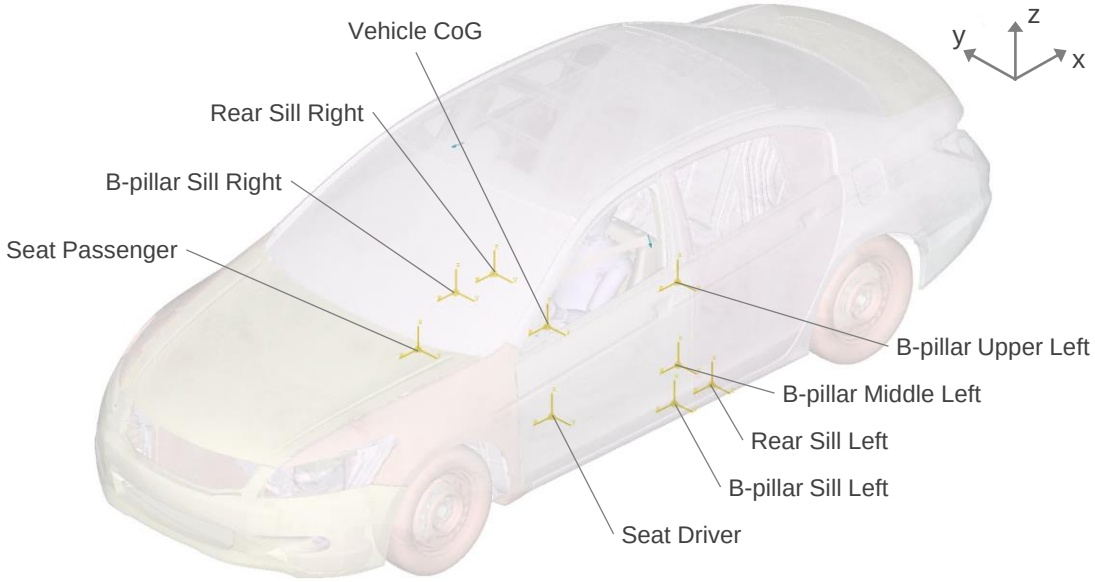


Figure 3.1: Location of Acceleration Sensors of the Honda Accord's BIW

Subsequently, Singh et al. (2018a) extended the reference report with the pure structural FE-model by interior, restraint systems, and ATDs. The ATD responses were correlated with data from the hardware-model in the Oblique Movable Deformable Barrier (OMDB) load cases. As compared to the ODB, in the OMDB load case, the barrier is not fixed in space but is movable.

Following the reference reports, the FE-models were created by analysing the components of the hardware-model. A satisfactory agreement between the FE-model and hardware-model, regarding the vehicle's overall kinematics, was reported for Oblique Movable Deformable Barrier Left (OMDB-L) and Oblique Movable Deformable Barrier Right (OMDB-R) load cases. This judgement was underlined by CORrelation and Analysis (CORA) ratings (see section 2.1, page 29 f) over 0.8 for X-acceleration and over 0.9 for the X-Y-velocities. Similar results were found for the other load cases.

The FE-model restraint systems were validated by single component hardware-models and finally correlated with the full hardware-model in the OMDB load cases. The B3-seatbelt forces showed good agreement, expressed by CORA ratings over 0.8 (see section 2.1, figure 2.1 on page 8 and ISO 18571 on page 29 ff).

Insufficient Passenger Airbag (PAB) positioning was observed, as presented in figure 3.2. In the OMDB-R load case, the ATD head missed the PAB in the FE-simulation completely, whereas in the hardware-model, the PAB catches the passenger. In the OMDB-R load case, the passenger seatbelt slips off the shoulder. The seatbelt slip occurs in the FE-model as well as in the hardware-model. Consequently, the CORA ratings of ATD head accelerations range from 0.6 to 0.8 and the pelvis accelerations from 0.5 to 0.7. For femur forces, CORA ratings under 0.6 were achieved only.

Figure 3.2: Hardware Test versus FE-model in Oblique Movable Deformable Barrier Right (OMDB-R) Load case @ 90 ms in Reference report (adapted from Singh et al., 2018a)

The FE-simulations were conducted with LS-DYNA Version 7.1. No information on the computer-hardware was given. Furthermore, the parallelization technique (see section2.1,

page 18) was not specified. For those reasons alone, it seemed necessary to conduct an initial comparison between simulations ran by the author and those from the reference report.

3.1.2.2 Finite Element Model Analysis

The original FE-model contains THOR-50M on each driver and passenger seat. Both ATDs were restrained by a seatbelt. The FE-model contains a DAB, PAB, and on each side a CAB. The airbags were modelled by using LS-DYNA hybrid formulation (see section 2.1, page 20). For the DAB, the jetting technique was added. The dashboard was modelled in high detail using shell elements. Carpets and seats are also detailed by using solid elements for the foam components. In total, the FE-model consists of approximately 3.5 million elements with an average edge length of ca. 6 mm.

The LS-DYNA FE-model was split into one file per component, such as ATD, airbag, and barrier. Each file represents a sub-model, which is independent of other parts of the FE-model and is located in its distinct numbering range. Interrelations like contacts and connections are defined in separate files. The whole model is assembled using one central file. This data organization strategy is in accordance with the industry's best practice and enables the changes to a single sub-model without affecting other parts of the whole FE-model.

During the studies, only LS-DYNA versions higher than version 12 were available. All new FE-simulations were conducted with LS-DYNA version 12 using the multi-node parallelization technique Message Passing Parallel (MPP). Despite LS-DYNA backwards compatibility, changes in the algorithms have to be assumed. In this context, a potential problem is the usage of default values as blanks. In this case, the solver is setting the version-specific defaults. Even by checking the published release notes, due to the sheer mass of changes, potential conflicts could not be fully ruled out.

Another pitfall was found in the integration interval control setting. The original FE-model was distributed with the setting of deletion of elements whose size shrank to a integration interval reducing size. While this setting can mimic meaningful material failure in a pure crash FE-simulation, in the case of ATD FE-simulation, contact failure due to deleted elements in the ATD were observed. Disabling the feature was required for stable simulations.

The defined mass scaling option led to an initial mass increase of 8 kg per hand and 7 kg per arm. Mass scaling is a technique to keep the integration interval stable by increasing the mass of an element by a small amount (see section 2.1, page 18). The additional mass corresponds with approximately 40 % of the physical mass of the respective component; a significant influence on the related kinematics was suspected. Additionally, notoriously instable simulations enforced the update of the ATD FE-model. The reference report stated issues with instability as well. The proposed countermeasures seemed to be integrated, but details were not specified.

To stabilize the contact situation while keeping a reasonable integration interval, the mass scaling option seemed necessary. At the same time, the resulting local mass increase of the THOR-50M FE-model were not acceptable. The solution was the update of the FE-model from

version 2.1 to the version 2.7 released in 2019, provided by University of Virginia School of Engineering and Applied Science, Center for Applied Biomechanics (2024).

Comparing both THOR-50M FE-model versions, the version 2.1 resembles the 2010 drawing package released by National Highway Traffic Safety Administration (NHTSA), whereas the version 2.7 represents the update from 2018. The update contained slight changes in shoe and upper leg geometry, and a pelvis wrapping jacket. Following Panzer et al. (2019), further updates were general remeshing, fidelity, stability, and an extension of the certification cases. In figure 3.3, the improved mesh quality and the fully modelled jacket is shown. In summary, minor posture adjustments were necessary and slight differences in the ATD response in comparison to the reference report are expectable.

Figure 3.3: Visual Comparison of THOR-50M FE-models – Versions 2.1 (left) vs 2.7 (right)

FE-simulations for the OMDB load cases were carried out with THOR-50M version 2.7. Additionally, the Full Width Barrier (FWB) load case, and the ODB load case, as defined by the Insurance Institute for Highway Safety (IIHS), also known as moderate overlap, were added. For all load cases, the reference reports by Singh et al. (2018a,b) provided vehicle crash characteristics. ATD responses were only available for the OMDB load cases.

The not specified signal filter procedure and the not fully defined positioning of the evaluated vehicle's acceleration sensor can be identified as the limitations of the comparisons. For further comparisons, the sensor under the rear sill on the driver or passenger side was deemed appropriate. This sensor position minimizes the interference of local deformations with the overall vehicle motion. An alternative is the sensor location at the vehicle's CoG. In general, all signals will be filtered by a Channel Frequency Class (CFC) 60 filter according to SAE J211 (see section 2.1, page 31).

A comparably strong filter as the CFC 60 is only proposed for overall vehicle motion comparisons as described in the standard. However, as shown in figure 3.4 on the following page, in this particular FE-model, some signals were oscillating heavily. To enable comparisons, a

strong filter had to be selected. This is justified because the overall tendency was considered more important than the analysis of minor effects. One reason for the strong oscillation is the contact of the ATD head to the chest. When the head is not properly caught by the front airbag, the nodding motion is not restraint and, hence, a contact can occur.

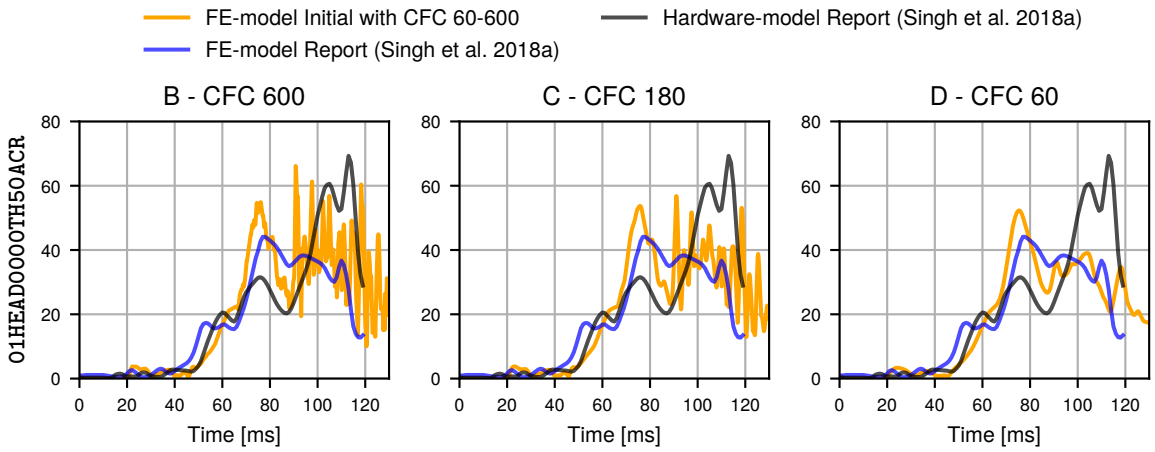


Figure 3.4: Different CFC Filter Class Strength on the Example of Oblique Movable Deformable Barrier Right (OMDB-R) Load case with Oscillating Signal of Initial FE-model versus Unchanged Data from Reference report by Singh et al., 2018a

During this work, additional model updates were released. An update for the PAB was published by National Highway Traffic Safety Administration (2024). Nevertheless, as described below, this work's focus is on fast simulations. Hence, the PAB model had to be simplified, and the update was not implemented. In March 2024, THOR-50M 2.9 was released by the University of Virginia School of Engineering and Applied Science, Center for Applied Biomechanics (2024). Following Panzer et al. (2024), the major geometry remained the same. The primary purpose of the update were stability and validity improvements. Because this work's focus was not on a Test Device for Human Occupant Restraint (THOR) but fast simulations, a later change of the ATD type was considered anyway. Thus, the numerical stability in the tested cases alone was seen sufficient and further updates were neglected.

3.1.2.3 Load Case Selection

For the subsequent tracking of the influences of each simplification, load cases had to be selected. Ideally, both, vehicle-pulse and ATD signals would be available from the reference report. Table 3.1 on page 72 shows possible candidates. Only for the OMDB load cases, the ATD data was available for validation. But here, as described above, partially bypassing the restraint systems, the ATD undergoes a strong lateral motion.

The OMDB load case, introduced around 2015, has not been implemented by the regulatory body. As the Honda Accord was released already in 2014, this new load case was most likely not considered during the development. Following Singh et al. (2018a), the interior model was

generated on the base of the original vehicle. Thus, the special implications of that new load case were not yet addressed.

To characterize the new OMDB load case, Barbat and Li (2017) conducted an early study. They observed a higher crash severity for medium-weighted vehicles compared to the classic FWB load case. Additionally, they compared the longitudinal and lateral acceleration of the vehicle for OMDB and ODB load cases on the example of the Hyundai-Elantra midsized saloon. In the X-direction, the ego-vehicle-pulse for the OMDB load case tended to be equivalent to the FWB load case. In contrast, the ODB load case led to a relatively late timestep for the maximum of the acceleration signal. The lateral acceleration in the ODB load case was approximately $10\,g$ over a longer period, which is significant but low in comparison to the approximately $20\,g$ in the OMDB load case. The Honda Accord, which is used in this thesis, can be considered as the same vehicle class as the Hyundai.

Figure 3.5 on the following page confirms the findings of Barbat and Li (2017) for the Honda. A shortcoming is the incomplete data for this load case in the reference report, for which reason the new FE-model data was used here. The vehicle motion for OMDB, ODB, and FWB load cases were evaluated at CoG. All vehicle-pulses were taken from the initial FE-model.

The Occupant Load Criterion (OLC) (see section 2.1, page 36) for the X-motion was calculated. Based on the start and end times of the idealized restraint-phase, a lateral OLC from the Y-velocity was determined. This lateral OLC has no direct interpretation like the original definition with the free-motion and restraint phases but enables a quantitative comparison of the lateral vehicle-pulses.

Visually and numerically, the FWB load case resulted in the vehicle-pulse with the highest OLC. In lateral cases, the impact on the left side resulted in a higher OLC and less lateral motion compared to the impact on the right side. The ODB load cases resulted for both impact directions in ca. $2\,g$ lower OLC in X-direction than in the FWB load case. In addition, the Y-motion of the OMDB load cases were higher. Typically for ODB load cases, the vehicle-pulse are a long-stretched ones. In comparison, both, OMDB and FWB load cases resulted in an early hard vehicle-pulse.

Other alternative load cases for which the original FE-model was developed were the FWB from NHTSA and the small overlap defined by the Insurance Institute for Highway Safety. No ATD data was available for both, but the vehicle-pulse was provided for the FWB load case. Moreover, the main objective of the data generation was to fuel the machine learning. Here, functional restraint systems are necessary. Other load cases like the small overlap, European New Car Assessment Programme (Euro NCAP) Movable Progressive Deformable Barrier (MPDB) or different impact speed were neglected as they do not contribute with any additional benefits in the context of this work's scope and were not explicitly part of the FE-model reference report.

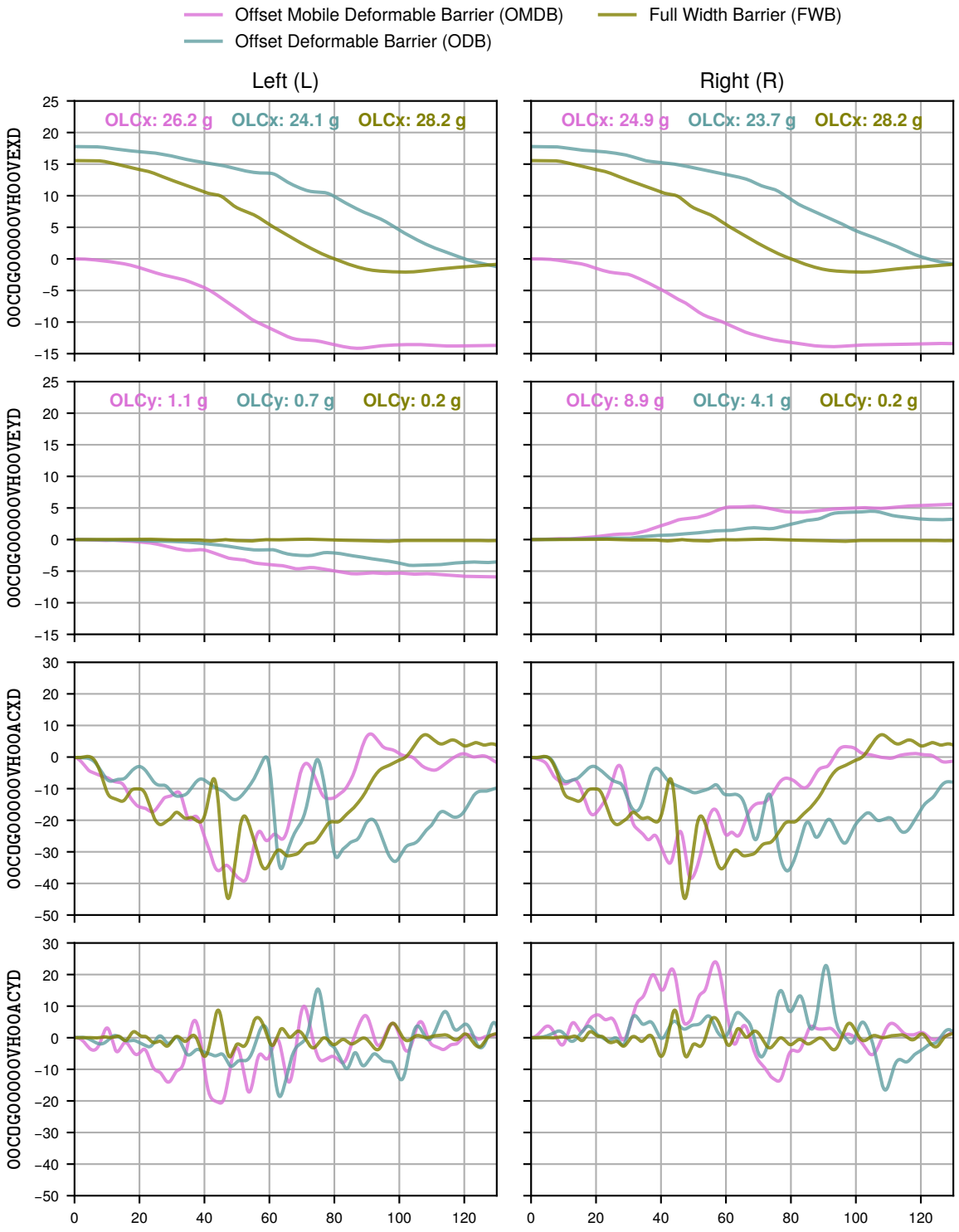


Figure 3.5: Vehicle Motion Comparison with Initial FE-model at BIW CoG for different Load cases (Single Precision, 48 Processes, CFC 60)

Since one of the objectives of this research was to explore the limits of the method, lateral motion could not be neglected completely. Thus, the OMDB load cases were used to ensure that no major issues occurred in comparison to the reference report. Since the Honda Accord was not optimized for the OMDB load case, a less intense lateral vehicle-pulse seemed necessary to ensure stable FE-simulations. Consequently, the ODBs load case were selected as a load case with a reasonable but less intense lateral motion to evaluate further steps for FE-model simplification. In addition to the load cases with lateral kinematics, the FWB load case was used for the direct comparison with the reference report and the subsequent evaluations.

3.1.3 Occupant Sitting Position

To satisfy the purpose of delivering data for machine learning, occupants on the driver, passenger, or both positions could be used. Considering the effort necessary to develop a fast and robust FE-simulation environment, the selection of one ATD position seemed beneficial. Another implication of the machine learning architecture is the requirement of same or at least dependant vehicle and restraint system settings for different ATD sizes. This major limitation excludes all systems with an adaptive restraint system whose configuration is dependent on any indication of the ATD anthropometry.

In practice, a robust occupant detection is assumed, and the restraint system settings are optimized individually. The seat's adjustment as a predictor of the occupant's size is mentioned as the state-of-the-art by Schöneburg (2023). An extended system using additional weight sensors and an occupant's head position detector was described by Takata Corporation (2017). As per the author's experience, the simple seat position sensors seem to be the more common ones.

In the section 2.1 on page 11 ff, the collected literature and data on probable sitting positions are presented. Following the data from Crash Investigation Sampling System (CISS), overall, all anthropometrical parameters have a wide spread on each seat track position. For drivers, it was found that there was a correlation to height, and a weaker correlation to weight. For passengers, the forward most position was populated predominantly by smaller occupants. No trend was recognized for the positions behind. Especially around the middle seat track position, all types of occupants were found.

Following the above argumentation, passengers with the seat track position "middle" were found to cover a wide spread of anthropometries. This position and no adaptivity of the restraint systems to different occupants were selected as the base for further investigations.

3.1.4 Initial Model Comparison

To investigate the numerical stability of the FE-model, each FE-simulation ran 4 times. In each repetition, either 48 or 96 parallel processes were used. Additionally, the computation in single precision or double precision was set (see section 2.1, page 18). Both, parallelization and precision can change the results of a FE-simulation. A good model should be robust against the

spacial decomposition strategy. For reasons of lower computational and storing cost, single precision is preferable.

Instead of CORA values, the more recent ISO-rating defined in ISO/TS 18571:2014(E) (see section 2.1, page 29 ff) was used for the time window of 20 ms to 120 ms. In this time window, the major event, from the occupant's perspective, happens. In the time before 20 ms, the signal of a sensor differs only minimal between different tested FE-models. Hence, the ISO-rating would unreasonably improve. After 120 ms, mainly the rebound of the ATD takes place, which is not viewed relevant for the conducted study. If not different indicated, the ISO-ratings are reported as median in the four computational configurations. Since for this study mainly big effects were of interest, only responses resulting in ISO-ratings lower than 0.8 were inspected in detail. The ISO-ratings were reported either as the numerical value or by the corresponding ISO-grade (see table 2.3 on page 30).

In figure 3.6, the resulting computation times for the selected load cases were displayed. For single precision, doubling the parallel processes almost halves the computation time. Simulations in double precision required more time and larger storage in general. The FWB load case resulted in the shortest computation time because of the simplest barrier setup. The higher overhead from transferring bigger data, however, reduces the effect of parallelization. In case of double precision, the overhead effect was so dominant that the computation time even increased.

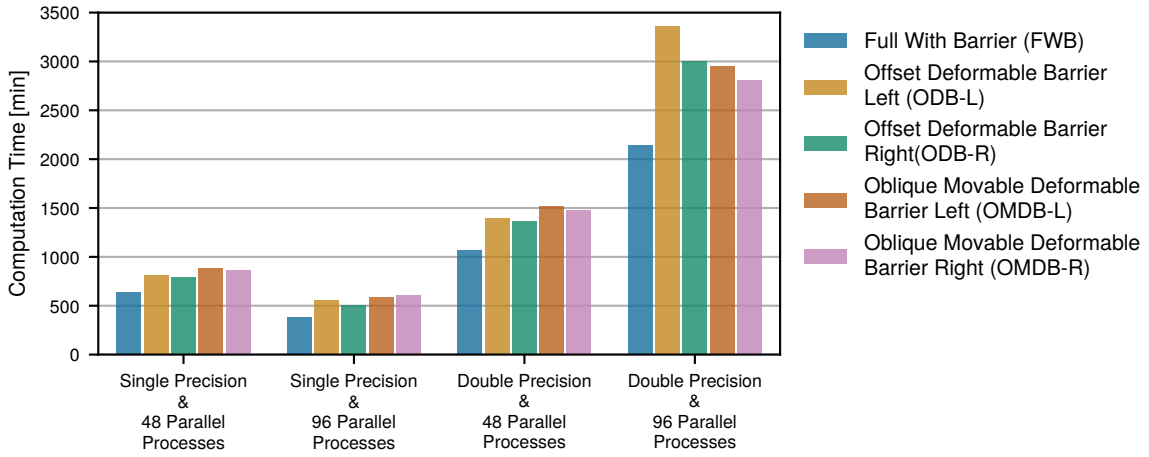


Figure 3.6: Computation Time of Initial FE-model for Full-crash FE-simulations with THOR-50M as Driver and Passenger

In the comparison of the initial FE-simulations with those in the reference report, the FWB load case for vehicle-pulse evaluation and the OMDB load cases for vehicle-pulse and ATD assessment were used. ISO-ratings are reported using the FE-model data in the reference report as reference. The median of the four ISO-ratings from each numerical setup was reported. The selected channels correspond to the ones available in the reference report.

Different line colours correspond to the three data sources. Different line styles indicate the four numerical setups:

- Continuous are FE-simulations Double Precision (DP) with 48 processes.
- Dotted are FE-simulations Double Precision (DP) with 96 processes.
- Dash-Dotted are FE-simulations Single Precision (SP) with 48 processes.
- Dashed are FE-simulations Single Precision (SP) with 96 processes.

When comparing two ATD FE-simulations, first the outer conditions should be correlated. If the vehicle-pulse deviates too strong, no other signal can match any more. In the second step, the restraint system should be correlated. Only if their responses are similar to each other, the ATD sensor's output is comparable.

In figure 3.7 on the next page, all vehicle-pulse data and the shoulder B3-seatbelt force according to the conventions defined in section 2.1 table 2.5 on page 32 and figure 2.1 on page 8 are presented. In the FWB load case, only the X-direction of the vehicle-pulse was provided by the reference report. The ISO-rating for the X-velocity was 0.94 and the X-acceleration 0.51. In general, the strong oscillation of the acceleration signals carries the chance of local deviations and hence lower ISO-ratings. The better alignment after 60 ms in the X-accelerations of the initial FE-model versus the hardware-model from the reference report in comparison to the FE-model from the reference report lead to the much higher ISO-rating of 0.72 while the signals before 60 ms were relatively similar.

The ISO-ratings in figure 3.7 on the following page of the vehicle's velocity in the OMDB-L load case were found to be almost perfect, with 0.96 in X- and 0.77 in Y-direction. The X-acceleration was ISO-rated at 0.73, while the Y-acceleration's higher amplitude of resulted in 0.41. However, the FE-simulation of the reference report was ISO-graded only as Poor with regard to the hardware-model as well. For the impact of the OMDB load case on the passenger side, similar ISO-ratings were found. For the acceleration, the average signal seems to match well to the reference report, but the oscillation around the average had a significant higher amplitude.

The vehicle motion in the OMDB-R load case matched to a similar degree as in the OMDB-L load case. The ISO-ratings were overall found to be only slightly lower – 0.94 for the X-velocity, 0.64 for the Y-velocity, 0.66 for the X-acceleration, and 0.33 for the Y-acceleration. The kinematics in the Y-direction seems to match better to the initial FE-model than to the reference report FE-model.

For the restraint system in figure 3.7 on the next page, only shoulder seatbelt B3-forces were available. Additionally, in the reference report in a constant pull-out test for the load limiter, increasing forces in the retractor force ${}^{\mathrm{B0}}F_{R,t}$ from 2.3 kN to 2.5 kN were measured. This behaviour was defined in the FE-model accordingly. No information regarding the friction between the seatbelt and D-ring was provided in the reference report.

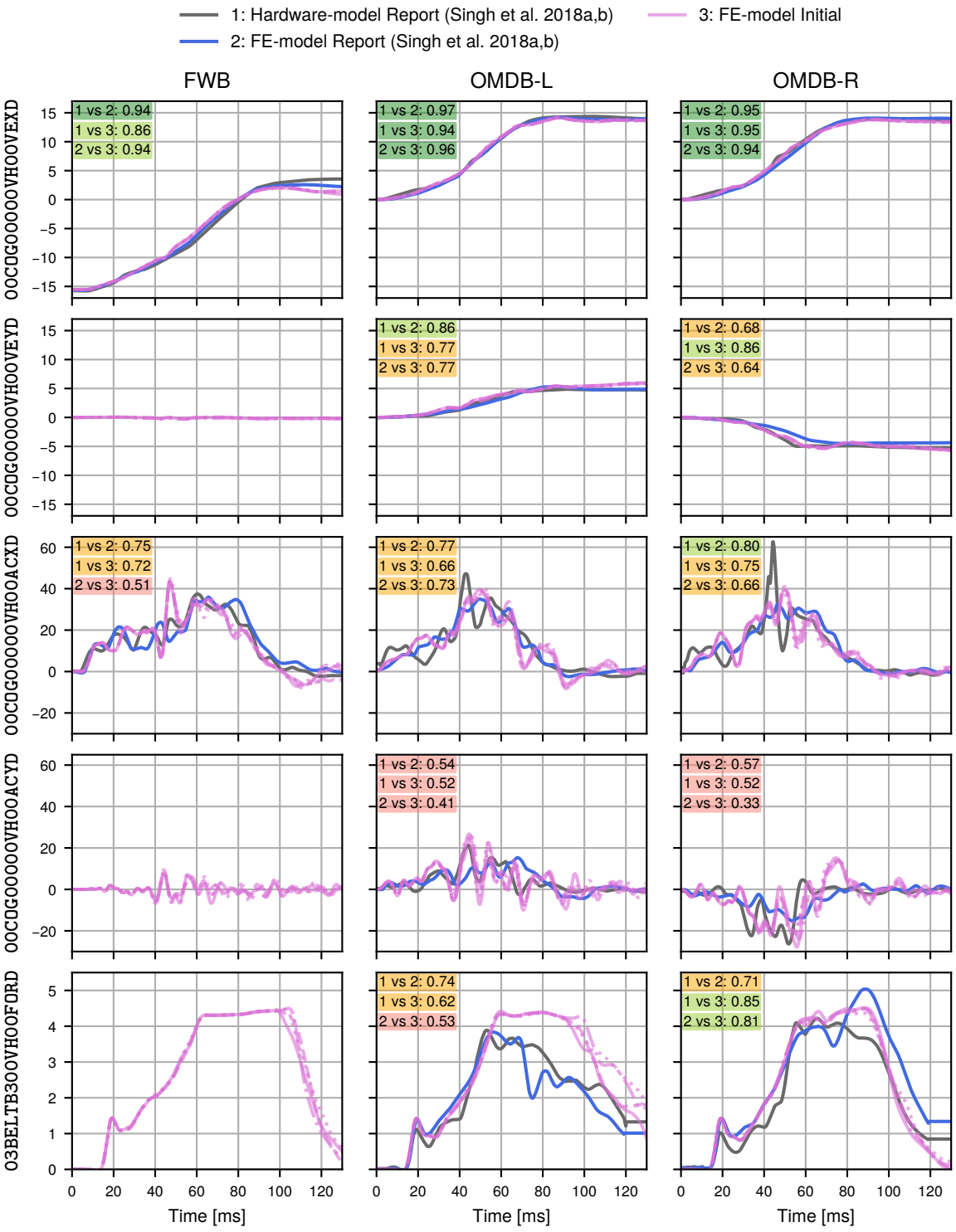


Figure 3.7: Vehicle-pulse at CoG and Seatbelt B3-force of Initial FE-model in Comparison with Reference reports by Singh et al., 2018a,b

As displayed in figure 2.1 on page 8, the D-ring redirects the seatbelt from the retractor to the shoulder as shown in section 2.1 in figure 2.1 on page 8. In the FE-model, a constant friction coefficient μ of 0.35 was defined. Measuring the angle β between the seatbelt sections before and after the D-ring, an approximate shoulder B3-force ${}^{\text{B3}}F_{R,t}$ can be calculated using from the Euler-Eytelwein derived equation

$$ {}^{\text{B3}}F_{R,t} = {}^{\text{B0}}F_{R,t} \cdot \exp\left(\mu \cdot \gamma_t\right). \tag{3.1} $$

The formula requires the wrapping angle γ which can be calculated by

$$ \gamma_t = (180° - \beta_t) \cdot \frac{\pi}{180°}, \tag{3.2} $$

from the inner angle sum of the quadrangle formed by the two tangents – the two seatbelt sections – and the respective radii. The angle $\boldsymbol{\beta}$ was measured in degrees, while $\boldsymbol{\gamma}$ is defined in radians. In the FE-simulations, angles β_t between 60° to 70° were measured in the initial FE-simulation during the time window of 60 ms to 80 ms. According to those values, a B3-force level ranging from 4.5 kN to 5.0 kN is expectable.

In the initial FE-simulation, a B3-force of 4.5 kN was observed in figure 3.7 on the facing page after the coupling phase. In the OMDB-R load case, an ISO-rating of 0.81 was measured. For the OMDB-L load case, the correlation was ISO-graded Poor. The seatbelt slips during the lateral motion of the shoulder of the THOR-50M, reducing the B3-force in the reference report data. In the case of the reference report FE-simulation, a sharp drop around 70 ms is the outcome. In the hardware-model, a drop but slower is observable too. In the initial FE-simulations, however, the seatbelt B3-force reaches its plateau and the seatbelt is continuously pulled out until the rebound begins. From visual analysis, a clamping of the seatbelt in the shoulder joint seem to take place.

For the accelerations of head and pelvis in figure 3.8 on the next page, the agreement of both FE-simulations was ISO-graded Fair to Poor. In the pelvis, for both OMDB load cases, the magnitudes between the FE-simulations are well-matched. The phase of the pelvis acceleration, on the other hand, matches better with the hardware-model.

The head accelerations are well aligned. In the case of the OMDB-R presented in figure 3.9 on page 85, the initial FE-simulation report approximately $10\,g$ higher maximum resultant acceleration. Comparing the directional components, the Z-direction accounts for the higher values. In all cases, the PAB could not act properly and following, the head nods down resulting in the Z-component as major head acceleration. Considering the timing, the peak coincides with an unusual behaviour in shoulder seatbelt force in the reference report FE-model: instead of remaining at its constant level, the B3-force increases again and reaches the maximum together with the head acceleration around 90 ms.

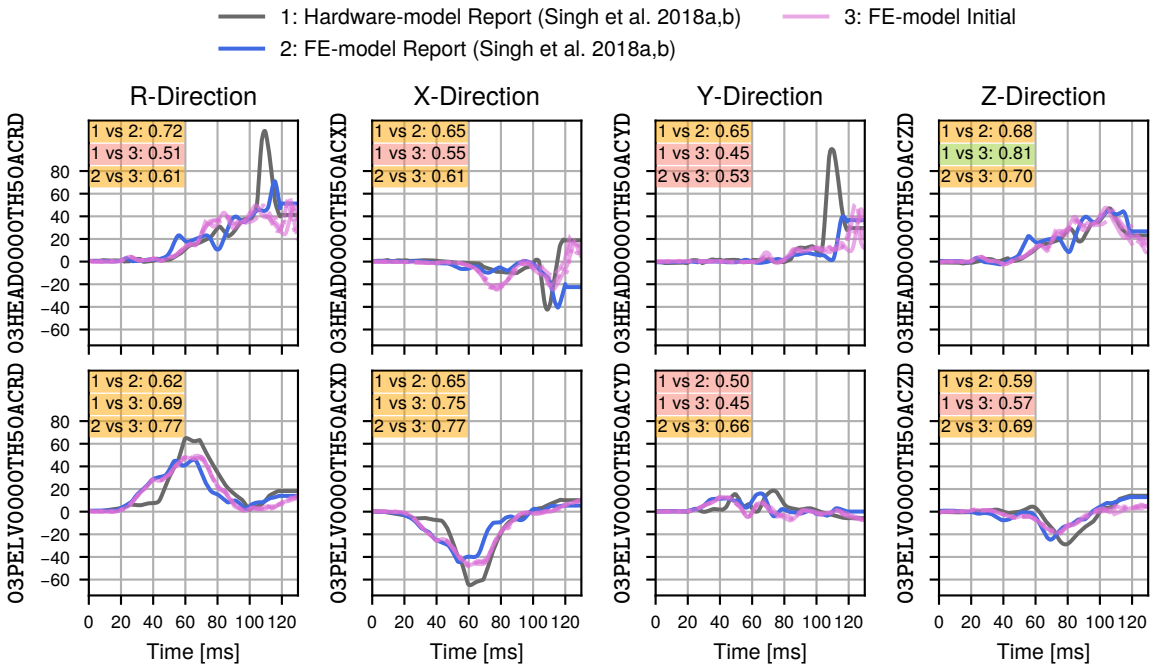


Figure 3.8: Head and Pelvis Acceleration in Oblique Movable Deformable Barrier Left (OMDB-L) Load case of Initial FE-model in Comparison with Reference reports by Singh et al., 2018a

One explanation for the difference in the acceleration magnitude could be the higher energy dissipation due to the higher seatbelt forces over the thorax. Higher B3-forces would result in a stronger nodding motion of the head if not properly restrained by the PAB. Unfortunately, the data available in the reference report does not allow a more detailed insight to support or reject this explanation.

The signals of the femur forces are shown in figure 3.10 on the next page. In all combinations and load cases, the ISO-grade of the correlations were Poor.

In summary, the important responses from the initial FE-simulations ISO-grades were Good to Fair regarding the reference report FE-model. The seatbelt slip of the shoulder in the OMDB-L load case and the instable PAB positioning in general reduced the comparability of the cases. Due to numerous unknowns with the published FE-model, it was possible to reason the observed differences, only for few sensors. The FE-model from the reference report already does not match perfectly the hardware-model.

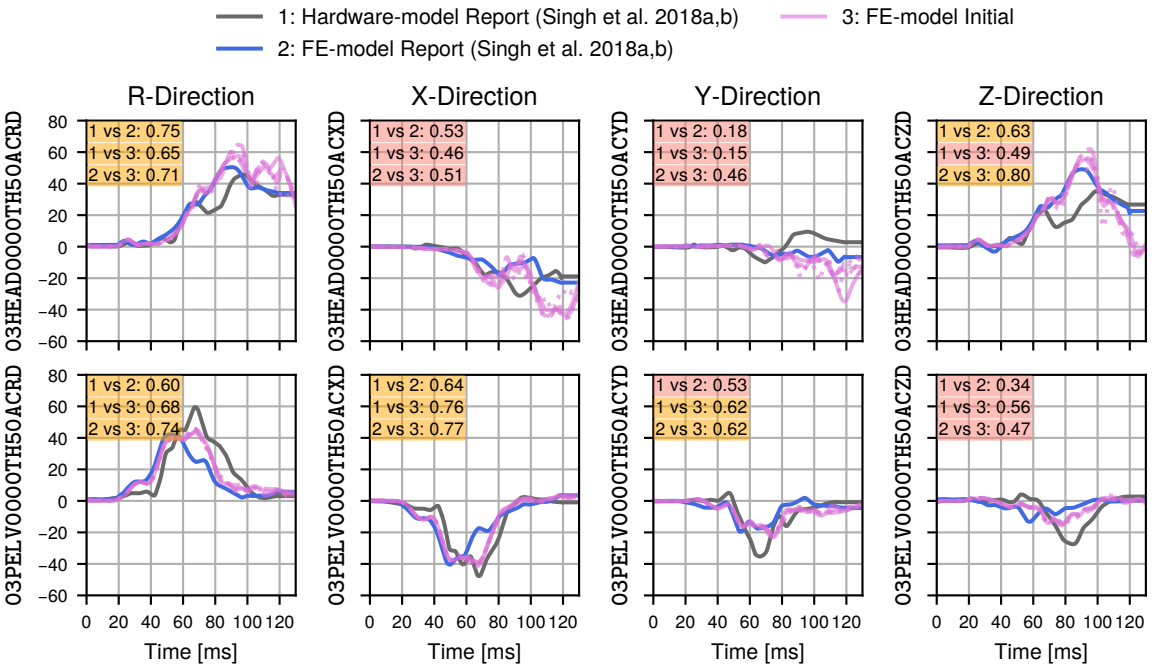


Figure 3.9: Head and Pelvis Acceleration in Oblique Movable Deformable Barrier Right (OMDB-R) Load case of Initial FE-model in Comparison with Reference reports by Singh et al., 2018a

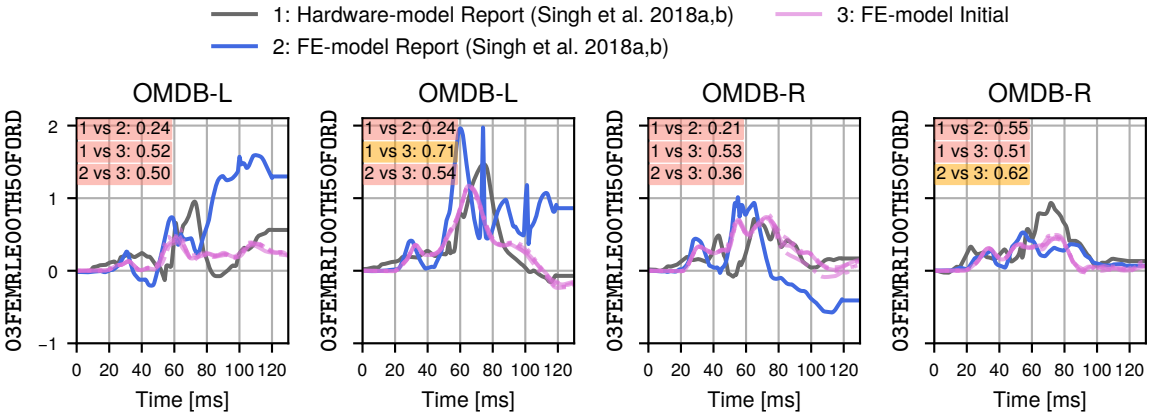


Figure 3.10: Femur Forces in Oblique Movable Deformable Barrier Left (OMDB-L) and Oblique Movable Deformable Barrier Right (OMDB-R) load cases of Initial FE-model in Comparison with Reference reports by Singh et al., 2018a

3.2 Model Simplification for Efficient Data Generation

3.2.1 Simplification Potential Analysis

As described above on page 69, the FE-model must have a reasonably short computation time on the available computer-hardware resources. Furthermore, numerical robustness is required. This second requirement was achieved by adjustments made in the FE-model and the selection of the Channel Frequency Class (CFC). However, the computation time of over 8 h observed in figure 3.6 on page 80 is too high. Further steps of FE-model alteration seemed necessary. To avoid completely losing the benefits of having a reference reports, all changes should be as minimal as possible. Using a stepwise approach and AB-comparisons for each change, the foundation of the Reference reports should be maintained.

Because the computer-hardware was mainly fixed, simplifications on the FE-model was the main driver. The established strategies were discussed above in section 2.1 on page 16 ff and on page 58 ff. The vehicle-pulse in a full vehicle crash is dependent on the vehicle's mass. In other words, simplifications like removing single components must not compromise the mass distribution within the vehicle. The submodel approach provides here a higher degree of freedom. Following, all entirely for the driver side relevant components could be removed since only the passenger is considered.

The original model was meshed with an average element size of 6 mm. Utilizing the Hybrid III 50th Percentile Male Dummy (HIII-50M) instead of Test device for Human Occupant Restraint, 50th Percentile Male (THOR-50M) can only bring its full potential, if the rest of the model is simplified and the critical integration interval increased (see section 2.1, page 17). The used Hybrid III (HIII) family from Guha et al. (2011) comes with an average element size of 25 mm. This size was used as a goal for the other components.

The relevant components were dashboard, door, seatbelt, airbags, and seat. The Body In White (BIW) was considered rigid and hence not part of the integration interval calculation (see section 2.1, page 17). Given a reduced lateral motion of the Anthropomorphic Testing Device (ATD) due to the switch from Oblique Movable Deformable Barrier (OMDB) to the Offset Deformable Barrier (ODB) load case, less to no interaction with the Curtain Airbag (CAB) and the door was observed. Hence, those components can be dropped instead of being simplified. Anyway, the two components had to be only considered for the impact on the right side.

As stated above in section 3.1 figure 3.2 on page 73, the Passenger Airbag (PAB) positioning was found to be instable. Hence, any comparison would be obfuscated by this effect. Therefore, as an early step, the PAB behaviour must be improved. During the remodelling of the PAB FE-model, the simplification occurred alongside.

Due to the smaller pelvis of the HIII-50M compared to the THOR-50M, a change in the interaction with the seat was expected. In the original sub-FE-model, the ATD was pressed into the seat's foam by a pre-FE-simulation. Unfortunately, the original cushion shape could not be restored. For comparison purpose, the upper body of both ATDs should be aligned as possible.

This, together with the described seat modelling, led to a distance between the pelvis and seat cushion. Simplifying the seat in an earlier step, while taking this geometric issue into account, can reduce the implied risk of submarining.

To reduce the storage and update effort, the BIW will be meshed coarser. Dropping the intrusions and modelling the BIW as completely rigid allowed simplifying the interior.

All displayed considerations led ultimately to the stepwise simplifications listed in table 3.2. For all steps, the changes are explained and AB-comparisons presented. The FE-models and the ISO-ratings were provided as supplement by Plaschkies (2024c). The ISO-ratings were either reported as numerical value or by the corresponding ISO-grade as defined in section 2.1 table 2.3 on page 30. The step 6 was only performed for the Offset Deformable Barrier Right (ODB-R) load case because the influence of the CAB and door is only given if the ATD moves towards them.

Table 3.2: Steps of Model Simplifications

Step	FE-model Version	FWB	ODB-L	ODB-R	Page
1	Initial FE-model	✓	✓	✓	79
2	Prescribed Motion and Intrusion	✓	✓	✓	87
3	Simplified Front Airbag	✓	✓	✓	91
4	Simplified Seat	✓	✓	✓	95
5	THOR-50M Replaced by HIII-50M	✓	✓	✓	98
6	Neglected CAB and Door			✓	103
7	Simplified Seatbelt	✓	✓	✓	104
8	Simplified BIW & Neglected Bulkhead Intrusion	✓	✓	✓	106
9	Simplified Interior: Rigidized Dashboard & Carpet	✓	✓	✓	109
10	Adjusted for Data Generation	✓	✓	✓	114

3.2.2 Simplification by Prescribed Motion and Intrusions

3.2.2.1 Method of Prescribing Motion and Intrusions

As elaborated in the theory chapter (see section 2.3, page 58), the technique using sub-FE-models is commonly used for FE-simulations with ATDs. Only for ATD interaction, relevant parts of the complete vehicle FE-model were used. The actual crash is replaced by a prescribed vehicle-pulse and applied intrusions as explained theoretically in section 2.1 on page 20. In total, 4 sensors on the sills of the passenger and driver side were used for the prescribed accelerometer: 2 sensors in the middle of the rear door and 2 sensors at the B-pillar. To avoid unnecessary instabilities, all signals were filtered by CFC 60. Alternatively and in case of a perfect sensor location in terms of no local deformations, a simple prescribed motion for the BIW rigid body on e. g. its Centre of Gravity (CoG) could also be used. Stable FE-simulation were achieved with interface files recorded with 5 kHz.

To identify the relevant parts to keep from the vehicle, the author took the ATD perspective.

> *What are the likely areas to interact with given the selected load cases?*

In the Honda Accord FE-model, they were the dashboard, the carpet, the restraint system, and the seat.

> *Which parts of the vehicle affect the ATD or the relevant components directly but are not affected vice versa?*

This second question led to the BIW. Here, the intrusions from the crash will override any impact from the much softer components like an ATD head. At this level, the recorded intrusions used for constraining nodes in the simplified model are sufficient. All remaining parts of the BIW can be neglected.

The interface, if recorded in the global coordinates, contains the rigid body motion and the deformations. Applying the interface approach to the whole BIW would make the prescribed motion unnecessary. However, a huge amount of data has to be stored and evaluated every cycle. Selecting only those parts which undergo significant deformation during the crash and whose deformations are relevant for the ATD, can reduce the overhead. The interface file with the reduced content, shown in figure 3.11, was the size of about 5.5 GB.

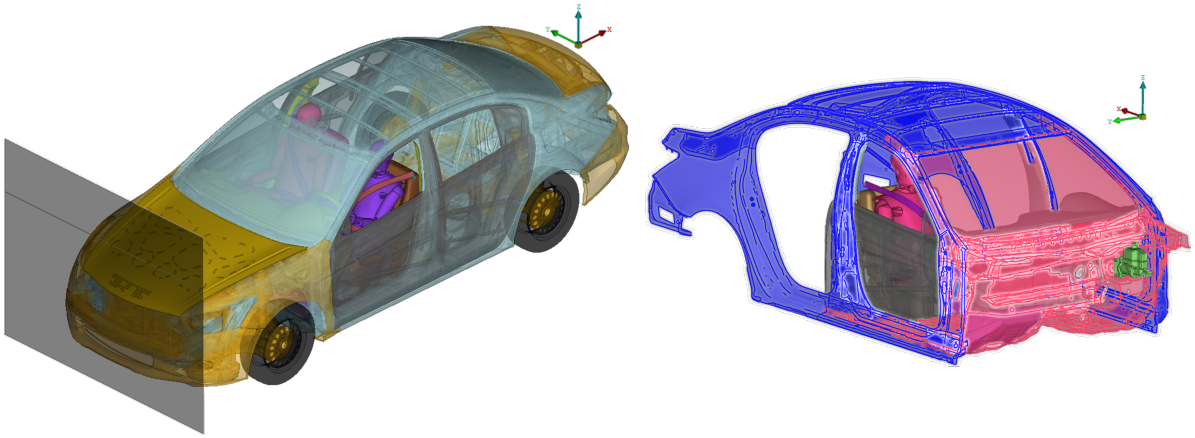


Figure 3.11: Full FE-model (Left) vs Reduced FE-model (Right) with Intrusion Zone in Red and Rigid Body in Blue)

All parts, which are constrained only by the rigid body motion, must be modelled as rigid using the rigid material type. In section 2.1 on page 18 ff, it was explained that Livermore Software technology corporation DYNA (LS-DYNA) calculates the contact stiffness from the involved partner's elastic modulus. Consequently, the as rigid defined parts should maintain their definition of Young's modulus.

Parts with the interface constrained nodes should not carry any inner stiffness by using the null material. Otherwise, the internal energies would be unreasonably high. This can lead to numerical issues even if the FE-simulation terminates without an error message. Nevertheless, similar to the rigid definition, the elastic modulus still has to be defined for the contacts.

Another potential pitfall comes from the fact that the interface constraint is not a real constraint but an applied force vector. The implication of this definition is that any real constraint overrides the interface locally. An example in the Honda Accord FE-model is the connection of the seat to the BIW. In the reference report FE-model, a nodal rigid body containing nodes from both components was used. Since the null material does not carry any structural stiffness and the rigid body definition overrides the interface, the seat would be practical unconstrained. The solution was to change the definition from hard constraints to soft constraints via the tied contact.

3.2.2.2 Results of Prescribing Motion and Intrusions

Due to the averaging nature of applying the vehicle-pulses by the described method, deviations were expected. Additionally, since the vehicle's acceleration sensor at the CoG is attached to the rigid part of the guided BIW and not to deformable parts as in the full crash model, less oscillation should be the result.

The X-velocity for all three load cases were reported with ISO-ratings greater than 0.95 compared to the initial FE-model. The Y-velocity in the Offset Deformable Barrier Left (ODB-L) load case had the ISO-grade Good, too. The Y-velocity of the ODB-R load case, however, got only an ISO-rating of 0.71. The X-acceleration of the Full Width Barrier (FWB) was ISO-graded as Good, too. Both ODBs were ISO-rated around 0.8 for the X-acceleration, but ISO-graded Poor for the respective Y-acceleration. The major reason for those low ISO-ratings can be inspected in figure 3.12 on the next page. The likely reason is the much higher amplitude in the acceleration's oscillations in the initial FE-model. Due to the averaging, those oscillations were reduced since sensor locations on the left and right side of the vehicle were used to generate the guiding vehicle-pulse. Second, in the initial model, the area around the evaluated CoG is deformable, and thus, can undergo its individual oscillations. In contrast, in the case of the guided BIW, the sensor is rigidly connected to the structure. The numerical configurations as described on page 3.1.4 shows that the oscillations were not stable in the initial FE-model.

The comparison of the restraint system was extended by the pressure of the PAB and all typical seatbelt forces and the pull-out. For all three load cases and evaluated channels, the correlations were greater than 0.8 and therefore ISO-graded Good.

The accelerations were evaluated at head, chest, and pelvis. All resultant accelerations got a Good ISO-grade compared to the initial FE-model. All accelerations in the Y-direction were ISO-graded only Poor or Fair. And yet, no major difference could be identified. One reason could be, similar to the vehicle's lateral motion, the ATD experience comparably low kinematic in the Y-direction. As described above, the ISO-rating tends to judge stricter than a human would do based on the plots, if the magnitudes are very low in relation to the X-direction. The femur forces of the left and right legs were ISO-rated consequently higher than 0.8.

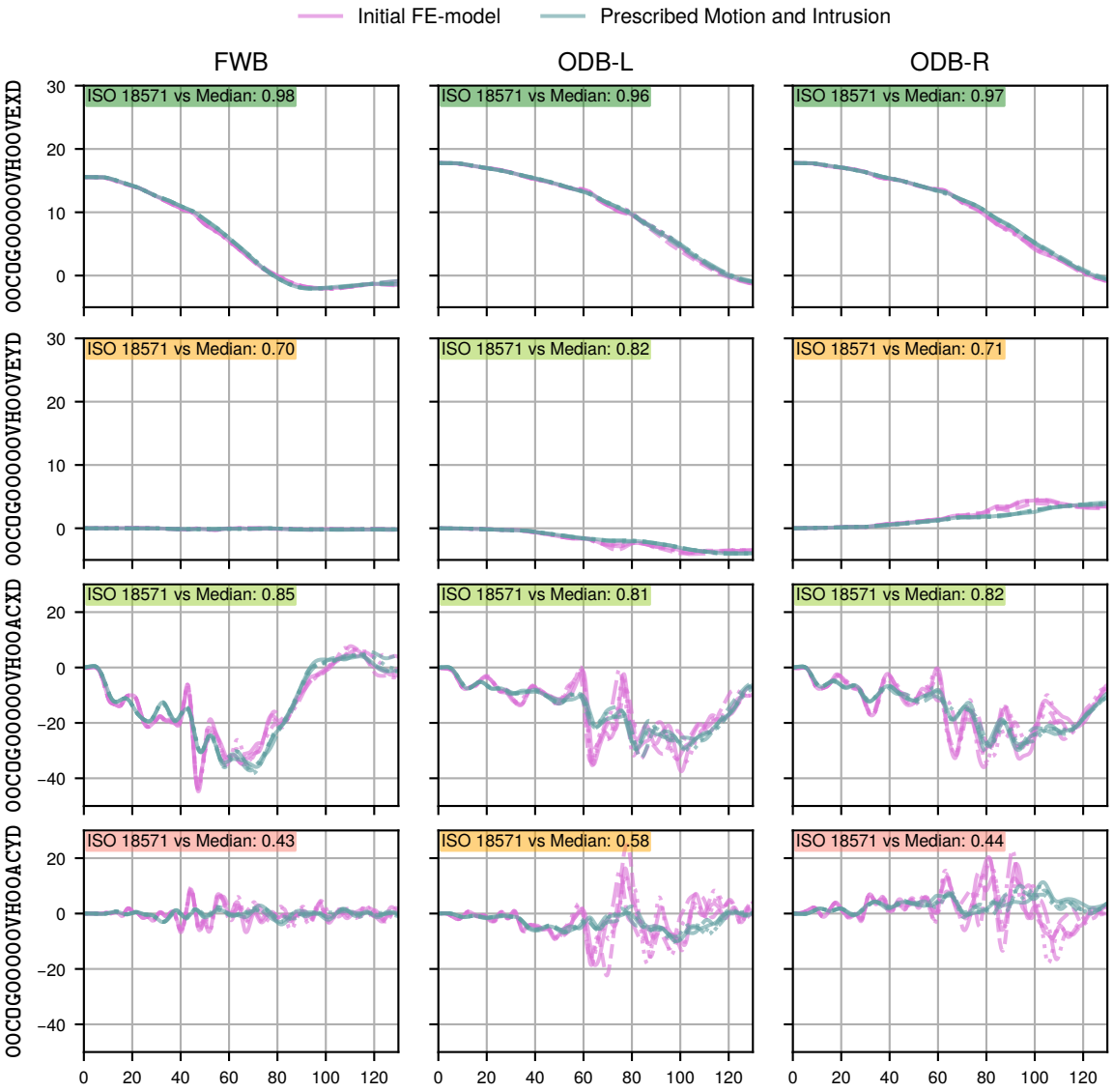


Figure 3.12: Vehicle-pulse for Initial FE-model and Prescribed Motion at BIW CoG

3.2.3 Front Airbag Simplification

3.2.3.1 Method of Simplifying the Passenger Airbag

In the case of the airbags, only the PAB was simplified. The CAB was planned to not be considered in later steps. The Driver Airbag (DAB) was out of scope, as only the passenger position was considered (see page 79). The PAB of the Honda Accord was modelled using elements with an average size of 5 mm. The airbag was compressed into its housing. When firing the inflator, the PAB would inflate. In the folded state, it is impossible to remesh the airbag. The complete remodelling procedure was as follows: (i) inflate PAB, (ii) remesh the resulting shape, (iii) shrink or fold the new PAB, and finally (iv) adjust the parameters if necessary.

Due to the interaction of the ATD with the PAB and numerical variance in a previous FE-model version, the alteration of the PAB was performed in an isolated environment, the so-called standing test. This environment consisted only of a rigidized version of the dashboard and the windscreen. The undeformable surfaces of both seemed sufficient, since the airbag alone cannot and must not apply deformations to the dashboard by an extent of airbag being influenced. One drawback of this method was that a PAB is usually developed to be clamped between the dashboard and the ATD thorax. If the ATD is missing, the airbag will not find a stable position but bounce.

A point where the positioning of the shape and its smoothness are acceptable was found at 60 ms. Since an airbag gets first inflated from the folded state and later deflated due to fabric leakage and the venting holes, the surface is only smooth, if the inner pressure is around its highest value during the fully inflated phase. In agreement with the reference reports by Singh et al. (2018a), the inner volume of the inflated bag was approximately 86 L.

By extracting the node coordinates at the time stamp 60 ms of the simulation, those coordinates were now used to replace the nodes in the folded state. After cleaning and smoothing of the elements, the automatic remeshing algorithm of the Finite Element Analysis (FEA) pre-processing software ANSA was utilized to generate elements of the targeted size. By applying smoothing algorithms and manual adjustments, a satisfactory mesh quality was achieved.

To avoid the initial intersections of the PAB with the ATD described in section 2.1 on page 19, the PAB had to be folded or shrank. The author decided for the second approach because the coarse mesh could undergo too strong distortions when folded. Geometrically, shrinking of the bag can be achieved using morphing tools. To ensure the inflation back to the original shape, two adjustments were necessary: (i) define a reference geometry, and (ii) adjust the inflator parameters.

The concept of a reference geometry as described in section 2.1 on page 20 adds the possibility to add the information on a deformed state at the beginning of a FE-simulation. In the author's experience, this approach is commonly used for folded airbags because the folding process deforms the fabric. However, no such definition was available in the Honda Accord FE-model. Hence, the coordinates of the PAB nodes before shrinking had to be used. Another

advantage of the reference geometry approach is that the element size for the integration interval determination is calculated based on the reference geometry and not the shrank element. The final result is displayed in figure 3.13.

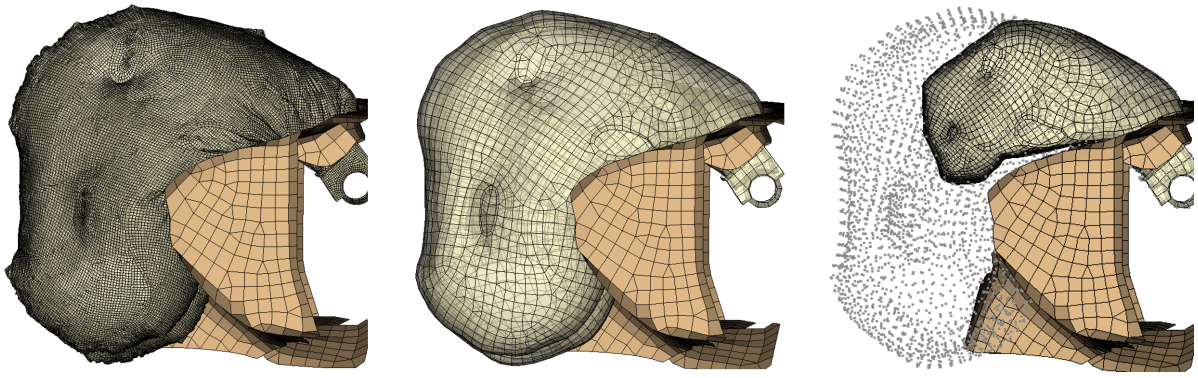


Figure 3.13: Airbag Standing Test with (from Left to Right) Original PAB @ 60 ms, Coarsely Meshed PAB @ 60 ms, and Shrank PAB with Reference Geometry on the Simplified Dashboard

In the case of a proper folded airbag, the initial volume would be near zero. Since here the PAB was only shrunk, the initial volume was greater than zero. By reducing the mass flow of the inflator, similar pressure was achieved during the time window of ATD restraint. The inflator was defined as hybrid, containing fractions of CO_2, H_2O, and N_2 which are released as a shared mass flows. For the shrunk PAB, the same fraction was assumed to be in the initial volume at normal pressure and the inflator. To compensate for this, the mass flow from the inflator definition was reduced by approximately 30 %.

Additionally, the use of LS-DYNA fabric material required the following adjustments. The liner and the fabric's compressive stress elimination had to be activated. Both reduced the wrinkling of the fabric. The elimination of compression prevented the implosion of the airbag at the beginning of the simulation.

In figure 3.14 on the next page, the comparison in the standing test is displayed. The time around 60 ms is relevant as the PAB geometry was taken from here. One simulation on a local workstation with 4 parallel processes took ca. 7 min. With the same computer-hardware input, the original airbag required 4 h.

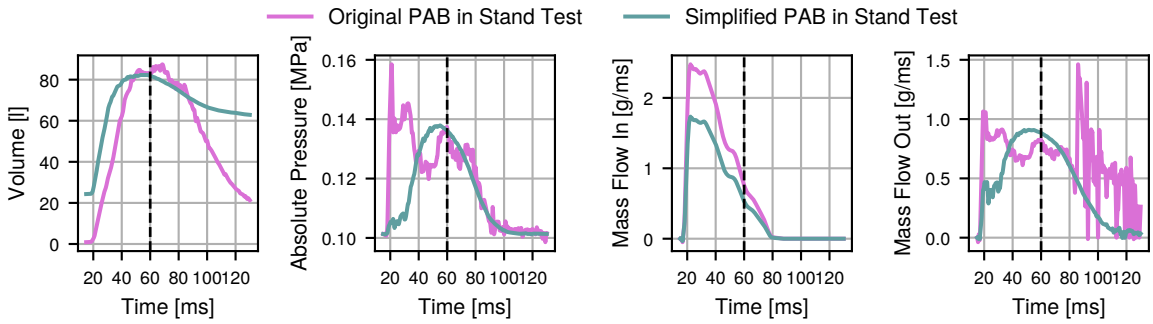


Figure 3.14: Unfiltered Signals from Standing Test with Original PAB versus Simplified PAB

3.2.3.2 Results of Simplifying the Passenger Airbag

Since the method for applying the vehicle-pulse was unchanged, the measured vehicle-pulses matches nearly perfectly with the previous FE-model version. Apart from negligibly small numerical deviations, in all subsequent models the vehicle-pulse remained unchanged.

All seatbelt signals also got the ISO-grade Good. Naturally, the inner pressure of the PAB got only the ISO-grade Fair. As shown in figure 3.15, the initial pressure and the changed mass flow altered the signal in the evaluated range. In the working time window, however, the pressure was found to be higher. The reason becomes obvious in figure 3.16 on the next page. The PAB original positioning was very skew; the simplified PAB middle remains approximately in the sagittal plane, and hence the PAB is properly loaded.

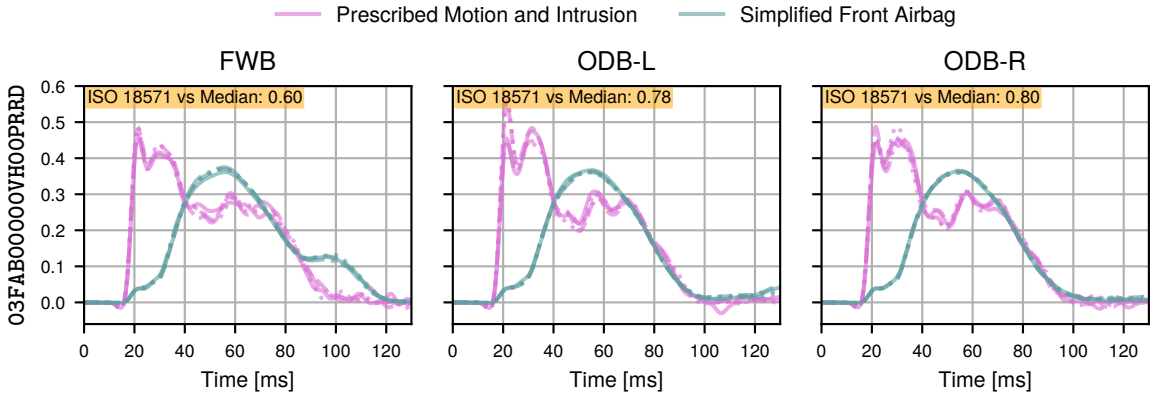


Figure 3.15: PAB Relative Pressures in [bar] for Simplified Airbag versus previous FE-model Version with Original Airbag

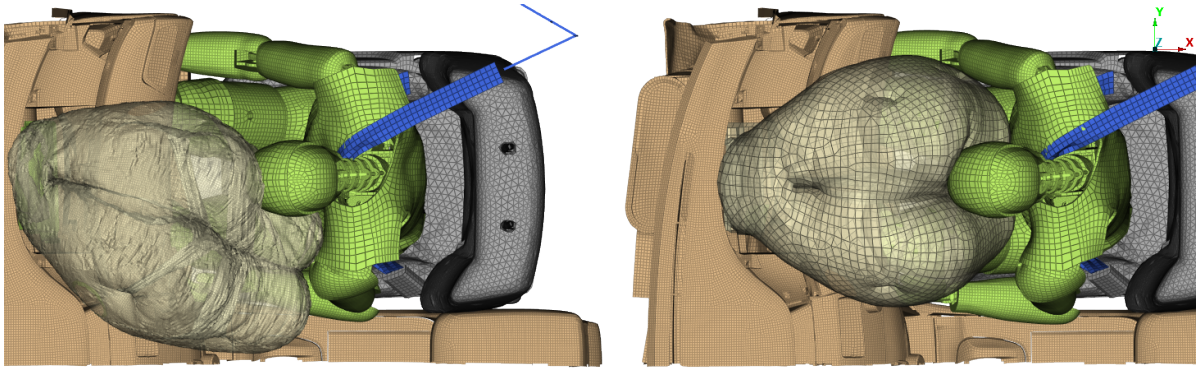


Figure 3.16: Original PAB (left) versus Simplified PAB (right) in Full Width Barrier (FWB) Load case @ 80 ms

The kinematics of ATD is influenced by different interaction of PAB with ATD. However, the ISO-ratings of head, chest, and pelvis resultant accelerations were overall found to be ISO-graded Good. For chest and pelvis, in the X- and Z-directions of the respective sensors, light changes were observed, but the correlations were ISO-graded Good. All accelerations in the Y-direction were ISO-graded Fair, without any particular effect identified. Natural assumptions for reasons are the low magnitude of the acceleration signal in the FWB load case and the significant different interaction of the ATD with the PAB in the lateral ODB load cases.

The head acceleration was ISO-graded Fair, as shown exemplary for the FWB load case in figure 3.17 on the facing page. In the initial FE-model, the X-component drops from 80 ms due to missing interaction with the PAB. Since the simplified airbag remains in the ATD sagittal plane, the X-acceleration remains on a high level. However, a strong deviation in this time window was observed between the on page 81 defined four numerical setups.

The femur forces were only marginally affected by the change in the PAB. The ISO-grade was found to be Good.

In summary, the simplification introduced a major change in the overall model. From a safety perspective, the positioning of the PAB was altered fundamentally towards the better. An PAB which does not remain in a stable position and in the trajectory of the ATD head would be useless. Additionally, the comparison of simulations would not be reliable. Despite changing the initial FE-model characteristics, the restraint system can now only be declared functional.

The computation time was reduced in median from 418 min to 387 min. This delta vanished in comparison to the enormously faster simulation in the stand test. The reason is that the integration interval of the simulation was not affected by the change in the PAB, since other components like the door or seat contained smaller elements.

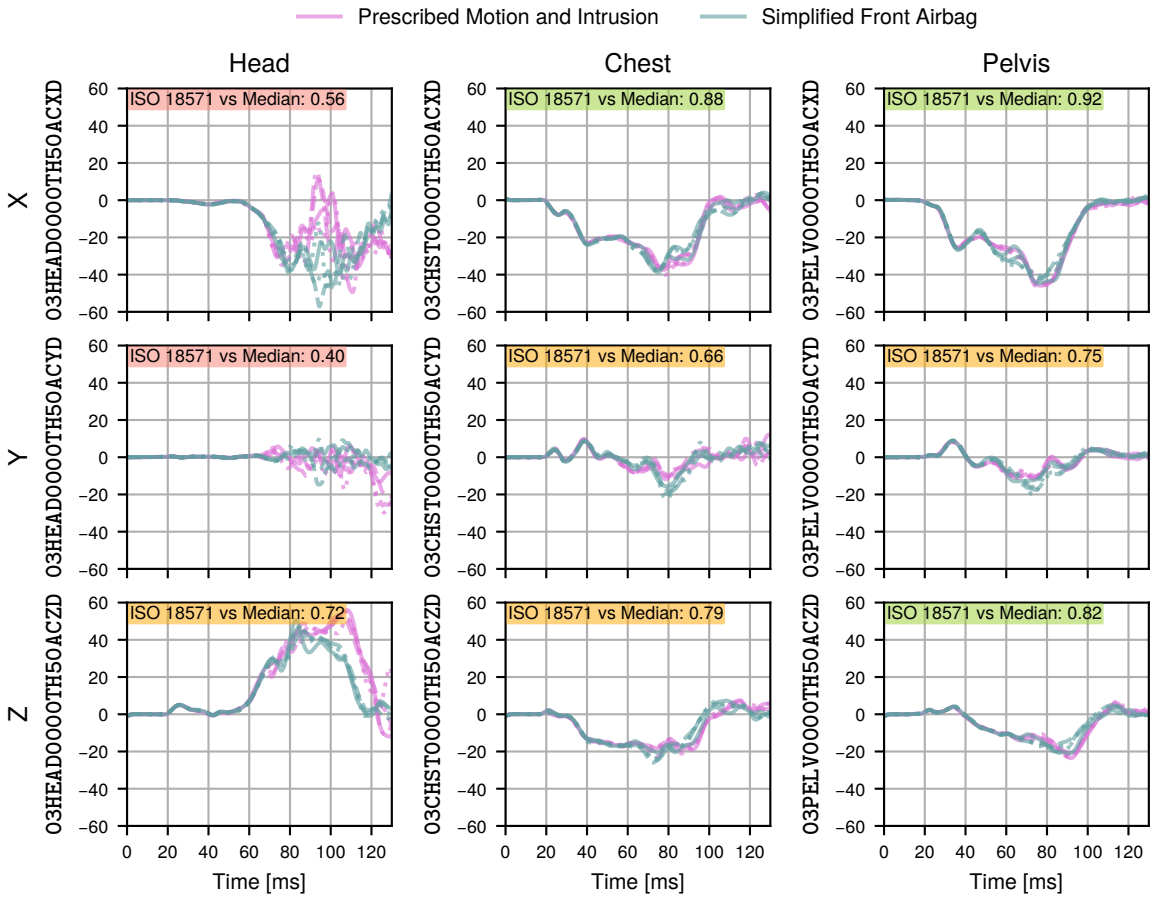


Figure 3.17: ATD Accelerations in Full Width Barrier (FWB) Load case for Simplified PAB versus previous FE-model Version with Original PAB

3.2.4 Seat Simplification

3.2.4.1 Method of Simplifying the Seat

A simplified seat can accelerate the simulation and extend the parametrizability of a FE-model. The complex shaped and pre-deformed cushions of the Honda Accord FE-model were replaced by a simple geometry. The deformable material definition was kept from the original cushions. But, the deformable seat structure was replaced by rigid plates. The surface of the cushions were oriented on the surface of the original seat. The bottom of the new cushion was defined with a distance under the original seat structure to take the seat's deformation into account.

In figure 3.18 on the next page, both seats are shown. Two characteristics of the original seat were not covered. First, the seat structure under the pillow is not homogeneous. Modelling this area homogeneous can change the way the ATD sinks into the cushion. This can increase the forward displacement but can reduce the risk of submarining. Second, the sides were modelled flat in the simplified model. In the case of a high lateral vehicle-pulse, the ATD pelvis is less

restricted. However, the OMDB load cases with their high lateral vehicle-pulses were replaced and hence it was anticipated that the missing lateral guidance would be insignificant.

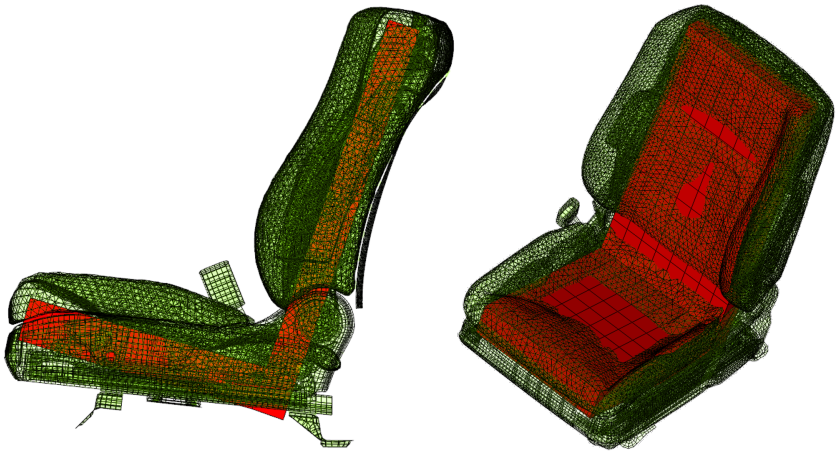


Figure 3.18: Original Seat (green) vs Simplified Seat (red)

3.2.4.2 Result of Simplifying the Seat

The vehicle motion remained in ISO-graded Excellent regarding the agreement with the previous FE-model version. For the FWB load case, all evaluated restraint system signals were of Good or Excellent ISO-grades.

In general, the movement of the seatbelt through the seatbelt-buckle was observed to be higher with the simplified seat than with the original seat. The higher Z-motion towards the seat cushion was identified in the case of the simplified seat as the main reason. Due to this effect, the lap seatbelt is slightly unloaded; seatbelt webbing is more pulled towards the thorax. This effect was especially strong in the ODB-L load case. In figure 3.19 on the facing page the thorax seatbelt B4-force, the lap seatbelt B5-force, and the pull-through at the seatbelt-buckle are shown. The B4-force were ISO-graded Good regarding the previous FE-model version. The B5-force, however, suddenly increases at 85 ms from an excellent match with the previous FE-model version to a force similar as measured at B4. The explanation can be found in the location of the B5 sensor. The force was measured in LS-DYNA as a so-called section force, initially located 93 mm away from the seatbelt-buckle across the seatbelt. This distance is pulled through the seatbelt-buckle around the same time window. Hence, the force-sensor location swaps from the lap-section to the thorax. In other words, the mismatch between the versions can be considered an artefact. However, the higher pull-through is a real effect. The seatbelt force-sensor locations were depicted in section 2.1 figure 2.1 on page 8.

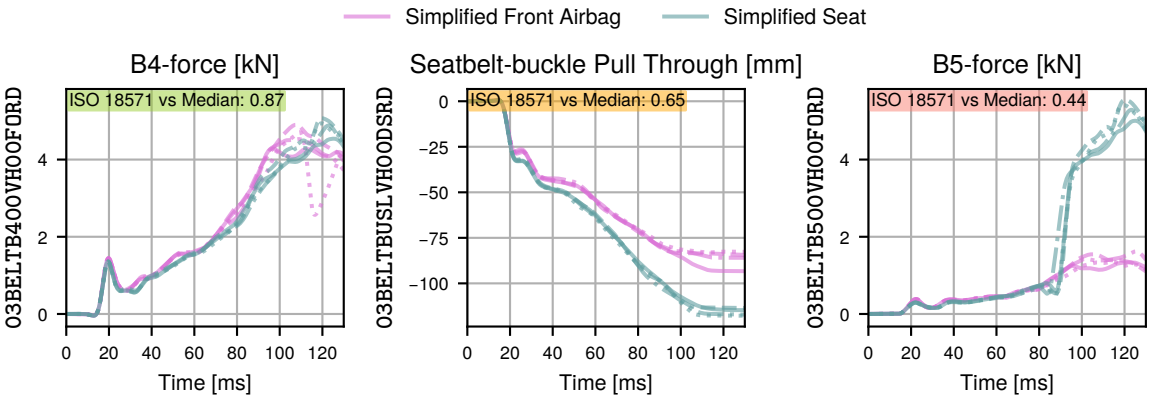


Figure 3.19: Comparison of Seatbelt Signals in Offset Deformable Barrier Left (ODB-L) Load case with Simplified Seat versus previous FE-model Version with Original Seat

The general increase in stiffness of the seat led especially to significantly increased accelerations in the Z-direction, as shown in the figure 3.20 on the following page. The overall head kinematics did not change much. However, the pelvis on the simplified seat first slides further forward, stopping then harder. As ATD is a kinematic chain, despite being partially elastic, the higher accelerations also affect the chest. The head was less influenced because the connection over the neck leads to less coupled behaviour.

The higher forward motion of the pelvis led to increased forces along the femurs from the contact with the dashboard. The height of the force peaks was determined where the knee hits the dashboard. Higher local stiffness can lead to higher forces. Consequently, between the three load cases, the impact intensity varied from unchanged to plus 1 kN. The ODB-L load case was least impacted, and the ODB-R load case the most.

The seat simplification led to a reduction in the computation time in a median of approximately 8 %. Considering the significant increase in the pelvis and chest accelerations, this simplification appeared to be unsuccessful. On the contrary, the median computation time of almost 6 h remains far from the desired duration defined above. Additionally, the new seat seemed to be a necessity to utilize the HIII-50M in the subsequent steps. Furthermore, in the next chapter it will be observed that for HIII-50M, the pelvis kinematics were different, leading accelerations back to the level of the original seat.

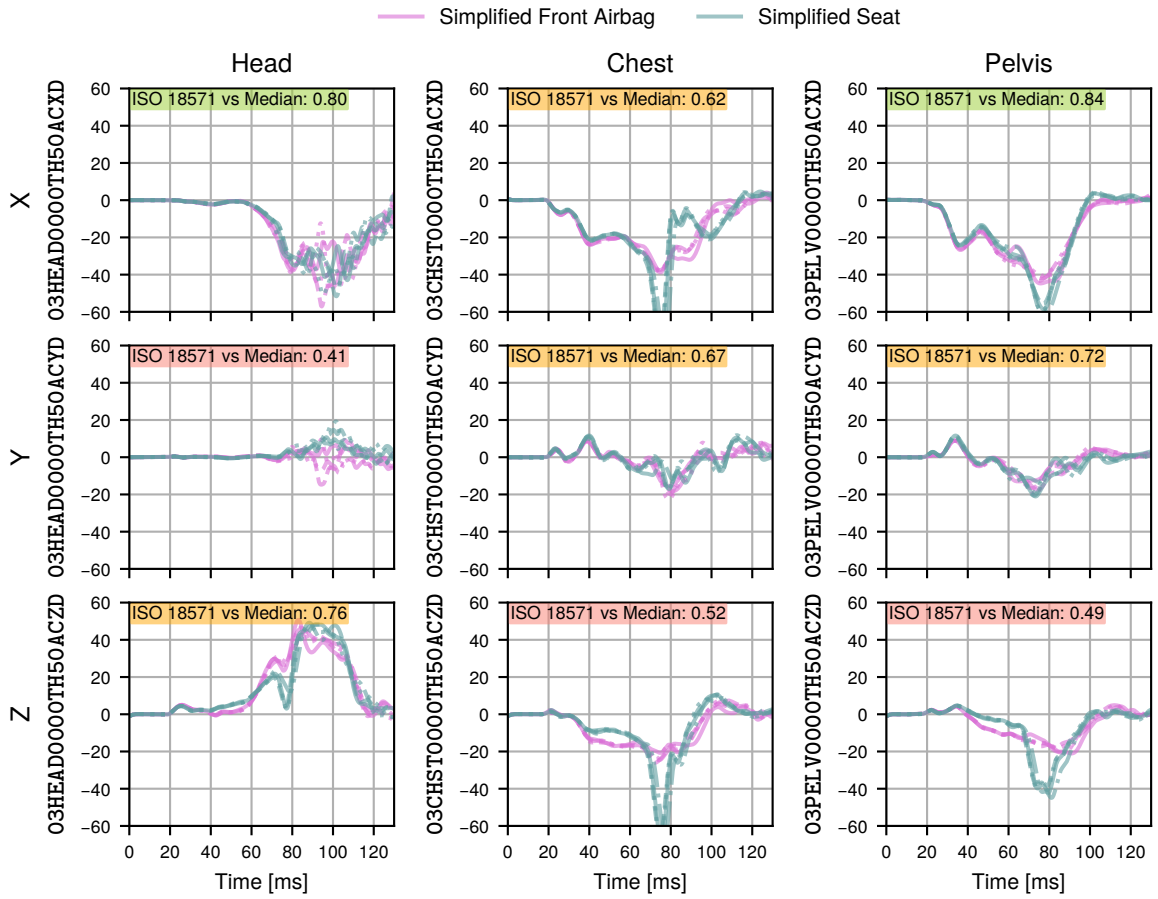


Figure 3.20: ATD Accelerations in Full Width Barrier (FWB) Load case with Simplified Seat versus previous FE-model Version with Original Seat

3.2.5 Exchanging THOR-50M with Hybrid-III-50M FE-model

3.2.5.1 Method of the ATD Exchange

In general, the THOR-50M and HIII-50M are based on similar geometry. Hence, to align the pelvis centres (H-points), it should be possible to approximately align the centres of the heads as well. In figure 3.21 on the next page, both ATDs are shown in the position used. To achieve this alignment, (i) both H-points were aligned, (ii) the back was aligned, (iii) the arms and legs were aligned, and (iv) the joint range limits of the HIII-50M were respected, and thus limited the degree of alignment in the spine.

The deeper thorax of the HIII-50M made the chest area of HIII-50M being significantly closer to the dashboard than THOR-50M. Due to the limited flexibility of the HIII-50M lumbar spine and neck, it was not possible to align the heads perfectly.

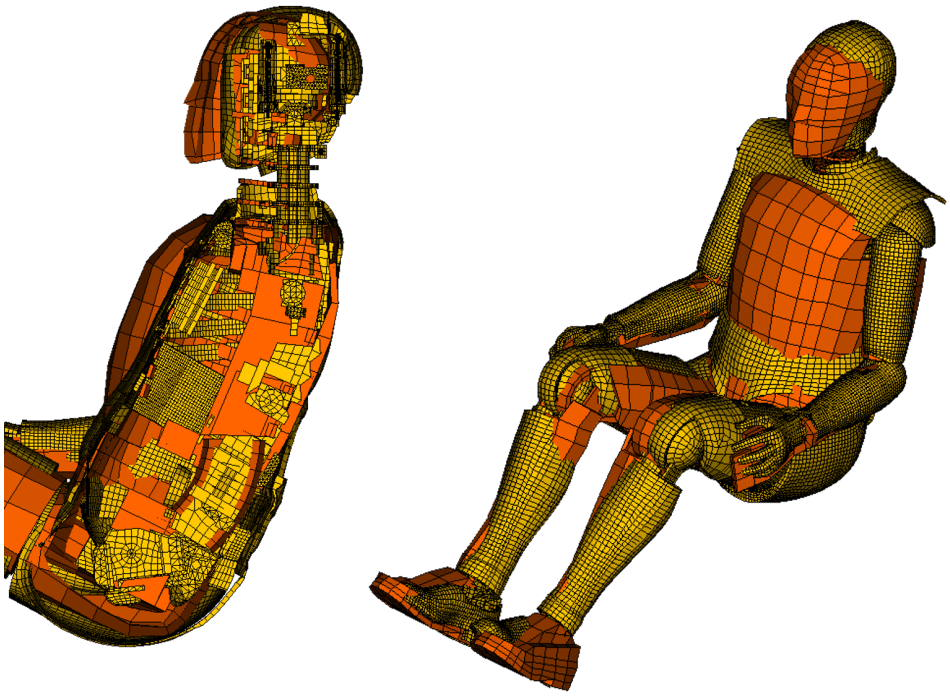


Figure 3.21: Comparison of THOR-50M (Yellow) and HIII-50M (Orange) as Cut in Sagittal Plane (Left) and as Isometric Perspective (Right)

As an alternative to the H-point alignment, the focus could have been the alignment on the seat cushion. However, since seat simplification allowed for cushion fitting to individual pelvis shapes, the upper body alignment was weighted higher.

The change in the measurement device can lead to fundamental changes in the recorded ATD kinematics and the evaluated sensors. In section 2.1 on page 23 ff, the known differences of both ATDs were laid out.

3.2.5.2 Result of Exchanging THOR-50M with Hybrid-III-50M

Interestingly, most of the signals received only Fair or Poor ISO-grades with respect to the THOR-50M and HIII-50M FE-simulations. As shown in figure 3.22 on the following page, the shoulder seatbelt force and the PAB characteristics were barely affected by the ATD exchange, although the HIII-50M couples slightly faster to the shoulder seatbelt. However, in the exemplary shown FWB load case was the lap seatbelt force found to be approximately 2 kN higher with the HIII-50M. The characteristics of the shoulder and lap seatbelt force are in accordance with Albert et al. (2023).

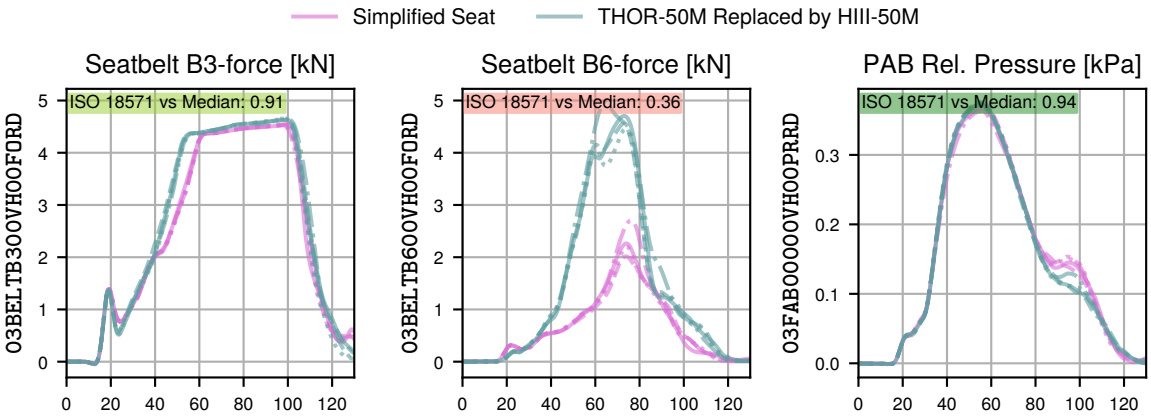


Figure 3.22: Restraint System of HIII-50M in Full Width Barrier (FWB) Load case versus previous FE-model with THOR-50M

In figure 3.23 on the next page, on the example of FWB load case, the ATDs' kinematics are displayed. Surprisingly, the head resultant acceleration correlations were ISO-graded as Good between both ATDs. In contrast, for both ODBs load cases, the peak head acceleration was observed to be ca. 10 g higher for THOR-50M than for the HIII-50M, which is in agreement with the 25 % reported by O'Connor et al. (2022).

Comparing the chest acceleration in figure 3.23 on the facing page, the peak acceleration for THOR-50M is ca. 40 % higher than for HIII-50M. Unlike O'Connor et al. (2022), who observed no significant differences, Xu et al. (2000) observed a similar deviation in pendulum tests. Fitting to the lower chest acceleration, the chest X-displacement relative to the vehicle is higher. That means, in combination with the on the same level remaining pelvis X-displacement, a stronger rotation of the thorax of the HIII-50M occurs. Both effects were observed less pronounced for both ODBs load cases also.

The femur forces in figure 3.24 on page 102 were for the left leg ca. 2 kN higher for HIII-50M than for THOR-50M in the FWB and ODB-L load cases. This is the opposite of the findings of O'Connor et al. (2022). As a reason, the different position of the left knee when in contact with the dashboard was identified. In figure 3.25 on page 102, it becomes obvious that the THOR-50M hits in an area of relatively low geometrical stiffness, while the HIII-50M hits a vertical member. The right femur forces are similar between both ATDs. For the ODB-R load case, the femur force of HIII-50M is lower than in the other load cases, since due to vehicle-pulse, the knee will move towards the less stiff areas of the dashboard.

In summary, most observed differences between both ATDs were in agreement with the findings from the literature presented in section 2.1 on page 23 ff. The significant deviation in the femur area could be reasoned by the different contact location of knee and dashboard.

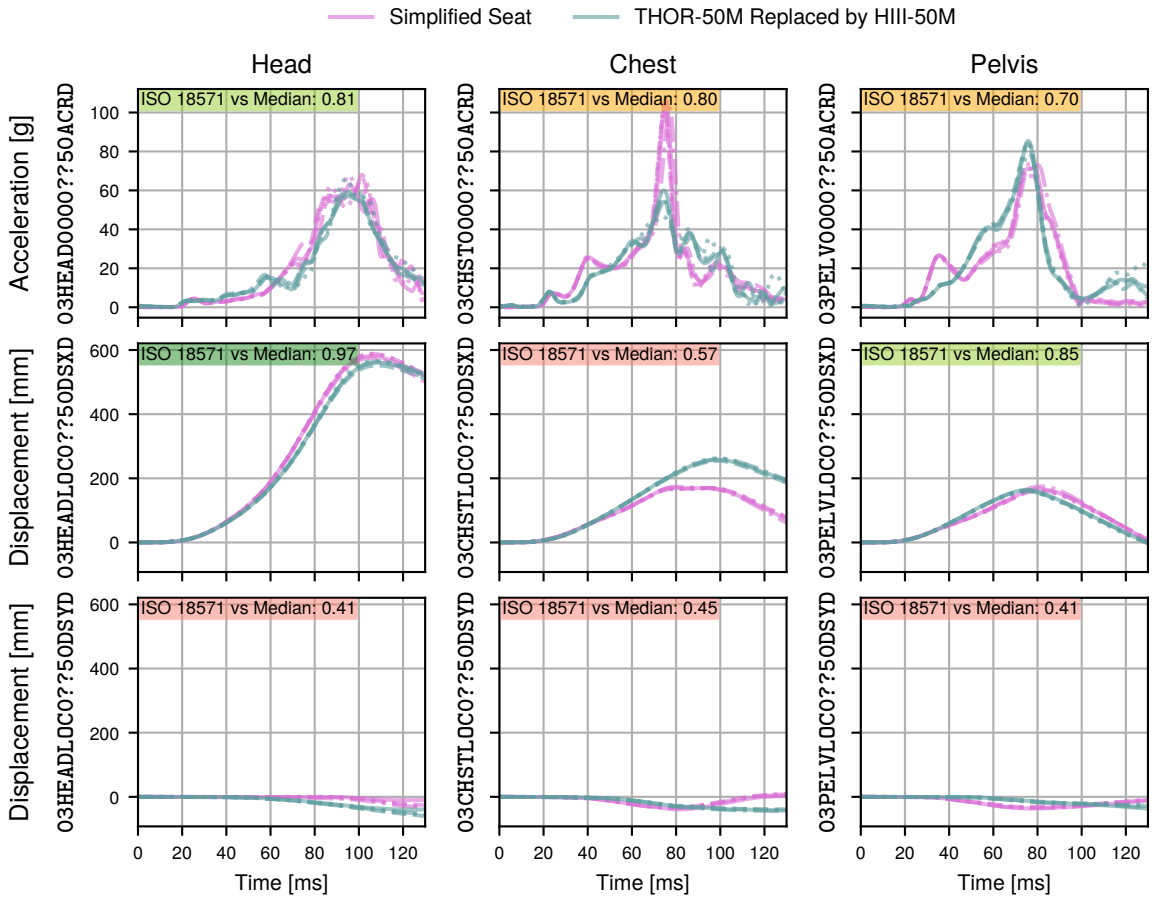


Figure 3.23: Resultant Acceleration and Local Displacements in Full Width Barrier (FWB) Load case with HIII-50M versus previous FE-model Version with THOR-50M

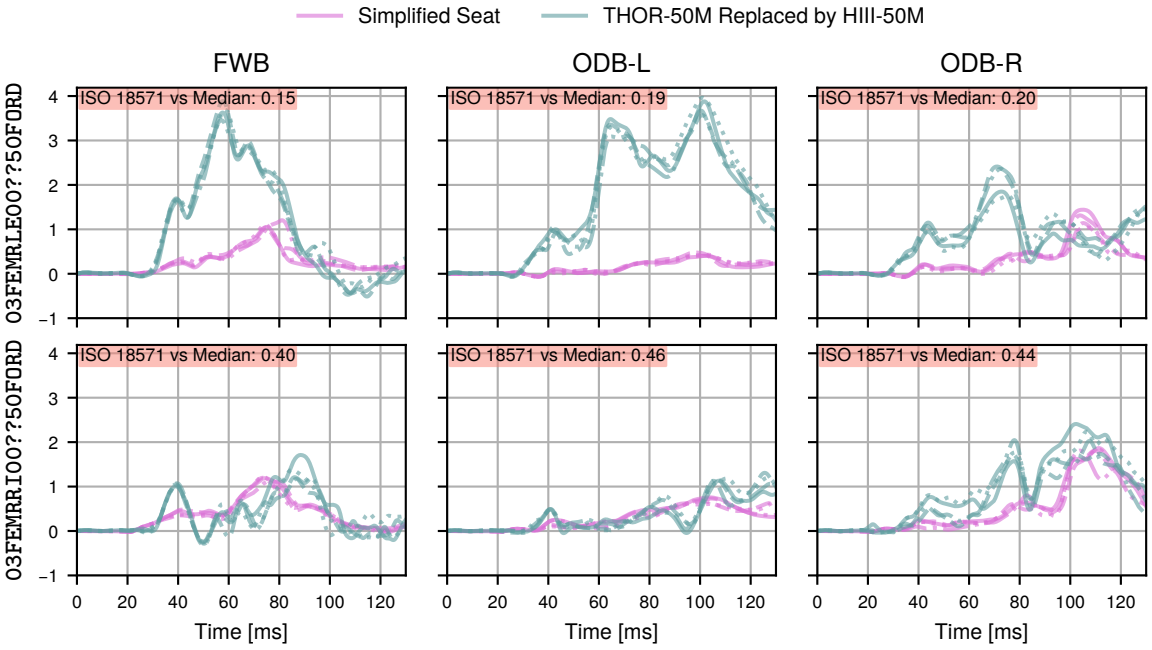


Figure 3.24: Femur Forces of HIII-50M versus previous FE-model Version with THOR-50M

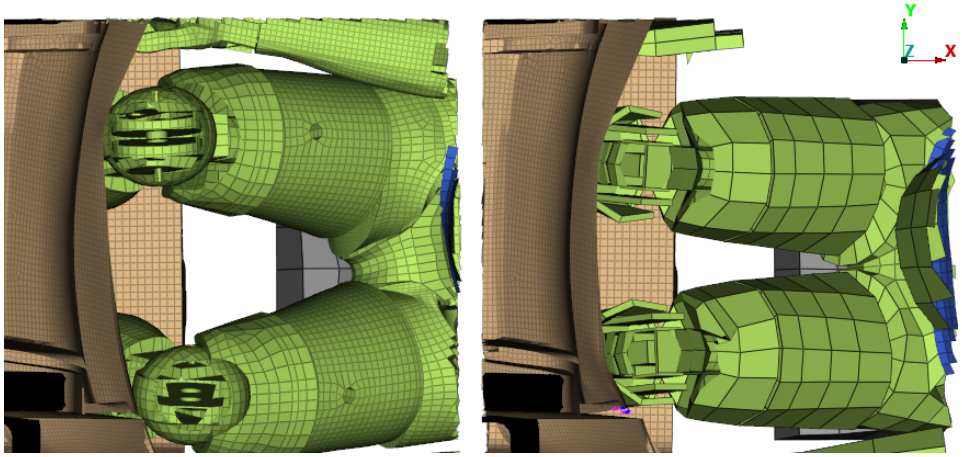


Figure 3.25: THOR-50M (left) vs HIII-50M (right) Knee Positions @ 60 ms in Full Width Barrier (FWB) Load case

3.2.6 Neglecting Curtain Airbag and Door

3.2.6.1 Method of Neglecting the Lateral Components

The vehicle side was simplified in two steps, where first the CAB was removed and then, the door. Both cases were only tested for the ODB-R load case because no influence on the other load cases.

If the vehicle were to experience strong lateral movement, removing the side restraint would have a significant impact on the ATD lateral motion. Due to the limitation to moderate lateral cases, such an influence was not anticipated.

3.2.6.2 Result of Neglecting the Lateral Components

Both simplification steps did not affect the prescribed vehicle motion. Hence, the ISO-grade was Excellent. The responses of all evaluated restraint systems were only minimally affected. All ISO-grades with the previous FE-model version were found to be at least Good.

The accelerations measured at the pelvis were found to be changed only slightly. The ISO-grades were Good. The same is true for the femur forces.

For head, chest, and pelvis, the resultant acceleration correlations were found to be ISO-graded as Good. The head's Z-acceleration and the chest's and pelvis' X- and Y-acceleration were found to have the ISO-grades Good too. The same is true for the pelvis' Z-acceleration.

Removing the CAB led to a Fair ISO-grade compared to the previous version for the head X- and Y-acceleration and the chest Z-acceleration in the ODB-R load case. However, the ISO-rating should not be overstated because the maxima of those accelerations were only $10\,g$. As recognizable in figure 3.26 on the next page, in the deviation time window, the numerical spread of the setups defined on page 81 is significant as well. Additionally, the oscillation was high. In this unstable phase, the correlation evaluation can be corrupted.

Two effects led to the low accelerations in the ODB-R load case. First, the seatbelt slips to the neck. That behaviour, induced by the lateral vehicle-pulse, is partly due to the seatbelt routing and partly due to the simplified modelling of the HIII-50M. The hardware-model has a metal ring around the neck area, mimicking the clavicle. In the author's experience, the shoulder seatbelt tends to remain at that ring's edge instead of getting into contact with the neck itself. The effect of that seatbelt behaviour is a late rotation of the ATD thorax, which leads to the second reason for the low accelerations. In combination with the relatively low acceleration but long stretched vehicle-pulse and the late upper body rotation, the PAB is reached so late that due to the venting, the airbag started deflating before the ATD even reached it.

The reduction of the computation time was significant. From in median 238 min with door and CAB, removing the CAB reduced to 187 min, and removing additionally the door to 165 min. In total, the simplification reduced the computation by 30 %. In comparison, for the light lateral content of the load case, the impact on the ATD response was ISO-graded Fair.

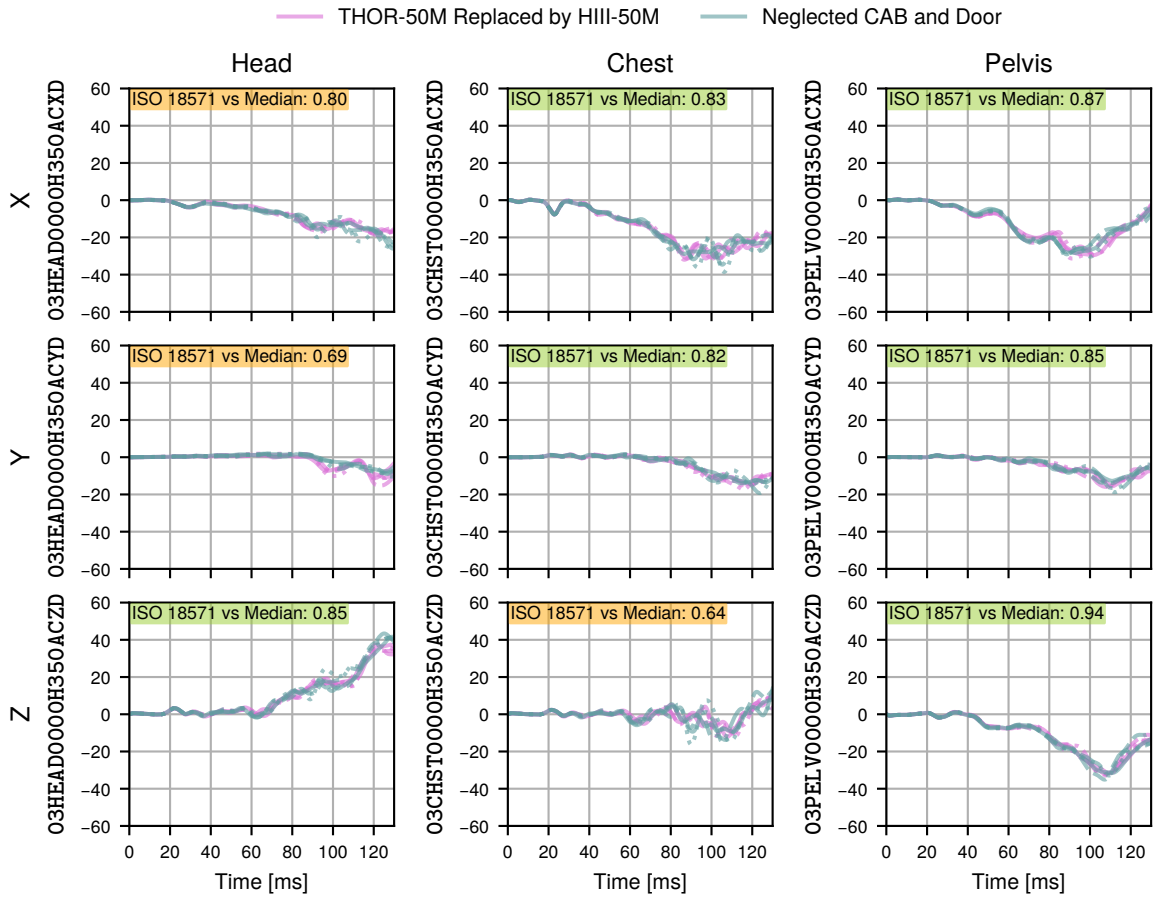


Figure 3.26: Removal of CAB and Door versus previous FE-model Version with CAB and Door in Offset Deformable Barrier Right (ODB-R) Load case

3.2.7 Simplifying the Seatbelt

3.2.7.1 Method of Seatbelt Simplification

The foremost change was the remeshing of the seatbelt itself. Before, each seatbelt shell element had an edge length of approximately 12 mm. After remeshing, the new length was 20 mm. As a second simplification, seatbelt anchor points were directly attached to the rigid and guided BIW. The seatbelt-buckle was set rigid, too.

3.2.7.2 Results of Seatbelt Simplification

The situation regarding vehicle motion remained unchanged. The agreements of the restraint system responses were ISO-graded Good.

Only an as Fair ISO-graded correlation was found for the pull-through of the seatbelt-buckle. Figure 3.27 on the facing page reveals a lower pull-through as the reason for the ISO-rating on

the example of the FWB load case. The deviation was mainly due to the change in stiffness of the seatbelt anchor points. In the original seatbelt, the seatbelt-buckle was attached to a deformable part of the seat and the component itself was deformable, too. Due to this modelling, the seatbelt-buckle was pulled towards the pelvis during the seatbelt tensioning. The simplified modelling features a pure rotational joint between the anchor point and the seatbelt-buckle; the seatbelt-buckle remains in its Y-Z-plane. Additionally, the bigger element size can affect both, the contact recognition to the ATD and the slipping behaviour. Over the different numerical setups of parallelization and precision, especially the B5-force seems to vary.

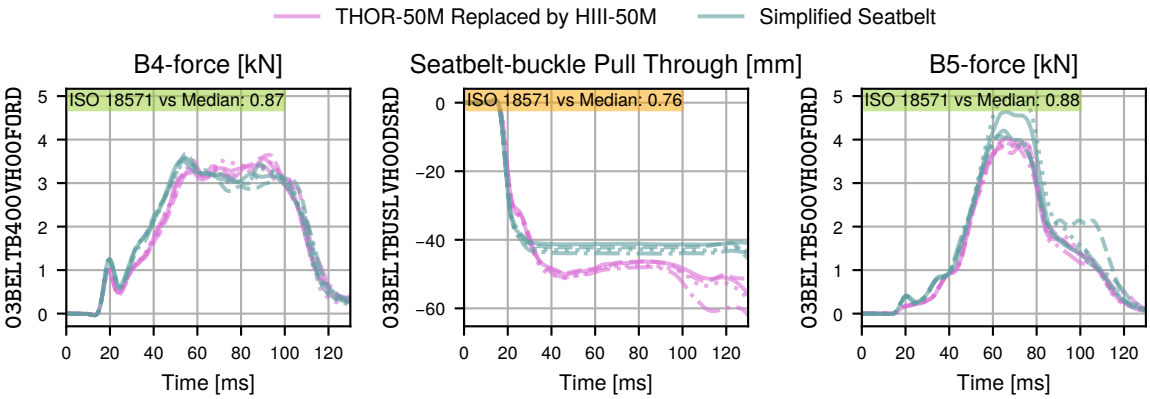


Figure 3.27: Effect of Seatbelt-buckle in Full Width Barrier (FWB) Load case with Simplified Belt versus previous FE-model Version with Original Belt

However, the slightly faster coupling of the thorax to the seatbelt and the different pull-through at the seatbelt-buckle influenced the whole ATD only minimal, as figure 3.28 on the next page shows. An observed effect was a reduced lateral motion of the entire ATD in the FWB load case. This effect can be traced again to the stiffer anchoring of the seatbelt-buckle. With the original seatbelt, the deformable anchor point at the seat was pulled in the X-direction. The anchor at the seatbelt end gets less deformed as a result of a lateral motion of the ATD. Since in the new model all anchor points were defined rigid, these characteristics were omitted.

The simplification step altered the ATD response minimal. The accelerations and femur forces ISO-graded Good regarding the agreement with the previous model, given the limitations described in the previous simplification steps. Only a minimal speed-up of the simulation was achieved. However, increasing the element size was another step to increase the integration interval ruling critical integration interval.

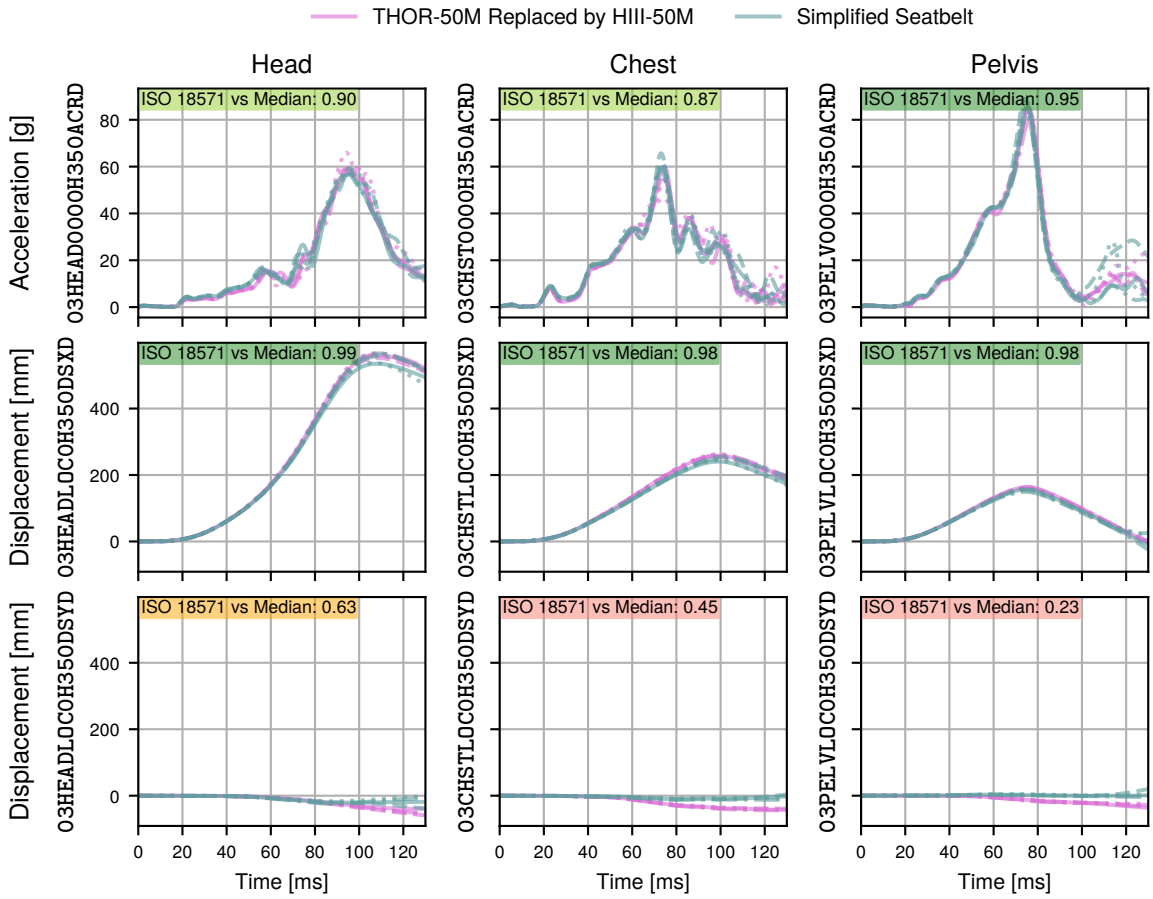


Figure 3.28: Resultant Acceleration and Local Displacement in Full Width Barrier (FWB) Load case with the Simplified Seatbelt versus previous FE-model Version with Original Seatbelt

3.2.8 Neglecting the Bulkhead's Intrusions

3.2.8.1 Method of Intrusion Neglection

In the reached FE-model degree of simplicity, the deformations of the BIW were only relevant in the form of the intrusions of the bulkhead. The interaction with the carpet ultimately affects the ATD lower extremities. The interaction with the dashboard can move the PAB towards the ATD. In figure 3.29 on the facing page intrusions are displayed as the combined maximum values of the FWB, ODB-L, and ODB-R load cases.

The intrusions D were calculated as relative displacements of a point ${}^{\text{Node}}\boldsymbol{U}$ with $\boldsymbol{U} \in \mathbb{R}^{3\times\mathcal{D}}$ of the bulkhead or the dashboard. The reference point ${}^{\text{Root}}\boldsymbol{U}$ was located in the vehicle c-column and another point ${}^{\text{Axis}}\boldsymbol{U}$ at approximately the same height on the Z-axis in the vehicle b-column. Both points formed a straight, on which ${}^{\text{Node}}\boldsymbol{U}$ was projected to measure the distance.

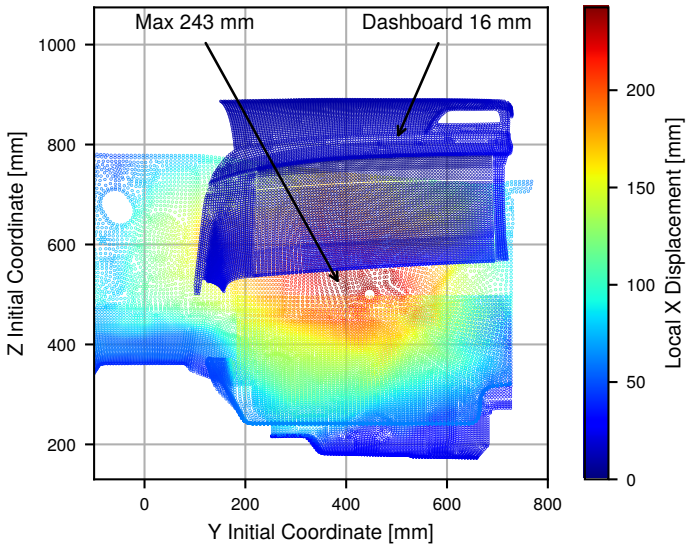


Figure 3.29: Combined Maximum Intrusions from Full Width Barrier (FWB) and Offset Deformable Barrier Left (ODB-L) & Offset Deformable Barrier Right (ODB-R) Load cases on Bulkhead, Carpet, and Dashboard

Given the assumption that the reference point located at the vehicle c-column is always behind the evaluated points at bulkhead or dashboard, the distance calculation simplifies to

$$ {}^{\text{Node}}D = \max_{0 \le t \le 120\,\text{ms}} \left\{ \frac{\sum_{d \in \{X,Y,Z\}} \left(\left| {}^{\text{Node}}U_{d,t} - {}^{\text{Root}}U_{d,t} \right| \cdot \left| {}^{\text{Axis}}U_{d,t} - {}^{\text{Root}}U_{d,t} \right| \right)}{\sqrt{\sum_{d \in \{X,Y,Z\}} \left({}^{\text{Axis}}U_{d,t} - {}^{\text{Ref}}U_{d,t} \right)^2}} \right\}. \tag{3.3} $$

The equation was derived from the projection of a point ${}^{\text{Node}}\boldsymbol{U}$ on a straight defined by ${}^{\text{Root}}\boldsymbol{U}$ and ${}^{\text{Axis}}\boldsymbol{U}$. The coordinates of the reference point ${}^{\text{Root}}\boldsymbol{U}$, the direction point ${}^{\text{Axis}}\boldsymbol{U}$, and the evaluated point ${}^{\text{Node}}\boldsymbol{U}$ were functions of time and their coordinates corresponded to a node of the FE-model. From each evaluated point, the maximum over time was used.

In the graphic of figure 3.29, the maximum displacement of the dashboard towards the ATD is small. The intrusions acting on the carpet, however, range from 100 mm to 240 mm. An influence on the ATD leg kinematics was expected. Regarding intrusions, mainly the dashboard influences the ATD upper body kinematics because it mounts the PAB and knee bolster. Due to the small maximum displacement of the dashboard, only a minimal influence of removal of the applied intrusions was expected.

To reduce the overhead, the BIW was remeshed to a 20 mm target size. Theoretically, this step could have been done before, since the interface algorithms of LS-DYNA support the mapping between different sized meshes. However, to support a stable contact situation between the BIW and the fine meshed dashboard and CAB, this step was moved until the removal of the intrusions.

3.2.8.2 Result of Neglecting the Bulkhead's Intrusions

The agreement of the vehicle-pulses continued to be ISO-graded Excellent. The evaluated accelerometer was mounted in the area of the always rigidized part of the BIW. Consequently, the model conditions remained unchanged here. All sensor outputs of the restraint system were at least ISO-graded Good, in agreement with the previous FE-model version. All acceleration sensors were in good agreement as well, considering the above declared limitations.

The femur forces, shown in figure 3.30, were affected as anticipated. However, no clear trend over the load cases was recognized. The ISO-grades ranged from Fair to Poor.

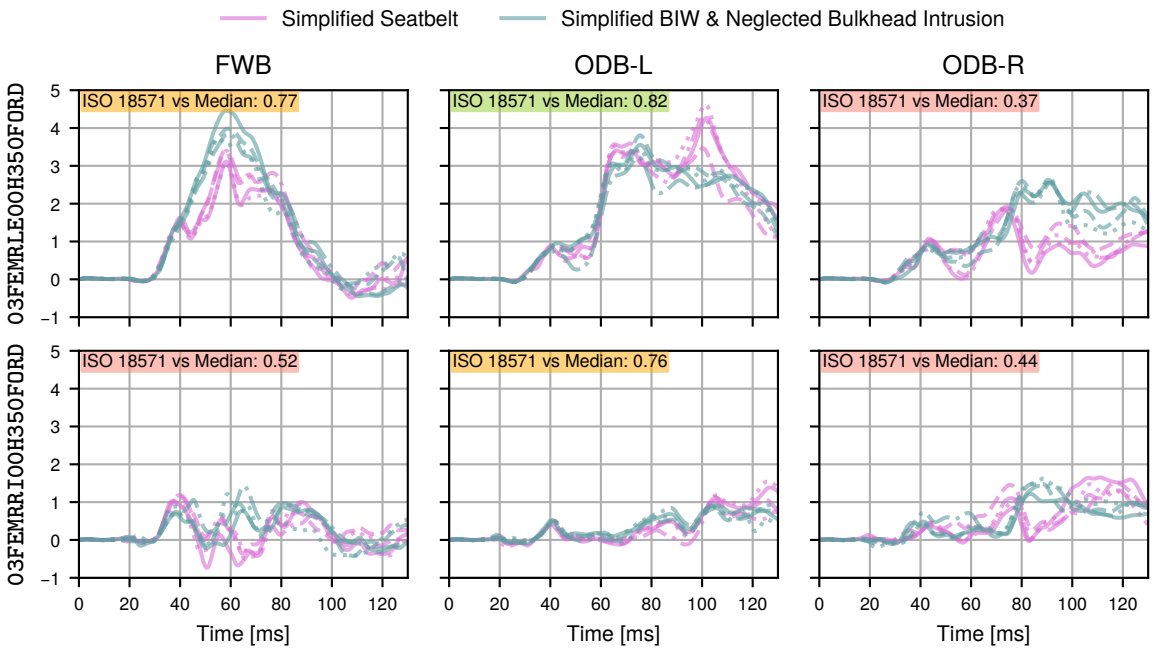


Figure 3.30: Femur Forces when Intrusions are Removed versus previous FE-model Version with Applied Intrusions

The simulation's computation time was mostly affected by the simplification of the BIW and was reduced from 163 min to 91 min. Additionally, neglecting the intrusions, the reduction to 88 min in median was achieved. This last reduction seemed insignificant, especially considering the change in the femur forces. However, as stated above, the femur was considered a lower priority in this study. Furthermore, the measure was the prerequisite for the subsequent steps.

3.2.9 Interior Simplifications

3.2.9.1 Method of Simplifying Dashboard and Carpet

Since the intrusions are neglected, the dashboard and carpet can be simplified now. For both components, the stiffness is especially relevant if the ATD is impacting.

The carpet was modelled as completely rigid. Despite a foam-like elastic modulus for the contact calculation, an unrealistic hard impact must be anticipated. However, it was assumed that the feet are supported by it but do not impact. Furthermore, the evaluation of sensors on the lower extremities was not in the scope of this study.

For the dashboard, the same assumptions as for the carpet were chosen, and the component was modelled as rigid. Due to the comparably compressible door of the glove compartment, a knee bolster seemed necessary. In this area, the knees are impacting the dashboard. High deformations and hence higher forward displacement of the pelvis can affect the overall body kinematics. The knee bolster was approximately 100 mm thick and for the material a carpet foam was selected. The rest of the dashboard's elastic modulus definitions for contact calculation remained as in the initial FE-model.

3.2.9.2 Result of Simplifying Dashboard and Carpet

The application strategy of the vehicle motion remained unchanged. Hence, the related sensor signal remain same too.

The restraint system responses are overall at ISO-graded as Good. The exceptions are the lap seatbelt forces. Due to the change in dashboard stiffness, their match with the previous version was only ISO-graded as Fair to Poor. The simplification of the carpet did not introduce a significant deviation within the restraint responses. In figure 3.31 on the next page are, as an example, the pelvis seatbelt forces in the FWB load case shown. The simplified dashboard with the integrated knee bolster led to a reduction in the seatbelt forces.

The femur forces were affected as in figure 3.32 on the following page. More homogeneous stiffness of the dashboard's knee zone led to more equally distributed femur load on the left and right leg. The local stiffness for the left leg seems to be in a range similar to the initial FE-model. The local stiffness in the area of the right knee, however, seems to be modelled stiffer.

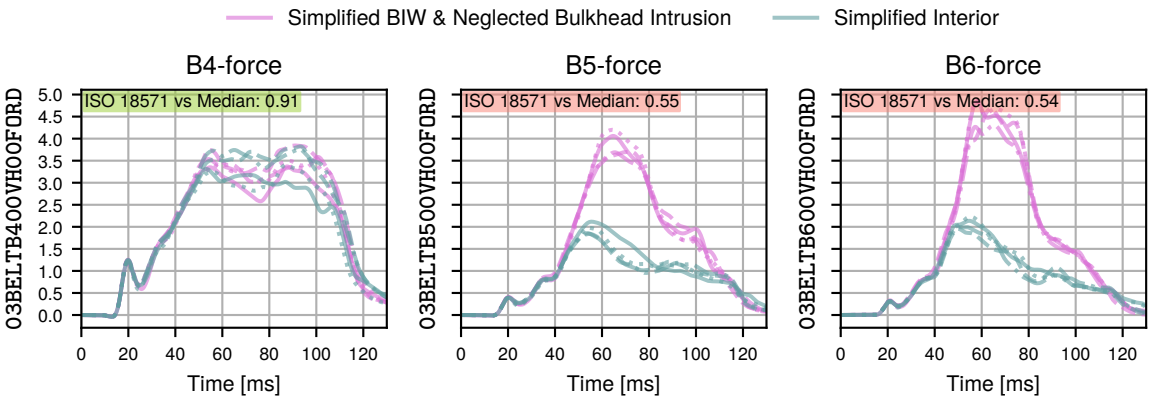


Figure 3.31: Change of Pelvis Seatbelt Forces in Full Width Barrier (FWB) Load case with Dashboard Simplification versus previous FE-model Version with Original Interior

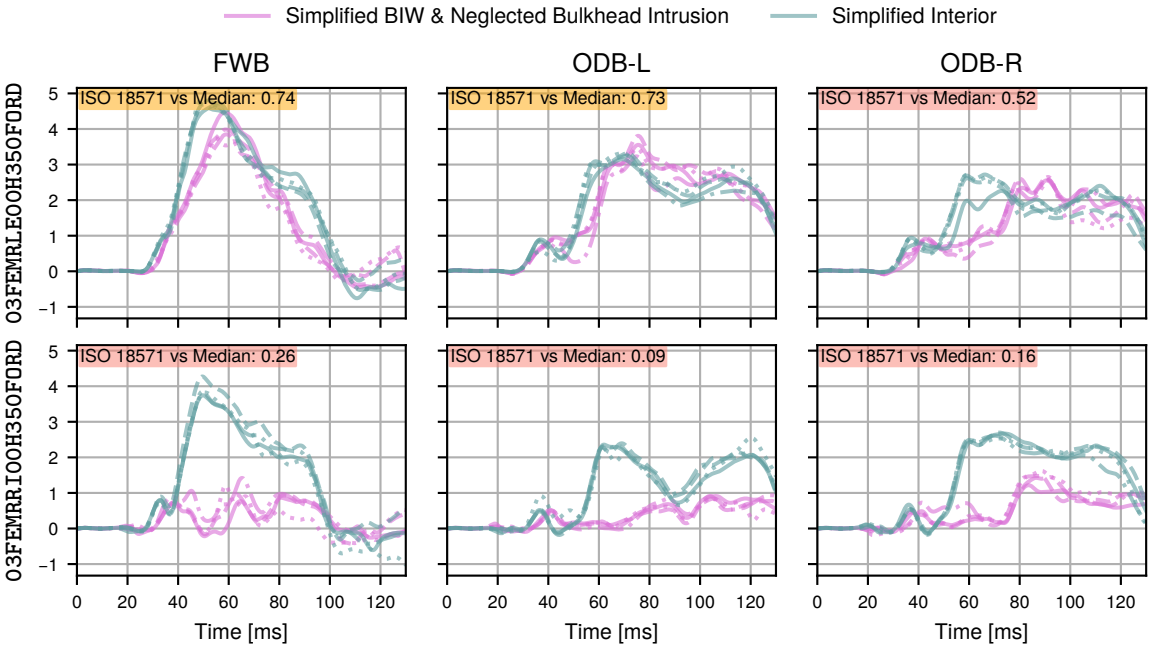


Figure 3.32: Femur Forces with Simplified Dashboard and Rigid Carpet versus previous FE-model Version with Original Interior

3.2.10 Evaluation of Final Simplified Model

The entire simplification process was motivated by enabling the creation of an extensive database by a reasonable short computation time. To validate the machine learning approaches, the more realistic and less simplified version of the vehicle model should be used. The initial FE-model with THOR-50M seemed not useful for this task due to the very long computation time and the unavailability of a 95th percentile Test Device for Human Occupant Restraint (THOR). The FE-model version with introduction of the HIII-50M seemed a good compromise.

In figure 3.33 on the next page, the ISO-ratings are displayed as the median over the four numerical setups for each of the evaluated channels. Since the reference model was already a guided BIW, all vehicle-pulse related channels were identical. Almost identical were the prescribed load limiter B0-force and the PAB characteristics. With the introduction of the simplified dashboard on page 109 ff, the lap seatbelt load dropped from ca. 4 kN to 2 kN, resulting in the measured Poor ISO-grade. The deviation in the pull-through of the seatbelt at the seatbelt-buckle was caused by the rigid anchor points, as discussed on page 104 ff. The different pull-through affected the retractor pull-out in the ODB load cases, too.

Most Y-accelerations and local Y-displacements were ISO-graded Poor. The reason was the relatively low magnitude and high oscillations of the signals, which resulted in an overcritical assessment by the ISO-rating. Furthermore, the simplified seatbelt influenced the Y-motion (see page 104 ff).

The stiffer impact area on the dashboard influenced for the right leg, the chest, and pelvis acceleration in the ODB-R load case as discussed on page 109. By this measure, the ISO-grade was reduced to Fair.

The ISO-rating is a mixture of different metrics. Hence, single aspects can get obfuscated despite being relevant for the calculation of injury-criteria as they were defined in section 2.1 on page 34 ff. Table 3.3 on page 113 displays the resulting deviation of injury-criteria between the two FE-models. For each load case was the Mean Absolute Error (MAE) (see section 2.2, equation 2.44 on page 44) for the four numerical setups calculated within the time interval from 20 ms to 120 ms. The limit-values were taken from the regulation UN ECE R137 except for Head Injury Criterion for up to 15 ms (HIC_{15}) and Chest Acceleration over 3 ms Criterion (CAC_3), which were taken from the FMVSS 208. The MAEs were found to be lower than 15 % of the related limit, except for the CAC_3 for the FWB.

In figure 3.34 on page 113 the difference in the two rating approaches becomes clear. Because of the high similarity of the signals in the FWB load case, the ISO-grade for the chest acceleration was Good. The sharp peak had only a small influence on ISO-rating, while for the CAC_3 a deviation of 22 % was observed. Comparing the Head Acceleration over 3 ms Criterion (HAC_3) for the ODB-R, the deviation was only 4 % while the ISO-rating for the related head acceleration was with 0.84 similar to the 0.83 of the discussed chest acceleration.

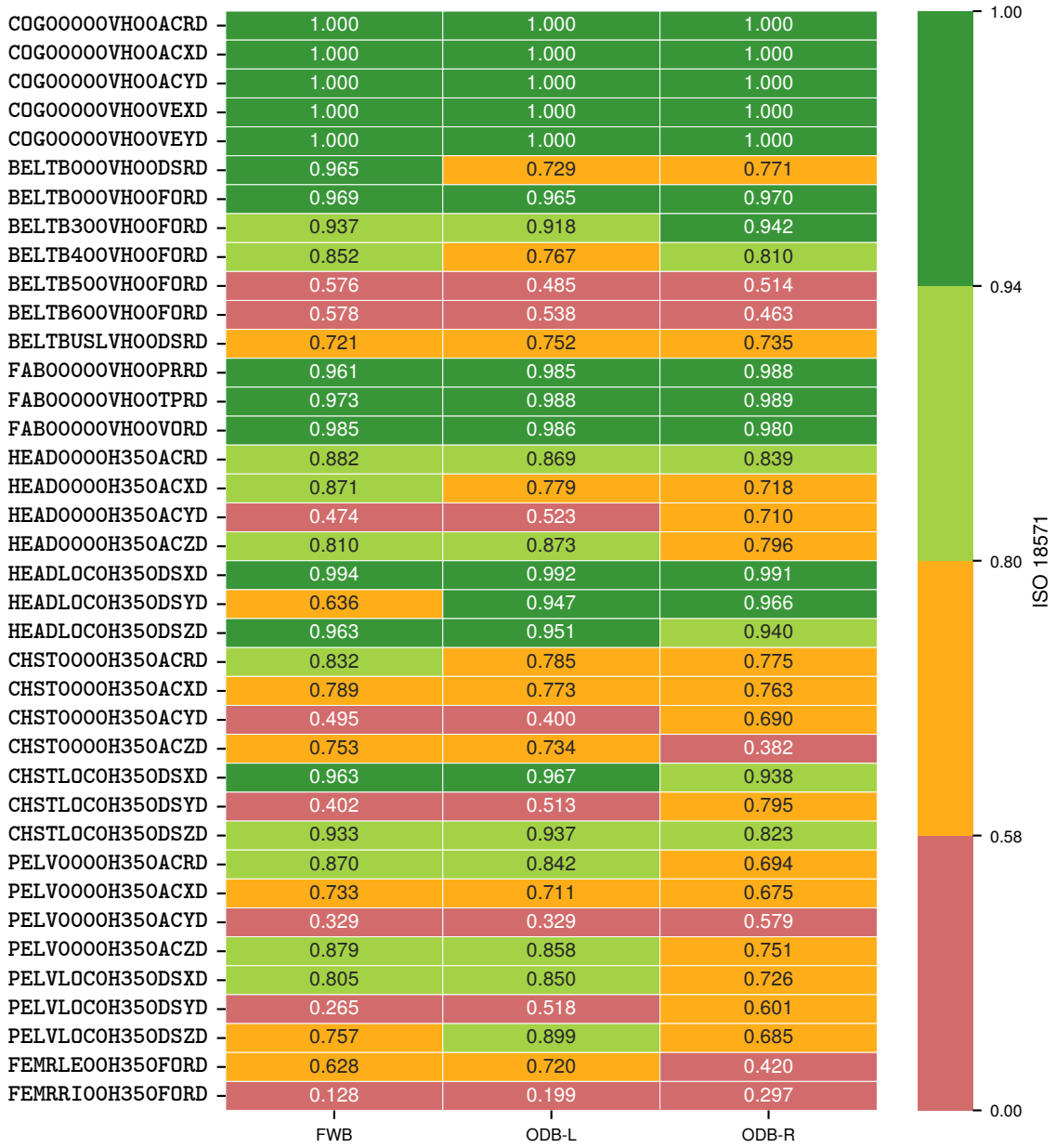


Figure 3.33: ISO-ratings of Channels from HIII-50M Reference versus Fully Simplified FE-model in Multiple Load cases

Table 3.3: Mean Absolute Error (MAE) for Numerical Setups of HIII-50M Reference vs Simplified FE-model

Region	Injury-criterion	Unit	Limit	MAE per Load case		
				FWB	ODB-L	ODB-R
Head	HIC_{15}		700	38.9	22.5	30.1
	HAC_3	g	80	2.80	3.32	5.09
Neck	NEC	N m	57	6.88	1.25	3.39
	NTC	kN	3.3	0.37	0.19	0.31
	NSC	kN	3.1	0.11	0.09	0.14
Chest	CDC	mm	42	3.43	6.47	6.75
	CVC	$m\,s^{-1}$	1.0	0.04	0.04	0.03
	CAC_3	g	60	13.5	3.54	4.11
Femur	FCC	kN	6.8	1.10	0.51	0.33

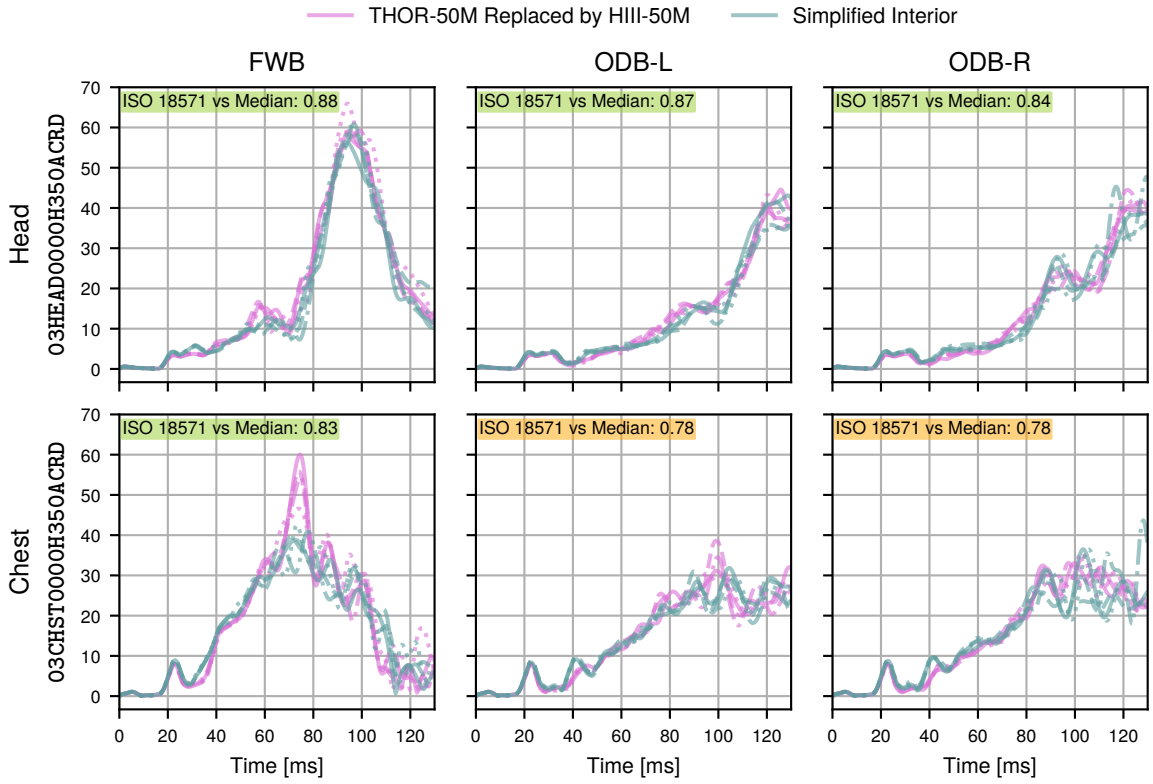


Figure 3.34: HIII-50M Reference versus Simplified FE-model with Resultant Accelerations of Head and Chest

In summary, the major differences were introduced by altering the stiffness of the dashboard's knee bolster and the rigid mounting of the seatbelt anchor points. Both measures affected the ATD. But actually, from the perspective of a restraint system developer, the system was improved. The reference simulation took ca. 2 h to be computed using single precision and 48 processes. The massively simplified FE-model had a computation time for the same setup of 10 min. The speed-up of ca. 90 % seems to justify the deviations between the two FE-model versions.

3.2.11 Model Adjustments for Data Generation

3.2.11.1 Method to Prepare FE-model for Data Generation

The purpose of the previous section was to explain the rationale behind the simplification steps. The final FE-model had sufficiently low computational costs, but was not yet configured for a high number of FE-simulation. The vehicle environment had to be parametrized and the Hybrid III 5th Percentile Female Dummy (HIII-05F) and Hybrid III 95th Percentile Male Dummy (HIII-95M) should also be seated.

To parametrize the restraint system, the airbag mass flow curves over time were assigned with a Y-scaling factor. The opening time of the adaptive vent defined two dependant X-values in the mass-out-time-curve. The retractor and airbag time-to-fires were already factors in the original model. As a conclusion from the observed tendency of the PAB being relatively soft, the airbag time-to-fire was delayed by 4 ms. For the retractor load limiter, two dependant support points were defined using the existing support points, controlled by one factor for the limiting force.

To parametrize the vehicle motion, the vehicle-pulse application method was simplified. The vehicle-pulse had to match the full-crash motion first. However, the scalability in X- and Y-direction seemed desirable. To achieve matching and scalability, the "prescribed accelerometers" were replaced by a "prescribed motion" applied to the CoG of the vehicle. The X-acceleration was recorded in the previous FE-model version at the CoG in the FWB load case in the numerical setup of single precision with 48 parallel processes. The Z-acceleration and the 3 rotational Degree of Freedoms (DoFs) were set to zero throughout the simulation. The Y-acceleration was scaled from the X-acceleration. A vehicle-pulse angle factor was introduced for scaling, applied as cosine for the X- and sine for the Y-direction on the original X-vehicle-pulse. The approach is the same as the standard procedure for emulating a lateral motion from a full crash test on 1 DoF sled test device. In these applications, the BIW is mounted Z-rotated on the sled. In addition to the initial Z-rotation factor, the Y-value of the vehicle-pulse was scalable as a whole, too.

Lastly, the three ATDs had to be seated. Following the argumentation on page 79, all percentiles were placed on the same seat-position. The middle position remains the same as in all previous simulations. The actual seat adjustment field of the Honda Accord 2014 is not known to the author. Nevertheless, since the original test position for THOR-50M is the middle setting on the seat tracks, it seemed a reasonable choice. In preparation for potential later studies with more anthropometrical configurations and backrest angles, a reproducible seating procedure

was defined. In the following procedure, “near” means 3 mm to 10 mm orienting on typical shell thicknesses, given the modelling on the shell’s mid-surface. The procedure contains the steps:

1. Perform the initial ATD pre-positioning (keep all internal angles constant).
   (a) H-point Y-coordinate in seat-lane (387 mm for passenger in the Honda Accord).
   (b) ATD sagittal plane equals vehicle’s X-Z-plane.
   (c) ATD direction of view in driving direction (negative X).
   (d) Rotate pelvis around Y-axis until the back (line of pelvis and thorax) parallel to the seat’s backrest.
   (e) Translate the ATD in X- and Z-direction until the ATD back is near to the seat’s backrest.
   (f) Translate the ATD in X- and Z-direction along the seat’s backrest until the pelvis is near to the seat’s sitting-pillow.
2. Adjust the legs iteratively, until:
   - Heels are on the floor, and
   - Feet are parallel to the carpet-wall and as near as possible to the carpet wall.
3. Rotate arms around the respective shoulder-joints Y-axis to 23°.
4. Adjust the elbow joint until the hands and knees are approximately in one line.
5. Rotate the neck bracket around Y as near as possible to “head is in water”.
6. Check minimum distance for all surrounding interior components.

After determining the ATDs position and inner angle, the marionette method was used to generate the positioned model. The method is implemented in the software ANSA. In an initial position of the ATD, landmark points like elbow and knee are moved to the desired position. For this work, a force-based definition was used.

The overall procedure oriented on typical processes as they are for example defined by Euro NCAP (2023c). Following the new procedure, the HIII-50M now deviates from the position during the simplification steps. During those steps, the comparability was paramount. Consequently, it was attempted to position the HIII-50M as close as possible to the THOR-50M.

In summary, the changes did not alter the vehicle model. Hence, the validity status of the simplification chain would not be compromised. All changes were within the usual scope of optimization of the restraint system. However, the response of the reference HIII-50M will be altered.

3.2.11.2 Result of Preparing FE-model for Data Generation

In figure 3.35, the new sitting position is compared to the previous one. The previous position was related to the THOR-50M, while the new one is the result of the defined procedure. The most apparent difference is the head position nearer to PAB and levelled in the global X-Y-plane. Second, because of the different pelvis positioning strategy, the new knees are significantly further away from the dashboard.

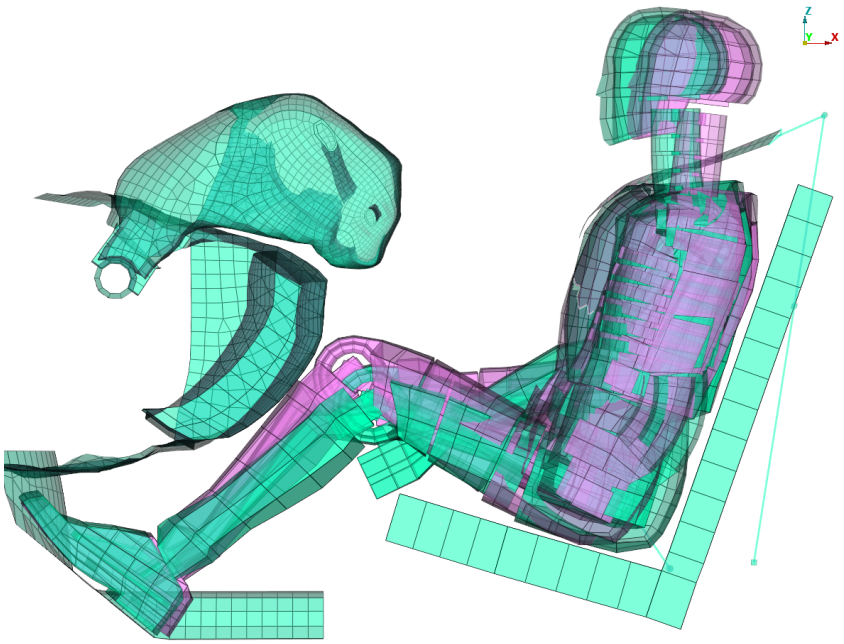


Figure 3.35: Initial Sitting Position of HIII-50M from Simplification Chain (violet) and from Sitting Procedure (turquoise)

In figure 3.36 on the facing page, the effect of the different vehicle-pulse application approach is shown for the FWB load case. Minimal deviation can be recognized for the X-acceleration. The minor oscillations around zero in the Y-acceleration are with the alternative approach removed. In the simplification chain, the acceleration was applied using multiple locations corresponding to the recording nodes in the initial full-crash FE-model. In the new FE-model, the parametrized acceleration signal was applied directly to the vehicle's CoG.

The seatbelt signals in figure 3.37 on the next page show a high similarity for the shoulder seatbelt but reveal a higher slack in the lap part. The retractor pulls more seatbelt from the lap through the seatbelt-buckle. The lap seatbelt force peak is ca. 4 kN higher due to the initially higher distance of the knees to the dashboard's knee bolster and the further located pelvis.

Only minimal deviation can be observed between both airbag characteristics in figure 3.38 on page 118. The slightly later inflator activation and the different sitting position seem to result in a more homogeneous airbag unloading.

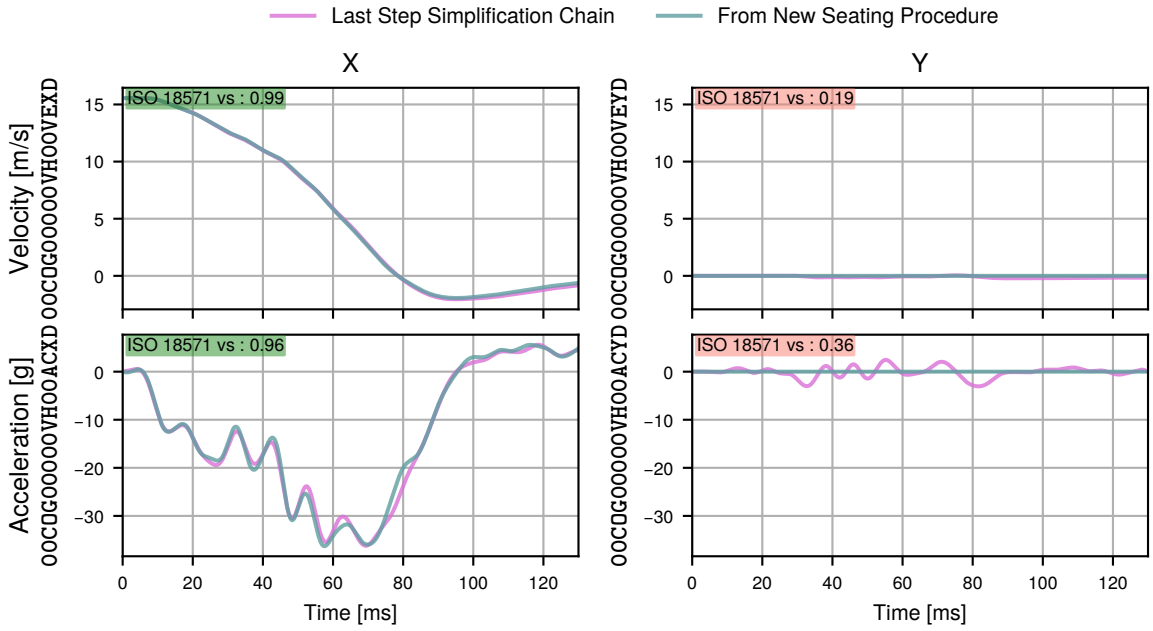


Figure 3.36: BIW Kinematics at Vehicle CoG in Full Width Barrier (FWB) Load case for Final FE-model versus Last FE-model of Simplification Chain

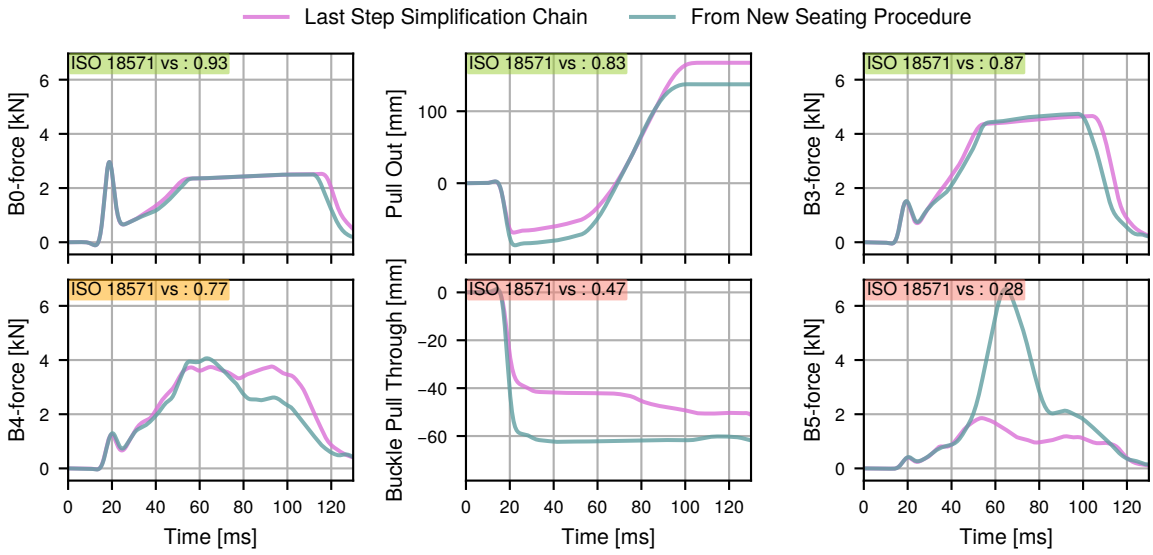


Figure 3.37: Seatbelt Characteristics in Full Width Barrier (FWB) Load case for Final FE-model versus Last FE-model of Simplification Chain and CFC 60

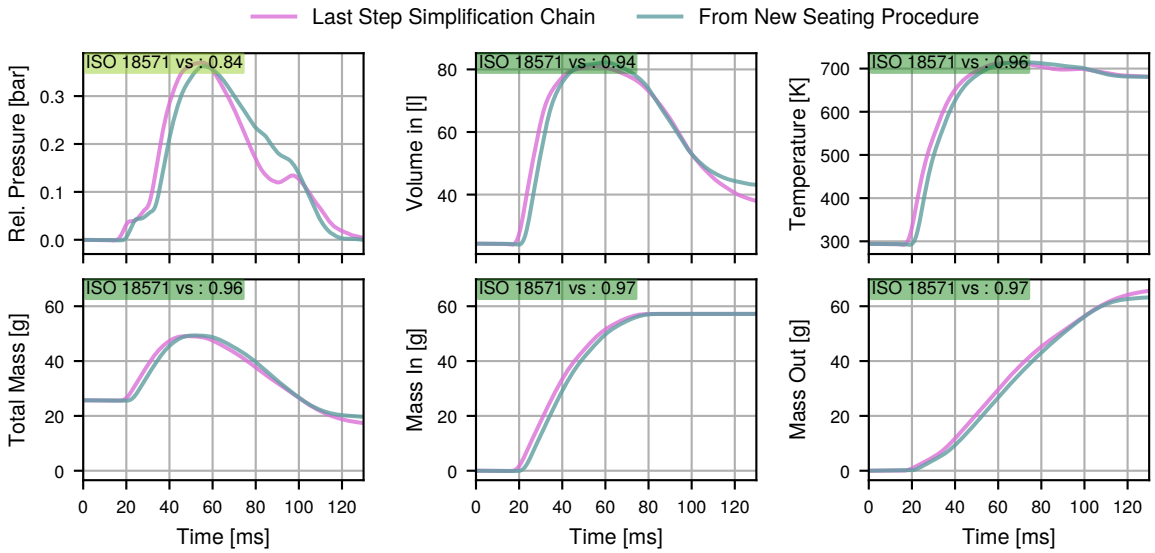


Figure 3.38: PAB Characteristics in Full Width Barrier (FWB) Load case for Final FE-model versus Last FE-model of Simplification Chain and CFC 60

Figure 3.39 on the next page compares the ATD kinematics. Only minimal changes are observable in the local Z-displacements. The head appears to be more restrained; the X-acceleration is slightly higher, while the X-displacement is ca. 80 mm less. The higher restraint is also reflected in the minimum distance between the head and the dashboard. In the final FE-model version from the simplification chain, the rest depth of the airbag was approximately 50 mm and in the new model ca. 65 mm. Corresponding with the higher distance of the knees to the dashboard, the pelvis accelerations are significantly higher.

The comparison of the three ATD sizes is displayed in figure 3.40 on page 120 for the FWB load case. The X-accelerations of the heads have all a peak around 95 ms with ca. 50 g. The local maximum X-displacements were sorted by the ATD sizes. No clear tendency was found in the pelvis Z-acceleration. The effects of the longer upper legs of the HIII-95M and the lower weight of the HIII-05F compared to the HIII-50M both reduce the maximum Z-acceleration.

In summary, especially the new sitting position influenced the kinematics of the HIII-50M. The effect of the higher distance of the knees to the dashboard seemed dominant. The FE-model appeared numerically stable and the restraint system functional for the additional percentiles.

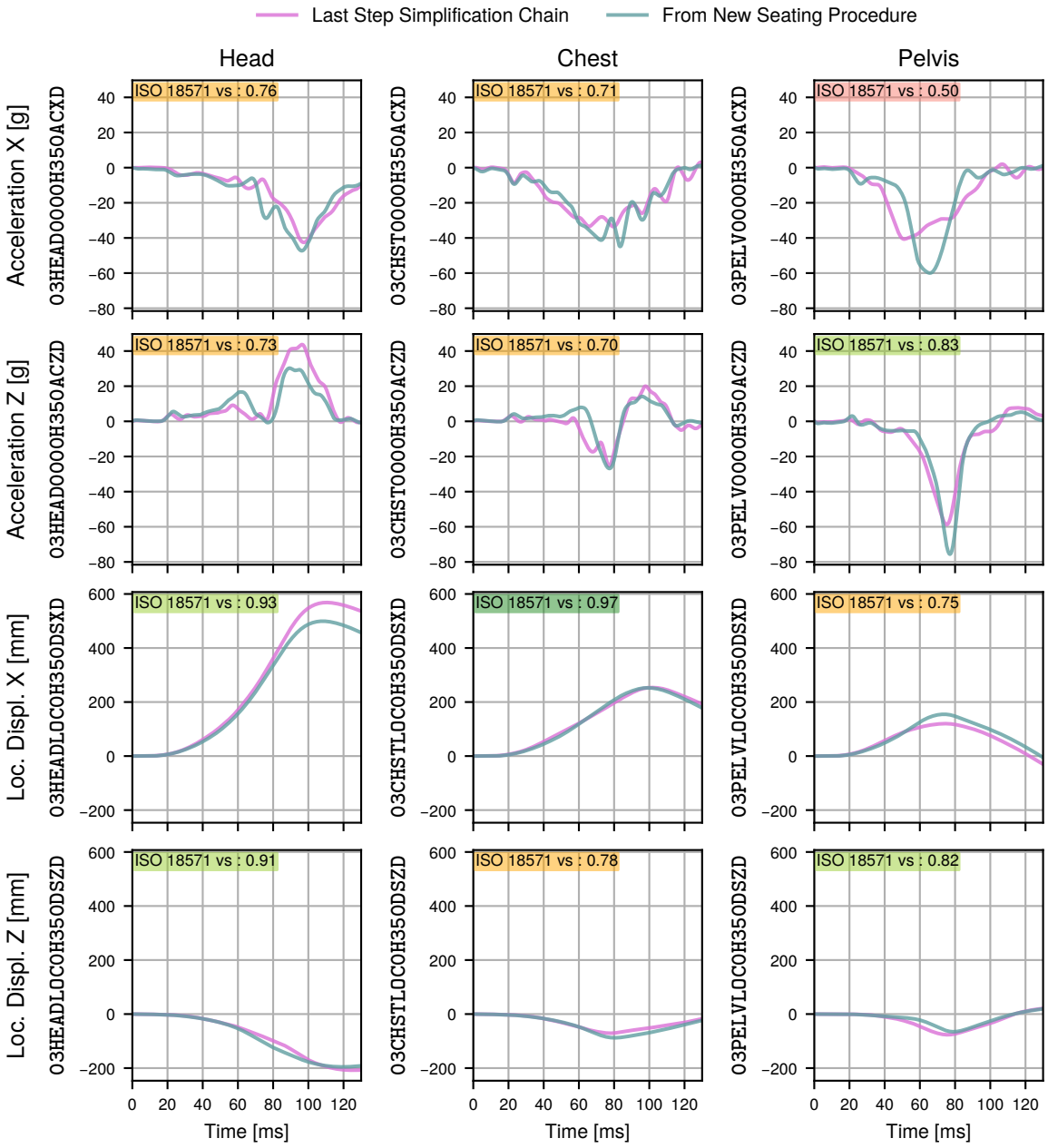


Figure 3.39: ATD Kinematics in Full Width Barrier (FWB) Load case for Final FE-model versus Last FE-model of Simplification Chain

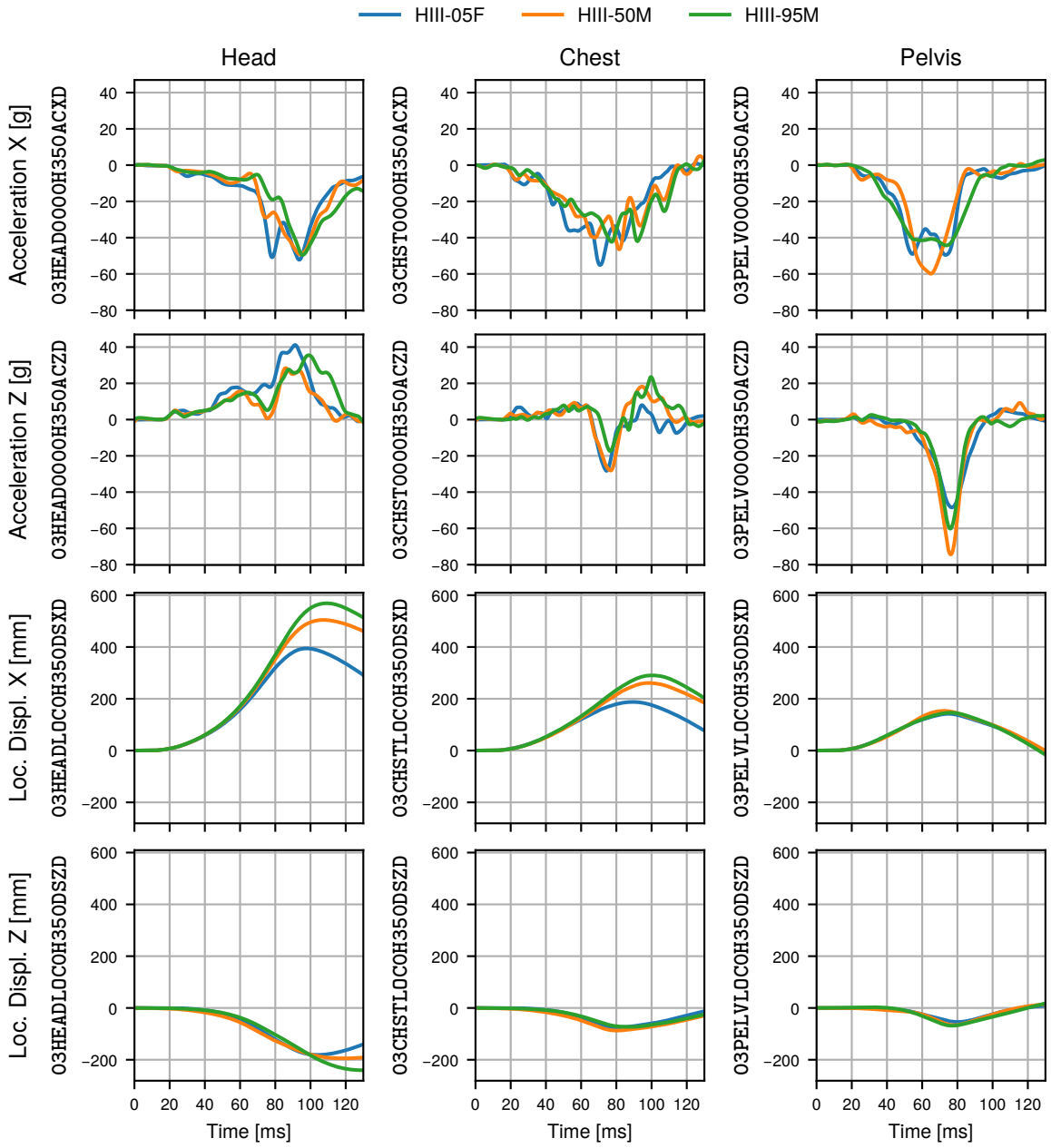


Figure 3.40: Comparison of HIII-05F, HIII-50M, and HIII-95M in Full Width Barrier (FWB) Load case

3.3 Alternative Model Generation – VIRTHUMAN

3.3.1 Motivation and Approach

For the potential application of transfer learning strategies, another more complex human surrogate FE-model should be used. A metamodel, trained on an extensive database filled with FE-simulations from one Anthropomorphic Testing Device (ATD) FE-model, should be used to be tuned on a smaller database with FE-simulations with a complex human surrogate. Natural candidates would be the Test Device for Human Occupant Restraint (THOR), or a more detailed version of the Hybrid IIIs (HIIIs), or Human Body Models (HBMs). THOR and the highly detailed HBMs like Total Human Model for Safety (THUMS) seemed computationally expensive for this work's purpose. The more detailed HIII FE-model could be somewhat too similar to the used HIII version. In contrast, the fast calculating but as HBM classified VIRTHUMAN (VHUM) seemed to be a sufficient example. The VHUM FE-model was introduced in section 2.1 on page 25.

To be useable in any transfer learning application, the metamodel input-space should be similar to the one it was originally trained on. Consequently, the VHUM FE-model should be seated in a similar vehicle environment. Since the VHUM FE-model exists only within the software Virtual Performance Solution (ESI) (VPS) but the above described FE-model was developed in Livermore Software technology corporation DYNA (LS-DYNA), the FE-model had to be translated between both software packages. The concept for the transfer between both FE-models are shown in figure 3.41.

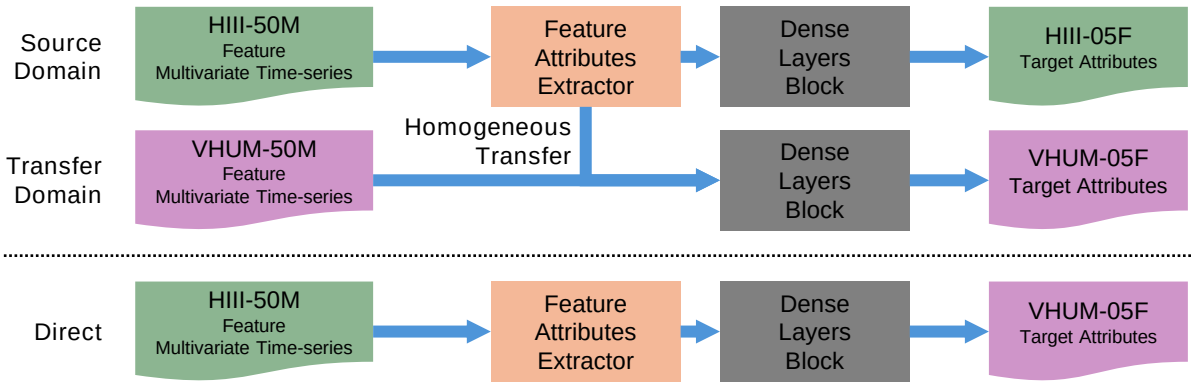


Figure 3.41: Concept of Transfer Learning between HIII and VHUM of the Example of Feature Attributes Regression for the 5th Percentile Human surrogate

A second approach was inspired by the work presented by Plaschkies (2023). Instead of using the complex FE-model version as the reference, the simplified one should take its place. In this case, the Hybrid III 50th Percentile Male Dummy (HIII-50M) should be used to estimate the results of FE-simulations with VIRTHUMAN 50th Percentile Male Human Body Model adapted to HIII-50M (HIII-50M). The other combinations of percentiles are imaginable, too.

The FE-model translation was carried out in the frame of the EU project AWARE2ALL[1] in collaboration with the ESI Group[2]. The author provided the initial FE-model and guidance in the validation approach. ESI facilitated the actual translation. Since both partners used different FE-simulation tools, no partner had full access to the detailed data from the FE-models of both software packages, limiting the explainability of observed effects.

In the first step, the vehicle FE-model was exchanged and translated. The HIII-50M FE-model was taken from the ESI in-house pool. The FE-simulation results from the LS-DYNA and VPS FE-models were compared and the factors of the VPS were adjusted to fit to the LS-DYNA reference. First the vehicle-pulse was matched, then the restraint system and finally the ATD response. After achieving a satisfactory correlation, the HIII-50M was replaced by VHUM-50M. The validation process was carried out in the Full Width Barrier (FWB) load case.

3.3.2 Validation of the VPS FE-model with HIII-50M

The pulse comparison in figure 3.42 on the facing page showed a perfect correlation between the model in LS-DYNA and in VPS. In both FE-models, the vehicle-pulse was applied to the rigid Body In White (BIW) on its Centre of Gravity (CoG). The constraining algorithms seem to lead to the same result in both software packages, given the identical constraining vehicle-pulse.

The next validation step was regarding the seatbelt system. The comparison is displayed in figure 3.43 on the next page. Well-matched are the load limiter plateau and the shoulder seatbelt B3-force as defined in section 2.1 in figure 2.1 on page 8. For the retractor, a simple bar-model was implemented in the VPS FE-model. In contrast, in LS-DYNA, the seatbelt tensioning was modelled with a force impulse. The impulse resulted in the retractor peak while the bar applied a constant force. The peak is more in line with a real retractor. However, for the scope of the study, the manifested force plateau seemed more relevant.

The comparison of the seatbelt routing explained the ca. 70 mm higher seatbelt pull-in at tensioning. In the VPS FE-model, the seatbelt was, especially in the lap area, routed tighter to the body's surface than in the LS-DYNA FE-model. Since in both cases the slack was completely pulled out, the difference was not interpreted as problematic.

The different retractor modelling and seatbelt slack can explain the faster coupling of the ATD to the seatbelt in the VPS FE-model. Due to limited insight to the VPS FE-model, the differences in the lap seatbelt forces could not be fully explained by the author. The analysis of the pelvic-kinematics below hinted to a stiffer seat cushion or ATD pelvis rubber in the VPS FE-model.

[1] https://aware2all.eu/

[2] https://www.esi-group.com/

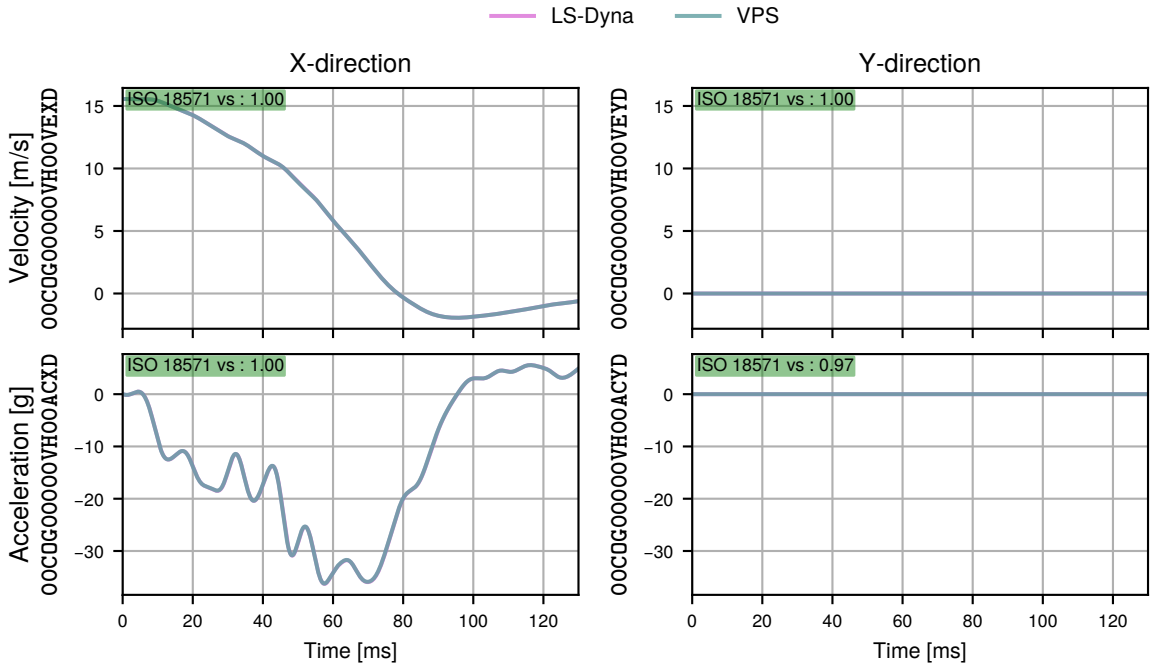


Figure 3.42: Vehicle-pulse in Full Width Barrier (FWB) Load case at BIW CoG with LS-DYNA versus VPS FE-model

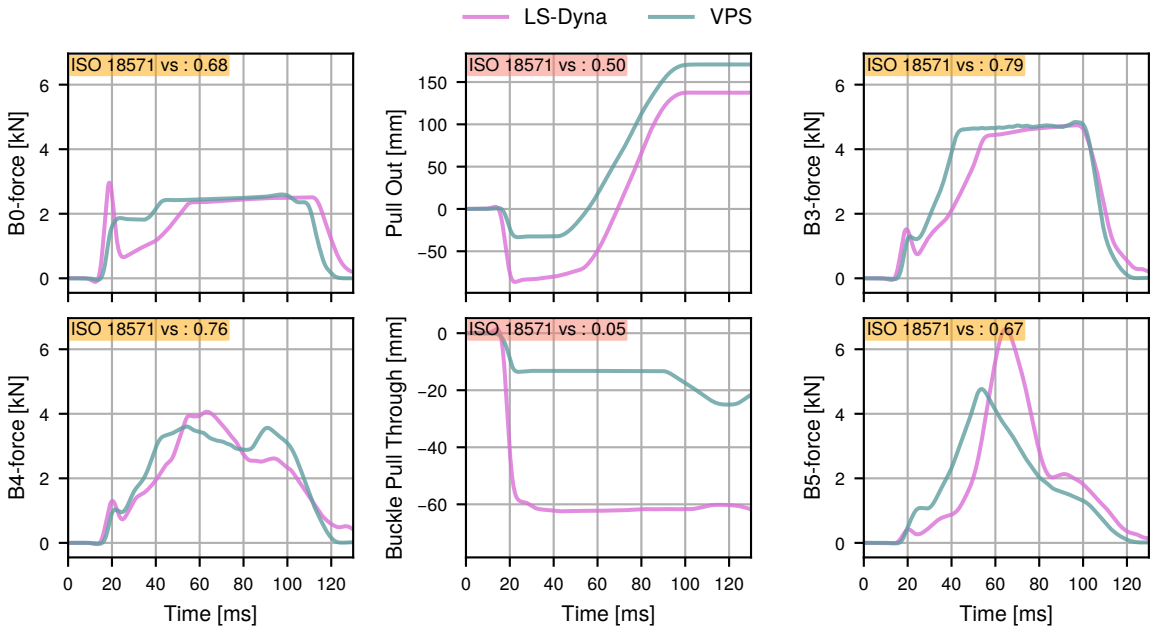


Figure 3.43: Seatbelt System in Full Width Barrier (FWB) Load case with LS-DYNA versus VPS FE-model using HIII-50M and CFC 60

The most challenging task was the translation of the Passenger Airbag (PAB) FE-model. Both software packages utilize different algorithms. In the finally accepted FE-model, displayed in figure 3.44, the initial gas mass in the pre-inflated airbag was higher in the VPS FE-model. This was balanced with lower in and out mass flows and the slightly lower temperature. The most important characteristics considered were the pressure and volume in the working time window around 60 ms. Here, the pressure was well-matched. The volume in the LS-DYNA FE-model was slightly higher.

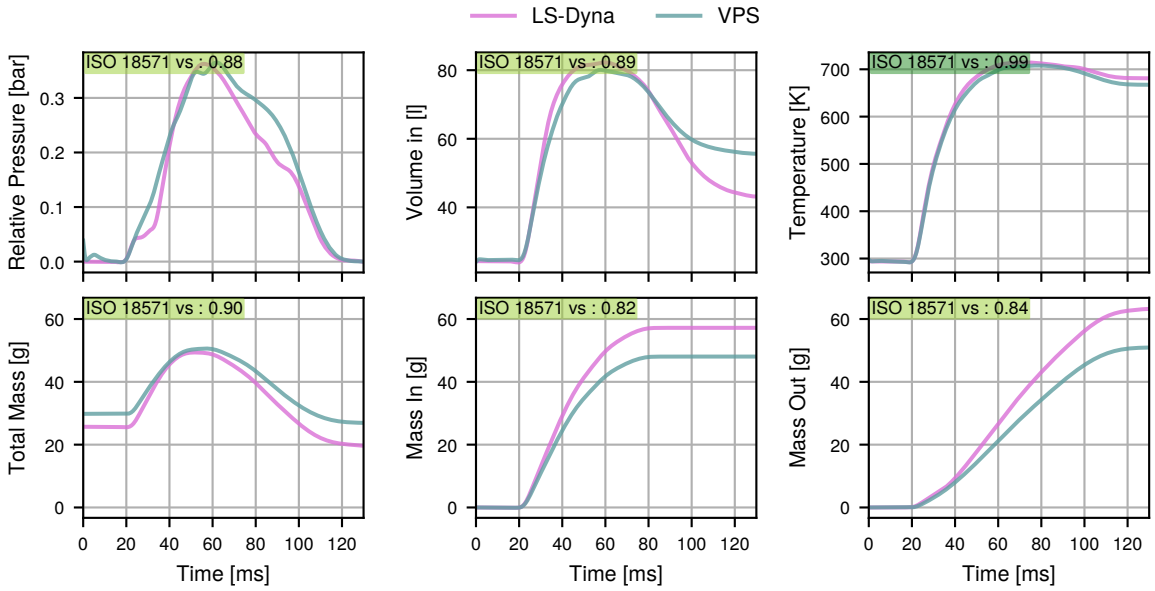


Figure 3.44: PAB Characteristics in Full Width Barrier (FWB) Load case with LS-DYNA versus VPS FE-model using HIII-50M and CFC 60

In figure 3.45 on the facing page the ATD kinematics at head, chest, and pelvis is displayed as acceleration and relative displacement in X- and Z-directions. The chest kinematics are well-matched. The head displacements are matching until 100 ms. From there on, in the rebound phase, the LS-DYNA HIII-50M seems to pull back the head as a horizontal motion, while the VPS one moves on a more circular path. The pelvic accelerations, especially in Z-direction, were higher. One reason can be a less restraining seat in the VPS FE-simulation. The LS-DYNA ATD experiences ca. 30 mm higher displacement, which results in a harder stop on the seat's structure.

In summary, the FE-model was successfully translated from LS-DYNA to VPS. Most differences between the FE-models in LS-DYNA and VPS could be reasoned. The remaining uncertainties are due to the different HIII FE-models and the limited insight of the project partners in the respective FE-model. Ultimately, the partners accepted the differences.

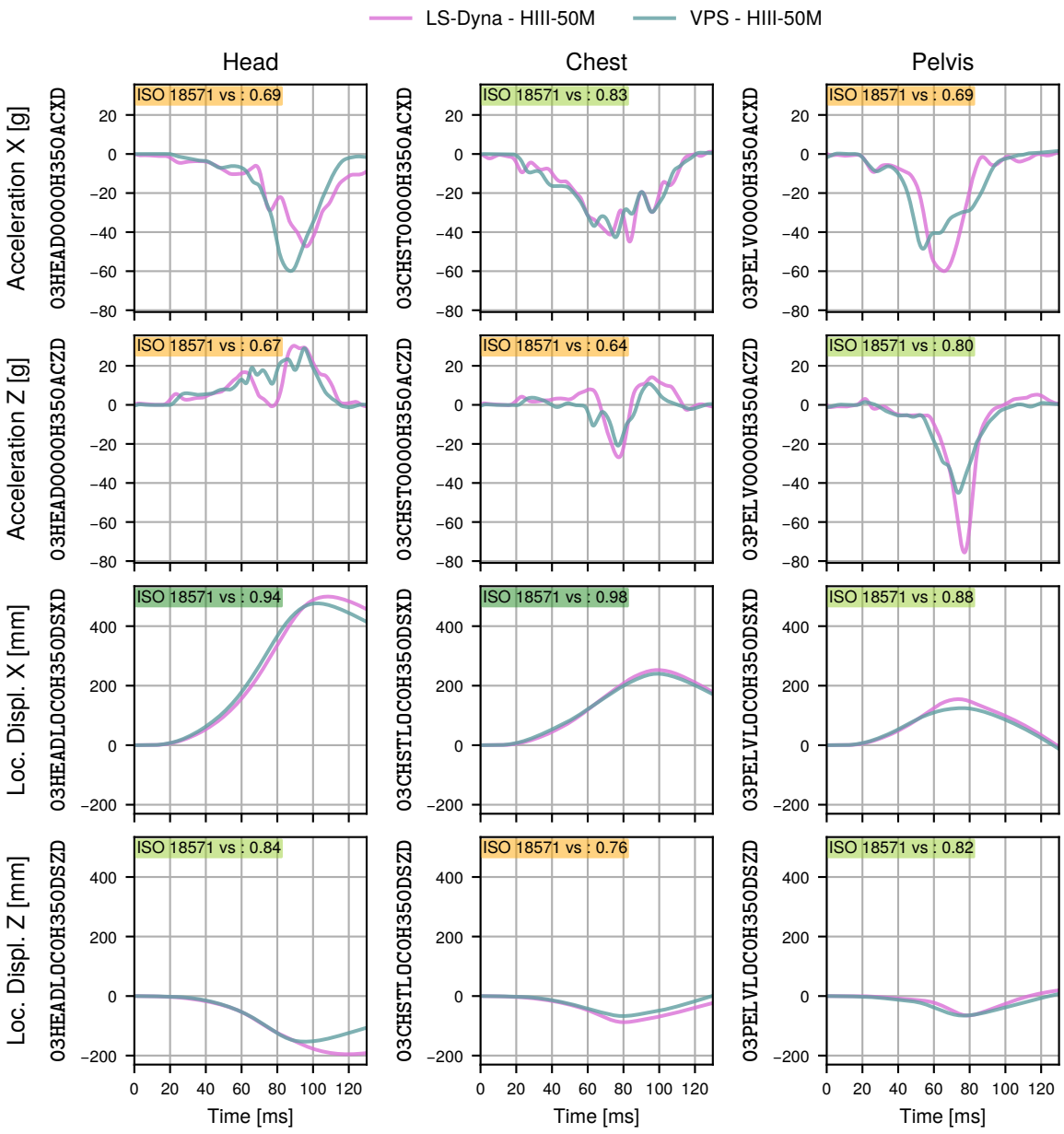


Figure 3.45: ATD Kinematics in Full Width Barrier (FWB) Load case with LS-DYNA versus VPS FE-model

3.3.3 Introduction of VIRTHUMAN to the VPS FE-model

The next step was the introduction of the HBM VHUM-50M. The kinematics responses are displayed in figure 3.46 on the next page. The heads' X-acceleration is overall similar, but the initial restraint seemed approximately 10 ms delayed. The Y-accelerations are similar. In agreement with the later head deceleration, the VHUM-50M local head-displacement is ca. 100 mm higher.

For the chest accelerations, similar relationships as for the head were observed; only the timings of the chest X-acceleration seemed better aligned. The pelvis accelerations in Z-direction are well aligned, but the VHUM-50M experiences ca. 25 g higher peak-deceleration. Additionally, the pelvis restraint appears to be effective for the HIII-50M significantly earlier. The differences in the local Y-displacement reflect the higher flexibility of the VHUM model regarding torsion of the thorax. The HIII-50M remains approximately within its sagittal plane during the forward motion in the straight FWB load case. In contrast, the VHUM-50M bends around the thorax seatbelt.

In figure 3.47 on page 128 the VIRTHUMAN 5th Percentile Female Human Body Model adapted to HIII-05F (HIII-05F) and VIRTHUMAN 95th Percentile Male Human Body Model adapted to HIII-95M (HIII-95M) were introduced. Similar to the three HIIIs, the pelvis accelerations for the VHUM-50M were highest, followed by the VHUM-95M. The tendencies for the local X-displacements are the same also, since the HBM mass is the primary cause here. Only small differences were observed in the head X-acceleration.

In summary, for the HIII-50M and the VHUM-50M, similar kinematics were observed. The most significant differences originate from the more flexible spine of the VHUM-50M. The tendencies between the VHUM-05F, VHUM-50M, and VHUM-95M were found similar to the tendencies between Hybrid III 5th Percentile Female Dummy (HIII-05F), HIII-50M, and HIII-95M.

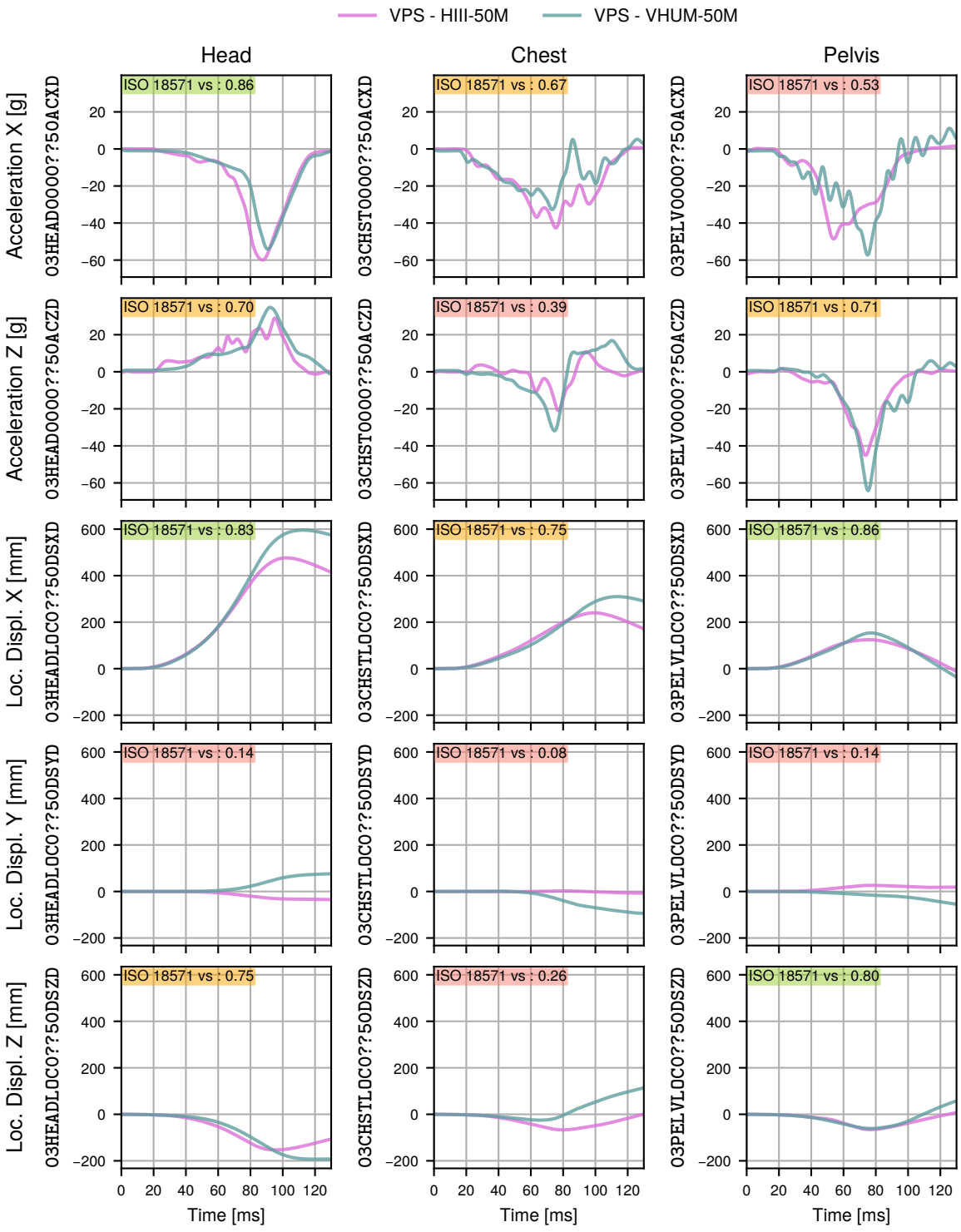


Figure 3.46: Kinematic Comparison in Full Width Barrier (FWB) Load case with HIII-50M versus VHUM-50M

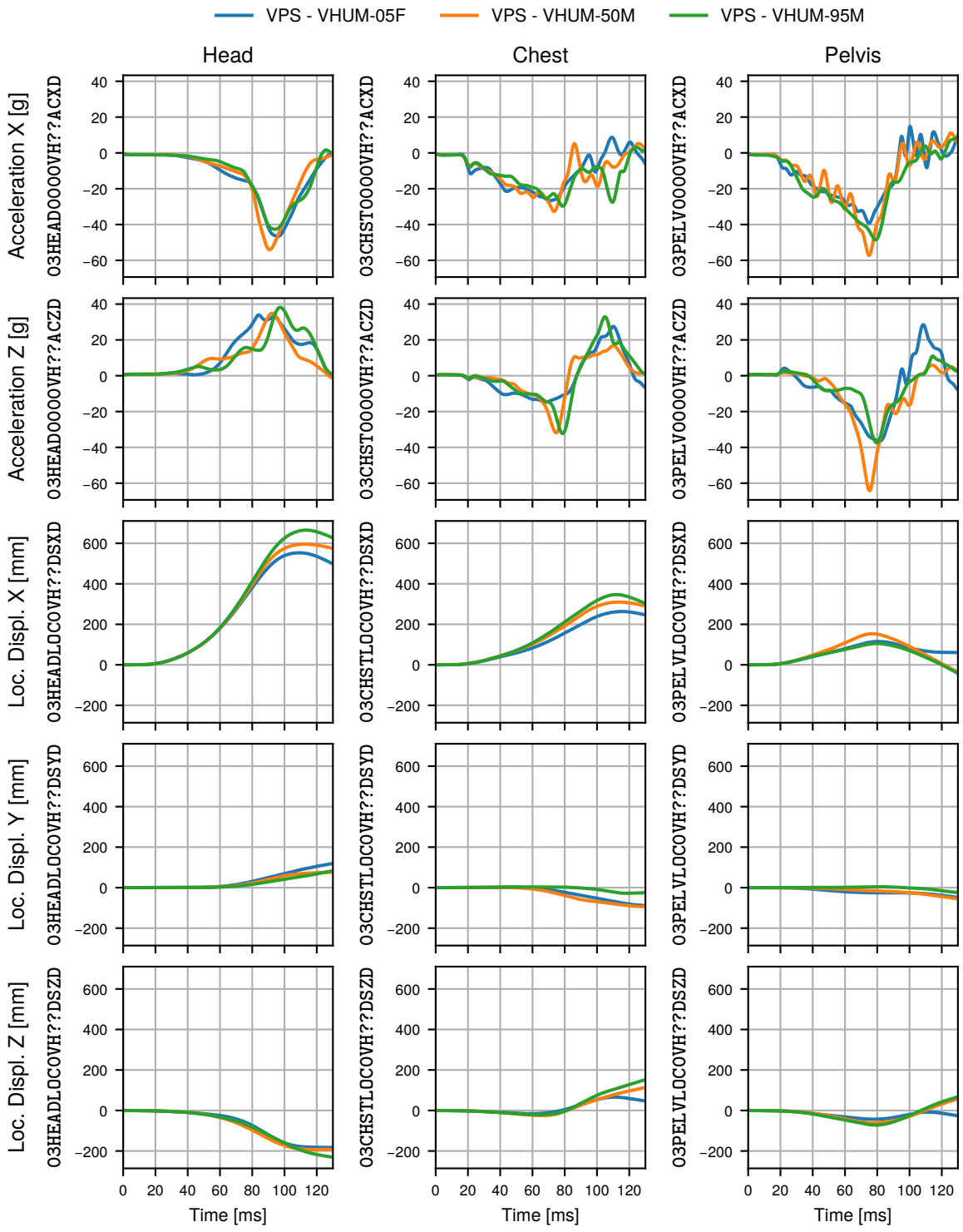


Figure 3.47: Kinematic Comparison of VHUM-05F, VHUM-50M, and VHUM-95M in Full Width Barrier (FWB) Load case

3.4 Database Creation

3.4.1 Factor Selection for the Designs of Experiments

For a Design of Experiments (DoE), from all options to vary the vehicle setup, it is important to select meaningful factors and their varied ranges. In the given FE-model, variations are possible but require different degrees of complexity. Varying the mass-flow of an airbag or the force levels of a constant of switchable load limiter can be achieved by a direct parameter. A variation of the shape and volume of an airbag would require a parametrized morphing technique.

The author's selection was guided by the expected influence of the variations on the Anthropomorphic Testing Device (ATD) response. The airbag shape is known to influence the kinematics of ATD. However, this influence is difficult to quantify. For that reason, the more direct potential factors were preferred. Furthermore, the variation of the restraint system was in focus. Other characteristics like the cushion stiffness of the seat may influence the ATD responses, but are a fundamental part of the ergonomics and comfort and therefore not a primary variable for the restraint system development unless the safety is not compromised.

The variations were limited to three fundamental areas of passive safety: the vehicle's crumble zone, the airbag, and the seatbelt system. From an ATD perspective, the characteristics of the crumble zone are defined as the vehicle-pulse with the 6 Degree of Freedoms (DoFs). A clear and direct influence can be expected, especially from scaling the prescribed translational X-motion and the rotation around the global Z-axis. As described above on page 114, a simple way of implementing such a variation would be via factors for the vehicle-pulse scaling and initial Z-rotation of the Body In White (BIW).

The seatbelt system as it exists in the FE-model has manifold potential factors like the material properties. Another pair is the time of tensioning and the force with which the seatbelt is pulled in. The next characteristics are the seatbelt load limiting level or the multiple limiting levels in case of a Switchable Load Limiter (SLL) with its switching timings. Ultimately, the maximum pull-out from the retractor can be considered.

An SLL is complex to balance with the frontal airbag and its effect depends on this balance. Hence, it was decided to limit the retractor capabilities to a single load level. Second, the activation time of the retractor is vehicle-pulse dependant. An irreversible system can only be activated, when the algorithms are sure about the crash event based on sensor data. Consequently, the time-to-fire is usually not part of the restraint system optimization loop and, therefore, not considered a factor in this work.

Because the webbing material and the maximum pull-out length are usually due to technical constraints defined, they were not considered either. Ultimately, the scaling value for the vehicle-pulse, the initial vehicle rotation, the force level limiting the seatbelt force, and the mass-flow scaling value for the inflator plus the opening time of the adaptive vent of the airbag were selected as factors.

The neutral setting of the five selected factors was the configuration of the developed and adjusted FE-model from see page 114 ff. The individual value-range of each factor was determined by experience and trial simulations. The SLL is commonly set at a level that leads to a shoulder seatbelt B3-force of 4.5 kN. The energy scaling of vehicle-pulse uses the original vehicle-pulse as the base. The rotation of the vehicle-pulse assumes a non-skew motion as the base and a skewness as the factor range, corresponding to the above considered Offset Deformable Barrier (ODB) load cases. The Passenger Airbag (PAB) mass-flow base setting is the result from the tuning of the restraint system above. The final factors with their naming and value ranges are summarized in table 3.4.

Table 3.4: Factors and Factor Ranges

Factor	Unit	Minimum (-1)	Middle (0)	Maximum (1)	Description
PABSCAL	1	0.9	1.0	1.1	Scaling of inflator in-mass-flow
PABVENT	s	0.07	0.1	0.13	Start of additional out-mass-flow
PSCAL	1	0.8	1.0	1.2	Scaling of vehicle-pulse
ALPHA	°	-10	0	10	Initial Z-rotation of vehicle-pulse
SLL	kN	2.14	2.34	2.54	Seatbelt force limiting level

To assess the individual influence of factors and to ensure a reasonable variation within the ATD response, each factor was altered over 3 levels while keeping the other factors constant. The results of the in total 33 FE-simulations are presented in figure 3.48 on the facing page. Each row represents another factor with its 3 levels on the X-axis. Each column contains a selected injury-criterion on the Y-axis, normed by a given reference. The individual effects of a factor in the nominal setting −1, 0 and 1 on a particular injury-criterion $\mathcal{I}$ were calculated, like defined by

$$\text{Effect} = \max\left[|\mathcal{I}(-1) - \mathcal{I}(0)|, |\mathcal{I}(0) - \mathcal{I}(1)|, |\frac{\mathcal{I}(-1) - \mathcal{I}(1)}{2}|\right], \tag{3.4}$$

as the maximum slope. Each line colour in the figure is associated with an ATD percentile, and the value behind the legend entry is the effect. The Head Acceleration over 3 ms Criterion (HAC_3) was selected to represent effect in the head, the Neck Injury Criterion (Nij) for the neck, the Chest Acceleration over 3 ms Criterion (CAC_3) and Chest Deflection Criterion (CDC) the chest, and the Femur Compression Force Criterion (FCC) effects in the femurs. The injury-criteria were normalized by the related biomechanical threshold value as indicated in the figure and in relation to table 3.3 on page 113 in section 3.2.

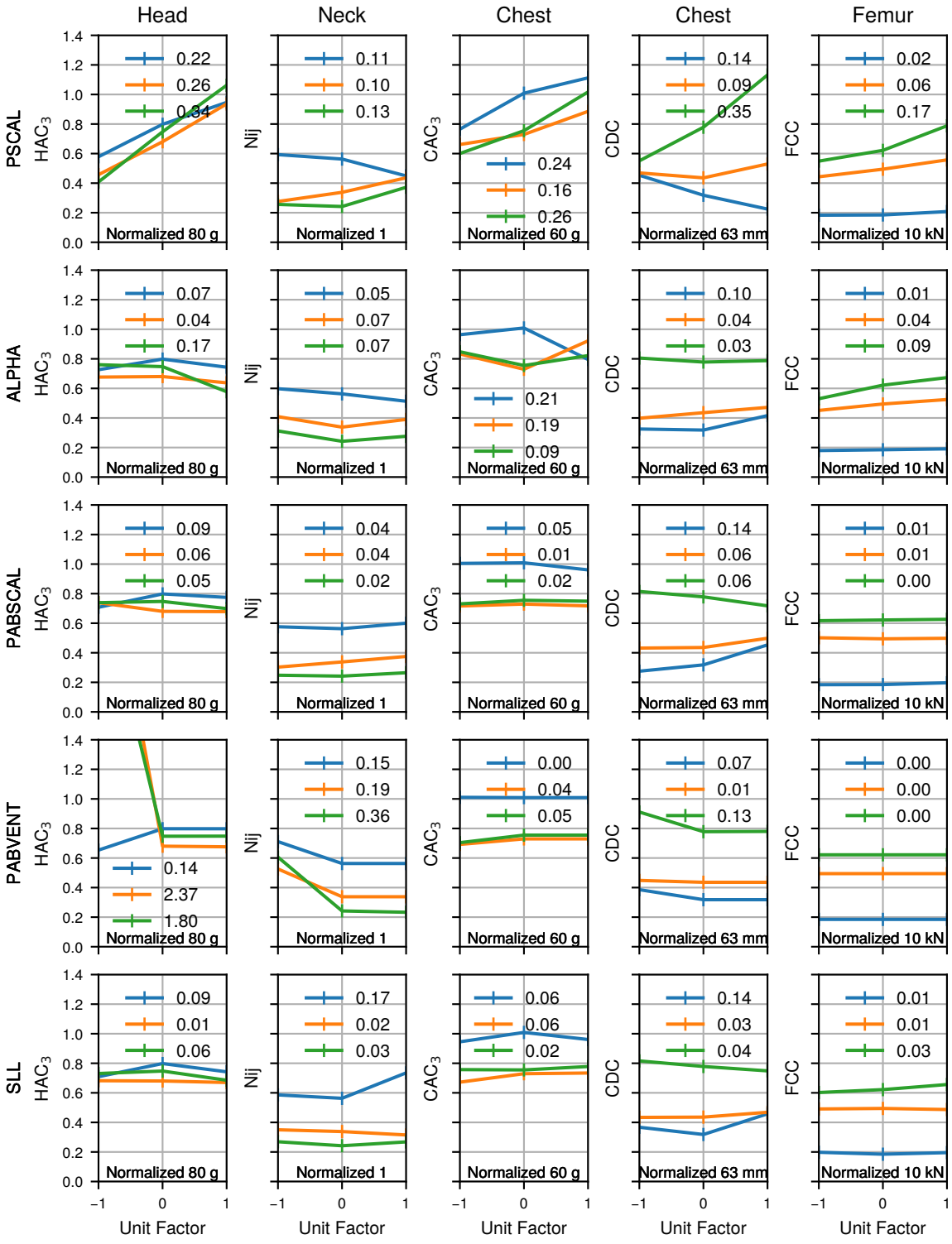


Figure 3.48: Factor Influences on Selected Injury-criteria with HIII-05F (Blue), HIII-50M (Orange), HIII-95M (Green)

In general, for all factors, significant changes within the responses of at least one ATD were observed. The highest effect had the scaling of the vehicle-pulse by the factor PSCAL. With higher energy in the system, the values of all injury-criteria increase. Smaller but clearer effects were observed for the initial rotation controlled by the factor ALPHA. For the early opening time of the adaptive vent controlled by the factor PABVENT, a hit through the PAB on the dashboard of the Hybrid III 50th Percentile Male Dummy (HIII-50M) and Hybrid III 95th Percentile Male Dummy (HIII-95M) were recorded. In addition, the CDC showed a clear effect. The SLL variation by the factor SLL showed a small effect, but no clear trend.

In summary, the selected factors had a clear influence on the inspected injury-criteria. The extreme edge cases were computed to check the stability. For the combination of a skew, high-energy vehicle-pulse with a soft restraint system, all simulations ran stable.

3.4.2 Data Generation by Designs of Experiments

After selecting the factors and determining their ranges, DoE had to be defined. In section 2.3 on page 59 ff were common methods introduced. A typical experiment generation method, which precisely covers the edge cases, is the discrete all-vs-all full factorial algorithm. A better coverage of the space can be achieved by the semi-randomly and continuous algorithm, by the Latin Hypercube approach and the Sobol algorithm.

The Sobol algorithm was selected for the main database because of

i the good coverage of the factor space,

ii the possibility to incrementally increase the factor value-density, and

iii the possibility to add new factors was not a priority.

Using the FE-model from section 3.2 on page 114 ff, the Sobol algorithm and the factors with their respective ranges defined in table 3.4 on page 130 resulted in 8192 experiments. Repeating each experiment per each of the three percentiles led to in total, 24 576 FE-simulations. This database will be hereafter referred to as the **development-set**. The resulting database was made available by Plaschkies (2024f). The size of the database was mainly limited by the available resources.

As an alternative approach to the Sobol algorithm, the full factorial algorithm was applied to generate an **alternative development-set**. The factor ranges were the same as before, just now each factor was treated as discrete instead of continuous. For 5 factors with 5 levels, 3125 experiments were generated and made available by Plaschkies (2024d). This alternative development-set was motivated by the better coverage of edge cases and by the possibility to have a comparison to the Sobol method.

For the **testing-set**, the FE-model version at the introduction of the Hybrid III (HIII) in section 3.2 on page 98 ff was selected from the simplification chain. The varied factors were the same as before, but their ranges broadened by 10 %. Due to the wider ranges, the extrapolation performance of metamodel can be assessed. As an exception, the range of the factor ALPHA

was kept the same. Since there were already instabilities such as the sliding seatbelt position observed for the skew cases, it did not seem beneficial to increase that lateral kinematics even more. By choosing a different FE-model than the metamodel was trained on, the whole process, including simplifications, can be evaluated. The testing-set can be smaller than the development-set. A number of experiments of 256 was selected. The resulting database was made available Plaschkies (2024e).

The last dataset was generated using the FE-model containing the VIRTHUMANs (VHUMs) from section 3.3 on page 121 ff. The factor ranges of this **transfer-set** were defined the same as in the development-set. Because in the FE-model comparison, the VIRTHUMAN 50th Percentile Male Human Body Model adapted to HIII-50M (HIII-50M) responses were roughly correlated with those of the HIII-50M, similar response ranges in the transfer-set and the development-set were expected. By the project partners who generated that last dataset, the full factorial approach was selected. The main motivation was the superior coverage of the edge cases. The number of experiments was limited by the available resources and oriented on the testing-set. Due to the specifics of the full factorial method, 243 experiments were generated. The resulting database was made available by Plaschkies and Müller (2024). Due to tuning of the FE-model for the Virtual Performance Solution (ESI) (VPS) environment, the numerical values of the factors differ from those of the alternative development-set. Yet, the 3 levels of the DoE correspond to the minimum-, maximum-, and mid-values of each factor of the alternative development-set.

After the data generation, the development-set had 8192 experiments, the testing-set contained 256 experiments, and the transfer-set 243 experiments. The alternative development-set featured 3125 experiments. Each experiment was repeated for the three percentiles of the respective human surrogate corresponding to the ATD sizes as FE-simulation.

3.4.3 Data Analysis and Summary Statistics

3.4.3.1 Applied Methodologies for Data Analysis and Cleaning

The generated FE-simulations were checked as a first step for an irregular termination type. For a few cases, restarting the FE-simulation with a different number of parallel processes was necessary due to light numerical instabilities. Ultimately, all FE-simulations except for 2 in the testing-set ran successfully. The two exceptions had both a very low value for the factor PABVENT. This early opening of the adaptive vent resulted in a collapse of the PAB while the ATD head was still in contact with the airbag fabric. However, both the FE-simulations aborted around 130 ms and the sensor signals seemed normal. Hence, the two cases were kept in the database and their missing 10 ms was set constant from the signals' end values.

Usually, relevant numerical instabilities within a particular FE-simulation reflect in the ATD responses; the result differs from the expectation. Thus, in a database, suspicious FE-simulations can be detected if the factor values in the DoE are dense. Consequently, screening the values of typical injury-criteria would give a rough idea. For the screening, the European New Car Assessment Programme (Euro NCAP) criteria defined in section 2.1 on page 34 ff and addi-

tionally the Chest Acceleration over 3 ms Criterion (CAC_3) were evaluated. The added chest acceleration was motivated by fitting to the evaluated channels during the simplification efforts.

Below, only the analysis of the development-set is presented. The steps and findings for the other sets were similar, and the detailed figures are provided as a digital supplement in the respective repository.

For each dataset, a letter plot as introduced in section 2.2 on page 2.2.5 was generated. There, the extreme statistical outliers determined were the Interquartile Range (IQR) method introduced in section 2.2 on page 55. The extreme outliers are indicated in red. In addition, the experiment IDs of the letter plot's fliers were provided. To be suspicious, the evaluated FE-simulation had to be an extreme outlier by the IQR method and the letter plot, and subjectively far from the next FE-simulation. For the suspicious FE-simulations, the related channels were checked in the context of the ten highest other FE-simulations. If a physically plausible reason could be identified, the FE-simulation was kept, representing the non-linearity of the system. Else, a numerical or model artefact was assumed as a cause, and the experiment was not further considered.

3.4.3.2 Analysis of the Development-set

In figure 3.49 on the facing page, the distributions for the development-set were plotted. For all injury-criteria, the extreme outliers are in the high-value range. One exception are the values of the CAC_3 for the Hybrid III 5^{th} Percentile Female Dummy (HIII-05F). Due to the occurrence of the seatbelt slipping of the seatbelt towards the neck in some cases, the seatbelt did not load the chest. The neck and head related criteria for the HIII-50M and HIII-95M are affected by the hits though the PAB on the dashboard. A high number of extreme outliers were identified by the IQR method. However, a hit-through cannot be considered a numerical effect or a modelling artefact. This interpretation is supported by the high density of the high values. Suspicious are only those FE-simulations, whose injury-criteria are with a distance from the others. The characteristics highlighted below were selected because they were significantly different to the peer group of FE-simulations.

From the **head acceleration** related injury-criteria, the Head Injury Criterion for up to 15 ms (HIC_{15}) and Head Acceleration over 3 ms Criterion (HAC_3), for the HIII-05F, the IDs 102, 600, 1388 and 5100 were identified as not plausible. Figure 3.50 on page 136 shows the 4 dropped signals in comparison to the normal FE-simulations. All affected FE-simulations had a sharp peak of the head accelerations in all directions around 100 ms. Due to its size, the HIII-05F does not reach the dashboard. Hence, a sharp peak is likely to be caused by a modelling artefact or instability.

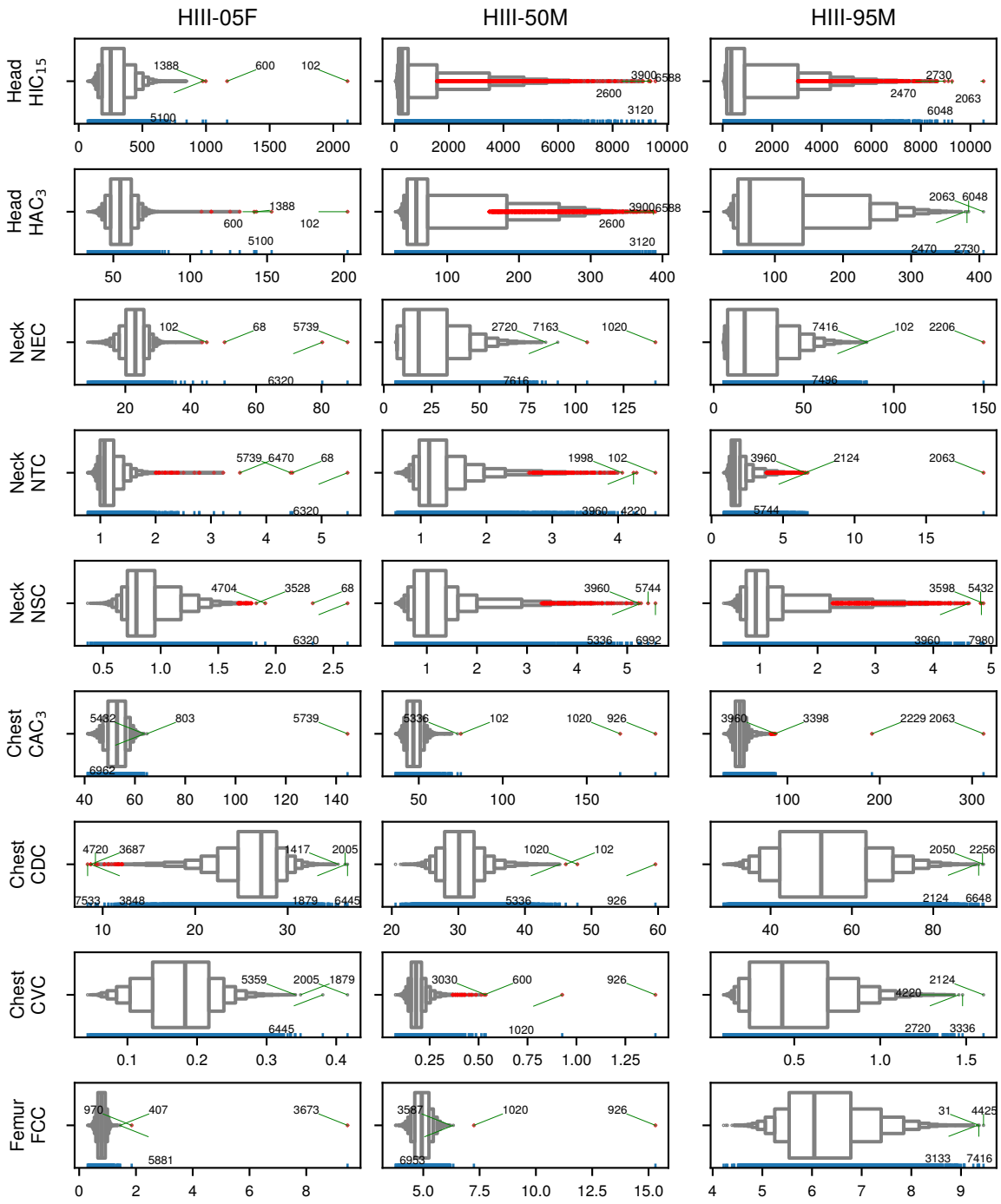


Figure 3.49: Letter-value-plots of Raw Data from Development-set with 8192 FE-simulations per Percentile with Extreme Outliers (red) and Suspicious IDs

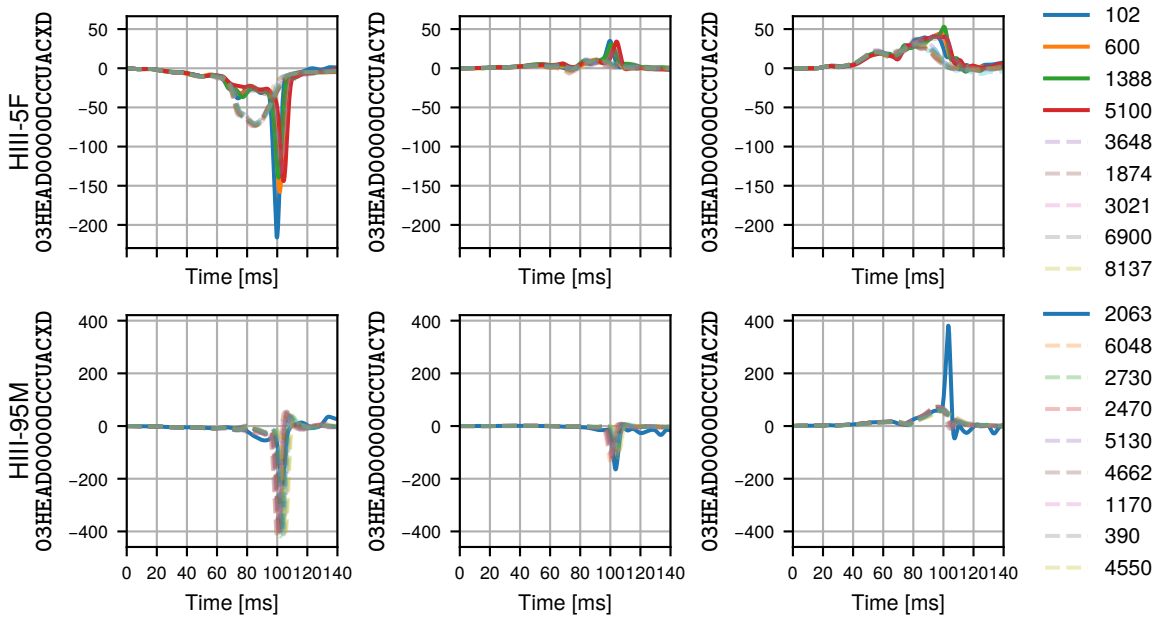


Figure 3.50: Head Acceleration Channels Related to Highest 20 Values of HIC_{15} and HAC_3 for HIII-05F and HIII-95M with Dropped IDs (continuous) and up to 8 Reference IDs (dashed)

As described above, both bigger ATDs suffered from the risk hitting through the PAB on the dashboard. Consequently, many FE-simulation reported head accelerations far above the biomechanical limits. This behaviour seemed plausible. The ID 2063 of the HIII-95M was an exception as presented in figure 3.50. In the other inspected FE-simulations, the maximum resultant acceleration was mainly influenced by the X-acceleration. In this particular FE-simulation on the other hand, the peak occurred in the Z-direction.

The **neck loads** were found to be for all HIIIs within a reasonable range. When interpreting the channel overview in figure 3.51 on the next page, it must be considered that for the injury-criteria not all load-directions are relevant (see section 2.1, page 34 ff). For the HIII-05F, the ID 5739 was dropped due to peaks early in the crash. Additionally, the IDs 68 and 6320 were neglected due to an unusual drop of the neck moment around 120 ms. For the HIII-95M, the ID 2206 was dropped because of an unusual increase in the neck extension moment after 130 ms. The ID 2063 of the HIII-95M was dropped too, due to very low Z-forces in the neck during the main restraint phase and a subsequent overly high increase of that force. This behaviour deviates significantly from all other cases and hints at a not properly working restraint system. No ID was dropped from the HIII-50M FE-simulations because the signals in figure 3.51 do not deviate significantly and are in alignment with the signals from the HIII-95M. Consequently, a hit through the airbag on the dashboard seems a valid assumption.

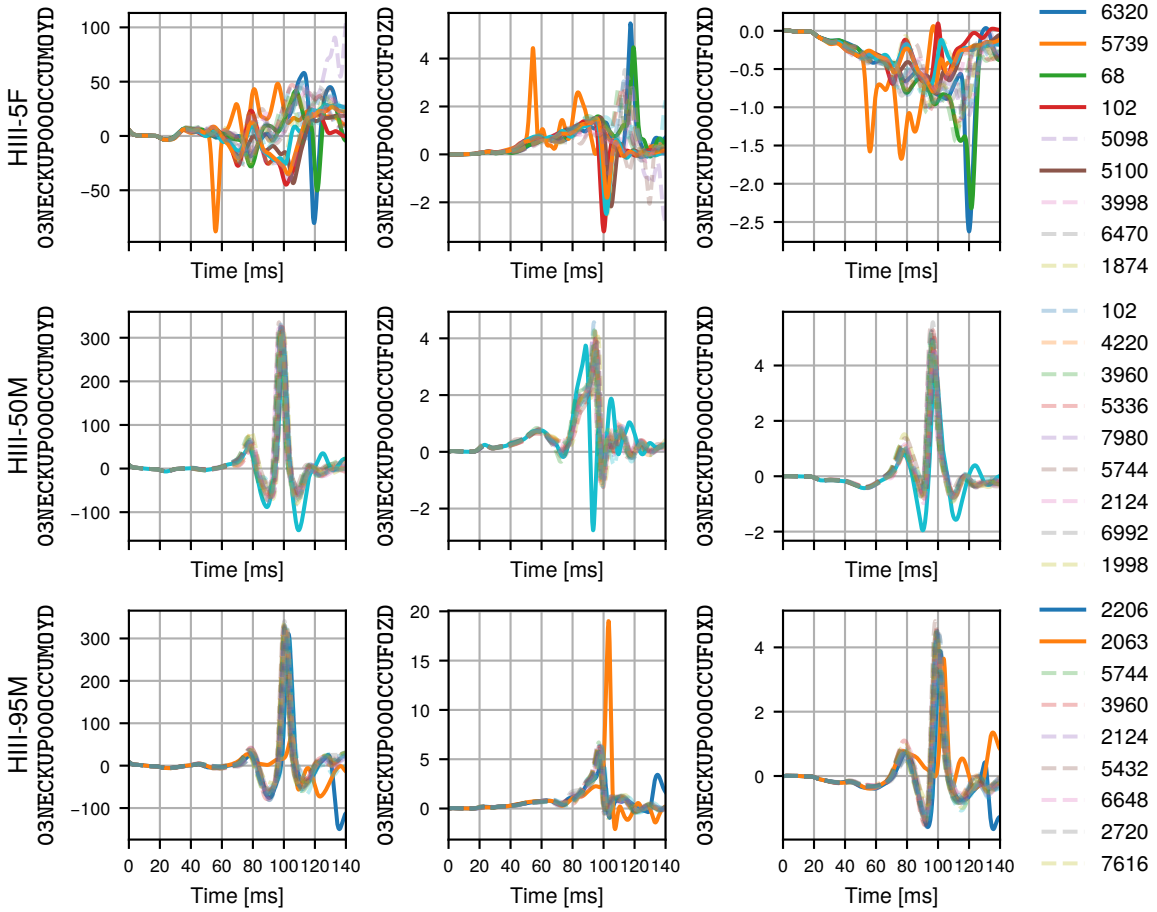


Figure 3.51: Neck Load Cell Channels Related to Highest 20 Values of NEC, NSC and NTC for HIII-05F, HIII-50M and HIII-95M with Dropped IDs (continuous) and up to 8 Reference IDs (dashed)

For the **chest acceleration**, the CAC_3 indicated the insufficient FE-simulations shown in figure 3.52 on the following page. For the HIII-05F, the ID 5739 was again identified due to a sharp peak around 50 ms. The IDs 926 and 1020 were dropped from the HIII-50M FE-simulations because of oscillations in the acceleration signals with a peak-to-peak amplitude over 150 g, being three times higher than in most FE-simulations. In both IDs, unusual oscillations in the chest deflection signal were observed too. Finally, the IDs 2063 and 2229 were dropped from the HIII-95M FE-simulations. The reasons were sharp and high peaks after 100 ms.

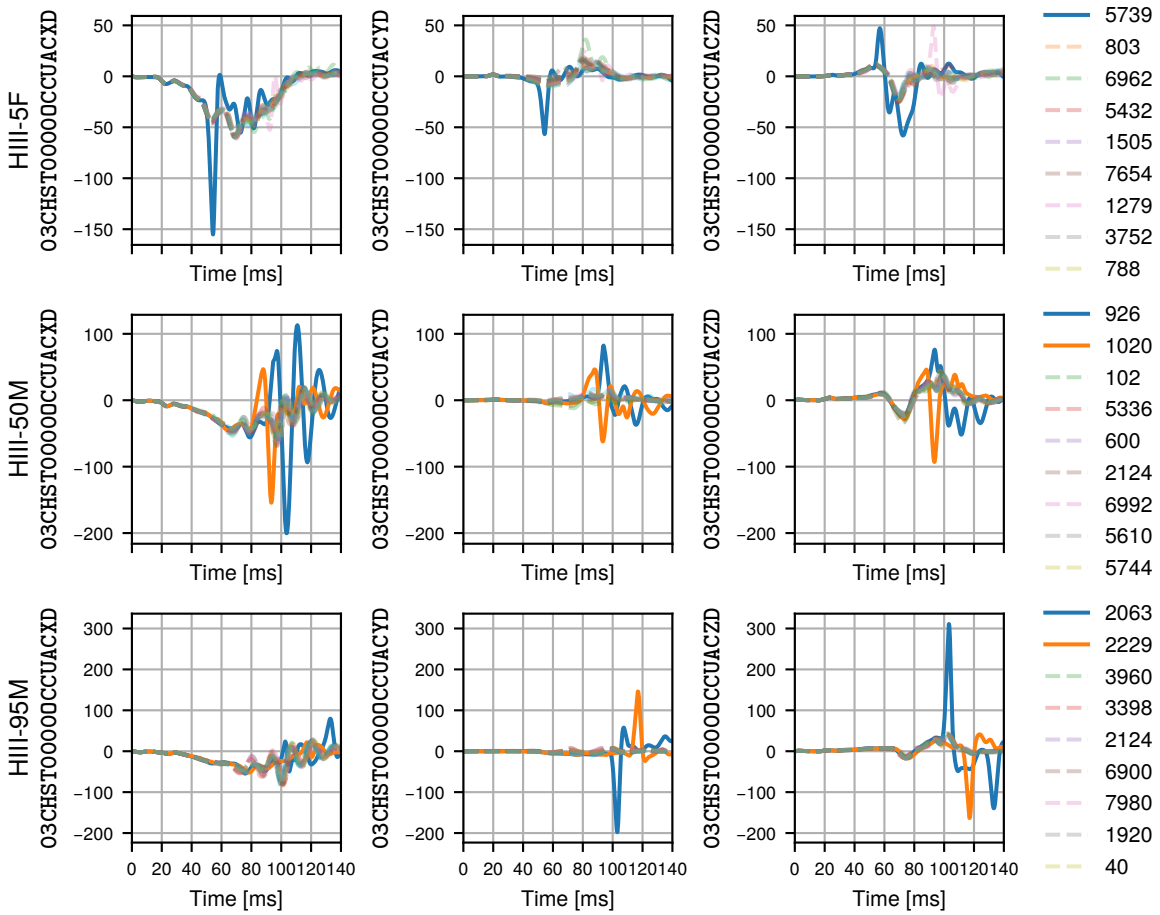


Figure 3.52: Chest Acceleration Channels Related to Highest 20 Values of CAC_3 for HIII-05F, HIII-50M and HIII-95M with Dropped IDs (continuous) and up to 8 Reference IDs (dashed)

The extremely low **chest deflection** of few HIII-05F FE-simulations due to the seatbelt slip had to be considered as a nonlinearity. However, there was a gap in the points around 10 mm. This gap motivated to drop the 6 IDs with a chest deflection lower than 10 mm. The IDs 3687, 3848, 4720, 5665, 5733, 6802 and 7533 were the affected FE-simulations. In figure 3.53 on the next page it can be observed that for those IDs the chest deflection signal does not increase after the coupling like it would be expected when the seatbelt is located near the sensor-point.

Due to the **femur loads** represented by the FCC, the ID 3673 was removed from the HIII-05F FE-simulations and the IDs 926 and 1020 from the HIII-50M FE-simulations. In the HIII-05F FE-simulation, a sharp negative peak in the left femur force at 115 ms was observed. Similar, but of the right femur were the observed signals of the HIII-50M. Figure 3.54 shows, how heavily the suspicious FE-simulations deviate from the others with high femur loads.

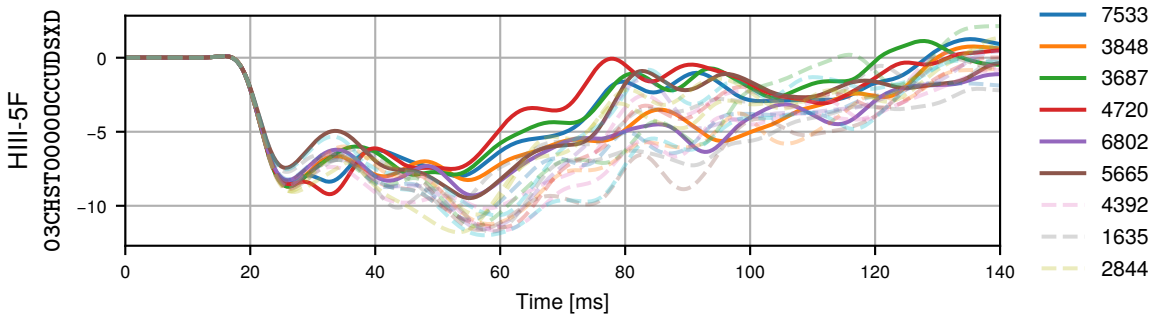


Figure 3.53: Chest Deflection Channel Related to Lowest 20 Values of CDC for HIII-05F with Dropped IDs (continuous) and up to 8 Reference IDs (dashed)

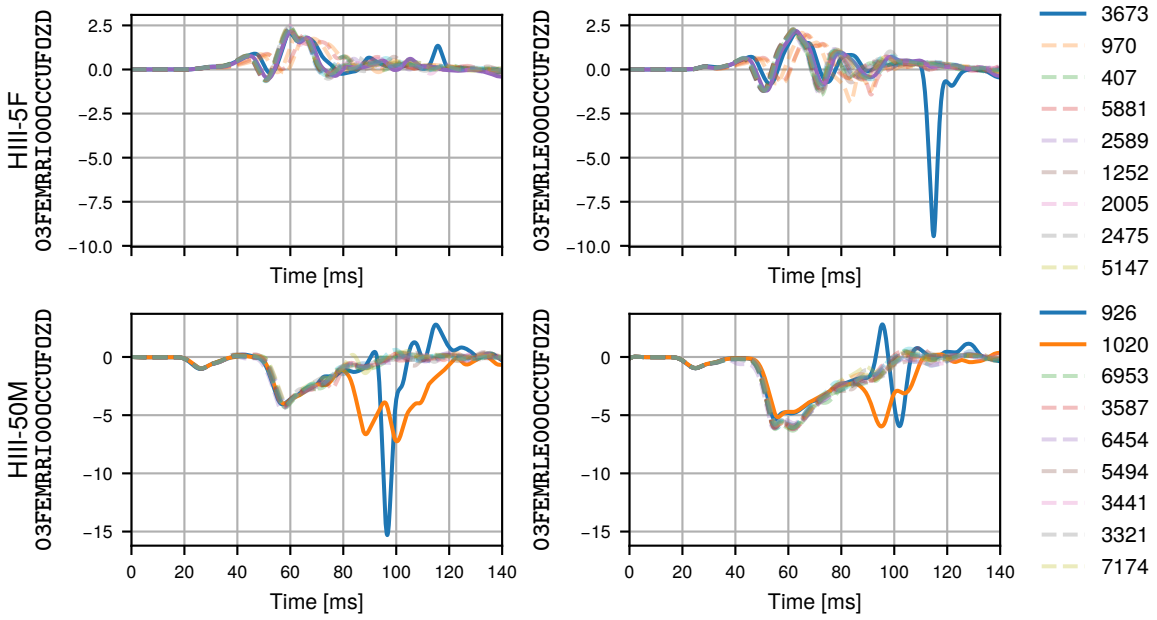


Figure 3.54: Femur Load Cell Channels Related to Highest 20 Values of FCC for HIII-05F, HIII-50M and HIII-95M with Dropped IDs (continuous) and up to 8 Reference IDs (dashed)

In total, 15 FE-simulations from the HIII-05F, 2 FE-simulations from the HIII-50M, and 3 FE-simulations from the HIII-95M were dropped. The drop rate is equal to a maximum drop fraction of 0.2 % per percentile and 0.1 % of the total database. In figure 3.55 is the final cleaned Development-set shown.

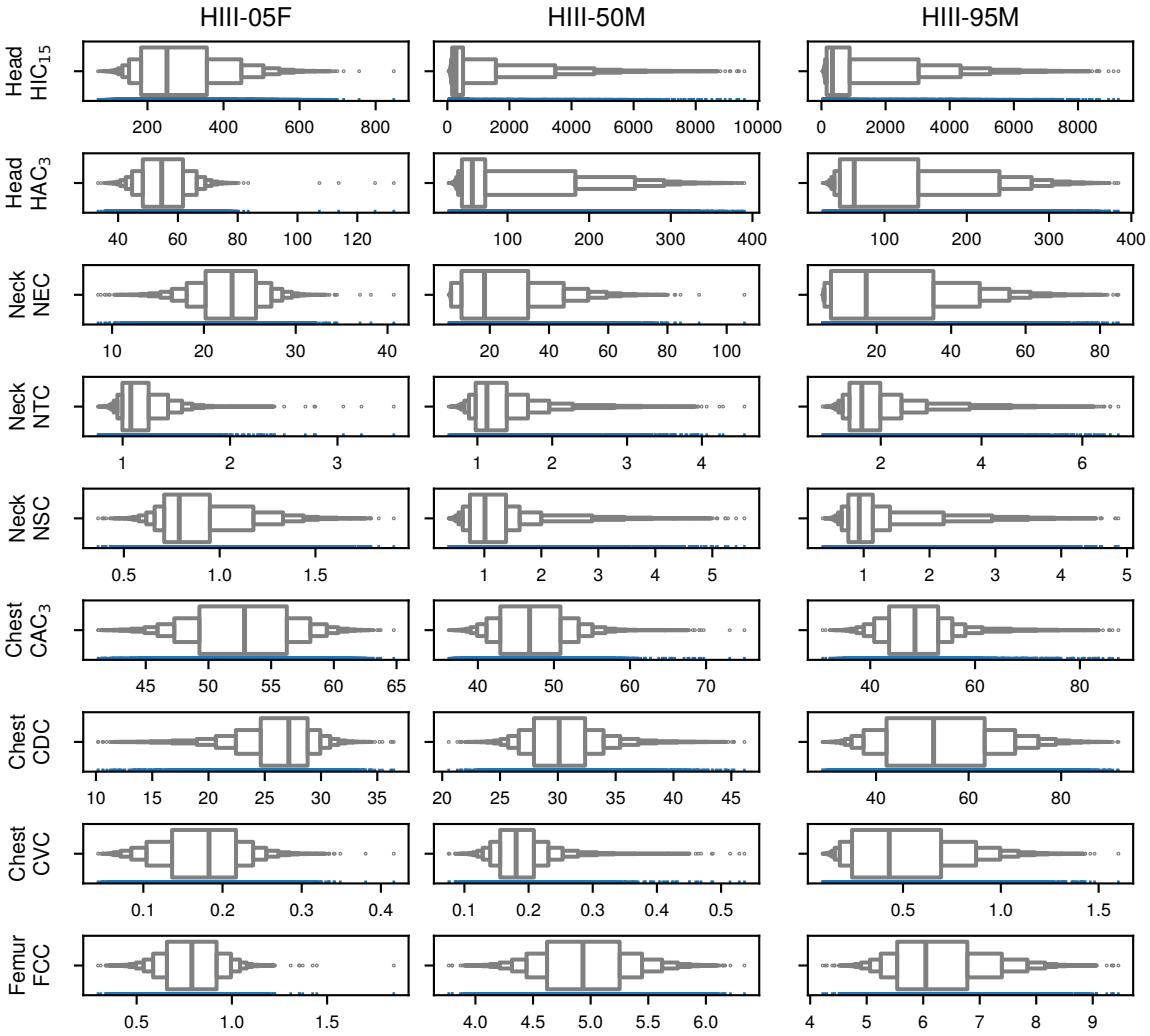


Figure 3.55: Cleaned Data as Injury-criteria from Development-set

3.4.3.3 Analysis of the Alternative Development-set

In the alternative development-set, 7 FE-simulations with the HIII-05F and 3 FE-simulations with the HIII-95M were dropped. The causes of the drop were similar to those mentioned above, the implausible peaks in the head accelerations and the very low values of CDC of the HIII-05F, and the sharp peaks in the HIII-95M chest acceleration after 100 ms.

During the data analysis, one major drawback of the full factorial method in comparison with the Sobol algorithm was confirmed. The factor PABVENT did not influence the HIII-05F

responses in the two higher levels. This insight is in accordance with the factor pre-study in figure 3.48 on page 131. FE-simulations, whose only difference was in the factor PABVENT in the two upper levels, were found identical. The literature calls this pseudo-repetition (compare Siebertz et al., 2017).

3.4.3.4 Analysis of the Test-set

In the testing-set, 15 FE-simulations with the HIII-05F, 8 FE-simulations with the HIII-50M, and 2 FE-simulations with the HIII-95M had to be removed. Again, the reasons were similar as above.

In figure 3.56 on the following page are the data distributions of the development-set and the testing-set compared. The Y-axis is shared for each injury-criterion over the three percentiles. The shown value ranges of each criterion were determined by the minimum and maximum values from the testing-set. The full data of the development-set was shown in figure 3.55 on the preceding page. The factor ranges were broader in the testing-set. Furthermore, lateral protection was part of the testing-set FE-model by the door and Curtain Airbag (CAB). Consequently, less overly high head acceleration and a broader spread of the system's responses were expected. On page 111 were the two FE-models directly compared. The analysed differences are expected to be reflected in the median of both datasets.

The first look at figure 3.56 on the following page confirms the expectation of HIC_{15} and HAC_3 in the testing-set. The median for all HIII percentiles for the HAC_3 were found to be approximately $10\,g$ higher in the testing-set than in the development-set. The accelerations were found in the direct single comparison slightly higher, too. The stronger difference can be due to the adjusted restraint system. Across the three ATD sizes, the head, and neck loads were found to be increasing with the ATD size.

The values of Neck Extension Moment Criterion (NEC) were similar for the HIII-05F and HIII-50M between the two datasets. The HIII-95M NEC values were in median ca. 10 kN higher for the testing-set. A reason can be the more effective lateral restraint, leading to a stronger back bending of the ATD head. The values of Neck Tension Force Criterion (NTC) were found to be 1 kN to 3 kN higher in the testing-set for all ATDs. In contrast, the Neck Shear Force Criterion (NSC) values were quite similar. The higher extreme values in the development-set are height and cause in agreement with the head acceleration.

The CAC_3 values were within the expected result. The spread in the testing-set was a bit wider, while the median remained similar. The direct FE-model comparison indicated stronger differences, yet not with a clear tendency. The CDC value distribution did not vary significantly for the HIII-05F in the testing-set. The higher median hints at a more stable seatbelt location. The Chest Viscous Criterion (CVC) values were comparable between the databases.

The approximately 2 kN lower values of FCC of the simplified FE-model were reflected in the development-set for the HIII-50M, too. The different forces were mainly reasoned by the different stiffness distribution of the dashboards of both FE-model versions. With the increased masses and upper leg lengths of the ATDs the femur forces increase.

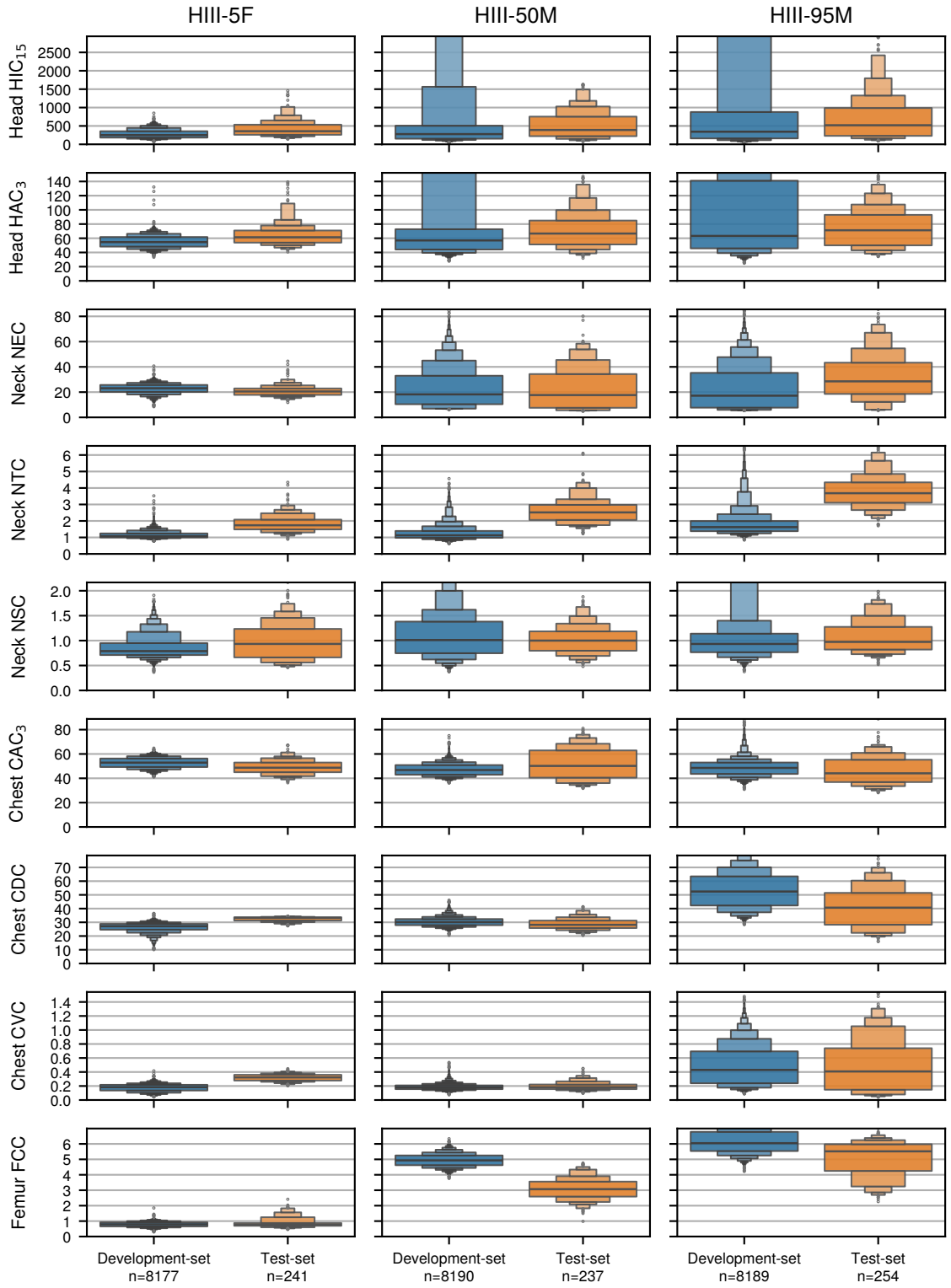


Figure 3.56: Data Distribution of Development-set vs Testing-set for Selected Injury-criteria

3.4.3.5 Analysis of the Transfer-set

In the transfer-set, 7 FE-simulations with the VIRTHUMAN 5th Percentile Female Human Body Model adapted to HIII-05F (HIII-05F) had to be removed. The reason was, again, very sharp peaks in the head acceleration after 100 ms.

As a reason for the sharp peaks, the hit of the head on the dashboard was identified. While the kinematics, shown in figure 3.57, does not appear to be completely implausible, it can be classified as a modelling artefact. No real restraint system would be accepted, if the smallest Human Body Model (HBM) in the tested factor range hits the dashboard with its head. For that reason and the dropped FE-simulations being extreme outliers, the removal seemed justified. The effect was observed for high values of the factor PSCAL in combination with a low value of the factor SLL.

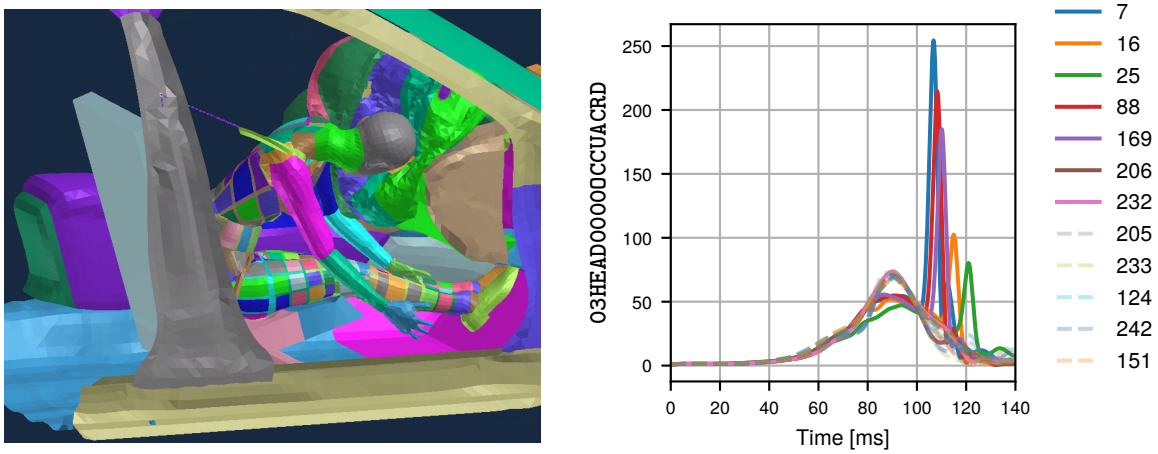


Figure 3.57: VHUM-05F Kinematics with High-energy Vehicle-pulse and Low Seatbelt Restraint (left) and Head Acceleration (right) Related to Highest 20 Values of HIC_{15} and HAC_3 with Dropped IDs (continuous) and 5 Reference IDs (dashed)

In the figure 3.58 on the following page, the results from the alternative development-set with the HIII family and the results from the transfer-set with the VHUM family are shown. The expected general differences were discussed on page 121. Because not all factors contribute equally to the human surrogate response, the experiments were matched. In the alternative development-set, the factors were varied in 5 levels. In the transfer-set on the other hand, only 3 factors were used. For the figure 3.58 on the next page, only those experiments of the alternative development-set were used, which have an equivalent experiment in the transfer-set.

The head acceleration related criteria HIC_{15} and HAC_3 were in a similar range in both databases. In the transfer-set, the hits on the dashboard seem to cause less high peaks.

The neck responses NTC, NEC, NSC of the 5th percentiles were in a similar range. For the bigger VHUMs, no clear trend was observed. Due to the more flexible thorax of the VHUM, differences, especially in the FE-simulations with higher lateral vehicle-pulse, were expectable.

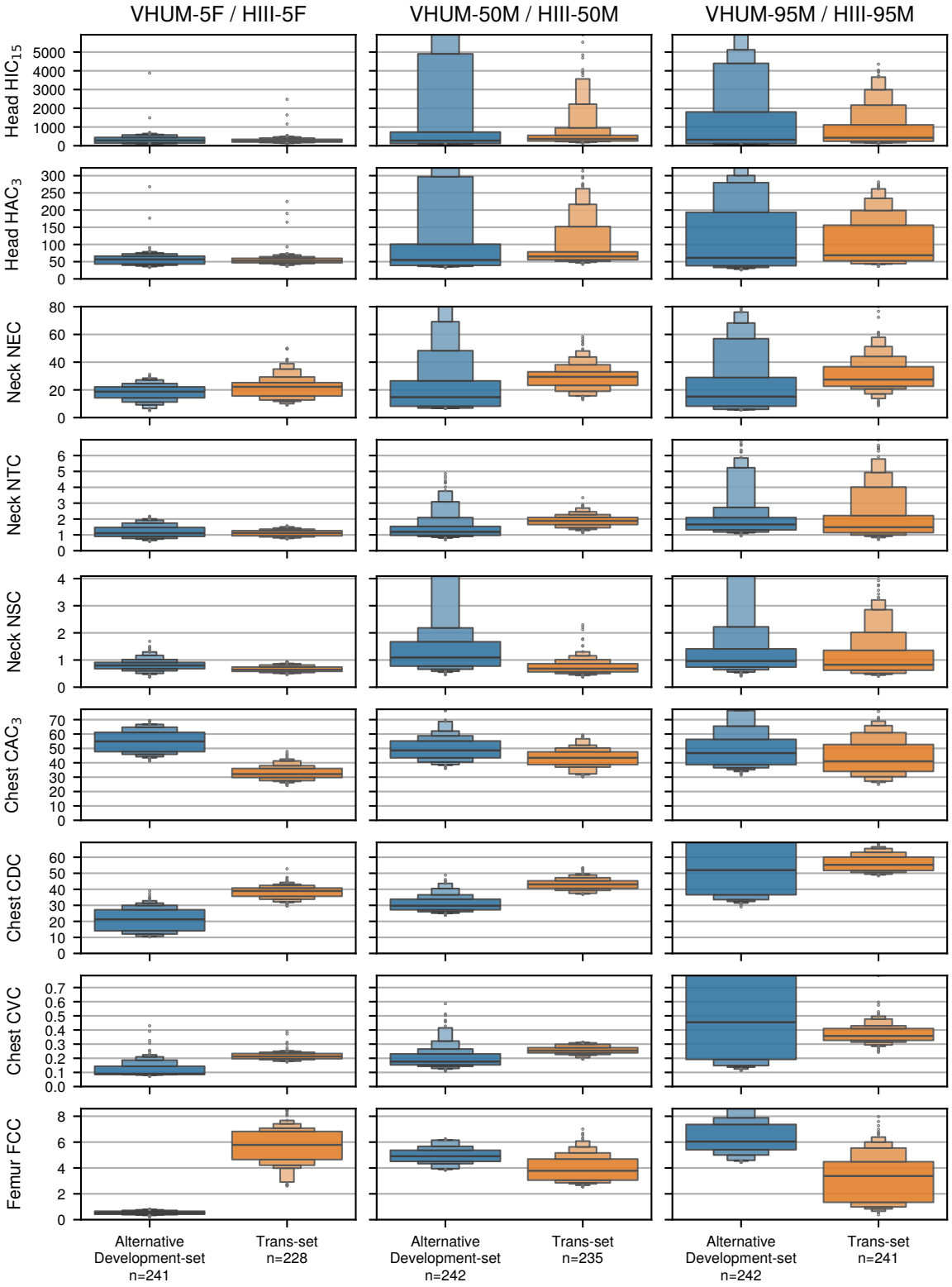


Figure 3.58: Reduced Alternative Development-set vs Transfer-set for Selected Injury-criteria

The CAC_3 values were lower for all HBM sizes in the transfer-set. For the CDC and CVC values, the opposite was the case. Again, the higher flexibility of the thorax would lead to higher compression, too. The tendency of the CDC values to increase with the VHUM size was observable in both datasets.

The FCC values are strongly dependent on the upper leg length and the location of the contact between the knees and the dashboard. Consequently, the results for the HIII and VHUM FE-models differ significantly. The HIII FCC values increase with the VHUMs size. The VHUM femur loads, on the contrary, decrease.

3.5 Summary and Discussion on the Data Generation

The above described procedure can be interpreted as a methodology to simplify FE-model reasonably. The key steps were (i) defining the targeted level of quality of the FE-model and its acceptable computational cost, (ii) selecting the initial FE-model, (iii) analysing the initial FE-model regarding computational cost and numerical stability, and (iv) applying incremental simplifications to the FE-model in a pre-defined manner until the targets are met.

The database generation was conducted by executing the steps by starting (i) defining the targeted size and complexity, then (ii) selecting the factors and their respective ranges, while (iii) ensuring the factors generate variance in the data and the FE-model runs stable, ultimately (iv) defining the Design of Experiments (DoE), and (v) conducting the FE-simulations. In the final step, (vi) the database is cleaned and analysed.

The selection of the initial FE-model, or even the preceding step of selecting the appropriate simulation method, imposes indeterminable risks. In particular, the simplification efforts are prone to become significantly more complex than anticipated. In the incremental approach, each step results in a working FE-model version. Changes in the order or strategy of the measure due to unforeseen challenges can be implemented in the chosen approach.

The biggest advantage of the incremental approach is the trackability of a single measure's impact. The utilized ISO-rating provided a valuable overview of the areas of interest. However, in future use, the ISO-rating should be calibrated to the particular use-case. In this study, the default settings were used. Signals of low magnitude like such from the kinematics in Y-direction seem to rate overly strict. At the same time, key characteristics for the injury assessment like maxima were obfuscated by the averaging nature of the ISO-rating; this holds true especially for the selected threshold of 0.8 associated with the ISO-grade Good. A calibration of the ISO-rating parameters and weights can take such preferences into account.

In the context of the signal comparison, the chosen filter class was relevant. Due to artefact effects of the FE-model, the Channel Frequency Class (CFC) 60 was chosen. This seemed justified, since for this study's objective, the overall tendency was considered more important than the analysis of minor effects.

To compare and contextualize the vehicle-pulses, the Occupant Load Criterion (OLC) was used. The default parameters for the Anthropomorphic Testing Device (ATD) displacements during the unrestrained phase and the restrained phase were used unchanged. However, these settings were developed for a particular vehicle. Rabus et al. (2022) proposed alternative approaches of adjusting the original methodology to the actual vehicle. A second well-known limitation of the OLC is the modelling of the ATD kinematics only within its sagittal plane. Consequently, for vehicle-pulses with signifiant lateral motion, the OLC is not meaningful. Although these limitations should be known when interpreting the OLC, they seemed to be acceptable for the insights sought in this research.

Two characteristics affected the quality of the database – (i) unwanted seatbelt sliding towards the neck, and (ii) relatively low airbag pressure caused the partially very low values of Chest Deflection Criterion (CDC) for the Hybrid III 5th Percentile Female Dummy (HIII-05F) and high acceleration peaks, especially for the Hybrid III 95th Percentile Male Dummy (HIII-95M). Both effects would not be accepted for a developed restrained system. However, they are physically reasonable and can appear during the development process. In future use of the discussed approaches, only FE-simulations with acceptable restraint performance should be used in the database. Alternatively, if such cases should be kept, strategies must be developed, to predict such effects.

The general seatbelt routing in the FE-model used for the databases can be improved. Within the context of the presented research, no negative effect on the FE-simulation results were observed. However, from the system development perspective, the seatbelt slack should be reduced. The FE-simulations with VIRTHUMAN (VHUM) for the transfer-set were correlated with the original FE-model with Hybrid III (HIII) and Livermore Software technology corporation DYNA (LS-DYNA). Unfortunately, the limited insight in the respective FE-models of the project partners reduced the depth of analysis.

The analysis and cleaning of the database seemed suitable. However, the high-level approach of screening the values of the injury-criteria and analysing the output of related sensor, only for the suspicious case, carries the risk of overseeing numerical artefacts. In the future work, a more elaborate approach should be developed. For example, Wu and Hartmann (2020) proposed an approach to assess the plausibility by storing an updated mean value and the standard deviation of each response channel. If a new simulation deviates significantly, the system sends a warning. Recently, Kracker (2024) proposed a process of generalizing the finite element mesh as an intermediate step as a base for higher processing. This generalized representation allowed the comparison of inner energies and the deformation mode over changing vehicle geometries. Furthermore, the global energies should be monitored since they can indicate numerical artefacts, too.

In summary, using the presented method, two development-sets, one testing-set, and one transfer-set were generated for further use. During the cumbersome process, the methodology proved usable and stable. The generated data was of acceptable quality and contained reasonable variance for the subsequent machine learning tasks. The discussed limitations should be considered if the approach is transferred to industrial use.

4 Estimation of Different Anthropometrics

In this chapter, different approaches for generating a suitable metamodel are studied. The core scenario is the estimation of results from a FE-simulation using the Hybrid III 50th Percentile Male Dummy (HIII-50M) as input and the HIII-05F or HIII-95M as output. Different architectures are discussed and complemented by strategies to control the size of the database. Furthermore, the transfer between ATDs and Human Body Models (HBMs) in the form of VHUM is presented. The source code was published by Plaschkies (2024a), and the generated metamodels by Plaschkies (2024b).

4.1 Description of the Used Architecture & Framework

The basic architecture was outlined in the initial problem statement of section 1.2 in figure 1.3 on page 4 and is based on the idea of characterising a vehicle in all its properties implicitly. As a result from the chapter 3, the Hybrid III (HIII) is used. Consequently, the Hybrid III 50th Percentile Male Dummy (HIII-50M) was used to assess a new environment. Its output, determined by vehicle characteristics, restraint system, and load case, is used as the metamodel input space. The Hybrid III 5th Percentile Female Dummy (HIII-05F) or rather the Hybrid III 95th Percentile Male Dummy (HIII-95M) was used as the metamodel output space. The metamodel itself is a mapping between these two spaces.

The selection of the predicted output oriented on the defined injury-criteria in Euro NCAP (2023b) like presented in section 2.1 on page 34 ff. Following injury-criteria were selected:

- the Head Injury Criterion for up to 15 ms (HIC_{15}),
- the Head Acceleration over 3 ms Criterion (HAC_3),
- the Neck Extension Moment Criterion (NEC),
- the Neck Tension Force Criterion (NTC),
- the Neck Shear Force Criterion (NSC),
- the Chest Deflection Criterion (CDC),
- the Chest Viscous Criterion (CVC), and
- the Femur Compression Force Criterion (FCC).

For classification tasks, the limits from table 2.6 on page 36 were used. To generate the multiclass classification target attributes, first the limits of the injury-criteria were linear interpolated for 1 to 4 points. Additionally, the threshold to reach 1 point, respectively 4 points were buffered with 10 %, 20 %, 30 %, 40 % and 50 %. The point-buffer-related intervals used to derive the classes are presented in table 4.1. For the binary case, the threshold for the HIII-05F is 4 points because this Anthropomorphic Testing Device (ATD) is used in the tests defined by legal entities and consumer-oriented organizations. The HIII-95M is used primarily in a due-care context, and hence the biomechanical limit seemed more appropriate. The more granular multiclass thresholds were built on this logic, but are tuned for better distribution. Unfortunately, for some injury-criteria, only a very skew distribution was achievable as indicated in section 4.2, table 4.3 on page 152. The selected numbers of classes 2, 3, 5 and 7 were based on practical considerations with respect to typical interests in the development process of a restraint system. Furthermore, it is in agreement with the definitions by Plaschkies et al. (2023) for the application of Multi-Criteria Decision Making (MCDM).

Table 4.1: Intervals for Injury-criteria Classification Related to Euro NCAP Points

Classes	ATD	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
2	HIII-05F	[0, 4.0)	[4.0, ∞)					
	HIII-95M	[0, 1.0)	[1.0, ∞)					
3	HIII-05F	[0, 4.0)	[4.0, 4.3)	[4.3, ∞)				
	HIII-95M	[0, 0.8)	[0.8, 4.0)	[4.0, ∞)				
5	HIII-05F	[0, 1.0)	[1.0, 4.0)	[4.0, 4.2)	[4.2, 4.4)	[4.4, ∞)		
	HIII-95M	[0, 0.8)	[0.8, 1.0)	[1.0, 4.0)	[4.0, 4.3)	[4.3, ∞)		
7	HIII-05F	[0, 1.0)	[1.0, 2.0)	[2.0, 3.0)	[3.0, 4.0)	[4.0, 4.2)	[4.2, 4.4)	[4.4, ∞)
	HIII-95M	[0, 0.8)	[0.8, 1.0)	[1.0, 2.0)	[2.0, 3.0)	[3.0, 4.0)	[4.0, 4.3)	[4.3, ∞)

Additionally, the Chest Acceleration over 3 ms Criterion (CAC_3) was added for regression tasks, as it is a central injury-criterion in many other testing protocols. For relevant sensor signals, the corresponding channels were identified.

As features, the injury-criteria and channels similar to the target were selected. In addition, the displacement of the head's, chest's, and pelvis' Centres of Gravity (CoGs) relative to the rigid body motion of the Body In White (BIW) was added, as they characterize together the overall kinematics of the ATD. Furthermore, injury-criteria from legislation and other consumer protocols, such as the Neck Flexion Criterion (NFC) and Neck Injury Criterion (Nij), were considered as well. The injury-criteria are feature attributes, and the channels feature multivariate time-series.

In section 2.2 on page 45, a framework was presented. This framework was adapted for this thesis. The evaluation metric-functions as defined in section 2.2 on page 43 ff are reported as the median of the 5-folds for the training-sets and validation-sets respectively and for each target attribute. To enable the utilization of extended evaluation methods like presented on page 46, all metric-functions should report in the interval [0 , 1], where 1 is best. For target

attributes regression tasks, the to 0 capped R2-score was used. For target attributes classification tasks, the F1-score was chosen. For the multiclass case, the score for each class is calculated individually and combined as an average weighted by the class occurrence. For the target multivariate time-series regression task, the in chapter 3 on page 69 applied ISO-rating rating was used.

For adaptive querying strategies, an evaluation of the performance of an individual instance would be necessary. However, the above described metric-functions evaluate the score as a statistical value related to the evaluated set. Apart from the ISO-rating, they are not defined for a single instance. Hence, as the loss-function for classification and regression tasks, for a pair $\hat{\boldsymbol{\mathcal{Y}}}$ and $\boldsymbol{\mathcal{Y}}$ of a single instance ς and single attribute τ, the absolute error in the form

$$^{\text{Absolute Error}}\mathcal{L}_{\varsigma,\tau} = \left|\hat{\mathcal{Y}}_{\varsigma,\tau} - \mathcal{Y}_{\varsigma,\tau}\right| \tag{4.1}$$

was selected. Hence, $\boldsymbol{\mathcal{L}}$ reports in the interval $[0\,,\infty)$, where 0 is best. To achieve the same direction of optimisation for ISO-rating, the ISO-rating value is subtracted from 1. The target attributes are normalized to achieve a weighted arithmetic mean similar to the calculation of the cost-function.

Raw data is required within the framework to be stored in the PARQUET format as introduced in section 2.2 on page 41, carrying the human surrogate percentile and the experiment ID as indices of tabular data. For temporal data, the timestep as an additional index is needed. The multivariate time-series are within the framework referred as temporal data and pre-calculated values as tabular data. The PARQUET format seemed preferable due to the fast assessing time and compressed storage. Predictions are stored in the same format with the same motivation.

The pipeline configuration, evaluated key metric-functions, and other diagnostic and book-keeping data are stored in the JSON format specified in section 2.2 on page 41. This format is human-readable, and the list and dictionary structures allow quick insights to an trial. The evaluated metric-functions for each fold are stored in a CSV format specified in section 2.2, page 41. This format features a compact representation of tabular data.

To ensure the tracking of changes, a multilayer tracking system was deployed. GIT was used as a version control system for the code base. The version and branch from which an trial was run, is stored. As GIT was developed for many but relatively small files, the database has to be tracked differently. Alongside the PARQUET files, another JSON file carries the hash strings (see section 2.2, page 40) for each database file and, in addition, the file paths and hashes of all finite element simulation files are used to generate the database. In a trial, the documented hash is compared with the one from the actual file, and equality is enforced.

As analysed above, the features are from the category multivariate time-series. The available algorithms were collected in section 2.2 on page 57 ff. Time-series data can be either directly treated as attributes, or attributes have to be extracted. In the case of the direct usage, each sample of each time-series represents an attribute. The drawback of that approach lies in the assumed independence of each attribute. Of course, for sensor signals, single samples are heavily correlated with each other. Consequently, the direct application appears not expedient.

In terms of performance and flexibility, the Artificial Neural Networks (ANNs) architectures explained in section 2.2 on page 50 ff were deemed to have the highest potential. Especially, the prediction of complete target multivariate time-series is only efficiently implementable with ANNs. In figure 4.1 is the processing of feature multivariate time-series with an exchangeable block is presented. To keep the possibility of different outputs, the dense layers and the output layers of the preferred ANN should be used to process the generated features attributes.

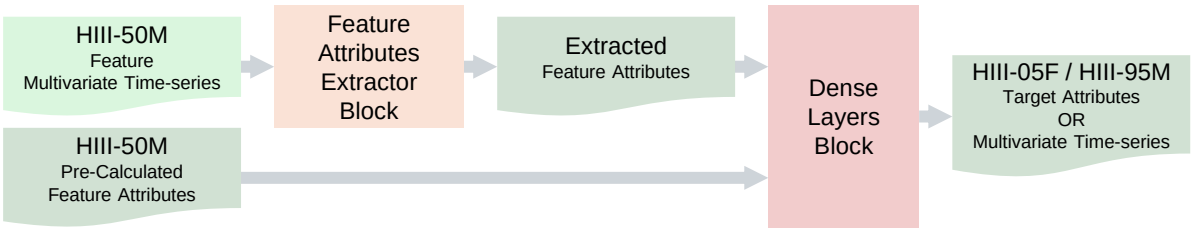


Figure 4.1: Base Architecture for the Example with Feature Attributes Extractor Block and Estimator Block

As deterministic feature attributes, the injury-criteria should always be considered. In the specific study on the domain-agnostic deterministic feature attributes extractors, the wrapper of TSFRESH and CATCH22 in SKTIME, and the library's own ROCKET algorithm were used. In addition, the Principal Component Analysis (PCA) implementation in SCIKIT-LEARN should be studied.

Due to limited resources, not all algorithms were tested with all output types. The development scenario is the prediction of the European New Car Assessment Programme (Euro NCAP) relevant injury-criteria. Only the most promising approaches will be studied with other outputs. Two workstations were available to perform the trials. In the first, the Central Processing Unit (CPU) was an Intel Core i7-11800H with 2.3 GHz and 8 cores. The Graphics Processing Unit (GPU) was an NVIDIA GeForce RTX 3080 Laptop GPU with 8 GB dedicated storage. The Random Access Memory (RAM) was 32 GB. The second workstation contained an Intel Xeon W-2123 CPU with 3.6 GHz and 4 cores. The GPU was an NVIDIA Quadro P2000 with 5 GB dedicated storage. The Random Access Memory (RAM) was 64 GB. Both GPUs were equipped with a CUDA Core enabling the computation of ANN on the faster GPU instead of the CPU. Since both workstations were used, the computation times are provided as ranges.

4.2 Baseline Performance

A baseline performance was reported to determine if a metamodel learned anything meaningful. Like in medical research, where therapies should have an effect stronger than the placebo effect, a metamodel too must predict better than a constant but reasonable guess from the training-set.

For target attributes regression tasks, the median for each injury-criterion within the training-set was chosen. For target attributes classification tasks, the most frequent value of a class, the

mode, set was determined. For target multivariate time-series regression, the median time-series for each target time-series was calculated.

The evaluation of the capped R2-scores is shown in table 4.2. For each injury-criterion, the R2-score resulted in a value of zero. This holds true for both, the predictions for Hybrid III 5th Percentile Female Dummy (HIII-05F) and Hybrid III 95th Percentile Male Dummy (HIII-95M). In any case, it can be expected that a real learning algorithm will outperform the baseline.

Table 4.2: Baseline R2-score for Target Attributes Regression Task for Injury-criteria

		HIII-05F		HIII-95M	
		Median	Capped R2-score	Median	Capped R2-score
Head	HIC_{15}	251	0	340	0
	HAC_3	55 g	0	63 g	0
Neck	NEC	23 kN	0	17 kN	0
	NTC	1.1 kN	0	1.6 kN	0
	NSC	0.8 kN	0	0.9 kN	0
Chest	CAC_3	53 g	0	49 g	0
	CDC	27 mm	0	52 mm	0
	CVC	0.2 m s^{-1}	0	0.4 m s^{-1}	0
Femur	FCC	0.8 kN	0	6 kN	0

For the target attributes classification tasks, the problematic class distributions are reflected in the assessed metric-functions shown in table 4.3 on the following page. The mode is the most frequent class. The mode-frequencies indicate that many classes are filled with only few or even no instances. To display very low instance counts in the table, the frequencies are cut after two decimals instead of rounded.

The decision for class-occurrence-based weighting for generating the baseline estimations resulted in high F1-scores for highly imbalanced datasets. However, as a baseline this is desirable since a metamodel would primarily learn the most frequent class in such cases. Random correct guesses of the minority class would improve the score unreasonably.

In table 4.3 on the next page, the colours of the F1-score cells indicate whether a class is useful for further predictions. F1-scores over the in this work in general used threshold over 0.8 will likely not be exceeded by a metamodel. Especially, the Femur Compression Force Criterion (FCC) does not seem to be usable in the given data distribution. For the HIII-95M, the Neck Tension Force Criterion (NTC) and Neck Shear Force Criterion (NSC) indicate similar condition.

Table 4.3: Baseline F1-scores for Target Attributes Classification Task for Injury-criteria Classes

Region	Criterion	Classes	HIII-05F			HIII-95M		
			Mode	Frequency	F1-score	Mode	Frequency	F1-score
Head	HIC_{15}	2	1	0.93	0.89	1	0.74	0.63
		3	2	0.73	0.62	2	0.64	0.50
		5	4	0.63	0.49	4	0.51	0.34
		7	6	0.63	0.49	6	0.51	0.34
	HAC_3	2	1	0.97	0.96	1	0.67	0.54
		3	1	0.63	0.49	2	0.60	0.45
		5	3	0.53	0.37	4	0.32	0.15
		7	5	0.53	0.37	6	0.32	0.15
Neck	NEC	2	1	0.99	0.99	1	0.98	0.97
		3	2	0.70	0.58	2	0.87	0.82
		5	3	0.58	0.42	4	0.73	0.61
		7	5	0.58	0.42	6	0.73	0.61
	NTC	2	1	0.97	0.96	1	0.97	0.96
		3	2	0.70	0.57	2	0.96	0.95
		5	3	0.50	0.34	4	0.90	0.85
		7	5	0.50	0.34	6	0.90	0.85
	NSC	2	1	0.88	0.83	1	0.99	0.99
		3	2	0.61	0.46	2	0.95	0.93
		5	3	0.47	0.30	4	0.91	0.88
		7	5	0.47	0.30	6	0.91	0.88
Chest	CDC	2	0	0.97	0.96	0	0.55	0.40
		3	0	0.97	0.96	1	0.67	0.54
		5	1	0.97	0.96	2	0.44	0.26
		7	2	0.55	0.40	0	0.32	0.15
	CVC	2	1	1.00	1.00	1	0.93	0.90
		3	2	0.99	0.99	2	0.56	0.41
		5	4	0.99	0.99	4	0.40	0.23
		7	6	0.99	0.99	6	0.40	0.23
Femur	FCC	2	1	1.00	1.00	1	1.00	1.00
		3	2	0.99	0.99	1	0.95	0.92
		5	4	0.99	0.99	2	0.95	0.92
		7	6	0.99	0.99	4	0.81	0.72

Table 4.4 summarises the ISO-ratings for the median time-series for each panel. The evaluated time window was selected shorter than for the evaluation of the FE-model in chapter 3 on page 69 ff. The samples before 60 ms and after 120 ms were omitted. During those time intervals, as displayed exemplary in figure 4.2, the variation due to the factor changes in the database was minimal; the Interquartile Range (IQR) (see section 2.2, page 55) at each timestep is small. As a result, the ISO-rating would have unreasonable high values. The channel naming and units follow the ISO-codes defined in table 2.5 on page 32 of section 2.1.

Table 4.4: Baseline Performance as ISO-rating for Target Multivariate time-series Regression of Selected Channels

ISO-code	ISO-rating HIII-05F	HIII-95M
03HEAD0000OCCUACRD	0.68	0.54
03CHST0000OCCUACRD	0.70	0.57
03CHST0000OCCUDSXD	0.81	0.77
03NECKUP00OCCUMOYD	0.64	0.52
03NECKUP00OCCUFOZD	0.63	0.66
03NECKUP00OCCUFOXD	0.70	0.54
03FEMR0000OCCUFOZD	0.71	0.88

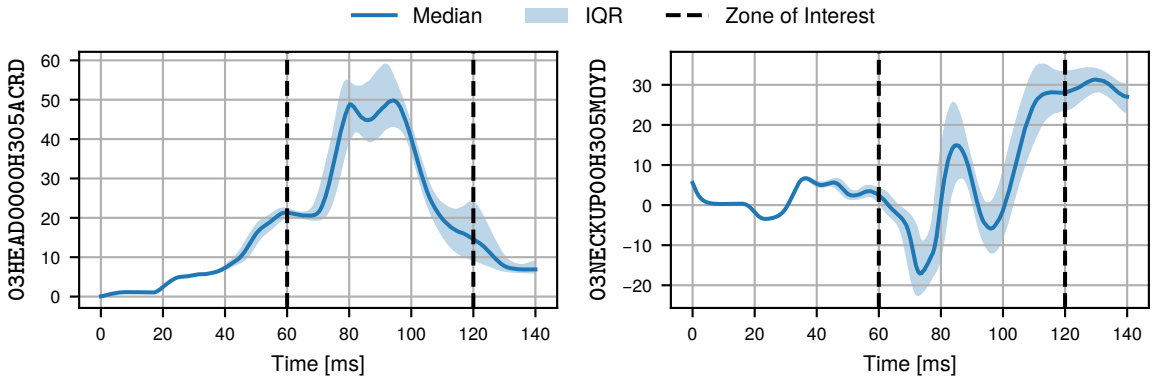


Figure 4.2: Median and IQR Corridor of Head Acceleration and Neck Moment for HIII-05F from Development-set (8177 FE-simulations)

Overall, the ISO-rating and the amount of variation of the sensor signals in the database match. For the HIII-05F, in the chest deflection and for the HIII-95M in the femur force, the variation is too low to be assessed by the selected metric-functions. As a result, an improvement by a trained metamodel cannot be expected.

In summary, the baseline for the tasks were established. For the target attributes regression task, a clear improvement is expected. For the target attributes classification task, too, improvements are expected. However, in some cases, the class distribution will not allow appropriate learning. Furthermore, for target multivariate time-series regression, an improvement is also expected.

4.3 Learnable Feature Attributes Extractors

4.3.1 Study with Convolutional Neural Networks

4.3.1.1 Architecture of the Convolutional Neural Network

The Convolutional Neural Network (CNN) was built using TENSORFLOW and KERAS. Next to the hyperparameters of each level of the framework, the architecture itself is controllable by hyperparameters. The architecture is displayed in figure 4.3 and elaborated below.

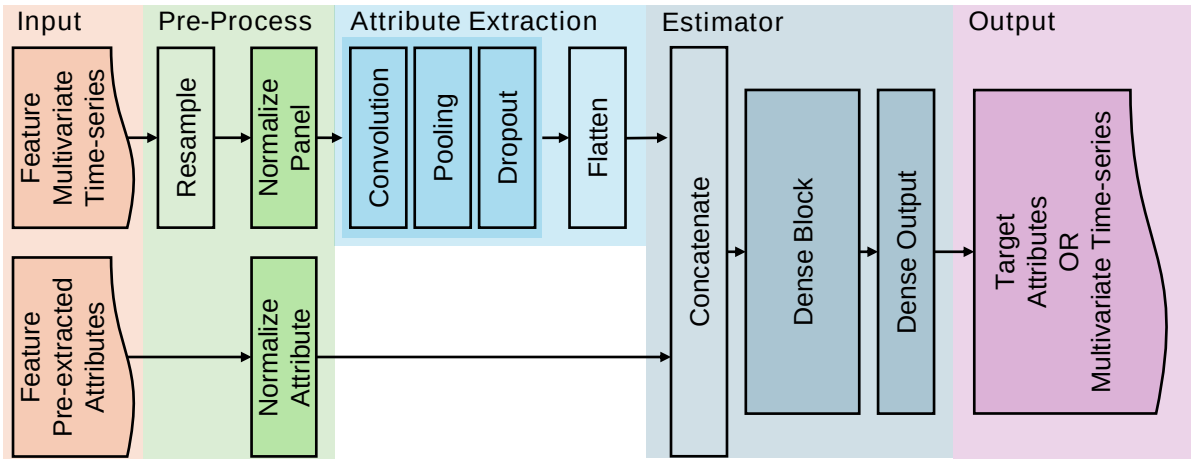


Figure 4.3: Base Architecture of CNN

The Artificial Neural Network (ANN) can process tabular and temporal input. The temporal input is passed through a feature attributes extractor block. The time-series data can be downsampled using the rolling and averaging time window as implemented in PANDAS. The effect of the sampling rate on the accuracy of the representation of the time-series can be inspected in figure 4.4 on the next page. The time-series from the FE-simulation were stored with a sampling rate of 10 kHz. The time-series can be downsampled up to 0.5 kHz without a significant deviation. The flattened output of the feature attributes extractor and the pre-calculated feature attributes are concatenated and passed in to the prediction block.

The temporal and the tabular input were normalised before passed into the ANN. For the training-set, the medians and Interquartile Ranges (IQRs) were calculated. The medians were subtracted and, the data scaled, resulting in an IQR of 1 and a median of 0 for the training-set. For all sets not being the training-set, the calculated parameters were used.

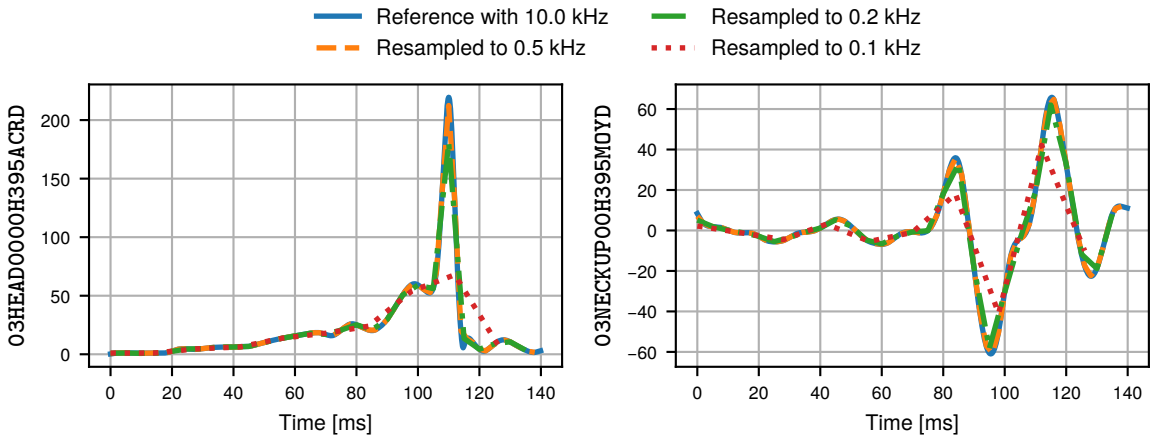


Figure 4.4: Example of Resampling Levels of Head Acceleration and Neck Moment for HIII-95M for ID 66

For multivariate time-series, the median and IQR were calculated individually per panel. For attributes, the parameters were calculated individually per attribute. For target multi-multivariate time-series and attributes regression tasks, the same strategy was applied to achieve the same weight during the fitting process.

The order of layers in the feature attributes extractor block is oriented on Zhao et al. (2017). A pooling layer follows each convolutional layer. Optional, a dropout layer was stacked behind the pooling layer. Inspired by the inception architecture in Ismail Fawaz et al. (2020), in addition to the horizontal stacking of convolutions, vertical dimensions can also be added. In a horizontal stack, the data is passed forward through the layers. Each vertical input is a copy of the temporal data. By variation of the convolutional kernel sizes, different time window sizes can be evaluated.

The prediction block consists of dense layers with Rectified Linear Unit (ReLU) activation functions. The multi-output is realized by either having one stack of dense layers and multiple output neurons. Alternatively, the features are copied into individual prediction blocks with a single output neuron each. The latter approach is related to the so-called multitask learning, a subcategory of transfer learning.

For all regression tasks, the Mean Absolute Error (MAE) as defined in section 2.2 in equation 2.44 on page 44 and for all classification tasks, the cross entropy as defined in the equations 2.61 and 2.62 on page 54 was used. Furthermore, in ANNs, an additional validation-set can be taken from the training-set to evaluate the generalization performance during the parameter optimization over the epochs. This monitoring allows the optimizer not to stop after a defined number of epochs, but at the indicated point of overfitting.

4.3.1.2 Hyperparameter Study with Convolutional Neural Network

Design of the Hyperparameter Study with Convolutional Neural Network

The computation time of the CNN based ANNs was about 20 min per fit. Hence, a complete trial required in total approximately 1.5 h. This relatively inexpensive algorithm enabled the use of advanced hyperparameter tuning algorithms. The Tree-structured Parzen Estimator (TPE) algorithm from Watanabe and Hutter (2023) implemented in the PYTHON library OPTUNA by Akiba et al. (2019) seemed to be a promising trial generator. During the iterations, two Gaussian Process Regressions (GPRs) are repeatedly fitted; one to the set of hyperparameter values associated with the best result and one to the remaining set of hyperparameters values. The algorithm will select the next hyperparameters values of a trial by maximising the ratio of both metamodels. In the OPTUNA documentation, for the TPE algorithm, around 200 iterations are considered suitable.

The hyperparameter space was defined in relation to the variability described above of the ANN. For the input by the hyperparameter (file_names_ai_in), temporal data, injury-criteria or both could be used. Additionally, the trial generator could select the channels actually used and injury-criteria.

The block of convolutional layers was allowed to have up to four independent lanes (conv_width) with up to four stacked convolutional layers each (conv_depth). The number of convolutional kernels and the convolutional kernel sizes were suggestible for each convolutional layer separately. Allowed values for the number of convolutional kernels were from 10 to 100 with a step width of 10. The same step with, but for the range of 1 to 50, the convolutional kernel size was specified. By the convolutional kernel size, characteristics in the signals of different widths should be captured. E. g., a convolutional kernel width of 50 would cover a characteristic of width 50 ms, if the signal was sampled with 1 kHz over 140 ms. The pooling size (pooling_size) was set globally between two and five, and the pooling strategy (pooling_strategy) to max-pooling or average-pooling. Finally, the dropout layer could be turned on and off (dropout_conv) and its dropout rate ranging from 0.1 to 0.5 be suggested (spatial_dropout_rate).

The shape of the dense layer block was suggestible, too. The number of layers (n_dense_layers) were in the range of 2 to 5. The width of the first layer (fst_dense_layer_shape) and last layer (last_dense_layer_shape) could take all values from 10 to 100 divideable by 10. The width for all layers in between were linearly interpolated and converted to integers. For the dense layers, L1 regularization terms or L2 regularization terms as explained in section 2.2 in equation 2.34 and equation 2.34 on page 39) were activatable (dense_regularizer).

The time-series were resampled to any length of 70, 140 and 1400 (temporal_feature_n_tsps), where 1400 corresponds to the 10 kHz sampling rate. Further, the dense layers can be either shared within the output layer, or cloned as individual lanes (share_dense). Finally, the learning rate could take the typical values of 1×10^{-3}, 1×10^{-4} and 1×10^{-5} (learning_rate).

Results of the Hyperparameter Study with Convolutional Neural Network

In table 4.5 the prediction performance for the values of injury-criteria for Hybrid III 5th Percentile Female Dummy (HIII-05F) and Hybrid III 95th Percentile Male Dummy (HIII-95M) is presented. The R2-scores for HIII-05F result from a single trial. To generate the reported spans for the HIII-95M R2-scores, the same trial was repeated four times. This stability assessment shows that for hyperparameter tuning changes in the R2-score smaller than 2×10^{-3} are more likely due to the randomness of the algorithm than to the changed hyperparameters. By the comparison of the R2-scores from the training-set and validation-set, no overfitting is indicated; all differences are smaller than 1×10^{-2}. The only exception is the Neck Extension Moment Criterion (NEC) of the HIII-05F with 2×10^{-2}, but here, the R2-score of the training-set is already unsatisfactory.

Table 4.5: CNN in Final Configuration – R2-score for Target Attributes Regression Task for Injury-criteria

		HIII-05F		HIII-95M	
		Training-set	Validation-set	Training-set	Validation-set
Head	HIC_{15}	0.93	0.93	0.9390 ±0.0005	0.9292 ±0.0005
	HAC_3	0.85	0.84	0.9400 ±0.0005	0.9333 ±0.0005
Neck	NEC	0.58	0.56	0.9275 ±0.0010	0.9219 ±0.0002
	NTC	0.72	0.71	0.8932 ±0.0013	0.8887 ±0.0006
	NSC	0.80	0.79	0.9310 ±0.0013	0.9240 ±0.0002
Chest	CDC	0.25	0.24	0.9329 ±0.0020	0.9308 ±0.0002
	CVC	0.35	0.34	0.9389 ±0.0007	0.9357 ±0.0005
	CAC_3	0.91	0.90	0.7586 ±0.0016	0.7526 ±0.0003
Femur	FCC	0.66	0.66	0.9292 ±0.0005	0.9268 ±0.0002
Total	Median	0.72	0.71	0.9310 ±0.0010	0.9268 ±0.0003

The R2-scores in table 4.5 for HIII-95M were found to be with more than 0.9 high. An exception is the Chest Acceleration over 3 ms Criterion (CAC_3) with 0.75 that is almost in the acceptable range.

More mixed results were observed for the HIII-05F. The predictions for the head related injury-criteria and the CAC_3 were with above 0.8 found to be acceptable. For the Chest Deflection Criterion (CDC) and the directly related Chest Viscous Criterion (CVC), only scores below 0.4 were achieved. The low scores result from the instable seatbelt behaviour at the HIII-05F neck, as elaborated in section 3.4 on page 134 ff. The cross plot in figure 4.5 on the next page reveals that the metamodel predicts values for the CDC as they would be expectable when the instable behaviour would not exist. The ground truth contained values lower than 20 mm. This value range was completely neglected in the predictions. The instable seatbelt behaviour affects the reliability of the neck kinetic as well, because of the interaction of the seatbelt and the neck's surface. The R2-score of the Femur Compression Force Criterion (FCC) was relatively low due

to some extreme outliers. Interestingly, the CAC_3 was not influenced by the instable seatbelt behaviour. This indicates that in principle, the seatbelt works successfully as a restraining system for the whole thorax. Practically, this must be considered an artifact of the FE-model because a seatbelt that applies the restraining force on the throat is not acceptable.

The results with the final setting of hyperparameters from the tuning procedure were discussed above. In figure 4.6 on the facing page, the hyperparameters as defined on page 156 and the trials are presented as a parallel coordinates diagram. The hyperparameter values are indicated row wise. The last row is the resulting R2-score of the respective trial. The lower half of the hyperparameters correspond to the convolutional layers. The absence of the convolutional layers and, hence, temporal features resulted in inferior performance of the metamodel. Furthermore, a small learning rate and individual dense layers for each target attribute seem to be beneficial. For the number of stacked dense layers and their widths, there is no clear trend. Finally, the generalization capabilities and the R2-score were improved by adding the L2 regularization terms and a dropout rate from 0.2 to 0.4.

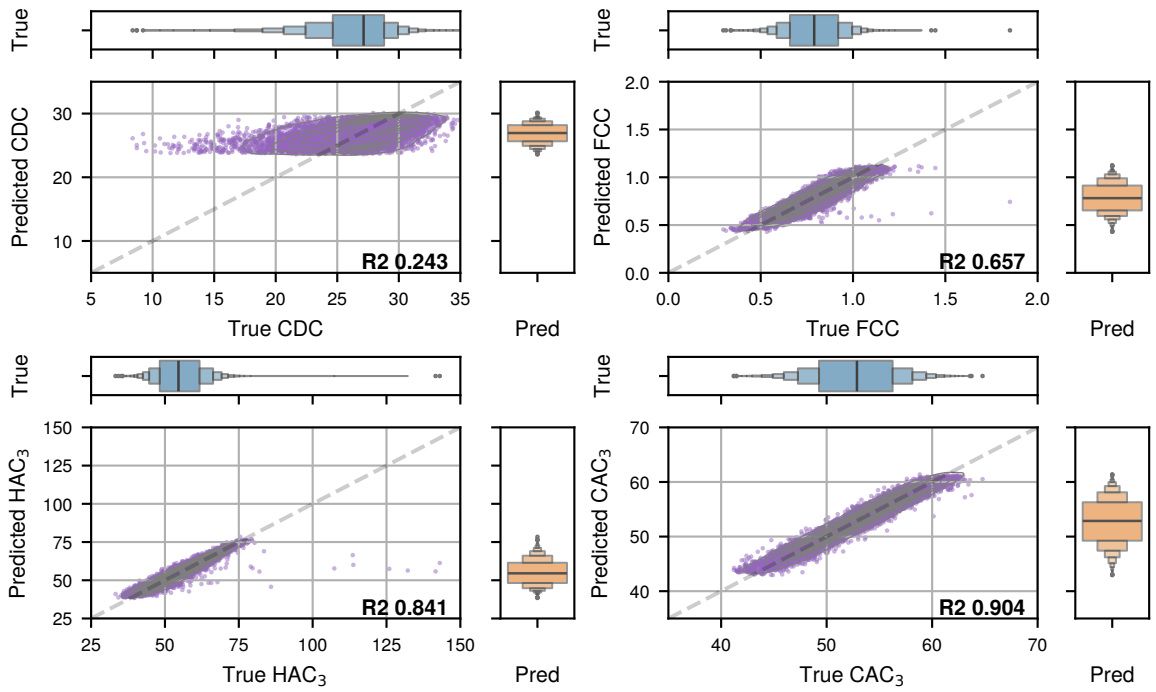


Figure 4.5: Cross plots with Prediction of Selected Injury-criteria of Head, Chest, and Femur for HIII-05F using Final CNN and Concatenated Validation-set

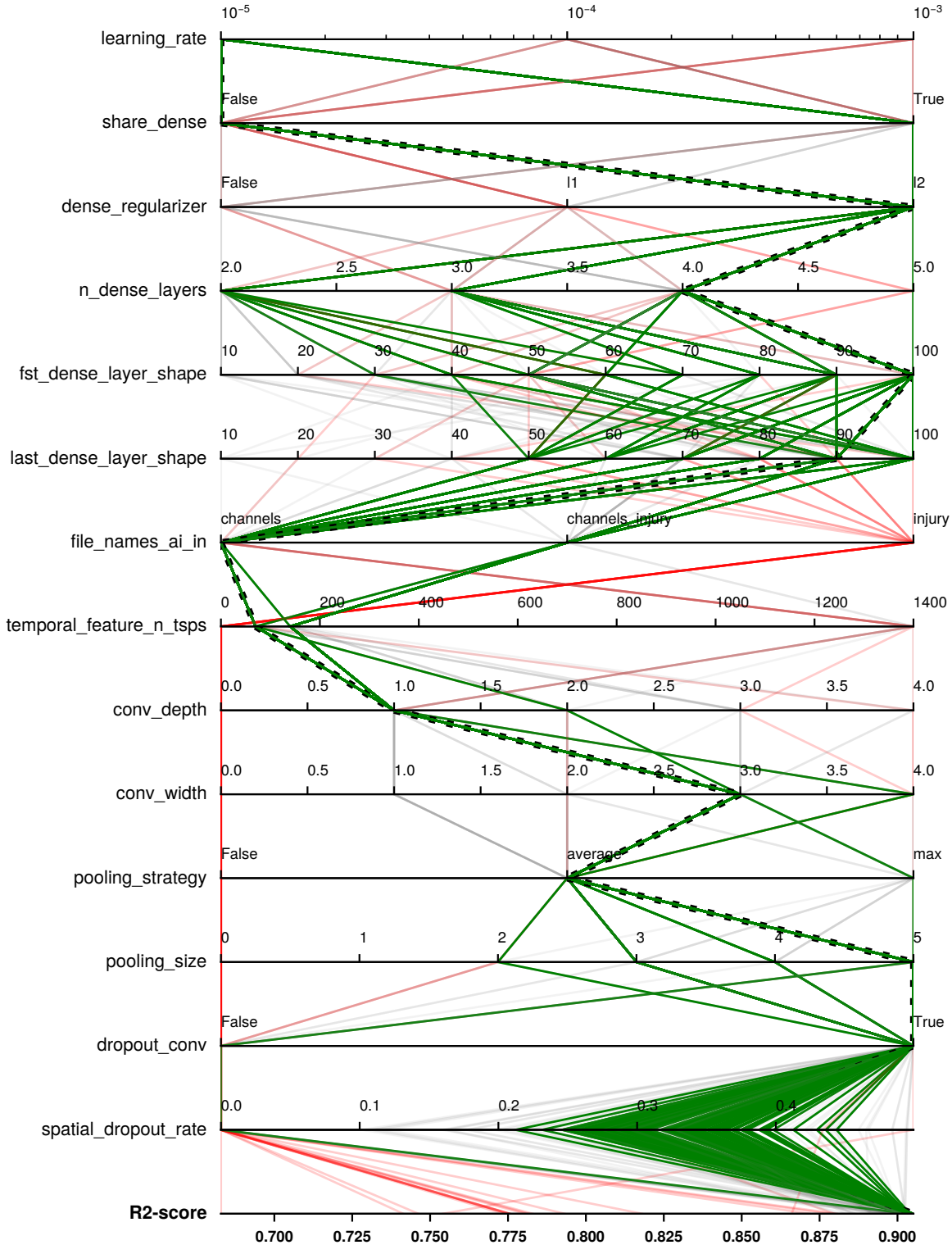


Figure 4.6: Parallel Coordinates for CNN Hyperparameter Tuning for 200 Trials (Black: Best Trial, Green: Best in Tolerance, Red: Bottom 20 Trials)

In agreement with figure 4.4 on page 155 is the preference of smaller sampling rates. Since there is an insignificant loss of information from the downsampling, this physic-related information does not have to be modelled within the ANN but can be provided directly as observational bias as mentioned in section 2.3 on page 61. From the tuning-algorithm tested trials, three parallel convolutional layers seemed beneficial. To keep the diagram readable, the convolutional kernel size is not included in figure 4.6. In the 10 best performing hyperparameter settings, there were always 2 convolutional kernels with a length of around 40 and 1 convolutional kernel with the length of 10. In combination with the favoured number of samples of 70, the lengths correspond to a time window with the width of 80 ms and 20 ms respectively. The average pooling with 2 to 5 elements seemed preferred.

4.3.2 Study with Recurrent Neural Networks

4.3.2.1 Architecture of the Recurrent Neural Networks

The basic architecture and data normalization strategy was shared with the module described for the CNN above on page 154 ff. As representative of Recurrent Neural Networks (RNNs), the Long Short-Term Memory (LSTM) architecture was selected. In figure 4.7, the feature attributes extractor block from the CNN was replaced by a LSTM block, while the input layer and the parallel output layers remained the same.

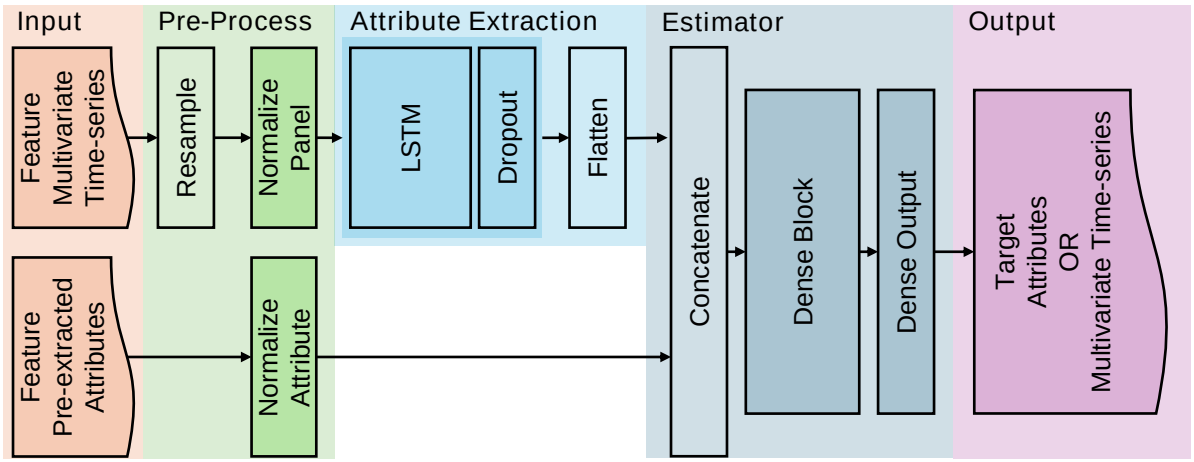


Figure 4.7: Base Architecture of LSTM Based RNN

The number of cells in a KERAS LSTM layer is determined by the signal lengths passed as input. As output of a layer, either each cell's output as hidden state can be used or only the last cell's output. Another hyperparameter controls if the signal is used only in its original direction or additionally in the opposite direction. In the latter case, the layer is duplicated, and the additional outputs are concatenated with the outputs from the forward pass. Other hyperparameters are the number of stacked LSTM layers, the number, and shape of the dense layers and the learning rate. By default, multivariate time-series are directly passed into one LSTM layer. A switch can generate an individual LSTM stack per time-series.

4.3.2.2 Hyperparameter Study with Recurrent Neural Networks

Design of Hyperparameter Study with Recurrent Neural Networks

The first observation during the training of LSTM based architectures was the comparably high computation time of around 8 h for one complete trial. The high computation time limited the number of iterations for hyperparameter tuning. Instead of extensive search strategies, a tournament-style approach was used; the results for few settings of one hyperparameter were compared, and the best was chosen for the subsequent rounds. The evaluated metric-function was the arithmetic mean of the R2-scores for the target attributes regression task for injury-criteria of the HIII-95M.

In the initial hyperparameter setting was the bidirectional treatment (is_bidirectional) of the time-series deactivated. There were 140 parallel LSTM lanes with 300 units each (lstm_units). Depending on the hyperparameter use_sequence_output, Either only the last output of each LSTM lane was used or all hidden states passed. The LSTM lanes shared the temporal input. Each target attribute had its individual block of dense layers (share_dense) with 50 layers of 20 neurons each (dense_layer_shapes). The initial learning rate was 1×10^{-3} (learning_rate). In addition to the feature multivariate time-series, the pre-computed injury-criteria could be provided, too (ai_in).

Results of Hyperparameter Study with Recurrent Neural Networks

In figure 4.8 on the next page, the result of the hyperparameter tuning was evaluated. As first insight, the R2-score over 0.9 must be stated for all iterations. For the scope of this research, this rating was defined as sufficient. Providing the pre-extracted injury-criteria, in addition to the time-series with 1 kHz sampling rate, slightly improved the R2-score. Furthermore, it seemed beneficial to pass the outputs of each LSTM unit to the block of dense layers. A transfer learning inspired split of the dense layers into independent blocks per target attribute while sharing the output of the feature attributes extractor, did not improve the R2-score.

The bidirectional evaluation of the time-series and the shape of the dense layers had only minimal influence on the R2-score. However, adding the additional LSTM line increased the computation time by a factor of 1.5. A similar increase was observed for the very complex dense block. No clear tendency was observed for the structure of the LSTM block, apart for an increased computation time with a more complex feature attributes extractor block.

The most significant improvement was found for the learning rate of the hyperparameter optimizer. A smaller learning rate improved the R2-score. Due to the early stopping strategy, the computation time was not significantly influenced.

In table 4.6 on the following page, the prediction performance for the values of injury-criteria for HIII-05F and HIII-95M are presented. The R2-scores are almost identical with the ones from the CNN presented in table 4.5 on page 157.

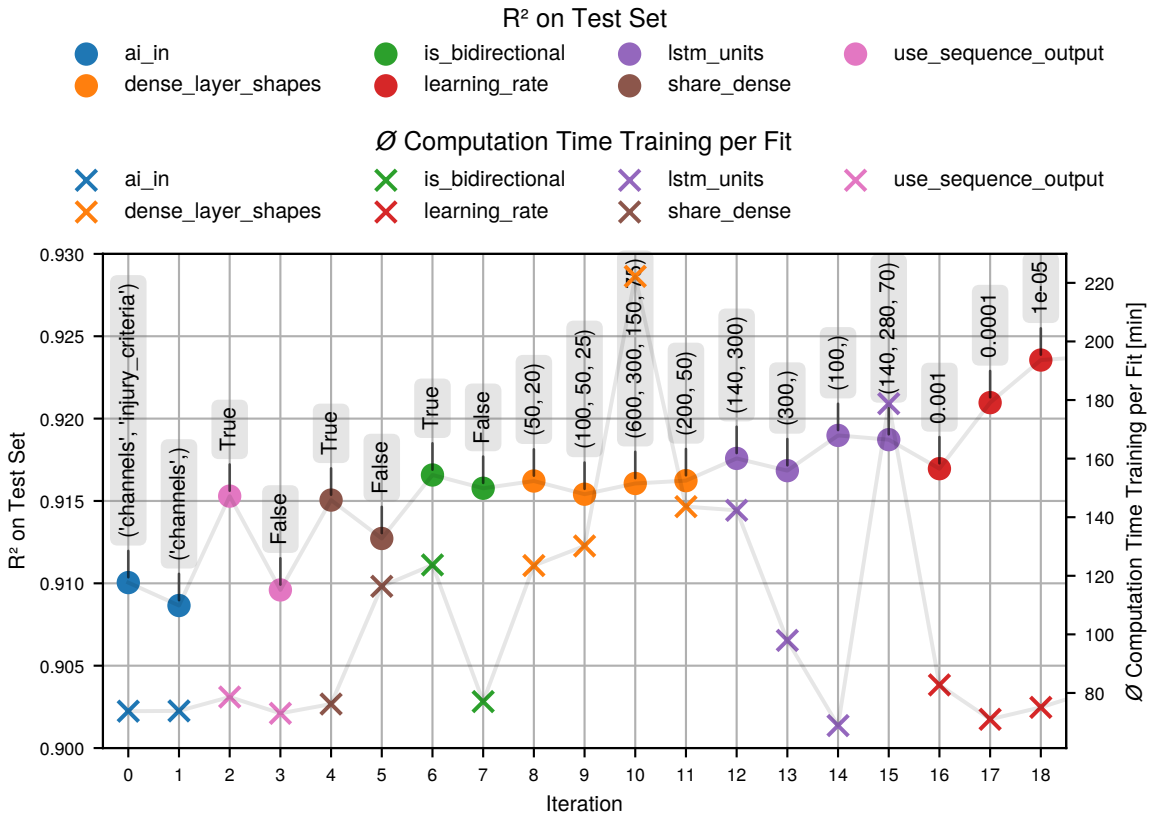


Figure 4.8: Tournament Iterations for LSTM based RNN Hyperparameter Tuning for the Target Attributes Regression Task for Injury-criteria of HIII-95M

Table 4.6: LSTM Based RNN in Final Configuration – R2-scores for Target Attributes Regression Task on Injury-criteria

		HIII-05F		HIII-95M	
		Training-set	Validation-set	Training-set	Validation-set
Head	HIC_{15}	0.93	0.92	0.94	0.92
	HAC_3	0.84	0.84	0.94	0.92
Neck	NEC	0.58	0.56	0.93	0.92
	NTC	0.72	0.71	0.90	0.89
	NSC	0.80	0.78	0.94	0.92
Chest	CDC	0.27	0.22	0.93	0.92
	CVC	0.36	0.32	0.94	0.93
	CAC_3	0.91	0.90	0.76	0.75
Femur	FCC	0.66	0.65	0.93	0.92
Total	Median	0.72	0.71	0.93	0.92

4.3.3 Comparison and Discussion of CNN and RNN

The hyperparameters of CNN and the LSTM based RNNs were tuned. A clear tendency for a smaller learning rate, the utilization of the feature multivariate time-series and regularization terms was observed. For the other hyperparameters, both approaches allowed a broad value range without significant change in the assessed R2-scores.

Regarding the observed R2-scores, both approaches were comparable. However, the computation time for one single trial with 1 h to 2 h for the CNN is clearly superior to the around 8 h to 20 h of the LSTM based RNN. The required disc space to store the trained metamodel is for the CNN ca. 34 MB and for the LSTM based RNN ca. 28 MB. Despite both values are slightly different, the contrast with the size of the development-set with 6.4 GB makes the difference insignificant.

In summary, sufficient results were observed for the learnable feature extractors. The CNN seems the preferable approach. Consequently, this architecture will be further used as the reference.

4.4 Studies on Deterministic Attributes Extraction

4.4.1 Studies on Pre-extracted Target Independent Attributes

As discussed above, three approaches of pre-extracted feature attributes were tested – the injury-criteria, the TSFRESH, and the CATCH22 methodologies. The base was the Convolutional Neural Network (CNN) as established and tuned above. The feature attributes extractor block was removed and effectively only the dense layers used.

The feature attributes of the TSFRESH methodology were pre-calculated. The approach evaluates each instance independently, what enables the pre-calculation. The algorithm is computational expensive; the extraction of attributes from the development-set required about 6 h on the cluster using 48 processes parallel. For the feature time-series, the same channels selected as above were used, and the sampling rate remained at 10 kHz. After the extraction process, all attributes containing undefined numbers (NaN) or being constant (variance smaller than 1×10^{-6}) were dropped. From the 25 channels, 18 212 attributes were extracted.

The same procedure was applied to the CATCH22 methodology. For TSFRESH and CATCH22, the wrapper of SKTIME was used. By the CATCH22, 427 attributes were extracted. For comparison, the pre-extracted values of the injury-criteria were used. The pre-extracted attributes were regularly passed through the evaluation framework using the dense layers from the final Artificial Neural Network (ANN) from above and the regular tabular data processing pipeline.

The predictions based on both, the TSFRESH and CATCH22 resulted in a zero-capped R2-score of 0.0 in the target attributes regression task for all monitored injury-criteria. In contrast, for the pre-extracted values of the injury-criteria, R2-scores with 0.62 for the Hybrid III 5th Percentile Female Dummy (HIII-05F) and 0.87 for the Hybrid III 95th Percentile Male Dummy (HIII-95M) were recorded.

4.4.2 Study with ROCKET Feature Extractor

The implementation of SKTIME of the ROCKET algorithm as published by Dempster et al. (2020) (see section 2.2, page 57) was used again in combination with the dense layers from above. The algorithm's hyperparameter (n_kernel) controls the number of convolutional kernels, which are randomly initialised and not trainable. In the trials, the values 20, 50, 100, 1000 and 10 000 were evaluated. The 10 000 convolutional kernels were the default value.

For 20, 50 and 100 convolutional kernels the R2-score on the validation-set ranged from 0.90 to 0.92. For 100 convolutional kernels, a tendency to overfitting was observed. For 1000 and 10 000 convolutional kernels, the zero-capped R2-score was 0.0.

4.4.3 Study with Attributes from Principal Component Analysis

For Principal Component Analysis (PCA), the SCIKIT-LEARN implementation was used. The library supports the classic linear version and the implementation of a kernel function based PCA. To screen the capabilities of the methods, the linear PCA, and the kernel function based PCA with a linear kernel function and a radial basis function as kernel function as described in section 2.2 equation 2.52 on page 49 were used. For the three cases, the numbers of extracted feature attributes, for PCA also known as Principal Components (PCs), were set to 20 and 40.

Similar to the temporal feature processing pipeline of the framework, the time-series were downsampled and then normalized over the panel. With the time-series downsampled to 1 kHz, the 140 samples of the 25 time-series resulted in 3500 attributes as input to the PCA. Next, the resulting PCs were regularly passed through the evaluation framework using the dense layers from the final ANN from above and the regular tabular data processing pipeline. All parameters from the normalizer and PCA were fitted to the respective training-set. For linear PCA and the full development-set, the 20 PCs (n_components) correspond to 95 % and the 40 PCs correspond to 98 % of the variance explained by the extracted PCs regarding the 3500 attributes.

In figure 4.9 on the next page, the R2-scores for the feature attributes extraction by PCA in the target attributes regression task are presented as the median of the injury-criteria. All metamodels generalize well. In general, more PCs lead to slightly higher scores. The best configuration for both percentiles was the PCA with a linear kernel function using 40 PCs.

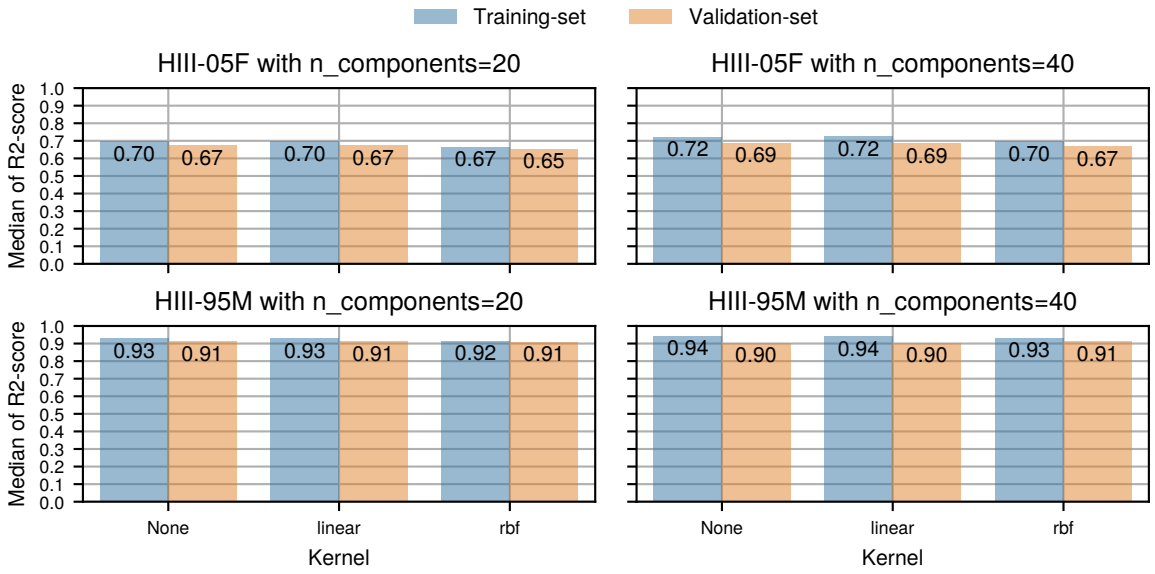


Figure 4.9: Median of R2-score of PCA as Feature Attributes Extractor and ANN for Target Attributes Regression Task

4.5 Extended Predictions with Final Architecture

4.5.1 Selection of Final Learning Algorithm

Above, as learnable feature attributes extractors, the Convolutional Neural Network (CNN) and the Long Short-Term Memory (LSTM) based Recurrent Neural Networks (RNNs) were discussed. In addition, the deterministic domain independent feature attributes extraction was performed by the TSFRESH, CATCH22, and ROCKET algorithms. Furthermore, the Principal Component Analysis (PCA) feature attributes extraction was presented.

The approaches of TSFRESH and CATCH22 in combination with the pre-defined Artificial Neural Network (ANN) architecture without further tuning of the hyperparameters were not successful. Especially, the high number of feature attributes could not be used by the ANN. With a high number of convolutional kernels, the ROCKET algorithm underperformed similarly. For the PCA-based approach and the ROCKET algorithm with a few convolutional kernels were comparable. For both algorithms, a light indication of overfitting was detected.

The CNN and the LSTM based RNNs were compared above on page 163. The CNN was found to be superior due to the significantly lower computational cost. The computational cost for the CNN, the PCA-based approach and the ROCKET algorithm was similar in the range of 3 h to 6 h for a complete trial. The CNN seems to generalize the data slightly better. Furthermore, the assessed R2-score as median over the target attributes with 0.93 was slightly better than for the other two, which ranged from 0.90 to 0.92. Consequently, the CNN will be used for all further work.

4.5.2 Prediction Performances with Final Algorithm

The final selected CNN was used to evaluate whether target types besides the continuous target attributes can be successfully predicted. The results for the target attributes classification of single or multiple classes are presented in table 4.7 on the facing page. The number of classes and the conversion of the continuous target attributes were defined in section 4.1 on page 147. Similar to the presentation of the baseline in section 4.2 on page 150, in table 4.7 the occurrence frequency (Freq.) of the most common class for each target attribute is indicated. The cell-colours should provide an orientation on how a F1-score should be rated in the context of the baseline performance and class-distribution. The results will be discussed in the context of the findings from the target attributes regression presented in table 4.5 on page 157. By common sense, a classification task is less complex than a regression task, and hence, given a valid class-distribution, a better performance can be expected.

Positive results were observed for the target attributes classification for the head-related Head Injury Criterion for up to 15 ms (HIC_{15}) and Head Acceleration over 3 ms Criterion (HAC_3) for both Anthropomorphic Testing Devices (ATDs). The target attributes classification for the neck-related Neck Extension Moment Criterion (NEC), Neck Shear Force Criterion (NSC), and Neck Tension Force Criterion (NTC) for the Hybrid III 5th Percentile Female Dummy (HIII-05F) was found to be better than when predicted by regression. Similar good as the regression performance was the target attributes classification for the chest-related Chest Deflection Criterion (CDC) and Chest Viscous Criterion (CVC) for the Hybrid III 95th Percentile Male Dummy (HIII-95M). For most of the injury-criteria mentioned in this paragraph, a few number-of-class scenarios suffered from the problematic class distribution.

The estimation of CDC and CVC class values of the HIII-05F were expected to be not successful due to the instable seatbelt behaviour. For the 7 classes of the CDC was this expectation fulfilled. All other scenarios could not be evaluated regarding this expectation because of the class distributions. Similarly, the Femur Compression Force Criterion (FCC) for both ATDs and the neck-related injury-criteria for the HIII-95M were not suitable for classification.

In the table 4.8 on page 168 are the ISO-ratings for the target multivariate time-series regression of selected channels presented. The channels naming follows the ISO-code defined in table 2.5 on page 32 in section 2.1. Similar to the classification, the baseline performance (base) presented in table 4.4 on page 153 must be considered in the evaluation.

In all cases, the metamodel performed better than the baseline. No indication of overfitting was observed. For most channels, the ISO-grade Good or better was achieved. The HIII-05F chest deflection's (`03CHST0000H305DSXD`) relatively small variance over the panel was reflected in the baseline ISO-rating of 0.81. The metamodel could not extract additional knowledge in this particular case, as the ISO-rating of 0.82 suggests. For the neck extension moment channel's (`03NECKUP000CCUMOYD`) estimations of both ATDs only the ISO-grade Fair was achieved.

In general, it seems that the tuned CNN performed well. Values of the evaluated metric-functions significantly lower than the stipulated 0.8 could be explained by the deficits of the development-set. The classification tasks were especially affected by the class-distributions.

Table 4.7: Final CNN – F1-score for Target Attributes Classification Task for Injury-criteria Classes

Criterion	Classes	HIII-05F			HIII-95M		
		Freq.	Training-set	Validation-set	Freq.	Training-set	Validation-set
HIC_{15}	2	0.93	0.98	0.98	0.74	0.96	0.96
	3	0.73	0.96	0.95	0.64	0.93	0.92
	5	0.63	0.92	0.91	0.51	0.88	0.88
	7	0.63	0.90	0.90	0.51	0.86	0.85
HAC_3	2	0.97	0.98	0.98	0.67	0.95	0.94
	3	0.63	0.91	0.90	0.60	0.93	0.93
	5	0.53	0.88	0.86	0.32	0.88	0.87
	7	0.53	0.86	0.85	0.32	0.86	0.85
NEC	2	0.99	1.00	1.00	0.98	0.99	0.99
	3	0.70	0.84	0.82	0.87	0.96	0.96
	5	0.58	0.76	0.76	0.73	0.91	0.90
	7	0.58	0.76	0.76	0.73	0.87	0.86
NTC	2	0.97	0.97	0.97	0.97	0.99	0.99
	3	0.70	0.91	0.91	0.96	0.98	0.98
	5	0.50	0.77	0.76	0.90	0.96	0.96
	7	0.50	0.77	0.76	0.90	0.96	0.96
NSC	2	0.88	0.94	0.93	0.99	1.00	1.00
	3	0.61	0.83	0.82	0.95	0.99	0.99
	5	0.47	0.75	0.74	0.91	0.97	0.96
	7	0.47	0.73	0.72	0.91	0.96	0.95
CDC	2	0.97	0.96	0.96	0.55	0.94	0.94
	3	0.97	0.96	0.96	0.67	0.94	0.94
	5	0.97	0.96	0.96	0.44	0.89	0.88
	7	0.55	0.53	0.52	0.32	0.78	0.77
CVC	2	1.00	1.00	1.00	0.93	0.97	0.97
	3	0.99	1.00	1.00	0.56	0.93	0.93
	5	0.99	0.99	0.99	0.40	0.85	0.84
	7	0.99	0.99	0.99	0.40	0.75	0.74
FCC	2	1.00	1.00	1.00	1.00	1.00	1.00
	3	0.99	1.00	1.00	0.95	0.97	0.97
	5	0.99	1.00	1.00	0.95	0.97	0.97
	7	0.99	1.00	1.00	0.81	0.93	0.93

Table 4.8: Final CNN – ISO-rating for Multivariate time-series Regression

	HIII-05F			HIII-95M		
ISO-code	Base	Training-set	Validation-set	Base	Training-set	Validation-set
03HEAD0000OCCUACRD	0.68	0.88	0.88	0.54	0.87	0.87
03CHST0000OCCUACRD	0.70	0.90	0.90	0.57	0.89	0.88
03CHST0000OCCUDSXD	0.81	0.82	0.82	0.77	0.92	0.92
03NECKUP00OCCUMOYD	0.64	0.78	0.77	0.52	0.78	0.78
03NECKUP00OCCUFOZD	0.63	0.83	0.83	0.66	0.85	0.84
03NECKUP00OCCUFOXD	0.70	0.85	0.85	0.54	0.82	0.83
03FEMR0000OCCUFOZD	0.71	0.88	0.88	0.88	0.93	0.93

4.5.3 Assessment of Test-set Performance

When discussing data leakage strictly, the validation-set has leaked into the metamodel because it was part of the algorithm selection process. Consequently, a testing-set should be held out during the whole process. Furthermore, the FE-model simplification chain in section 3.2 on page 86 ff had the purpose of efficient data generation, while the actual utilization of a metamodel would use the FE-model used in the restraint system development. Therefore, the testing-set with a reference FE-model was generated in section 3.4 on page 141 ff. In figure 3.56 on page 142 was the comparison of the values of the injury-criteria for development-set and testing-set presented. The discussed differences accumulate with the inaccuracy of the metamodel.

To assess the performance, the final selected CNN was trained for the target attributes regression of injury-criteria values using the complete development-set. In table 4.9 on the next page are the evaluated R2-scores for the estimation of the results within the testing-set presented. Additionally, Mean Absolute Error (MAE) was calculated to enable a relation with the direct FE-model comparison. The MAE for the FE-models were aggregated from table 3.3 on page 113 in section 3.2.

Evaluating the R2-scores in table 4.9 on the next page, only the HIII-95M CVC seems to be predictable reliably. When analysing the MAE values, the error from the FE-model simplification in combination with the metamodel seems higher than the FE-model simplification alone. Next to the different FE-model, the testing-set's Design of Experiments (DoE) had a broader factor range than the development-set. In figure 4.10 on page 170, the loss over the factor values is presented. The loss-function

$$^{\text{Error}}\mathcal{L}_{\varsigma,\tau} = \hat{\mathcal{Y}}_{\varsigma,\tau} - \mathcal{Y}_{\varsigma,\tau} \tag{4.2}$$

was evaluated individually for each instance.

Table 4.9: Testing-set Performance – R2-score and MAE of Target Attributes Regression Task for Injury-criteria

		Metamodel Trained with Development-set				FE-simulation
		HIII-05F		HIII-95M		HIII-50M
Injury-criterion	Unit	R2-score	MAE	R2-score	MAE	MAE
HIC_{15}		0.43	88	0.00	1175	31.1
HAC_3	g	0.34	6.0	0.00	83	3.74
NEC	N m	0.14	3.08	0.69	7.83	3.84
NTC	kN	0.00	0.39	0.00	1.56	0.29
NSC	kN	0.03	0.32	0.00	0.29	0.11
CDC	mm	0.00	4.81	0.00	14.3	5.55
CVC	$m\,s^{-1}$	0.00	0.13	0.81	0.12	0.03
CAC_3	g	0.00	6.28	0.43	7.81	7.06
FCC	kN	0.00	0.24	0.00	1.21	0.65

The R2-scores were calculated individually as summary statistics within the factor range of the development-set DoE and for the extrapolated values. It must be noted that both sets were generated using the Sobol algorithm. Consequently, it is not possible to completely separate the interpolation and extrapolation spaces completely. However, the factors affect the results of FE-simulations not equally.

Due to the described obfuscation of the interpolation and extrapolation spaces and the overall low R2-scores, only in few cases a clear tendency is recognisable. There is a light linear correlation between the loss for NEC and the Passenger Airbag (PAB) related factors PAB_M_SCAL and PAB_Vent_T. For the energy of the vehicle-pulse controlled by Pulse_X_Scale, the metamodel estimates higher loads than the ground truth. In figure 3.56 on page 142, a reduced amount of hard hits through the PAB on the dashboard in the HIII-95M data of the testing-set was concluded. The metamodel was trained on the occurrence of the high peaks, and consequently estimates such values.

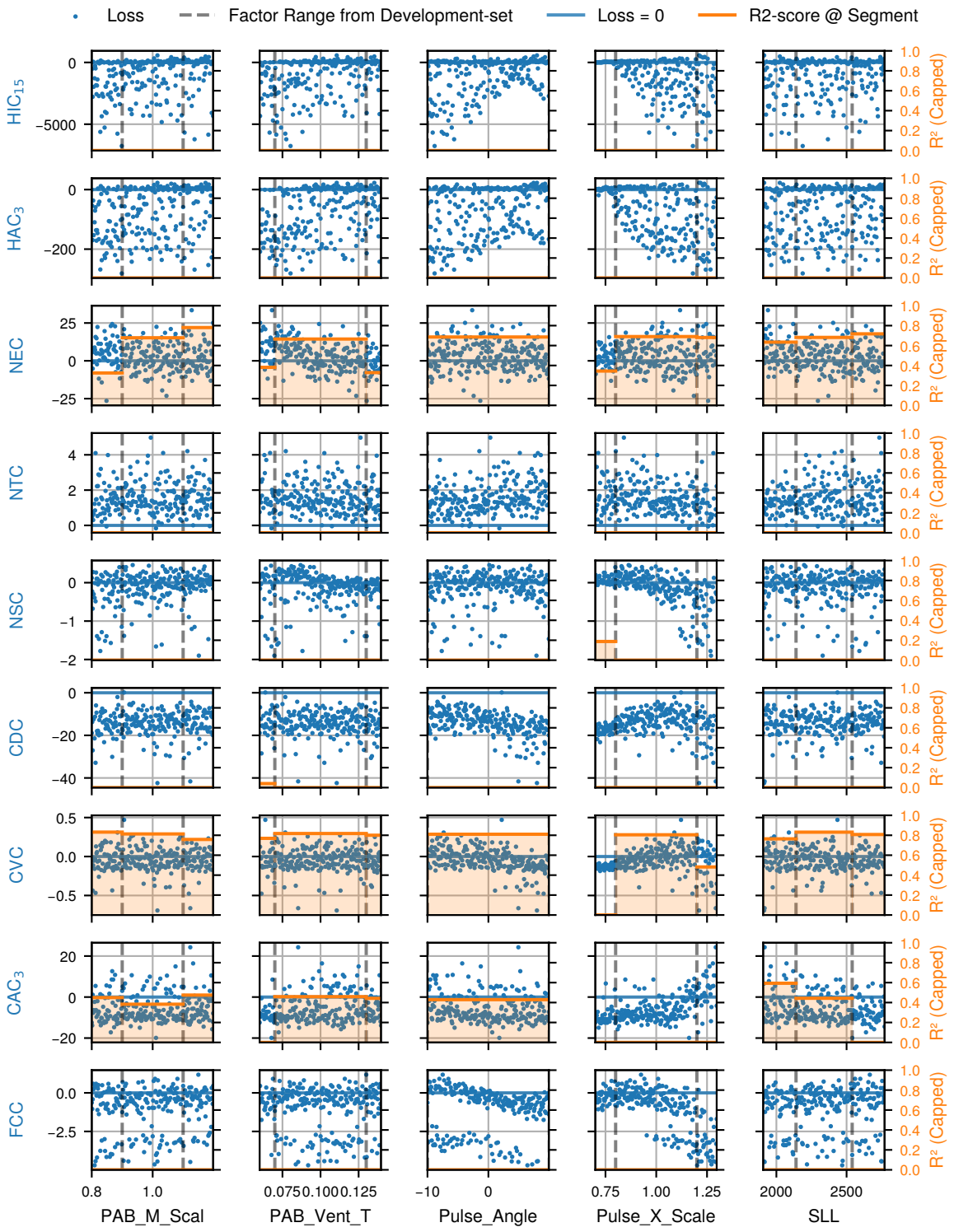


Figure 4.10: Loss $^{Error}\mathcal{L}$ and R2-score over DoE Factors for Target Attributes Regression for HIII-95M within the Testing-set

4.6 Efficient Data Utilization

4.6.1 Study on the Database Size

With the tuned Convolutional Neural Network (CNN) based architecture on page 163, the effect of different sizes of the database was explored. For this study, the characteristic of the Sobol algorithm was used: one Design of Experiments (DoE) contains smaller DoEs. Revisiting the rule that a Sobol's number of instances must equal to two to the power of the exponent s in section 2.3 equation 2.66 on page 60), the given number of instances of 8192 and a minimum number of instances of 16, enable a size variation for exponents s from 4 to 13 in a step size of 1. The minimum number of instances results in 5 instances per fold, which leads to the absolute minimum of 1 instance in the validation-set for the fitting of the Artificial Neural Network (ANN).

In figure 4.11 the influence of the number of instances was plotted for the predictions on Hybrid III 5th Percentile Female Dummy (HIII-05F) and Hybrid III 95th Percentile Male Dummy (HIII-95M). The arithmetic mean of the R2-scores increases fast with the number of instances. After around 2000 instances for the HIII-05F and approximately 500 instances, a plateau was reached. For 4000 instances the R2-scores for the Femur Compression Force Criterion (FCC) of the HIII-05F and the Neck Shear Force Criterion (NSC) and Chest Acceleration over 3 ms Criterion (CAC_3) of the HIII-95M decreased and recovered with the next iteration where ca. 4000 instances are added.

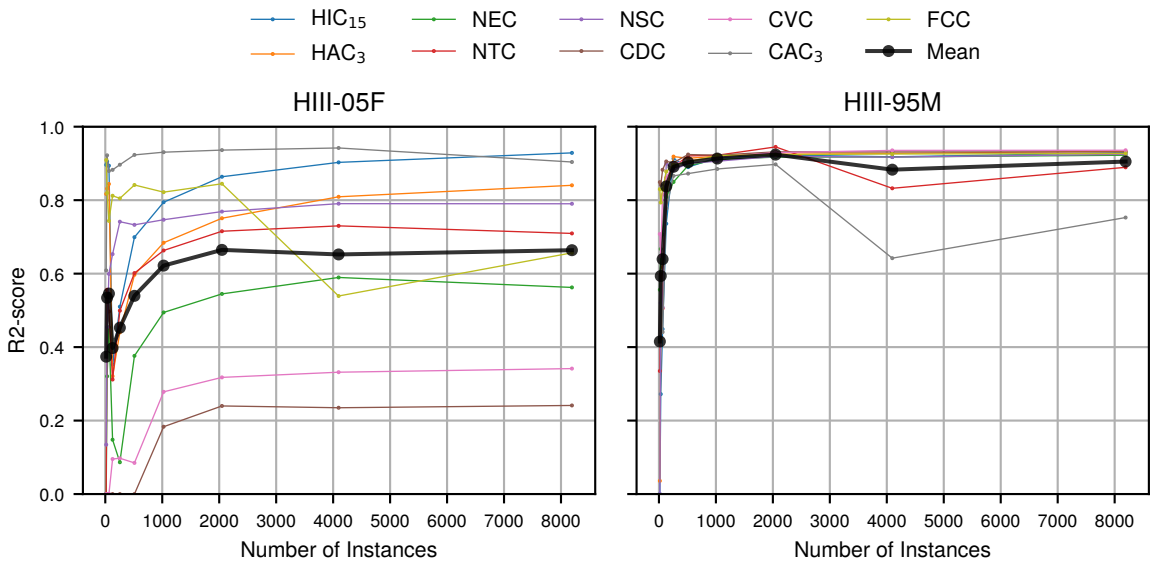


Figure 4.11: Number Instances and R2-scores of Target Attributes Regression of Injury-criteria

To gain a more profound insight into the data distribution, it is important to be aware that in fact not the distribution of the DoE factors are relevant for the prediction but the distribution of the responses from the FE-simulations of the Hybrid III 50th Percentile Male Dummy (HIII-50M).

An approach to characterize the HIII-50M responses can be using the Principal Component Analysis (PCA). The key kinematics of the Anthropomorphic Testing Device (ATD) were selected – the global accelerations and local displacements of the Centres of Gravity (CoGs) of the head, chest, and pelvis. Additionally, the chest acceleration and the femur forces were considered central. The time-series were resampled from 10 kHz to 1 kHz using the method as described above. Each of the 140 samples of each of the 25 channels were treated as independent feature attributes, resulting in 3500 attributes. Using the implementation of linear PCA in SCIKIT-LEARN, the first four Principal Components (PCs) account for approx. 80 % of the variance in the data.

In figure 4.12 on the facing page are the first 4 PCs as bimodal plots presented. Each column represents a different number of instances resulting from the Sobol exponents. Obviously, the point density increased with the number of instances. Especially in the fourth PC, the point density for values over 5 is significantly lower than for values lower than 5. The corresponding values of the loss-function $^{\text{Absolute Error}}\mathcal{L}$ from equation 4.1 on page 149 indicate that the dependency of the estimation performance depends on the local density. In combination of the 3$^{\text{rd}}$ and 4$^{\text{th}}$ PC is this effect the best noticeable.

As an alternative approach to assess the data distribution, in figure 4.13 on page 174, the discrepancy of the space coverage was calculated using the implementation in SCIPY. The discrepancy measures the distance of a point in the attribute space from an ideal continuous uniform distribution. For the Sobol algorithm, the discrepancy in the factor space decreases with the number of instances. This observation is plausible, since the progressing Sobol DoE keeps the factor ranges while increasing the density. The HIII-50M responses in the PC space, however, do not cover the space well. Hence, the increased local density does not improve the space coverage and consequently, the discrepancy is reduced only minimal. This intuition is supported by the observation that the discrepancy increases with the number of considered PCs. The discrepancy for 8192 instances and 4 PCs was 0.16 while for 22 PCs, the discrepancy was 3.67. Also in figure 4.12 on the next page, the decrease in space coverage in the form of increased local density is visible.

In summary, a rapid improvement in the arithmetic mean of the R2-scores was observed for both ATDs. However, a single injury-criteria benefited from an increased number of instances. The main reason for the overall stagnating R2-score was the sketchy coverage of the PC space representing the feature space. With the increasing number of instances, formerly only sparsely populated areas get more instances. First, this leads to a drop of the scores, but later to improvements.

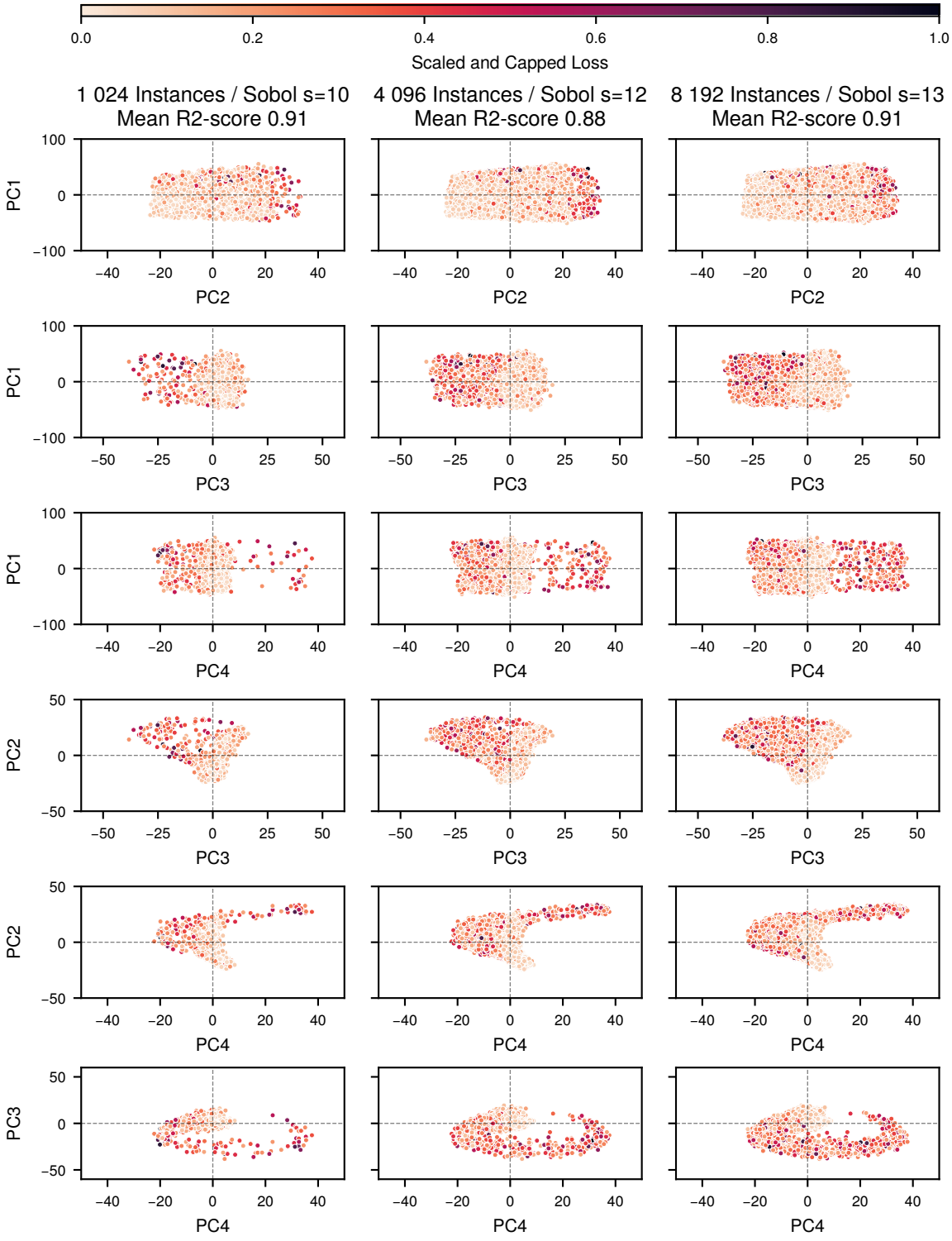


Figure 4.12: First Development-set 4 Principal Components (PCs) of HIII-50M Characteristics for Different Number of Instances by Sobol s with Loss of Target Attributes Regression for HIII-95M

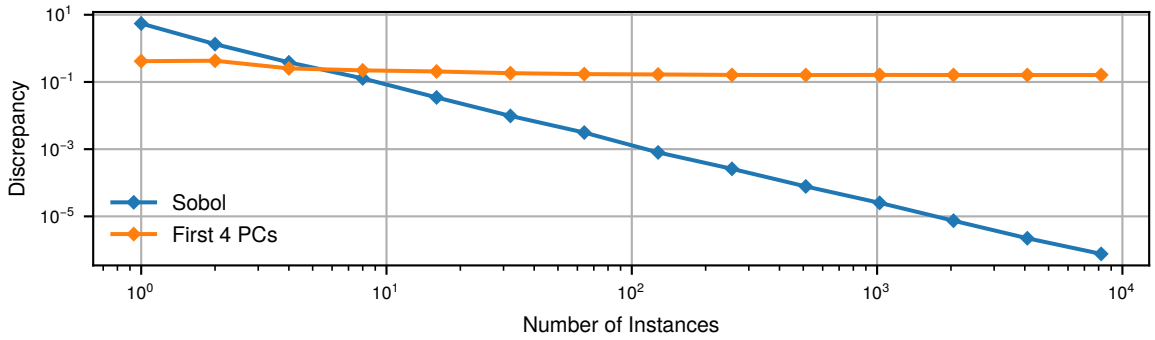


Figure 4.13: Space Coverage with 4 PCs versus DoE factor Space for Different Number of Instances

4.6.2 Adaptive Querying Strategies

4.6.2.1 Method for Adaptive Querying Strategies

For the presented research, the databases were generated in advance. Of course, not all instances contribute by the same amount to the metamodel fit. In other words, some FE-simulations were not necessary. In a practical application, those not-contributing FE-simulations account for an unnecessary cost. The adaptive DoE and active learning approaches introduced in section 2.3 on page 60 ff promise a solution.

An example of a typical implementation of adaptive DoEs was proposed by Ballal et al. (2023). DoE factors were used to fit a Gaussian Process Regression (GPR) metamodel explained in section 2.2 on page 49 using a set of instances with an existing ground truth. A second set of instances without ground truth and with grid-like spaced factors was used to determine the uncertainty of the metamodel. For instances with the highest uncertainties, a new ground truth would be generated by FE-simulations. Those instances become part of the training-set.

However, the method using uncertainties was not applicable to the problem discussed in this dissertation. Unlike in the above case, here, the DoE factor space was not relevant for the predictions but the responses from the FE-simulations of the HIII-50M as latent space. As discussed above to figure 4.12 on the preceding page, the HIII-50M responses may be a function of the DoE factors, but their space coverage was different.

The method using GPR would still be applicable for extracted feature attributes. The quality of the mapping into the latent space is dependent on the training-set itself for learnable feature extraction attributes and statistical methods like PCA. After each fit of the ANN, its feature attributes extractor must be applied as a transformer on the FE-simulations of the HIII-50M.

The subsequent training of the GPR is again limited to target attributes regression. To calculate the uncertainties, the database must contain as many FE-simulations as possible with HIII-50M and as an initial batch, few FE-simulations with HIII-05F and HIII-05F.

The limitation of the metamodel output type and the dependence of the representation of the

HIII-50M latent space on the number of instances with existing ground truth, are shortcomings of the GPR based adaptive query strategies. A relatively simple alternative would be the utilization of the loss-function on the validation-set.

The proposed approach assumes that two instances which are near to each other in the DoE factor space are near in the latent space of the HIII-50M and the target spaces of HIII-05F and HIII-95M, too. Second, it is assumed that an instance with a low value of its loss-function, equals an area in the latent space of high uncertainty. Consequently, adding more instances in such an area to the training-set could improve the metamodel performance.

The declared assumption makes an extensive database with FE-simulations of HIII-50M not necessary, since the new instances were picked directly from a DoE. The generalized method is proposed as follows:

1. Generate a DoE with Sobol scheme and an exponent ${}^{\text{Main}}s$ of at least 5.
2. Calculate the distance matrix for the DoE experiments using the factor values.
3. Generate an initial sub-DoE with $4 \leq {}^{\text{Initial}}s < {}^{\text{Main}}s$.
4. Generate experiments as instances with ground truth for initial sub-DoE (FE-simulations with HIII-05F, HIII-50M, and HIII-95M) as initial development-set.
5. Train metamodel and evaluate loss-function ${}^{\text{Absolute Error}}\mathcal{L}$ from equation 4.1 on page 149.
6. Identify the ${}^{\text{Worst}}\mathcal{S}$ instances with the lowest loss-function values.
7. Identify the ${}^{\text{Seed}}\mathcal{E}$ experiments not used for training in the DoE with the nearest distance in the DoE space to each in the previous step identified instance.
8. Generate the ground truth for the identified ${}^{\text{Worst}}\mathcal{S} \cdot {}^{\text{Seed}}\mathcal{E}$ instances by FE-simulations with HIII-05F, HIII-50M, and HIII-95M and add them to the development-set.
9. Repeat from step five until all experiments are used or an end criterion like budget is reached.

The limitation of ${}^{\text{Main}}s$ is due to the 5-fold cross-validation strategy within the framework. Of course, the DoE should be significantly bigger than the initial number of instances. However, a very high density of instance attributes can compromise the efficiency of the method. The distance calculation can be performed using the Euclidean norm. The factors should be normalized.

The simplest alternative for a query algorithm would be the incremental increase in the number of experiments as DoE, like it was studied above (see page 171). Of course, this approach suffers from potentially unnecessary FE-simulations. Nevertheless, it can serve as a reference for the adaptive querying study.

4.6.2.2 Study with Adaptive Querying Strategy

The adaptive querying study was performed using the final CNN. The initial Sobol exponent ${}^{\text{Initial}}s$ was 4. The neighbouring number of experiments ${}^{\text{Seed}}\mathcal{E}$ and the number of instances ${}^{\text{Worst}}\mathcal{S}$ were set equal to 15. The resulting number of queried experiments of 225 was selected to study the effects while keeping the needed computation time reasonable.

In figure 4.14, the result is presented. For both percentiles, R2-scores were stabilised around 2000 instances. Like in previous studies, the HIII-05F R2-score for the individual injury-criteria varies, while the HIII-95M R2-scores are very similar. The HIII-05F CDC, CVC, NEC, and NTC stabilize around 3000 to 4000 instances. All scores oscillate for small numbers of instances. The oscillations and the later increase in the R2-scores of some injury-criteria had only a small influence on the arithmetic mean for the HIII-05F. In contrast, the R2-scores for the HIII-95M stabilized earlier.

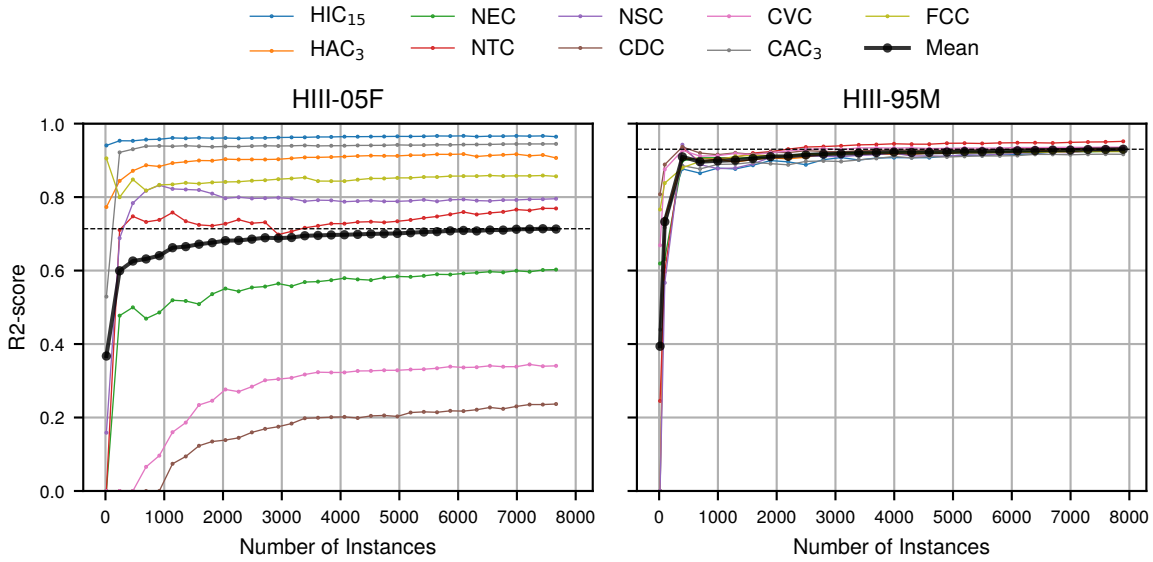


Figure 4.14: R2-score on Increasing Number of Instances in the Validation-set by Adaptive Querying in Target Attributes Regression Task

After the stabilization, the R2-scores improved monotonously with an increasing number of instances for both percentiles. In comparison, the simple usage of larger DoEs was discussed above and presented in figure 4.11 on page 171. In this study, a decline of the R2-score was observed around 4000 instances. The monotonous improvements seem very desirable as it enables a granular control on the utilized resources. Additionally, the results can appear to be more reliable.

The figure 4.15 on the facing page, the development of the local instance density is presented. Similar to figure 4.12 on page 173, the space is expressed in PCs, calculated from the results of the FE-simulations in the whole development-set. For the diagrams, only such instances were printed, which were used in the respective querying iteration. Starting with the 16 instances from the DoE with Sobol s equal to 4, the R2-score improved by adding the first 225 instances dramatically. The general shape of the different dimensions is already recognizable. Compared for approx. 1000 instances to the DoE based increase in the database from figure 4.12, the 4^{th} PC is denser populated when the adaptive querying is used. For approx. 4000 instances no difference is recognizable despite the R2-score is with 0.93 significantly higher for adaptive querying than the 0.88 in the DoE study.

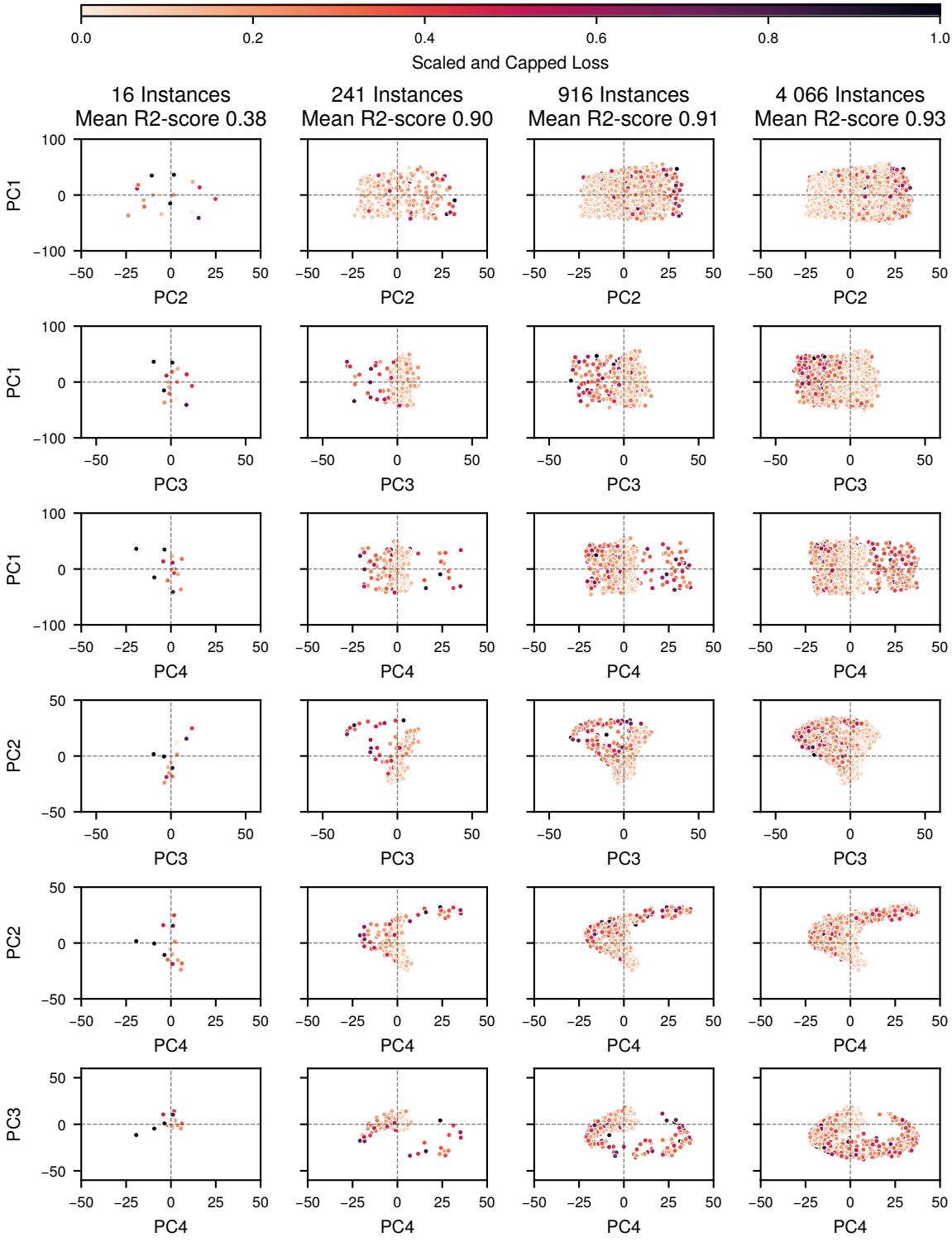


Figure 4.15: First Development-set 4 Principal Components (PCs) of HIII-50M Characteristics for Different Number of Queried Instances with Loss for HIII-95M

4.6.3 Transfer Learning Inspired Approaches

4.6.3.1 Study on Homogeneous Transfer Learning

Description of the Homogeneous Domain Transfer

In section 2.3 on page 63 ff were strategies of transfer learning discussed and summarized in figure 3.41. One of such strategies was used already in the studies above: the multitask learning. The output of one feature attributes extractor is passed to multiple blocks of dense layers. For the regression of values of injury-criteria, this approach was found to be inferior to having a single dense block feeding the output layer.

Another approach described in section 2.3 was the homogeneous transfer between domains. In this work, the domain of FE-simulations with Hybrid III (HIII) is used as a source and the domain of FE-simulations with VIRTHUMAN (VHUM) as transfer. A feature attributes extractor with fitted hyperparameter should be transferred for predictions within the VHUM domain. The approach uses the same architecture for both domains. First, the whole ANN is fitted with the data from the HIII. Second, the parameters from the feature attributes extractor are saved. Next, the second ANN is generated and the parameters of its feature attributes extractor block frozen. The now untrainable parameters are overwritten by those of the HIII metamodel. Afterwards, the ANN is trained with the VHUM data. In a last step, the feature extractor is unfrozen, and the whole ANN continues fitted with a smaller learning rate.

To assess whether the transfer learning approach is beneficial, a set of 3 trials was conducted. The first trial used the transfer-set with VHUM and a fully trainable ANN to establish a reference. In the second trial, the development-set with HIII was used to train another metamodel. This metamodel feature attributes extractor was used in the last trial, the procedure described above utilizing the pre-trained feature attributes extractor to fit a metamodel to the transfer-set. Is the performance of the last trial better than the one from the first, the transfer learning approach would be beneficial. The architecture was the CNN with the hyperparameters from above.

Findings of the Homogeneous Domain Transfer Study

The results of the transfer learning are presented in table 4.10 on the facing page. The R2-scores were calculated for the target attributes regression on the training-set (seen) and the validation-set (unseen). For each injury-criterion, there is the score for a direct prediction. “Direct” means, the whole ANN was trained using only the transfer-set with VHUM data. For the predictions in the “transfer” rows, the feature attributes extractor was first trained using the entire HIII based development-set and then frozen during the training with the transfer-set. As described above, in the fine-tuning step, the feature attributes extractor was unfrozen. The learning rate during the first training-steps was 1×10^{-5} and for fine-tuning 1×10^{-6}.

Table 4.10: Transfer Learning with VHUM-50M as Input – R2-scores for Target Attributes Regression Task for Injury-criteria

Region	Injury-criterion	Mode	VHUM-05F			VHUM-95M		
			Seen	Unseen	Diff.	Seen	Unseen	Diff.
Head	HIC_{15}	Direct	0.47	0.37	0.10	0.93	0.73	0.20
		Transfer	0.30	0.29	0.01	0.73	0.76	0.03
		Fine Tune	0.31	0.28 ↓	0.03	0.89	0.74 ↑	0.13
	HAC_3	Direct	0.45	0.34	0.11	0.95	0.79	0.16
		Transfer	0.26	0.25	0.01	0.90	0.77	0.13
		Fine Tune	0.26	0.25 ↓	0.01	0.90	0.77 ↓	0.13
Neck	NEC	Direct	0.98	0.82	0.16	0.88	0.56	0.32
		Transfer	0.73	0.69	0.04	0.79	0.57	0.22
		Fine Tune	0.77	0.71 ↓	0.06	0.78	0.60 ↑	0.18
	NTC	Direct	0.99	0.96	0.03	0.99	0.91	0.08
		Transfer	0.93	0.90	0.03	0.97	0.85	0.12
		Fine Tune	0.95	0.91 ↓	0.04	0.97	0.87 ↓	0.10
	NSC	Direct	0.98	0.91	0.07	0.97	0.85	0.12
		Transfer	0.86	0.80	0.06	0.94	0.78	0.16
		Fine Tune	0.91	0.82 ↓	0.09	0.94	0.77 ↓	0.17
Chest	CDC	Direct	0.86	0.64	0.22	0.92	0.60	0.86
		Transfer	0.70	0.59	0.11	0.82	0.59	0.23
		Fine Tune	0.73	0.60 ↓	0.13	0.83	0.58 ↓	0.25
	CVC	Direct	0.56	0.40	0.16	0.71	0.03	0.68
		Transfer	0.38	0.32	0.06	0.42	0.11	0.31
		Fine Tune	0.42	0.30 ↓	0.12	0.43	0.09 ↑	0.34
	CAC_3	Direct	0.96	0.90	0.06	0.98	0.92	0.06
		Transfer	0.92	0.89	0.03	0.96	0.91	0.05
		Fine Tune	0.92	0.89 ↓	0.03	0.96	0.91 ↓	0.05
Femur	FCC	Direct	0.84	0.67	0.17	0.82	0.44	0.38
		Transfer	0.66	0.62	0.04	0.75	0.52	0.23
		Fine Tune	0.69	0.65 ↓	0.04	0.75	0.53 ↑	0.22
Total	Median	Direct	0.86	0.67	0.22	0.93	0.73	0.20
		Transfer	0.70	0.62	0.08	0.90	0.76	0.14
		Fine Tune	0.73	0.65	0.08	0.89	0.74	0.15

A first analysis of table 4.10 reveals a general tendency of the metamodel for overfitting. The difference of the median between the VHUM training-set and validation-set calculated for all injury-criteria is for the VIRTHUMAN 5th Percentile Female Human Body Model adapted to HIII-05F (HIII-05F) reduced by 0.2 and for the VIRTHUMAN 95th Percentile Male Human Body Model adapted to HIII-95M (HIII-95M) by 0.06 with the pre-trained feature extractor. The fine-tuning led partially to an again bigger difference between the R2-scores. These improvements indicate a reduction of the overfitting, and therefore a positive effect of the transfer learning. However, in most cases, the actual R2-score was lower with transfer learning than for the direct prediction.

One reason for non-successful transfer learning applications is the mismatch between the source and target domains. In such a case, the learnt attributes from the source domain would render meaningless. The data analysis in chapter 3 on page 143 ff, indicated that the linked domains seemed reasonably similar. The similarities in the scores of most injury-criteria between the direct learning and the transfer learning before fine-tuning support this point.

Another intuitive explanation for the absence of clear improvement can be the relatively low complexity of the data. In typical cases of successful applied transfer learning, the highly complex feature attributes extractor was trained using millions of instances. For example, the popular AlexNet described by Krizhevsky et al. (2017) was trained by approx 1.2 million images. In contrast, the comparably simple feature attributes extractor trained by only a few thousand of instances seems not to carry information which cannot be extracted from the transfer-set directly also.

4.6.3.2 Alternative Method – Direct Prediction of VIRTHUMAN Responses

Description of the Direct Prediction of VIRTHUMAN Responses by Hybrid III

Similarly to the above approaches and inspired by a study of Plaschkies (2023), the input data of the HIII should be used directly to estimate the responses from FE-simulation of VHUM. The HIII represents, as in the transfer learning study, a FE-model of low degree of detail and, hence, low computational cost. The VHUM represents a highly complex and computationally expensive FE-model. As this is a feasibility study, especially the VHUM is a placeholder for a real, more complex and costlier FE-model like the Test Device for Human Occupant Restraint (THOR) or the Total Human Model for Safety (THUMS).

This approach requires pairwise FE-simulations of both ATDs. The development-set was generated using the Sobol method, while the transfer-set used the full factorial design. The continuous nature of the first database and the discrete nature of the second database are incompatible. Hence, the alternative development-set had to be used. In this DoE, factor values were generated in 5 levels, while the transfer-set used 5 levels. However, since the factor ranges matched, the equal environment configurations could be extracted from the alternative development-set.

The CNN was used in the configuration developed above in section 4.3 on page 154 ff. As the metamodel input the HIII-50M was used for the target attributes regression of injury-

criteria of VHUM-05F, VIRTHUMAN 50th Percentile Male Human Body Model adapted to HIII-50M (HIII-50M), and VHUM-95M. For comparison, the HIII-05F was used for target attributes regression of VHUM-05F and the HIII-95M for the VHUM-95M. The drawback of the percentile-matched approach is that the FE-simulations of all three percentiles must be available, instead of only the FE-simulations with HIII-50M.

Findings from the Direct Prediction of VIRTHUMAN Responses by Hybrid III

The direct prediction between the same percentiles of HIII and VHUM is presented in table 4.11 on the next page. Input into the metamodel is the FE-simulation results of HIII and output the FE-simulation results of VHUM. The baseline is again the constant prediction of the median of the target data. The data distribution did not cause the capped baseline R2-score to get any bigger than 0.

In the R2-scores on the validation-set for the individual injury-criteria and pairings, cases may reach a satisfactory level of over 0.8. However, only for the target attributes regression of the VHUM-05F Head Injury Criterion for up to 15 ms (HIC_{15}) and CAC_3, no indication for overfitting was observed. A comparison with the cross-plots in figure 4.16 on page 183 shows that the head acceleration related injury-criteria of the 5th percentile has the simplest interrelation. For other injury-criteria, highly non-linear to not existing interrelations between the two human surrogates are observed and reflected in the R2-scores of table 4.11 on the next page

The above concept requires for a prediction for a percentile of VHUM the corresponding FE-simulation with HIII. A higher usage value would be achieved if the prediction required only the HIII-50M as input. The results are presented in table 4.12 on page 184. Overall, the performance seems similar for both concepts.

The target attributes regression task of injury-criteria values of VHUM using HIII was successful for the acceleration of head and chest related injury-criteria. Additionally, the prediction of neck related injury-criteria for the VHUM-05F can be called a success with the limitation of a light indication for overfitting. All other tasks tested suffered from insufficient R2-scores or significant overfitting.

Table 4.11: HIII to VHUM with Matched Percentiles – R2-score for Target Attributes Regression Task for Injury-criteria

Region	Injury-criterion	HIII	VHUM	Baseline	Training-set	Validation-set	Difference
Head	HIC_{15}	05F	05F	0.00	1.00	0.98	0.02
		50M	50M	0.00	0.97	0.91	0.06
		95M	95M	0.00	0.89	0.80	0.09
	HAC_3	05F	05F	0.00	0.99	0.97	0.02
		50M	50M	0.00	0.95	0.91	0.04
		95M	95M	0.00	0.91	0.82	0.09
Neck	NEC	05F	05F	0.00	0.98	0.94	0.04
		50M	50M	0.00	0.77	0.55	0.22
		95M	95M	0.00	0.81	0.48	0.33
	NTC	05F	05F	0.00	0.99	0.93	0.06
		50M	50M	0.00	0.93	0.82	0.11
		95M	95M	0.00	0.98	0.91	0.07
	NSC	05F	05F	0.00	0.99	0.87	0.12
		50M	50M	0.00	0.94	0.74	0.20
		95M	95M	0.00	0.97	0.86	0.11
Chest	CDC	05F	05F	0.00	0.89	0.69	0.20
		50M	50M	0.00	0.85	0.56	0.62
		95M	95M	0.00	0.88	0.58	0.29
	CVC	05F	05F	0.00	0.61	0.46	0.15
		50M	50M	0.00	0.93	0.85	0.08
		95M	95M	0.00	0.65	0.12	0.53
	CAC_3	05F	05F	0.00	0.97	0.91	0.06
		50M	50M	0.00	0.95	0.88	0.07
		95M	95M	0.00	0.98	0.94	0.04
Femur	FCC	05F	05F	0.00	0.87	0.62	0.25
		50M	50M	0.00	0.74	0.37	0.37
		95M	95M	0.00	0.79	0.51	0.28
Total	Median	05F	05F	0.00	0.98	0.91	0.07
		50M	50M	0.00	0.93	0.82	0.11
		95M	95M	0.00	0.89	0.80	0.09

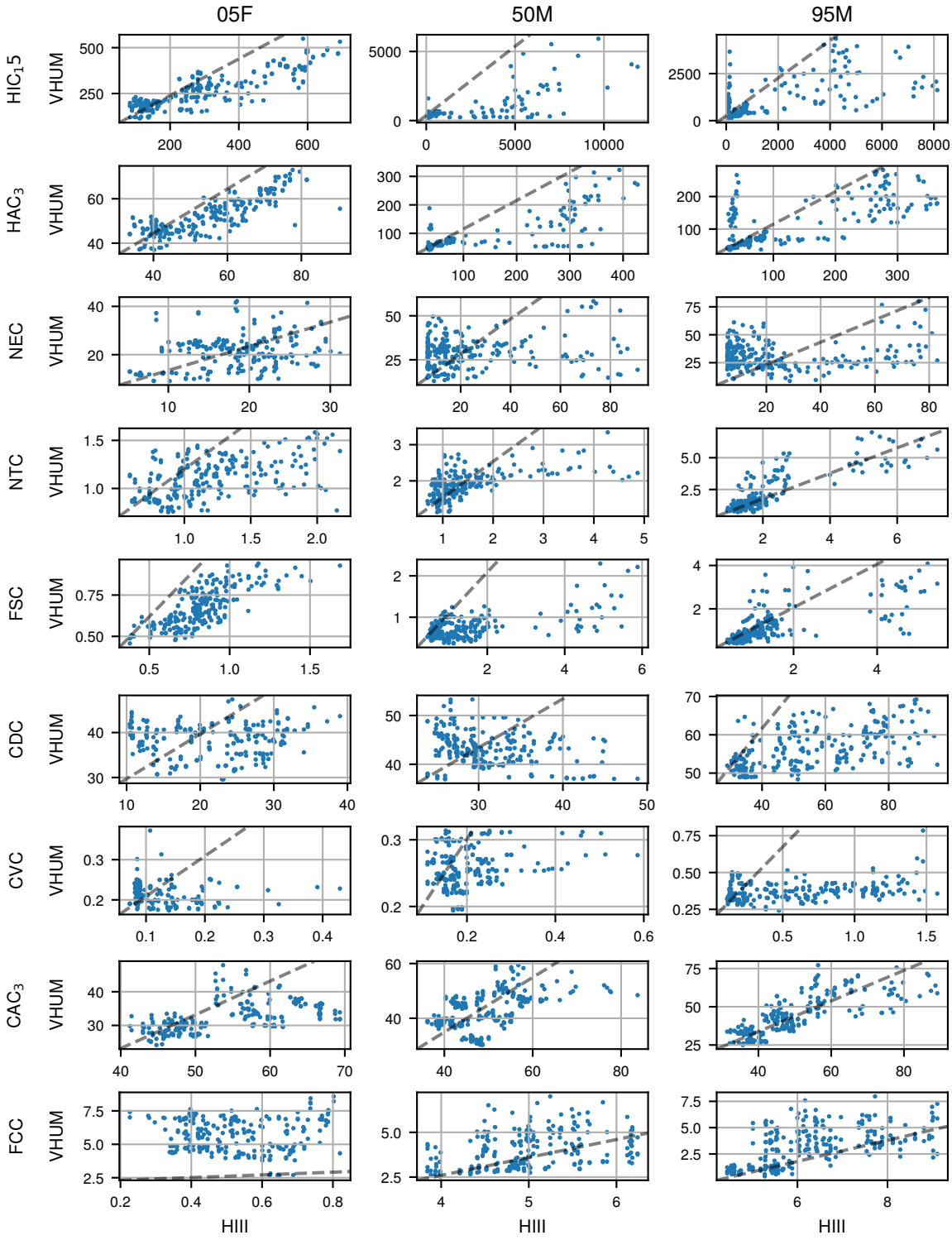


Figure 4.16: Cross Plot VHUM over HIII for the Ground truth Injury-criteria

Table 4.12: HIII to VHUM with HIII-50M as Input – R2-scores for Target Attributes Regression Task for Injury-criteria

Region	Injury-criterion	Output	Training-set	Validation-set	Difference
Head	HIC_{15}	VHUM-05F	1.00	0.99	0.01
		VHUM-95M	0.91	0.82	0.09
	HAC_3	VHUM-05F	0.99	0.98	0.01
		VHUM-95M	0.93	0.85	0.08
Neck	NEC	VHUM-05F	0.98	0.93	0.05
		VHUM-95M	0.84	0.53	0.31
	NTC	VHUM-05F	0.99	0.96	0.03
		VHUM-95M	0.98	0.90	0.08
	NSC	VHUM-05F	0.98	0.90	0.08
		VHUM-95M	0.95	0.83	0.12
Chest	CDC	VHUM-05F	0.88	0.68	0.20
		VHUM-95M	0.90	0.62	0.28
	CVC	VHUM-05F	0.58	0.45	0.13
		VHUM-95M	0.77	0.33	0.44
	CAC_3	VHUM-05F	0.95	0.88	0.07
		VHUM-95M	0.98	0.92	0.06
Femur	FCC	VHUM-05F	0.82	0.62	0.20
		VHUM-95M	0.78	0.49	0.29
Total	Median	VHUM-05F	0.98	0.90	0.08
		VHUM-95M	0.91	0.82	0.09

4.6.3.3 Summary and Conclusion for the Transfer-set

The transfer learning approach of the homogeneous domain transfer reduced the tendency of the metamodel to overfit the data in the transfer-set. For most injury-criteria were R2-scores smaller than 0.8 achieved, which could not be improved by the transfer approach. In conclusion, the minimal reduction of overfitting cannot justify the effort to create a source database for the transfer.

As a reason for the ineffectiveness of the transfer learning, the relatively low complexity of the CNN feature extractor was pointed out. Another reason was the overfitting. Strategies against overfitting are discussed many times. For example, Hastie et al. (2009) presented approaches directly related to different learning algorithms. In summary, common strategies are (i) the selection of less respective better features, (ii) the increase in the regularization, and (iii) the adding of more instances. The feature selection in terms of channels and injury-criteria was done during the hyperparameter tuning in section 4.3 on page 156 ff. Additional attributes as inspected in section 4.4 on page 163 ff did not support the metamodel. The studies to compare

the number of instances in section 4.6 on the pages 171 ff and 174 ff revealed a necessary number of instances of 500 to 1000. In comparison, the transfer-set with the number of instances of approximately 240 instances leaves room for improvement. Due to time constraints and the dependency on external partners in the framework of the AWARE2ALL project, it was not possible to generate more instances.

The direct cross-domain prediction from HIII to VHUM seemed to be successful for injury-criteria related to the head and chest acceleration of all percentiles, and the neck of the VHUM-05F. The other target attributes regression tasks could not be solved sufficiently. This shortcoming can be partially reasoned by the low number of instances in the transfer-set, too. Furthermore, especially the chest compression appears to be too different between the two types of human surrogate.

4.7 Discussion of Learning Approaches and Outlook

4.7.1 Overview, Framework, and Research Questions

In this chapter, multiple studies were conducted as an attempt to answer the research questions. The evaluation of the estimation of individual injury-criteria provided an insight to the capabilities to cover the FE-model non-linearities in a metamodel. The studies on the number of instances, the adaptive experiment querying strategy, and the utilization of the transfer-set illustrate different aspects of efficient data generation and utilization.

The developed framework proved invaluable over the course of the work. The standardized evaluation procedure allowed for an efficient and flexible development and testing of the machine learning modules. The automated check of the database's version prevented failures when working on different workstations. Furthermore, the standardized documentation of hyperparameters and main results allowed the reliable evaluation of the studies. The framework base was developed and tested during the arTIco[1] project together with the IDIADA Fahrzeugtechnik GmbH[2] by a small team of engineers and later adapted by a single researcher for this work. The applicability of the framework for larger teams must be proven in future projects.

The major research of this chapter used the development-set. A baseline was established for the regression of target attributes and target multivariate time-seriess, and for the classification with different numbers of classes. Different strategies to extract from the feature multivariate time-seriess meaningful feature attributes were compared. All extracted feature attributes were passed through dense layers for the final predictions.

[1] https://www.thi.de/forschung/carissma/c-isafe/artico/
[2] https://www.applusidiada.com/

4.7.2 Feature Attribute Extraction Strategies

As deterministic database independent and domain-agnostic methods, the pre-calculation of injury-criteria, the TSFRESH attributes, and the CATCH22 attributes were considered. The latter two resulted in predictions on a level of random guessing. Since the authors Christ et al. (2018) and Lubba et al. (2019) claimed excellent prediction performances in their databases, future work should analyse the differences in those databases and the development-set. The use of domain-knowledge in the form of pre-calculated values of the injury-criterion produced acceptable results, but were as alone-features in context of a hyperparameter tuning study inferior to other approaches. Similar rated was the randomized database-independent ROCKET transformer. All studies with the database-independent feature attributes extractors were conducted using a fixed Artificial Neural Network (ANN) as the actual learning algorithm. Small improvements can be expected if its hyperparameters are tuned for the specific architecture.

The deterministic and database-dependent method was tested with the feature attributes extraction by Principal Component Analysis (PCA). The measured performance was competitive. Like for the database-independent methods, small improvements can be expected if its hyperparameters are tuned with the specific architecture. Overall, the performance in combination with the computational cost was inferior to other methods.

From the algorithm group of feature attribute extractors with learnable parameters, the Convolutional Neural Network (CNN) and the Long Short-Term Memory (LSTM) based Recurrent Neural Network (RNN) were selected. For both approaches was an hyperparameter tuning study performed and significant tendencies stated. The Tree-structured Parzen Estimator (TPE) algorithm used to generate trials for the CNN limits the statistical evaluation of the effects of hyperparameter, since they are not independent of each other. In fact, this evaluation is performed within the algorithm. Furthermore, the tournament-style hyperparameter tuning study used for the LSTM based RNN does not allow the evaluation of interrelations between the hyperparameters. Despite these shortcomings, both tuning algorithms were found to be sufficient given the computational cost of one trial.

Alone due to the significantly higher computational cost of 1 h to 2 h for the CNN versus the around 8 h to 20 h of the LSTM based RNN led to the assessment of the CNN method as superior. Compared with the other approaches, the CNN was comparable in terms of its computational cost and superior in terms of estimation performance. The infusion of domain knowledge in the form of additionally provided pre-calculated injury-criteria did not benefit the metamodel significantly.

4.7.3 Estimation of Different Targets and Test-set Assessment

With the tuned CNN, an overall satisfactory performance for the target attributes regression of the injury-criteria values of the HIII-95M was achieved. More mixed results were observed for the HIII-05F. The predictions for the head related injury-criteria and the Chest Acceleration over 3 ms Criterion (CAC_3) were found to be acceptable. For the Chest Deflection Criterion (CDC) and the directly related Chest Viscous Criterion (CVC), only low values of the metric-

function were observed. The instable seatbelt behaviour at the HIII-05F neck was identified as the reason. The predicted values for the CDC were shown to be as they would be when the instable behaviour would not exist. The instable seatbelt behaviour affected the reliability of the neck kinetic as well. This non-linearity could not be covered by the metamodel. On the contrary, since this behaviour is partially an artifact of the FE-model, this shortcoming may be more related to the development-set itself. In future work, the affected FE-simulations should be replaced. Alternatively, another metamodel could be trained to estimate the HIII-05F instability and provide this information to the user. Unfortunately, due to the instable nature of the problem, the first attempts were not successful.

Although it did not appear in the statistical evaluation of the metamodel performance for the HIII-95M, this bigger Anthropomorphic Testing Device (ATD) suffered from hits on the dashboard and the resulting peaks in the channels. The metamodel tends to predict lower values than those generated by such peaks. Similarly to the above limitation of the HIII-05F, in future work, an additional tool should be developed to indicate the risk of hitting. A potential approach might be the evaluation of the distance between the head and the dashboard.

The extended predictions showcased the flexibility of the ANN architecture. The target attributes classification task was successful for some injury-criteria, showing the general capability of the metamodel. There were many classes with only a few to no instances. With this shortcoming of the development-set, no learning is possible. In future development, it can be attempted to predict the factor values to generate loads on the human surrogate because they are necessary to belong to the under-represented class. The adaptive querying strategy, tested in this work for the regression task, has potential in this context, too. The thesis performed by Salike (2024) and co-supervised by the author highlighted the challenge of such a task, especially if there are very few to no instances in a class.

For the prediction of target multivariate time-series an overall satisfactory performance was achieved. Due to the channel averaging nature of the used metric-function the peaks of the HIII-95M and the instable seatbelt of the HIII-05F did not significantly affect the metric-function values. In future work, the ISO-rating shall be tuned further to the specifics of the individual channels.

The assessment of the FE-model used for the development-set in combination with the trained metamodel on the testing-set contained multiple challenges. As expected, the deviation between the FE-models increased the inaccuracies of the metamodel. Additionally, the intended evaluation of the factor spaces which can be associated with an interpolation or respectively an extrapolation lowered the values of the metric-function further. Due to the randomized nature of the Sobol Design of Experiments (DoE), both spaces overlap. An improvement could potentially be achieved, when the testing-set DoE match the factor values ranges of the development-set. The extrapolation performance must be evaluated by performing a second testing-set DoE with the extended ranges only.

4.7.4 Efficient Data Utilization and Generation

Two approaches were followed to determine how many instances are actually necessary for a sufficiently trained metamodel, and if there is a more efficient way to generate the data. In the first approach, the main advantage of the Sobol DoE was leveraged. The sub-databases had the same statistical properties regarding the distribution of their factor values, as the whole development-set. In the second approach, a distance-based method for querying new instances for the training-set was successfully implemented and tested.

For both approaches, a rapid improvement of the evaluated metric-function was observed and a critical minimum number of instances of around 2000 established. To at least stabilize the assessed metric-function, approximately 500 instances were required. Unlike the simple increase of the DoE in size, with adaptive querying, a monotonous improvement of the performance of the metamodel with the increasing number of instances was observed. The identified major reason was the sketchy coverage of the feature space. With the increasing number of instances, formerly only sparsely populated areas get more instances. First, this leads to a drop in the evaluated metric-function, but later to its improvement. Since the adaptive querying utilized the local loss-function, the sparser areas seemed to get more populated early with enough instances.

Introducing the VIRTHUMAN (VHUM) FE-simulations as a transfer-set, the application of a FE-model with low computational cost for the estimation of results for a more expensive FE-model was mimicked. One study was conducted for the homogeneous transfer of the feature attribute extractor. Another study featured the direct prediction of the complex FE-model results from the simpler FE-model. As a baseline, a metamodel was trained using only the transfer-set. Due to the relatively small number of instances, overfitting was observed. The metamodel architecture and the hyperparameters values were the same as in the studies with the development-set where no overfitting was indicated. Hence, the main reason must be the much smaller number of instances in the transfer-set. Furthermore, the approximately 200 instances of the transfer-set are significantly less than the above discussed required 2000 instances. In future work, the size of the transfer-set should be increased.

The homogeneous transfer could reduce the overfitting slightly. On the contrary, most values of the evaluated metric-function worsened. The reasons were the mismatch of both domains and the relatively low complexity of the data compared to other successful transfer learning applications in combination with the relatively low number of instances.

The prediction across domains from Hybrid III (HIII) to VHUM seemed to be successful for injury-criteria related to the head and chest acceleration of all percentiles and to the neck of the VHUM-05F. The other tasks could not be solved sufficiently. This shortcoming can also be partially explained by the low number of instances in the transfer-set. Additionally, especially the chest compression appears to be too different between the two types of human surrogate.

5 Summary and Outlook

This thesis described the application of machine learning in the context of development and assessment methods in the domain of safety systems for occupant safety. Supervised learning methods were applied to generate rapid predictions for FE-simulations results in the domain of the assessment of occupant safety systems. The key concept revolved around the question of how a metamodel could be designed without explicit characterization of the vehicle. In this context, the following research questions were addressed.

1. Can a human surrogate response sufficiently characterize a vehicle?
2. To what extent limit the system's non-linearity the capabilities of the metamodel?
3. What are the cost and use aspects of such a machine-learning-based system?
4. How can the required data be generated and used efficiently?
5. Can domain knowledge be utilized to reduce the required amount of data?
6. What is the related work?

The related work was presented in chapter 2. The insights covered the topics of machine learning and occupant safety, with a special focus on human body diversity.

The base machine learning architecture used the result of a reference FE-simulation to predict the result of a FE-simulation in the same vehicle configuration but different occupant anthropometrics. In chapter 4, metamodels were generated and evaluated. The input for the metamodels are feature multivariate time-series. The learnable feature attributes extracting architectures using CNN and the LSTM based RNNs were compared. As an alternative to the learnable feature attributes extractors, data transformations were discussed. The vehicle characterizing a human surrogate was a 50^{th} percentile. The task was the target attributes regression of injury-criteria for 5^{th} and 95^{th} percentile. The task was extended by target attributes classification of different number of classes for the criticality of a crash and for target multivariate time-series regression of complete channels.

Overall, the predictions for the Hybrid III 5^{th} Percentile Female Dummy (HIII-05F) and Hybrid III 95^{th} Percentile Male Dummy (HIII-95M) using the Hybrid III 50^{th} Percentile Male Dummy (HIII-50M) as input were acceptable. The high values of the evaluated metric-function for e. g. the head acceleration and the HIII-95M chest deflection showcased the functionality of the base

machine learning architecture. The inability to sufficiently predict the chest deflection for the HIII-05F was reasoned by the instable seatbelt slip towards the neck. Additionally, there was some inaccuracy in predicting high and sharp peaks in the head acceleration of the bigger ATDs due to a hard contact of the head to the dashboard. Those two limitations represent typical non-linearities of occupant safety systems and correspond to the second research question.

The fourth and fifth research questions were directly related. The chapter 3 focused on efficient data generation by FE-simulations. Starting from the detailed FE-model of the Honda Accord 2024, using an incremental and structured approach, a heavily simplified FE-model version was derived. Each simplification step was reasoned and evaluated for the reduction of model quality in relation to the shorter computation time. The initial FE-model required 12 h on the cluster using 48 nodes for one FE-simulation. The final FE-model version required only 20 min on one single Central Processing Unit (CPU) while being comparable to the original FE-model version to some extent. Regarding the research questions, profound domain knowledge was the enabler for the simplifications with the proposed method. The method can be adapted in terms of order and stopping point depending on the available budget.

A second take on the fourth and fifth research questions was in chapter 4. A study on the influence of the number of instances implied a required minimum number of instances around 2000 to achieve an acceptable estimation performance for the studied data domain. The direct training with this size of database generated sufficient results. A more stable and monotonous progression of the monitored metric-function was shown for the adaptive querying strategy over the classic DoE approach.

Transfer learning approaches as a promising method to deal with limited data were also studied. Neither multitask learning nor the homogeneous domain transfer could improve the metamodel performance. Multitask learning was studied during the hyperparameter tuning of the ANNs. The homogeneous transfer learning could reduce the overfitting slightly. On the contrary, most of the values of the evaluated metric-function worsened. The direct cross-domain prediction from HIII to VHUM seemed to be successful for the head and chest acceleration of all percentiles and the neck related injury-criteria of the VHUM-05F. The other tasks could not be solved sufficiently. This shortcoming can be partially explained by the low number of instances in the transfer-set. Additionally, especially the chest compression appears to be too different between the two types of human surrogate.

Important studies derived from the presented research would be the prediction of FE-model instabilities such as the seatbelt slip and the independence from pre-generated and structured data. Furthermore, as a finding of the arTIco project published by Plaschkies et al. (2024a), the explainability of the metamodel would be crucial for a successful implementation. Moreover, the abstract representation of the vehicle's crash characteristics by the latent space of the reference ATD responses holds the potential of being applied to existing unstructured data. In future work, the studies on the predictions between a simple and a complex FE-model should be repeated and extended with significantly more data.

The generation of high-quality labelled data emerged as the major challenge when introducing machine learning in the domain of assessment methods for passive occupant safety. In comparison, the application of machine learning algorithms appeared to be relatively convenient, thanks to highly developed libraries in PYTHON. For future development, the effort to generate a reasonable database should be of high priority in budgeting a related project.

From this research, multiple databases were published. All databases were released under the Creative Commons Attribution 4.0 International licence.

- The development-set based on the Sobol algorithm and HIII with 24 576 FE-simulations was published by Plaschkies (2024f).
- The alternative development-set based on the full factorial scheme and HIII with 9375 FE-simulations was published by Plaschkies (2024d).
- The testing-set based on the Sobol algorithm and HIII with 768 FE-simulations was published by Plaschkies (2024e).
- The validation-set based on the full factorial scheme and VHUM with 729 FE-simulations was published by Plaschkies and Müller (2024).

Furthermore, the publishable source code was released by Plaschkies (2024a). The machine learning results were published by Plaschkies (2024b).

The presented method for an incremental simplification of the FE-model and the detailed analysis of each step can be a valuable guideline for engineers and researchers. A reasonable FE-model simplification is a central task in the virtual assessment of vehicle safety systems. The insights from the analysis can provide inspiration for simplifications and prevent potential pitfalls. The FE-models from each simplification step, published by Plaschkies (2024c), can be used as a base for the further creation of databases.

Pre-Publications and Participations

Pre-Publications

- Plaschkies, F. and Vaculín, O. (2020): "Estimation of the Impact of Human Body Variation on its Crash Behavior Using Machine Learning Methods". In: *FISITA Web Congress* (Online). DOI 10.46720/F2020-PIF-051.
- Plaschkies, F., Vaculín, O., and Schumacher, A. (2021): "Assessment of the Influence of Human Body Diversity on Passive Safety Systems. A State-of-the-art Overview". In: *FISITA World Congress* (Online). DOI: 10.46720/F2021-PIF-071.

  Parts of the chapter 2 were extracted from this publication. The subsubsections 2.1.2.1, 2.1.2.2, 2.1.4.1, and 2.1.4.2 and the section 2.4 were with minimal adjustments content of the publication. Furthermore, the subsections 2.1.2.3, 2.1.3.1, 2.1.4.4, 2.1.5.2 are based on the publication but were extended and updated. A small fraction of subsection 2.3.4 is based on the publication as well.
- Plaschkies, F.; Vaculín, O.; Pelisson, A. and Schumacher, A. (2022): "Schnelle Abschätzung des Crashverhaltens von Insassen unter Berücksichtigung der Vielfalt des Menschen: Robustheit, Datenintensität und Vorhersagekraft von Metamodellen". In: *VDI Fahrzeugsicherheit* (Berlin, Germany), DOI: 10.51202/9783181023877-313.

  Angelo Pelisson was an exchange student through the German Academic Exchange Service (DAAD) in the project enGlobe, which supported the research in the frame of the Bavarian Center for Applied Research and technology with Latin America (AWARE). He worked under the supervision of the author during his stay in the winter semester 2020/21.
- Plaschkies, F., Possoli, K., Vaculín, O., Schumacher, A., and Andrade Junior, P. d. (2023): "Evaluation Approach for Machine Learning Concepts in Occupant Protection Based on Multi-Attribute Decision Making". In: *27th International Technical Conference on the Enhanced Safety of Vehicles* (Yokohama, Japan).

  Ketlen Possoli was an exchange student through the DAAD in the project enGlobe, which supported the research in the frame of the AWARE. She worked under the supervision of the author during her stay in the summer semester 2022.

- Plaschkies, F.; Stocker, F.; Yuan, Z. (2024): “arTIco Framework”. In: *Zenodo* (Online) DOI: 10.5281/zenodo.11657448.

  The software framework was part of the project arTIco – Artificial Intelligence and Correlations, a research project in the funding line Digitalization of the Bavarian Cooperative Research Programme (BayVFP) of the Free State of Bavaria under the funding number DIK-2110-0025//DIK0404/02. The author was the researcher in the arTIco project. The design and implementation was majorly done by the first author. The co-authors contributed by continuous testing and feedback. The framework was the base for the framework used in chapter 4.

Contributions to Academic Education

The following master theses were conducted under the author’s guidance:

- Christoph Alexander Noll (2021): “Simulative Ermittlung eines dimensionsreduzierten Pulses für einen Frontaufprall mit Offset”. Masterthesis. Technische Hochschule Ingolstadt, Germany.
- Sai Srijan Salike (2024): “Data Abstraction and Analysis for Universalizing Passenger Cell Model Interface for Occupant Protection Optimization”. Masterthesis. Technische Hochschule Ingolstadt, Germany.
- Sai Sharan Ravichandran (2024): “Transparent Rating of Multivariate Time Series Classification Using Explainable AI”. Masterthesis. Technische Hochschule Ingolstadt, Germany.

The author was part of the programme “Simulations- und Testverfahren für automatisierte Fahrfunktionen” each summer semester between 2021 and 2024 at Technische Hochschule Ingolstadt. The content was the requirements for vehicles with different levels of automation, from type approval to the specification sheet regarding passive safety, active safety, and automated driving functions. Procedures regarding physical and virtual testing, simulation approaches, vehicle models, sensor models, and the validity of models were presented. The author was responsible for passive safety and all related procedures. The module was mandatory for students of the master course “Automatisiertes Fahren und Fahrzeugsicherheit” and elective for other master programmes.

Datasets and Code-Repositories

From this research, multiple databases were published. All databases were released under the Creative Commons Attribution 4.0 International licence.

- The publishable source code was released by Plaschkies (2024a).
- The FE-models from the simplification chain was published by Plaschkies (2024c).
- The development-set based on the Sobol algorithm and Hybrid III (HIII) with 24 576 FE-simulations was published by Plaschkies (2024f).
- The alternative development-set based on the full factorial scheme and HIII with 9375 FE-simulations was published by Plaschkies (2024d).
- The testing-set based on the Sobol algorithm and HIII with 768 FE-simulations was published by Plaschkies (2024e).
- The validation-set based on the full factorial scheme and VIRTHUMAN (VHUM) with 729 FE-simulations was published by Plaschkies and Müller (2024).
- The results from the machine learning chapter 4 were published by Plaschkies (2024b).

Bibliography

Abrams, M. Z. and Bass, C. R. (2020). "Female vs. Male Relative Fatality Risk in Fatal Crashes". In: *Proceedings of the International Research Council on the Biomechanics of Injury Conference* 2020, pp. 47–85. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc20/pdf-files/13.pdf.

Ackland, T. R., Henson, P. W., and Bailey, D. A. (1988). "The Uniform Density Assumption: Its Effect upon the Estimation of Body Segment Inertial Parameters". In: *International Journal of Sport Biomechanics* 4.2, pp. 146–155. ISSN: 0740-2082. DOI: 10.1123/ijsb.4.2.146.

Agnew, A. M., Kang, Y.-S., Moorhouse, K., Herriott, R. G., and Bolte IV, J. H. (2013). "Age-Related Changes in Stiffness in Human Ribs". In: *Proceedings of the International Research Council on the Biomechanics of Injury Conference* 2013, pp. 257–269. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc13/pdf_files/32.pdf (visited on Feb. 5, 2024).

Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M. (2019). "Optuna: A Next-generation Hyperparameter Optimization Framework". In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (Anchorage, USA), pp. 2623–2631. ISBN: 9781450362016. DOI: 10.1145/3292500.3330701.

Albert, D. L., Beeman, S. M., and Kemper, A. R. (2023). "Comparative Biofidelity of the Hybrid III and THOR 50th Male ATDs Under Three Restraint Conditions During Frontal Sled Tests". In: *Traffic Injury Prevention* 24, pp. 41–46. DOI: 10.1080/15389588.2023.2176710.

ANSYS, Inc. (2024). *LS-DYNA Theory Manual*. URL: https://ftp.lstc.com/anonymous/outgoing/web/ls-dyna_manuals/DRAFT/DRAFT_Theory.pdf (visited on Dec. 10, 2024).

Apache Software Foundation (2024). *Apache Arrow*. URL: https://arrow.apache.org/docs/index.html (visited on Jan. 4, 2024).

Araújo, C. G. S. de (2008). "Flexibility Assessment: Normative Values for Flexitest from 5 to 91 Years of Age". In: *Arquivos Brasileiros de Cardiologia* 90.4, pp. 257–263. ISSN: 1678-4170. DOI: 10.1590/S0066-782X2008000400008.

Azzimonti, D., Ginsbourger, D., Chevalier, C., Bect, J., and Richet, Y. (2021). "Adaptive Design of Experiments for Conservative Estimation of Excursion Sets". In: *Technometrics* 63.1, pp. 13–26. ISSN: 0040-1706. DOI: 10.1080/00401706.2019.1693427.

Bach, C., Song, L., Erhart, T., and Duddeck, F. (2019). "Dimensionality Reduction of Crash and Impact Simulations using LS-DYNA". In: *12th European LS-DYNA Conference* (Koblenz, Germany). URL: https://www.dynalook.com/conferences/12th-european-ls-dyna-conference-2019/model-reduction-and-analysis/bach_bmw.pdf (visited on Feb. 5, 2024).

Bagnall, A., Lines, J., Bostrom, A., Large, J., and Keogh, E. (2017). "The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances". In: *Data Mining and Knowledge Discovery* 31.3, pp. 606–660. ISSN: 1384-5810. DOI: 10.1007/s10618-016-0483-9. eprint: 30930678.

Ballal, N., Soot, T., Dlugosch, M., Trube, N., and Fressmann, D. (2023). "A Method for Efficient Generation and Optimization of Simulation-Based Training Data for Data-Driven Injury Prediction in VRU-Vehicle Accident Scenarios". In: *27th International Technical Conference on the Enhanced Safety of Vehicles* (Yokohama, Japan). URL: https://www-nrd.nhtsa.dot.gov/departments/esv/27th (visited on Mar. 6, 2024).

Bance, I. and Nie, B. (2019). "A Framework for Near Real-Time Occupant Injury Risk Prediction using a Sequence-to-Sequence Deep Learning Approach". In: *Proceedings of the International Research Council on the Biomechanics of Injury Conference* 2019, pp. 86–87. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc19/pdf-files/20.pdf (visited on Jan. 29, 2024).

Bance, I., Yang, S., Zhou, Q., Li, S., and Nie, B. (2021). "A Framework for Rapid On-board Deterministic Estimation of Occupant Injury Risk in Motor Vehicle Crashes with Quantitative Uncertainty Evaluation". In: *Science China Technological Sciences* 64.3, pp. 521–534. ISSN: 1674-7321. DOI: 10.1007/s11431-019-1565-9.

Barbat, S. D. and Li, X. (2017). "Analysis of the Proposed Frontal Oblique Crash Test". In: *25th International Technical Conference on the Enhanced Safety of Vehicles* (Detroit, USA). URL: https://www-esv.nhtsa.dot.gov/Proceedings/25/25th%7B%5C_%7DESV%7B%5C_%7DProceedings%7B%5C_%7DBooklet.pdf (visited on Apr. 4, 2024).

Beillas, P. and Berthet, F. (2018). "Development of Simulation Based Liver and Spleen Injury Risk Curves for the GHBMC Detailed Models". In: *Proceedings of the International Research Council on the Biomechanics of Injury Conference* 2018, pp. 369–381. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc18/pdf-files/61.pdf (visited on Feb. 5, 2024).

Belaid, M. K., Rabus, M., and Krestel, R. (2021). "CrashNet: An Encoder–decoder Architecture to Predict Crash Test Outcomes". In: *Data Mining and Knowledge Discovery* 35.4, pp. 1688–1709. ISSN: 1384-5810. DOI: 10.1007/s10618-021-00761-9.

Brans, J. P. and Vincke, P. (1985). "Note – A Preference Ranking Organisation Method". In: *Management Science* 31.6, pp. 647–656. ISSN: 0025-1909. DOI: 10.1287/mnsc.31.6.647.

Brownlee, J. (2019). *Loss and Loss Functions for Training Deep Learning Neural Networks*. Machine Learning Mastery. URL: https://machinelearningmastery.com/loss-and-loss-functions-for-training-deep-learning-neural-networks/ (visited on Nov. 27, 2024).

Brumbelow, M. L. (2019). "Front Crash Injury Risks for Restrained Drivers in Good-rated Vehicles by Age, Impact Configuration, and EDR-based Delta V". In: *Proceedings of the International Research Council on the Biomechanics of Injury Conference* 2019, pp. 561–575. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc19/pdf-files/81.pdf (visited on Feb. 5, 2024).

Butz, K. D., Spurlock, C. M., and Lister, K. (2020). "Validation of the CAVEMAN Human Body Model Lumbar Spine Response to Vertical Accelerative Loading and the Effect of Spine Curvature on Skeletal Fracture". In: *Proceedings of the International Research Council*

on the Biomechanics of Injury Conference 2020, pp. 113–122. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc20/pdf-files/17.pdf (visited on Feb. 5, 2024).

Calders, T. and Assche, D. van (2018). "PROMETHEE is Not Quadratic: An O(qnlog(n)) Algorithm". In: *Omega* 76, pp. 63–69. ISSN: 03050483. DOI: 10.1016/j.omega.2017.04.003.

Carter, P. M., Flannagan, C. A. C., Reed, M. P., Cunningham, R. M., and Rupp, J. D. (2014). "Comparing the Effects of Age, BMI and Gender on Severe Injury (AIS 3+) in Motor-Vehicle Crashes". In: *Accident Analysis and Prevention* 72, pp. 146–160. ISSN: 0001-4575. DOI: 10.1016/j.aap.2014.05.024.

CATARC, ed. (2024). *C-NCAP Management Regulation*. URL: https://cncap.obs.cn-north-4.myhuaweicloud.com/cms/files/779220639339855872.zip (visited on Nov. 19, 2024).

Chen, W., Zhou, Q., Gan, S., and Nie, B. (2020). "Generating a Large-scale Numerical Database of Motor Vehicle Crashes for Rapid Injury Severity Prediction". In: *Proceedings of the International Research Council on the Biomechanics of Injury Conference* 2020.Asia, pp. 25–28. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc20-asia/pdf-files/2016a.pdf (visited on Feb. 5, 2024).

Chollet, F. et al. (2015). *Keras*. URL: https://keras.io (visited on Dec. 16, 2024).

Chollet, F. (2023). *Transfer Learning and Fine-tuning*. URL: https://keras.io/guides/transfer_learning/ (visited on Dec. 16, 2024).

Chong, M., Abraham, A., and Paprzycki, M. (2005). "Traffic Accident Analysis Using Machine Learning Paradigms". In: *Informatica* 29.1, pp. 89–98. ISSN: 1854-3871. URL: https://informatica.si/index.php/informatica/article/download/21/15 (visited on Feb. 5, 2024).

Christ, M., Braun, N., Neuffer, J., and Kempa-Liehr, A. W. (2018). "Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests (tsfresh – A Python package)". In: *Neurocomputing* 307, pp. 72–77. ISSN: 09252312. DOI: 10.1016/j.neucom.2018.03.067.

Collette, A., Kluyver, T., Caswell, T. A., Tocknell, J., Kieffer, J., Jelenak, A., Scopatz, A., Dale, D., Chen, Vincent, T., Einhorn, M., payno, juliagarriga, Sciarelli, P., Valls, V., Ghosh, S., Pedersen, U. K., Kittisopikul, M., jakirkham, Raspaud, M., Danilevski, C., Abbasi, H., Readey, J., Mühlbauer, K., Paramonov, A., Chan, L., Schepper, R. de, Solé, V. A., jialin, and Guest, D. H. (2023). *h5py: 3.8.0*. DOI: 10.5281/zenodo.7560547.

Cormier, J. M. (2008). "The Influence of Body Mass Index on Thoracic Injuries in Frontal Impacts". In: *Accident Analysis and Prevention* 40.2, pp. 610–615. ISSN: 0001-4575. DOI: 10.1016/j.aap.2007.08.016.

Cox, A. E. and Cicchino, J. B. (2021). "Continued Trends in Older Driver Crash Involvement Rates in the United States: Data Through 2017-2018". In: *Journal of Safety Research* 77, pp. 288–295. ISSN: 1879-1247. DOI: 10.1016/j.jsr.2021.03.013. eprint: 34092320.

Craig, M. J., Scarboro, M., and Ridella, S. A. (2011). "Predicting Occupant Outcomes with EDR Data". In: *22nd International Technical Conference on the Enhanced Safety of Vehicles* (Washington, DC., US). URL: https://www-nrd.nhtsa.dot.gov/departments/esv/22nd/files/22nd%7B%5C_%7DESV%7B%5C_%7DProceedings%7B%5C_%7DBooklet.pdf (visited on Feb. 5, 2024).

Cronin, D. S., Fice, J. B., DeWit, J. A., and Moulton, J. A. (2012). "Upper Cervical Spine Kinematic Response and Injury Prediction". In: *Proceedings of the International Research Council on the Biomechanics of Injury Conference* 2012, pp. 225–234. ISSN: 2235-3151.

URL: https://www.ircobi.org/wordpress/downloads/irc12/pdf_files/30.pdf (visited on Feb. 5, 2024).

Dempster, A., Petitjean, F., and Webb, G. I. (2020). "ROCKET: exceptionally fast and accurate time series classification using random convolutional kernels". In: *Data Mining and Knowledge Discovery* 34.5, pp. 1454–1495. ISSN: 1384-5810. DOI: 10.1007/s10618-020-00701-z.

Diaz-Rodriguez, M., Valera, A., Page, A., Besa, A., and Mata, V. (2016). "Dynamic Parameter Identification of Subject-Specific Body Segment Parameters Using Robotics Formalism: Case Study Head Complex". In: *Journal of Biomechanical Engineering* 138.5, pp. 1–8. ISSN: 1528-8951. DOI: 10.1115/1.4032997.

Diederich, A., Bastien, C., and Blundell, M. (2023). "The Prediction of Autonomous Vehicle Occupants' Pre-crash Motion During Emergency Braking Scenarios". In: *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering* 237.14, pp. 3304–3312. ISSN: 0954-4070. DOI: 10.1177/09544070231153262.

Duma, S. M., Moorcroft, D. M., Stitzel, J. D., and Duma, G. G. (2004). "Evaluating Pregnant Occupant Restraints: The Effect of Local Uterine Compression on the Risk of Fetal Injury". In: *Annual Proceedings - Association for the Advancement of Automotive Medicine* 48, pp. 103–114. ISSN: 1540-0360. URL: https://pubmed.ncbi.nlm.nih.gov/15319120.

Duma, S. M., Moorcroft, D. M., Stitzel, J. D., and Duma, G. G. (2005). "A Computational Model of The Pregnant Occupant: Effects of Restraint Usage and Occupant Position on Fetal Injury Risk". In: *19th International Technical Conference on the Enhanced Safety of Vehicles* (Washington DC, United States). URL: https://www-esv.nhtsa.dot.gov/Proceedings/19/19th%7B%5C_%7DESV%7B%5C_%7DProceedings%7B%5C_%7DBooklet.pdf (visited on Feb. 5, 2024).

Dumas, R., Cheze, L., and Verriest, J.-P. (2007). "Adjustments to McConville et al. and Young et al. Body Segment Inertial Parameters". In: *Journal of Biomechanics* 40.3, pp. 543–553. ISSN: 1873-2380. DOI: 10.1016/j.jbiomech.2006.02.013.

Dummy Testing and Equipment Committee (2022). *User's Manual for the 50th Percentile Male Hybrid III Dummy*. J2856. DOI: 10.4271/J2856_202211.

"Hans Newsletter Issue #1" (2023). In: ed. by DYNAmore GmbH. URL: https://www.dynamore.de/de/news/hans-newsletter/hans-newsletter-issue-1 (visited on Nov. 19, 2024).

Elvebakk, B. (2007). "Vision Zero: Remaking Road Safety". In: *Mobilities* 2.3, pp. 425–441. ISSN: 1745-0101. DOI: 10.1080/17450100701597426.

Engin, A. E. and Chen, S.-M. (1987). "Kinematic and Passive Resistive Properties of Human Elbow Complex". In: *Journal of Biomechanical Engineering* 109.4, pp. 318–323. ISSN: 1528-8951. DOI: 10.1115/1.3138687.

Eppinger, R. H., Sun, E., Kuppa, S. M., and Saul, R. (2000). *Supplement: Development of Improved Injury Criteria for the Assessment of Advanced Automotive Restraint Systems -II*. URL: https://rosap.ntl.bts.gov/view/dot/14739/dot_14739_DS1.pdf (visited on Feb. 5, 2024).

Euro NCAP (2020). *THOR Specification and Certification*. TB 026. Version 1.3. URL: https://cdn.euroncap.com/media/77148/tb-026-thor-specification-and-certification-v13.pdf (visited on Feb. 15, 2024).

"Euro NCAP Vision 2030: A Safer Future for Mobility" (Nov. 9, 2022). In: ed. by Euro NCAP. URL: https://www.euroncap.com/en/press-media/press-releases/euro-ncap-vision-2030-a-safer-future-for-mobility/ (visited on May 20, 2024).

Euro NCAP, ed. (2023a). *Assessment Protocol – Adult Occupant Protection*. Version 9.3. URL: https://www.euroncap.com/media/79871/euro-ncap-assessment-protocol-aop-v93.pdf (visited on Nov. 19, 2024).

Euro NCAP, ed. (2023b). *Data format and Injury Criteria Calculation*. TB 021.

Euro NCAP, ed. (2023c). *MPDB Frontal Impact Testing Protocol*. Version 1.1.4. URL: https://www.euroncap.com/media/79875/euro-ncap-mpdb-testing-protocol-v114.pdf (visited on Nov. 19, 2024).

Euro NCAP (2023d). *Pedestrian Human Model Certification*. TB 024. Version 4.0. URL: https://cdn.euroncap.com/media/77683/tb-024-pedestrian-human-model-certification-v40.pdf (visited on Feb. 15, 2024).

Euro NCAP (2023e). *Vulnerable Road User Testing Protocol*. Version 9.1. URL: https://www.euroncap.com/media/79878/euro-ncap-vru-testing-protocol-v91.pdf (visited on Feb. 15, 2024).

Evers, C. (2021). "Gurte, Kindersitze, Helme und Schutzkleidung – 2021". In: ed. by Bundesanstalt für Straßenwesen. URL: https://www.bast.de/DE/Publikationen/DaFa/2022-2021/2022-02.html (visited on Nov. 9, 2024).

Forman, J. L., Lopez-Valdes, F. J., Lessley, D. J., Kindig, M. W., Kent, R. W., and Boström, O. (2009). "The Effect of Obesity on the Restraint of Automobile Occupants". In: *Annals of Advances in Automotive Medicine / Annual Scientific Conference* 53, pp. 25–40. ISSN: 1943-2461. eprint: 20184830.

Gan, N. and Gu, J. (2019). "Hybrid Meta-model-based Design Space Exploration Method for Expensive Problems". In: *Structural and Multidisciplinary Optimization* 59.3, pp. 907–917. ISSN: 1615-147X. DOI: 10.1007/s00158-018-2109-x.

Gartner (2024). *Explore Beyond GenAI on the 2024 Hype Cycle for Artificial Intelligence*. Ed. by A. Jaffri. Gartner. URL: https://www.gartner.com/en/articles/hype-cycle-for-artificial-intelligence (visited on Nov. 11, 2024).

Gayzik, F. S., Moreno, D. P., Geer, C. P., Wuertzer, S. D., Martin, R. S., and Stitzel, J. D. (2011). "Development of a Full Body CAD Dataset for Computational Modeling: A Multi-modality Approach". In: *Annals of Biomedical Engineering* 39.10, pp. 2568–2583. DOI: 10.1007/s10439-011-0359-5.

George Mason University, Center for Collision Safety and Analysis (2016). *2010 Toyota Yaris. Finite Element Model Validation, Coarse Mesh*. DOI: 10.13021/G8JS5D.

George Mason University, Center for Collision Safety and Analysis (2022). *2020 Nissan Rogue FE Model. Version 2*. DOI: 10.13021/xb7g-8z06.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. Adaptive computation and machine learning. Cambridge, Massachusetts and London, England: The MIT Press. 775 pp. ISBN: 0262035618. URL: https://www.deeplearningbook.org/.

Gordon, C. C., Blackwell, C. L., Bradtmiller, B., Parham, J. L., Barrientos, P., Paquette, S. P., Corner, B. D., Carson, J. M., Venezia, J. C., Rockwell, B. M., Mucher, M., and Kristensen, S. (2014). *2012 Anthropometric Survey of U.S. Army Personnel. Methods and Summary*

Statistics. NATICK/TR-15/007. Natick, Massachusetts. URL: http://tools.openlab.psu.edu/publicData/ANSURII-TR15-007.pdf (visited on Feb. 5, 2024).

Gordon, C. C., Churchill, T. D., Clauser, C. E., Bradtmiller, B., McConville, J. T., Tebbetts, I., and Walker, R. A. (1989). *1988 Anthropometric Survey of U.S. Army Personnel. Summary Statistics Interim Report*. NATICK/TR-89/027. Natick, Massachusetts. URL: https://apps.dtic.mil/sti/pdfs/ADA209600.pdf (visited on Feb. 15, 2024).

Graz University of Technology (2023). *Objective Rating Metric ISO 18571*. URL: https://openvt.eu/validation-metrics/ISO18571 (visited on Oct. 10, 2024).

Guha, S., Bhalsod, D., and Krebs, J. (2011). *LSTC Hybrid III 50th Fast Dummy Positioning & Post–Processing. Dummy Version: LSTC.H3_50TH_FAST.111130_V2.0*.

Hao, Z., Liu, S., Zhang, Y., Ying, C., Feng, Y., Su, H., and Zhu, J. (2022). *Physics-Informed Machine Learning: A Survey on Problems, Methods and Applications*. arXiv Preprint. DOI: 10.48550/arXiv.2211.08064.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., Del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E. (2020). "Array programming with NumPy". In: *Nature* 585.7825, pp. 357–362. DOI: 10.1038/s41586-020-2649-2. eprint: 32939066.

Härtel, B. (2024). "Entwicklung eines Biofidelen Dummys zur Darstellung Komplexer Verletzungen beim Fahrzeugcrash". Faculty of Mechanical, Process and Energy Engineering. Dissertation. Technische Universität Bergakademie Freiberg. URL: https://nbn-resolving.org/urn:nbn:de:bsz:105-qucosa2-919620 (visited on Nov. 19, 2024).

Hasija, V. and Takhounts, E. G. (2019). "Machine Learning Based Model for Predicting Head Injury Criterion (HIC)". In: *SAE Technical Paper Series*. ISSN: 0148-7191. DOI: 10.4271/2019-22-0016.

Hasija, V. and Takhounts, E. G. (2022). "Deep Learning Methodology for Predicting Time History of Head Angular Kinematics from Simulated Crash Videos". In: *Scientific reports* 12.1, p. 6526. DOI: 10.1038/s41598-022-10480-w.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning*. New York, NY: Springer New York. ISBN: 978-0-387-84857-0. DOI: 10.1007/978-0-387-84858-7.

Hofmann, H., Wickham, H., and Kafadar, K. (2017). "Letter-Value Plots: Boxplots for Large Data". In: *Journal of Computational and Graphical Statistics* 26.3, pp. 469–477. ISSN: 1061-8600. DOI: 10.1080/10618600.2017.1305277.

Hu, J., Fanta, A., Neal, M. O., Reed, M. P., and Tai-Wang, J. (2016). "Crash Simulations with Morphed GHBMC Human Models Vehicle Crash Simulations with Morphed GHBMC Human Models of Different Stature, BMI, and Age". In: *Proceedings of the 4th International Digital Human Modeling Conference* (Montreal, Canada). URL: https://api.semanticscholar.org/CorpusID:42313083 (visited on Feb. 5, 2024).

Hu, J., Zhang, K., Fanta, A., Jones, M. L. H., Reed, M. P., Neal, M. O., Wang, J.-T., Lin, C.-H., and Cao, L. (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 the Biomechanics of Injury Conference 2017, pp. 656–667. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc17/pdf-files/85.pdf (visited on Feb. 5, 2024).

Hu, J., Zhang, K., Reed, M. P., Wang, J.-T., Neal, M. O., and Lin, C.-H. (2019). "Frontal Crash Simulations using Parametric Human Models Representing a Diverse Population". In: *Traffic Injury Prevention* 20.sup1, pp. 97–105. DOI: 10.1080/15389588.2019.1581926.

Humanetics Innovative Solutions Inc (2024a). *Anthropomorphic Test Devices (ATD)*. URL: https://www.humaneticsgroup.com/products/anthropomorphic-test-devices (visited on May 20, 2024).

"Next-Gen Safety Simulations with the HBM Connect™ Family" (2024b). In: ed. by Humanetics Innovative Solutions Inc. URL: https://www.humaneticsgroup.com/perspectives/next-gen-safety-simulations-hbm-connecttm-family (visited on Nov. 19, 2024).

Hunter, J. D. (2007). "Matplotlib: A 2D Graphics Environment". In: *Computing in Science & Engineering* 9.3, pp. 90–95. ISSN: 1521-9615. DOI: 10.1109/MCSE.2007.55.

Hynčík, L., Cechova, H., Kovar, L., and Blaha, P. (2013). "On Scaling Virtual Human Models". In: *SAE Technical Paper Series*. ISSN: 0148-7191. DOI: 10.4271/2013-01-0074.

Imam, F., Musilek, P., and Reformat, M. Z. (2024). "Parametric and Nonparametric Machine Learning Techniques for Increasing Power System Reliability: A Review". In: *Information* 15.1, p. 37. DOI: 10.3390/info15010037.

International Organization for Standardization, ed. (2014). *Road Vehicles - Objective Rating Metric for Non-ambiguous Signals*. ISO/TS 18571:2014(E). British Standards Institution.

International Organization for Standardization, ed. (2015). *Road Vehicles — Measurement Techniques in Impact Tests — Instrumentation*. ISO 6487:2015.

International Organization for Standardization, ed. (2019). *Road Vehicles — Multimedia Data Exchange Format for Impact Tests*. ISO/TS 13499:2019.

Ismail Fawaz, H., Lucas, B., Forestier, G., Pelletier, C., Schmidt, D. F., Weber, J., Webb, G. I., Idoumghar, L., Muller, P.-A., and Petitjean, F. (2020). "InceptionTime: Finding AlexNet for time series classification". In: *Data Mining and Knowledge Discovery* 34.6, pp. 1936–1962. ISSN: 1384-5810. DOI: 10.1007/s10618-020-00710-y.

Iwamoto, M., Nakahira, Y., and Kimpara, H. (2015). "Development and Validation of the Total Human Model for Safety (THUMS) Toward Further Understanding of Occupant Injury Mechanisms in Precrash and During Crash". In: *Traffic Injury Prevention* 16 Suppl 1, S36–48. DOI: 10.1080/15389588.2015.1015000. eprint: 26027974.

Izumiyama, T., Nishida, N., Iwanaga, H., Chen, X., Ohgi, J., Asahi, R., Sugimoto, S., and Fukushima, M. (2020). "Identification of Influential Factors for Seatbelt Kinematics in a Collision and Analysis of their Influence Degree to the Kinematics". In: *Proceedings of the International Research Council on the Biomechanics of Injury Conference* 2020, pp. 456–469. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc20/pdf-files/57.pdf (visited on Feb. 5, 2024).

Izumiyama, T., Nishida, N., Iwanaga, H., Chen, X., Ohgi, J., Mori, K., Hayashi, T., Sakuramoto, I., Asahi, R., Sugimoto, S., and Ueno, M. (2018). "The Analysis of an Individual Difference in Human Skeletal Alignment in Seated Posture and Occupant Behavior Using HBMs". In: *Proceedings of the International Research Council on the Biomechanics of Injury Conference*

2018, pp. 549–560. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc18/pdf-files/84.pdf (visited on Feb. 5, 2024).

El-Jawahri, R. E., Laituri, T. R., Ruan, J. S., Rouhana, S. W., and Barbat, S. D. (2010). "Development and Validation of Age-dependent FE Human Models of a Mid-sized Male Thorax". In: *Stapp Car Crash Journal* 54, pp. 407–430. ISSN: 1532-8546. DOI: 10.4271/2010-22-0017.

Jehle, D., Gemme, S., and Jehle, C. (2012). "Influence of Obesity on Mortality of Drivers in Severe Motor Vehicle Crashes". In: *The American Journal of Emergency Medicine* 30.1, pp. 191–195. DOI: 10.1016/j.ajem.2010.10.017.

Jin, X., Kalra, A., Hammad, A., Khandelwal, P., Porwal, V., Shen, M., and Yang, K. H. (2018). "Development and Validation of Whole–Body Finite Element Occupant and Pedestrian Models of a 70–Year–Old Female". In: *Proceedings of the International Research Council on the Biomechanics of Injury Conference* 2018, pp. 261–283. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc18/pdf-files/41.pdf (visited on Feb. 5, 2014).

Jolivet, E., Lafon, Y., Petit, P., and Beillas, P. (2015). "Comparison of Kriging and Moving Least Square Methods to Change the Geometry of Human Body Models". In: *Stapp Car Crash Journal* 59, pp. 337–357. ISSN: 1532-8546. DOI: 10.4271/2015-22-0013.

Joodaki, H., Gepner, B. D., and Kerrigan, J. R. (2021). "Leveraging Machine Learning for Predicting Human Body Model Response in Restraint Design Simulations". In: *Computer Methods in Biomechanics and Biomedical Engineering* 24.6, pp. 597–611. DOI: 10.1080/10255842.2020.1841754. eprint: 33179985.

Jorlöv, S., Bohman, K., and Larsson, A. (2017). "Seating Positions and Activities in Highly Automated Cars – A Qualitative Study of Future Automated Driving Scenarios". In: *Proceedings of the International Research Council on the Biomechanics of Injury Conference*, pp. 13–22. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc17/pdf-files/11.pdf (visited on June 4, 2024).

Kaneko, H. (2021). "Adaptive Design of Experiments Based on Gaussian Mixture Regression". In: *Chemometrics and Intelligent Laboratory Systems* 208, p. 104226. ISSN: 01697439. DOI: 10.1016/j.chemolab.2020.104226.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., and Yang, L. (2021). "Physics-informed Machine Learning". In: *Nature Reviews Physics* 3.6, pp. 422–440. DOI: 10.1038/s42254-021-00314-5.

Kato, D., Nakahira, Y., Atsumi, N., and Iwamoto, M. (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 the Biomechanics of Injury Conference*, pp. 207–223. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc18/pdf-files/32.pdf (visited on June 4, 2024).

Kayvantash, K. (2018). "Model Reduction Techniques for On-board and Parametric Crash and Safety Simulations". In: *German LS-DYNA Forum 2018* (Bamberg, Germany). URL: https://www.dynamore.de/de/fortbildung/konferenzen/vergangene/2018-ls-dyna-forum/agenda/e-05-occupant-safety-ii_kayvantash_cadlm.pdf (visited on Feb. 5, 2024).

Keyser, W. de and Peeters, P. (1996). "A Note on the Use of PROMETHEE Multicriteria Methods". In: *European Journal of Operational Research* 89.3, pp. 457–461. ISSN: 03772217. DOI: 10.1016/0377-2217(94)00307-6.

Kim, J.-E., Kim, I. H., Shum, P. C., Shih, A. M., Pintar, F. A., Shen, W., Ma, X., Laud, P. W., Heymsfield, S. B., Allison, D. B., and Zhu, S. (2014). "A Computational Study of Injury Severity and Pattern Sustained by Overweight Drivers in Frontal Motor Vehicle Crashes". In: *Computer Methods in Biomechanics and Biomedical Engineering* 17.9, pp. 965–977. DOI: 10.1080/10255842.2012.728589.

Kingma, D. P. and Ba, J. (2015). "Adam: A Method for Stochastic Optimization". In: *3rd International Conference on Learning Representations (ICLR)* (San Diego, CA, USA).

Klinich, K. D., Schneider, L. W., Moore, J. L., and Pearlman, M. D. (2000). "Investigations of Crashes Involving Pregnant Occupants". In: *Annual Proceedings - Association for the Advancement of Automotive Medicine* 44, pp. 37–56. ISSN: 1540-0360. eprint: 11558095.

Köllner, C. (2023). "Meilensteine der passiven Sicherheit". In: *ATZ*. URL: https://www.springerprofessional.de/125-jahre-atz/fahrzeugsicherheit/meilensteine-der-passiven-sicherheit/24085200 (visited on Nov. 9, 2024).

Kovar, L. and Hlucha, J. (2019). "ESI VIRTHUMAN Models for Impact". In: *DHM and Posturography*. Elsevier, pp. 169–185. ISBN: 9780128167137. DOI: 10.1016/B978-0-12-816713-7.00015-5.

Kracker, D. (2024). *Automatisierte Auswertung von Crashsimulationen unterschiedlicher Fahrzeug-Entwicklungsstände mit Methoden des maschinellen Lernens*. Dissertation. University of Wuppertal, School of Mechanical Engineering and Safety Engineering, Shaker Verlag. ISBN: 978-3-8440-9424-4.

Kracker, D., Garcke, J., Schumacher, A., and Schwanitz, P. (2020). "Automatic Analysis of Crash Simulations with Dimensionality Reduction Algorithms such as PCA and t-SNE". In: *16th International LS-DYNA Conference* (Online). URL: https://www.dynalook.com/conferences/16th-international-ls-dyna-conference/automotive/t1-1-a-automotive-011.pdf (visited on Feb. 5, 2024).

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2017). "ImageNet Classification with Deep Convolutional Neural Networks". In: *Communications of the ACM* 60.6, pp. 84–90. ISSN: 0001-0782. DOI: 10.1145/3065386.

Kübler, L., Gargallo, S., and Elsäßer, K. (2009). "Bewertungskriterien zur Auslegung von Insassenschutzsystemen". In: *ATZ - Automobiltechnische Zeitschrift* 111.6, pp. 426–433. ISSN: 0001-2785. DOI: 10.1007/BF03222079.

Kunitomi, S. and Konosu, A. (2018). "Prediction of Head Injury Severity for Pedestrians in Car–Pedestrian Accidents using Deep Learning Methodology". In: 2018, pp. 309–310. URL: https://www.ircobi.org/wordpress/downloads/irc18/pdf-files/46.pdf (visited on Feb. 5, 2024).

Laakmann, F., Zink, L., and Seyffert, M. (2019). "New Interior Concepts for Occupant Protection in Highly Automated Vehicles". In: *ATZ worldwide* 121.4, pp. 48–53. DOI: 10.1007/s38311-019-0012-8.

Larsson, E., Iraeus, J., Fice, J. B., Pipkorn, B., Jakobsson, L., Brynskog, E., Brolin, K., and Davidsson, J. (2019). "Active Human Body Model Predictions Compared to Volunteer Response in Experiments with Braking, Lane Change, and Combined Manoeuvres". In:

Proceedings of the International Research Council on the Biomechanics of Injury Conference, pp. 349–369. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc19/pdf-files/50.pdf (visited on May 21, 2024).

Lasso GmbH (2022). *Open LASSO Python*. URL: https://open-lasso-python.github.io/lasso-python/ (visited on Jan. 4, 2024).

Lee, Y., Cho, E., Park, M., Kim, H., Choi, K., and Park, S. (2018). "A Machine Learning Approach to Prediction of Passenger Injuries on Real Road Situation". In: *2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS)*. Institute of Electrical and Electronics Engineers (IEEE), pp. 1087–1090. ISBN: 978-1-5386-2633-7. DOI: 10.1109/SCIS-ISIS.2018.00174.

Leva, P. de (1996). "Adjustments to Zatsiorsky-Seluyanov's Segment Inertia Parameters". In: *Journal of Biomechanics* 29.9, pp. 1223–1230. ISSN: 1873-2380. DOI: 10.1016/0021-9290(95)00178-6.

Li, X., Yuan, Q., Lindgren, N., Huang, Q., Fahlstedt, M., Östh, J., Pipkorn, B., Jakobsson, L., and Kleiven, S. (2023). "Personalization of Human Body Models and Beyond via Image Registration". In: *Frontiers in Bioengineering and Biotechnology* 11, p. 1169365. ISSN: 2296-4185. DOI: 10.3389/fbioe.2023.1169365. eprint: 37274163.

Limousin, V., Delgerie, X., Leroy, E., Ibáñez, R., Argerich, C., Daim, F., Louis Duval, J., and Chinesta, F. (2019). "Advanced Model Order Reduction and Artificial Intelligence Techniques Empowering Advanced Structural Mechanics Simulations: Application to Crash Test Analyses". In: *Mechanics & Industry* 20.8, p. 804. ISSN: 2257-7777. DOI: 10.1051/meca/2020009.

Liu, L., Zhang, X., Liu, Y., Zhu, W., and Zhao, B. (2020). "An Ensemble of Multiple Boosting Methods Based on Classifier-Specific Soft Voting for Intelligent Vehicle Crash Injury Severity Prediction". In: *2020 IEEE Sixth International Conference on Big Data Computing Service and Applications (BigDataService)* (Oxford, UK), pp. 17–24. DOI: 10.1109/BigDataService49289.2020.00011.

Livermore Software Technology (2020a). *LS-DYNA Keyword Users's Manual. Volume I*. LS-DYNA R12. URL: http://ftp.lstc.com/anonymous/outgoing/marleigh/manual/LS-DYNA_Manual_Volume_I_R12.pdf (visited on Dec. 10, 2024).

Livermore Software Technology (2020b). *LS-DYNA Keyword Users's Manual. Volume II Material Models*. LS-DYNA R12. URL: http://ftp.lstc.com/anonymous/outgoing/marleigh/manual/LS-DYNA_Manual_Volume_II_R12_Ver2.pdf (visited on Dec. 10, 2024).

Löning, M., Király, F., Bagnall, T., Middlehurst, M., Ganesh, S., Oastler, G., Lines, J., Walter, M., Kaz, V., Mentel, L., Holder, C., Tsaprounis, L., Kuhns, R., Parker, M., Owoseni, T., Rockenschaub, P., danbartl, jesellier, eenticott-shell, Gilbert, C., Bulatova, G., Lovkush, Schäfer, P., Khrapov, S., Buchhorn, K., Take, K., Subramanian, S., Meyer, S. M., Brooke, A. R., and Rice, B. (2022). *sktime: v0.13.4*. Zenodo. DOI: 10.5281/zenodo.7117735.

Lubba, C. H., Sethi, S. S., Knaute, P., Schultz, S. R., Fulcher, B. D., and Jones, N. S. (2019). "catch22: CAnonical Time-series CHaracteristics". In: *Data Mining and Knowledge Discovery* 33.6, pp. 1821–1852. ISSN: 1384-5810. DOI: 10.1007/s10618-019-00647-x.

Lugner, R., Inderst, M., Sequeira, G., Schneider, K., and Brandmeier, T. (2020). “Collision Prediction for Irreversible Pre-Crash Safety Measures”. In: *FISITA Web Congress* (Online). DOI: 10.46720/F2020-PIF-033.

Manary, M. A., Reed, M. P., Flannagan, C. A. C., and Schneider, L. W. (1998). “ATD Positioning Based on Driver Posture and Position”. In: *SAE Technical Paper Series*. ISSN: 0148-7191. DOI: 10.4271/983163.

Martin, P. G. and Shook, L. (2007). “NHTSA´s THOR-NT Database”. In: *20th International Technical Conference on the Enhanced Safety of Vehicles* (Lyon, France). URL: https://www-esv.nhtsa.dot.gov/Proceedings/20/20th%7B%5C_%7DESV%7B%5C_%7DProceedings%7B%5C_%7DBooklet.pdf (visited on Feb. 5, 2024).

Maurath, C., Guha, S., Bhalsod, D., Burger, M., Krebs, J., Bala, S., Stahlschmidt, S., D’Souza, R., Mohan, P., and Marzougui, D. (2021). “Overview of LSTC’s LS-Dyna Anthropomorphic Models”. In: *11th International LS-DYNA Conference* (Detroit, US). URL: https://lsdyna.ansys.com/wp-content/uploads/attachments/OccupantSafety-4.pdf (visited on Mar. 19, 2024).

McGrath, J. F. and McConville, J. B. (1998). *Introduction to ADAMS Theory*.

Mertz, H. J. (2002). “Injury Risk Assessments Based on Dummy Responses”. In: *Accidental Injury*. Ed. by A. M. Nahum and J. W. Melvin. New York, NY: Springer New York, pp. 89–102. ISBN: 978-1-4419-3168-9. DOI: 10.1007/978-0-387-21787-1_5.

Moorcroft, D. M., Stitzel, J. D., Duma, G. G., and Duma, S. M. (2003). “Computational Model of the Pregnant Occupant: Predicting the risk of Injury in Automobile Crashes”. In: *American Journal of Obstetrics and Gynecology* 189.2, pp. 540–544. ISSN: 0002-9378. DOI: 10.1067/S0002-9378(03)00519-2.

Moorcroft, D. M., Stitzel, J. D., Duma, S. M., and Duma, G. G. (2004). “The Effect of Pregnant Occupant Position and Belt Placement on the Risk of Fetal Injury”. In: *SAE Technical Paper Series*. ISSN: 0148-7191. DOI: 10.4271/2004-01-0324.

Morris, A. P., Welsh, R., and Hassan, A. (2003). “Requirements for the Crash Protection of Older Vehicle Passengers”. In: *Annual Proceedings - Association for the Advancement of Automotive Medicine* 47, pp. 165–180. ISSN: 1540-0360. eprint: 12941224.

Motmans, R. (2005). *DINBelg*. Ergonomie RC. URL: http://www.dinbelg.be/adultstotal.htm (visited on Apr. 7, 2021).

National Health Interview Survey (NHIS) (2022). *2021 Survey Description*. URL: https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHIS/2021/srvydesc-508.pdf (visited on Apr. 5, 2024).

National Highway Traffic Safety Administration (2019). *Crash Simulation Vehicle Models*. URL: https://www.nhtsa.gov/crash-simulation-vehicle-models (visited on Oct. 1, 2023).

National Highway Traffic Safety Administration (2022). *Interim Report to Congress — Crash Test Dummies*. DOI: 10.21949/1530264.

National Highway Traffic Safety Administration (2024). *Crash Simulation Vehicle Models*. URL: https://www.nhtsa.gov/crash-simulation-vehicle-models (visited on Oct. 1, 2024).

NHTSA, ed. (2015). *New Car Assessment Program*. NHTSA-2015-0119. URL: http://federalregister.gov/a/2015-31323.

NHTSA (2024). *FMVSS No. 208 THOR-50M Compliance Option*. URL: https://www.reginfo.gov/public/do/eAgendaViewRule?pubId=202404&RIN=2127-AM21 (visited on Nov. 19, 2024).

O'Connor, C., Kim, A., Barrette, T., and Dix, J. (2022). "A Comparison of the Mid-Size Male THOR and Hybrid III ATDs in Vehicle Frontal Crash Tests". In: *Stapp Car Crash Journal* 66, pp. 143–173. ISSN: 1532-8546. DOI: 10.4271/2022-22-0005. eprint: 37733824.

Östh, J., Mendoza-Vazquez, M., Linder, A., Svensson, M. Y., and Brolin, K. (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 the Biomechanics of Injury Conference*, pp. 443–466. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc17/pdf-files/60.pdf (visited on Feb. 15, 2024).

Östling, M. and Larsson, A. (2019). "Occupant Activities and Sitting Positions in Automated Vehicles in China and Sweden". In: *26th International Technical Conference on the Enhanced Safety of Vehicles* (Eindhoven, Netherlands). Vol. 19-0083. URL: https://www-nrd.nhtsa.dot.gov/pdf/ESV/Proceedings/26/26th%7B%5C_%7DESV%7B%5C_%7DProceedings%7B%5C_%7DBooklet.pdf (visited on May 17, 2024).

Pandas Development Team (2024). *pandas-dev/pandas: Pandas*. Zenodo. DOI: 10.5281/zenodo.10957263.

Panzer, M. B., Caudillo, A. F., Gepner, B. D., and Parent, D. (2024). *THOR 50th Male Finite Element Model User Manual. Model Version 2.9 for LS-Dyna*.

Panzer, M. B., Caudillo, A. F., and Parent, D. (2019). *THOR 50th Male Finite Element Model User Manual. Model Version 2.7 for LS-Dyna*.

Parenteau, C. S., Smedley, J., Carhart, M., and Dibb, A. (2020). "The Effect of Obesity on Rollover Ejection and Injury Risks". In: *SAE Technical Paper Series*. ISSN: 0148-7191. DOI: 10.4271/2020-01-1219.

Park, B.-K. and Reed, M. P. (2016). *Human Shapes. Realistic Human Body Shape Modeler Based on Real Data*. University of Michigan Transportation Research Institute. URL: http://humanshape.org/ (visited on Jan. 31, 2020).

Park, G., Kim, T., Crandall, J. R., Arregui–Dalmases, C., and Luzon–Narro, J. (2013). "Comparison of Kinematics of GHBMC to PMHS on the Side Impact Condition". In: *Proceedings of the International Research Council on the Biomechanics of Injury Conference*, pp. 368–379. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc13/pdf_files/41.pdf (visited on June 4, 2024).

Park, J., Ebert, S. M., Reed, M. P., and Hallman, J. J. (2016). "Statistical Models for Predicting Automobile Driving Postures for Men and Women Including Effects of Age". In: *Human Factors* 58.2, pp. 261–278. ISSN: 0018-7208. DOI: 10.1177/0018720815610249.

Patterson, J. and Gibson, A. (2017). *Deep Learning. A Practitioner's Approach*. First Edition. Beijing et al.: O'Reilly. 507 pp. ISBN: 9781491914250.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., VanderPlas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É. (2011). "Scikit-learn: Machine Learning in Python".

In: *Journal of Machine Learning Research* 12.85, pp. 2825–2830. ISSN: 1532-4435. URL: http://jmlr.org/papers/v12/pedregosa11a.html.

Picheny, V., Ginsbourger, D., Roustant, O., Haftka, R. T., and Kim, N.-H. (2010). "Adaptive Designs of Experiments for Accurate Approximation of a Target Region". In: *Journal of Mechanical Design* 132.7. ISSN: 1050-0472. DOI: 10.1115/1.4001873.

Plaschkies, F. (2023). *Utilizing AI to Improve Prediction Quality for Reduced Run Time Dummy Models*. In cooperation with Humanetics Innovative Solutions. Würzburg, Germany: Safety Week, carhs.

Plaschkies, F. (2024a). *Codebase AiHIII*. DOI: 10.5281/zenodo.14610182.

Plaschkies, F. (2024b). *Collection of Trained Metamodels for the Prediction of Injury Data of Human Surrogates in Car Crashes*. DOI: 10.5281/zenodo.14608924.

Plaschkies, F. (2024c). *FE-model Simplification Chain of Honda Accord 2014 for Occupant Safety Assessment in Front Load Cases*. DOI: 10.5281/zenodo.13998073. URL: https://doi.org/10.5281/zenodo.13998073.

Plaschkies, F. (2024d). *Occupant Simulation Data based on Honda Accord 2024 Simplified Passenger Model and Full-factorial Sampling with 3,125 samples and HIII05F, HIII50M, HIII95M*. DOI: 10.5281/zenodo.14000052. URL: https://doi.org/10.5281/zenodo.14000052.

Plaschkies, F. (2024e). *Occupant Simulation Database and FE-Model based on Honda Accord 2024 Simplified Passenger Model and SOBOL Sampling with 256 samples and HIII05F, HIII50M, HIII95M*. DOI: 10.5281/zenodo.14000145. URL: https://doi.org/10.5281/zenodo.14000145.

Plaschkies, F. (2024f). *Occupant Simulation Database and FE-Model based on Honda Accord 2024 Simplified Passenger Model and SOBOL Sampling with 8,192 samples and HIII05F, HIII50M, HIII95M*. DOI: 10.5281/zenodo.13998257. URL: https://doi.org/10.5281/zenodo.13998257.

Plaschkies, F. and Müller, S. (2024). *Occupant Simulation Data based on Honda Accord 2024 Simplified Passenger Model and Full-factorial Sampling with 243 samples and VIRTHUMAN 5, 50, 95 Percentiles*. DOI: 10.5281/zenodo.14000200. URL: https://doi.org/10.5281/zenodo.14000200.

Plaschkies, F., Possoli, K., Vaculín, O., Schumacher, A., and Andrade Junior, P. d. (2023). "Evaluation Approach for Machine Learning Concepts in Occupant Protection Based on Multi-Attribute Decision Making". In: *27th International Technical Conference on the Enhanced Safety of Vehicles* (Yokohama, Japan). URL: https://www-nrd.nhtsa.dot.gov/departments/esv/27th (visited on Mar. 6, 2024).

Plaschkies, F., Stocker, F., Vaculin, O., Fischer, P., Menzel, T., and Yuan, Z. (2024a). *Workshop arTIco - Artificial Intelligence and Correlations*. DOI: 10.5281/zenodo.11656662.

Plaschkies, F., Stocker, F., and Yuan, Z. (2024b). *arTIco Framework*. Zenodo. DOI: 10.5281/zenodo.11657449. URL: https://doi.org/10.5281/zenodo.11657449.

Plaschkies, F. and Vaculín, O. (2020). "Estimation of the Impact of Human Body Variation on its Crash Behavior Using Machine Learning Methods". In: *FISITA Web Congress* (Online). DOI: 10.46720/F2020-PIF-051.

Powers, D. M. W. (2011). “Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation”. In: *Journal of Machine Learning Technologies* 2.1, pp. 37–63. ISSN: 2229–3981.

Praxl, N. (2011). “How Reliable Are Injury Risk Curves?” In: *22nd International Technical Conference on the Enhanced Safety of Vehicles* (Washington, DC., US). URL: https://www-nrd.nhtsa.dot.gov/departments/esv/22nd/files/22nd%7B%5C_%7DESV%7B%5C_%7DProceedings%7B%5C_%7DBooklet.pdf (visited on Feb. 5, 2024).

Python Software Foundation (2024). *The Python Language Reference*. Version Version 3.11. URL: http://www.python.org (visited on Jan. 4, 2024).

Rabus, M., Belaid, M. K., Maurer, S. A., and Hiermaier, S. (2022). “Development of a model for the prediction of occupant loads in vehicle crashes: introduction of the Real Occupant Load Criterion for Prediction (ROLCp)”. In: *Automotive and Engine Technology* 7.3-4, pp. 229–244. ISSN: 2365-5127. DOI: 10.1007/s41104-022-00111-x.

Radja, G. A., Noh, E.-Y., and Zhang, F. (2022). *Crash Investigation Sampling System. 2020 Analytical User’s Manual*. DOT HS 813 243. URL: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813243 (visited on Apr. 5, 2024).

Rahman, N. I. A., Dawal, S. Z. M., Yusoff, N., and Kamil, N. S. M. (2018). “Anthropometric Measurements Among Four Asian Countries in Designing Sitting and Standing Workstations”. In: *Sādhanā* 43.10, pp. 1–9. ISSN: 0973-7677. DOI: 10.1007/s12046-017-0768-8.

Raja, T. T. and Swain, A. K. (2017). “Investigation of the Vehicle Restraint System in a Frontal Impact”. In: *International Journal of Crashworthiness* 22.6, pp. 662–675. ISSN: 1358-8265. DOI: 10.1080/13588265.2017.1301071.

Ramachandra, R., Kang, Y.-S., Stammen, J. A., Moorhouse, K., Murach, M., Bolte IV, J. H., and Agnew, A. M. (2019). “Evaluation of Skeletal and Soft Tissue Contributions to Thoracic Response of GHBMC M50-O Model in Dynamic Frontal Loading Scenarios”. In: *Proceedings of the International Research Council on the Biomechanics of Injury Conference*, pp. 332–348. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc19/pdf-files/49.pdf.

Reed, M. P. and Ebert, S. M. (2013). *Elderly Occupants: Posture, Body Shape, and Belt Fit*. UMTRI-2013-26. URL: https://deepblue.lib.umich.edu/handle/2027.42/134392 (visited on Feb. 15, 2024).

Reed, M. P. and Flannagan, C. A. C. (2000). “Anthropometric and Postural Variability: Limitations of the Boundary Manikin Approach”. In: *SAE Technical Paper Series*. ISSN: 0148-7191. DOI: 10.4271/2000-01-2172.

Reed, M. P. and Rupp, J. D. (2013). “An Anthropometric Comparison of Current ATDs with the U.S. Adult Population”. In: *Traffic Injury Prevention* 14.7, pp. 703–705. DOI: 10.1080/15389588.2012.752819.

Richardson, R., Kerrigan, J. R., Forman, J. L., Shaw, G. C., Gepner, B. D., Pipkorn, B., Overby, B., Kopp. Kevin, Östling, M., Mroz, K., Donlon, J. P., Chastain, K., Sochor, S., and Jayathirtha, M. (2019). “Test Methodology for Evaluating the Reclined Seating Environment with Human Surrogates”. In: *26th International Technical Conference on the Enhanced Safety of Vehicles* (Eindhoven, Netherlands). URL: https://www-nrd.nhtsa.dot.gov/pdf/

ESV/Proceedings/26/26th%7B%5C_%7DESV%7B%5C_%7DProceedings%7B%5C_%7DBooklet.pdf (visited on May 17, 2024).

Robbins, D. H. (1983). *Anthropometric Specifications for Mid-Sized Male Dummy. Volume 2*. UMTRI-83-53. URL: https://deepblue.lib.umich.edu/handle/2027.42/260 (visited on Feb. 15, 2024).

Robin, S. (2001). "HUMOS: HUman MOdel For Safety - A Joint Effort Towards The Development Of Refined Car Occupant Models". In: *17th International Technical Conference on the Enhanced Safety of Vehicles* (Amsterdam, Netherlands). URL: https://www-esv.nhtsa.dot.gov/Proceedings/17/17th%7B%5C_%7DESV%7B%5C_%7DProceedings%7B%5C_%7DBooklet.pdf (visited on Feb. 15, 2024).

Robinette, K. M., Blackwell, C. L., Daanen, H., and Boehmer, M. (2002). *Civilian American and European Surface Anthropometry Resource (CAESAR). Final Report*. Vol. 1. AFRL-HE-WP-TR-2002-0169. URL: https://apps.dtic.mil/sti/pdfs/ADA406704.pdf (visited on Feb. 15, 2024).

Rupp, J. D., Flannagan, C. A. C., Leslie, A. J., Hoff, C. N., Reed, M. P., and Cunningham, R. M. (2013). "Effects of BMI on the Risk and Frequency of AIS 3+ Injuries in Motor-vehicle Crashes". In: *Obesity* 21.1, E88–97. DOI: 10.1002/oby.20079.

Ryb, G. E. and Dischinger, P. C. (2008). "Injury Severity and Outcome of Overweight and Obese Patients After Vehicular Trauma: A Crash Injury Research and Engineering Network (CIREN) Study". In: *The Journal of Trauma* 64.2, pp. 406–411. DOI: 10.1097/TA.0b013e31802beff9.

Safety Companion (2024). Version 1.5. Alzenau, Germany: carhs.training GmbH.

Safety Test Instrumentation Standards Committee, ed. (Aug. 19, 2022). *Instrumentation for Impact Test - Part 1 - Electronic Instrumentation*. J211/1_202208. Standard. Version 1. SAE International. Warrendale, PA, United States. DOI: 10.4271/J211/1_202208.

Salike, S. S. (2024). "Data Abstraction and Analysis for Universalizing Passenger Cell Model Interface for Occupant Protection Optimization". Masterthesis. Technische Hochschule Ingolstadt.

Schäffer, M., Sturm, R., and Friedrich, H. E. (2019a). "Automated generation of physical surrogate vehicle models for crash optimization". In: *International Journal of Mechanics and Materials in Design* 15.1, pp. 43–60. ISSN: 1569-1713. DOI: 10.1007/s10999-018-9407-8.

Schäffer, M., Sturm, R., and Friedrich, H. E. (2019b). "Methodological approach for reducing computational costs of vehicle frontal crashworthiness analysis by using simplified structural modelling". In: *International Journal of Crashworthiness* 24.1, pp. 39–53. ISSN: 1358-8265. DOI: 10.1080/13588265.2017.1389631.

Schmitt, K.-U., Niederer, P. F., Cronin, D. S., Morrison III, B., Muser, M. H., and Walz, F. (2019). *Trauma Biomechanics*. Cham: Springer International Publishing. ISBN: 978-3-030-11658-3. DOI: 10.1007/978-3-030-11659-0.

Schneider, B., Kofler, D., D'Addetta, G. A., Freienstein, H., Wolkenstein, M., and Klug, C. (2022). "Approach for Machine Learning Based Design of Experiments for Occupant Simulation". In: *Frontiers in Future Transportation* 3. DOI: 10.3389/ffutr.2022.913852.

Schneider, K. (2023). “Prospektive Approximation der Unfallschwere für Integrale Fahrzeugsicherheitssysteme”. Electrical Engineering and Information Technology. Dissertation. Otto von Guericke University Magdeburg. URL: 10.25673/101902.

Schneider, L. W., Robbins, D. H., Pflüg, M. A., and Snyder, R. G. (1983). *Development of Anthropometrically Based Design Specifications for an Advanced Adult Anthropomorphic Dummy Family. Volume 1*. UMTRI-83-53-1. URL: https://deepblue.lib.umich.edu/handle/2027.42/259 (visited on Feb. 15, 2024).

Schneider, S., Peldschus, S., Fuchs, T., Fürst, F., and Yigit, E. (2017). “Factors Influencing the Validation of FE Human Body Models”. In: *12. VDI-Tagung Fahrzeugsicherheit* (Berlin, Deutschland). VDI Wissensforum GmbH. VDI Wissensforum GmbH.

Schoell, S. L., Weaver, A. A., Urban, J. E., Jones, D. A., Stitzel, J. D., Hwang, E., Reed, M. P., Rupp, J. D., and Hu, J. (2015). “Development and Validation of an Older Occupant Finite Element Model of a Mid-Sized Male for Investigation of Age-related Injury Risk”. In: *Stapp Car Crash Journal* 59, pp. 359–383. ISSN: 1532-8546. DOI: 10.4271/2015-22-0014.

Schöneburg, R. (2023). *Integrale Sicherheit von Kraftfahrzeugen*. Wiesbaden: Springer. ISBN: 978-3-658-42805-1. DOI: 10.1007/978-3-658-42806-8.

Schwartz, D., Guleyupoglu, B., Koya, B., Stitzel, J. D., and Gayzik, F. S. (2015). “Development of a Computationally Efficient Full Human Body Finite Element Model”. In: *Traffic Injury Prevention* 16 Suppl 1, pp. 49–56. DOI: 10.1080/15389588.2015.1021418.

Sequeira, G. J. A. (2022). “Prediction based activation of vehicle safety systems - A contribution to improve occupant safety by validation of pre-crash information and crash severity plus restraint strategy prediction”. Electrical Engineering and Information Technology. Dissertation. Magdeburg, Germany: Otto von Guericke University Magdeburg. URL: 10.51202/9783186817129.

Sequeira, G. J. A., Patel, A., Afraj, S., Lugner, R., and Brandmeier, T. (2020). “FEM-based Methodology for Crash Severity Estimation in Frontal Crash Scenarios”. In: *IOP Conference Series: Materials Science and Engineering* 831, p. 012019. DOI: 10.1088/1757-899X/831/1/012019.

Serre, T., Brunet, C., Bruyere, K., Verriest, J.-P., Mitton, D., Bertrand, S., and Skalli, W. (2006). “HUMOS (Human Model for Safety) Geometry: From One Specimen to the 5th and 95th Percentile”. In: *SAE Technical Paper Series*. ISSN: 0148-7191. DOI: 10.4271/2006-01-2324.

Settles, B. (2010). *Active Learning Literature Survey*. TR1648. URL: http://digital.library.wisc.edu/1793/60660 (visited on Feb. 15, 2024).

Shi, X., Cao, L., Reed, M. P., Rupp, J. D., and Hu, J. (2015). “Effects of Obesity on Occupant Responses in Frontal Crashes: a Simulation Analysis Using Human Body Models”. In: *Computer Methods in Biomechanics and Biomedical Engineering* 18.12, pp. 1280–1292. DOI: 10.1080/10255842.2014.900544.

Shigeta, K., Kitagawa, Y., and Yasuki, T. (2009). “Development of Next Generation Human FE Model Capable of Organ Injury Prediction”. In: *21th International Technical Conference on the Enhanced Safety of Vehicles* (Stuttgart, Germany). URL: https://www-esv.nhtsa.dot.gov/Proceedings/21/21st%7B%5C_%7DESV%7B%5C_%7DProceedings%7B%5C_%7DBooklet.pdf (visited on Feb. 15, 2024).

Siebertz, K., Bebber, D. van, and Hochkirchen, T. (2017). *Statistische Versuchsplanung. Design of Experiments (DoE)*. 2nd ed. VDI-Buch. Berlin, Heidelberg: Vieweg. 1518 pp. ISBN: 978-3-662-55742-6.

Siemens Digital Industries Software (2024a). *Simcenter*. URL: https://plm.sw.siemens.com/en-US/simcenter/simulation-test/occupant-safety-simulation/ (visited on Apr. 8, 2024).

Siemens Digital Industries Software (2024b). *Simcenter MADYMO Example Model 4-point Harness Belt*. URL: https://community.sw.siemens.com/s/article/Simcenter-Madymo-Example-model-4-point-seat-belt (visited on Nov. 19, 2024).

Singh, H., Davies, J., Paramasuwom, M., and Gradischnig, L. (2018a). *Vehicle Interior and Restraints Modeling Development of Full Vehicle Finite Element Model Including Vehicle Interior and Occupant Restraints Systems For Occupant Safety Analysis Using THOR Dummies*. DOT HS 812 545. URL: https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/812545_edagvehicleinteriorandrestraintsmodelingreport.pdf (visited on Mar. 19, 2024).

Singh, H., Ganesan, V., Davies, J., Paramasuwom, M., Gradischnig, L., Wood, P., and Mogal, V. (2018b). *Structural Countermeasure/ Research Program: Mass and Cost Increase Due to Oblique Offset Moving Deformable Barrier Impact Test*. DOT HS 812 523. URL: https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/13377-edag_crashdeformation_040418_v3_tag.pdf (visited on Mar. 19, 2024).

Suzuki, T. (2022). *dtwalign*. Version 0.1.1. URL: https://pypi.org/project/dtwalign/.

Takata Corporation (2017). *Advanced Adaptive Restraint Systems*. DOT HS 812 432. URL: https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/812432_adaptiverestraintreport.pdf (visited on Apr. 5, 2014).

"A2MAC1 and TECOSIM Have Joined Forces to Offer Unique Synergies Driving Significant Value for their Customers" (Oct. 7, 2022). In: ed. by TECOSIM Group GmbH and A2MAC1 Group. URL: https://www.a2mac1.com/newsroom/a2mac1-and-tecosim-have-joined-forces-to-offer-unique-synergies-driving-significant-value-for-their-customers/ (visited on Mar. 19, 2024).

Tekkaya, A. E. and Soyarslan, C. (2014). "Finite Element Method". In: *CIRP Encyclopedia of Production Engineering*. Ed. by L. Laperrière and G. Reinhart. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 508–514. ISBN: 978-3-642-20616-0. DOI: 10.1007/978-3-642-20617-7_16699.

Tennant, J. A., Jensen, R. H., and Potter, R. A. (1974). "GM-ATD 502 Anthropomorphic Dummy - Development and Evaluation". In: *SAE Technical Paper Series*. ISSN: 0148-7191. DOI: 10.4271/740590.

TensorFlow Developers (2024). *TensorFlow*. Zenodo. DOI: 10.5281/zenodo.12586505.

THUMS User Community (2019). *Frontal Sled Gold Standard*. URL: https://tuc-project.org/frontal-sled-gold-standard/ (visited on Mar. 19, 2024).

Thunert, C. (2017). *CORAplus. User's Manual*. Version 4.0.4. PDB and GNS.

Toyota (June 16, 2020). "Toyota Offers Free Access to THUMS Virtual Human Body Model Software. Greater use of THUMS for Analysis of Vehicle Collision-related Injuries to Enhance Vehicle Safety". In: URL: https://global.toyota/en/newsroom/corporate/32665896.html (visited on Jan. 6, 2024).

Toyota (Feb. 28, 2023). “Toyota’s Total Human Model for Safety Version 7 Released”. In: URL: https://www.toyota-europe.com/news/2023/total-human-model (visited on Jan. 26, 2024).

“TMC Adds New Variations to Virtual Crash Dummies” (2011). “Small Female, Large Male Join THUMS Lineup”. In: ed. by Toyota Motor Corporation. URL: https://global.toyota/en/detail/244409 (visited on Nov. 19, 2024).

U.S. Department of Health & Human Services, ed. (2021). *National Health and Nutrition Examination Survey*. URL: https://www.cdc.gov/nchs/nhanes/index.htm (visited on Apr. 5, 2021).

“THOR 50th Male Finite Element Model” (2024). In: ed. by University of Virginia School of Engineering and Applied Science, Center for Applied Biomechanics. URL: https://www.engineering.virginia.edu/centers-institutes/center-applied-biomechanics/blogs/thor-50th-male-finite-element-model (visited on Mar. 21, 2024).

Vandenhende, S., Georgoulis, S., van Gansbeke, W., Proesmans, M., Dai, D., and van Gool, L. (2022). “Multi-Task Learning for Dense Prediction Tasks: A Survey”. In: *IEEE transactions on pattern analysis and machine intelligence* 44.7, pp. 3614–3633. DOI: 10.1109/TPAMI.2021.3054719. eprint: 33497328.

Vavalle, N. A. (2015). “The Application of a Radial Basis Function Morphing Technique to a Validated Full Human Body Finite Element Model”. School of Biomedical Engineering & Sciences. Dissertation. Winston-Salem, North Carolina: Virginia Tech – Wake Forest University.

Vezin, P. and Verriest, J.-P. (2005). “Development of a Set of Numerical Human Models for Safety”. In: *19th International Technical Conference on the Enhanced Safety of Vehicles* (Washington DC, United States). URL: https://www-esv.nhtsa.dot.gov/Proceedings/19/19th%7B%5C_%7DESV%7B%5C_%7DProceedings%7B%5C_%7DBooklet.pdf (visited on Feb. 5, 2024).

Vink, R., Gooijer, S. de, Beedie, A., Gorelli, M. E., Guo, W., Peters, O., Zundert, J. van, nameexhaustion, Hulselmans, G., Grinstead, C., Burghoorn, G., Marshall, chielP, Mitchell, L., Turner-Trauring, I., Santamaria, M., Heres, D., Magarick, J., Harbeck, H., ibENPC, Genockey, K., Wilksch, M., deanm, Leitao, J., Gelderen, M. van, Barbagiannis, P., Koutsouris, I., Haag, J., and Borchert, O. (2024). *polars: Python Polars 1.7.1*. Zenodo. DOI: 10.5281/zenodo.13754515.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., and van Mulbregt, P. (2020). “SciPy 1.0: fundamental algorithms for scientific computing in Python”. In: *Nature methods* 17.3, pp. 261–272. DOI: 10.1038/s41592-019-0686-2. eprint: 32015543.

Vychytil, J., Manas, J., Cechova, H., Spirk, S., Hynčík, L., and Kovar, L. (2014). “Scalable Multi-Purpose Virtual Human Model for Future Safety Assessment”. In: *SAE Technical Paper Series*. ISSN: 0148-7191. DOI: 10.4271/2014-01-0534.

Wang, J. (2023). "An Intuitive Tutorial to Gaussian Process Regression". In: *Computing in Science & Engineering* 25.4, pp. 4–11. ISSN: 1521-9615. DOI: 10.1109/MCSE.2023.3342149.

Wang, S. C., Holcombe, S. A., Derstine, B. A., Goulson, R. L., Grenda, D., Ruan, J., Rabban, P., Wang, N. C., Friedm, J. F., Cron, D. C., Henderson, B. C., Sullivan, J. A., Kohoyda-Inglis, C., Su, G. L., Ejima, S., and Zhang, P. (2016). "Reference Analytic Morphomics Population (RAMP): A Reference to Measure Occupant Variability for Crash Injury Analysis". In: *Proceedings of the International Research Council on the Biomechanics of Injury Conference*, pp. 582–591. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc16/pdf-files/80.pdf (visited on Feb. 15, 2024).

Waskom, M. (2021). "seaborn: statistical data visualization". In: *Journal of Open Source Software* 6.60, p. 3021. DOI: 10.21105/joss.03021.

Watanabe, S. and Hutter, F. (2023). "c-TPE: Tree-structured Parzen Estimator with Inequality Constraints for Expensive Hyperparameter Optimization". In: *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI-23)* (Macau, China). Ed. by P. Stone and E. Elkind, pp. 4371–4379. ISBN: 978-1-956792-03-4. DOI: 10.24963/ijcai.2023/486.

Wielens, S. (2022). *Automatische Erstellung von Submodellen für die Crashoptimierung von Fahrzeugkarosserien*. Dissertation. University of Wuppertal, School of Mechanical Engineering and Safety Engineering, Shaker Verlag. ISBN: 978-3-8440-8717-8.

Wikipedia Contributors (2024). *Comparison of Data-serialization Formats – Wikipedia, The Free Encyclopedia*. Ed. by Wikipedia Contributors, 2024. URL: https://en.wikipedia.org/w/index.php?title=Comparison_of_data-serialization_formats&oldid=1247693950 (visited on Dec. 12, 2024).

Wu, D. and Hartmann, O. (2020). "Automatic Evaluation of LS-DYNA® Simulation Results Using Statistical Database and Python". In: *16th International LS-DYNA Conference* (Online).

Xu, L., Jensen, J., Byrnes, K., Kim, A., Agaram, V., Davis, K. L., Hultman, R. W., Kostyniuk, G., Marshall, M. E., Mertz, H., Nusholtz, G., Rouhana, S. W., and Scherer, R. (2000). "Comparative Performance Evaluation of THOR and Hybrid III". In: *SAE Technical Paper Series*. ISSN: 0148-7191. DOI: 10.4271/2000-01-0161.

Yang, K. H., Hu, J., White, N. A., King, A. I., Chou, C. C., and Prasad, P. (2006). "Development of Numerical Models for Injury Biomechanics Research: A Review of 50 Years of Publications in the Stapp Car Crash Conference". In: *Stapp Car Crash Journal* 50, pp. 429–490. ISSN: 1532-8546. DOI: 10.4271/2006-22-0017.

Yang, S., Zhou, Q., Yasuki, T., Yamamae, Y., Katsuhara, T., and Nie, B. (2020). "A Look–up Diagram for Rapid Estimation of Occupant Injury Risk Applied to Frontal Motor Vehicle Crashes". In: *Proceedings of the International Research Council on the Biomechanics of Injury Conference*, pp. 29–32. ISSN: 2235-3151. URL: https://www.ircobi.org/wordpress/downloads/irc20-asia/pdf-files/2017a.pdf (visited on Feb. 15, 2024).

Yoganandan, N., Pintar, F. A., Zhang, J., and Baisden, J. L. (2009). "Physical Properties of the Human Head: Mass, Center of Gravity and Moment of Inertia". In: *Journal of Biomechanics* 42.9, pp. 1177–1192. ISSN: 1873-2380. DOI: 10.1016/j.jbiomech.2009.03.029.

Zhao, B., Lu, H., Chen, S., Liu, J., and Wu, D. (2017). “Convolutional neural networks for time series classification”. In: *Journal of Systems Engineering and Electronics* 28.1, pp. 162–169. ISSN: 10044132. DOI: 10.21629/JSEE.2017.01.18.

Zhao, Z., Alzubaidi, L., Zhang, J., Duan, Y., and Gu, Y. (2024). “A Comparison Review of Transfer Learning and Self-supervised Learning: Definitions, Applications, Advantages and Limitations”. In: *Expert Systems with Applications* 242, p. 122807. ISSN: 09574174. DOI: 10.1016/j.eswa.2023.122807.

Zhao, Z., Jin, X., Cao, Y., and Wang, J. (2010). “Data Mining Application on Crash Simulation Data of Occupant Restraint System”. In: *Expert Systems with Applications* 37.8, pp. 5788–5794. ISSN: 09574174. DOI: 10.1016/j.eswa.2010.02.029.

A Appendix – Used Software and Packages

In table A.1, the essential PYTHON packages are listed. The selection was limited to the packages, directly used by the author. Packages used by the listed ones were excluded.

The MIT (Massachusetts Institute of Technology) licence is highly permissive and allows reusing the code, even in commercial applications. The BSD (Berkeley Software Distribution) licences are highly permissive, too. In both cases, the copyright of the code should not be removed. The BSD-3 licence has an additional clause that prohibits the use of the package-author's name or affiliation for promotion and advertising purposes. Similar is the Apache licence. The most notable difference is the necessity to mention changes to the original code. Again, similar is the PSF (Python Software Foundation) licence.

Table A.1: Essential Top-level PYTHON Packages

Package	Version	Licence Name
CAMELOT-PY	0.9.0	MIT
DOEPY	0.0.1	MIT
DTWALIGN	0.1.1	MIT
H5PY	3.10.0	BSD-3
IPYKERNEL	6.29.4	BSD-3
IPYTHON	8.24.0	BSD
IPYWIDGETS	8.1.3	BSD
KERAS	3.0.5	Apache Software Licence 2.0
LASSO-PYTHON	2.0.3	BSD-3
MATPLOTLIB	3.9.0	PSF
NUMPY	1.26.4	BSD
OPENPYXL	3.1.2	MIT
OPTUNA	3.6.1	MIT
PANDAS	2.2.2	BSD

Continued on next page

Table A.1: Essential Top-level PYTHON Packages (Continued)

Package	Version	Licence Name
CAMELOT-PY	0.9.0	MIT
PLOTLY	5.22.0	MIT
POLARS	1.7.1	MIT
PYARMOR	9.0.7	Free To Use But Restricted
PYARROW	16.1.0	Apache Software Licence
PYDOE	0.3.8	BSD
PYDOT	3.0.1	MIT
PYPDF	4.2.0	BSD
SCIKIT-LEARN	1.5.1	BSD
SCIPY	1.13.0	BSD
SEABORN	0.13.2	BSD
SKTIME	0.31.0	BSD
TENSORFLOW	2.16.1	Apache Software Licence
TSFRESH	0.20.2	MIT
XARRAY	2024.6.0	Apache Software Licence 2.0
XGBOOST	2.1.1	Apache Software Licence 2.0

In table A.2, relevant software for generating and analysing the technical results is listed. Special thanks go to the companies BETA CAE Systems and ANSYS for providing the software with an academic discount.

Table A.2: Used Essential Software

Software	Version	Provider	Licence
ANSA	24.1.1	BETA CAE Systems	Academic
CORAPLUS	4.0.5	Partnership for Dummy Technology and Biomechanics GbR	Free
LS-DYNA	12.0	ANSYS, Inc	Academic
META	24.1.1	BETA CAE Systems	Academic
PYTHON	3.11	Python Software Foundation	PSF

B Appendix – Overview Codebase

The PYTHON codebase was published by Plaschkies (2024a) and is available on GitHub https://github.com/frapla/aihiii.git as well. The codebase is structured in multiple areas. The "data" section contains the databases used for machine learning. In "experiments", the results of the trials are located. The "jupyter" directory contains all the jupyter notebooks with the file extension ipynb. "Source" contains the actual code base with all files with the extension py.

The central packages in the "source" section are "data" to deal with FE-simulations and database generation, "experiments" as control files to run machine learning trials, and "evaluate" as the heart of the framework.

To enable other researchers to use the code base, the studies which were presented in this thesis should be linked with the scripts. In the tables B.1, B.2, and B.3 is the mapping presented. During the development activities, the nomenclature has shifted. The author is positive that with the provided tables a researcher can orient.

Table B.1: Used Essential Software in Chapter 2

Study	Generation & Evaluation
Subsubsection 2.1.5.1	400_ISO_Validation.ipynb
Subsubsection 2.1.2.3	014_FARS_third_insight.ipynb
Subsection 2.2.2	persistant_storage.ipynb

Table B.2: Used Essential Software in Chapter 3

Study	**Generation** (in “data”)	**Evaluation**
Subsection 3.1.4	054_Eval_ValChain_THOR_HIII_vs_HIII _Speed_Corridors_Selection.ipynb	
Section 3.2	054_Eval_ValChain_THOR_HIII_vs_HIII _Speed_Corridors_Selection.ipynb	
Subsection 3.2.8	049_Intrusions.ipynb	
Section 3.3	054_Eval_ValChain_THOR_HIII_vs_HIII _Speed_Corridors_Selection.ipynb	
Subsubsection 3.4.1	080_Eval_Parameter_Influence.ipynb	
Subsubsection 3.4.3.2	doe_generation/SobolDoe.py	analyze_injury_data _sobol_final.ipynb
Subsubsection 3.4.3.3	doe_generation/DoeGenerators.py	analyze_injury_data _full_fact_final.ipynb
Subsubsection 3.4.3.4	doe_generation/SobolDoe.py	analyze_injury_data _test_final.ipynb
Subsubsection 3.4.3.5	analyze_injury_data_virthuman_final.ipynb	

Table B.3: Used Essential Software in Chapter 4

Study	**Generation** (in “experiments”)	**Evaluation**
Section 4.2	baseline.py	baseline.ipynb
Subsubsection 4.3.1.2	cnn_optuna_2.py	cnn_evaluation.ipynb
Subsubsection 4.3.2.2	lstm_treesearch.py	evaluate_lstm_treesearch.ipynb
Subsection 4.4.1	ann_preextracted.py	
Subsection 4.4.2	rocket_ann.py	
Subsection 4.4.3	pca_ann.py	eval_pca_features.ipynb
Section 4.5	pure_cnn.py	cnn_evaluation.ipynb
Subsection 4.5.3	test_set.py	eval_test_set.ipynb
Subsection 4.6.1	sobol_size.py	n_samples_eval.ipynb
Figure 4.12		characterize_50th.ipynb
Subsection 4.6.2	knn_sampler.py	n_samples_eval.ipynb
Figure 4.15		characterize_50th.ipynb
Subsubsection 4.6.3.1	transfer_learner.py	
Subsubsection 4.6.3.2	cnn_hiii2vh.py	virthuman_hiii_intercorrelation.ipynb