Microscopic and Macroscopic Risk Metrics for the Safety Validation of Automated Driving

Vom Fachbereich Maschinenbau an der
Technischen Universität Darmstadt
zur Erlangung des Grades eines
Doktor-Ingenieurs (Dr.-Ing.)
genehmigte

Dissertation

eingereicht von

Philipp Matthias Junietz, M.Sc.
aus Münster

Berichterstatter:	Prof. Dr. rer. nat. Hermann Winner
Mitberichterstatter:	Prof. Dr.-Ing. Christoph Stiller
Tag der Einreichung:	04.07.2019
Tag der mündlichen Prüfung:	06.11.2019

Darmstadt 2019
D 17

Danksagung

Nach vier Jahren FZD ist es nun geschafft. Ich möchte mich bei meinen Unterstützern sehr herzlich bedanken:

Prof. Winner, Sie haben mich während der gesamten Zeit immer mit fachlichem Rat unterstützt und durch Ihre intensive und vorbildliche Betreuung diese Dissertation mit ermöglicht.

Prof. Stiller, Danke für die Übernahme des Korreferats und die Gelegenheit meine Forschungsarbeit inhaltlich zu diskutieren.

Allen Projektpartnern des Projekts PEGASUS des BMWi für die fachliche Diskussion während des Projektes und die gute Zusammenarbeit.

Meinen Kollegen bei FZD, Ihr habt für ein gutes Arbeitsklima gesorgt und hattet immer ein offenes Ohr bei Problemen. Das Korrekturlesen, Eure Tipps und kritischen Fragen in Doktorandenseminaren waren ungemein wertvoll.

Meinen Studenten bei FZD, Ihr habt während Eurer Abschlussarbeiten viele Stunden über Metriken und Risiko diskutiert. Als HiWi habt Ihr viele kleinere und größere Aufgaben erledigt und Forschung und Projekt so weitergebracht.

Meinen Freunden, Ihr habt immer ein offenes Ohr und einen guten Rat für mich übrig. Durch nette Abende und interessante Gespräche sorgt Ihr immer für Entlastung nach stressigen Phasen.

Meiner Familie und ganz besonders meinen Eltern, Sabine und Volker, Ihr habt mich immer unterstützt und hattet ein offenes Ohr.

Meiner Freundin Elena, Du unterstützt mich immer und stehst mir jeden Tag mit Rat und Tat zur Seite. Du gibst mir Halt und Ausgleich, nicht nur wenn es stressig wird, sondern in jeder Lebenslage.

Table of Contents

Table of Figures

Table of Abbreviations

Abbreviation	Description
TTC	Time-to-Collision
TTX	Time-to-X
TTB	Time-to-Brake
TTS	Time-to-Steer
TTR	Time-to-React
TTEC	Time-to-Edge-Crossing
PET	Post-Encroachment-Time
MiR	Microscopic Risk
MaR	Macroscopic Risk
AD2-	Assisted Driving at SAE Level 2 and lower
AD3+	Automated Driving at SAE Level 3 and higher
TTCE	Time-to-closest-Encounter
THW	Time Headway
OuT	Object under Test
SP	Safety Performance
SSM	Surrogate Safety Model
NDS	Naturalistic Driving Study
FOT	Field Operational Test
EVT	Extreme Value Theory
GAMAB	Globalement au moin aussi bon
ALARP	As low as reasonably possible
MEM	Minimum Endogenous Mortality
TCI	Trajectory Criticality Index

Formula Symbols and Indices

Symbol	Unit	Description
$\underline{u}$		Input vector
$\underline{x}$		State vector
a	m/s²	Acceleration
C	./.	Criticality (component)
d	m	Relative distance
f	1/a	yearly accident frequency
$\mathcal{F}$	1/m	Distance-based accident frequency
g	m/s²	Gravitational constant
G	./.	Generalized Pareto Distribution
I	./.	Index (as in criticality index)
i	./.	discrete step number
J	./.	Cost Function
k	./.	number of events e.g. accidents
K	€	Monetary Costs
m	./.	*m*-observation return level
N	./.	Number of Years
n	./.	Number of measurements
P	./.	Probability distribution
s	m	Travelled distance
t	s	Time
T	s	Duration of period
v	m/s	Velocity
V	./.	Variance
w	./.	weighting factor
x	m	Position (longitudinal)
z	./.	Variable in Generalized Pareto Distribution
λ	./.	Probability for an observation over threshold
η	./.	Market share
κ	1/m	Curvature
λ	./.	Location parameter in Generalized Pareto Distribution
μ	./.	Friction coefficient
ξ	./.	Shape parameter in Generalized Pareto Distribution
σ	./.	Scaling parameter in Generalized Pareto Distribution
τ	s	Timespan e.g. until a reaction
ψ	rad	Yaw angle

Φ	./.	Cumulative distribution function

Index	**Description**
0	initial
2016	year of 2016
a	acceleration
a	annual
acc	accepted
d	fatal accidents
e	ego perspective (as left index)
ego	referring to ego vehicle
exp	expected (e.g. from extrapolation)
gamab	Derived from principle “Globalement au moins aussi bon”
GER	Germany
i	*i*-th element
inv	for involuntary exposure
lat	lateral
lin	based on linear behavior
long	longitudinal
m	Corresponding to *m*-observation return level
MEM	Derived from principle “Minimal Endogenous Mortality”
new	for new risks
nI	without Injuries
obj	referring to object vehicle
occ	occured
old	previous value
P	Precision
R	reaction
rel	relative
RP	reaction due to disturbance in precision of course angle
soc	from society’s perspective
T	trajectory (as in Trajectory Criticality Index)
TB	to brake
TC	to collision
thr	threshold
tot	total
TR	to react
TS	to steer
User	from User perspective
w	world fixed (as left index)

wI	with Injuries
x	in x-direction (longitudinal)
y	in y-direction (lateral)
κ	due to curvature
σ	standard deviation
τ	combined reaction time-based components
Φ	based on cumulative distribution

Deutsche Zusammenfassung

Automatisiertes Fahren ist einer der großen Trends in der Automobilindustrie. Neueste technologische Entwicklungen und Prototypen deuten darauf hin, dass die Einführung des automatisierten Fahrens technisch schon bald möglich ist. Trotz der Fortschritte in der Systementwicklung selbst, ist der Sicherheitsnachweis immer noch ungelöst. Ohne weitere Forschung und Weiterentwicklung von Test- und Validierungsmethoden ist eine sichere Einführung des automatisierten Fahrens nicht nachzuweisen. Derzeit verlangt die ECE-Typgenehmigung, dass das System für den Straßenbetrieb sicher sein muss und kein inakzeptables Risiko besteht. Dies basiert auf der Annahme, dass ein menschlicher Fahrer das Fahrzeug steuert und in kritischen Szenen die Kontrolle übernimmt. Für höher automatisierte Systeme ist die Überwachung durch den Fahrer jedoch nicht mehr vorgesehen. Daher müssen neue Methoden des Sicherheitsnachweises entwickelt werden, die die Sicherheit in automatisierter Fahrt ohne menschliche Überwachung gewährleisten.

Zunächst werden bestehende Validierungsmethoden für das automatisierte Fahren analysiert und strukturiert, um daraus offene Forschungsfragen zu Risikometriken abzuleiten. In dieser Arbeit liegt der Schwerpunkt dabei auf szenariobasiertem Testen und Feldtests. Ein Feldtest ist ein Test im realen Verkehr. Anhand der auftretenden Unfallhäufigkeit wird das Risiko des Systems in einem statistischen Ansatz abgeschätzt. Szenariobasiertes Testen erfordert hingegen die Identifizierung von Testfällen, um das Testen von der Straße in die Simulation oder in ein Testgelände zu verlagern.

Ziel jeder Sicherheitsvalidierung ist, das Risiko der Technologie abzuschätzen und Sicherheit gemäß den Anforderungen aller Gesellschaftsgruppen nachzuweisen. In dieser Arbeit werden zwei verschiedene Risikobegriffe verwendet. Das durchschnittliche Risiko eines Systems, z.B. die Häufigkeit tödlicher Unfälle, wird als *makroskopisches Risiko* (MaR) bezeichnet. Das Risiko in einer einzelnen Verkehrsszene wird als *mikroskopisches Risiko* (MiR) bezeichnet. Aufgrund der großen Entfernung zwischen zwei Unfällen im heutigen Verkehr kann das MaR nicht ohne umfangreiche Datenanalyse bestimmt werden. Da die durchschnittliche Entfernung zwischen zwei Unfällen hoch ist, Eine sehr große Testdistanz ist erforderlich, um genügend Unfalldaten für eine aussagekräftige statistische Auswertung zu sammeln. Daher ist die Frage entscheidend, wie MiR-Metriken, angewendet auf Szenen einer Messfahrt, zur Extrapolation von MaR verwendet werden können. Ein weiterer Verwendungszweck von MiR-Metriken ist die Identifizierung kritischer Szenen in aufgezeichneten Daten oder online während Testfahrten. Aus diesen Daten können Testfälle für den szenariobasierten Test abgeleitet werden.

Die drei wichtigsten Forschungsfragen dieser Arbeit befassen sich mit der Definition von MaR-Anforderungen und einem Top-Down-Ansatz zur Definition von MiR-Metriken, mit

denen das MaR aus kritischen Szenen extrapoliert und Testfälle identifiziert werden können. Diese drei Fragen werden im Verlauf der Arbeit noch in weitere Unterfragen aufgeteilt.

Was sind die Anforderungen an MiR-Metriken und wie kann die Eignung der Metriken für die Extrapolation von MaR und die Identifikation von Testfällen falsifiziert werden?

Bereits heute sind zahlreiche MiR-Metriken vorhanden, die zum Teil auch schon für die beiden beschriebenen Ansätze verwendet werden. Daher wird in dieser Arbeit ein Bewertungsprozess etabliert, der beide Ansätze abdeckt. Die Metrik soll geeignet sein, Szenarien zu identifizieren, die für den Fahrer oder die Automatisierung sehr anspruchsvoll sind, um daraus Testfälle abzuleiten. Gleichzeitig soll die Metrik das Risiko in einer Szene beschreiben, sodass eine Extrapolation von MiR zu MaR möglich ist. Um zu bewerten, ob eine Metrik für diese Zwecke geeignet ist, werden Anforderungen definiert. Da die Erfüllung aller Anforderung und damit die Eignung der Metrik nicht immer nachweisbar ist, wird stattdessen eine Falsifizierungsstrategie vorgeschlagen, die zwei Schritte enthält: Zunächst werden Testszenen definiert, die von einer Metrik korrekt bewertet werden müssen. Wenn keine Falsifizierung durch die Testszenen erreicht wird, wird die Metrik auf aufgezeichnete Daten von vom Menschen gefahrenen Fahrzeugen angewendet. Wenn die tatsächliche Unfallrate der Hochrechnung des Risikos auf der Grundlage der Metrik innerhalb statistischer Toleranzen widerspricht, ist die Eignung der Metrik falsifiziert. Da der letzte Falsifizierungsschritt sehr aufwändig ist, werden Design-Richtlinien festgelegt, die zu einer aussagekräftigen Metrik führen, wenn sie bei der Entwicklung der Metrik befolgt werden.

Welche Methoden und Metriken können verwendet werden, um das MaR des automatisierten Fahrens aus kritischen Szenen in Feldtests zu extrapolieren?

Bei Feldtests wird die Sicherheit eines Fahrsystems basierend auf der Auftretensrate von Ereignissen bestimmter Kategorie, z.B. Unfälle mit Todesfolge, mit Hilfe eines statistischen Nachweises ermittelt. Handelt es sich bei dem untersuchten Ereignis nicht um einen Unfall, sondern um eine kritische Szene, erhöht sich die Auftrittsrate; und bei gleicher statistischer Signifikanz sind weniger Kilometer erforderlich. Das Auftreten von kritischen Szenen allein gibt jedoch keine Auskunft über das Unfallrisiko. Wenn davon ausgegangen wird, dass der Fahrer eine Kritikalität über einer bestimmten Schwelle zu verhindern versucht und dass das Auftreten einer Kritikalität über diesem Wert aus der Beobachtung in einem Feldtest extrapoliert werden kann, wäre es möglich auf das Auftreten von Szenen mit noch höherer Kritikalität und sogar Unfällen hochzurechnen. Zunächst wird untersucht, ob bestehende Metriken, Datenerfassungen und Extrapolationsmethoden die abgeleiteten Anforderungen erfüllen. Anschließend wird die Extremwerttheorie als statistisches Instrument zur Extrapolation des Risikos ausgewählt. Voraussetzung ist, dass es sich bei hochkritischen Ereignissen um Extremereignisse handelt, die nicht durch eine Approximation einer Verteilung aller auftretenden Kritikalitätswerte angenähert werden können. Abschließend wird eine Metrik basierend auf diesen Designrichtlinien entwickelt. Die Metrik verwendet eine modellprädiktive Optimierung, um die Trajektorie mit den minimalen Fahranforderungen in einer bestimmten

Szene zu finden. Die Metrik erfüllt alle definierten Testfälle und vergleicht die Fahranforderungen mit dem geschätzten menschlichen Fahrkönnen. Nach den festgelegten Designrichtlinien wird eine Sensitivitätsanalyse für willkürlich festgelegte Parameter und den Einfluss der Parameterdefinition auf das Ergebnis durchgeführt. Zu diesem Zweck wird der highD-Datensatz analysiert, der aus Drohnenvideos aufgenommen wurde.

Was ist das akzeptable MaR für automatisiertes Fahren?

Die allgemeine Erwartung ist, dass die Einführung von automatisierten Fahrzeugen die Anzahl der Unfälle zumindest langfristig und pro Fahrleistung verringert. Gleichzeitig ist es offensichtlich, dass die Einführung neue Risiken für die Gesellschaft mit sich bringt, wie bei fast jeder neuen Technologie. Vermutlich werden Erstkunden aufgrund der neuen Erfahrungen und des persönlichen Nutzens bereitwillig Risiken oder Ungewissheiten in Bezug auf Risiken akzeptieren. Andererseits haben Passanten, die keinen persönlichen Vorteil verspüren, vermutliche höhere Anforderungen. In dieser Arbeit werden akzeptable Risiken aus Unfallstatistiken, Risikoakzeptanzstudien und dem Vergleich mit anderen Technologien abgeleitet. Basierend auf den Anforderungen werden Einführungsstrategien unter Ungewissheit diskutiert. Die zugrundeliegenden Annahmen dabei sind, dass Benutzer die Hypothese akzeptieren, dass Fahrzeuge sicher sind, während Passanten und die Gesellschaft die Hypothese eher ablehnen.

Summary

Automated Driving is one of the trends in the automobile industry. Latest developments in technology and prototypes suggest that the introduction of automated driving is near. Despite the advances in the systems themselves, the safety approval is still unsolved. Without further research and improvement in verification and validation methods, a safe introduction of automated driving is not justifiable. Currently, the ECE type approval certifies that the system is safe for road operation and that there is no unacceptable risk involved, based on the assumption that a human driver is able to control the vehicle and decides about trajectory in critical scenes. For higher automation, human surveillance is not available any more. Hence, new methods of safety approval and risk assessment need to be installed to substitute the current type approval.

To begin with, state of the art safety validation methods for automated driving are analyzed and structured to derive open research questions concerning risk metrics. In this thesis, the focus is on scenario-based testing and field-testing. Field-testing means straightforward testing in real traffic. The risk of the system can be estimated in a statistical approach by the occurrence rate of accidents. Scenario-based testing requires the identification of test cases to shift testing from the road to simulation or proving grounds. However, the target of each safety validation is to estimate the risk of the technology ultimately proofing that the safety exceeds the required safety according to all viewpoints in the whole society. In this thesis, two different terms of risk are used. The average risk of a system, e.g. the occurrence rate of fatal accidents, is called *macroscopic risk* (MaR). The risk in a single traffic scene is called *microscopic risk* (MiR). Due to the high distance between two accidents in today's traffic, MaR cannot be estimated without an extensive amount of data. Thus, it requires an enormous mileage to gather enough accident data for significant statistic evaluation. Hence, an important research question is how MiR metrics that evaluate the risk of single scenes without accidents can be used to extrapolate MaR. Another use of MiR metrics is the identification of critical scenes in recorded data or online, during test-drives. These data can be used to derive test cases for the scenario-based testing approach. The three most crucial research questions address the definition of MaR requirements and a top-down approach for defining MiR metrics that are eligible to extrapolated MaR from critical scenes and identify test-cases. They will be further refined in the course of this dissertation. The three questions are:

What are the requirements for MiR metrics and how can their eligibility for the extrapolation towards MaR and the identification of critical scenes be falsified?

As there are many MiR metrics available as state of the art, an assessment process is established in this thesis that evaluates both use-cases. The metric shall identify scenarios that are highly demanding for the driver or the automation. At the same time, the metric shall describe the risk in a scene, so extrapolation from MiR towards MaR is possible. To evaluate if a metric is eligible for those purposes requirements are defined. As it is challenging to verify if a metric fulfills all requirements, a falsification strategy is established instead that contains two steps: First, test scenes are defined that must be assessed correctly by a metric. If falsification by the test scenes is not achieved, it is applied on recorded data of human driven traffic. If the true accident rate corresponds to the extrapolation of risk based on the metric within statistical tolerances, the metric's eligibility is not falsified. As the last falsification step has a high effort, design guidelines are established that lead to a potent metric if followed in the development process.

Which methods and metrics can be used to extrapolate MaR of automated driving from critical scenes in field-testing?

In field-testing, the safety of a driving system is derived based on the occurrence rate of certain events, e.g. fatal accidents. If the event under investigation is a critical scene instead of an accident, the occurrence rate increases and less mileage is required for the same statistical significance. However, the occurrence of critical scenes alone has no information on accident risk. If it is assumed that criticality above a certain threshold is prevented by the driver if possible, and that the occurrence of criticality above this value can be extrapolated from the observation in a field test, the occurrence of scenes of higher criticality and even accidents could be extrapolated. First, it is investigated if state of the art metrics, data collections and extrapolation methods fulfill the derived requirements. As a result, extreme value theory is selected as statistical tool for extrapolation of risk, assuming that highly critical events are extreme events that cannot be approximated by a fit of all occurring criticality values. Finally, a metric that has proven itself so far, is presented. The metric uses model predictive optimization to find the trajectory with the minimum driving requirements in a given scene. The metric fulfills all defined test cases and compares driving requirements with the estimated human driving skill to describe criticality. Following the established design guidelines, a sensitivity analysis is conducted on uncertain parameters to research the influence in parameter choice on the extrapolation result. For this purpose, the highD-dataset that was recorded from drone footage is analyzed.

What is the acceptable MaR for automated driving?

The common expectation is that the introduction of automated driving will reduce the number of accidents at least long-term and per mileage. At the same time, it is obvious that the introduction will induce new risks to the live of modern society, as almost every new technology does. Probably, early adopters will willingly accept risks or uncertainty about risk due to the new experience and the personal benefit. In contrast, passers-by that feel no personal benefit likely have higher requirements. In this thesis, acceptable risks are derived from accident statistics, risk acceptance studies and comparison with other technologies. Based on the requirements, introduction strategies under uncertainty are discussed with the assumption that user are likely to accept the hypothesis that vehicles are safe and fulfill their individual requirements, while passers-by and the society are more likely to reject the hypothesis.

1 Introduction and Scope

Recently, demonstrations from car manufacturers, suppliers and tech companies suggest that automated vehicles are technologically feasible and will be introduced to customer in the near future. The common expectation is that the introduction of automated driving will reduce the number of accidents at least long-term. The *Vision for Safety*[1] program of the National Highway Traffic Safety Administration (NHTSA) of the United States as well as the *Visio Zero*[2] of the European Union rely on the expected safety benefit of automated driving. However, the introduction of new technology is not without risk. At the same time, we know that a statistical proof of superior safety based on accident statistics cannot be performed without introducing automated vehicles to the market due to the high required mileage[3,4]. State of the art development and testing procedures from advanced driver assistance systems are insufficient because they rely on the active supervision of the driver at all time.[5] The problem of a statistic proof of safety is that it is based on accident statistics and as accidents are relatively rare events, it requires a high mileage to come to a statistically significant result. This raises the first question:

How to monitor the current risk level in test drives and after the introduction besides accident statistics?

As statistic on accidents is not feasible before an introduction, the question rises if other metrics could be used to derive the accident risk. If the accident risk in critical scenarios can be described, the total accident risk could be extrapolated. Metrics for the description of proximity to accidents exist, e.g. the well-known Time-to-Collision[6]. However, the connection to accident probability is unknown. Additionally, the definition of a general criterion that covers all types of scenarios is challenging. To conquer this challenge, a top-down approach that starts with the definition of requirements and test cases is the scope of this dissertation.

Besides statistical safety approval in real-world tests, other test concepts are available to assess driving functions. Guidelines for the development of safe systems are given in ISO 26262:2018 for functional safety and ISO/PAS 21448:2019 for the safety of the in-

[1] U.S. Department of Transportation NHTSA: A Vision for Safety 2.0 (2017).

[2] Tingvall, C.; Haworth, N.: Vision Zero (2000).

[3] Winner, H.; Weitzel, A.: Die Freigabefalle des autonomen Fahrens (2011).

[4] Kalra et al.: How Many Miles of Driving Would It Take? (2016).

[5] Donner, E. et al.: RESPONSE 3 - Code of Practice (2007).

[6] Hayward, J. C.: TTC (1972).

tended functionality (SOTIF). At the end of the development, the system validation is demanded. ISO 26262:2018-4 suggests execution of tests in the validation process, including simulation and long-term tests. However, the test case generation is based on expert knowledge, so completeness is unlikely to be achieved. In SOTIF, the term of unknown scenarios is of importance. Due to the open world application, a test can never cover all possible scenes. Some might even be unknown because there is no data or because the scenario does not happen in current traffic, but will in the future. As an approach to reduce the number of unknown scenarios, a data driven test case generation process[7] is suggested to discover unknown unsafe scenarios to add them to a test catalogue.

Numerous studies collected data for human driven traffic in the past and critical scenarios were identified (see Benmimoun[8] for an overview). Metrics to identify critical scenarios are well established. As the amount of information collected in past studies includes only ego vehicle information and sometimes information about one preceding object, new opportunities result when the data are enhanced. Highly sophisticated datasets cannot only be accumulated in test drives with advanced environment perception. Due to the advances in image processing, vehicle trajectories can be recorded from video footage as it was done in the highD-dataset[9].

Which metric can match the detail of modern datasets and identify critical scenarios effectively?

Those datasets offer the opportunity to improve identification-metrics, because the whole traffic scenario can be analyzed and ideally, no relevant object is missing. It will be discussed, which requirements a metric that is designed to analyze large scale data have to fulfill.

Based on the scenarios, SOTIF suggests a process in which the function is analyzed and improved iteratively. SOTIF demands that the process shall be repeated and the function improved until the residual risk is acceptable. So quantitative risk requirements are necessary as well as a method to quantitatively estimate the risk of the system. It is commonly discussed that automated driving functions will improve traffic safety. However, there is also much skepticism in society and individuals. Hence, the following question is addressed in this thesis:

What is the acceptable risk for all groups of our society?

It is out of question that today's traffic is not without risk and likely, this will not diminish completely with the introduction of automated driving. Gasser et al.[10] argue that accidents

---

[7] Wachenfeld, W. et al.: Safety Assurance based on an Objective Identification of Scenarios (2016).

[8] Benmimoun, M.: Automatisierte Klassifikation von Fahrsituationen (2015), pp. 30–36.

[9] Krajewski, R. et al.: The highd dataset (2018).

[10] Gasser, T. M. et al.: Rechtsfolgen zunehmender Fahrzeugautomatisierung (2012).

that are known from today's traffic will be partly prevented due to the lack of human error, but new types of accidents will happen due to new automation risks. Wachenfeld et al.[11] add that it is also unknown, whether the distribution of accidents over different severity classes will change (Figure 1-1). However, it has to be assumed that the future distribution and frequency of accidents are highly relevant for risk acceptance.

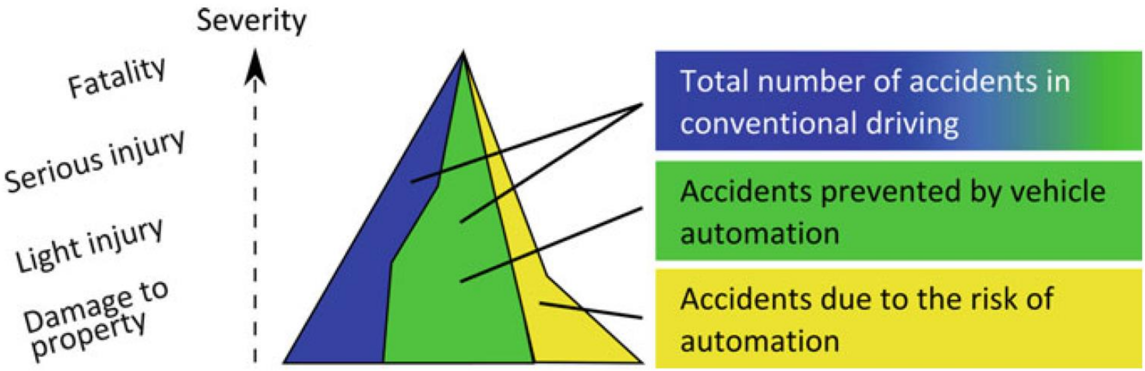

Figure 1-1 (Unknown) future distribution of accidents (Wachenfeld et al.[11] after Gasser et al.[12])

Besides the general expectation for increased safety, there are no defined quantitative acceptable risk levels. For the derivation of the acceptable risk, all viewpoints in society should be addressed because once deployed, autonomous vehicle will likely influence passengers as well as passers-by. An important baseline is the risk in the current traffic that is known from accident statistics. Another consideration is the benefit for passenger or society. If benefits are overwhelming, higher risks are acceptable.[13]

The three questions introduced above are the key questions that include the scope of this thesis. The estimation of risk is always based on an estimation under uncertainty. Hence, the different approaches of safety assessment will be analyzed together with their individual strengths and weaknesses. From the requirements side, the risk requirements are defined and compared with the extrapolation methods and the expected uncertainty level.

1.1 Methodology and Structure

This thesis follows a question driven top-down methodology. Its structure is depicted in Figure 1-2:

- In this chapter, three basic questions that are in scope of this thesis were discussed above. Below, key terms that are used throughout this thesis are defined.

[11] Wachenfeld, W.; Winner, H.: The Release of Autonomous Vehicles (2016), p. 426.

[12] Gasser, T. M. et al.: Rechtsfolgen zunehmender Fahrzeugautomatisierung (2012).

[13] Grunwald, A.: Societal Risk Constellations (2016).

- In chapter 2, the state of the art of validation approaches of automated driving are analyzed with attention to real world data evaluation and risk requirements. As there are limited data from automated test drives available, the research is extended towards data analysis of manual and assisted driving. The metrics that are used in the process of extrapolating risk and identification of critical scenarios are in focus. Additionally, risk levels that are normative or accepted at present in other technologies are discussed.
- In chapter 3, the basic research questions introduced above are further refined and extended based on the findings in the state of the art. Nine refined research questions are deduced.
- In chapter 4, the risk requirements are derived based on an analysis of the requirements from the different stakeholders and the market penetration in relation to conventional traffic. The research questions concerning macroscopic risk are addressed and the consequences for the introduction phase discussed.
- In chapter 5, requirements for metrics that describe risk in traffic scenes are derived. Based on the requirements a falsification strategy for risk metrics is derived and applied on state of the art metrics. As falsification for advanced metric is challenging, design guidelines are established that lead to an eligible metric. Different approaches to define a metric in accordance with the requirements and guidelines are discussed and a suitable metric is suggested.
- In chapter 6, the developed metric is applied using the falsification method. First, test cases derived in chapter 5 are applied. Additionally, data from a German motorway that were collected by drone footage with a total mileage of about 50,000 km are evaluated. It is discussed, whether the results are eligible to extrapolate the accident risk from. Research questions concerning risk extrapolation are addressed here.
- In chapter 7, the results are discussed and summarized, with focus on the research questions, the developed method and its implication on a future safety validation. Open research questions are formulated.
- In chapter 8, a conclusion about the achieved progress is given together with an outlook towards the next steps in safety validation.

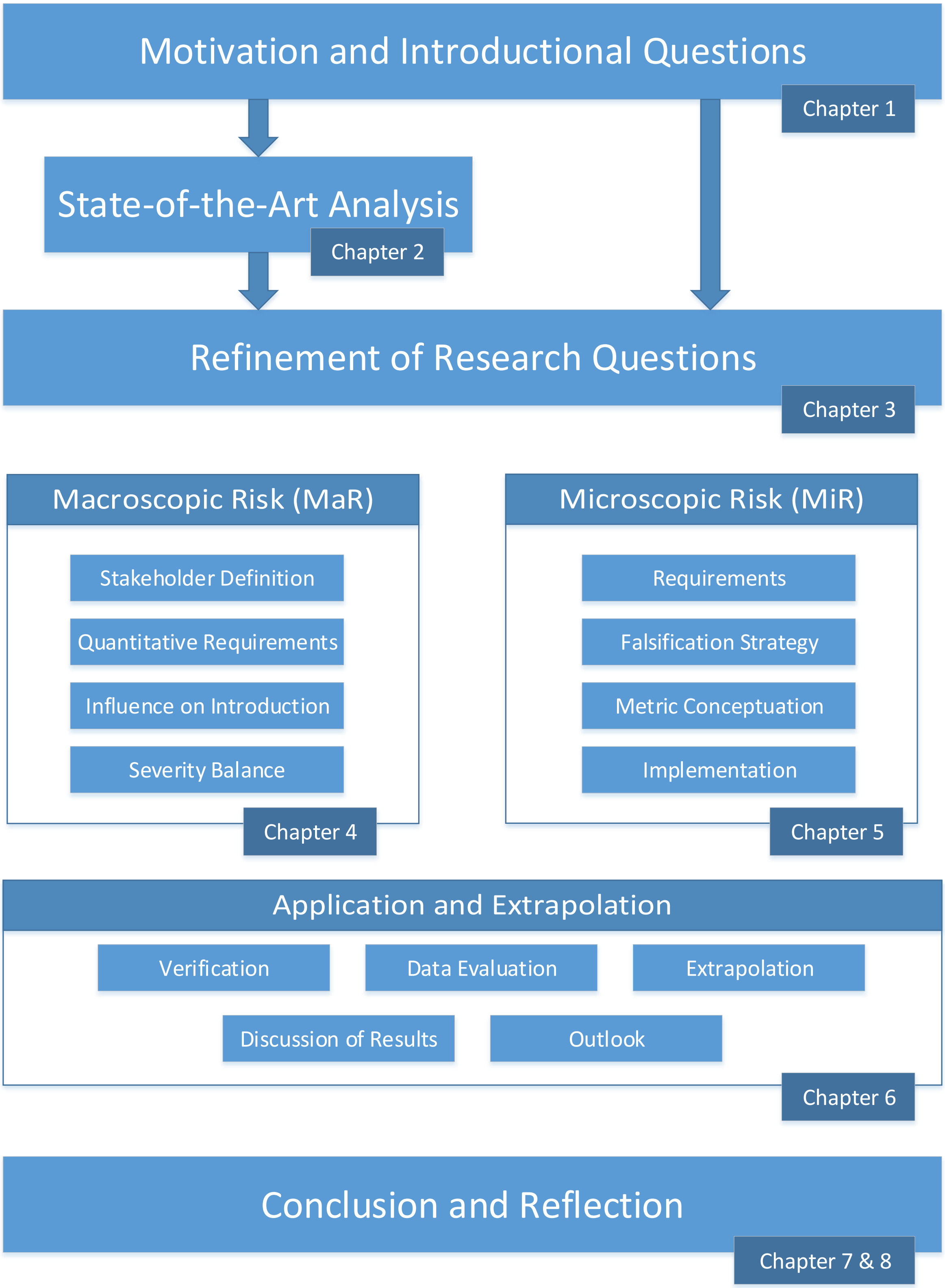

Figure 1-2 Methodology and Structure

1.2 Definition of Terms

In this chapter, key terms used in the following parts are introduced and defined.

1.2.1 Automated Driving

The scope of this thesis is the validation of automated driving. Therefore the taxonomy according to SAE J3016 is used (Figure 1-3). For this thesis, it is sufficient to differentiate between systems with SAE level two and lower (AD2-) and systems with SAE level three and higher (AD3+). Current series-production vehicles provide assistance systems of AD2-. Their validation is established and out of scope of this thesis. A key concept here is that all scenarios must be controllable by the average driver. The driver is always in charge of the driving tasks and monitors the system. Intervention is either necessary when the AD2- systems reaction requires it, as discussed by Weitzel[14], or when a fault of the electric and electronic system (E/E system) is present, which is covered by ISO26262 on functional safety. The safety validation of AD3+ systems on the other hand requires new approaches as the driver does not need to monitor the driving environment.

SAE Level	SAE Name	SAE Narrative Definition	Execution of Steering/ Acceleration/ Deceleration	Monitoring of Driving Environment	Fallback Performance of Dynamic Driving Task	System capability (driving modes)	BASt Level	NTHSA Level
Human Driver monitors the driving environment								
0	No Automation	the full-time performance by *the human driver* of all aspects of the *dynamic driving task*	Human Driver	Human Driver	Human Driver	N/A	Driver only	0
1	Driver Assistance	the *driving mode-specific* execution by a driver assistance system of either steering or acceleration/deceleration	Human Driver and Systems	Human Driver	Human Driver	Some Driving Modes	Assisted	1
2	Partial Automation	Part-time or driving mode-dependent execution by **one or more driver assistance systems** of both steering and acceleration/deceleration. Human driver performs all other aspects of the *dynamic driving task*.	**System**	Human Driver	Human Driver	Some Driving Modes	Partially Automated	2
Automated driving system ("system") monitors the driving environment								
3	Conditional Automation	*driving mode*-specific performance by an *automated driving system* of all aspects of the *dynamic driving task* - *human driver* **does** respond appropriately to a *request to intervene*	System	**System**	Fallback-ready User	Some Driving Modes	Highly Automated	3
4	High Automation	*driving mode*-specific performance by an *automated driving system* of all aspects of the *dynamic driving task* - *human driver* **does not** respond appropriately to a *request to intervene*	System	System	**System**	Some Driving Modes	Fully Automated	3/4
5	Full Automation	**full-time** performance by an *automated driving system* of all aspects of the *dynamic driving task* under all roadway and environmental conditions that can be managed by a *human driver*	System	System	**System**	Some Driving Modes		

Figure 1-3 SAE level of automated driving[15]

[14] Weitzel, D. A.: Kontrollierbarkeit nicht situationsgerechter Reaktionen (2013).

[15] SAE International Standard J3016: Taxonomy for Automated Driving Systems (2014).

1.2.2 Scene, Situation and Scenario

The term scene was defined by Geyer et al.[16] and further extended by Ulbrich et al[17]. In this thesis following the definition by Ulbrich is used:

> *"A scene describes a snapshot of the environment including the scenery and dynamic elements, as well as all actors' and observers' self-representations, and the relationships among those entities. Only a scene representation in a simulated world can be all-encompassing (objective scene, ground truth). In the real world it is incomplete, incorrect, uncertain, and from one or several observers' points of view (subjective scene)."* [17]

The definition is further explained in Figure 1-4.

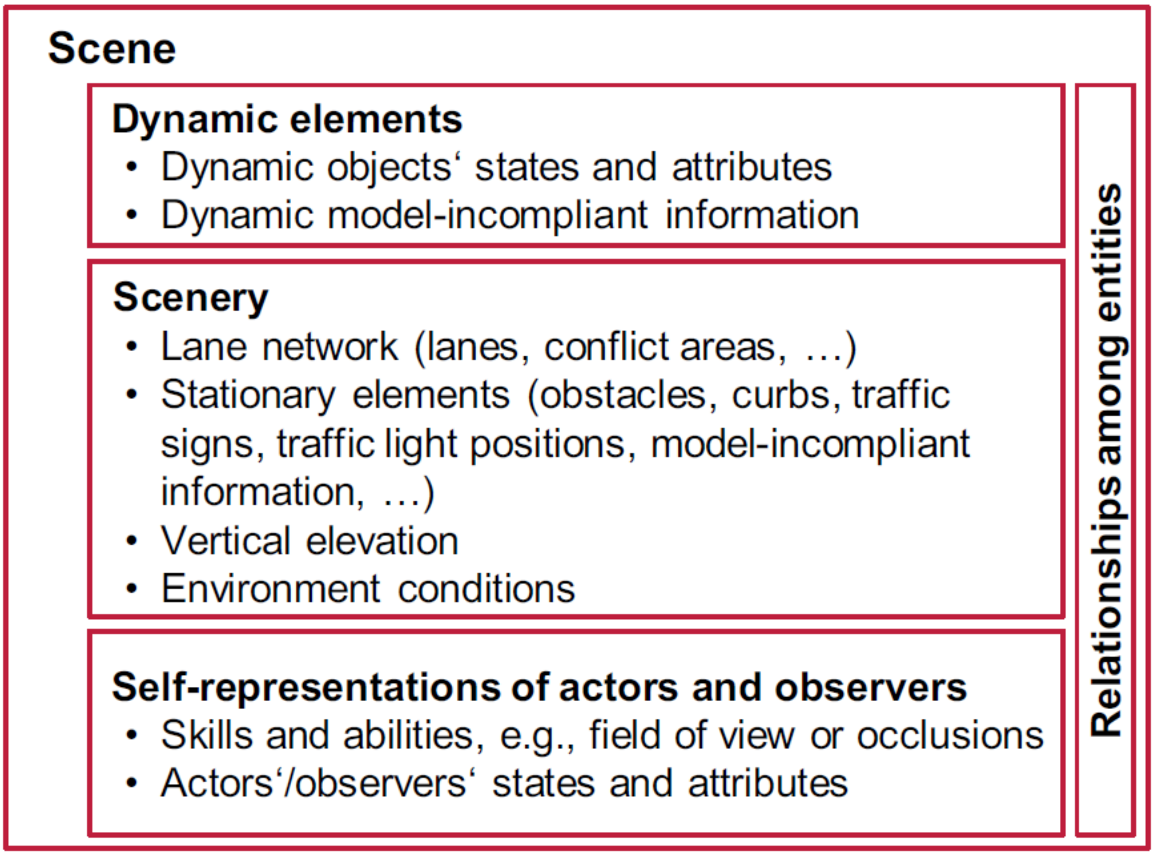

Figure 1-4 Definition of a scene[17]

In literature, the term situation is often used as a synonym for the term scene. However, according to Ulbrich, a situation also includes goals and values of the function in addition to

[16] Geyer, S. et al.: a unified ontology for generating test and use-case catalogues (2013).

[17] Ulbrich, S. et al.: Defining and Substantiating the Terms Scene, Situation, and Scenario, p. 983.

all aspects of a scene that are relevant for those goals and values. The difference between scene and situation is of minor importance in this thesis. Hence, the term scene will be used.

While a scene is only a snapshot of the vehicles' condition and the surroundings, a scenarios spans over a timespan. It either can be defined by several scenes that follow each other with changing states of dynamic objects, or can have an initial scene with definitions of the following maneuvers that evolve form the starting point. In this thesis, the term scenario always addresses a scenario with data points at each time step, so the first specification will be used.

1.2.3 Safety Performance and Risk

To describe the safety level of a system or of the current traffic. The term Safety Performance (SP) according to Wachenfeld[18] is used. It is described as the average distance until an accident occurs. As it is necessary to differentiate between different severity categories when discussing risk and not only accident likelihood, the SP can be given for each severity category (e.g. fatal accidents: index f; accidents with injuries: index wI; accidents without injuries: index nI (no injuries)).

According to ISO 26262 risk is the severity of an accident multiplied with the frequency or likelihood of the accident. The frequency of accidents is either given per year (symbol f) or per distance (symbol $\mathcal{F}$).

So $\mathcal{F}$ is the reciprocal of SP when the severity level is defined. In ISO26262, different severity levels are combined by assuming a factor of ten between severity categories, meaning that an accident with injuries is handled as ten times more severe than an accident without injuries and an accident with fatalities is handled as 100 times more severe. The combination of different severity level into a single value of risk or SP is further discussed in subchapter 4.3.

In this thesis, the risk in a single traffic scene and the average risk of a fleet of vehicles are relevant. SP and risk according to ISO26262 address the average risk of a vehicle or a fleet of identical vehicles. This is called Macroscopic Risk (MaR). The risk in a single scene is called Microscopic Risk (MiR). It is described by MiR metrics when applied on data.

1.2.4 Criticality and Criticality Metric

The term criticality metric is refered to a broad family of metrics in literature (see section 2.2.2.2.1). In this thesis, the term criticality is defined as the temporal or spatial closeness to a potential collision in a driving scene or scenario, or the magnitude of a driving dynamic reaction required to prevent an accident. A criticality metric describes the criticality or aspects of criticality in a scene or scenario. The reaction can either be assessed a posteriori by

[18] Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017), p. 25.

measuring the vehicle state or a priori by environment perception and prediction of the required reaction. A criticality metric does not necessarily deliver information about risk or accident probability. However, a risk metric is also a criticality metric but not vice versa. Often the purpose of the metric is the classification into critical and uncritical scenes. This requires definition of a threshold above or underneath of which the scene would be critical. Examples for a classification in major Naturalistic Driving Studies (NDS) and Field Operational Tests (FOT) are given in Appendix A.

All metrics handled in this dissertation describe risk or criticality on action level of driving. Scenes were the action itself is easy but the decision making or the information reception is difficult are not considered. An example would be driving in a sandstorm were the action level is not directly influenced because driving straight might be still possible. Amersbach[19] [20a] defines a total of six layers to decompose the driving task (Figure 1-5) and criticality could be described on all levels. Metrics on actions level are necessary and sufficient to describe risk, as insufficient behavior on lower levels will result in risky action later on: A late detection of a front object could be detected on information processing level, but results in an emergency braking on action level.

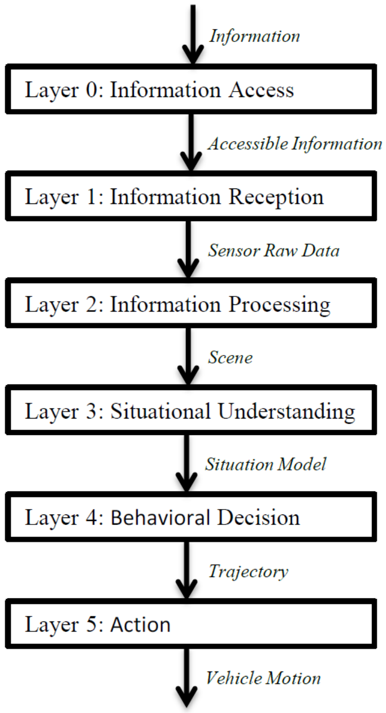

Figure 1-5 Decomposition levels[20b]

[19] Amersbach, C.; Winner, H.: Functional Decomposition (2017).

[20] Amersbach, C.; Winner, H.: Funktionale Dekomposition (2018).a: -; b: p.4

1.2.5 Coordinate Systems and Notation

In this thesis, two different two-dimensional Cartesian coordinate systems are used: a world coordinate system that is used to describe positions and directions on a road plane and a (horizontized) ego-vehicle coordinate system that describes positions of objects in relation to the vehicle as well as the accelerations and velocities of ego-vehicle and objects. The position of an objects always refers to the geometric center of the vehicle, length, width and yaw angle must be known in order to describe a bounding box around the object. Whenever it is necessary to differentiate between the two coordinate systems, the letter *e* for ego-vehicle centered system or *w* for world centered system is indicated at the lower left of the variable as depicted in Figure 1-6 (e.g. ${}_{w}x$ for an x-position in world coordinates). As the slip angle is assumed to be zero, yaw angle and course angle are identical and the velocity ${}_{e}v_{y}$ of the ego-vehicle is zero.

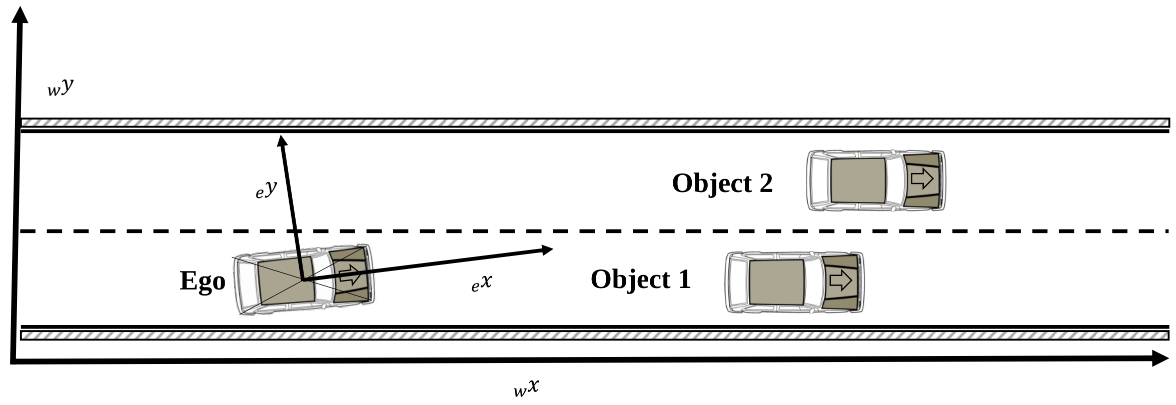

Figure 1-6 Coordinate systems

In addition to the two Cartesian coordinate systems, a curvilinear coordinate system with the index *c* is introduced in section 5.3.2.8. The road curvature is used for the longitudinal axis in order to efficiently handle driving that follows the lane on motorways with small curvature.

2 State of the art Requirements and Approaches for Safety Validation

This chapter analyzes the current state of safety validation. First, the current expectations and requirements towards AD3+ are introduced and comparisons to other technologies are drawn. Second, the state of the art of safety validation for AD3+ is summarized and structured with focus on risk metrics.

2.1 How Safe is Safe Enough?

In this subchapter, the basis for the definition of safety requirements is discussed. First, publication for the derivation of safety requirements for AD3+ are briefly discussed. Next, the fundamental basis for safety requirements that is used in norms is summarized, before a comparison with the requirements in aviation and pharmacy is drawn.

2.1.1 Existing Considerations for AD3+

It is broadly assumed that AD3+ shall increase the safety on public roads and reduce the number of traffic victims compared to human driven traffic. The Ethics Commission Automated and Connected Driving of the German Federal Ministry of Transport stated that "*the licensing of automated systems is not justifiable unless it promises to produce at least a diminution in harm compared with human driving, in other words a positive balance of risks*"[21]. Besides this statement, other approaches exist:

Wachenfeld introduces a limited introduction[22] (that is also called risk-limited introduction[23]), where AD3+ is introduced without a sufficient proof (but still with the expectation) of superior safety. The driven distance and the accidents from the sold series vehicles are monitored, gaining additional knowledge about the performance, and ideally allowing more vehicles to be deployed, because the uncertainty of the safety performance assessment decreases.

---

[21] Federal Ministry of Transport and Digital Infrastructure: Report of the Ethics Commission (2017), p. 4.

[22] Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017), pp. 102ff.

[23] Winner, H.; Wachenfeld, W.: Risk-limited Introduction of Automated Driving, (2017).

Kalra et al.[24] even suggest that with the deployment of AD3+, we can accelerate the development of the systems due to the gained knowledge and the recorded data that can be used in machine learning. Hence, AD3+ could be introduced even with inferior safety, assuming a certain improvement rate per mileage. In the end, the system would outperform human driver due to this learning effect. If the improvement rate is high enough, lives are saved at the cost of a higher risk directly after the introduction. A balance of risk that is positive in total, compared to a delayed and careful introduction, would permit this increased initial risk.

Though, the approaches of Kalra and Wachenfeld do not demand a proof of safety before the introduction. Nevertheless, they do not contradict the abovementioned statement of the ethics commission, because the system ultimately promises increased safety.

Another question is how to evaluate a shift in different severity categories. It might be possible that the accident rate of minor accidents would increase dramatically, but the accident rate of more severe accidents would decrease. Hydén suggests that a reduction of minor accidents typically results in a reduction of more severe accidents. This is also called thy Hydén triangle.[25] However, it is possible that this does not hold true for AD3+.[26] Possible reasons include minor accidents due to unexpected slow-speed maneuvers in mixed traffic or high-velocity crashes without braking due to false-negative obstacle detection. The first would increase the minor accident rate with no influence on severe accidents. The latter would lead to the opposite, if the false-negative detection only happens at higher velocities. To cope with the different severity levels, Wachenfeld suggest a monetary approach motivated from insurance practices to combine the different accident rates.[27] Its application and consequences is further discussed in section 4.3.2.

2.1.2 Fundamental Safety Requirements[28]

This section discusses normative approaches to describe risk requirements. Defining risk requirements is challenging in nature as the assessment and comparison of risks is challenging for humans. It is illustrated from a small excursion:

The recent publication of the German Federal Statistical Office[29] again reports an increase in average life expectancy of newborns. Since the recording of mortality tables began in

---

[24] Kalra, N.; Groves, D. G.: The Enemy of Good (2017).

[25] Hydén, C.: Method for traffic safety evaluation (1987).

[26] Wachenfeld, W.; Winner, H.: The Release of Autonomous Vehicles (2016), p. 426.

[27] Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017), pp. 26ff.

[28] This section and its sub-sections is taken from Junietz, P. et al.: Macroscopic Risk Requirements (2019) and only modified slightly.

[29] Destatis: Kohortensterbertafeln für Deutschland (2017).

1871, we observed a doubling of the average life expectancy. One should assume that people are very lucky with this development. But a look into daily media reporting shows that people are not only very skeptical about technical achievements but they are even afraid of effects that are obviously responsible for the aforementioned increase in life expectancy - for example medical and agricultural advances. People are concerned, for example, about man-made radiation and air pollution by industrial plants or transportation but also about the side-effects of medicine, the consumption of meat of uncontrolled origin, bacteria in green salad, dioxin in free-range eggs, etc.

Nevertheless, the facts speak for themselves: The most common natural causes of death in Germany - and this is representative for industrial countries - are cardiovascular diseases with 39%, followed by cancer with nearly 25% and, well behind, by diseases of respiratory and digestive system with 7% and 4%. It is interesting to note that non-natural causes of death, i.e. mainly suicides and accidents, contribute only 4%[30 31]. From the medical point of view there is no doubt regarding the factors that really kill us - smoking, overweight, high blood pressure, diabetes, and physical inactivity. Everybody is able to control those factors and to prevent the consequences but why are we not doing this consistently? Moreover, why are we so concerned regarding other factors that are less risky but not readily controllable by ourselves? Why do our risk perception and risk acceptance seem contradictory? (comp. Fritzsche[32a])

There is obviously a discrepancy between objectively existing risks on the one hand and their perception and acceptance by individuals as well as by society on the other. It is important for that purpose, among other things, whether people enter into the risk voluntarily or not, whether they feel a personal benefit, and whether the risk is natural or synthetic. Moreover, risk perception depends on risk communication[32b,33,34].

We have to ask ourselves, whether it is even possible to deal with risk in an objective manner. What the consequences of the described difficulties with risk perception and acceptance are for the introduction of new, complex technologies like for example highly automated driving.

2.1.2.1 Quantitative Risk Assessment

The usual quantitative risk definition “risk equals frequency times severity” is illustrated in Figure 2-1 (left). Due to the double logarithmic scale, constant risk is represented by a line.

---

[30] Destatis: Pressemitteilung Nr. 022 (2017).

[31] Destatis: Gesundheit - Todesursachen in Deutschland (2017).

[32] Fritzsche, A. F.: Wie sicher leben wir? (1986).a: pp.1-3; b: -

[33] Grunwald, A.: Societal Risk Constellations (2016).

[34] Slovic, P.: The perception of risk (2011).

Considering the occurrence of unintended events (frequency) and extent of damage (severity), ideally two areas are defined. In the green area, the system is in a safe state; the corresponding risk is accepted. In the red area, the system is in an unsafe state; the corresponding risk is not accepted. The borderline between these two areas is probably not sharp; there can be a kind of transition area.

Although this simple definition is very useful for many questions in technology and insurance industry, it neglects aspects like aversion against high severity, lack of controllability, and personal benefit, which are relevant for risk perception and acceptance by individuals and society. Intensive research on risk perception and risk acceptance started in the second half of the last century. Different authors analyzed risk acceptance and risk-benefit constellations in various studies (Douglas[35], Crouch[36], Gibson[37], Kinchin[38], Kletz[39]. Starr[40,41] and Webb[42]). Fritzsche discussed risk acceptance relating to voluntary nature of exposure based on the above mentioned studies[43]. Slovic concludes similar results in a more recent, updated publication.[44] Their results are summarized in Figure 2-2. It is interesting that both authors conclude similar risk numbers despite the major gap of several decades. The reason might be that risk perception studies reached their peak in the 70's with the introduction of nuclear power. Junietz et al.[45] suggest correcting the numbers with a factor derived from the change in mortality rate.

For voluntary activities, Fritzsche found that the willingness to accept risks is nearly unlimited, depending on the experienced personal benefit. We can see this by the example of high-risk sport or other leisure activities, e.g. free climbing, motorcycling etc.. Job-related activities are important for a deeper understanding of the subject. Acceptance is relatively well investigated in this field and there is a common understanding of accepted individual mortality risk in the order of 10^{-5} per person and year, for example by professional associations and insurance companies, on the one hand. On the other hand, job-related risks are useful to bridge the gap between voluntary and involuntary risks.

---

[35] Douglas, M.; Wildavsky, A.: How can we know the risks we face? (1982).

[36] Crouch, E. A.; Wilson, R.: Risk/benefit analysis (1982).

[37] Gibson, S. B.: Risk criteria in hazard analysis (1976).

[38] Kinchin, G. H.: Design Criteria, Concepts and Features Important to Safety and Licensing. ANS (1979).

[39] Kletz, T. A.: Hazard analysis, its application to risks to the public at large (1978).

[40] Starr, C.: Social benefit versus technological risk (1969).

[41] Starr, C.: Benefit-cost relationships in socio-technical systems (1971).

[42] Webb, G. A.; McLean, A. S.: Insignificant levels of dose (1977).

[43] Fritzsche, A. F.: Wie sicher leben wir? (1986).

[44] Slovic, P.: The perception of risk (2011).

[45] Junietz, P. et al.: Macroscopic Safety Requirements for Highly Automated Driving (2019), 7.

Fritzsche found that for involuntary risks, e.g. death of passengers due to a train or airplane crash, the acceptance level is an order of magnitude lower than for job-related risks. Moreover, acceptance decreases another order of magnitude, if the risk is caused by major technology, e.g. chemical industry or nuclear power generation. Beside the fact that the experienced personal benefit of those technologies is low (at least from a subjective point of view), the low degree of self-determination or rather controllability by individuals plays an important role for the low acceptance level as well as the potentially high number of mortalities (severity). Nevertheless, the studies, summarized in Figure 2-2, show that it is generally possible to deal with risk, risk perception, and acceptance in a quantitative manner.

To implement safety requirements based on risk acceptance, several concepts have been developed in different application areas. Because of the relationship between railway and road traffic, it is useful to refer to the CENELEC safety standard EN 50126. The development of this standard has been started in the 1990's, where safety requirements based on quantitative risk analysis have been implemented and ALARP, MEM, and GAMAB have been introduced as principles for risk acceptance. Those principles are briefly explained in the following sections.

2.1.2.2 As low as reasonably practicable (ALARP)

ALARP tries to assess what is technically feasible considering economic sense and social acceptance. Between the two regions of generally unaccepted and broadly accepted risk, there is a tolerance range where risk is undertaken only if a benefit is desired and where each risk must be made as low as reasonably practicable.

Derivation of risk figures is not directly applicable because EN 50126 failed to give certain values for generally unaccepted and broadly accepted risk. However, other authors, for example Risk & Reliability Associates, deliver both values[46]: The two key levels seem to be located around road death statistics (about 10^{-4} per person and year) and the chances of being struck by lightning (about 10^{-7} per person and year). If something is more dangerous than driving a car, the risk is unacceptable. If something is less dangerous than being struck by lightning, it is not required to be reduced further. In the range between these two figures, cost benefit studies are appropriate to reduce the risk to as low as reasonably practicable. Especially this lower ALARP limit corresponds very well with the acceptance criterion for major technology risks shown in Figure 2-2.

[46] Risk & Reliability Associates Pty Ltd, Consulting Engineers: Risk and Reliability (2004).

2.1.2.3 Minimum Endogenous Mortality (MEM)

MEM is based upon age- and gender-specific mortality rates[47]. Although the absolute values of the mortality rates change with birth cohort, they show a typical development over age as well as a significant minimum at an age of about 10 years, which is about $5 \cdot 10^{-5}$ per person and year. The related mortality at an age of 10 years is defined as “minimum endogenous mortality”[48]. The MEM principle demands that a new system does not significantly contribute to the existing minimum endogenous mortality. EN 50126 specifies that the individual risk due to a certain technical system must not exceed $1/20^{th}$ of the minimum endogenous mortality, considering that people are normally exposed to the risk of several technical systems. This means that the acceptable individual risk of a certain technical system is $2.5 \cdot 10^{-6}$ per person and year, when using latest mortality rates as a basis (EN 50126 uses mortality rates from the 80’s). This value corresponds very well with the acceptance criterion for involuntary risks shown in Figure 2-2.

2.1.2.4 Globalement au moins aussi bon (GAMAB)[49]

GAMAB, (or GAME globalement au moins équivalent), requires, unlike MEM, the existence of a reference system with – currently – accepted residual risks. According to GAMAB, residual risks caused by a new system must not exceed those of the reference system. In other words: a new system must offer a level of risk generally at least as good as the one offered by any equivalent existing system. Hence, it is necessary to identify the risk of an equivalent existing system.

Looking for the acceptable risk of highly automated driving in a certain application area according to GAMAB, the current risk of an equivalent existing system in the same application area needs to be identified. To derive acceptance requirements for a controlled-access highway pilot, the current risk on German Autobahn during manual driving is analyzed. Table 2-1 shows average distances between two accidents referring to severity levels according to ISO 26262. However, for other AD3+ systems, there might be no reference system because the use-case is not covered by today’s vehicle fleet.

---

[47] Destatis: Kohortensterbertafeln für Deutschland (2017).

[48] Although the mortality statistics does not differentiate between endogenous and exogenous causes for the death.

[49] English: generally at least as good as

Table 2-1 Accidents on German Autobahn[50]

Severity	ISO 26262 Severity level	Average distance between two accidents of this level	Accident rate per driven distance
Fatal	S3	$660 \cdot 10^6$ km	$1.52 \cdot 10^{-9}$/km
Severe Injuries	S2	$53.2 \cdot 10^6$ km	$1.88 \cdot 10^{-8}$/km
Injuries	S1	$12.5 \cdot 10^6$ km	$8.00 \cdot 10^{-8}$/km
w/o Injuries	S0	$7.5 \cdot 10^6$ km	$1.33 \cdot 10^{-7}$/km

Figure 2-1 (right) shows the observed accident rates on Autobahn versus severity levels according to ISO 26262. Comparing risks of the different severity categories requires weighting of the different levels. However, there is no standardized way (see also Baum[51], Hydén[52] and Wachenfeld[53]). The lines of constant risk in Figure 2-1 assume that the difference between adjacent severity levels is one order of magnitude, or in other words that an accident with fatalities is assessed with a ten times higher severity than an accident with severe injuries. However, that does not mean that a fatal accident is definitely ten times worst. The factor of ten fits to the current accident frequency, which could change in the future e.g. due to enhanced passive safety. However, this assumption allows defining a band of constant risk in current traffic, which can be used as reference. As already discussed in section 2.1.2.2, the risk will not be accepted above the upper envelope. Beyond the lower envelope, the risk might be accepted. Between both lines is a transition area. In accordance with the ALARP principle, this is a tolerance zone where risk is undertaken if a benefit is desired and where each risk must be made as low as reasonably practicable.

In Figure 2-2, the results of the application of the different risk acceptance principles are displayed related to the risk acceptance limits of the different expositions explained previously. The different approaches for the mortality risk are summarized. On the one hand, it shows that the application of different risk acceptance principles delivers comparable and consistent results. On the other hand, it demonstrates that we have to deal with a relatively broad range of applicable acceptance criteria.

---

[50] Steininger, U. et al.: Validation of Assisted and Automated Driving (2016).

[51] Baum, H. et al.: Volkswirtschaftliche Kosten durch Straßenverkehrsunfälle in Deutschland (2010).

[52] Hydén, C.: Method for traffic safety evaluation (1987).

[53] Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017), pp. 26ff.

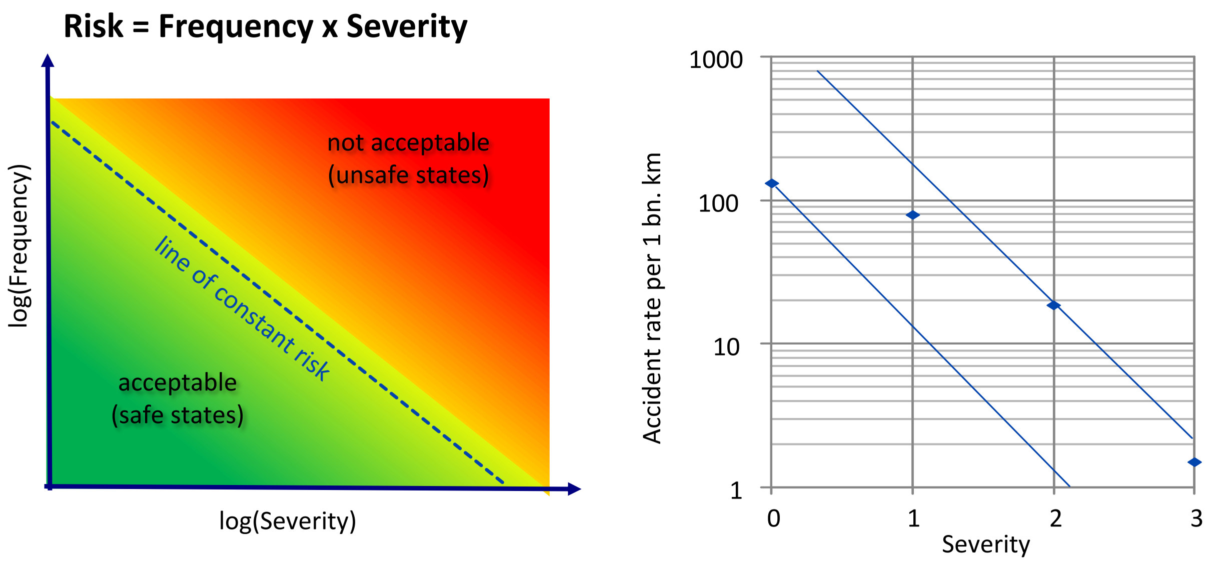

Figure 2-1 Left: Illustration of risk; Right: Quantitative accident risk on German highways[54,55]

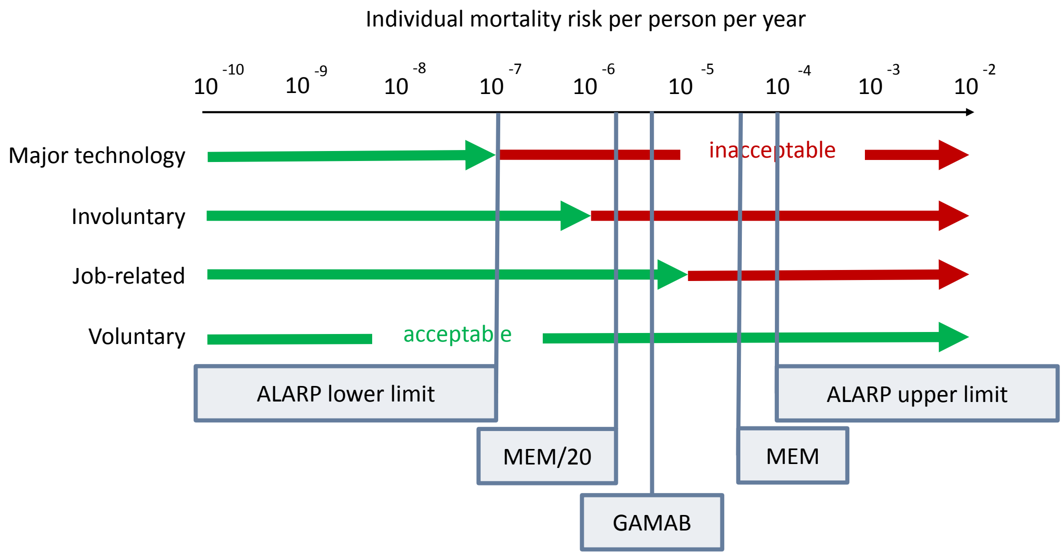

Figure 2-2 Application of different risk acceptance principles to highway accidents, translated from Steininger and Wech[56], based on Fritzsche[57], GAMAB is based on the risk on German controlled-access highways, MEM is based on the current mortality tables.

Considering the impact of voluntary exposure, different groups of users have to be distinguished, for example users of highly automated driving systems and other traffic participants. Finally, a comparison with other technologies – especially other traffic systems and

[54] Steininger, U. et al.: Validation of Assisted and Automated Driving (2016).

[55] Schöner, H.-P.: Challenges and Approaches for Testing of Highly Automated Vehicles (2014).

[56] Steininger, U.; Wech, L.: Wie sicher ist sicher genug? Sicherheit und Risiko zwischen Wunsch und Wirklichkeit (2013).

[57] Fritzsche, A. F.: Wie sicher leben wir? (1986).

technologies that deliver a high personal benefit – seems to be useful. Additionally, a decrease in total mortality risk is expected in the future, following the trend in the last decades and centuries. Therefore, risk acceptance might change over time.

2.1.3 Introduction Strategies and Risk-handling in other Domains

Automated driving challenging is not the only technology with a challenging safety validation. In the following risk-handling in pharmacy and aviation is briefly discussed. The example of aviation is obvious because it is another traffic system, where more and more automation was introduced in the past. Pharmaceutics are especially interesting because weighting risks and benefits is crucial because every (new) medicine has side-effects.

2.1.3.1 Pharmaceutical Risk/Benefit-Analyses

When introducing new medicines to the market a pharmaceutical risk/benefit-analyses is mandatory and established. It is typically done by clinical studies in comparison to no treatment (absolute risk/benefit) and treatment with alternative drugs (relative risk/benefit).[58] Additionally, the supervision of the product in the market is mandatory. When the estimation of the risk/benefit changes, the product needs to be re-evaluated. In analogy to the introduction of automated driving, this would mean a supervision of accident statistics. If the system does not meet the requirements, a mandatory update or even a prohibition would follow. This would happen for example, if side effects become known that were not discovered in clinical studies before market introduction.

Another aspect is product liability. For pharmaceutical products, the manufacturer is not liable when damages occur that are considered acceptable during the risk/benefit analysis. Even more, if the threats were not considered during the analysis but would not have changed the results (e.g. because the occurrence rate is low or the benefits are overwhelming), there is no liability.[59a]

There is no mandatory quantitative measure for conducting the risk/benefit analysis because weighting the different factors is challenging and varies from case to case. Instead the analysis shall be assessed using scientific findings according to the legislation in Germany (§84(2) No. 1, AMG[59b]). It is also practice weighing fatal risks with very low occurrence rate but benefits in curing a non-fatal diseases.[60,59c] However, the introduction of AD3+ will even influence passers-by that have no direct benefit. In pharmacy, it concerns usually the same

---

[58] Hart, D.: Die Nutzen/Risiko-Abwägung im Arzneimittelrecht (2005).

[59] Besch, V.: Produkthaftung für fehlerhafte Arzneimittel (2000), p. 54.a: p. 54; b: p. 58; c: pp. 61-62

[60] Dieppe, P. et al.: Balancing benefits and harms (2004).

individual. Therefore, it is questionable if the transfer of risk to others due to societal benefits is reasonable.

If the same logic were applied to AD3+, that would mean that a risk/benefit analysis shall be conducted comparing additional risks e.g. due to systematic failures to the benefit of reducing accidents. Similar to pharmaceutics, it is impossible to guarantee that all risks are known before the introduction. Hence, a field surveillance should be mandatory, re-evaluating the risk/benefit analysis at a regular basis. The manufacturer would not be liable, even if new risks occur that do not change the overall assessment.

2.1.3.2 New Technologies in Aviation[61]

In aviation, passengers are exposed to a technical system without having personal control. Although severe accidents happen, its safety is accepted by most of the population. Aviation has become increasingly automated in the past, although today's systems are still comparable to SAE level 2 because they are supervised by the crew. However, it is not directly comparable because the safe state is very difficult to obtain. This increased difficulty is compensated by reaction time that is usually longer than in a vehicle. Due to the long travelling distance and the fact that accidents mostly happen during take-off and landing, accident rates are typically given per flight and not per travel distance. Accidents and critical scenes are strictly reported and collected in a database, so we have even more profound data compared to road traffic. To compare annual risks for the average person, the average driving distance and the number of flights per year is used. One fatal accident happens about once per ten million flights[62]. With a typical exposure of two flights per year, the risk of a fatal accident would be lower than the risk of involuntary exposure f_{inv} and about one order of magnitude lower than driving on a highway. However, with 20 flights per year, one would be exposed to a risk that is in the same order of magnitude. Therefore, the levels of risk are in fact comparable, if only driving on Autobahn is considered. However, users typically drive on all types of roads. The risk of car accident is at least one order of magnitude higher in total, so the superior reputation of air traffic is justified. Additionally, accidents of lesser severity are much more frequent in vehicles traffic compared to aviation.

As mentioned before, aviation has become increasingly automated over the past decades. The detailed collection of data in aviation allows an analysis per generation of airplanes, which was summarized by Airbus Industries[62]. As depicted in Figure 2-3, with every introduction of a new technology generation, the fatal accident rate for this new generation was higher than state of the art. Due to the low number of new airplanes at introduction, this trend

---

[61] This section is taken from Junietz, P. et al.: Macroscopic Risk Requirements (2019)

[62] Airbus: Commercial Aviation Accidents 1958-2016 (2017).

cannot be observed in the total accident rate. Nevertheless, the introduction was clearly beneficial to society in total because after an introduction phase of five to ten years, the new generation achieved the lowest accident rate of all.

Judging from this data, new generations of airplanes are not tested in a way to prove statistically that the system is superior to the former. In fact, this is impossible; because the knowledge about the new system's behavior is incomplete and only field experience can reduce the unknowns. Like AD3+, statistical testing is neither economically feasible nor necessary because the strict supervision of air traffic allows efficient improvement in case of critical scenes or accidents. However, the highest automation in commercial air traffic is still comparable to SAE level 2, so human error is still a factor. Nevertheless, the leap in accident rate occurred with the introduction of technology, either because of flaws in human-machine-interaction or in the technology itself.

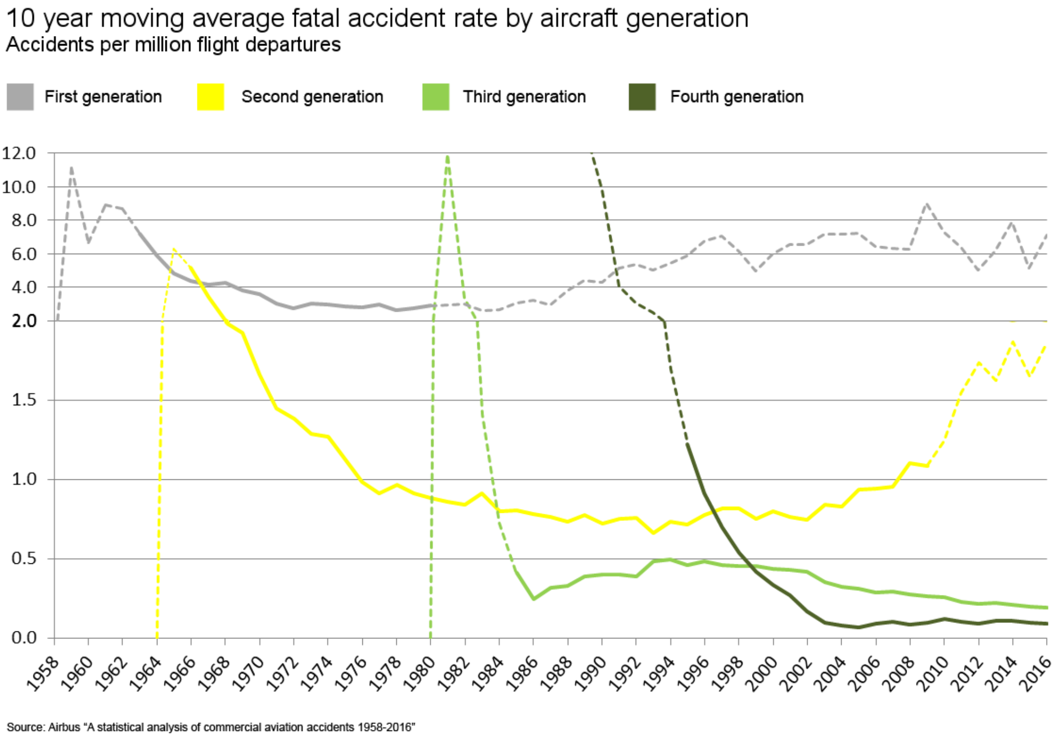

Figure 2-3 Fatal accidents with different generations of airplanes in commercial traffic. Dotted line means less than one million flights a year. First generation: Early commercial jets, Second generation: More integrated Auto Flight System, Third generation: Glass cockpit and Flight-Management-System, Fourth generation: Fly-By-Wire with flight envelope protection.[63]

[63] Airbus: Commercial Aviation Accidents 1958-2016 (2017), p. 17.

2.1.4 Summary for Existing Safety Requirements[64]

Most approaches try to define safety requirements focused on the requirements for a whole society while requirements of individuals are often neglected. However, there are studies that establish risk acceptance criteria dependent on the type of exposure or the individual benefits.

Comparing the risks in other domains shows that directly after introduction the risk is higher than expected. Additionally, in pharmaceutics, a market observation and a re-evaluation of the risk is mandatory. In aviation, accidents and critical situations are monitored. One could argue that it is unethical to release a system that is not tested in the best way possible. However, one could also argue the opposite. First, it is impossible to completely test a system operating in an uncontrolled environment because there might still be scenes that the tester was not aware of. These "unknown unknowns" cannot be tested. Second, a stricter approval process would prevent technical progress because a profit-oriented development would become impossible. It seems possible if not likely that the accident rate of automated vehicles will behave in a similar to aviation. We should be aware of that possibility and focus on the improvement of the system in case of a detected critical scene or accident. A similar thought is also expressed by Kalra and Groves[65]. The delayed introduction of AD3+ could in fact risk the lives of many people because the system is believed to improve safety over time.

With a combined testing strategy of simulation, proving ground tests, and real traffic tests, it is still unlikely to complete a logical proof of safety because every validation test has certain underlying assumptions. In order to deal with this uncertain safety performance, accidents, unexpected critical scenes, and near misses must be monitored similar to air traffic, in order to find flaws in the system (including infrastructure and human interaction) with the chance to improve them.

Research questions that can be derived for the definition of macroscopic safety requirements are discussed in subchapter 3.1.

2.2 Safety Validation Approaches

This subchapter discusses existing approaches for test and safety validation. The focus is on the dynamic product test that is defined as a test where the object-under-test (OuT) is executable and able to produce an assessable result[66].

---

[64] Parts of this subchapter are taken from Junietz et al.: Macroscopic Risk Requirements (2019)

[65] Kalra, N.; Groves, D. G.: The Enemy of Good (2017).

[66] Nörenberg, R.: Effizienter Regressionstest von E/E-Systemen (2012).

2.2.1 Categorization of Dynamic Testing of AD3+

In order to compare the different approaches, the assessment methodology introduced by Junietz et al.[67] is used. It is explained briefly in this section. As this is a validation test, the estimation of the safety performance SP is the overall target instead of the improvement of the function as in testing during development. The test is separated into three dimensions with seven sub-dimensions in total (Table 2-2) that are explained in the following. Each sub-dimension has a best-case value for a full test approach. Every deviation from the full test approach should be justified and the consequences on the test outcome discussed.

Object under Test (A): What is tested?

The OuT is divided into two sub-dimensions. First the OuT itself (A.1), tests can be executed with the whole vehicle or on system, sub-system, or component level. All tests might be equally valuable in order to estimate the SP, but it should be discussed and reasoned, whenever the test is not on vehicle level. Hence, the scale level is nominal. The OuT can be tested on different abstraction levels (A.2). This can be the real OuT or a simulation. In general, the test in the real OuT is most valid, while validity of deviation and the consequences for the trust in the result must be discussed.

Stimulus (B): How is tested?

The Stimulus is divided into three sub-dimensions containing all test elements that are not part of the OuT. For the test case selection itself, the full test approach would be the full coverage of all possible scenarios without any effort in reducing the mileage or the number of test scenarios. A reduction could be achieved by using knowledge about the OuT (B.1) or the surroundings (B.2). However, even without this reduction, the test might not cover all possible driving conditions due to a lack of information about the current or future traffic scenes. If the test scope is reduced e.g. due to a limited use-case, this reduction should be justified. In many cases, tests are executed in an artificial environment, a simulation or on a proving ground. This abstraction of the stimulus is the third sub-dimension (B.3).

Assessment Criterion (C): How is assessed?

The assessment is divided into direct measurement of the accident risk or the abstraction using an estimated accident probability and severity e.g. as in ISO 26262. This is often done because of insufficient data to derive an exact accident risk. In addition, the scale level can be ordinal instead of a ratio scale because the exact value is difficult to estimate.

[67] Junietz, P. et al.: Evaluation of Safety Validation Approaches (2018).

Table 2-2 Categories for the Assessment of Safety Validation Methods[68]

Test dimension	Sub-dimension	Scale level	Full test approach
OuT representative (A)	OuT (A.1)	nominal	Whole vehicle
	OuT abstraction level (A.2)	ordinal[a]	Real entity
Stimulus (B)	Used OuT knowledge for selection (B.1)	ordinal[a]	Not any / full coverage[b]
	Used surround system knowledge (B.2)	ordinal[a]	Not any / full coverage[b]
	Stimulus abstraction level (B.3)	ordinal[a]	Real world
Assessment criterion (C)	Abstraction of assessment (direct or indirect) (C.1)	ordinal	Accident risk / SP
	Assessment scale level (C.2)	ordinal	ratio

[a.] An improvement of the scale level to an interval scale level might be achievable due to an application of appropriate scientific methods (measurements, analyses etc.)

[b.] Although full coverage might not be reached, it should be mentioned here for reasons of completeness.

2.2.2 Real World Testing

Testing with a real vehicle in the real traffic comes closest to a straightforward full test approach in dimension A and B. In the following, section 0 describes a deduction of the required testing strategy and effort for a full test approach in all dimensions. Furthermore, it is discussed, why it is impossible to achieve a full test in dimensions B.1 and B.2. Section 2.2.2.2 discusses existing studies that use criticality metrics as an indirect assessment criterion (C.1).

2.2.2.1 Estimating the Effort for a Full Test Approach

A full test approach offers the chance of valid test environment and high completeness of driving scenes, if the test is executed on sufficient driven distance. Based on today's frequency of severe accidents, Winner and Weitzel[69] calculated a minimum testing distance of more than 100 million km to prove superior safety of the OuT compared with the reference frequency of severe accidents. They assumed that the occurrence of an accident is Poisson distributed, that the OuT is in fact twice as safe as the reference, and a 5% error probability. If statistical proof of a reduction of fatal accident on German controlled-access highways is required, the testing mileage increases to more than 6.6 billion km.[70] Kalra and Paddock

[68] Junietz, P. et al.: Evaluation of Safety Validation Approaches (2018), p. 3.

[69] Winner, H.; Weitzel, A.: Die Freigabefalle des autonomen Fahrens (2011).

[70] Wachenfeld, W.; Winner, H.: The Release of Autonomous Vehicles (2016).

calculate a similar mileage for the US[71]. Wachenfeld[72a] generalizes the calculations by introducing the distance factor a_d that is multiplied with the reference distance (e.g. average distance between fatal accidents) to determine the required mileage to proof increased safety. Figure 2-4 visualizes the distance factor that is dependent only on the true OuT's SP and the desired error probability,. The steepness of the dotted line is dependent from the SP of the OuT compared to the reference (index bench). The figure also shows that a proof of decreased safety also requires high mileage, if the OuT is not significantly worse than the reference.

With typical distances between fatal accidents as reference, the required testing mileage is too high to allow an economically feasible proof of safety. Even with a perfectly safe OuT a distance factor of three is derived for an error probability of 5% indicated by the vertical axis.

Instead of field tests on random routes in an unsupervised test, the route could be chosen according to known challenges for the automation (e.g. direct sunlight during dusk) in a supervised field test. So knowledge about the OuT (B.1) would be used and deviation in test result should be discussed.

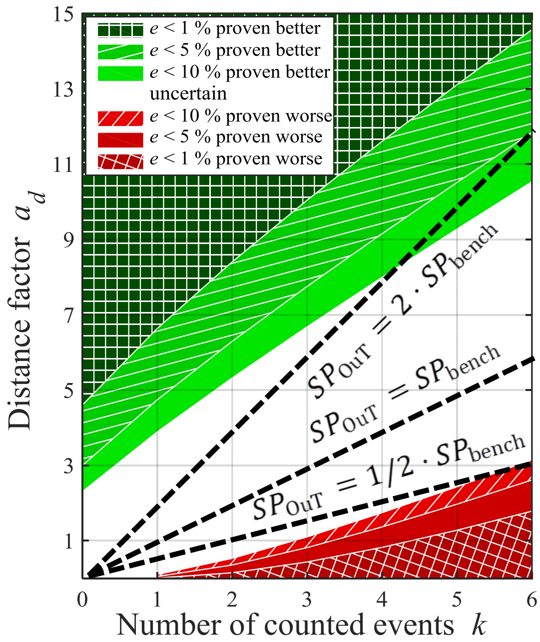

Figure 2-4 Distance factor for different SP (modified from Wachenfeld[72b])

---

[71] Kalra et al.: How Many Miles of Driving Would It Take? (2016).

[72] Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017), p. 47.a: p. 47; b: p. 78

2.2.2.2 Extrapolating the Accident Risk from Critical Scenes

One of the reasons why the required driving distance in real world testing is so high is the low frequency of accidents from which the SP is estimated. Faster evaluation is desirable not only to reduce the testing effort, but also to accelerate the response to road traffic abnormalities after the introduction based on field monitoring. If, instead of accidents, critical scenes could be used as a basis for statistical evaluation, evaluation could be done earlier or faster, as the number of events per distance is higher.

Typically, the metric algorithms assess recorded data. Then either scenarios are identified and their occurrence rate used for safety evaluation[73], or the distribution of metric values is used directly to assess the safety. This section introduces state of the art metrics and extrapolation methods.

2.2.2.2.1 State of the art Metrics

The so called time-headway (THW)[74] or time gap (as in ISO 15622) describes the time that the front of the ego-vehicle needs with the current velocity to reach the point where the rear of the preceding vehicle is at the same point in time. There is no information whether there might be a conflict with the current motion state or not.

The by far most known class of metrics are the Time-to-X (TTX) metrics. They typically assume a constant vehicle motion. As the most simple prediction method, constant velocity and course angle is assumed. Models that assume constant acceleration and turn rate also exist. Then, the time to a certain event is calculated assuming no other interaction. This event can either be a collision to calculate the Time-to-Collision (TTC)[75], or the departure of the road to calculate the Time-to-Edge-Crossing (TTEC)[76]. Both metrics only provide the time left until the incident without any indication, if the prevention of the incident is physically possible or requires increased driving skill. However, they are fast to compute and require only few measurements from the environment. These metrics should only have a correlation to accident probability, if the driver is inattentive, because there is no information about the difficulty of accident prevention.

Metrics[77] that calculate the available reaction time of the ego-vehicle are Time-to-Brake (TTB), Time-to-Steer (TTS), and Time-to-Kickdown. The maximum value of the three metrics is also called Time-to-React (TTR). The main difference to the aforementioned metrics is that they include information about the difficulty of the prevention of an accident. However, there is no information about the difficulty of the maneuver once the driver reacts. As

---

[73] Guo, F. et al.: Near Crashes as Crash Surrogate (2010), pp. 9-12.

[74] Forbes, T. W.: Human factors in highway design, operation and safety problems (1960).

[75] Hayward, J. C.: TTC (1972).

[76] Gordon, T. J. et al.: Multivariate analysis (2013).

[77] Hillenbrand, J. et al.: Situation assessment algorithm (2005).

adaption to the aforementioned time-based metrics, Winner et al.[78] suggest using the reciprocal as measure for criticality. Using the reciprocal, results in a monotone increase of criticality with increasing value.

At least for braking maneuvers, the metrics required deceleration (AREQ)[79] and brake thread number (BTN)[80] provide information about the difficulty of the reaction. Both calculate the required deceleration to prevent an accident that would have happened with constant movement of the involved vehicles, similar to the TTX metrics. BTN is additionally normalized with the maximal physically possible deceleration that is dependent from the available friction coefficient. However, it is uncertain how the driving difficulty and hence the accident probability is correlated with those metrics. Another disadvantage of BTN is that only longitudinal scenarios, where the accident is prevented by braking only, are considered. Jansson[81] suggests a combined metric that calculates the necessary acceleration, also in multiple-object scenarios. However, all aforementioned metrics (except THW) can only be applied in scenes, where constant driving would result in an accident. Scenarios where the object and the ego-vehicle would miss each other are not considered because no reaction is necessary. Nevertheless, they might be critical because minor disturbance would result in a collision.

The Time-to-closest-Encounter (TTCE)[82] calculates the time to the closest approach of the vehicles with constant driving or other prediction models. A similar mechanism is used for the Post-Encroachment-Time (PET)[83] that calculate the time difference between a vehicle leaving an area and the next car entering the same area. In other words, the time gap between two vehicles is calculated when the vehicles drive in an angle up to 90 degrees e.g. in crossing scenarios. Hence, PET can be considered as the extension of THW beyond longitudinal scenarios.

A separate class of metrics uses the measured movement of the ego-vehicle that is evaluated a posteriori without any prediction of the scenario outcome. This is helpful in particular when the environment information is incomplete. In large studies, vehicles are often only equipped with basic sensor setups to reduce costs. Unusual high values of the recordings of the vehicle's motion (e.g. acceleration and steering angle) are an indication for a critical scene. With a single driver, the individual driving behavior would influence the result. Hence, those metrics are typically used in the evaluation of large datasets such as naturalistic driving studies

[78] Winner, H. et al.: Maße für den Sicherheitsgewinn von Fahrerassistenzsystemen (2013).

[79] Karlsson, R. et al.: Decision making for collision avoidance application (2004).

[80] Brannstrom, M. et al.: Situation and threat assessment algorithm (2008).

[81] Jansson, J.: Collision Avoidance Theory (2005).

[82] Damerow, F.; Eggert, J.: Predictive risk maps (2014).

[83] Allen, B. L. et al.: Analysis of traffic conflicts and collisions (1978).

(NDS, e.g. SHRP2[84]) or field operational tests (FOT, e.g. euroFOT[85]), typically together with other metrics (e.g. TTC). Appendix A summarizes the used thresholds that indicate an extreme maneuver . On an even larger scale, accelerations from smartphone data is used[86] to locate areas with increased braking activity.

Metrics that are used in trajectory planning evaluate the whole environment and are feasible for multi-object scenarios and cannot only rely on constant behavior. The target is to find a suitable trajectory that fulfills the driving mission of the ego-vehicle, while assessing the risk due to the environment information.

Yi[87] separates trajectory planning methods into geometric based, sampling based, potential field methods and optimization methods. In sampling-based methods, the prediction of the environment is often integrated. Otherwise, it is predicted in a previous step.

Geometry or maneuver based sampling is a simple and fast way to find suitable evasion or braking maneuvers. Schmidt[88] describes the computation of evasive trajectories in multi-object scenarios. Using the driving dynamic potential, the reachable areas are computed depending on the environment. Stählin[89] parameterizes sigmoid curves to find suitable evasion trajectories. Similar approaches are described by Rodemerk[90] and Käfer[91].

Sampling methods that use varying inputs for acceleration or steering, are suitable to cover complicated maneuver combinations but they are also used for assessing the criticality of a scene:

The most common sampling method is Monte-Carlo sampling[92], where model inputs are generated at random following a given distribution function. The results are different episodes based on the identical scene. Broadhurst et al.[93] assess the criticality of future outcomes, where the inputs are distributed following a normal distribution. Eidehall and Petersson[94] add an additional probability distribution for a reaction to other objects depending on

[84] Hankey, J. M. et al.: Description of the SHRP 2 Naturalistic Database (2016).

[85] Benmimoun, M.: Automatisierte Klassifikation von Fahrsituationen (2015).

[86] Stipancic, J. et al.: Vehicle manoeuvers as surrogate safety measures (2018).

[87] Yi, B.: Integrated planning and control for collision avoidance systems (2018), pp. 10–13.

[88] Schmidt, C.: Fahrstrategien zur Unfallvermeidung (2014).

[89] Stählin, U.: Eingriffsentscheidung für ein Fahrerassistenzsystem zur Unfallvermeidung (2008).

[90] Rodemerk, C. et al.: Development of a general criticality criterion (2012).

[91] Käfer, E.: Situationsklassifikation und Bewegungsprognose in Verkehrssituationen (2013).

[92] Mohamed, M.; Saunier, N.: Motion Prediction Methods for Surrogate Safety Analysis (2013).

[93] Broadhurst, A. et al.: Monte carlo road safety reasoning (2005).

[94] Eidehall, A.; Petersson, L.: Statistical Threat Assessment for General Road Scenes (2008).

their relative position. With Gaussian distribution of inputs, higher acceleration are underrepresented, so the computational effort is high in order to cover all possible trajectories. Zhao et al.[95] therefore use the method of importance sampling, where a different density function is used that favors rare events. After the sampling, criticality is assessed by the relative proportion of accident-free trajectories. Though, one could argue that the majority of trajectories that are generated is more or less straight, but there is no information about the difficulty of the accident-free trajectories. There is also no guarantee that all relevant trajectories are included, especially in scenarios that require more than one maneuver such as a double lane-change. Furthermore, the effort increases when more trajectories shall be covered. To conquer this challenge, Stumper et al. [96] use Chebyshev polynomials with sampled acceleration inputs and varying initial positions of object vehicles. The resulting criticality assessment is used as feature for a machine learning approach in order to assess criticality directly from the initial position.[97]

Instead of sampling, the different future trajectories can be represented by probabilistic methods as in Eggert et al.[98,99] They present a Gaussian prediction that uses different distributions for lateral and longitudinal prediction and a survival analysis that is based on likelihood of collision based on trajectory prediction. Uncertainties are addressed by defining the steepness factor of the resulting risk distribution depending on environment, and maneuver prediction uncertainties.

To find the single best trajectory regarding certain conditions, optimization or potential field methods can be used. A potential field uses a distribution of costs, typically based on surrounding objects that reflect the ego-vehicle. Optimization methods include the driving mission and criticality in a joint cost function to find the trajectory with the minimal costs. Sometimes, constraints are set to prevent the collision with objects.[100,101,102]

2.2.2.2.2 Accident risk extrapolation – existing studies

The primary use case for aforementioned metrics in this thesis is the extrapolation of accident risk or at least the accident probability. This section summarizes existing studies applying metrics together with extrapolation methods. To extrapolate the accident probability from the metrics described above, there are two different methods that are analyzed in this section,

[95] Zhao, D. et al.: Evaluation of lane-change scenarios based on importance sampling techniques (2017).

[96] Stumper, D. et al.: Towards Characterization of Driving Situations (2016).

[97] Stumper, D.; Dietmayer, K.: Towards Criticality Characterization of Situational Space (2018).

[98] Eggert, J.: Predictive risk estimation for intelligent adas functions (2014).

[99] Eggert, J.: Risk estimation for driving support and behavior planning in intelligent vehicles (2018).

[100] Yi, B. et al.: Real time integrated vehicle dynamics control and trajectory planning (2016).

[101] Arrigoni, S. et al.: Safety path planner with collision risk estimation (2016).

[102] Ulbrich, S.; Maurer, M.: Towards tactical lane change behavior planning for automated vehicles (2015).

highlighting the metrics used and possible consequences. The two methods are correlation methods and statistical extrapolation of extreme values.

Stipancic et al.[103] use correlation between the numbers of critical braking maneuvers collected from smartphone data. Maneuvers with an acceleration of less than -2 m/s² are classified as “high braking events”, acceleration of more than 2 m/s² classified as “high acceleration events”. Especially at intersections, the correlation between the frequency of these events and crashes with a Spearman’s rank correlation coefficient of about 0.5 - 0.6 was observed, suggesting a monotonous relation, but strongly depending on the selected threshold and filter parameters of the GPS data. An approach to link the severity of the crash was inconclusive.

Songchitruksa and Tarko[104] first suggested using Extreme Value Theory (EVT) to derive crash frequencies from criticality metrics. They used PET to analyze conflicts during right turn in crossing scenes. For most crossings, the four-year average of collisions was in the 95% confidence interval of the observed counts. However, for some crossings there were overestimations and underestimations outside the interval. The authors assume that the same values of PET can occur in different scenes with different inherent risk, so there is no monotonous relation between PET and risk. However, EVT appears to be a valuable method for extrapolating the crash frequency. The authors contributed four key considerations for the application of EVT:

> “1. *Crash proximity measure corresponding with the studied type of crash must be defined.*
>
> 2. *A valid crash proximity measure must be observable and possess a continuous characteristic that can represent crash-free operations as well as characterize a collision.*
>
> 3. *A definitive boundary between crash and non-crash must exist.*
>
> 4. *The risk estimation method should include a bias-variance trade-off, a choice of r value [the choice of a threshold for asymptotic observations], and identification of nonstationarity and associated covariates.”* [104b]

The first consideration in particular is crucial for the design of criticality metrics and is often neglected. The type of accidents predicted depends on the metric used. If all types of accidents shall be predicted, one or more metrics that address all types of accidents are required.

---

[103] Stipancic, J. et al.: Vehicle manoeuvers as surrogate safety measures (2018).

[104] Songchitruksa, P.; Tarko, A. P.: The EVT approach to safety estimation (2006).a: -; b: p. 821

Guo et al.[105] analyzed near-crash events in the 100-Car NDS. The metrics that determined the near-crash were metrics that measure the driver's action a posteriori (acceleration, yaw rate) and TTC. They found that the causes of crashes are also found in near-crashes and analyzed six causes (distraction, surface conditions, traffic density, lighting, weather and visual obstruction). They found a correlation between the frequency of occurrence of near-crashes and crashes in all six categories. However, the relationship was strongly dependent on the scenario (e.g. conflict with lead vehicle: 380 near-crash; 15 crash, single-vehicle conflict: 48 near-crash; 24 crash). This suggests that the metrics may not be sophisticated enough, e.g. critical scenes solved by other vehicles or requiring only a minor driver action are not detected.

Jonasson and Rootzén[106] analyzed data from the same study using the minimum TTC in a scenario as criticality metric and applied EVT to estimate rear-striking crashes from near-crash scenes. They estimated the crash probability to be $2\cdot 10^{-5}$ in a rear-striking near crash-scene. However, 14 rear-striking crashes and 384 rear-striking critical scenes indicate that the probability is 175 times larger than the estimate. The authors argue that the crashes occurred in slow moving traffic, while the near-crashes occurred in free-flowing traffic. Again, the circumstances of the critical scene vary, so the metrics do not apply to both scenes. Najm et al.[107] used their methodology to predict a crash probability of $1.3\cdot 10^{-5}$ using data from SHRP2. The result is in the same order of magnitude as before. A further indication of the discrepancy of the prediction compared to the true events could be the entity of TTC to return a comparable criticality only with similar relative velocity (comp. Junietz et al.[108]), since the necessary reaction time or TTB develops as follows:

$$\tau_{\mathrm{TB}} = -\frac{d_{\mathrm{long}}}{v_{\mathrm{rel}}} + \frac{v_{\mathrm{rel}}}{2\mu g} = \tau_{\mathrm{TC}} + \frac{v_{\mathrm{rel}}}{2\mu g} \tag{2.1}$$

TTB has a component that is proportional to the relative velocity that must be reduced to zero to prevent a collision. This is not covered by TTC, which proves that scenes with different relative velocities are incomparable.

Gordon et al.[109] first analyzed various criticality metrics for lane departure crashes in the SHRP2 NDS. Using a statistic method called Seemingly Unrelated Regression Models, they identified TTEC as the best metric (out of three) to predict the number of lane departure crashes. They used EVT together with TTEC and compared the estimated rate of lane departure events with the events that had occurred on a monitored highway over four years. According to the estimate, departure events were about seven times more likely than the actual

[105] Guo, F. et al.: Near Crashes as Crash Surrogate (2010), p. 9.

[106] Jonasson, J. K.; Rootzén, H.: Internal validation of near-crashes in naturalistic driving studies (2014).

[107] Najm, W. G. et al.: Analysis of light vehicle crashes and pre-crash scenarios (2003).

[108] Junietz, P. et al.: Metrik zur Bewertung der Kritikalität von Verkehrssituationen und -szenarien (2017).

[109] Gordon, T. J. et al.: Multivariate analysis (2013).

lane departure accident rate. However, not every lane departure necessarily leads to an accident. Therefore, TTEC contradicts the third consideration from Songchitruksa and Tarko (see above), as there is no clear limit in TTEC when an accident occurs. The authors also pointed out that there is a difference between crash risk and injury risk. Hence, two different methods for estimating the potential severity in the near-crash scenes are presented, the estimation by the relative velocity weighted with the object's masses (impulses) and the what-if crash simulation when no evasive action was taken.

Asljung et al.[110,111] applied both TTC and BTN to two different data sets and further applied EVT with both metrics. In a validation step, they compared the estimated crash rate with the MaR known from accident statistics. It was expected that the actual accident rate was within the 5% error interval of the estimated value. However, in this study only BTN was accepted as a valid metric. No collisions were estimated with TTC alone. The authors conclude that criticality metrics describing proximity to an inevitable collision state are better suited than metrics describing proximity to collision.

In conclusion, EVT in particular promises valid extrapolation if an adequate metric is found. When applying EVT, the validity of the metric for the type of accident should be discussed. Some requirements for these metrics are available, but seem to require refinement, as some have proven to be more useful than others, although they are appropriate metrics according to the requirements.

2.2.2.2.3 Introduction to Extreme Value Theory

In the previous section, it was shown that EVT became the dominant method in extrapolating crash likelihood based on critical scenes. This section explains why EVT is useful and how it is applied. The origin of EVT lies in financial statistics and the prediction of natural disasters. It is designed to extrapolate the occurrence of rare events based on measurements or observation of events of lesser severity. Popular examples are the daily losses of stock portfolios or the occurrence rate of high sea levels exceeding dike height.[112,113] By analogy with these examples, the accident occurrence rate could be extrapolated from observations that are scenes of a particular criticality or proximity to an accident. In contrast to the two examples, however, it is unclear for most metrics whether an increased value is actually always closer to an accident. Nevertheless, EVT should be applicable, but the choice of metric influences the result. Until now, only metrics that address a particular type of accident, e.g. front-to-rear collisions was used with EVT.

If the values of a criticality metric applied to the data are considered as independent observations, a distribution function could be fitted to determine the occurrence rate of accidents.

[110] Asljung, D. et al.: Comparing Collision Threat Measures using EVT (2016).

[111] Asljung, D. et al.: EVT for Vehicle Level Safety Validation (2017).

[112] Haan, L. de; Ferreira, A.: Extreme value theory (2007), p. 13.

[113] Coles, S.: An introduction to statistical modeling of extreme values (2001), p. 1.

However, there is insufficient data on the outline of the fitted distribution, so the fit is dominated by uncritical scenarios, while the interesting part, the occurrence rate of accidents, is underrepresented. When the whole data is used, a fit is dominated by the uncritical part due to the high occurrence rate, which is problematic especially when the tail of the distribution is a so-called "fat-tail", i.e. the occurrence rate is (much) higher than with a normal distribution. In these cases, a normal distribution could be the most accurate approximation to the overall observation of criticality values. Nevertheless, it still underestimates the occurrence rate in the tail. The approximation of the extreme values is done with generalized extreme value (GEV) distributions of three different types (Gumbel distribution for exponential decrease, Fréchet distribution for polynomial decrease and Weibull distribution that reaches the occurrence rate of zero)[114a].

In the following, the underlying mathematical principles are described, which lead from observation to the approximated distribution:

If we have a criticality metric I that describes an accident condition, when it exceeds the value I_c. The probability distribution that the maximum of all n observations is less than the value z is described as:

$$P\{\max(I_1, \dots, I_n) \leq z\} = \prod_{i=1}^{n} P\{I_i \leq z\} = \{F(z)\}^n \tag{2.2}$$

The standard approach in statistical modelling would be to estimate F from data, which is a good fit when many data points are available that are similar to z but inaccurate when the density of data points is low. As an alternative approach, EVT directly estimates F^n.

Three different types of extreme value distributions are known (Gumbel, Fréchet and Weibull) that can be combined into a single formula whose shape-parameter ξ defines one of the three types defined for $\left\{z: 1 + \frac{\xi(z-\lambda)}{\sigma} > 0\right\}$ and the three parameters satisfying the conditions $-\infty < \lambda < \infty, \sigma < 0$ and $-\infty < \xi < \infty$.[114a] In this chapter, only formulas for $\xi \neq 0$ are given. The formulas resulting for $\xi \to 0$ and are given by Coles[114b].

$$\{F(z)\}^n \approx G(z) = \exp\left\{-\left[1 + \xi\left(\frac{z-\lambda}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{2.3}$$

The three parameters are fitted from data, typically using a maximum likelihood approach, which is neither further described nor used in this thesis. This approach according to equation (2.3) is the so-called block-maxima approach, since the maximum of the measurements is used. For frequent data collection, the so-called peak-over-threshold approach is advantageous.[114c] Instead of equation (2.2), the probability distribution for a value greater than $I_{\text{thr}} + z$ is formulated as follows under the condition that I_{thr} is exceeded:

---

[114] Coles, S.: An introduction to statistical modeling of extreme values (2001).a: pp. 45-48; b: -; c: p. 74

$$P\{I - I_{\text{thr}} > z | I > I_{\text{thr}}\} = \frac{1 - F(I_{\text{thr}} + z)}{1 - F(I_{\text{thr}})}, z > 0 \tag{2.4}$$

For I_{thr} that are large compared to majority the recordings I:[115a]

$$P\{I - I_{\text{thr}} > z | I > I_{\text{thr}}\} \approx 1 - \left(1 + \frac{\xi z}{\tilde{\sigma}}\right)^{-1/\xi}$$
$$\text{for } y > 0 \text{ and } \left(1 + \frac{\xi z}{\tilde{\sigma}}\right) > 0, \text{where } \tilde{\sigma} = \sigma + \xi(I_{\text{thr}} - \lambda) \tag{2.5}$$

Again, there are various methods for estimating parameters. In the following, the maximum likelihood approach $\ell(\tilde{\sigma}, \xi)$ is used for n observations maximizing the following equation:[115b]

$$\ell(\tilde{\sigma}, \xi) = -n\log\tilde{\sigma} - (1 + \frac{1}{\xi})\sum_{i=1}^{n}\log(1 + \frac{\xi I_i}{\tilde{\sigma}}) \tag{2.6}$$

A procedure for selecting the threshold value I_{thr} is also required.[115c,116] There are no definitive criteria for selection, as it is a trade-off between high variance when the threshold is high and there are few observations above the threshold, and bias when the threshold is too low and the asymptotic basis of the model is violated.[115d] Two methods are available. The first method uses the assumptions that the mean value of observation over the threshold is linear with increasing threshold, as described by Coles[115e]. The plot of the mean values is called the *Mean Residual Life Plot*. The threshold can be estimated before the parameter is estimated. This method is used in chapter 6. The second method is carried out after the parameter estimation and assumes that parameter estimates are stationary. The threshold is selected as the highest value where the parameters remain approximately constant. For both methods, there are no limits that define which threshold is still valid.

When the parameters are estimated, the occurrence rate $1/m$ of observation beyond a certain value $P\{I > I_m\}$ is of interest, which is computed as follows[115f]:

$$P\{I > I_m\} = \gamma_{\text{thr}}\left[1 + \xi\left(\frac{I_m - I_{\text{thr}}}{\tilde{\sigma}}\right)\right]^{-1/\xi} = 1/m$$
$$\text{where } \gamma_{\text{thr}} = P\{I > I_{\text{thr}}\} \tag{2.7}$$

Which can be rearranged to:

$$I_m(m) = I_{\text{thr}} + \frac{\tilde{\sigma}}{\xi}[(m\gamma_{\text{thr}})^{\xi} - 1] \tag{2.8}$$

[115] Coles, S.: An introduction to statistical modeling of extreme values (2001).a: pp. 75-77; b: p. 80; c: pp. 80-85; d: p. 78; e: p. 79; f: p. 81

[116] Asljung, D. et al.: EVT for Vehicle Level Safety Validation (2017), p. 5.

The probability γ_{thr} can be determined directly from the recordings. However, the variance should not be neglected, as well as the variances of the other parameter estimates.[117a] The variances for the parameters $\tilde{\sigma}$ and ξ can be approximated by numerical means by maximizing the equation (2.6), which leads to the variance-covariance matrix, or by analytical calculation of the inverse Hessian matrix.[117b] The variance of I_m is than approximated by the delta-method.[117a]

Plotting I_m over m yields the so called return level plot, which shows the occurrence rate of observations of the value I_m. If the observations are regular, this can be transformed into the average distance between two occurrence of this level and the distance between two occurrences I_c indicates the estimated distance of accidents. An example is given in Figure 2-5, where the green solid line is the estimate, dashed-red lines are the 95% confidence interval of parameter estimate and the value one of the BTN represents the accident.

For safety validation, the worst-case estimate of the safety performance is of particular interest. This is the return period in the lower 95% confidence interval on the return level, which corresponds to an accident (comp. Figure 2-5).

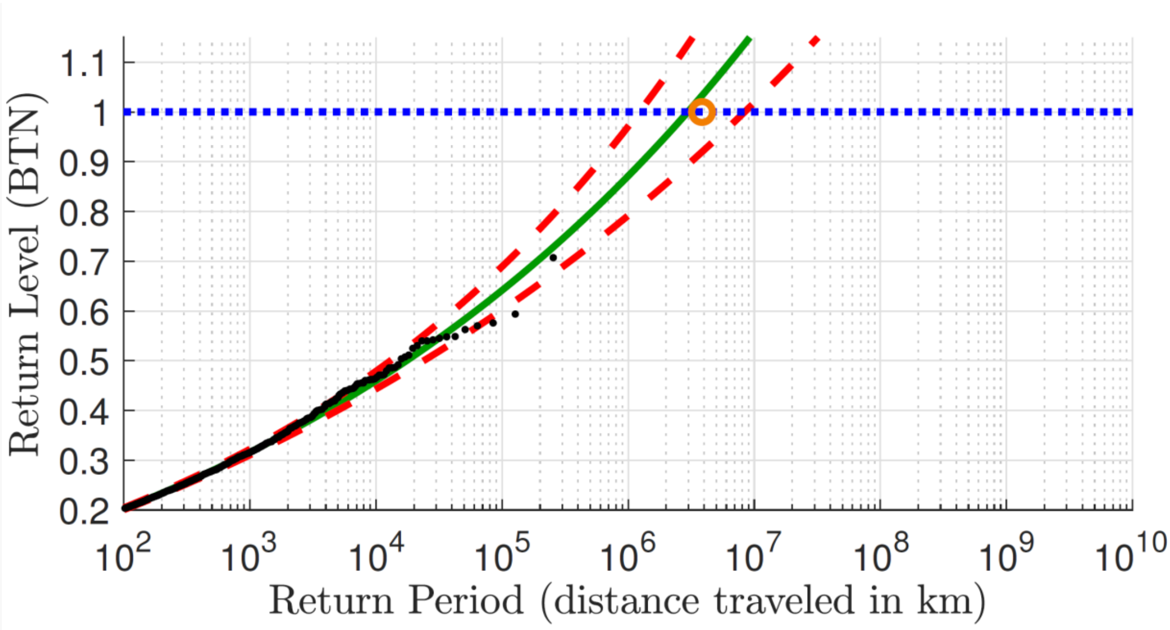

Figure 2-5 Return level plot by Asljung et al.[118]

2.2.3 Formal Verification

Formal verification is technically not a dynamic testing method, because the OuT is not executed. Instead, it is proven that the implemented function does not violate formalized rules. The following assumption need to hold true in order to prove safety by formal verification:

---

[117] Coles, S.: An introduction to statistical modeling of extreme values (2001).a: p. 82; b: p. 40

[118] Asljung, D. et al.: EVT for Vehicle Level Safety Validation (2017), p. 6.

- The rules must prevent accidents
- Proof that the model does not violate the rules exists
- The system and the model show identical behavior regarding the rules

Especially if the system is tested on subsystem level, formal verification could be helpful verifying the planning algorithm. If it is guaranteed that the sensor input is sufficient and the planned trajectory is executed accordingly, safety on planning level could be guaranteed. Using formal verification for AD3+ validation was mainly suggested by Shalev-Shwartz et al.[119] from Intel/Mobileye but also other authors have discussed the concept[120]. They showed that with formalized traffic rules and sensor inputs that are ideally correct or their maximal errors known, a formally safe system could be designed regarding the trajectory planning. With formal verification of the planning, the role of dynamic testing would reduce to the environment perception sensors and the trajectory control. However, especially the validation of the sensors is still a challenge because it is highly dependent on the different environment conditions. Additionally, the other traffic participants' behavior must be formalized and it is unclear if this is feasible in mixed traffic.

2.2.4 Scenario-based Testing

The idea of scenario-based testing is that instead of real world driving, only relevant test cases are extracted and tested in simulation or on test track. A scenario contains several scenes with the scenery (static objects), dynamic objects and the self-representation (e.g. skills, field of view) of the objects. The scenario in a test case could be defined with given object trajectories, or with object goals and a modelled object behavior.

The scenario based testing has several deviations from the full test approach. First is the selection of scenarios that are tested. As not all possible scenarios can be tested due to economic reasons, a selection or identification is required. Typically, scenarios that are challenging for the systems are selected as well as scenarios that are likely to happen. In other words, information about the OuT and the surroundings are used. Both are part of the category stimulus (B.1 and B.2). Additionally, there is abstraction of the stimulus (B.3) because the representation of the environment is artificial and dependent on the simulation tool or the conditions on the test track. It should be argued if deviation from the real entity change the test result. In simulation, the environment is discrete and often modelled without information about material and surface entities that are important for perception sensors. On proving grounds, the variety of the scenario might be limited e.g. because the road surface is always the same and the test track only offers limited objects to be placed on and at the side of the road. In simulation the OuT is abstracted and deviation should be considered as well

[119] Shalev-Shwartz, S. et al.: On a Formal Model of Safe and Scalable Self-driving Cars (2017).

[120] Rizaldi, A. et al.: A Formally Verified Checker of the Safe Distance Traffic Rules (2016).

(A.2). Especially modelling sensors together with the environment to generate realistic effects is still unsolved as summarized for Radar sensors by Holder et al.[121].

Scenario-based testing involves the identification of test scenarios, the execution in a controlled environment, and the assessment of the test result. The project PEGASUS[122] aims to collect driving scenarios in a database in order to derive test cases. In Figure 2-6, the general method is depicted. It is separated into three main parts: the processing of the available information sources, the database where different information is collected, and the assessment and testing process, which is based on the scenario database. In all three parts further progress in research is required that is out of scope of this thesis. In the following the focus is especially on the use of MiR metrics in scenario-based testing.

MiR metrics can be applied in two substeps of the overall method: first in the identification of critical scenarios and second, in the assessment of simulated test runs. However, the assessment of test runs towards the estimation of SP is challenging because the exposure to the scenario is artificial and a human or a capable function could have anticipated the threat to avoid it. It remains unknown how often the scenario would occur in real-world driving. Also, the test scenarios are likely known during development so they are never surprising for the function and testing only the identified scenarios might overestimate the SP.

Therefore, scenario based testing is eligible for comparing the performance of different driving functions or human drivers but not necessarily the overall safety. If all scenarios are handled accident-free, a high confidence in SP might still be established depending on the number of scenarios and the coverage of real driving. Nevertheless, the maturity level of the test process should be qualified, which will be discussed in section 2.2.5. In the following sections, the scenario identification will be discussed further, be it with metrics or other means. Besides the extraction from data, scenarios can also be identified using knowledge about the system and its surroundings or by using ontologies to derive scenarios that cover all possible driving scene.

---

[121] Holder, M. et al.: Measurements revealing Challenges in Radar Sensor Modeling (2018).

[122] PEGASUS stands for **P**roject for the **E**stablishment of **G**enerally **A**ccepted quality criteria, tools and methods as well as **S**cenarios and **S**ituations for the release of highly-automated driving functions, https://www.pegasusprojekt.de.

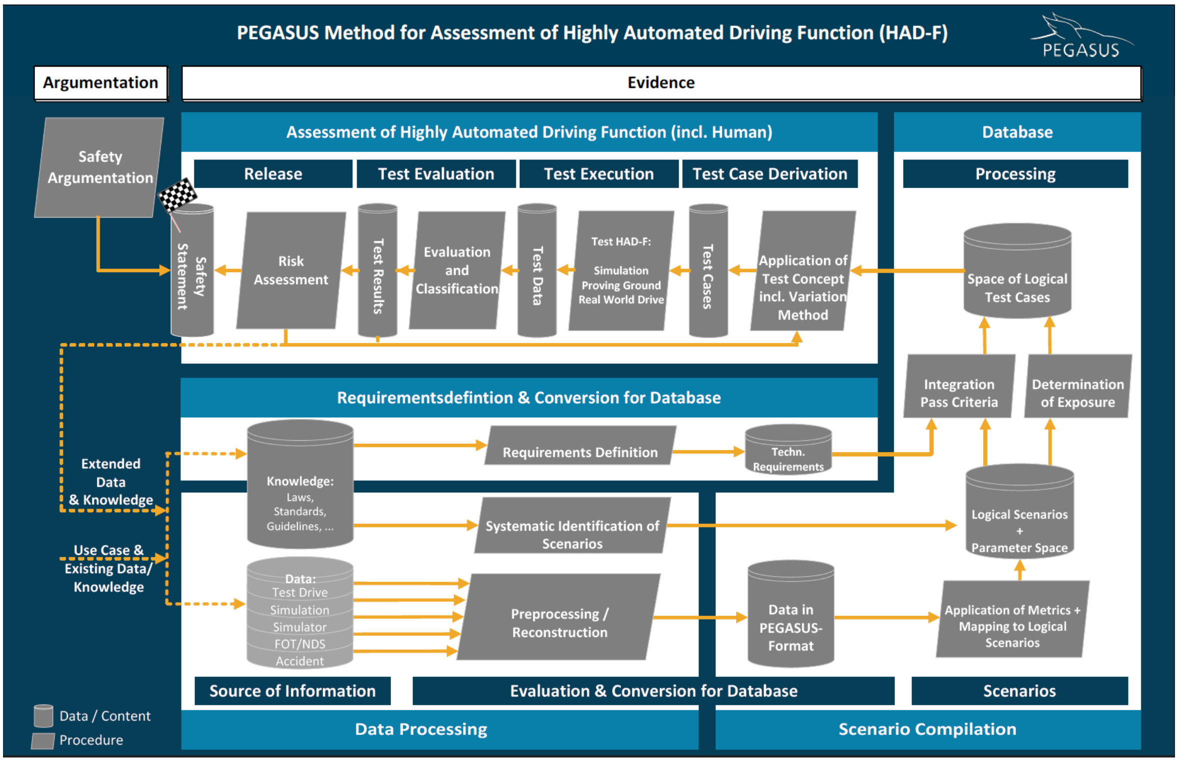

Figure 2-6 Scenario-based testing approach in PEGASUS project[123] after Wachenfeld et al.[124]

2.2.4.1 Identification from Knowledge

Test Scenarios can be derived from knowledge about the system or the occurring traffic conditions. Possible approaches involve the analyses of critical driving scenes that are known from human driven traffic or the analysis of future traffic in a hazard analysis and risk assessment (HARA) as suggested in ISO26262. Büker et al.[125] suggest an identification process based on HARA. They argue that it is impossible to find all scenarios that will be critical for AD3+ from data of today's traffic. The underlying causes for the HARA are separated into three classes: First is the influence of the environment on the system, meaning that the system is unable to handle the environment conditions. As a result errors in the environment perception occur. Second is the influence of the automation on other traffic participants. The driving behavior of AD3+ might differ from the human behavior other traffic participants are used to. Misunderstandings e.g. due to passive and careful driving might be the consequence causing distress in human drivers around the AD3+ vehicle and ultimately causing accidents. The third group contains errors caused by the interaction between the automation and the driver of the automated vehicles. Especially the so called mode confusion, a state in which the driver is unaware if the system is active or not might be an issue. After the HARA

[123] PEGASUS Symposium: Approach & Consistency (2017).

[124] Wachenfeld, W. et al.: Safety Assurance based on an Objective Identification of Scenarios (2016).

[125] Büker, M. et al.: Identifikation von Automationsrisiken hochautomatischer Fahrfunktionen (2019).

underlying causes in scenarios are derived based on an analysis of the chain of effects. Especially for the first two groups of underlying causes, the test scenarios for simulation are derived. The results from the third group could be used to derive guidelines for human machine interfaces.

Even when core entities of a scenario are defined there is still a great number of possible concrete scenarios because of the high number of parameters. Amersbach et al.[126] analyze typical scenarios and suggest a total number of possible scenarios in the order of magnitude of 10^{13} assuming reasonable discretization steps for the entities road width, infrastructure, moving objects and environment conditions. An approach to overcome the challenge of huge scenario numbers that cannot be generated or handled manually is the scenario creation with ontologies that are able to generate realistic scenarios based on formalized rules. Bagschik et al. [127] suggest an ontology using a 5-layer model as depicted in Figure 2-7.[128] The model does not include digital information yet, which are added in a sixed layer by Sauerbier et al.[129] The entities in the scenario are connected by logical constraints that prevent physically impossible combinations of scenarios and decide whether all traffic objects drive according to traffic rules (or not). The road level consists of elements described by the German guidelines for possible road layouts for motorways[130] (Autobahn Regelquerschnitt) and could be adapted in case deviation from these layouts should be allowed. In the infrastructure layer, other static elements such as boundaries or signs are arranged and the orientation towards each other described with logical constraints. The third layer contains temporary manipulation of the previous layer such as a construction site, which is not yet implemented in the cited publication. In the fourth layer, all static and dynamic objects are initialized including maneuvers and the relation to each other. Constraints prevent accidents and other illegal maneuvers (e.g. overtaking on the right, speeding). The fifth layer includes weather and lightning conditions. Constraints prevent combination of conditions such as direct sun light and fog. The layers are connected logically, e.g. heavy rain results in wet road conditions. Using the ontology with less than 1000 attributes, a very large number of scenarios can be generated automatically. However, a completeness in stimulus (B) cannot be reached as the discretization is finite and the environment representation is not (yet) detailed enough. Currently, environment entities such as the shape of cars and the reflectivity of surfaces is not covered.

[126] Amersbach, C.; Winner, H.: Funktionale Dekomposition (2018).

[127] Bagschik, G. et al.: Ontology based scene creation (2018).

[128] Bagschik, G. et al.: Wissensbasierte Szenariengenerierung (2018).

[129] Sauerbier, J. et al.: Definition von Szenarien zur Absicherung automatisierter Fahrfunktionen (2019).

[130] Forschungsgesellschaft für Straßen und Verkehrswesen: Anlage von Autobahnen (2009).

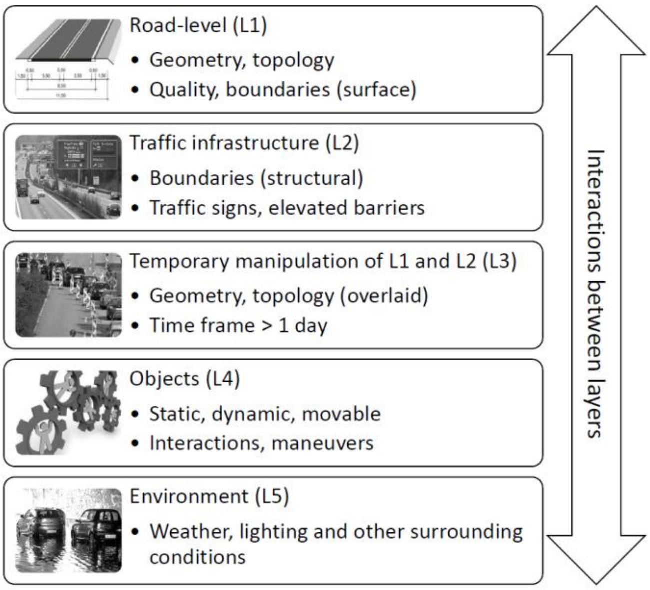

Figure 2-7 5-layer ontology by Bagschik et al.[131] adopted from Schuldt[132]

2.2.4.2 Identification from Data

Besides the scenario generation from knowledge, recorded data can be used. One reason might be to gather information about the relevance or the exposure of the scenarios that are identified from knowledge. The reason is that a huge amount of scenarios can be generated using ontologies, but without any information about relevance, their importance cannot be weighted. For this purpose Krajewski et al.[133] collected trajectory data using a drone and image processing. The occurrence rate of maneuvers such as lane change can be analyzed and the likelihood of whole scenarios derived. However, the data contains only AD2- traffic. As the likelihood of scenarios might change with AD3+ introduction, it is unknown, if the results can be extrapolated. Until AD3+ reaches a relevant field penetration, the occurrence rate from AD3+ perspective can only be assessed in real world test drives and even then only for today's traffic that might change over time. However, a scenario that is likely does not necessary need to be involved as a test case in a scenario-based test concept because the idea is to select test cases that are challenging for the system and hence represent an effective test. To identify critical test cases, MiR metrics, similar to those described in section 2.2.2.2.1 can be used. Wachenfeld et al.[134] suggest a step-wise filter of uncritical scenarios starting

[131] Bagschik, G. et al.: Ontology based scene creation (2018), p. 1817.

[132] Schuldt, F.: Methodischer Test mit virtueller Umgebung (2017).

[133] Krajewski, R. et al.: The highd dataset (2018).

[134] Wachenfeld, W. et al.: The worst-time-to-collision metric for situation identification (2016).

with simple metrics that process huge amount of data efficiently. Hallerbach et al. [135] describe how simulation data can be used to identify critical scenes by modifying object behavior and by injection of sensor or map errors using a smaller amount of data as baseline. For AD2- traffic, various studies were conducted which identified critical scenes using different trigger conditions (see Benmimoun[136a] for a detailed description). All studies use the drivers reaction (e.g. steering, vehicle motion sensors) to identify critical scenes a posteriori. Some also use environment perception sensors to calculate TTC or to detect lane departures. In all studies, the identified scenarios are later analyzed manually using the recorded video data. The largest study is SHRP2[137] with the predecessor 100-Car-Study[138]. The trigger conditions rely mainly on the vehicle motion sensor (e.g. acceleration, yaw rate) and on TTC. Benmimoun[136b] improved the findings further developing a heuristic that is based on similar information, but with more complex trigger condition depending on the driving state. They are summed up in Appendix A. In those studies, the trigger conditions are always designed based on the available data, which is limited to video image and one front object. Studies performing scenario identification based on a whole environment model from a sensor setup that would suit AD3+ systems are not available (or not published).

2.2.5 Maturity Level of the Test Process

Despite all effort in the design of the test suite, it will most likely be incomplete. This is because of a lack of experience of automated vehicles on the roads and because of the unknown future development of road traffic, as the vehicles shall operate in an open world. Winner et al.[139] describe this problem as the *dark matter problem*. An easy task is to derive test cases from recorded accident scenes because the scene can be reconstructed. These are known knowns. Critical scenes that do not result in accidents but that are relevant test cases require metrics to be identified. However, if there is data of those scenes so they are observable. They are named known unknowns because there is the general possibility for identification. Incompleteness of the test suite is the result. However, it is unlikely that there will be a complete test process testing every scene that will happen in the future. No-one knows how the traffic will evolve after the introduction of AD3+ either due to different behavior of the new vehicles or due to changes in road network. Scenarios that cannot be derived from observations of the current traffic are called unknown unknowns. Some of these unknown unknowns might be identified using knowledge based methods but completeness is unlikely.

---

[135] Hallerbach, S. et al.: Simulation-based identification of critical scenarios (2018).

[136] Benmimoun, M.: Automatisierte Klassifikation von Fahrsituationen (2015).a: pp. 30ff; b: -

[137] Hankey, J. M. et al.: Description of the SHRP 2 Naturalistic Database (2016).

[138] Dingus, T. A. et al.: The 100-car naturalistic driving study (2006).

[139] Winner, H. et al.: Validation and Introduction of Automated Driving (2018), p. 184.

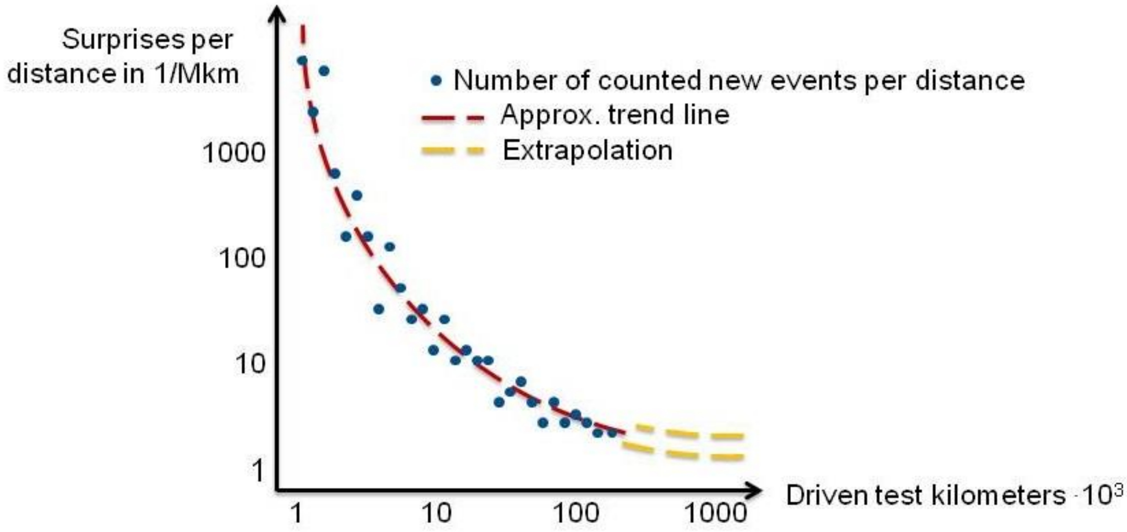

Figure 2-8 Decreasing Trend of new events[141]

Koopman at al.[140] describe the problem using the terms *edge case* and *corner case*. Where a corner case results of the combination of parameters that are not rare individually but the unique combination of several parameters results in a critical corner case. This case is easy to identify with knowledge based methods. Edge cases are scenes that were not expected before and are therefore difficult to gather based on knowledge-based methods.

The maturity level of the test process can be defined by the number and occurrence rate of new edge cases that the automation cannot deal with or at least that result in an unexpected system behavior. Winner et al.[141] suggest measuring the maturity level using real world driving measuring the occurrence rate of surprising scenes. New surprises can be collected and saved to complete the test catalog further. The remaining surprises can be estimated by extrapolation of the decay in new surprises (Figure 2-8). Strategies to find those surprises in test drives are virtual assessment or silent testing in field operation described by Wachenfeld et al.[142] and Junietz et al.[143] or by identification using MiR metrics.

However, it is difficult to estimate the remaining surprises because of the unknown distribution of new events. Koopman[144] argues that assuming an exponential decrease of new surprises instead of a long-tail distribution, the remaining unknowns might be underestimated. Especially if there are surprises that occur only in a very unique combination of scene parameters so that they are reasonably rare to neglect them in the safety approval on their own are problematic, if they occur in a high number. Identifying one new of those rare surprises

[140] Koopman, P. et al.: Credible Autonomy Safety Argumentation (2019), p. 17.

[141] Winner, H. et al.: Validation and Introduction of Automated Driving (2018), p. 191.

[142] Wachenfeld, W.; Winner, H.: Virtual Assessment of Automation in Field Operation (2015).

[143] Junietz, P. et al.: Gaining Knowledge on Automated Driving's Safety--The Risk-Free VAAFO Tool (2019).

[144] Koopman, P.: Edge Cases and Autonomous Vehicle Safety (2019).

(and improving the test catalogue by adding the scenario) does not improve the test catalogue significantly, because the surprises per distance do not significantly decrease resulting in a long-tail distribution.

2.2.6 Summary of Safety Validation Approaches

All presented testing strategies have deviations from a full test approach, so it is unknown whether the true safety performance can be determined with sufficient certainty or only with a large error band. Furthermore, apart from the qualification of the maturity based-on real-world surprise rates, there is no method to quantify the uncertainty (apart from the statistical uncertainty in field tests). Figure 2-9 gives an overview of the shortcoming of the different dynamic testing methods. The unsupervised field tests lead to a safety performance with quantifiable error probability but the required distance for small uncertainty is too large to be feasible. All other approaches have shortcomings in the different categories that lead to uncertainty that is often not quantified (or not quantifiable).

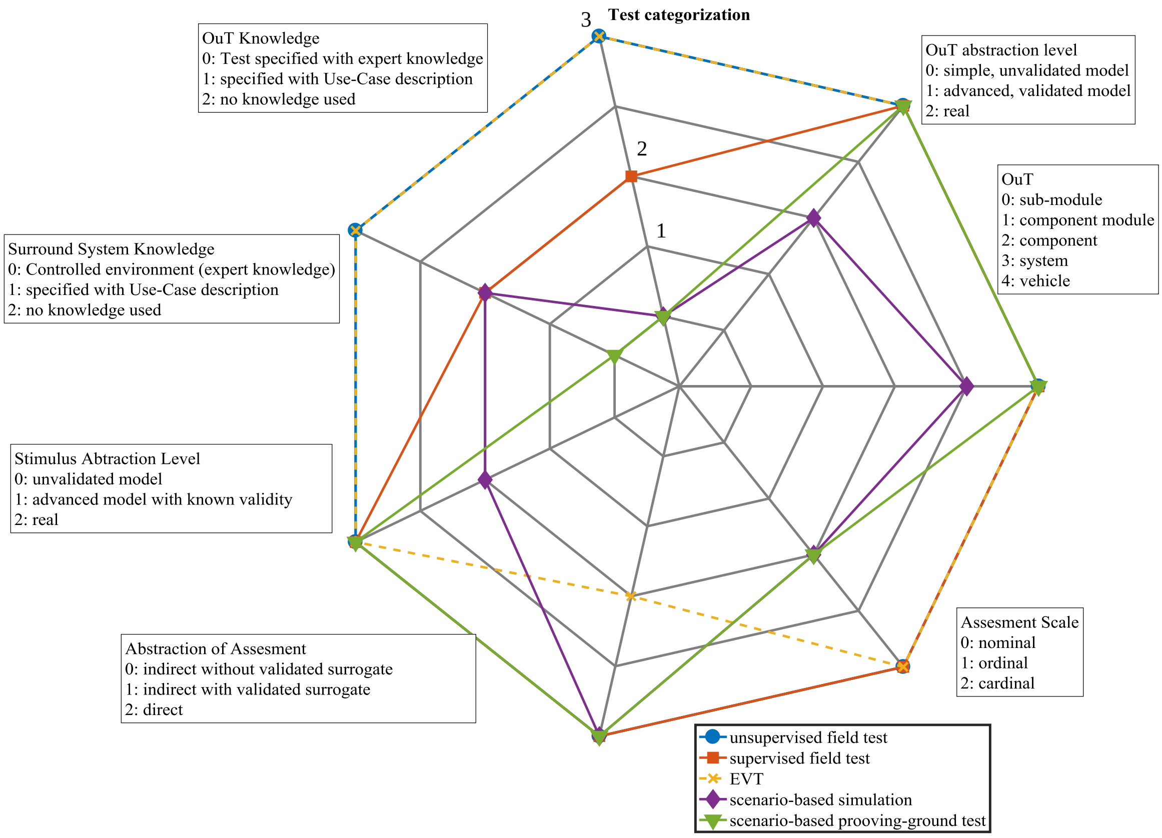

Figure 2-9 Assessment of different test strategies[145]

[145] Junietz, P. et al.: Evaluation of Safety Validation Approaches (2018), p. 4.

In several ways, risk metrics are of outmost importance. MiR metrics are used in the scenario based test for identification of relevant scenarios and for the indirect risk assessment of field tests with extrapolation of accident risk. MaR metrics are required for the description of safety requirements. Existing studies about the extrapolation of critical scenarios towards accident probability and crash risks show some promising results because the statistical uncertainty is quantifiable. However, this does not include and must not be mistaken with uncertainty that origins in the method and is mainly dependent of the used metrics and the data quality.

Those existing studies uniformly use relatively simple metrics. The metric BTN shows the most promising results and corresponds to the definition of criticality from section 1.2.4 because the requirements on use of the available friction coefficient are described. However, the metric only works in longitudinal traffic. While metrics for combined scenarios with multiple objects are available, there were not specified to fulfill requirements for extrapolation methods. Although, most large-scale datasets do not contain enough information to apply those metrics on.

Hence, the focus in this thesis is the improvement of MiR and the derivation of MaR requirements. Additionally, the quantification of uncertainty needs to be discussed. In the next chapter, research questions concerning metrics are derived.

3 Derivation of Research Questions

In the following, the start of the art is analyzed and research questions concerning risk metrics are derived.

3.1 Research Questions for Macroscopic Risk Metrics

Based on the findings in chapter 2.1, it is known that different groups in a society might have different risk perception and therefore different requirements towards technology. However, most customers or passers-by will not provide quantitate requirements for AD3+. Hence, quantitative requirements based on all different viewpoints shall be derived. Only if all necessary and justified viewpoints are considered, final acceptance of AD3+ is likely or even possible. This brings us to the fist research question:

Q 1 *Which viewpoints are necessary to be considered for a sufficient derivation of requirements for macroscopic risk?*

The viewpoints will likely differ based on the individual benefits (of AD3+), because it is known that with personal benefit comes higher risk acceptance (section 2.1.2). In section 2.1.3.1 it is explained that this is even applied when a non-vital benefit comes with the minor probability of severe or even fatal consequences with the example of pharmaceutics. In section 2.1.3.2 introduction of new technologies in aviation is analyzed with the result that risk might increase directly after introduction, but is reduced continuously by gaining knowledge after the introduction and by monitoring accidents and critical scenes. Nevertheless, there should be risk requirements for the first introduction, also if the total number of vehicles is small. However, it is impossible to give a high-certain estimation of the accepted risk when it comes to accidents after the introduction of AD3+. Instead, the acceptable risk is deduced. In addition to acceptable risk, other criteria might influence acceptance. Hence, the second research question is formulated as follows:

Q 2 Which acceptance criteria result, if all viewpoints identified in Q 1 *are analyzed scientifically?*

Currently, there is no quantitative requirements analysis based on the different lifecycle phases, viewpoints and field penetration rate. Wachenfeld[146] introduces the risk-limited introduction, names important stakeholders and analyses the requirements for the society.

[146] Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017).

However, this introduction concept does not include quantitative requirements for individuals. Depending on the introduction phase, different safety targets might apply. When all viewpoints are combined, the third research question is relevant:

Q 3 What is the influence on introduction concepts due to the acceptance criteria derived in Q 2?

With the now defined quantitative risk requirements, introduction strategies such as the risk-limited introduction need to be analyzed. Requirements likely increase with field penetration similar to aviation. The main question will be if the gained knowledge due to the experience in the field is sufficient to reduce the uncertainty of the estimated safety performance enough.

3.2 Research Questions for Microscopic Risk Metrics

A major challenge in approving AD3+ vehicles are the high requirements for preventing rare events with low occurrence rate, such as scenes that could result in fatal accidents. Statistical approval requires high effort due to the baseline of rare accidents. If only the driven distance between two accidents is used as observations for statistical extrapolation of the average occurrence rate, a high driving distance is required because a distance factor between mileage and distance between events higher than ten is realistically required for a statistical estimation (comp. section 0). If not only the distance between accidents is measured but also the distance between scenes with a certain criticality is used to extrapolate the likelihood of accidents, a smaller mileage would be sufficient. To describe criticality in scenes, metrics are required that are called microscopic risk (MiR) metrics. Another purpose of MiR metrics is the identification of critical scenes to derive test-cases in a scenario-based test approach. Here, high requirements apply on false-negative detection rate, i.e. Scenes that are not detected even if they should. Even a single scenario constellation that is missed by the MiR metrics in the evaluation has a severe influence on the total SP if this scenario ultimately results in an accident.

Chapter 2.2.2.2.1 introduces studies using metrics for extrapolation of accident risk as well as for identification of critical scenarios. Metrics that have been applied on large-scale datasets include the vehicle data and sometimes one additional object or information about road boundaries. Metrics that use only vehicle motion sensors such as acceleration or steering wheel angle can detect critical scenes and could even be sufficient to describe criticality on an ordinal scale. However, they do not assess whether the reaction was necessary. Hence, they cannot be applied on scenes where an accident was close but there was no reaction of the driver. Those metrics are developed in bottom-up approaches, as the available data does not allow complex computations because the majority of the environment is unknown. For the purpose of macroscopic risk extrapolation, similar metrics (as for identification) were used in past studies (comp. chapter 2.2.2.2.1) as the available dataset do not allow the application of more complex metrics. Findings of the studies were that the used metrics were not able to predict the MaR in all types of accidents: For example, Songchitruksa and Tarko[147] found that the used metric (PET) did not perform well when applied on data from different crossings. They assumed a non-ordinal connection between the metric and risk when comparing different scenes. Therefore, the first research questions for MiR metrics is:

Q 4 What are the requirements on microscopic risk metrics for extrapolation of MaR using EVT?

Highly detailed data that will be available in the future or is already available such as the highD-dataset[148] raise the opportunity to define metrics in a top-down approach. The definition of criticality for this purpose and the specification of requirements for the metric's performance is necessary. Additionally there are requirements on data that should be discussed as well.

Q 5 What are the requirements on data for application of microscopic risk metrics?

A side effect of metrics that fulfil the requirements derived for *Q 4* is that they can also be used for identification of critical scenarios to derive test cases. Currently, the identification process still requires manual inspection because of the high number of false positive detections due to thresholds that are designed not to miss potentially critical scenes. However, for scene identification, different requirements might apply compared to extrapolation:

Q 6 What are the requirements on microscopic metrics for identification of scenarios?

State-of-the-art metrics are not designed to be applied on data as detailed as described above. Hence, it is likely that the definition of metrics and criticality from section 1.2.4 needs to be redefined. Ideally, a metric or set of metrics that addresses both use-cases is found. Following a top-down approach in development of a new metric, tests that verify the eligibility shall be defined. However, a fulfilled test-case can never verify the eligibility of the metric for application in extrapolation or identification. Hence, according to Popper[149] and Zahar[150], the hypothesis that a metric is eligible for application should at least be falsifiable in order to hold up as a scientific statement. Therefore, a falsification strategy needs to be developed:

Q 7 How can the eligibility of MiR metrics for the use cases MaR extrapolation and identification of scenarios be falsified?

After defining a falsification strategy, the state-of-the-art should be analyzed before designing a new metric. Metrics that utilize environment information typically extrapolate future trajectories of objects. Different strategies include worst-case assumption such as WTTC,

[147] Songchitruksa, P.; Tarko, A. P.: The EVT approach to safety estimation (2006), pp. 819–821.

[148] Krajewski, R. et al.: The highd dataset (2018).

[149] Popper, K. R.: Conjectures and refutations (1969), p. 256.

[150] Zahar, E. G.: Falsifiability (2007), p. 106.

best guess assumption (e.g. constant velocity or constant turn and constant acceleration) or probabilistic predictions (especially in metrics used in trajectory planning). An overview and categorization of different prediction methods is given by Schreier[151]. As risk assessment for the purpose of this dissertation is not necessarily done online, an a posteriori assessment would also be possible, when the objects (future) trajectories are already known. Based on the start-of-the-art review, metrics for trajectory planning might be eligible as well. They should be analyzed first, before developing new metrics. The following research questions apply:

Q 8 Do state of the art metrics fulfil all requirements derived in Q 5 and Q 6?

Due to the lack of sufficient data, falsification might not be achieved for all metrics. There is, however, reason to believe that state-of-the-art metrics are insufficient especially for extrapolation because they are either not designed for highly detailed data or (in the case of trajectory planning) not for the addressed use case which is the assessment of criticality. As it will be equally challenging to falsify a new metric due to the lack of data, design guidelines are proposed that should be followed in order to define metrics that are eligible potentially.

Q 9 What are design guidelines for the development of MiR metrics according to the requirements derived in Q 5 and Q 6?

Based on the design guidelines, a metric will be designed and implemented that has to prove itself based on data that is available currently or in the future. All research questions mentioned in this subchapter will be addressed in chapter 5.

[151] Schreier, M.: Bayesian environment representation, prediction, and criticality assessment (2016), p. 136.

4 Macroscopic Risk

In this chapter, the research questions from subchapter 3.1 are addressed. These are:

Q 1 *Which viewpoints are necessary to be considered for a sufficient derivation of requirements for macroscopic risk?*

Q 2 Which acceptance criteria result, if all viewpoints identified in Q 1 *are analyzed scientifically?*

Q 3 What is the influence on introduction concepts due to the acceptance criteria derived in Q 2?

The first question is addressed in subchapter 4.1, the second in subchapters 4.2, 4.3 and 4.4, and the third in subchapter 4.5.

4.1 Viewpoints on Safety Requirements

Q 1 *Which viewpoints are necessary to be considered for a sufficient derivation of requirements for macroscopic risk?*

To address this question, the difference between viewpoints on safety requirements and stakeholder for safety validation has to be pointed out to begin with. A basic analysis of stakeholder for a safety validation in general is presented in Junietz et al.[152]. This includes the industry (OEM and supplier), and the three government branches: executive, legislature and judiciary. The state institutions shall represent the interest of the people, as safety validation requires expert knowledge. However, the state should address the fundamental requirements for safety out of this group and consider all viewpoints. The goal of this chapter is to find necessary viewpoints sufficient for definition of MaR. In subchapter 2.1, it was pointed out that the individual benefit influences risk acceptance. According to this, viewpoints should be distinguished between users or passengers and non-users or passers-by. Wachenfeld[153] already covers the viewpoints of users and of the whole society but not of other individuals. Additionally, for the user it might be a difference based on his type of exposure, whether it is voluntarily or job-related (comp. Figure 2-2), and for passers-by individual benefits might result as well, e.g. due to the increased overall road safety.

In addition to requirements by individuals, state institutions should define requirements that go beyond the fate of individuals and mainly care about total accident numbers over several

[152] Junietz, P. et al.: Evaluation of Safety Validation Approaches (2018).

[153] Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017), p. 14.

years or decades. In the following, this is described with the term "society" as a separate viewpoint. Obviously, the industry has an interest in safety as well, which should be also dependent on branding and marketing strategies. Other companies or state organs are additional stakeholders as discussed by Junietz et al.[154] They are out of scope of this thesis, because they are not directly threatened in terms of threat to their life or health.

To summarize, the three viewpoints that have to be addressed are:

- users or passengers (further called users)
- non-users or passers-by (further called passers-by)
- society as a whole

Different requirements might result for the three groups in the accident categories of fatal accidents (index f), accidents with injuries (index wI) and accidents without injuries (index nI).

From now on, the frequency of accidents per time (e.g. per year) is described by the symbol *f,* while requirements that are given per distance are described by the symbol $\mathcal{F}$. Assuming an average travel distance per year $\bar{v}$, the time-based frequency f can be transmitted to a distance-based frequency $\mathcal{F}$ and vice versa. In case of Germany, $\bar{v}_{2016,\mathrm{GER}} = 4000\ \mathrm{km/a}$ is assumed as an average travel distance according to recommendations for ISO26262 application[155] and assuming an average velocity of 100 km/h.

4.2 Fatal Accidents[156]

First, requirements for fatal accidents are derived because requirements on the occurrence rate will be the highest out of the three accident groups, and the requirement definition can be formulated based on a broad basis of existing studies (comp. section 2.1.2).

4.2.1 MaR for the User

The fatal risk for the user is assumed equivalent to the risk of a fatal accident of an AD3+ vehicle (neglecting a higher damage with more than one user at the same time). As depicted in Figure 2-2, the type of exposition is relevant for accepting risks. In most use cases, AD3+ functions are used voluntarily; they must be actively bought and activated. Professional use is also plausible but business usage is not expected during the first introduction phase, at

[154] Junietz, P. et al.: Evaluation of Safety Validation Approaches (2018).

[155] VDA: VDA 702 Situationskatalog E-Parameter nach ISO 26262-3 (2015), p. 5.

[156] This chapter was taken from Junietz, P. et al.: Macroscopic Risk Requirements (2019) and extended slightly.

least not as a mandatory prerequisite. Job-related use by choice (e.g. E-mails or video-conferences) is plausible from the beginning on, but is classified as voluntary use as long as not mandatory. It might become relevant with the first driverless systems for trucks, or if regulation allows longer driving duration for truck drivers using AD3+ systems. Involuntary use is excluded in typical use cases because it will not be mandatory to use the systems outside of professional use. Nevertheless, requirements for professional use should be derived from Figure 2-2, as professional use is plausible in foreseeable future.

Figure 2-2 concludes that the accepted frequency for a person's death per year f_{inv} is $10^{-6}/\mathrm{a}$ for involuntary exposure, $10^{-5}/\mathrm{a}$ for professional exposure (f_{prof}), and a theoretically unlimited risk for voluntary exposure with typical acceptance rates f_{vol} of up to $10^{-2}/\mathrm{a}$. In general, the accepted risk varies with the benefit for the user or focus group. Most of these considerations are from the 70's and 80's regarding discussions on safety of nuclear power plants. However, similar to MEM, these assumptions are still valid in general, but should be adapted to today's level of safety. A factor that compensates the increase in general safety based on the development of MEM is suggested and used in the following. (Reduction based on the decrease of accident rate would be another approach with a similar outcome.) In the following, the lower risk today compared to the numbers above is indicated by its index with year and country of the underlying statistic.

Reducing the accepted risk for professional exposure f_{prof} and using the average travelled distance per year $\bar{v}$[157], results in $\mathcal{F}_{\mathrm{prof,2016,GER}}$ equal to $1.4 \cdot 10^{-9}/\mathrm{km}$.. Similar rates are also present in the US ($\mathcal{F}_{\mathrm{prof,2013,US}} = 2.5 \cdot 10^{-9}/\mathrm{km}$)[158]. For other countries, data about mileage on different road types is not always available. The accident rate on all roads' combined mileage is in a similar order of magnitude for most developed countries.[159,160]

However, the substitution of conventional driving also suggests comparing the risk of today's driving with the suggested rate of $\mathcal{F}_{\mathrm{prof}}$. In the following, both considerations will be examined and compared. For today's driving risk, driving on controlled-access highway (Autobahn) in Germany will be taken as a reference. This has several advantages. First, driving on controlled-access highways is one of the safest, if not the safest way of travelling in a car, especially when taking the accident rate per mileage as reference. Second, it is possible that the first AD3+ system will drive on a controlled-access highway because of the reduced complexity. Third, accident data on highways are well documented. Even minor accidents often result in the involvement of police because of traffic disturbance. Data about estimated

157 As the annual driving distance might be higher for professional exposure the requirements might even be higher. Though, it is unknown which use-case for professional exposure will be the first.

158 U.S. Department of Transportation Federal Highway Administration: Fatality Rate per VMT (2013).

159 Oguchi, T.: Achieving safe road traffic—the experience in Japan (2016), p. 115.

160 U.S. Department of Transportation NHTSA: Comparison of Fatality Rates (2016), p. 10.

travelling distance is more accurate than on other roads because traffic surveillance measures the traffic density based on which the estimated travelled distance.

To derive the upper limit for tolerable frequency by users, the experienced benefits caused by the new technology justifies an increased risk (comp section 2.1.2). As a new technology is introduced replacing the former, the increased risk could be derived from the MEM principle from EN 50126. According to the principle the increased risk by the new technology should be smaller than 1/20th of the MEM. In the following, the GAMAB principle and the MEM principle are combined using GAMAB as a baseline risk and MEM as the increased risk. Note that only the risk for the user is considered in this section. This is not applicable to non-users or society at all as a whole, what will be discussed in the following sections. Depending on the individual, even higher risk is acceptable. Here, a careful person is assumed as user. As described above, voluntary risk acceptance might be as high as 10^{-2}/a.

$$\begin{aligned}
&f_{\text{f,User}} \leq f_{\text{GAMAB}} + f_{\text{MEM/20}} \\
&\Rightarrow f_{\text{f,,User}} \leq \mathcal{F}_{\text{f,gamab}} \cdot \bar{v} + f_{\text{MEM/20}} \\
&\Rightarrow \mathcal{F}_{\text{f,User}} \leq \mathcal{F}_{\text{f,gamab}} + f_{\text{MEM/20}}/\bar{v} \\
&\mathcal{F}_{\text{f,User,2016,GER}} \leq 2.15 \cdot 10^{-9}/\text{km};\;\; f_{\text{f,User,2016,GER}} = 8.6 \cdot 10^{-6}/\text{a}
\end{aligned} \tag{4.1}$$

Interestingly, the order of magnitude according to equation (4.1) corresponds to the acceptable frequency for professional exposure. This strengthens the hypothesis that both estimations result in acceptable values for users of automated vehicles. However, higher risk could be accepted by the user (similar to motorbikes or extreme sport) but the user should be aware of this potentially increased risk.

Obviously, not all users will have the same individual requirements. Early adopters and innovators are typically the first to use a new product and might have different risk acceptance. While the general risk acceptance might be similar to regular users, the uncertainty of the current risk is higher for the first users after the introduction. Rogers concludes:

> *"Because the innovator is the first to adopt, he or she must take risks that can be avoided by later adopters, who do not wish to cope with the high degree of uncertainty concerning the innovation when it is first introduced into the system."*[161]

In the following, the acceptable risk according to equation (4.1) will be used for all users but the acceptance of a higher uncertainty for the first users will be discussed.

4.2.2 MaR for Passers-by

For all other traffic participants, AD3+ has no direct benefit (besides the decreased total risk for all traffic participants assuming that AD3+ is safer than the average driver). However,

[161] Rogers, E. M.: Diffusion of innovations (1983), p. 252.

non-users could have a lower risk acceptance threshold because they are skeptical about the new technology or might even have (subjective) disadvantages e.g. due to slow vehicles on the road (comp. section 2.1.2 and Grunwald[162]). Wachenfeld[163] deduces risk requirements for the society and does not consider passers-by as a separate group. However, the risk of new types of accidents is extraordinary important because non-users would blame AD3+ systems for those accidents despite a potential reduction of the total number (comp. Gasser et al.[164]). New risks could be caused for example by systematic software failures, functional insufficiencies, or cyber-attacks. The total new risk of the technology for an individual non-user should be below f_{inv}, or in concrete figure for Germany and year 2016: $f_{\text{inv,2016,GER}} = 2.5 \cdot 10^{-7}/\text{a}$.

So, how can the individual risk for a non-user be calculated? As long as there are not many AD3+ vehicles on the roads, the exposure is very low and the probability that the individual traffic participant is involved in an AD3+ accident is low. Therefore, the risk is multiplied with the field penetration ratio η that will increase over time after the introduction. The risk for passers-by is diluted by the exposure to vehicles equipped with AD3+.

$$\begin{aligned} f_{\text{f,new,2016,GER}} \cdot \eta(t) &\leq f_{\text{inv,2016,GER}} = 0.25 \cdot 10^{-6}/\text{a} \\ \Leftrightarrow \mathcal{F}_{\text{f,new,2016,GER}} &\leq \frac{1}{\eta(t)} \cdot 6.25 \cdot 10^{-11}/\text{km} \end{aligned} \tag{4.2}$$

According to equation (4.2), the accepted risk for a single AD3+ vehicle decreases with increasing number of AD3+ vehicles in the field. This is intuitively obvious since exposure multiplies with the number of potential single threats. So when only few AD3+ vehicles are in the field, the passers-by's requirements are neglectable due to the dilution factor. With increasing field share, requirements increase.

This calculation deliberately neglects that the non-user also has benefits if the system is safer than the human driver it replaces. However, long distances of AD3+ systems must be driven before there is statistical proof of those benefits. Before this is an undeniable statistic fact, skepticism due to the (subjective) disadvantage of the new technology might stay dominant.

4.2.3 MaR for the Society

For society, the fate of individuals is of lesser importance. Benefits and costs of the introduction of AD3+ systems are measured by the total number of accidents and whether they are reduced over time. In general, a decreasing trend of accident rate throughout the years

---

[162] Grunwald, A.: Societal Risk Constellations (2016).

[163] Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017), p. 83.

[164] Gasser, T. M. et al.: Rechtsfolgen zunehmender Fahrzeugautomatisierung (2012), p. 11.

can be observed in Germany[165] and the US[166]. However, in Germany it is observed that this trend has been diminishing over the last 5 years for accidents with injuries and even had a slight increase during these years. For fatal accidents, a similar trend is observable as well.

When introducing AD3+ systems, there will likely be a non-zero risk of severe accidents and therefore it is likely that AD3+ will be involved in those severe or even fatal accidents. However, since the total number of AD3+ vehicles will be insignificant compared to the majority of AD2- vehicles, it will take a while until accident numbers and mileage increase in order to evaluate the systems with significance (comp. Figure 2-4). So, what are the requirements by society, if individual accidents do not influence the total number significantly?

What is the upper total accident rate limit accepted by society?

The overall target is to reduce the amount of accidents over time with the introduction of new technology. If we follow the argumentation of Wachenfeld[167] and Kalra[168], we should allow a certain risk in order to bring AD3+ systems to the market and allow gaining further knowledge. At the same time, it would not be acceptable for the whole society if the total risk increases in a noticeable way.

However, there is no way to check how accident numbers would have evolved without the technology as soon as it has entered the market. Wachenfeld interpolates the accident numbers of the years 1992-2014 and suggest a standard deviation of 39 fatal accidents per year as a reference[169] for a maximum deviation caused by AD3+. However, in the last decade, the decrease of fatal accidents and accidents with injuries diminished. At the same time, the annual travel distance increased. Hence, it seems justified to use recent numbers as reference. When using the accident rate for fatal accidents $\mathcal{F}_{\mathrm{d}}$, an exponential regression is a better fit than a linear regression. Interestingly, this is also the case for accidents in aviation (comp. section 2.1.3.2). The standard deviation of the exponential regression for all years since 2010 results in:

$$\sigma\mathcal{F}_{\mathrm{f}}(N, i_0) = \sqrt{\frac{1}{N - i_0}\sum_{i_0=2010}^{N=2016}\left(\mathcal{F}_{\mathrm{f,i}} - \mathcal{F}_{\mathrm{f,exp}}(i)\right)^2} = 9.4 \cdot 10^{-11}\,\frac{1}{\mathrm{km}} \tag{4.3}$$

Multiplying the standard deviation with the average annual mileage in 2016 results in 22.9 fatal accidents per year, which is slightly lower than what Wachenfeld calculated. However, it must be pointed out that the type of regression and the number of years influence the result. It is also possible to use other factors of the standard deviation as a measure. However, the

165 Destatis: Verkehrsunfälle - Fachserie 8 Reihe 7 - 2015 (2015).

166 U.S. Department of Transportation NHTSA: Fatality Analysis Reporting System (2017).

167 Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017), pp. 102ff.

168 Kalra, N.; Groves, D. G.: The Enemy of Good (2017).

169 Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017), p. 84.

results will be in a similar order of magnitude. Below, the result from equation (4.3) will be used.

The requirement by society should be a risk from AD3+ significantly lower than the described exponential trend observed in the latest data. As a result, AD3+ should be at least one standard deviation $\sigma\mathcal{F}_{\mathrm{f}}$ better than the predicted performance of conventional driving. However, society should give AD3+ systems time to reach this high safety reference. Similar to air traffic, it is necessary to monitor the performance to enable improvement in functions, infrastructure, and user experience. In the following scenario, it is suggested to allow additional risk of <u>one</u> standard deviation at the beginning of introduction and demand a risk three standard deviations lower than the extrapolation, when full field penetration is reached. Furthermore, the change between the beginning and the full field penetration is assumed linear with $\eta(t)$. Therefore, the acceptable risk not only depends on the development of the risk in conventional traffic over the years, but also on the field penetration ratio $\eta(t)$.

$$\begin{gathered} \mathcal{F}_{\mathrm{f,acc,soc}}(t)\cdot\eta(t)+\mathcal{F}_{\mathrm{f,exp}}(t)\cdot(1-\eta(t)) \\ \leq (\mathcal{F}_{\mathrm{f,exp}}(t)+\sigma\mathcal{F}_{\mathrm{f}})\cdot(1-\eta(t))+(\mathcal{F}_{\mathrm{f,exp}}(t)+3\cdot\sigma\mathcal{F}_{\mathrm{f}})\cdot\eta(t) \\ \Leftrightarrow \mathcal{F}_{f,\mathrm{acc,soc}}(t)\leq\mathcal{F}_{\mathrm{f,exp}}(t)+\sigma\mathcal{F}_{\mathrm{f}}\frac{1-4\eta(t)}{\eta(t)} \end{gathered} \tag{4.4}$$

To compare the different requirements with each other, a field share $\eta(t)$ is required to describe all requirements on a time-based level. It is assumed that $\eta(t)$ develops similarly to the field share of other driving functions such as electronic stability control (comp. Wachenfeld[170]). Full field share is assumed to be reached after 30 years and described by a cosine function from 0 to π:

$$\eta(t)=(1-\cos\,\pi\cdot t/T)/2;\ 0\leq t\leq T \tag{4.5}$$

However, the true development is highly speculative. A slower penetration is also likely e.g. due to high costs and remaining challenges in the validation.

4.2.4 Summary of Safety Requirements

In the previous sections, safety requirements based on three different viewpoints were deduced. For society, the acceptable risk depends on the market share of AD3+. The authors suggest allowing an increase in total risk by one standard deviation of the predicted accident rate, so AD3+ can be introduced although the knowledge about its safety level is not yet complete. In addition to society's requirements, passers-by (as part of society) have increased requirements for new risks that come with automation. For users, constant risk requirements are used in the following discussion although they might increase with the current traffic safety over the years. However, the user's requirements are only dominant in the early introduction phase (comp. Figure 4-1) when the market share is relatively low. From a

[170] Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017), p. 105.

market share of about 10%, the requirements of society (and non-users) are dominant. However, if the field share reaches 100%, there are no non-users remaining, at least on Autobahn. However, it has to be assumed that AD2- driving will coexist for a long time.

In the following table, the requirements are summed up. Since data for the accident rate on the whole road network is in the same order of magnitude for developed countries (see above), similar figures result. As discussed in section 4.2.2, early-adopters might have an increased acceptance for uncertainties of risk. As the risk requirements at the beginning of introduction are dominated by the user's requirements, introduction despite remaining uncertainty in risk might be feasible. This will be discussed in detail in subchapter 4.5.

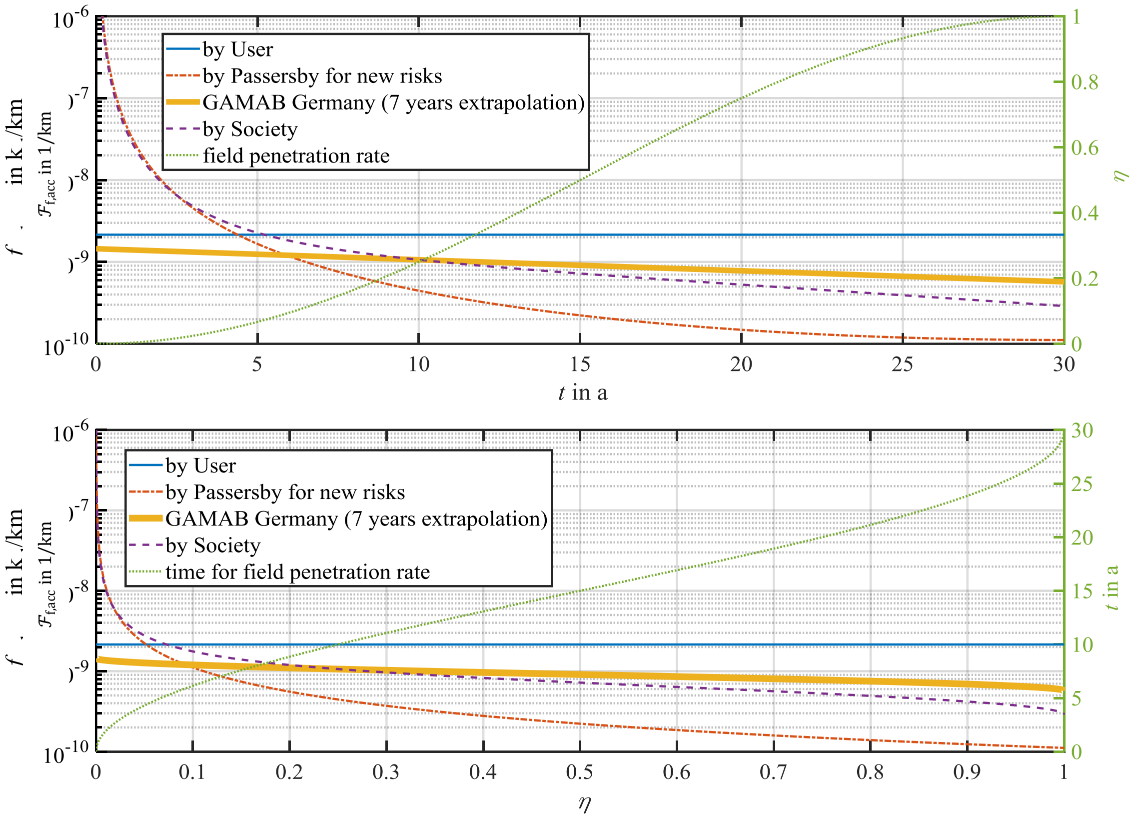

Figure 4-1 Safety requirements for the different focus groups ever time and field penetration rate[171]

[171] Junietz, P. et al.: Macroscopic Safety Requirements for Highly Automated Driving (2019), p. 8.

Table 4-1 Summary of Safety Requirements

Description	**Symbol**	**Value based on German data from 2016**
User requirements		
Per Distance	$\mathcal{F}_{\mathrm{d,User}}$	$2.2 \cdot 10^{-9}/\mathrm{km}$
Per Time	$f_{\mathrm{d,User}}$	$8.6 \cdot 10^{-6}/\mathrm{a}$
Passers-by requirements for new risks		
at η =0.1	$\mathcal{F}_{\mathrm{d,new}}$	$6.3 \cdot 10^{-10}/\mathrm{km}$
at η =0.1	$f_{\mathrm{d,new}}$	$2.5 \cdot 10^{-6}/\mathrm{a}$
at η =1	$\mathcal{F}_{\mathrm{d,new}}$	$6.3 \cdot 10^{-11}/\mathrm{km}$
at η =1	$f_{\mathrm{d,new}}$	$2.5 \cdot 10^{-7}/\mathrm{a}$
Society requirements		
In 5 years at η =0.095	$\mathcal{F}_{\mathrm{d,soc}}$	$1.8 \cdot 10^{-9}/\mathrm{km}$
In 5 years at η =0.095	$f_{\mathrm{d,soc}}$	$7.2 \cdot 10^{-6}/\mathrm{a}$
In 30 years at η =1	$\mathcal{F}_{\mathrm{d,soc}}$	$2.9 \cdot 10^{-10}/\mathrm{km}$
In 30 years at η =1	$f_{\mathrm{d,soc}}$	$1.2 \cdot 10^{-6}/\mathrm{a}$

4.3 Requirements for Accidents of Lower Severity

The requirements derived above are based on studies focusing on the acceptance of fatal risks. For accidents without fatalities there is no such profound background research about accepted risk for injuries or property damage based on type of exposure. For this reason requirements cannot be deduced in the same way. Only for society's requirements the relevant figures are available because they are based on accident statistics.

Instead, there are several possibilities to deduce requirements for crash frequencies of less severity based on above findings for fatal accidents. They are discussed in the following. For simplification, the possible severity levels are separate into fatal accidents, accidents with injuries (Index wI) and accidents without (no) injuries (Index nI).

4.3.1 Reduction according to Hydén Triangle

As introduced in subchapter 2.1, Hydén[172] assumes that a reduction of less severe accidents goes always hand in hand with a reduction of more severe accidents. His findings are based on accident statistics over the years and are supposedly accurate for AD2- traffic. So can this be expected for AD3+ vehicles as well and what are the consequences concerning MaR requirements? Due to the paradigm shift from human to machine driving, there is no guarantee that the same behavior as discovered by Hydén can be observed. This is easily clarified by a hypothetical systematic fault in scenes with lower speeds where there is no hazard of higher severity. The frequency of less severe accidents might even increase, while the other categories decrease. Obviously it is also possible that systematic faults happen only at high speed with the inverse effect.

However, it would be plausible to demand the same quotient between MaR requirements and current accident statistics for all severity levels. That means that the demand for a relative reduction in fatal accidents triggers the same demand for all categories. For users and passers-by it follows:

$$\frac{\mathcal{F}_{\mathrm{nI,user}}}{\mathcal{F}_{\mathrm{nI,exp}}} \equiv \frac{\mathcal{F}_{\mathrm{wI,user}}}{\mathcal{F}_{\mathrm{wI,exp}}} \equiv \frac{\mathcal{F}_{\mathrm{f,user}}}{\mathcal{F}_{\mathrm{f,exp}}} \tag{4.6}$$

$$\frac{\mathcal{F}_{\mathrm{nI,inv}}}{\mathcal{F}_{\mathrm{nI,exp}}} \equiv \frac{\mathcal{F}_{\mathrm{wI,inv}}}{\mathcal{F}_{\mathrm{wI,exp}}} \equiv \frac{\mathcal{F}_{\mathrm{f,inv}}}{\mathcal{F}_{\mathrm{f,exp}}} \tag{4.7}$$

For the society's requirements equations (4.3) and (4.4) are used with the extrapolated frequencies for accidents with property damage and with injuries $\mathcal{F}_{\mathrm{wI,exp}}$ and $\mathcal{F}_{\mathrm{nI,exp}}$ as basis.

4.3.2 Monetary Balance

The aforementioned approach requires the same relative reduction of accidents in all three categories. So if there were no fatalities at all, but no reduction in total accident numbers, the requirements would be failed. In order to weight an increased reduction in one category with a reduced reduction in another, Wachenfeld[173] suggests a monetary balance between different severity categories using average accident costs $\overline{K}$ of the different severity categories for the German national economy[174]. If the occurring accident frequency (index occ) is not below the number of acceptable accident frequency (index acc), balancing between the different costs could be done in the following way:

[172] Hydén, C.: Method for traffic safety evaluation (1987).

[173] Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017), pp. 26ff.

[174] Baum, H. et al.: Volkswirtschaftliche Kosten durch Straßenverkehrsunfälle in Deutschland (2010).

$$\mathcal{F}_{\text{nI,occ}} \cdot \overline{K}_{\text{nI}} + \mathcal{F}_{\text{wI,occ}} \cdot \overline{K}_{\text{wI}} + \mathcal{F}_{\text{f,occ}} \cdot \overline{K}_{\text{f}} \leq \mathcal{F}_{\text{nI,acc}} \cdot \overline{K}_{\text{nI}} + \mathcal{F}_{\text{wI,acc}} \cdot \overline{K}_{\text{wI}} + \mathcal{F}_{\text{f,acc}} \cdot \overline{K}_{\text{f}} \tag{4.8}$$

It might be ethically questionable to derive safety requirements that allow shift between the categories, as it allows the compensation of fatalities with monetary damages. Alternatively, it could be allowed to compensate increases in the categories of less severity with a decrease in higher severity but not vice-versa. This is also in accordance with the German constitution (Grundgesetz Article 2(2)) that protects life and physical integrity. In addition to equation (4.8), the following conditions apply:

An increased accident frequency of fatal accidents would not be tolerated:

$$\mathcal{F}_{\text{f,occ}} \leq \mathcal{F}_{\text{f,acc}} \tag{4.9}$$

Accidents with injuries can be compensated by an increased reduction of fatal accidents but not of accidents without injuries:

$$\mathcal{F}_{\text{wI,occ}} \cdot \overline{K}_{\text{wI}} + \mathcal{F}_{\text{f,occ}} \cdot \overline{K}_{\text{f}} \leq \mathcal{F}_{\text{wI,acc}} \cdot \overline{K}_{\text{wI}} + \mathcal{F}_{\text{f,acc}} \cdot \overline{K}_{\text{f}} \tag{4.10}$$

An increase in accidents with property damage without injuries could still be compensated according to equation (4.8).

4.3.3 No Requirements for Lower Severity Categories

Opposite to the approaches discussed above, one could also argue that no additional requirements for less severe accidents are necessary, because saving lives outweighs preventing injuries. However, it is questionable, if this is applicable when injury number or even personal damages increase dramatically. Hence, requirements following equation (4.6) and/or (4.8) to (4.10) are favorable. Alternatively, equations (4.8) could be used as soft requirements, with mandatory or optional system updates when (4.9) and/or (4.10) are violated.

4.4 Acceptable versus Accepted Risk

In the previous sections of this chapter, risk requirements that should be acceptable are derived. This was done using past studies of accepted risks, current accident statistics from road traffic as well as aviation and a dilution of the risk for passers-by because of the exposure that is dependent of the field penetration. However, there is no guarantee that those risk requirements are accepted later on. There exist studies that try to evaluate accepted behavior and risk of future AD3+ vehicles (Liu et al.[175]). However, other studies show that novices in

[175] Liu, P. et al.: How Safe Is Safe Enough for Self-Driving Vehicles? (2018).

risk assessment are unable to estimate the accurate fatality risks[176]. People tend to underestimate risk they are in control of and overestimate risk coming from an unknown or uncontrollable cause. So risk acceptance for a final product is not only dependent from facts but from the perceived risk that is also influenced by media and personal experience of individuals including the perceived benefit (comp. section 2.1.1 and Grunwald[177]).

4.5 Introduction Strategy

The findings from Figure 4-1 do not contradict risk-limited introduction. With the assumed market share development, the requirements by the user is dominant for the first five years after introduction. However, the statistical proof by unsupervised real-world driving is still unfeasibe before the introduction. After the introduction phase of five years, the requirements increase with market share. Assuming that the AD3+ vehicles drive the same annual distance $\bar{v}$ as the average vehicle, the distance driven is calculated as follows:

$$s_{\mathrm{AD3+}}\left(t_{\mathrm{AD3+}}\right) = \bar{v} \int_{0}^{t_{\mathrm{AD3+}}} \eta(t)\mathrm{d}t \tag{4.11}$$

After five years, the total mileage would add up to $2.74 \cdot 10^{7}$ km. Assuming that the initial safety performance was twice as good as the reference, the statistical proof of safety would be complete according to Figure 2-4. However, as the social requirements increase over time, a system that is twice as good as the current safety level, would be obsolete within five years. Another issue is that the society might decline the system in the initial state. Hence, further considerations are required.

As the society is skeptical about the system, it might demand proof for increased safety. As discussed earlier, this could be done by rejecting the hypothesis that the system's safety is inferior with a confidence interval of 5%. So the null-hypothesis ($H_{0,\mathrm{soc}}$) would be:

$H_{0,\mathrm{soc}}$: The systems safety is equal to the reference safety of the society.

The one-tailed directional alternative hypothesis of lesser safety that must be rejected would be:

$H_{1,\mathrm{soc}}$: The systems safety is inferior to the reference safety of the society.

As in Wachenfeld[178], the hypothesis could be rejected by evaluating the occurring accidents and the driven distance. Assuming that accidents occur exactly in the frequency that corresponds to the systems average performance, the minimal reference frequency that results in

[176] Slovic, P. et al.: Facts and fears: Understanding perceived risk (1980).

[177] Grunwald, A.: Societal Risk Constellations (2016).

[178] Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017), pp. 73–79.

a rejection of $H_{1,\text{soc}}$, can be determined. It can be regarded as the proven safety of the system as a worst-case assumption.[179]

$H_{1,\text{soc}}$ is rejected for the reference $\mathcal{F}_{\text{ref}}$, if the following condition is fulfilled with k_{f} as the occurring number of fatal accidents, assuming that the probability can be calculated with an Bernoulli process B:

$$\begin{aligned} &P(x \leq k_{\text{f}}) \leq 5\% \\ &\Leftrightarrow B(x \leq k_{\text{f}}(t) | \mathcal{F}_{\text{ref}}(t), s_{\text{AD3+}}(t)) \leq 5\% \end{aligned} \tag{4.12}$$

$\mathcal{F}_{\text{ref}}$ can be found by conducting a numerical search or by using tables documenting the Bernoulli process. Before a first accident can be expected or before any distance is driven, there is no initial reference safety rejecting $H_{1,\text{soc}}$. As a result, the requirements directly demand a certain test drive before introduction that is added to the distance in equation (4.11). In the following, an initial testing distance of 10 million km is assumed without any fatal accidents.

Figure 4-2 depicts the resulting safety level that is able to reject $H_{1,\text{soc}}$ with two different assumptions: First, a system, which is perfectly safe and will not cause any accidents is assumed. Second, the assumption that the safety matches the society's requirements after 30 years is taken into account.

Figure 4-2 assumes constant SP of the automated system. If the initial safety performance is worse, improvements of the system should be allowed as long as it seems promising that AD3+ will achieve the requirements with the improvements. For example, if the system initially fulfills only the user requirements, the society will reject the system after a short period of about four years so the successful deployment of such a system is only possible, if continues software updates improve the system over time. Figure 4-3 shows the development of $\mathcal{F}_{\text{ref}}$ under the assumptions that the accident rate fulfills the user requirements or that it initially fulfills only the requirements and improved by 10% per year. In the second case, society's requirements are fulfilled permanently.

The future safety of road traffic was analyzed by Kalra and Groves[180]. They conducted a parameter variation study with influence factors such as deployment of AD3+, improvement rate per distance and final safety performance (after improvement).

Especially the improvement rate is a key factor, hence it is of extraordinary importance not only to improve the function after accidents but use critical scenes as well. An additional indicator for the safety performance is the occurrence rate of critical scenes that is extrapolated towards crash risk. Those two use-cases for MiR metrics are addressed in the next chapter.

---

[179] Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017), p. 131.

[180] Kalra, N.; Groves, D. G.: The Enemy of Good (2017).

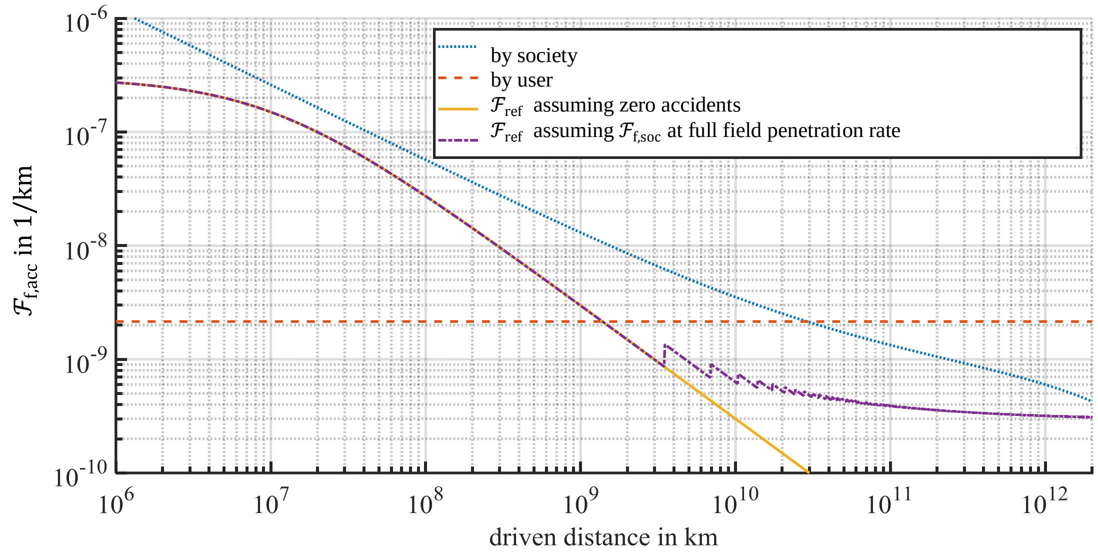

Figure 4-2 Introduction strategy assuming constant accident frequency and 10 million km of tests before introduction

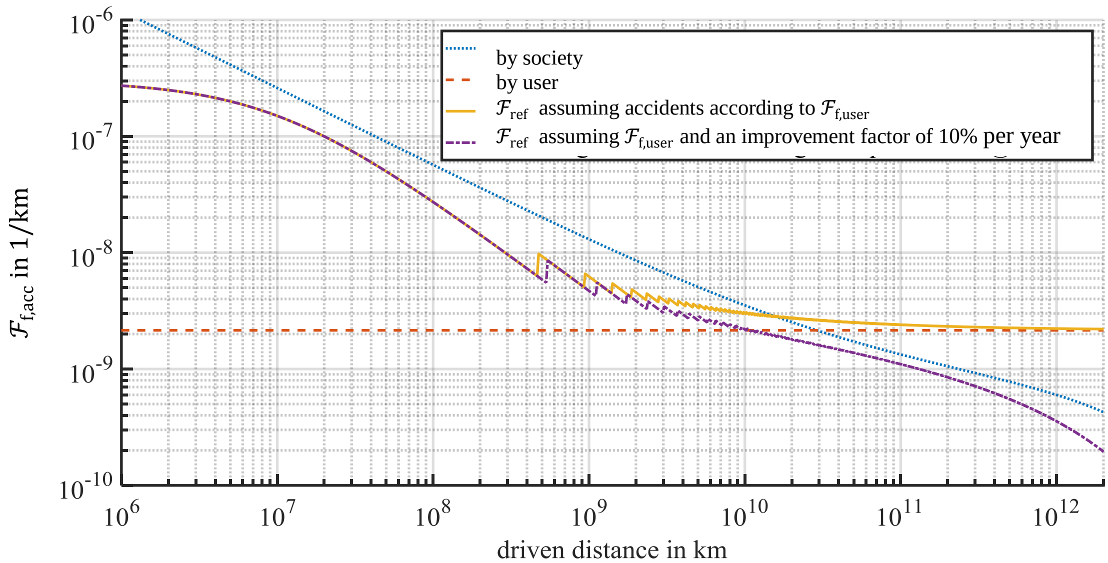

Figure 4-3 Introduction strategy assuming constant improvement factor and 10 million km of tests before introduction

5 Microscopic Risk

In this chapter, the research questions from subchapter 3.2 are addressed.

The chapter's structure follows a top-down approach. First, requirements are defined and test cases for verification and falsification deduced and applied with state of the art metrics. Then, possible approaches for a metric that complies with all requirements are discussed and an exemplary metric designed. Finally, the metric is applied on a dataset and the results discussed.

5.1 Requirements for Metrics and Data

To begin with the top-down approach, requirements shall be defined. They are separated into requirements for data and metrics. Data requirements are often neglected because recording highly detailed data is costly and therefore often reduced to lower quality and the metric then developed bottom-up.

In this subchapter, these research questions are answered:

Q 4 What are the requirements on microscopic risk metrics for extrapolation of MaR using EVT?

Q 5 What are the requirements on data for application of microscopic risk metrics?

Q 5 What are the requirements on data for application of microscopic risk metrics?

5.1.1 Requirements for Metrics

Q 4 What are the requirements on microscopic risk metrics for extrapolation of MaR using EVT?

Q 6 What are the requirements on microscopic metrics for identification of scenarios?

From literature, the requirements by Songchitruksa and Tarko for extrapolation are known[181] (see section 2.2.2.2). They are refined slightly in the following. There first requirement demands that the metric is eligible to describe proximity to a crash of a certain type.

Furthermore, the metric must be applicable in all scenes and describe crash-free operations. Collisions do not need to be covered by the metric because they could be detected by other means, e.g. reports by the user. However, the metric's value must increase if the criticality

[181] Songchitruksa, P.; Tarko, A. P.: The EVT approach to safety estimation (2006), p. 821.

of the scene is increased. Otherwise the extrapolation might over- or underestimate the macroscopic risk. Different kinds of scenarios must be comparable regarding the metric, meaning that the identical value of the metric in two different scenarios must correspond to the same crash likelihood. Due to the lower occurrence of highly critical scene, an overestimation of the metric is likely to have a severe impact on the result. The first requirement is formulated as follows:

RM 1 In case of an increased crash likelihood, the value of the metric must increase, independent of the type of accident.

In order to apply EVT, the value for a crash must exist and be known, so that the SP can be estimated using the return level of the observations (comp. section 2.2.2.2.3). The second requirement is formulated:

RM 2 The value of the metric shall either increase or decrease strictly with increasing crash proximity towards a known numeric value representing a collision.

In order to describe risk and not just accident probability, the severity of the potential accident is required. As severity is dependent on many factors (collision angle, type of accident participant, occupants), it is questionable if it can be estimated in collision-free driving. Nevertheless, it would be necessary in order to derive risk and not accident probability. It is further discussed in subchapter 5.4, which consequences for the extrapolation and the safety approval result from an imperfect severity estimation. The third requirement is formulated as follows:

RM 3 The severity of the potential accident shall be estimated.

For identification of critical scenarios, the most important requirement is the prevention of false-negative detection. As critical scenes are rare, missing those leads to an imperfect test suite despite the availability of data, so they remain known unknowns. To address this, critical scenes must not be missed and the fourth requirement is formulated.

RM 4 The metric shall detect all critical scenes without false-negatives.

These four requirements are mandatory in order to provide ideal metrics for safety validation. It is this discussed in the next section, if those metrics can be verified according to the requirements, or if they are at least falsifiable. Additional requirements that are not mentioned here might apply. For example, the computational effort is relevant, because a high data amount will be analyzed.

5.1.2 Requirements on Data

Q 5 What are the requirements on data for application of microscopic risk metrics?

Besides metrics, the data they are applied on needs to fulfil requirements (RD). From insufficient data, the MiR cannot be derived because the proximity to an accident cannot be computed, if important information is either missing or faulty. Obviously, MiR can be caused by incorrect environment representation and metrics working on this level would describe likelihood of a fault on this level.

The MiR metric on action level (comp. section 1.2.4) requires correct information about surrounding objects. This includes road boundaries or shoulders as well as moving or stationary objects on the road itself. As the metrics are not used for actual decision making, it is sufficient to detect objects later than normal, when the scene is already highly critical. The advantage for metrics on action level is that critical situations on other levels can result in critical situations on action level. For example, a faulty environment perception leads to a late reaction and therefore a high criticality on action level. As the metric will be applied in post-processing, it is sufficient to have the object information a posteriori. The environment could be corrected before applying the metrics as described in Junietz et al.[182]

Two requirements follow addressing road geometry and the existence of objects:

RD 1 The data should contain all information that increases the criticality of the scene concerning road geometry.

Which objects are relevant is dependent on the implementation of the metric. If it is assumed that objects never decrease criticality but can only increase it, it follows:

RD 2 No objects that are relevant for the current driving task and that increase the criticality of the used metric shall be missing in the data.

5.2 Evaluation of Metrics

In the following sections, an evaluation method for metrics is developed and existing metrics analyzed. These research questions are answered:

Q 7 How can the eligibility of MiR metrics for the use cases MaR extrapolation and identification of scenarios be falsified?

Q 8 Do state of the art metrics fulfil all requirements derived in Q 5 and Q 6?

5.2.1 Falsification Strategy

Q 7 asks for a falsification strategy in order to ensure refutability for the application of a metric especially in extrapolation with EVT but also for scenario identification. So it is discussed in the following if and how the requirements derived above can be falsified. First, the requirements are analyzed and a suitable falsification strategy is described. Then, state of the art metrics are evaluated following the developed strategy.

[182] Junietz, P. et al.: Gaining Knowledge on Automated Driving's Safety--The Risk-Free VAAFO Tool (2019).

RM 1 In case of an increased crash likelihood, the value of the metric must increase, independent of the type of accident.

In obvious cases, falsification is achieved by example scenarios. This can be done by changing details in a scenario that clearly must result in a different criticality. This could be an additional object that must be considered in an evasion maneuver or a reduced lane width at high velocity. Especially when comparing different accident types, it is not obvious which criticality is higher. Here, the metric's applicability is only falsifiable in data analysis. Assuming a correct metric, the true accident rate should correspond with the estimated accident likelihood using EVT (comp. Figure 2-5). If this is not the case, it could be caused either by an incorrect computation of the driving requirements or by an incorrect calibration of the driving skill. This falsification strategy leads to a dilemma because the true accident occurrence rate for AD3+ is unknown. For human traffic, the metric can be falsified because the accident statistic is known. For automation, there are no statistical data about accidents before the introduction, so the requirement is not falsifiable a priori. A correct functionality in human traffic is not necessarily transferable to automated traffic. Hence, design guidelines will be established in section 5.3.1 addressing *Q 9*. Following the guidelines will result in a metric that is likely to perform in EVT.

RM 2 The value of the metric shall either increase or decrease strictly with increasing crash proximity towards a known numeric value representing a collision.

If it assumed that *RM 1* holds up, this requirement can be falsified easily as it requires a known limit, when an accident occurs. This could be any arbitrary real value of the metric. This will be covered by the design guidelines as well. Many metrics fulfill the requirements, for example the value zero for TTX metrics.

RM 3 The severity of the potential accident shall be estimated.

This requirement is formulated vaguely because the estimation of severity of a prevented accident can never be precise. Similar to *RM 1*, falsification could be done by analyzing the accident statistics including the severity. If the total number of accidents is correct, but the severity distribution not, the severity is not estimated correctly. However, even if *RM 3* is falsified or simply not addressed, the metric might still be able to estimate the accident rate, just not the severity level and by this not the risk.

RM 4 The metric shall detect all critical scenes without false-negatives.

This requirement is essential if known unknowns shall be collected during real world driving. In some cases, falsification is achieved by a simple counter-example: a critical scene that is not detected by the metric. The requirement can never be fully achieved if the available data does not meet the requirements. If metrics are not falsified by known example scenarios, the falsification in real world traffic is still possible. If accidents happen and the metrics show no increased criticality before the accident, while the accident can still be prevented, the requirement is not fulfilled.

5.2.2 Evaluation in Test Scenarios

Metrics can be falsified according to *RM 4* and *RM 1* using exemplary test scenarios (TS). Exemplary scenarios are derived in this section. While additional scenarios are plausible, those suggested test scenarios are able to falsify the majority of existing MiR metrics as will be shown in section 5.2.3. Before, it needs to be defined which elements of a scenario are viable for the requirements. According to section 1.2.4 and 5.1.1, scenarios that impose increased driving requirements are relevant. Obviously, this includes scenarios with high-required driving dynamics and/or quick and precise reactions. It is assumed that scenarios that include just two objects on collision course are detectable according to *RM 4* with MiR metrics. Therefore, there will be no falsification attempts. This consideration would not cover driving scenes with small distance in lateral direction or small time gap in longitudinal direction. However, in everyday driving, such a behavior is regarded as highly critical especially at high velocities. So, are these scenes critical according to the definition?

Scenes with small time gap are commonly regarded as critical in daily driving because the time margin for successful reaction in case of sudden braking of the front object is small. Nevertheless, there will never be an accident if both objects hold their speed. To derive an accident likelihood in those scenes, a probabilistic prediction of the front object would be required because the scenes becomes critical as soon as the front objects decelerates because an imminent reaction is required. However, the model's parameterization is highly challenging as the action of the front objects is influenced by many factors that might be out of scope of the recorded data and also subject to the objects driving mission (e.g. in case of a nearly missed exit on the motorway). Another possibility would be a posteriori assessment. As long as the front vehicle is not braking in the recordings, the scene is not critical. The downside is that more mileage is required to assess driving behavior with small time gap as dangerous (increased accident probability) because only the scenarios containing decelerating front objects at small time gap are assessed as critical.

Lateral distances when passing objects however, are commonly considered critical because the requirements on the precision of the course angle are high. So even if the objects drive straight and do not approach the predicted ego path, driving requirements rise with increased ego velocity and reduced lateral distance. This should be covered by metrics.

Apart from collision with moving objects, collision with road boundaries should be considered. Hence, a cornering scenario shall be part of the test catalogue (Figure 5-1).

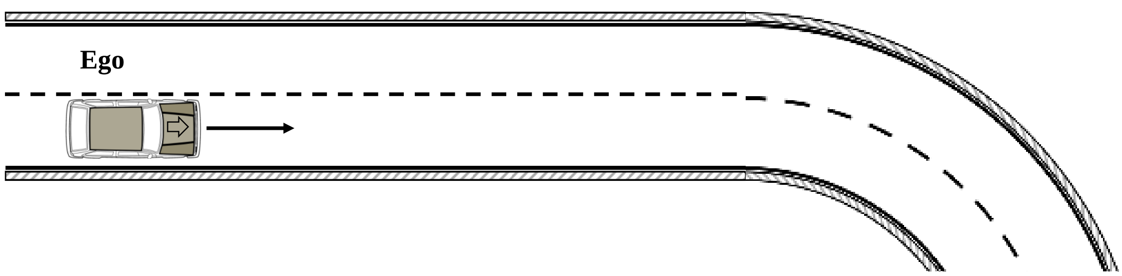

Figure 5-1 TS 1: Cornering Scenario

Depending on the velocity and the curvature of the corner, the test scenario cannot be solved without collision or only with very high lateral acceleration that would be difficult to handle.

Additionally, scenarios in which vehicles are not on collision course, but drive with small lateral distance are tested (see reasoning above). With higher velocity, the accident likelihood increases, because disturbances in the trajectory have to be corrected faster. In case of disturbance in the course angle, the velocity is integrated to lateral deviation. A suitable TS is depicted in Figure 5-2 (bottom). Metrics are falsifiable according to *RM 1*, when two scenario that have obviously different accident likelihood, result in the same assessment by the metric. When two scenarios that are uncritical apart from the lateral distance are compared, the scenario with smaller lane width should be assessed as more critical, at least at higher velocities. At low velocities, both scenarios might be equally uncritical. In Figure 5-2, the bottom scenario is more critical.

Another aspect is the combination of different influences on criticality, e.g. different objects. The two scenarios depicted in Figure 5-3 have different criticality, especially at higher velocities. At higher velocities, an evasion can be conducted later than a braking maneuver. [183] In the top scenario, evasion is impossible due to the second object on the left lane. Hence, the criticality is much higher. Depending on the velocities of the involved objects, the accident might not be preventable anymore at the top scenario, while the evasion is still possible at the bottom.

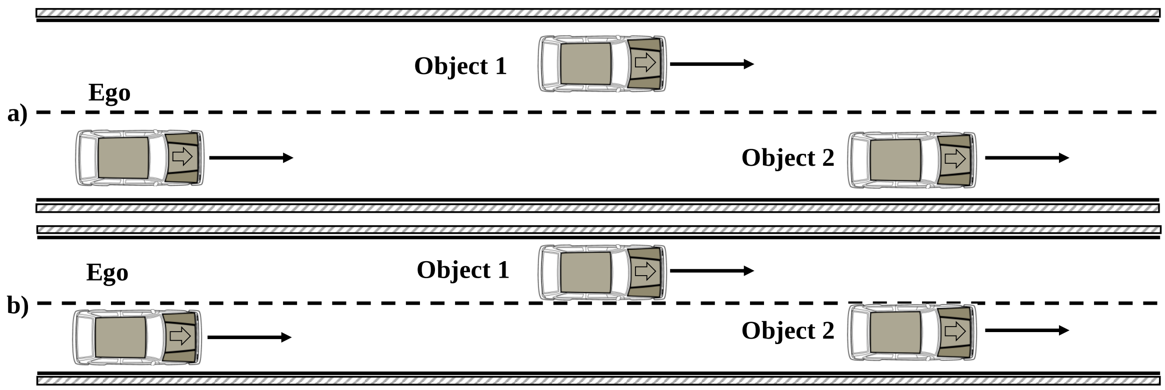

Figure 5-2 TS 2: Scenario with limited lane width

[183] Kühn, M.; Hannawald, L.: Driver Assistance and Road Safety (2016), p. 89.

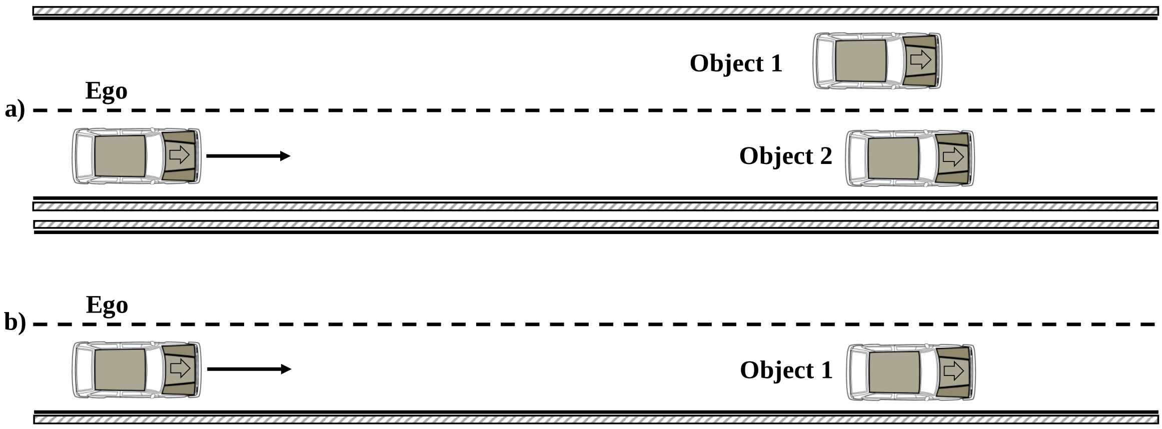

Figure 5-3 TS 3: Traffic jam scenarios with static obstacle and different criticality

5.2.3 Evaluation of Existing Metrics

In this section, metrics that were introduced in section 2.2.2.2.1 are evaluated according to the requirements and the falsification strategy. Most metrics do not consider the severity estimation (*RM 3*). It is only addressed by a factor of the squared velocity because the available kinetic energy is proportional (e.g. CI). However, a true severity estimation is missing. This will be further discussed in section 5.4.

5.2.3.1 Time-to-X Metrics

The metric Time-to-Lane-Crossing (TTLC) describes closeness to accidents because the road surface is left and could pass TS1, though the criticality is not well described. The prevention of leaving the road might be dependent on other factors besides the remaining time. Facing a rectangular corner at a high speed is extremely critical because a deceleration is required. The same TTLC might apply at straight driving with a minor deviation in course angle, which is quickly corrected by a steering maneuver. However, TTLC might be sufficient in highway driving, so it is not rejected due to TS1.

All other TTX metrics focus on object collision. So a combination of metrics is required to address all types of accidents. As the collision occurs at TTX equal zero, *RM 2* that demands a boundary when the collision occurs or is not preventable anymore is fulfilled. Time-to-Brake (TTB) and Time-to-Steer (TTS) require the friction coefficient, at least as a worst case estimation. However, it is obvious that the criticality cannot be described without this information, as the same scene changes its criticality dramatically if the friction is reduced to a minimum. Hence, this is not regarded as a disadvantage.

The metrics Time-to-Collision (TTC) and TTB are only designed to be applied regarding the front object. In this case, they would be falsified according to *RM 4* because sidewise accidents would be neglected. If the metrics would be extended to side-objects (e.g. as in Mages et al.[184]), the metrics would not be falsified because collision with all objects are covered.

However, falsification is achieved according to *RM 1* using TS3 (see Figure 5-3). To pass the TS, the TS3 a) must be assessed with higher criticality than TS3 b), where a relatively uncritical evasion maneuver could be an accident-free solution. Especially with higher velocities, the accident likelihood in the top scenario is higher because a hard braking is the only possible solution. However, TTX metrics only consider object-wise maneuvers. The simultaneous influence of several objects is not part of the metrics. TS 2 also falsifies the metrics as TTX metrics only assess scenes where objects are on collision course. To conclude, TTX metrics are not able to fulfill all requirements and their use for EVT is falsified.

5.2.3.2 Other Time-based Metrics

Time-based metrics apart from TTX metrics are namely Post-Encroachment-Time (PET), Time-Headway (THW) and Worst-Time-to-Collision (WTTC). Again, *RM 2* is fulfilled due to the accident at zero. Different than TTX metrics, PET and THW evaluate scenes where vehicles are not on a collision course. PET is designed for rectangular trajectories scenes, while THW is designed for longitudinal traffic. Both metrics are falsified according to TS3 because they do not assess lateral distance. TS1 is not passed as well because road boundaries are not considered. *RM 4* and *RM 1* are not fulfilled. WTTC is designed to find all critical scenes, so *RM 4* is not falsifiable according to the test cases. TS1 is only covered, if the metric also considers static objects. However, WTTC does not support ordinal classification of accident likelihood. The criticality of vehicles travelling in an adjacent lane is overestimated because worst-case behavior is assumed. TS3 is not passed because the second object and also different types of scenario are not comparable, e.g. when comparing TS2 and TS3 results. TS2 would be assessed with higher criticality even if strong baking is necessary.

5.2.3.3 Acceleration-based Metrics

Acceleration-based metrics describe the necessary acceleration to avoid collision with objects or road boundaries. If the friction coefficient can be estimated, *RM 2* is fulfilled because the driving dynamic limit is known. Sometimes the coefficient is included in the metric, so the transition to an inevitable accident has a fixed limit (see BTN, section 2.2.2.2.1). This would be required, if large-scale data is evaluated that contains different friction coefficients. However, the metrics only work on objects that are on collision course in the same way TTX metrics do, so TS2 is failed and *RM 4* not fulfilled. The metrics also do not consider multiple paths because each object is treated individually. As a result, TS3 is also failed, because the

[184] Mages, M. et al.: Intersection Assistance (2016), p. 1274.

second object does not increase criticality further. Especially on highways, the majority of accidents is in longitudinal traffic. Hence, BTN might lead to acceptable results in EVT (as in Asljung et al.[185,186]). However, some accident types are bound to be missed systematically.

5.2.3.4 Vehicle-movement-based metrics

In NDS- and FOT-Studies, the measured vehicles movements are used together with thresholds as an a posteriori assessment to find critical scenarios. As the environment is not further considered, *RM 2* cannot be fulfilled. It is unknown if the driver decided voluntarily for the maneuver or not, and if not whether maneuver was unreasonable extreme. Near-miss scenarios, in which no extreme movements were necessary are not detectable, so TS2 is failed. The metrics only assess the strength of a reaction but not the necessary precision.

5.2.3.5 Combination of several metrics

All aforementioned metrics are falsified because they do not fulfill one or two requirements according to the three test cases. Especially for scenario identification, several metrics are combined to reduce false-negatives, so *RM 4* can be fulfilled with state of the art metrics. However, when combining metrics and defining a threshold for each metric in order to reduce those, false-positive detection increase (Junietz et al.[187]). Nevertheless, TS3 is still not passed because the second object does not increase the criticality with any of the metrics. *RM 1* cannot be fulfilled.

5.2.3.6 Summary

In the analysis of conventional metrics that were used on large scale data in past studies, TS3 is not passed. A reason is the available data where only one front vehicle is recorded, so TS3 would not be included in the data as two vehicles are required. So the data does not fulfill *RD 2* because relevant objects are not included. Eligible metrics that fulfill the requirements must consider multiple objects together because the criticality from several objects cannot be superposed as only the combination increases the criticality. Metrics from the trajectory planning domain consider all available information. However, they are typically not used to compare different scenarios with each other. The goal is rather to give relative criticality of different trajectory decisions by predicting the objects in a probabilistic way. They are not falsifiable with the derived TS. However, it is questionable whether these metrics describe

[185] Asljung, D. et al.: Comparing Collision Threat Measures using EVT (2016).

[186] Asljung, D. et al.: EVT for Vehicle Level Safety Validation (2017).

[187] Junietz, P. et al.: Metrik zur Bewertung der Kritikalität von Verkehrssituationen und -szenarien (2017).

accident probability as they are not designed for this purpose. In the next subchapter the design of a new metric is described using concepts from trajectory planning metrics.

5.3 Design of a Criticality Metric

Q 9 What are design guidelines for the development of MiR metrics according to the requirements derived in Q 5 and Q 6?

According to the previous subchapter, multiple objects have to be considered simultaneously in order to fulfil the requirements. Metrics that origin from trajectory planning could not be falsified by simple test cases. Nevertheless, it is unlikely that those metrics would perform well according to *RM 1* when applied on data, because they were not designed to compare different traffic scenarios. Instead, they are designed to predict the environment and find the trajectory with the least costs. Their costs typically consider the fulfillment of the driving mission coupled with the risk of a collision. The collision risk is thereby sometimes included by using constraints that prevent the vehicle from getting to close to surrounding objects. As derived previously, applicability of metrics especially for EVT might not be falsifiable before the introduction of AD3+ because of insufficient data. Application with data from current traffic might be sufficient to gain trust in the metric regarding human traffic, but there is no guarantee that the metric will perform equally in AD3+ traffic. In order to find a metric that is likely to perform sufficiently in both, human and AD3+ traffic, design guidelines must be defined that aim towards the fulfillment of the requirements.

In order to address *RM 1*, the metric must provide an orthogonal scale that describes accident likelihood. This adopts the definition of criticality from section 1.2.4, where criticality is defined as proximity to an accident due to high driving requirements concerning reaction time, precision and driving dynamic reserve:

Precision is further refined as precision of course angle under disturbances. Longitudinal precision is neglected, because disturbances must increase the speed of the vehicle in order to increase criticality. As longitudinal distances are typically larger than lateral distances (at least at higher travelling speed) they are neglected in the following. Therefore, the focus is on lateral disturbances.

The driving dynamic reserve can be described by the usage of driving dynamic potential / tire-road friction. These are the driving requirements or the requirements for the driver on action level. However, if the driving skill is unknown, the proximity to an accident cannot be determined. Only the limit, where an accident happens is known, because of limiting driving dynamics conditions. If the driving skill is known, calibration towards accident probability would be possible. If it would be known which percentage of scenes with the respective requirements are solved without accident, a direct link towards accident likelihood could be established.

The required reaction time is the third element of criticality. It is the time until a braking or evasion maneuver must start at the latest, assuming maximal acceleration. Similar to the driving dynamic reserve, it must be calibrated using the assumed driving skill. Here, large differences between human and machine are expected.

It is of outmost importance to establish guidelines for the development of metrics because otherwise their applicability will likely be falsified only after the application on large scale data. As the driving skill is not precisely known, certain parameters of the criticality must be estimated. Whenever this is the case, they shall be justified in detail and arbitrariness shall be analyzed.

5.3.1 Discussion of Concepts for a Metric

In this section, different approaches for a metric are discussed. As falsification without the required data is not possible for advanced metrics, the theory behind the different concepts is analyzed in order to define a suitable approach.

5.3.1.1 Object Prediction

To assess the possible outcome of a scene, the future trajectories of the objects are required. This can be done without using information of the data recordings by different approaches. This is called a priori evaluation. In trajectory planning, probabilistic approaches are the most common that predict the movement of objects by underlying assumptions (e.g. by a statistic model such as Bayesian networks). When evaluating recorded data, it is also possible to use the true trajectories of the objects because the future of the scene in question is known. The possible ego-vehicle trajectories are assessed depending on the known outcome. However, this is only valid for a limited timespan as the objects movement might be influenced by a changing trajectory of the ego-vehicle. A disadvantage of an a posteriori assessment is that behaviors that are subjectively critical might be assessed differently because it did not turn out as critical in the future as it could have. One example is driving with very small time-gap on a motorway. From daily driving, it is well known that small time-gaps at high speeds are critical because the necessary reaction time on a deceleration of the front object is too small.

In an a posteriori assessment, the scene is not critical as long as the front-object does not decelerate. Nevertheless, a behavior like this would be punished by the metric in the long-run because deceleration will occur when enough data is analyzed. In other words, more data are required to assess this behavior with a lesser safety extrapolation.

An a priori assessment of the same scene would require a calibrated behavior model of the object. The scene would be assessed as critical because the probabilistic behavior model predicts deceleration with a (low) likelihood. The advantage is, that it requires less data to assess driving with small time-gap with a lesser safety extrapolation. Nevertheless, it requires data to be parameterized, so the advantage is questionable. As there is only limited data available and further parametrization is necessary for an a priori approach, it is not favorable. Another aspect is that an a posteriori approach is extendable when new data is available. With existing data, a model could be built and the results of both approaches compared. However, this might results in over-assessment of the criticality in scenes with small lateral distance. Without sufficient data, it is assumed that the driver's mental model of the surrounding traffic is more accurate than a model that is just based on the existing data.

To summarize it is stated that the advantage of an a posteriori assessment exceed their disadvantages. The fact that data about the future is available should be used with the benefit of relinquishing behavior models that are a source for errors is parameterization. This concept is obviously not found in trajectory planning because the trajectory is decided upon in real-time. A similar concept is found in silent testing or VAAFO[188,189], where incorrect environment representation is corrected a posteriori. With insufficient data, the safety extrapolation could underestimate the true safety compared to an a priori approach. However, the fact that additional data improves the extrapolation using a posteriori assessment, while all data must be reevaluate with an updated model in a priori assessment. This consideration also suggests using a posteriori assessment, as it is uncertain if enough data is available.

5.3.1.2 Criticality Assessment

From section 5.2.3.6 it is known that in order to fulfil *RM 1*, the criticality assessment shall consider different trajectories and assess the driving requirements in order to find their minimum. Different approaches known from literature that could be used with a posteriori assessment of the scenario are discussed in the following. In section 2.2.2.2.1, multi-object metrics where introduced in the categories geometry-based, sample-based, potential-field and optimization. Here, their potential for the fulfillment of the requirements, especially *RM 1* is discussed.

For simple evasion maneuvers, geometry-based methods find suitable trajectories. Though, there is no information whether the final solution is the least critical or not. The combination of different maneuvers is also not covered. Some scenarios might require parallel or serial combination of steering and braking maneuvers so the methods cannot be applied on every scenario. Hence, they are not eligible for EVT.

Sampling-based methods generate trajectories based on random inputs, usually following a certain probability distribution resulting in a high number of possible outcomes. In the literature, either the number of accident free trajectories defines the criticality[190] or the available

[188] Junietz, P. et al.: Gaining Knowledge on Automated Driving's Safety--The Risk-Free VAAFO Tool (2019).

[189] Wachenfeld, W.; Winner, H.: Virtual Assessment of Automation in Field Operation (2015).

[190] Stumper, D. et al.: Towards Characterization of Driving Situations (2016).

area around the trajectory[191]. These definition might fulfil *RM 1* in <u>some</u> scenarios but not in <u>every</u> as the following example shows: When straight driving is possible but the area or the number of trajectories sampled in the available lane are small, the criticality would be assessed as high. So only similar scenarios are comparable. Single trajectories could also be assessed according to their driving requirements but there might be a relatively uncritical trajectory that is not part of the sample.

Potential-field and optimization methods on the other hand, both result in a single trajectory. Especially, optimization methods could find the trajectory with the least criticality according to a defined cost function. In trajectory planning, the driving mission is always included in the cost function. If the cost function is modified in order to suit the definition of criticality in section 1.2.4, *RM 1* might be fulfilled. Hence, the cost function should contain the necessary acceleration, the reaction time and the necessary precision of the maneuver.

To conclude, out of all discussed approaches, optimization methods are likely most suitable for criticality metrics. However, the definition of the cost function that contains all criticality information into one equation without other information such as the driving mission ins unusual for trajectory optimization. Even though the approach seems promising, the fulfillment of the requirements cannot be proven but is falsifiable after being applied on data.

A remaining challenge is the combination of the different aspects of criticality. In order to define the cost function, they shall be combined into one function with output costs that are minimized in an optimization. This requires the definition of at least two parameters (shape and weighting) per component resulting in six parameters. As derived above, this parametrization and especially arbitrariness in the parametrization needs careful evaluation and justification. Ideally, the parameters are connected to the driving capabilities in a way that derivation of accident likelihood is possible. However, modelling the driving capabilities requires data that might not be available.

According to *RM 2*, the maximum value for the final metric shall represent a scene, in which the accident cannot be prevented. This could be used either by designing the cost function in a way that at each time step, the value range is below or equal to the maximum value, or by using part of the cost function as criticality assessment.

As the range of value for the final metric is arbitrary, it is hereby defined between zero and one so scenes without increased criticality are described by the value zero and scenes, in which the collision is not preventable are defined by the value one.

[191] Rodemerk, C. et al.: Development of a general criticality criterion (2012).

5.3.1.3 Summary

As summary, these design guidelines can be concluded:

- The metric shall consider the reaction time, precision of course angle and driving dynamic reserve.
- The three components shall be combined to a value that describes proximity to an accident using the estimated driving skills.
- Parameterization and weighting of the different factors shall be analyzed and justified.
- If arbitrary parametrization is chosen, the influence on the risk extrapolation shall be analyzed.
- Object trajectories should be assessed a posteriori, if there is no calibrated object prediction model available.
- All possible accident free trajectories should be considered resulting in the trajectory with the least accident likelihood. Trajectory optimization is one feasible approach.

In the following, a metric is developed according to these guidelines and the requirements derived above.

5.3.2 Definition of the Trajectory Criticality Index

In this section, a new metric for criticality assessment is developed based on trajectory optimization. It is called trajectory criticality index (TCI). An overview about the computation of TCI is given in Figure 5-4. Following the design guidelines it is based on trajectory optimization, finding the trajectory with the least driving requirements. The necessary requirements are then weighted according to the driving skill. As the metric will be applied on data of human driven traffic first, human driving skill is the basis for the parameterization. Nevertheless, calibration for AD3+ must be possible as soon as a system is defined and its driving skill known. The trajectory will further be optimized using a model predictive approach requiring a vehicle model (section 5.3.2.1) and a cost function that is dependent on the elements of criticality (reaction C_R, precision C_P and acceleration C_a) according to the design guidelines. This is described in sections 5.3.2.2 to 5.3.2.6. The total TCI is then derived from the costs of the optimal trajectory. The duration of the prediction should be long enough to actively reduce the criticality of the situation. A longer time does not influence the result because the most critical part is covered but increases computational effort.

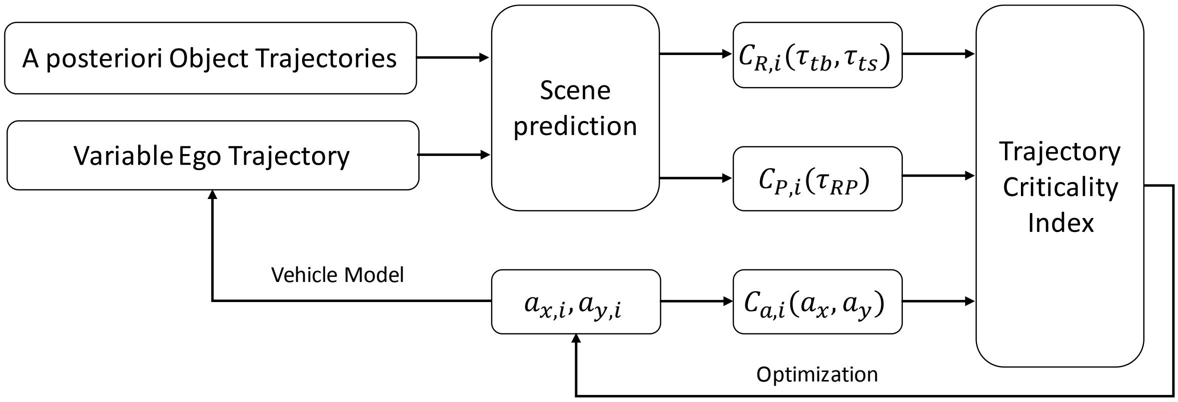

Figure 5-4 Overview of TCI computation

5.3.2.1 Vehicle model

In this section, it is discussed which vehicle model is used for the optimization. Parts of this section where previously published[192]. In addition to this paper, the different options and the influences on the final result are discussed.

The model is the link between the inputs (acceleration in lateral and longitudinal direction or steering angle and longitudinal acceleration) and the trajectory of the ego-vehicle. As input, the lateral acceleration is further used instead of the steering angle, because the transformation from steering angle to acceleration might vary for different vehicles. For the computation of criticality the position (including the yaw angle) in world coordinates (index w) and the velocity of the vehicle in natural coordinates (index v) are required. So the problem depends on the current ego velocity in natural vehicle coordinates ${}_{\mathrm{v}}v$, the yaw angle ${}_{w}\psi_c$, and the accelerations ${}_{\mathrm{v}}a_x$, ${}_{\mathrm{v}}a_y$ in natural vehicle coordinates. Therefore, a model with at least three degrees of freedom (DoF) is required. Additionally, the position in world coordinates is needed because the position of the objects is not updated into vehicle coordinates once the prediction horizon is initialized.

A state-space model with state vector $\underline{x} = \left[{}_{\mathrm{w}}x \;\; {}_{\mathrm{w}}y \;\; {}_{\mathrm{v}}v \;\; {}_{\mathrm{v}}\psi_c\right]^{\mathrm{T}}$ and input vector $\underline{u} = \left[{}_{\mathrm{v}}a_x \;\; {}_{\mathrm{v}}a_y\right]^{\mathrm{T}}$ is defined. For most MPC applications, a linear single-track-model is used. The simplest approach here would be a point-mass model with course angle as additional DoF. Compared to a more complex single-track or two-track model, it is more agile in lateral direction than a real car, because lateral acceleration directly results in a change in course angle because tire forces are not modeled and the wheelbase is infinitively small. As a result, the criticality computed by the metric is smaller in highly critical scenes than when using a more advanced model. The maximal lateral acceleration might be lower in reality as well as the model ignores load shift during cornering and self-steering effects. So the criticality in highly critical scenes is underestimated, which could be prevented using a more advanced

[192] Junietz, P. et al.: Criticality Metric for the Safety Validation of Automated Driving (2018).

model. As the metric should be universally applicable and the driving dynamically relevant parameters might vary with different vehicle models, a simple model is used here. As the model does not include a slip angle, yaw angle and course angle are identical. For better readability the letter ${}_e\psi_c$ is used instead of ${}_e\nu_c$.

The non-linear vehicle model is described by:

$$\begin{bmatrix} {}_\mathrm{w}\dot{x} \\ {}_\mathrm{w}\dot{y} \\ {}_\mathrm{v}\dot{v} \\ {}_\mathrm{w}\dot{\psi}_c \end{bmatrix} = \begin{bmatrix} {}_\mathrm{v}v \cdot \cos\left({}_\mathrm{w}\psi_c\right) \\ {}_\mathrm{v}v \cdot \sin\left({}_\mathrm{w}\psi_c\right) \\ {}_\mathrm{v}a_x \\ \frac{{}_\mathrm{v}a_y}{{}_\mathrm{v}v} \end{bmatrix} \tag{5.1}$$

Assuming small course angle changes and small changes in velocity allows linearization of the model.

Initializing the model at course angle zero and using the previous velocity ${}_e v_{\text{-}1}$ in each time-step results to:

$$\begin{bmatrix} {}_\mathrm{w}\dot{x} \\ {}_w\dot{y} \\ {}_\mathrm{v}\dot{v} \\ {}_\mathrm{w}\dot{\psi} \end{bmatrix} \approx \begin{bmatrix} {}_\mathrm{v}v_{\text{-}1} \\ {}_\mathrm{v}v_{\text{-}1} \cdot {}_\mathrm{w}\psi_c \\ {}_\mathrm{v}a_x \\ \frac{{}_\mathrm{v}a_y}{{}_\mathrm{v}v_{\text{-}1}} \end{bmatrix} \tag{5.2}$$

resulting in the system matrices:

$$\underline{A} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & {}_\mathrm{v}v_{\text{-}1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \underline{B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1/\ {}_\mathrm{v}v_{-1} \end{bmatrix} \tag{5.3}$$

and the system:

$$\underline{\dot{\mathrm{x}}} = \underline{A} \cdot \underline{x} + \underline{B} \cdot \underline{u} \tag{5.4}$$

Note that instead of a constant velocity over all time steps, the velocity of the previous time step is used. This results in more accurate but more expensive computation (comp. Yi et al.[193]) due to time-variant system matrices. As computational effort is not the main issue, the velocity of the previous step is used. Especially in critical scenes, braking is often necessary and the velocity decreases. From equation (5.2) the effect in combined maneuvers with lateral acceleration and braking is derived. Changes in course angle are smaller in the model as the previous velocity ${}_\mathrm{v}v_{\text{-}1}$ is used. For the computation of the lateral velocity ${}_W\dot{y}$, this influence is compensated partly because the lower magnitude of the course angle and the higher

[193] Yi, B. et al.: Real time integrated vehicle dynamics control and trajectory planning (2016).

velocity are multiplied. It is further assumed that the deviation is neglectable using typical stepsizes of about 0.1 s. However, the linearization might influence the performance in lane-change scenarios, so a comparison of results according to (5.1) and (5.2) will be conducted in subchapter 6.2.

5.3.2.2 Cost Function and Criticality – Preliminary Considerations

As argued in section 5.3.1.2, the minimization of a cost function should result in a trajectory with minimal criticality including the aspect of reaction strength, precision and reaction time. The three aspects should be calibrated against each other according to the available driving skill. As the driving skill can only be roughly estimated, the three components will be normalized to one to allow comparison. Requirements which result in the minimum value zero and the maximum value one, will be discussed in the following sections. The criticality assessment must fulfill the requirements of section 5.1.1 and does not necessarily have to be the minimized value of the cost function. For example, the cost function typically minimizes the sum of all time-steps in a trajectory, because the whole trajectory should be in focus of the optimization. For the criticality assessment on the other hand the maximum value is of importance because otherwise a very critical moment could be compensated by several uncritical following steps. What options remain in design of the cost function and the criticality will be discussed in section 5.3.2.6.

In conclusion, the following steps are required:

- Definition of the three elements including:
    - Highest Value
    - Value without criticality
    - Normalization and progression between lowest and highest value
- Design of the cost function and the criticality including:
    - Weighting factors
    - Algebraic combination of the elements

5.3.2.3 Acceleration

As mentioned before, the required accelerations of the optimal trajectory have an influence on criticality. Not only because the driver has to act, but also because the acceleration is required in this scene and a smaller reaction would either be insufficient or would require increased requirements in the other domains (precision and reaction time). A normalization of the acceleration can be done by dividing the acceleration by the available tire-road friction μ (that must be known or estimated) and the gravitational constant g. Neglecting aerodynamic effects, the maximum achievable acceleration is $\mu \cdot g$ according to Kamm's circle (Figure 5-5a). However, due to the non-linear characteristic of tire-road friction, handling with lateral accelerations of more than 0.4 g is challenging and not mastered by most drivers. Assuming that automated functions will work best outside of combined maneuvers with high absolute accelerations, especially because the available friction coefficient can only be estimated. Hence, instead of a friction circle, a rhombus (Figure 5-5b) or a cross/star (Figure 5-5c) could be used as a representation for the available friction coefficient in dependency of the acceleration in longitudinal and lateral direction. The driving dynamics of the vehicle could further influence the achievable accelerations especially in lateral direction, e.g. due to self-steering effect or in longitudinal direction for example due to limited available engine torque or two-wheel instead of four-wheel drive.

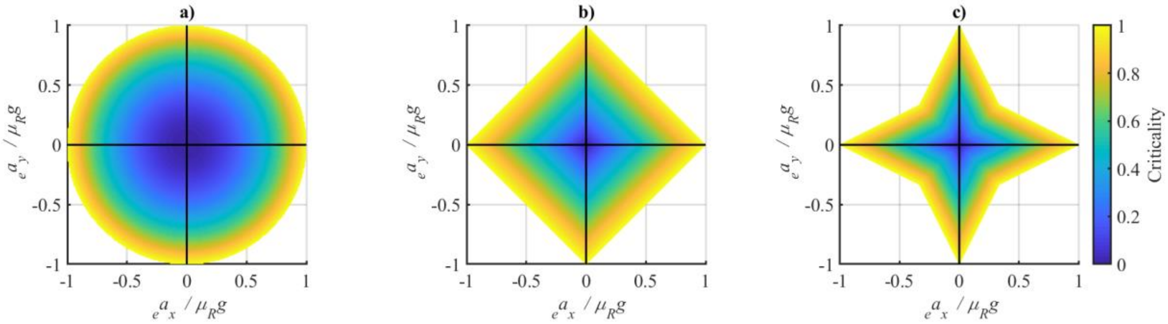

Figure 5-5 Different options to assess the criticality of the required acceleration

The formulas to calculate the criticality component acceleration C_a corresponding to Figure 5-5 are shown in equations (5.5) to (5.7). Different approaches or combinations are possible as well, but would require in depth studying of the driving skill of the human driver or the automation.

$$C_{a,\mathrm{a}} = \frac{{}_e a_x^2 + {}_e a_y^2}{(\mu \cdot g)^2} \tag{5.5}$$

$$C_{a,\mathrm{b}} = \frac{|{}_e a_x| + |{}_e a_y|}{\mu \cdot g} \tag{5.6}$$

$$C_{a,\mathrm{c}} = \frac{|{}_e a_x| + |{}_e a_y| + \min(|{}_e a_x|, |{}_e a_y|)}{\mu \cdot g} \tag{5.7}$$

To not overestimate the criticality, especially in critical scenes approach a) is favorable as long as the metric is used universally and not for a specific vehicle or automated system. It will be used in the following chapters. Another approach is only advisable, if it is certain that the vehicle or the driver cannot achieve parts of the circle.

5.3.2.4 Reaction Time

The implementation of the necessary reaction time is manifold because the reaction might be in longitudinal (TTB) or lateral direction (TTS). However, for evasion maneuvers, free space must be available on the neighboring lane so there are two possibilities to calculate the reaction time:

a) Assume only braking for the computation of TTB with the downside of overestimating the criticality in cases where evasion is possible.

b) Determine whether there is free space available and use the maximum of TTB and TTS with the downside that analytical computation is only possible for simple evasion maneuvers (e.g. based on sigmoid trajectories; comp. section 5.2.3.1).

As option b) results in uncertainty, because an analytical result of the free-space check is not trustworthy, option a) will be used later. This results in an increased criticality component reaction time based on TTB alone. To compensate this, optimizations resulting in an optimal solution of an evasion maneuver are detected and the ego trajectory is assed a priori to detect evasion maneuvers. If an evasion is detected, the optimization problem is reevaluated using option b) because the free space is available. The variable τ_{TR} in equation (5.15) is used for TTR meaning the respective value of TTB for the initial optimization and the maximum of TTB and TTS for the reevaluation.

As there is no further reaction time needed to use the current acceleration $a_{x,0}$, TTB and TTS are calculated assuming constant acceleration until the reaction. The geometric relations and variables are depicted in Figure 5-6. The formula for TTB is derived from the positions of ego and object vehicle assuming the braking duration t_b and the brake reaction time τ_{tB} and using the travelled distances s of both vehicles (comp. Hillenbrand et al.[194]).

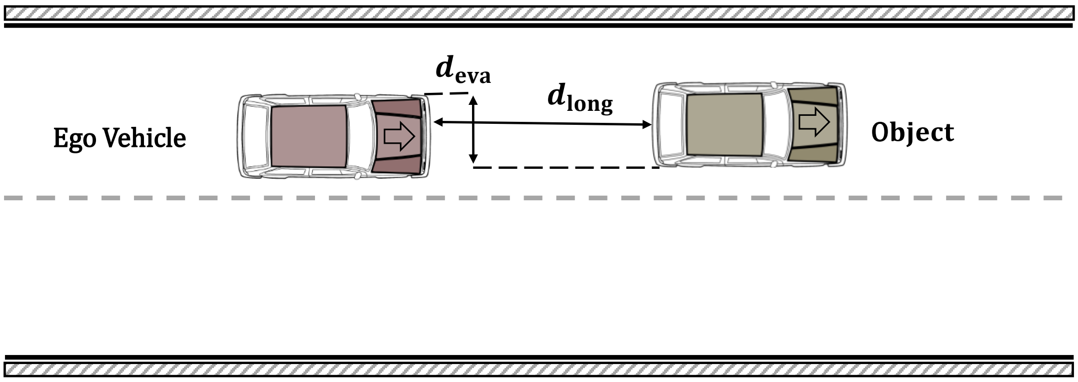

Figure 5-6 Geometric relations of an approaching scenario

The ego vehicle travels constantly with the initial speed and the current longitudinal acceleration until the reaction. After that, the ego vehicle decelerates with maximum deceleration. For the object vehicle, the velocity is assumed to be constant. Equations for the travelled distances s and the velocities v are combined with the condition that no relative velocity exists when the initial gap d_{long} is zero, eliminating the brake duration and resulting in the quadratic equation (5.11).

$$s_{\text{ego}} = v_{\text{ego},0}(t_b + \tau_{\text{tB}}) + \frac{1}{2}a_{x,0}\tau_{\text{tB}}^2 - \frac{1}{2}\mu_{\max}\, g t_b^2 \tag{5.8}$$

$$s_{\text{obj}} = v_{\text{obj}}(t_b + \tau_{\text{tB}}) \tag{5.9}$$

[194] Hillenbrand, J. et al.: Situation assessment algorithm (2005).

$$v_{\text{obj}} = v_{\text{ego}} + a_{x,0}\tau_{\text{tB}} - \mu_{\max} g t_b \tag{5.10}$$

with $s_{\text{ego}} = s_{\text{obj}} + d_{\text{long}}$ and equation (5.10):

$$\frac{1}{2}\left(a_{x,0} - \frac{a_{x,0}^2}{\mu_{\max} g}\right)\tau_{\text{tB}}^2 - v_{\text{rel}}\tau_{\text{tB}} - d_{\text{long}} + \frac{v_{\text{rel}}^2}{2\mu_{\max} g} = 0 \tag{5.11}$$

Solving equation (5.11) for τ_{tB} results in two real solutions with the smaller one being the actual TTB and the larger representing overtaking (or driving through) and decelerating backwards.

For TTS similar considerations apply, but now the last longitudinal evasion distance $d_{\text{long,eva}}$ depending on the necessary lateral evasion distance Δd_y and evasion time t_{eva}. When the lateral distance between the two vehicles is zero, the evasion must be completed (equation (5.13)).

$$t_{\text{eva}} = \sqrt{2\frac{\Delta d_y}{\mu_{\max} g}} \quad \text{and} \quad d_{\text{long,eva}} = (-v_{\text{rel}} + a_{x,0}\tau_{\text{tS}})t_{\text{eva}} \tag{5.12}$$

$$d_{\text{long,eva}} = d_{\text{long}} + v_{\text{rel}}\tau_{\text{tS}} + \frac{1}{2}a_{x,0}\tau_{\text{tS}}^2 \tag{5.13}$$

Combining (5.12) and (5.13) results into:

$$\frac{1}{2}a_{x,0}\tau_{\text{tS}}^2 + \left(v_{\text{rel}} - a_{x,0}\sqrt{2\frac{\Delta d_y}{\mu_{\max} g}}\right)\tau_{\text{tS}} + d_{\text{long}} + v_{\text{rel}}\sqrt{2\frac{\Delta d_y}{\mu_{\max} g}} = 0 \tag{5.14}$$

Again, solving the quadratic equation results in the actual TTS as the minimum of the two solutions due to reasons identical to the ones above.

For the normalization of the reaction time, different approaches are plausible. The upper limit is when no reaction time remains. In conclusion, the questions, which need to be addressed, are: What is the lower limit and how should the criticality increase with decreasing reaction time? Studies about brake reaction time of human driven vehicles are available and can be used as a basis for those considerations and for the calibration of the driving skill. Habenicht[195] analyzed studies focusing on reaction time. Burckhardt[196] found that the median of reaction times in his experiments was 0.86 s. Döhler and Nitsche found that that the 2% percentile was between 0.3 and 0.5 s and the 98% percentile was between 1.0 and 1.1 s.[197] In a more recent study for the European Association for Accident Research and Analysis, Hugemann[198] found that a Gamma-distribution would be the best fit for the data although more detailed experiments would be necessary to come to a profound fit.

---

[195] Habenicht, S.: Entwicklung eines manöverbasierten Fahrstreifenwechselassistenten (2012), pp. 35–37.

[196] Burckhardt, M.: Reaktionszeiten bei Notbremsvorgängen (1985).

[197] Döhler, H.; NITSCHE, K.: New mathematical findings on reaction times (2008).

[198] Hugemann, W.: Driver reaction times in road traffic (2002).

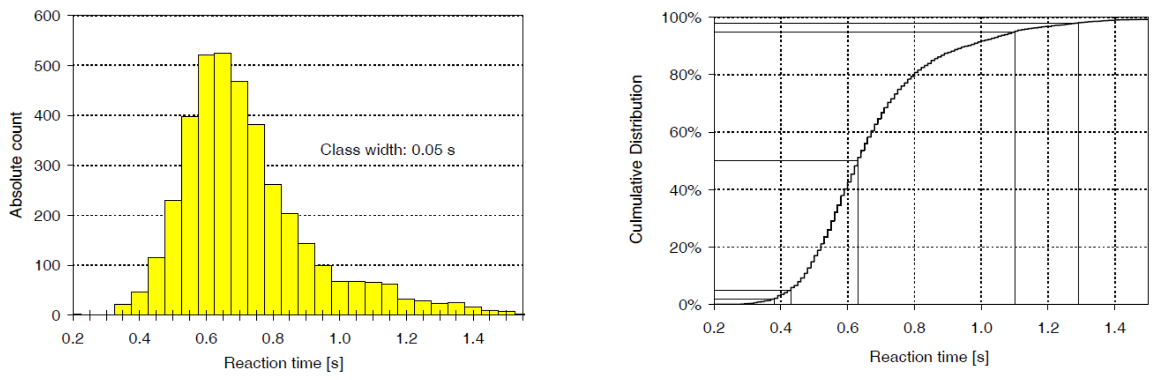

Figure 5-7 Data for brake reaction time by Hugemann[199a] based on Burckhardt[200]

A possible approach to normalize the reaction time into a criticality component reaction C_R would be to fit a distribution function (e.g. Gamma-distribution as suggested above) to reaction time studies. Its cumulative distribution function Φ_τ could then be approximated using the following formula that determines the likelihood that the reaction time of the driver τ_R is bigger than the required brake reaction time τ_tR.

$$C_{R,\Phi}(\tau_{tR}) = P(\tau_{tR} < \tau_R) = 1 - \Phi_\tau(\tau_\mathrm{tR}) \tag{5.15}$$

However, the available data is insufficient. Hugemann[199b] points out that in every day driving, reaction times of more than 1.5 s are common even when the driver is not distracted. For AD3+ reaction times, it is broadly assumed that the reaction times are faster, but object recognition and motion planning takes time as well and might be dependent on the scene's complexity. Nevertheless, the distribution and dependencies with the scene are unknown and do not allow calibration of the driving skill. It is only certain that criticality decreases from a reaction time of zero and that the lower limit should be somewhere about 2 s of remaining reaction time for human traffic. Consequently, a linear decrease is used in the following, as more detailed data would be required to justify the selection of a distribution function. Hence, the following formula is used further on:

$$C_{R,\mathrm{lin}}(\tau_{TR}) = 1 - \frac{\tau_{TR}}{2\ \mathrm{s}} \quad \text{for} \quad 0 < \tau_{TR} < 2\ \mathrm{s} \tag{5.16}$$

Linear representation of the cumulative distribution function according to equation (5.16) is probably a worse fit than a Gamma-distribution fit according to equation (5.15) regarding overall accuracy of the approximation. However, it underestimates criticality in the very critical region, which has a less severe influence than overestimating due to the low frequency of highly critical events. Additionally, events of lower criticality are overestimated,

[199] Hugemann, W.: Driver reaction times in road traffic (2002), p. 3.a: p. 3; b: p. 11

[200] Burckhardt, M.: Reaktionszeiten bei Notbremsvorgängen (1985).

so it also suits scenario identification purposes, where no critical scene shall be missed. In other words, there is a borderline $\tau_{tR,\mathrm{lim}}$ where:

$$\begin{array}{lcl} C_{R,\mathrm{lin}}(\tau_{tR}) > C_{R,\Phi}(\tau_{tR}) & \text{for} & 0 < \tau_{tR} < \tau_{tR,\mathrm{lim}} \\ C_{R,\mathrm{lin}}(\tau_{tR}) \leq C_{R,\Phi}(\tau_{tR}) & & \tau_{tR,\mathrm{lim}} \leq \tau_{tR} < \infty \end{array} \tag{5.17}$$

Above this borderline, the criticality is overestimated and below this borderline, criticality is underestimated.

5.3.2.5 Precision

The criticality component precision C_P is not as straightforward as the previous two. Hence, it is first analyzed using exemplary scenarios.

A high precision of a trajectory is required when the margins to other objects are small and minor disturbances to the trajectory would result in a collision. In theory, precision is necessary for speed and for the course angle as both is controlled by the driver. A precision in speed is especially required when driving behind another vehicle at the same speed with very small distance (where the speed must not be increased) or when cutting in between in front of a vehicles (where the speed must not be decreased). The two scenarios are depicted in Figure 5-8. Both cases are relatively easy to handle as long as the other vehicle does not change its speed because disturbances in longitudinal direction (wind, road unevenness) have only minor influence and it is relatively easy to increase the margin by decreasing the own speed in the “following” scenario and increasing the own speed at the “cut-in” scenario.

Is in conclusion the “following” scenario uncritical and if so, why is there a mandatory safety distance that must not be undercut? Obviously, the scenario becomes critical as soon as the preceding vehicle reduces its speed. However, this has nothing to do with precision but with reaction time. As soon as the vehicle brakes in the recorded data, C_R increases. If the vehicle does not brake, the scenario is uncritical.

Is therefore a behavior that favors small time-headway not considered critical? There are two approaches how a small time-headway influences the overall result and the frequency of critical scenes. To punish this driving behavior directly, a probabilistic trajectory prediction for the preceding vehicle could be used. The likelihood of braking with a certain deceleration would then be used to assess the criticality and short reaction time would lead to an increased criticality. The frequency of such a scene would then be weighted with the likelihood of such an event. However, this process would be error prone due to the calibration of the probabilistic model. Instead, an a posteriori assessment is used as discussed in section 5.3.1.1. In the long run, a systematic driving behavior that often drives with small time-headway will experience more scenes where the brake reaction time is low because of a decelerating preceding vehicle, but sufficient data is required to come to a statistically profound estimation of the average frequency. As this is already covered by C_R, there is no additional longitudinal component for C_P.

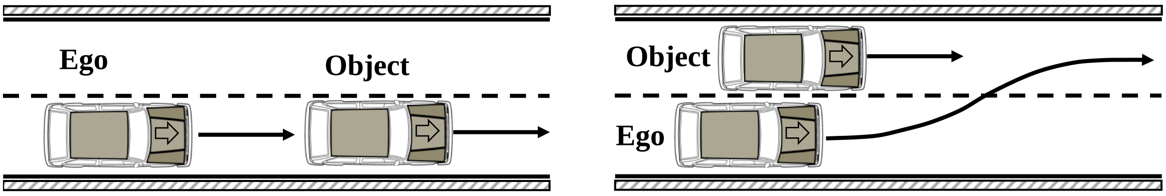

Figure 5-8 “Following” and “cut-in” scenario

As known from personal experience in every day driving, criticality from lateral margin is obviously dependent on the lateral distance and the velocity. It is obviously less demanding to drive through a gap at a lower speed compared to the same gap at a higher speed. So why is the velocity important to consider? It is typically not due to lack of steering ability or perception ability that causes the driver being unable to maneuver the vehicle through the gap. Instead, disturbances in the course angle due to side-wind, road unevenness, or non-precise control of the vehicle’s system require quick corrective reactions. Additionally, the speed of the reaction might increase overshoot especially in steering. Incorrect estimation of the available lateral distance further increases control errors. The driver must be ready to compensate these disturbances at all time when passing through a gap. Different from C_R, reaction time cannot be derived directly from data because disturbances are usually not recorded and there is also no data available to allow parameterization of a statistical model. Normalization of the known available reaction time would be ideal, that is however unknown for a human driver. When the lateral distance is zero, C_P should be maximal ($C_P = 1$). Thus, are there other scenarios that should be assessed as equally critical? The reaction time for the component precision τ_{RP} is the lateral distance divided by the lateral velocity that equals the velocity multiplied with the disturbance in course angle $\Delta\psi$. Ideally, the disturbance would be approximated with a statistical model as described above for preceding objects. In this case, an a posteriori assessment is impossible as well, because data about disturbance in course angle is usually not recorded. From personal experiences it can be derived, that driving on a highway with lane width of 3.5 m[201] at the recommended speed of 130 km/h is not critical and starts to be more challenging with increased speed. As no additional data is available the corresponding reaction time is arbitrarily defined as four seconds in this scene, while four seconds is the reoccurring limit to filter uncritical scenes (comp. section 2.2.4.2 and appendix A). This is equal to an assumed disturbance in course angle of 1.4° (0.0244 rad). It has to be pointed out that this parameterization is arbitrary and other considerations might come to other results. When the ego vehicle is approaching an object with a lateral distance as depicted in Figure 5-9, the reaction time due to disturbances is obviously higher than when the vehicle is beside the ego.

[201] Forschungsgesellschaft für Straßen und Verkehrswesen: Anlage von Autobahnen (2009).

In this case, the time until catching up must be added to the reaction time because the driver has more time to compensate the disturbance. It follows:

$$\tau_{RP}(d_{\mathrm{lat}}, v) = -\frac{d_{\mathrm{long}}}{v_{\mathrm{rel}}} + \frac{d_{\mathrm{lat}}}{v \cdot \sin(\Delta\psi)} \approx -\frac{d_{\mathrm{long}}}{v_{\mathrm{rel}}} + \frac{d_{\mathrm{lat}}}{v \cdot \Delta\psi} \tag{5.18}$$

Assuming that the occurring disturbance in course angle is small and not dependent on the velocity nor on the lateral distance, the reaction time is proportional to the lateral distance and anti-proportional to the velocity and the disturbance.

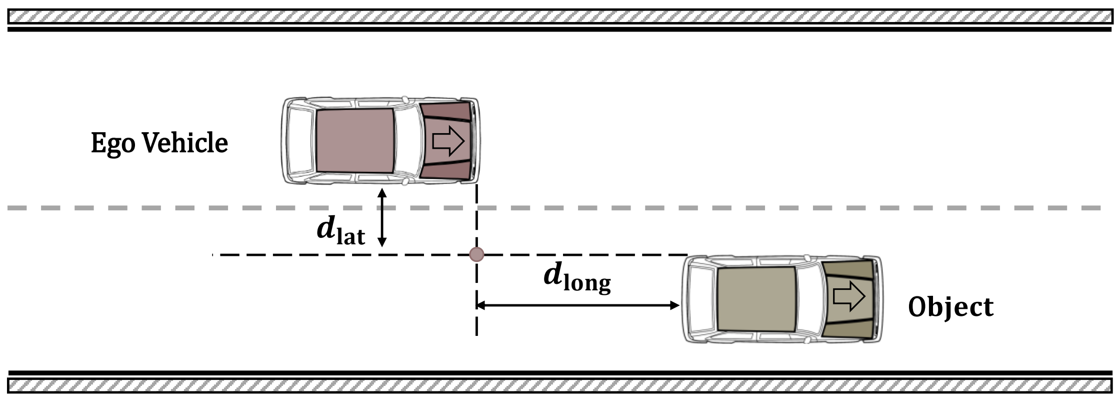

Figure 5-9 Approaching with lateral distance

The effect on the final result will be analyzed in subchapter 6.2.4. In accordance to equation (5.16), the following normalization will be used. As the parametrization is arbitrary and should not result in an overestimating in the region of higher criticality a weighting factor w_{P} is introduced that will be analyzed in the sensitivity analysis in section 6.2.4. Firstly, a weighting factor of 1/10 is used that is increased in the sensitivity analyses.

$$C_P(\tau_{RP}) = 1 - w_{\mathrm{P}} \frac{\tau_{RP}}{2\ \mathrm{s}} \quad \text{for} \quad 0 < \tau_{RP} < \frac{2\ \mathrm{s}}{w_{\mathrm{P}}} \tag{5.19}$$

5.3.2.6 Cost Function

In order to find the trajectory with the minimal criticality, a cost function must be formulated that contains the three elements: acceleration, reaction and precision. A straightforward approach would be the summation of the three elements with or without weighting factors. For the definition of the final criticality index there are basically two options:

a) The final value of the cost function is the criticality index
b) The criticality index is an algebraic combination of one or more components

For both options, a value of one of each component means that an accident cannot be prevented, so the criticality should be one as well. As reaction and precision are both time-based according to equation (5.16) and (5.19) and both describe a reaction time, the maximum value of both components could be used (as equivalent to the minimum reaction time). For the combination of acceleration and reaction time based components, the maximum function is insufficient because assuming constant required acceleration. The criticality should change with decreasing reaction time, also when the criticality from acceleration is dominant. Therefore, the combination function shall always increase when one of the inputs increases. When at least one input is already one, the combination function shall be one. Similar as for equation (5.16), different equations fulfill these requirements. Without additional information, a linear approach is favored. If it serves a purpose (e.g. an increased sensitivity in highly critical scenes), the combination function could be adapted in future work. The following combination formula for total criticality C_{tot} (also depicted in Figure 5-10) fulfills all requirements:

$$C_{\text{tot}}(C_a, C_\tau) = C_\tau + C_a - C_a C_\tau \quad \text{with } C_\tau = \max(C_{\text{P}}, C_{\text{R}}) \tag{5.20}$$

As discussed previously, criticality should be the maximum of the predicted driving requirements. With i as the current discrete step during the prediction horizon of N steps, the final criticality, the trajectory criticality index I_{TC} results:

$$I_{\text{TC}} = \max(C_{\text{tot},1}, \dots, C_{\text{tot},N}) \tag{5.21}$$

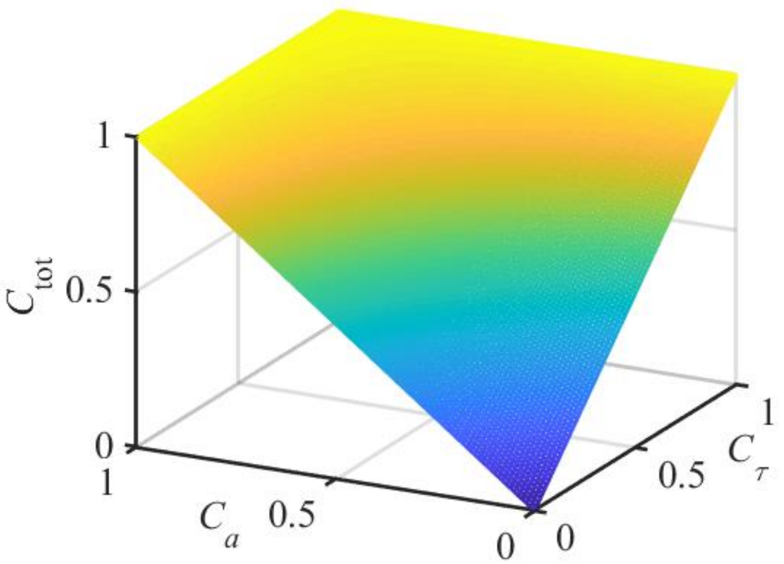

Figure 5-10 Combination function for criticality

Following option b), equation (5.21) is used as the cost function, which is minimized to find the trajectory with the minimal criticality. As only the maximal criticality of all N steps is relevant for the optimization, the resulting trajectories look arbitrary because most components of the input vector have no influence on the costs. As only the maximal criticality is of interest, this is not problematic.

The cost function J concludes to:

$$J = \max(C_{\text{tot},1}, \dots, C_{\text{tot},N}) \tag{5.22}$$

5.3.2.7 Programming Workflow and Parametrization

The cost function J is a nonlinear multivariate function. Additionally, the problem is not only nonlinear but can also be non-convex. Convexity is trivially excluded with an example scenario where a single object drives in front of the ego-vehicle on a three-lane road, where there is an additional vehicle on the right lane. When the first object vehicle is decelerating, staying in the center lane is most critical. Two local minima exist for passing on the left or right with passing left as the global minimum because the second object is on the right lane, so additional braking is required. Hence, the problem requires global optimization. Multiple initial trajectories are used to cover different emergency maneuvers including evasion to both sides, braking only and combined maneuvers. The used initial trajectories corresponding to the initial input accelerations are depicted in Figure 5-11. They are described in Appendix B. Additionally, the a posteriori trajectory, meaning the trajectory that the driver or automation has decided for, is used as additional initial trajectory. Finally, the solution, which converges to the smallest costs is selected.

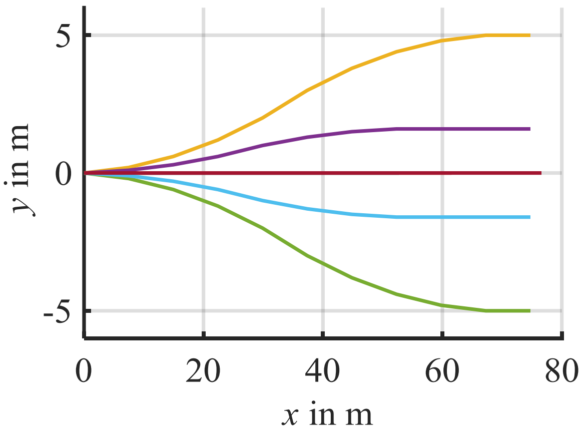

Figure 5-11 Initial trajectories for global minimization

Parameterization of the step size and the prediction horizon is highly relevant for the result as well. Both are also relevant for computational effort. The prediction horizon obviously influences the resulting TCI because a more critical scene might follow, if N is too small. However, this is compensated by analyzing a whole scenario and not just selected scenes. As the vehicle travels further, the optimization is started again with the short prediction horizon and assessed accordingly. A disadvantage of a longer prediction time is that the environment cannot change its behavior as a reaction to the optimized trajectory, while in reality, each deviation from the original recorded trajectory might have interdependencies with other objects. So the prediction horizon should be as short as possible. So what is its minimum duration? The prediction must be far enough into the future to decide about the criticality of the scene. This cannot be shorter than the typical reaction time because constant driving has to be assumed for this duration after an event that causes increasing criticality (e.g. a preceding vehicle at full brake). Also, it shall be long enough to execute a maneuver that reduces the criticality afterwards, e.g. braking or lane-change. While braking reduces criticality instantly due to equation (5.11), a lane change requires time to end the maneuver and criticality is reduced only after the whole maneuver. The minimum duration of lane changes on highways was found to be at least 1 s[202], so the prediction time should not be shorter. To allow the completion of slower lane changes, a duration of 2 s will be used further on.

The choice of the step size is mainly an issue of computational effort, so it should not be unnecessarily small because with each additional step during the prediction horizon, two additional input variables are part of the optimization problem. The same sampling frequency is also required for the data. However, the step size should be small enough, to cover the smallest reaction time that occurs as combination of the driver reaction time and the system delay (e.g. for braking). While the human driver typically does not react faster than 0.5 s[203], the reaction time of future systems is unknown. When hardware and software reaction times are summed up, the reaction time is estimated in the region of 0.2 s, which is used as step size in the following.

To avoid optimization outside of driving dynamic potential, constraints for the input arguments are required for limiting the maximum acceleration. The input acceleration is limited by approximating Kamm's circle with linear limitations as described by Yi et al.[204]. This allows for linear constraints for the optimization problem. Kamm's circle is over-approximated, so acceleration values up to the costs of one according to equation (5.5) are allowed. Additional constraints to avoid collisions as in Ulbrich et al.[205] are not required, as a collision check is included because the lateral and longitudinal distances in equations (5.11) and (5.13) become zero in case of a collision during the prediction horizon.

To reduce computational effort further, the optimization is only run when the initial input vector leads to a TCI that is larger than a threshold. Most scenes on the highway are relatively uncritical and constant driving or slight braking is adequate. Subsequent scenes in one scenario often have a similar criticality and similar TCI is expected. However, they should not be counted as separate occurring critical scenes. As an example: if once in 100 km of driving, a scenario occurs where TCI is 0.3 at the beginning and 0.5 in the next time step, this should only be evaluated as a scene of TCI 0.5 once in 100 km. The separation could be done by trigger events, e.g. driving without high acceleration or criticality for a defined timespan, or by using a defined length for each scenario and only using its maximum criticality. The application on the highD Dataset[206] in the next chapter, separate scenarios of 420 m travelling distance are available, so the maximum TCI of each of those scenarios is used further. The

[202] Toledo, T.; Zohar, D.: Modeling duration of lane changes (2007), p. 71.

[203] Wolff, C.: Grundlegendes zum Bremsvorgang (2017), p. 19.

[204] Yi, B. et al.: Real time integrated vehicle dynamics control and trajectory planning (2016).

[205] Ulbrich, S.; Maurer, M.: Towards tactical lane change behavior planning for automated vehicles (2015).

[206] Krajewski, R. et al.: The highd dataset (2018).

flowchart of the used programming flowchart is depicted in Figure 5-12. The value of the threshold $I_{TC,thr}$ is discussed in the next chapter.

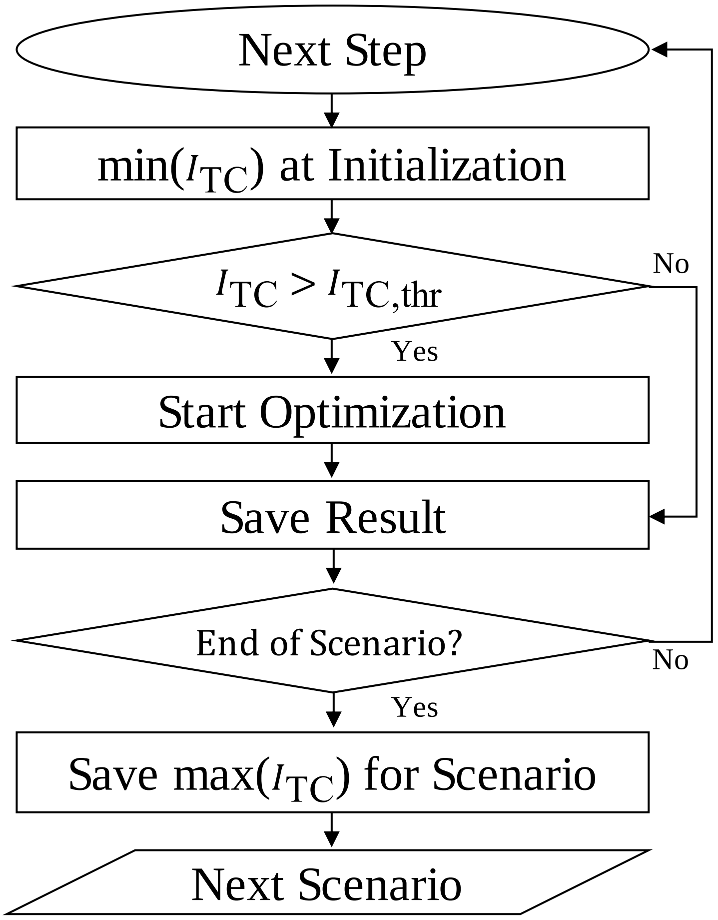

Figure 5-12 Flowchart for the application

5.3.2.8 Adaption to Curved Roads

The equations developed in this section are designed for straight road geometry, as it is common on highways. Curved roads influence the criticality because a part of the available tire-road-friction is used for holding the lane. Additionally, the reaction time is calculated under the assumption that the course angle is not changed before the reaction. However, in cornering scenes, the course angle is changed continuously. To address this issue, curvilinear coordinates are introduced and the data are transformed accordingly. The ${}_{c}x$-axis follows the road curvature with a perpendicular ${}_{c}y$-axis.

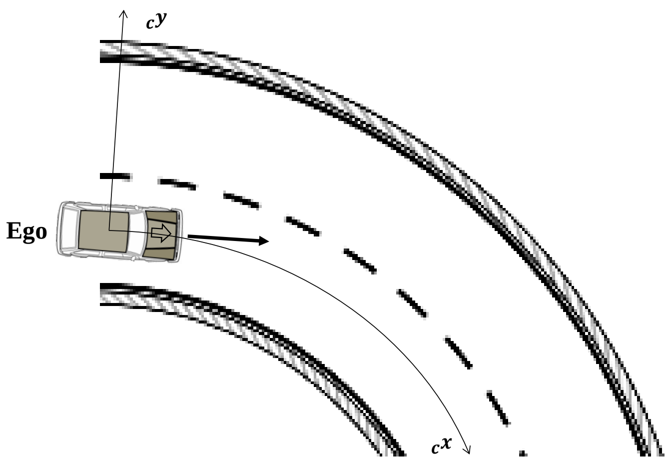

Figure 5-13 Curvilinear vehicle coordinates

Adaptions are now only required regarding TTS and because lane changes are not symmetrical to left and right side and to the assessment of lateral acceleration (because Kamm's circle in curvilinear coordinates is shifted sidewise). Adaption of TTS is not trivial because the underlying assumption for equation (5.14) is that lateral acceleration in both directions is possible. However, higher acceleration potential (in curvilinear coordinates!) towards the outside of the corner is compensated by lower potential in the second part of a lane change. So it is further assumed that the duration of the evasion maneuver is not significantly influenced by the curvature of the road.

The evaluation of the required acceleration however, needs to be adapted and corrected by the baseline lateral acceleration $_ca_{y,\kappa}$ due to curvature κ. Equation (5.5) is transformed to:

$$ {}_cC_a = \frac{{}_ca_x^2 + \left({}_ca_y + {}_ca_{y,\kappa}\right)^2}{(\mu_R \cdot g)^2} \qquad (5.23)$$

$$\text{with } {}_ca_{y,\kappa} = v^2 \cdot \kappa$$

A downside of this approach is that vehicle shapes that are approximated by rectangles in world- or ego-coordinates cannot be described in the same way in curvilinear coordinates because the rectangular shape is transformed to an arc segment. For small curvature on motorways, the effect can be neglected or compensated by over-approximating the dimension (comp. Figure 5-14) in a trade-off between accuracy and computational effort.

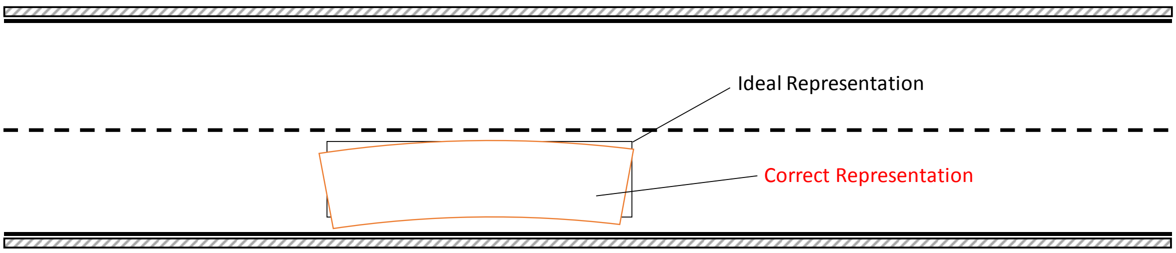

Figure 5-14 Comparison of ideal representation and the correct representation for cornering to left in curvilinear coordinates

5.4 Discussion and Evaluation of TCI

In this subchapter, the developed metric TCI is evaluated based on the requirements and design guidelines derived in sections 5.1.1 and 5.3.1.

RM 1 In case of an increased crash likelihood, the value of the metric must increase, independent of the type of accident.

TCI is developed according to the design guidelines in section 5.3.1 and is therefore capable of assessing all types of accidents. The test scenarios defined in section 5.2.2 do not falsify the metrics eligibility: Cornering and narrow lanes are covered with equations (5.23) and (5.18). Multi-object scenarios are covered because the trajectory is optimized in a predictive approach including all present objects. In addition to those theoretic considerations, the test scenarios are applied in subchapter 6.1.

Nevertheless, there is no proof for the fulfillment of the requirement. The arbitrariness of parametrization (e.g. in equations (5.16), (5.19) and (5.20)) suggest that the metric will need further calibration in order to fulfill the requirements for the most precise extrapolation of risk. However, verification is only possible using large amount of real world data and comparing the extrapolated accident risk with the occurring number of accidents. Before the introduction of AD3+, this is only achievable with human traffic and the application on AD3+ likely requires different parametrization (e.g. due to faster reaction time). When a final AD3+ function is available, the metric should be parameterized with functional knowledge and reevaluated during testing and market observation. Nevertheless, the metric should be able to fulfill the requirements if parametrized correctly because all design guidelines were followed. However, some parameters, e.g. regarding the extrapolation, will remain uncertain resulting in an uncertainty of the final result.

RM 2 The value of the metric shall either increase or decrease strictly with increasing crash proximity towards a known numeric value representing a collision.

This requirement is fulfilled, as the crash cannot be prevented at TCI of one increasing from a minimum value of zero. The direct link to crash proximity is established by following the design guidelines and implementing the three components of driving requirements. As explained above, the weighting and calibration of these three factors are still arbitrary but offer a basis for parameterization based on data or knowledge.

RM 3 The severity of the potential accident shall be estimated.

This requirement is not addressed by TCI. As mentioned before, the estimation of severity is manifold and depends not only on the velocity but also on collision angle, vehicle types and passengers. An estimation could be done by using the velocities of ego and object vehicles at the highest criticality as a worst case and estimate the severity based on accident statistics. Kusano et al.[207] pursued a similar approach for collision mitigation brake systems, however, the actual collision speed is known in their study because the collision is only mitigated. At the point of the highest criticality, it is still likely that the velocity would be reduced if a collision would follow, so that the actual collision speed remains unknown and worst-case approximation is likely to overestimate the severity severely. Nevertheless, the statistics of velocity at the highest criticality value could be evaluated. Conclusively, the metric is not eligible to estimate severity dependably, so the crash risk extrapolation is either a worst case estimation or estimates only the crash likelihood without distinguishing between different severity categories.

RM 4 The metric shall detect all critical scenes without false-negatives.

This requirement addresses the second purpose of the metric, the identification of critical scenarios for test case derivation. It has to be pointed out that the metric does not assess a scene as critical, if the driver reaction was appropriate. As an example, hard braking is not assessed as critical if the remaining reaction time is not low. An issue is that the environment data must be correct and the available tire-road friction must be known in order to identify all critical scenes. If this is fulfilled, the metric identifies the scenes correctly assuming that the driving skill is constant and not compromised (e.g. by dizziness or system failures). The metric only addresses the action level of the driving task (comp Figure 1-5). This is sufficient for an extrapolation of the crash likelihood, as inadequacies in precious levels will lead to criticality on the action level. However, for identification of test cases, it might be useful to develop metrics that identify critical scenarios on all levels of the driving task because known unknowns are identified faster using less data.

---

[207] Kusano, K. D.; Gabler, H. C.: Potential occupant injury reduction (2010).

6 Application and Verification

In this chapter, TCI is applied on artificial and real-world data. First, it is applied on the test cases from section 5.2.2 to verify its behavior by correctly assessing the criticality. Second, it is applied on the highD-dataset, which consist of recorded trajectories from motorway traffic in Germany. Thereby, the goal is to derive the distribution of scenes with different criticality and ultimately applying EVT to extrapolate crash risk.

6.1 Verification in Test Cases

In this subchapter, the metric is applied on the test cases introduced in section 5.2.2. As the metric was developed following the design guidelines, correct behavior in all three test cases is expected.

The first test case is a test for cornering scenarios. The metric shall be sensitive to road geometry and increase the criticality with curvature due to equation (5.23). According to the guidelines for German motorways RAA[208], the road curvature should be less than 1/(900 m), a small criticality is expected when travelling curvatures above this threshold. As test case, two different curvatures, above and below the threshold are selected. The test case is passed when the criticality increases.

The second test case deals with passing objects at a small lateral distance. When the distance is small, the criticality increases due to the necessary precision in course angle (comp. equation (5.18)). In the test, two lateral distances are compared: one with 0.1 m lateral distance and the second with 1 m lateral distance. The test case is passed, if the smaller distance results in a higher criticality. In the scenario, the ego-velocity is 140 km/h and the relative velocity is 60 km/h resembling in overtaking a truck on the motorway.

In Figure 6-1, the test results for the first two TS are presented. Both test cases are passed because the more critical variation of the TS is assessed correctly as more critical. The cornering test cases results in TCI values of 9.9 and 13.3 respectively. However, constant cornering at the given speed and curvature would require lateral acceleration of 1.5 and 1.9 m/s² and a respective TCI of about 15 and 19. The relatively short prediction horizon of 2 s prevents a higher TCI because during the prediction the vehicle can move to the outside of the road resulting in a smaller curvature of the vehicle trajectory compared to the road curvature. The second TS is also passed because the criticality is increased at the smaller lateral distance. Before reaching the object, the criticality increases continuously as an early evasion

[208] Forschungsgesellschaft für Straßen und Verkehrswesen: Anlage von Autobahnen (2009).

or a deceleration is still possible. The TCI for the higher distance that corresponds to a usual driving situation on the motorway is similar to the evenly typical cornering situation. This is a hint that the metric is in fact applicable on different types of scenarios.

The third TS evaluates two similar scenarios, where the ego-vehicles approaches slower or static vehicles. This is typical when approaching traffic jams (v_{rel} = 120 km/h) or trucks that drive at a lower speed (v_{rel} = 60 km/h). In the first scenario, just the right lane is blocked, where the ego-vehicle is travelling, so an evasion would be possible. In the second scenario, both lanes are blocked, so braking is the only accident free solution. In the test scenarios, the ego-vehicle shows no reaction, ultimately colliding with the object on the right lane. The test is fulfilled when the first scenario shows lower (or equal) criticality at all time. Additionally, the crash can be prevented at a smaller distance, so the maximum value of TCI, which is equal to 100, should be reached later. Additionally, an increase in criticality over time is expected. In Figure 6-2, the evaluation of TCI in four different scenarios is evaluated. Additionally, TTB is applied for comparison. As discussed in section 5.2.3.1, TTB does not pass the test case because it results in the same assessment independent of the possibility for evasion. TCI performs as expected. The scenario with two blocked lanes is much more critical than the scenario with only one blocked lane at the same relative velocity. In addition, the last moment where the accident could be prevented is identified corresponding to zero TTB. The difference between the scenarios with possible evasion is small. This is because the evasion trajectory requires similar accelerations in both cases.

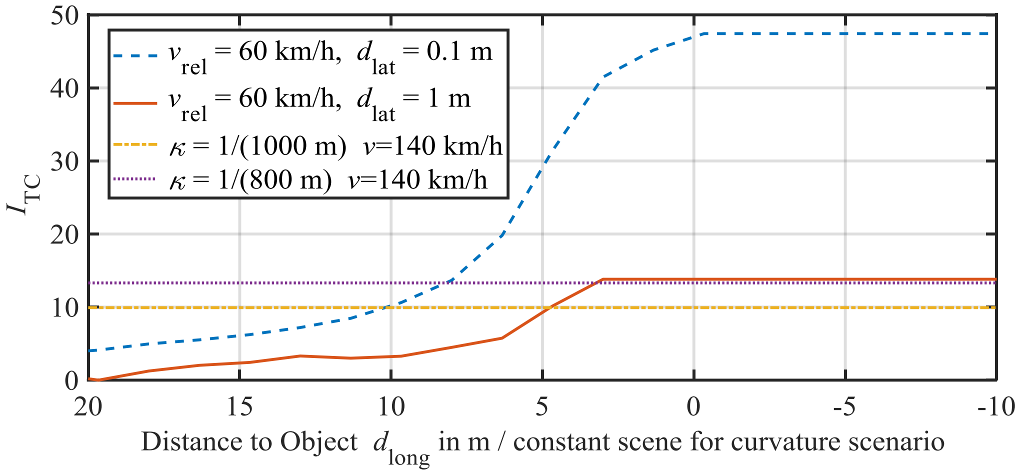

Figure 6-1 Evaluation of TS1 and TS2 with two similar scenarios each

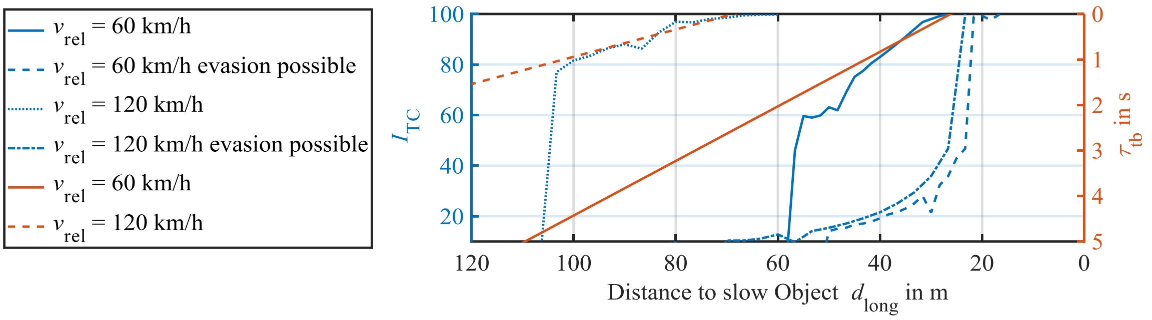

Figure 6-2 Evaluation of TS3; Approaching on a traffic jam with collision, comparison of one blocked lane (evasion possible) with two blocked lanes

The development of TCI over time is not as smooth as it should be expected. Presumably, the choice of initial conditions for the global optimization is the reason. If the global minimum is not found, an increased TCI is the result. The influence will be further analyzed in section 6.2.4.

6.2 Evaluation of the highD-Dataset

In this subchapter, TCI is applied on the data from the highD-dataset that is described by Krajewski et al.[209]. The data contains trajectories and object properties from over 110,000 vehicles travelling in two direction of a 420 m section of six different sections of straight motorways in Germany. The trajectories were acquired by video surveillance of a drone flying above the motorway. Trajectories and object dimensions are derived by image processing (Figure 6-3). The authors do not provide information on the accuracy of the image processing except that the typical position error is less than 10 cm[210]. The calculations for the evaluation of the data were conducted on the Lichtenberg high performance computer of TU Darmstadt.[211] The duration to analyze all vehicle trajectories was 1500 core hours, up to 70 cores were used simultaneously. The duration per vehicle was mainly dependent on the traffic density and the criticality of the scene with an average duration of 50 core seconds per vehicle.

[209] Krajewski, R. et al.: The highd dataset (2018).

[210] https://www.highd-dataset.com/details, accessed on 02.05.2019

[211] https://www.hhlr.tu-darmstadt.de/hhlr/index.en.jsp, accessed on 02.05.2019.

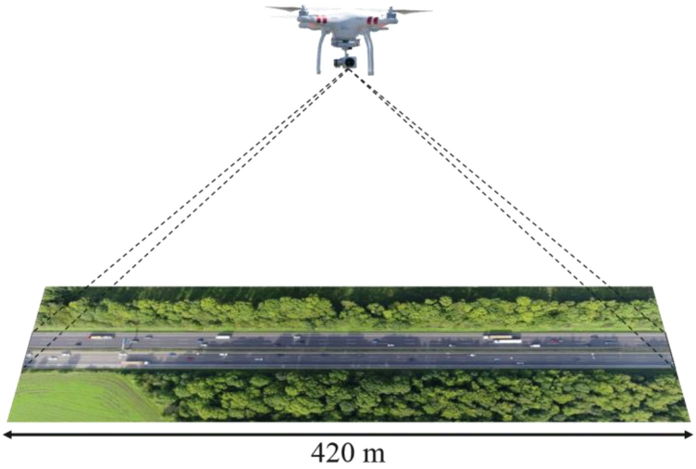

Figure 6-3 Data acquisition for the highD-dataset[212]

As TCI is designed to assess the criticality for an ego-vehicle, each vehicle in the dataset was analyzed separately resulting in 110,000 scenarios of 420 m length, so the total mileage is about $5 \cdot 10^4$ km. For each vehicle the environment data is transformed into vehicle coordinates. Transformation into curvilinear coordinates is not required because all data is recorded on straight roads. The data approximates all vehicles as rectangular objects giving the position, the dimensions and the velocities of all objects. Unfortunately, all yaw angles are assumed as zero because the image recognition could not determine the correct angle. As a course angle of zero overestimates the criticality especially in lane changing maneuvers, the angle must be estimated before application of the metrics. This could be done either by using information about the speed in lateral and longitudinal direction because the course angle is the direction of the velocity vector, or by comparing several subsequent positions. As the velocity information is already derived from subsequent positions, the latter approach is used. To prevent bias due to large jumps in one discretization step, the course angle is derived from linear approximation of the trajectory over five subsequent positions using the least squares method.

In 2016, the average distance between two accidents on the motorway in Germany was $7.6 \cdot 10^6$ km. So the difference between recorded mileage and extrapolation distance is about two orders of magnitude, which is similar to the factor from related studies[213]. However, the highD-dataset is not representative for motorway traffic in general, as there is no curvature, no motorway access, no construction site and no bad weather. It cannot be approximated with certainty from the available accident data, what proportion of accidents falls into the recorded conditions. Schoenawa found that about 35% of accidents in the accident database of the German In-Depth Accident Study (GIDAS) database occurred in similar driving conditions than the highD-dataset's conditions (straight driving, no entry/exits, and available

[212] Krajewski, R. et al.: The highd dataset (2018), p. 2119.

[213] Asljung, D. et al.: Comparing Collision Threat Measures using EVT (2016), p. 60.

lane markings).[214] However, cornering and reduction of lane numbers are included here, so the percentage is likely reduced even more. In the following the extrapolation will be compared with the average distance on motorways in general and additional to a best-case estimate of the average distance of 10^8 km. Note that the issues of representative data also exist when recording data with AD3+ vehicles (comp. section 2.2.2), though of lesser relevance because data collection is not limited to few locations.

6.2.1 Distribution of Critical Scenes

In order to reduce the computational effort, scenes, where driving straight with constant velocity, or slight braking results in a TCI of less than ten, no further optimization was conducted. All other scenes were analyzed. Consequently, when the final TCI was less than ten, the scene was not considered further. The table of all results is given in Appendix B. The scenario of one vehicle driving through the observed area of about 420 m was considered independent and the maximal TCI in this scenario was further processed. Additionally, it is assessed if a lane-change is followed by the critical scene indicating that the driver chose evading instead of braking. It is further indicated, if the vehicle is a car or a truck based on the classification that is part of the dataset. As depicted in Figure 6-4, scenarios that include lane changes are more critical than straight driving. However, it is unknown, if the scenario would have evolved less critical, if the driver decided for braking instead. It could also mean that drivers in general prefer evasion instead of hard braking, if still possible. The scenarios where the ego-vehicle is classified as truck are less critical compared to the whole dataset. However, this difference is not observed when focusing on lane change. However, lane changes by trucks are rare with less than 1500 scenarios in the dataset, so there is little significance. In total, 103 scenarios exceed the threshold of ten (< 0.1%). The highest criticality is at TCI 38%. The most critical scenes are analyzed in the next section.

[214] Schoenawa; Stefan: Critical Scenarios for Humand Drivers (2017).

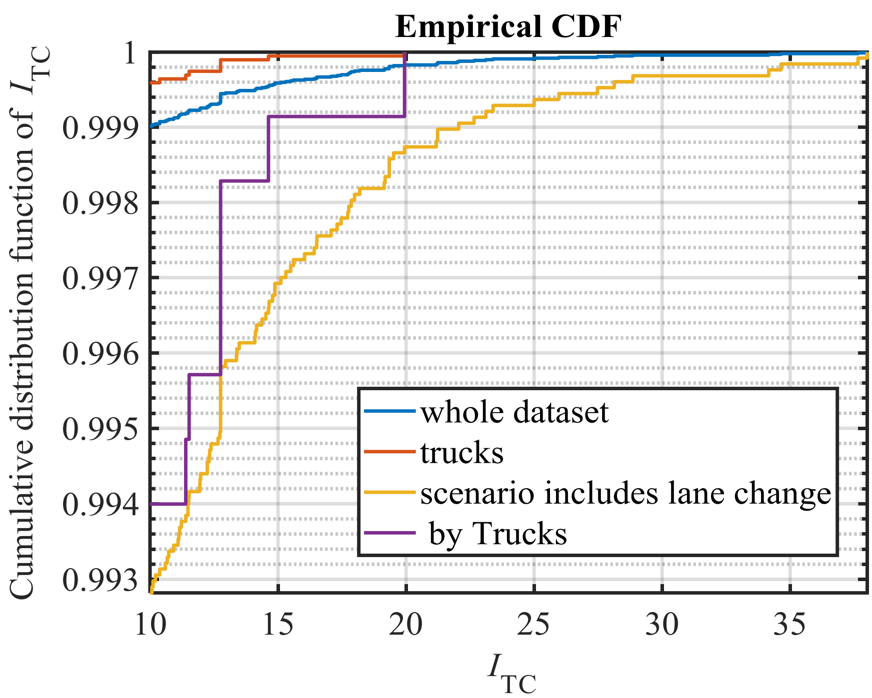

Figure 6-4 Cumulative distribution of TCI in the dataset

6.2.2 Analysis of Critical Scenes

In this section, the four most critical scenes are analyzed that all exceed a TCI of 30%. The scenes are depicted as a top-view in Figure 6-5 to Figure 6-8. The code for image generation is supplied together with the highD-dataset.[215] All scenes include a vehicle approaching on a slower object. In all scenarios, the driver decides to change lanes instead of a relatively hard braking. Subjectively, all scenarios are critical, however it is unknown, if the driver planned the maneuver long ahead. In this case, the reaction time should be assessed differently, resulting in a different TCI.

Furthermore, all four scenarios would probably not occur in an AD3+ system. Figure 6-6 and Figure 6-8 show relatively aggressive driving behavior of merging between two other vehicles. In the other scenes the distance to the front object before the evasion is smaller than the mandatory safety distance, which is legal to undergo for short timespan but not to be expected as a typical driving behavior for an AD3+ system as they would probably either change lanes earlier or decelerate.

---

[215] Krajewski, R. et al.: The highd dataset (2018).

Figure 6-5 Most critical scene in most critical scenario with TCI 38%, vehicle 76 is overtaking with v_{rel} = 18 km/h.

Figure 6-6 Most critical scene in second most critical scenario with TCI 37.7%, vehicle 104 is overtaking with v_{rel} = 7 km/h and merging between two other vehicles.

Figure 6-7 Most critical scene in third most critical scenario with TCI 34.6%, vehicle 87 is approaching with v_{rel} = 10 km/h and later overtaking after vehicle 93 has passed.

Figure 6-8 Most critical scene in forth most critical scenario with TCI 34.16%, vehicle 769 is approaching with v_{rel} = 63 km/h, scenario is later solved with combined braking and evasion to the left with total acceleration of 5 m/s².

6.2.3 Application of Extreme Value Theory

To estimate the accident occurrence rate, EVT is applied on the observations of TCI. The equations used in the section are given in appendix C based on state of the art EVT introduced in section 2.2.2.2.3. The average distance between two accidents on German motorway is about $8 \cdot 10^6$ km[216]. The distance covered in the highD-dataset is about $4.5 \cdot 10^4$ km without any accidents. A distance factor for statistical proof of three is derived from Figure 2-4. So the accident per distance information in the highD-dataset gives statistical evidence with 95% confidence, that the average distance between accidents is higher than $1.5 \cdot$

[216] Destatis: Verkehrsunfälle - Fachserie 8 Reihe 7 - 2015 (2015).

10^4 km. Using EVT and data that is enriched by TCI, a higher worst-case estimate should be achieved. However, the data only represents a small part of the actual road network. The true accident frequency on these small extracts is unknown. Nevertheless, the extrapolation is compared with the average accident frequency in the following. It is expected that the statistical worst-case estimation of accident distance using EVT is between $1.5 \cdot 10^4$ km as a worst-case assumption based on the statistical evidence and 10^8 km, as best-case estimation for the true accident distance for those sections.

To select a threshold over which the observation are used in the parameter estimation for EVT, the Mean Residual Life Plot is used (comp. section 2.2.2.2.3), plotting the mean of all excesses over the threshold. The plot should be approximately linear until the maximal valid threshold. Based on Figure 6-9, a threshold of 12.5 is further used, indicated by the vertical line. Until this value, the Mean Residual Life Plot is approximately linear. A higher threshold (e.g. until 23, leaving only eleven exceedances) would be justifiable as well based on linearity alone due to the confidence interval. However, increased variance would be the consequence, reducing bias from relatively uncritical scenarios.

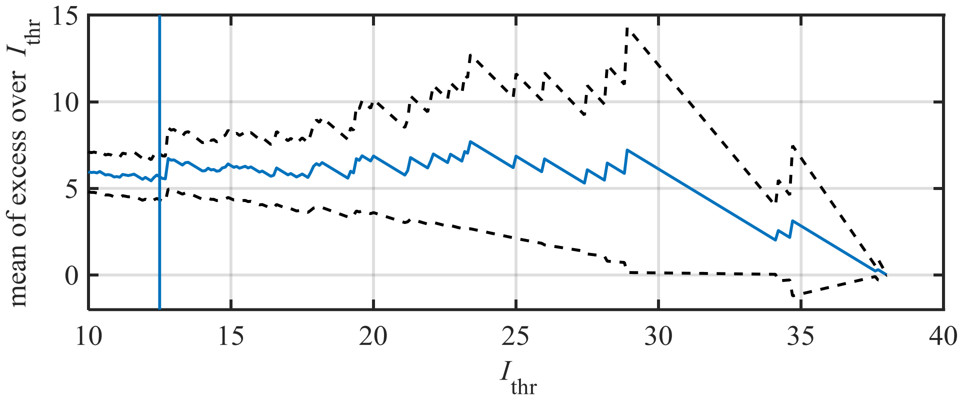

Figure 6-9 Mean residual life plot for TCI, dotted lines are 95% confidence intervals; selected threshold of 12.5 marked with vertical line

Using the threshold and the observations of TCI, the parameters estimations and 95% confidence intervals for $\hat{\xi}$ and $\hat{\sigma}$ are derived maximizing equation (2.6) resulting in $\hat{\xi} = 0.059 \pm 0.35$ and $\hat{\sigma} = 5.5 \pm 1.7$. To estimate the distance between accidents, the return level plot, and especially the return period on a value of 100 is of interest. As the confidence interval of $\hat{\xi}$ includes positive values, a standard deviation for the return level of 100% cannot be derived by the delta method, because it is unbound towards higher distances. Instead, maximizing and minimizing equation (2.8) with values of ξ and σ that fulfill the following condition of equation (6.1) and varying m, leads to the confidence intervals for $I_{\text{TC},m}(m)$ depicted in Figure 6-10. According to Coles[217], all values of ξ and σ must have a likelihood that is only half of the 95% quantile of the χ_1^2-distribution smaller than the maximum likelihood:

[217] Coles, S.: An introduction to statistical modeling of extreme values (2001), p. 42.

$$\ell(\sigma,\xi) \geq \ell(\hat{\sigma},\hat{\xi}) - 0.5 \cdot \chi^2_{1,0.05} \tag{6.1}$$

The resulting extrapolation and confidence intervals are depicted in Figure 6-10. The worst-case estimation for the distance between two accidents is $4.0 \cdot 10^5$ km, which is about 15 times higher, than derived from accident frequency directly. The extrapolated maximum likelihood distance of $1.3 \cdot 10^7$ km has little significance and is not considered further.

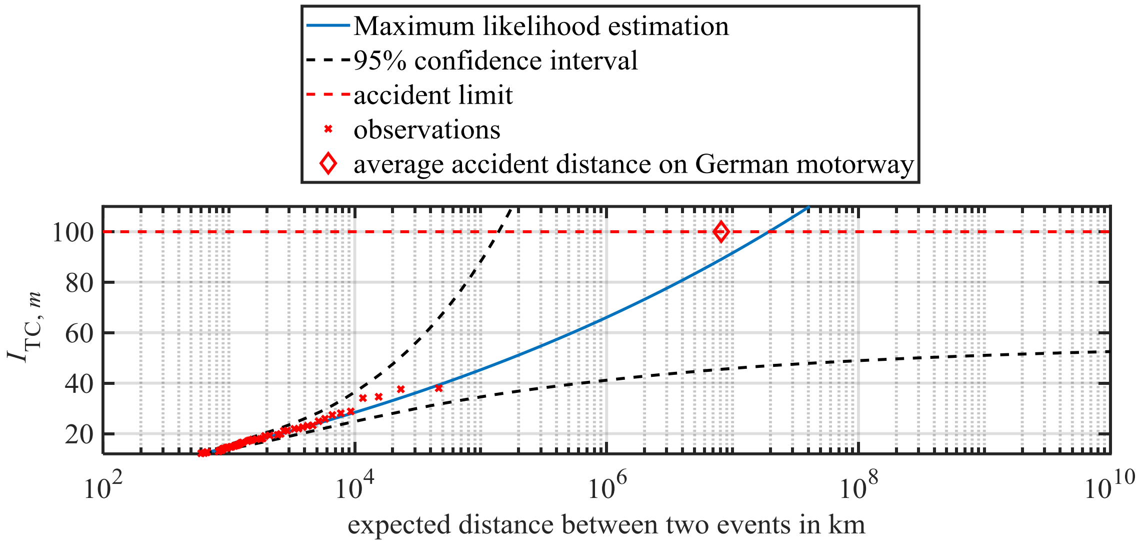

Figure 6-10 Return level plot for TCI with extrapolation and confidence intervals (dotted); red crosses are observations in the data, the diamond at TCI 100% marks the average distance between accidents on German motorways.

The results do not falsify the applicability of the metric. Furthermore, the lower bound of the 95%-confidence interval is more than one order of magnitude higher than by the statistical evaluation of driven accident-free mileage alone, indicating that less mileage is required for the estimation of safety performance compared to accident data alone.

6.2.4 Sensitivity Analysis

As demanded by the design principles, arbitrary or uncertain parameterization of the metric and the extrapolation method should be analyzed. Hence, a sensitivity analysis is conducted varying parameters in TCI and varying the threshold selected in the extrapolation.

6.2.4.1 Parameterization of Reaction Time

Here, the equation for reaction time (5.15) is used, which results into the criticality component C_R based on an assumed normal distribution of available reaction time instead of a linear representation of the CDF as suggested by equation (5.16). Based on the studies analyzed in section 0, the cumulative distribution is parameterized using the mean value $\bar{\tau} = 0.86$ s and standard deviation $\tau_\sigma = 0.2$ s. Both approaches are compared in Figure 6-11. For reaction times larger than 0.8 s, the criticality is now lower compared with the previous parameterization with a steeper increase between 1.4 and 0.4 s of reaction time. As a result, all scenarios evaluated before are now assessed with a lower criticality, because very critical scenes are not part of the dataset. With the new parameterization, only 37 scenarios excess the threshold of TCI ten or higher. Therefore, the extrapolation of distances between accidents results in increased safety as depicted in Figure 6-12. Even the worst-case estimation based on the 95%-confidence interval results in a distance higher than the average on German motorway. However, the average distance of the observed sections of the motorway are unknown so there is no definitive falsification. As derived in the beginning of this subchapter, an estimation of the accident distance for the selected road sections up to 10^8 km is reasonable.

Due to the shape of the CDF, very steep increase in criticality occurs, when the scene demands smaller reaction time. This is an advantage, if it holds true that reaction times of 1.4 s and higher are handled by almost all drivers, as most studies suggest. As pointed out by Hugemann[218], larger reaction times are common in every day driving different from common studies. Due to the shape of the CDF, increased TTR might lead to a very different extrapolation result because the criticality increases steeply. As a worst-case estimation it can be concluded, that the distance between two accidents should be above 10^7 km, if the available reaction times are distributed as expected. Judging by the available data, the previous used parameterization with the linear approach seems less biased compared to the real reaction times. However, the analysis of additional data might help finding the best parameterization.

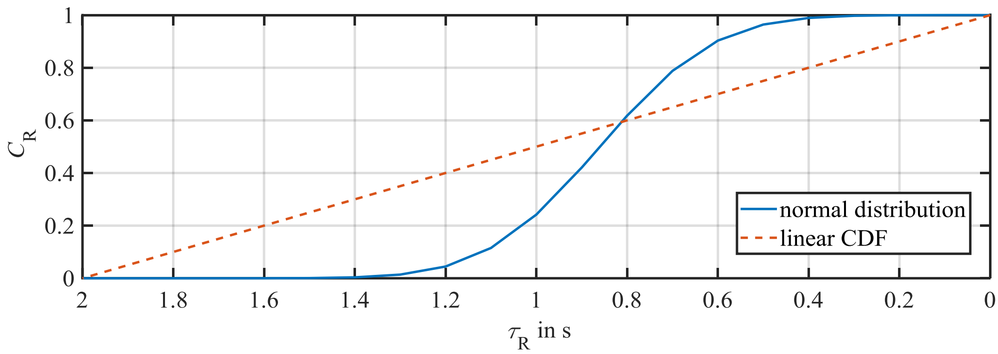

Figure 6-11 Comparison of different evaluation of reaction time

[218] Hugemann, W.: Driver reaction times in road traffic (2002).

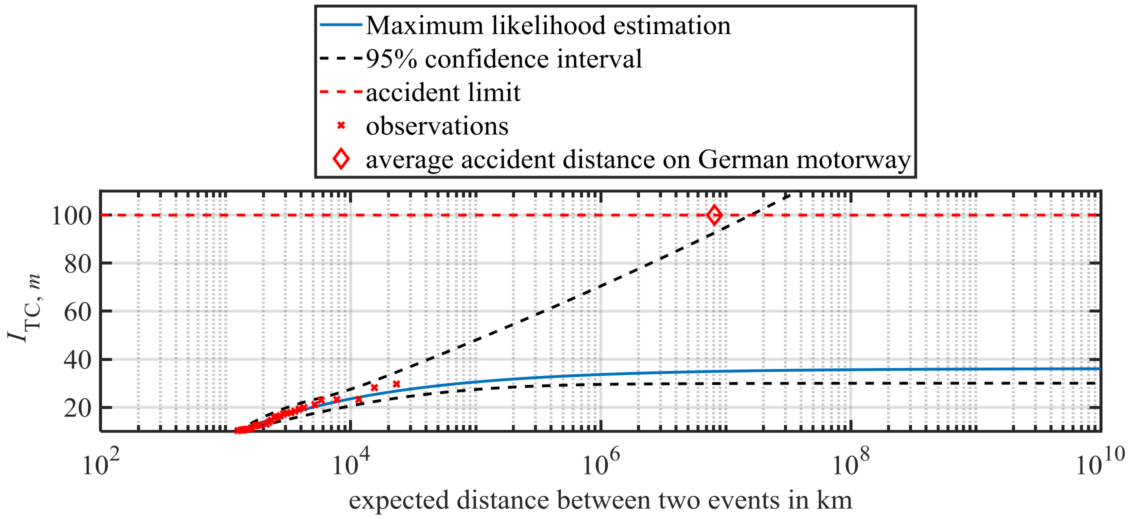

Figure 6-12 Return level plot for reaction time based on normal distribution CDF

6.2.4.2 Weighting of Precision

The weighting of the reaction time from precision in equation (5.19) was discussed previously in section 5.3.2.6. A careful parametrization of 1/10 was used before, also to reduce increased criticality based on uncertain object representation in position and yaw angle. Now, a weighting factor of one is used comparing the results.

As expected, the overall values of TCI are higher because of the increased weighting. Scenes with small lateral distances are most critical. The scene with the highest criticality of 63 is depicted in Figure 6-13. Here, an incorrect approximation of the course angle might lead to an increased assessment of TCI. As both vehicles trajectories are almost parallel during their lane changes, the scene might be less critical, though the data might suggest otherwise because the yaw angle is not directly measured.

Figure 6-13 Most critical scene with higher weighting of precision; vehicle 687 is overtaking truck number 682 after he is changing lanes.

Despite the higher assessment of TCI in individual scenarios, the extrapolation results in an increased safety comparing Figure 6-14 to Figure 6-10. The reason is the decreasing trend of higher TCI values that results in a shift of the shape factor estimation towards negative values. Based on the results, the parametrization is not falsified, but re-evaluation with data including correct course angle is advised as an erroneous course angle has a high influence on the result as derived from Figure 6-13.

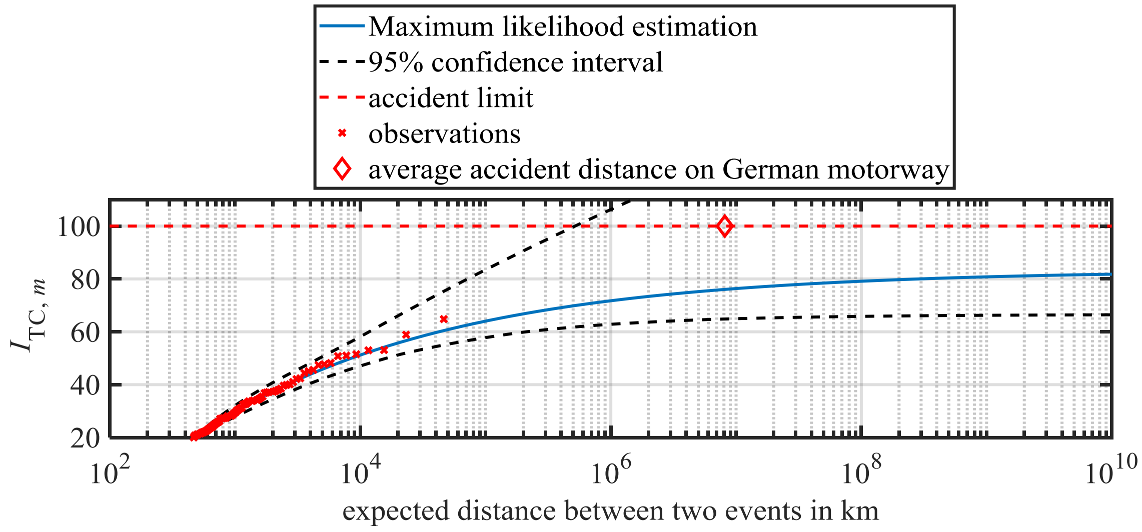

Figure 6-14 Return level plot for higher weighting of precision; I_{thr} = 20.

6.2.4.3 Acceleration weight

In this section, the influence of the acceleration weighting is analyzed. Instead of the radius of Kamm's circle, a rhombus is used now, so combined accelerations are assessed with higher criticality. The implementation is done using equation (5.6) instead of (5.5), which is the 1-norm of the acceleration vector instead of the 2-norm. In general, the criticality will be higher compared to the previous parametrization. For scenes with minor criticality, the absolute influence is small because no increased combined accelerations are required.

A total of 217 scenarios reached a criticality of TCI ten or higher. Hence, an increased threshold of 20 was used for EVT following the linear development in Figure 6-15. In Figure 6-16, it is observed that some scenarios have a highly increased criticality. The scenario with maximal criticality of 50 is the combined braking and lane-change described by Figure 6-8. The application of EVT did not falsify the parameterization as the average accident distance is within the extrapolated 95% confidence interval. The more critical assessment of combined accelerations leads to an estimated shape parameter $\hat{\xi} = 0.22 \pm 0.25$, which is greater zero with higher certainty compared to previous results, leading to a smaller 95%-confidence interval for TCI 100.

However, the results also show that the found maximum likelihood parameterization is not a great fit for the values of increased criticality, so it should be re-evaluated with more data.

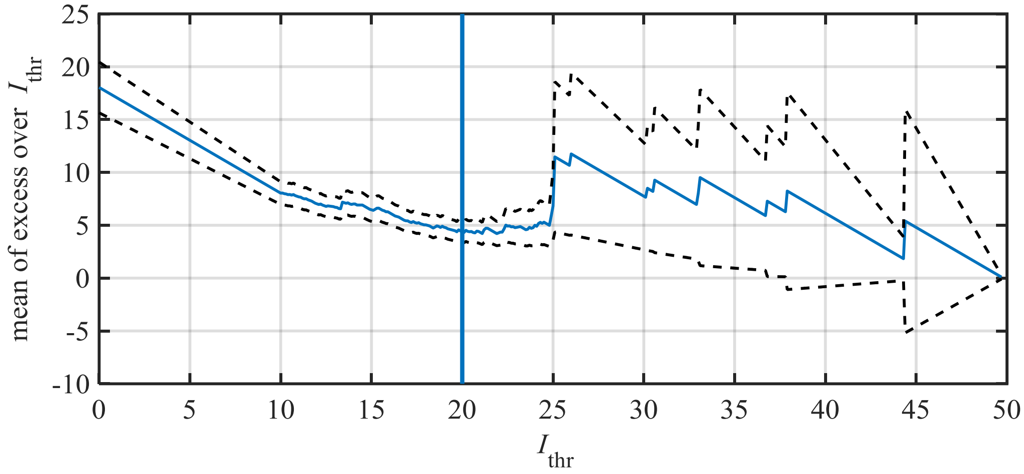

Figure 6-15 Mean residual life plot for TCI, dotted lines are 95% confidence intervals; selected threshold of 20 marked with vertical line

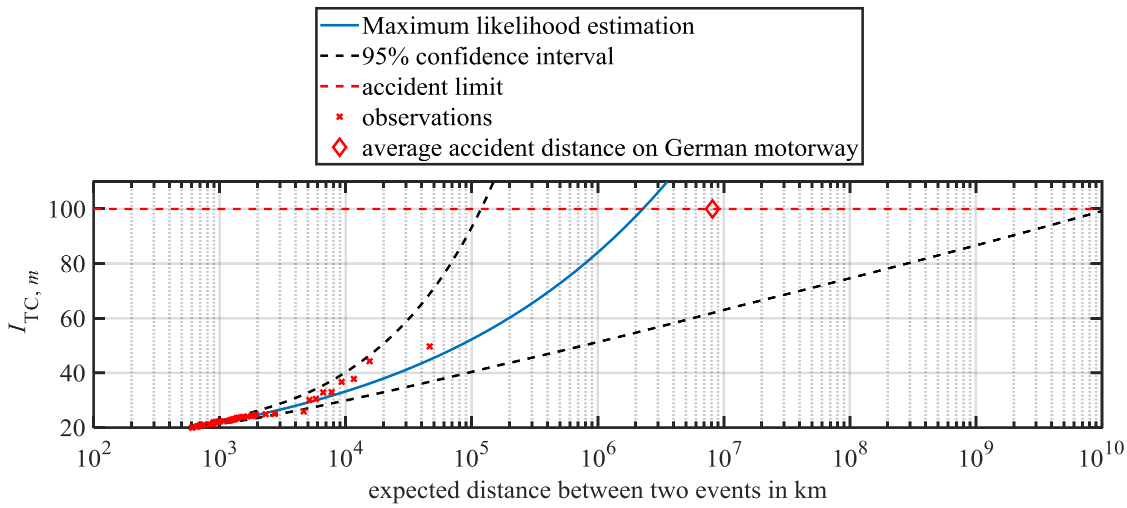

Figure 6-16 Return level plot for reduced combined accelerations; I_{thr} = 20.

6.2.4.4 Initial Conditions

To analyze the influence on initial conditions the data evaluation is repeated using less evasion trajectories to investigate the assumption that especially lateral evasion requires several initial conditions. Now, the initial conditions are parameterized including only one evasion maneuver (both sides): a fast evasion with maximal lateral acceleration.

An evaluation with additional initial conditions could be conducted as well, however, the computational effort increases with every additional initial condition.

In Figure 6-17 and Figure 6-18, the results of the evaluation are displayed. In total, 142 scenarios exceed the threshold of ten (0.14%) with a maximal TCI of 78. So at least in some cases, the global minimum was not found because the evaluation with more initial trajectories resulted in lower criticality. The strong lateral deceleration at the initial condition results in a collision with another object, so the optimization could not iterate towards a crash-free trajectory requiring a minor lateral acceleration. Nevertheless, the maximum likelihood estimation predicts an accident distance of 10^{10} km, despite the higher criticality. The reason is of course the different shape of the extrapolation suggesting a decreasing trend in criticality. Nevertheless, the worst-case extrapolation based on the 95% interval suggests more frequent accidents, but only by a factor of about two.

It is not surprising that despite having very different assessments in each single criticality values, the overall assessment is comparable. When crash-free evasion trajectories are difficult to find due to initial conditions, scenes, where braking is not possible or high braking is required, are highly critical.

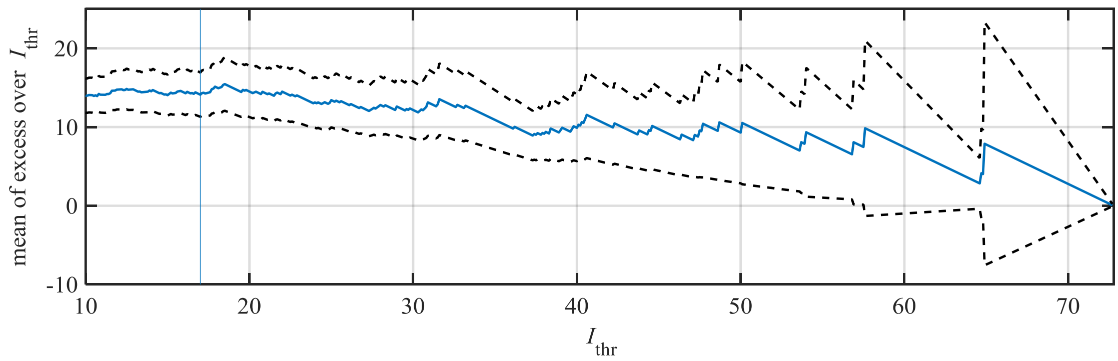

Figure 6-17 Mean residual life plot for TCI, dotted lines are 95% confidence intervals; selected threshold marked with vertical line

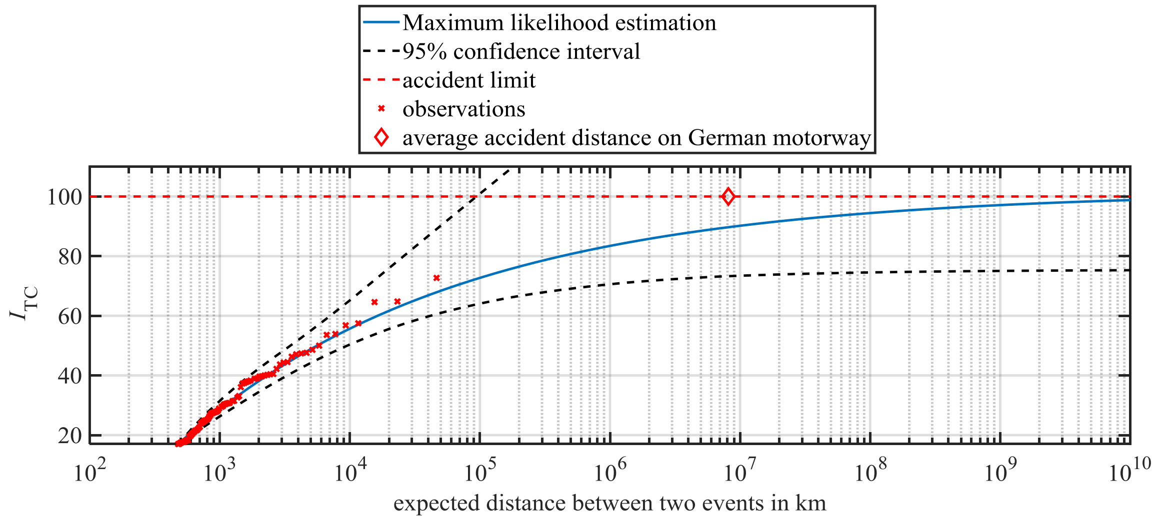

Figure 6-18 Return level plot for TCI with extrapolation and confidence intervals (dotted); red crosses are observations in the data, the diamond at TCI 100 marks the average distance between accidents on German motorways.

6.2.4.5 Threshold Selection

The selection of a threshold for EVT is usually not definitive and is discussed in the following. The threshold of 12.5 as suggested above already has a relatively small number of exceedances resulting in a high variance. Nevertheless, thresholds from 10 to 17 are evaluated in Figure 6-19. At a threshold of 12 and higher, especially the shape parameters estimation $\hat{\xi}$ changes with different thresholds, resulting in very different maximum likelihood estimations of average accident distances. However, the distance between accidents for the lower 95% confidence interval changes only between 10^5 and 10^6 km, indicating that at least this result is reliable. Nevertheless, more data is needed to reduce the variance and allow higher thresholds to reduce bias from uncritical results at the same time.

The results imply that different thresholds should be compared when evaluating the safety of AD3+, where the true average distance is unknown. While the maximum likelihood estimation has little significance, the 95% confidence interval could be used similar to the statistical proof of safety based on accidents alone.

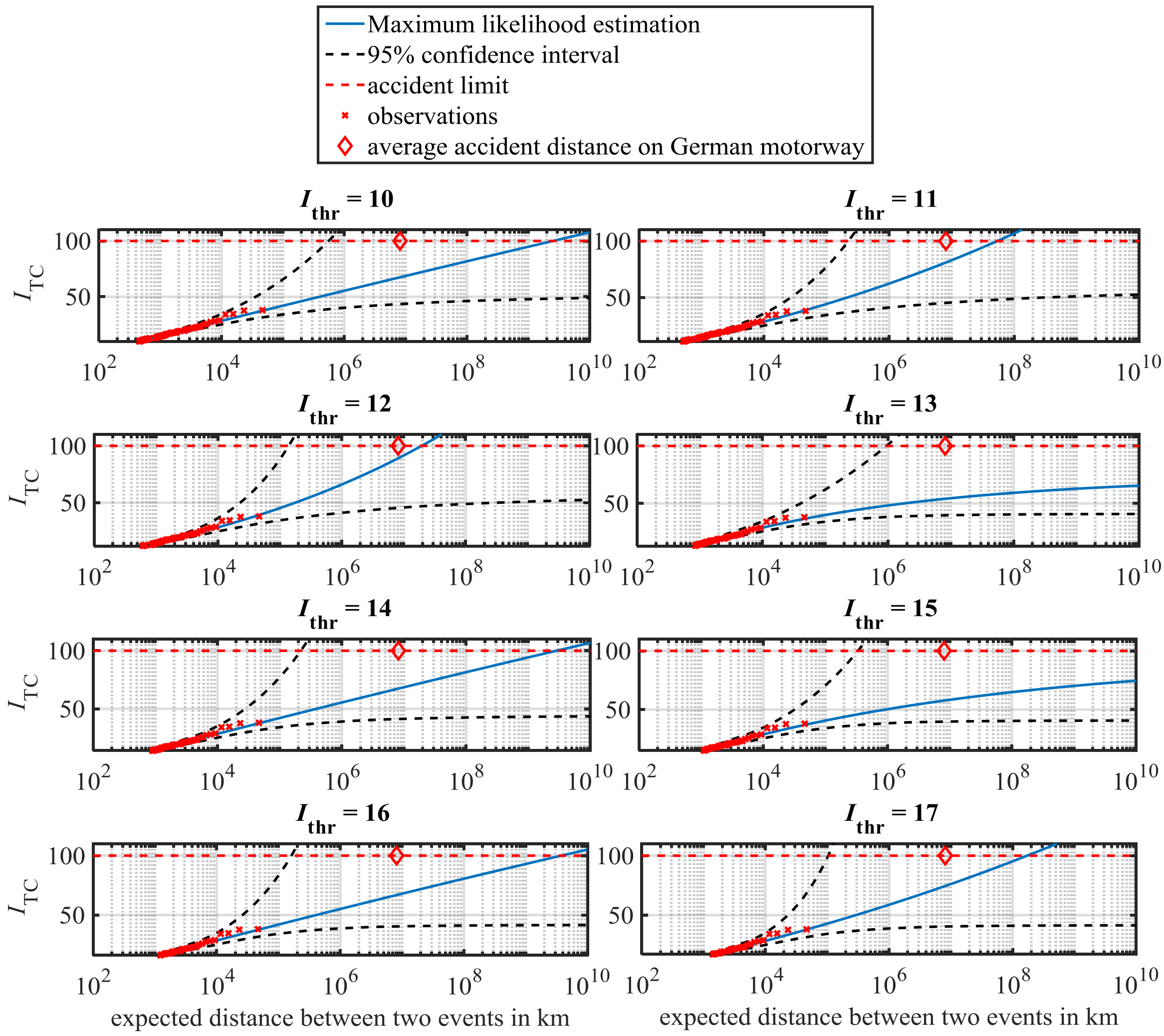

Figure 6-19 Comparison of different thresholds

6.3 Summary

To conclude, the application of EVT did not falsify TCI as a suitable metric for the extrapolation of average accident distances. This was expected as TCI was developed following the design principles derived in section 5.3.1. Nevertheless, the increase in trust in the metric's eligibility is limited as the statistical variance and the variance from uncertain parameterization is large. The sensitivity analysis shows that the maximum likelihood estimation is dependent of parameterization and threshold selection, but there is a certain robustness in worst-case estimation for average accident frequency based on the 95% confidence interval for TCI 100. Independent of the parameter choice, the minimal estimated distance of two accidents lies between 10^5 and 10^6 km, which is about one order of magnitude higher than based on the accident-free mileage alone. As soon as more data is available, the sensitivity analysis should be re-evaluated reducing uncertainty in parameterization because some parameter sets might be falsifiable when applied on more data from human driven vehicles.

For AD3+ vehicles, TCI computation should be adapted because parameters can be chosen based on the knowledge of the systems driving capabilities that are not available for human drivers. The correct TTR is derivable from the path planner as well as the required precision of the trajectory control. Additionally, the available reaction time will be known based on the time that is required by data processing and action. As a result, a metric is designed that is tailored to the specified driving function. Assuming that uncertainties based on arbitrary parameters do not remain, only statistical uncertainty influences the result that is deduced from data alone and thus produces reliable worst-case estimations. With more driven distance, the uncertainty will shrink, which could be compliant with the introduction strategy from subchapter 4.5.

Applying the metric with reduced initial conditions showed that the optimization function has several local minima. It is not guaranteed that the global minimum is found with the used initial conditions. Future research should also focus on improving optimization investigating the trade-off between computational effort and precision of the result. A pre-selection of suitable initial conditions could also accelerate the computation.

7 Reflection

In the following, the progress of this dissertation regarding the research questions is addressed. The key contributions are highlighted and remaining research questions are derived.

7.1 Macroscopic Risk

The three research questions that were derived in subchapter 3.1 are the following:

Q 1 *Which viewpoints are necessary to be considered for a sufficient derivation of requirements for macroscopic risk?*

Q 2 Which acceptance criteria result, if all viewpoints identified in Q 1 *are analyzed scientifically?*

Q 3 What is the influence on introduction concepts due to the acceptance criteria derived in Q 2?

The state of the art discussion about safety requirements lacked quantitative measures for macroscopic risk requirements involving all relevant viewpoints. In subchapter 4.1, the viewpoints of user, society and individual passers-by were derived. The viewpoint of an individual passers-by is often ignored because he is also part of the society as a whole. Nevertheless, it should be addressed individually because the society is a group of users and passers-by and its requirements might vary from individuals because the average safety of the whole group is addressed. It is impossible to derive risk requirements for each individual person because their individual requirements might vary by personal experience, personal benefit and affinity towards new technologies. Nevertheless, requirements can be derived, if each viewpoint is analyzed scientifically addressing the second research question.

To derive quantitative risk requirements, acceptance in other technologies and studies about general risk acceptance are analyzed to address the viewpoints of individuals. As personal risk acceptance varies with the experienced benefit, different requirements based on the type of exposure (voluntary/involuntary/job-related) are deduced. For passers-by, new risks that are unique to the technology are of importance because without personal benefit, those risks must be minimal also if another risk is reduced. Society's requirements focus on the risk reduction over time and total accident numbers in a time period. So an increase in accident numbers can be compensated by a stronger decrease afterwards. This temporary increase motivates Kalra et al.[219] to analyze how fast safety performance must increase to reduce the total number of accidents even if a system with lower safety is introduced at the beginning. While this concept might not function because of the user's requirements, which does not allow increasing risk, a similar idea that monitors the rate of market introduction while re-evaluating the risk based on new information from the field might be beneficial. The risk-limited introduction was formulated by Wachenfeld[220] and is motivated by the fact that current test methods are unable to determine the safety performance without a large uncertainty. The impact for total safety is diluted when the field penetration rate of AD3+ vehicles is low. The dilution of risk is used for the quantitative risk requirements for passers-by and society resulting in lower requirements at the beginning of introduction. Assuming an increase in penetration rate comparable to other mechatronic automotive systems, the society's and passers-by's requirements become stricter than the user's at a field penetration rate of about 12% or higher. Quantitate requirements for an accident occurrence rate with lesser than fatal severity cannot be derived in the same way. Comparison with today's traffic is possible due to detailed accident statistics, but other risk studies and principles such as the minimal endogenous mortality are only available for fatal risks. To derive requirements for other severity categories as well, a reduction of the accident rate equivalent to the reduction of fatal accidents is demanded. As it is questionable, if different severity categories can be weighed against each other, it is suggested that a decrease in one accident category can reduce requirements in categories with less severity but not vice versa. Weighting factors such as the factor ten as used in ISO26262 or a monetary balance based on the costs are possible candidates. However, further studies about risk acceptance or discussions with representatives on how to handle the different severity categories are advised.

Addressing the third research question, it is analyzed if the requirements, the risk-limited introduction, and the expected field penetration increase fit into an introduction scenario, where the uncertainty about the driving function's safety performance decreases fast enough. As the safety performance of the first AD3+ vehicles must comply with their requirements, an initial test distance of about 10 million km is required. If the system is in fact safe enough, the uncertainty is reduced due to field observation so that the requirements of society are fulfilled at all time. However, if the system's safety performance is lower (e.g. as required by the user who has benefits), an increase in performance over time (due to updates of the system or infrastructure) will be required to comply with all requirements. Assuming that the user follows the hypothesis of sufficient safety and rejects it only under sufficient statistical evidence, an introduction is possible as soon as a system is available that tests in an unsupervised field test with 10 million km without accident.

[219] Kalra, N.; Groves, D. G.: The Enemy of Good (2017).

[220] Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017), pp. 102–138.

The scope of this dissertation was the derivation of acceptable MaR requirements based on different viewpoints, other technologies and risk acceptance studies. However, it is not guaranteed if those risks are later accepted. The subjective assessment of risks by individuals is often not fact-oriented as described in section 2.1.2. Especially if accidents happen shortly after the introduction, or if two accidents happen within a short mileage, general acceptance is questionable, though the statistical evidence might still hint towards an increased safety and the fulfillment of acceptable MaR requirements.

However, it is unknown if the user and the industry will accept the high uncertainty after introduction. Here, other test strategies might increase trust but the effect on a decreasing uncertainty is unknown or not sufficiently researched yet.

7.2 Microscopic Risk

The second part of this dissertation focusses on microscopic risk metrics that evaluate the occurring critical scenes with the goal to extrapolate the occurrence rate of accidents. The derivation of uncertainty is in focus as well.

The idea to derive information about road safety from critical scenes started with the first large scale FOT and NDS studies (e.g. 100-Car-Study[221]). All past studies only record a certain amount of data, so the information about environment objects is incomplete. Hence, the application and development of metrics followed a bottom-up approach based on the available data. Critical scenes were identified using the front object and the reaction of the driver that was monitored (comp. appendix A). As the driver's reaction is often stronger than necessary, false positive detection is the consequence and manual annotation is required. Songchitruksa et al.[222] applied extreme value theory to extrapolate crash occurrence rates from critical scenes and concluded that the metrics that describe the criticality must fulfill certain requirements. The metric PET that describes criticality by time-based closeness to a collision was not eligible presumably because difficulty of collision avoidance was different depending on the scene even if the value of PET was identical. The metric BTN showed promising results in a study by Asljung et al.[223,224] but is only applicable on longitudinal traffic. Other types of accidents are ignored. Again, the metric is defined in a bottom-up approach based on the available data. Without additional information on environment and other objects, a top-down approach defining a metric that describes criticality regarding all types of accidents is impossible. With the current advances in environment perception and more and more AD3+ prototypes, the possibility to develop and apply a metric in a top-down

[221] Dingus, T. A. et al.: The 100-car naturalistic driving study (2006).

[222] Songchitruksa, P.; Tarko, A. P.: The EVT approach to safety estimation (2006).

[223] Asljung, D. et al.: Comparing Collision Threat Measures using EVT (2016).

[224] Asljung, D. et al.: EVT for Vehicle Level Safety Validation (2017).

approach rises. Another source of data is video monitoring by drones such as the recent highD-dataset[225] that was evaluated in this thesis. As explained in the previous subchapter, the derivation of uncertainty of the risk extrapolation method is a key factor. The computation of the statistical uncertainty is state of the art and was already done in previous studies. However, additional uncertainties rise from the parameterization of the metric and uncertainties of the data itself. In the following sections, the findings on the top-down methodology, the uncertainty of result, and the development and evaluation of the new metric are discussed.

7.2.1 Findings on the Top-Down Methodology

As explained above, the availability of detailed data motivates the development of metrics in a top-down approach. Based on the analysis of the state of the art, four requirements on metrics and two requirements on data are deduced. Verification of the requirements is only feasible by application on large scale data and if the true accident occurrence rate is known as it is for human traffic. For AD3+ systems, the true accident occurrence rate is unknown so verification before introduction is impossible. Hence, a falsification strategy is developed based on the requirements. A falsification for advanced metrics is again only possible in data evaluation. In a first step, falsification based on human data is suggested. If falsified, their eligibility for AD3+ is at least questionable. Otherwise, trust in the metric increases but verification for AD3+ vehicles cannot be achieved. Hence, design guidelines for metrics are derived from the requirements that lead to a metric that should fulfill the requirements. The progress in each research question is summarized in the following.

Q 4 What are the requirements on microscopic risk metrics for extrapolation of MaR using EVT?

Q 5 What are the requirements on data for application of microscopic risk metrics?

Q 6 What are the requirements on microscopic metrics for identification of scenarios?

For all three research questions, requirements were defined. Identification of critical scenarios must prevent underestimation of criticality in scenes that have already an only slightly increased criticality. Otherwise, false-negative detections will occur meaning that critical scenes are missed. As critical scenes are rare events, missing those, results in a higher driving distances required for collecting the scenes. For EVT application, the orthogonal characteristic between metric's value and criticality is most important. If the same value corresponds to different criticalities, EVT is not applicable. Especially in different types of scenes that are critical due to different reasons (e.g. lateral and longitudinal approach), this is difficult to establish. To retain the metric from subjective assessment, design guidelines are required and discussed below. Other requirements are the existence of a known limit that corresponds

[225] Krajewski, R. et al.: The highd dataset (2018).

to an accident and the estimation of the crash severity. If the last requirement is not fulfilled, only accident occurrence rates, but not the risk (that includes severity) can be extrapolated. This would be a relevant contribution towards safety validation but not a proof that the MaR requirements are fulfilled.

Q 7 How can the eligibility of MiR metrics for the use cases MaR extrapolation and identification of scenarios be falsified?

Q 8 Do state of the art metrics fulfil all requirements derived in Q 5 and Q 6?

As verification of a metric regarding the requirements cannot be achieved using single test cases, a falsification procedure is established. For the first requirement that demands an orthogonal characteristic towards crash likelihood, test cases are derived that contain two similar scenes, where it is obvious which scene is more critical. Only if the metric assesses the scenes correctly meaning that the more critical scene is also assessed with increased criticality, the metric passes the test case. Otherwise, it is falsified. In addition, falsification could be done by evaluating data as described above. If the true accident occurrence rate lies outside the statistical error bandwidth, the metric is not eligible. While most state of the art metrics are falsified using test cases, more advanced metrics that consider all relevant objects and the road geometry could not be falsified. Especially metrics from trajectory optimization seem eligible. Nevertheless, it is questionable if they are able to fulfill the orthogonality requirement because they are typically not designed to compare the criticality in different types of scenarios. To enable a development of metrics that fulfills this requirement, design guidelines are deduced.

Q 9 What are design guidelines for the development of MiR metrics according to the requirements derived in Q 5 and Q 6?

In order to describe criticality with an orthogonal connection towards accident occurrence, it shall derive the driving requirements in a scene and calibrate them with the driving skill. Theoretically, a metric could be applied on each level of the driving task decomposed according to Amersbach et al.[226]. However, as mistakes or failures on lower level will become critical on action level as a consequence, it is sufficient to describe the criticality on action level. The criticality is further separated into the three components: reaction time, precision of course angle, and driving dynamic reserve. All three components must be combined into a criticality value. Hence, a total of at least six parameters is required to define shape and weighting of each component. Based on the current scene and an a posteriori evaluation of all object trajectories, the possible trajectories of the ego vehicle shall be assessed and the trajectory with the least criticality selected. The requirements for this trajectory build the basis for the criticality assessment. A priori assessment and model based prediction of object trajectories is another option. However, model parametrization would require extensive amount of data. If the data is instead evaluated directly, different outcomes of very similar

[226] Amersbach, C.; Winner, H.: Functional Decomposition (2017).

scenes will be part of the data without additional assumptions and parameterization of models.

The parameterization of the three criticality components has severe influence on the result, as was shown in a sensitivity analysis. For human traffic, available studies about the performance of human drivers are insufficient to derive the parameterization from. Faulty parameterizations can be falsified because the true accident occurrence rate is known. Based on a set of feasible parameters for human driver, a first working hypothesis for AD3+ metric be established. Before data on AD3+ systems is available, parametrization cannot be falsified. Nevertheless, deviation to human traffic is likely e.g. regarding reaction time. The careful parameterization of parameters and the discussion of effects of arbitrary parameterization should be discussed.

7.2.2 Findings on Uncertainty of Risk Extrapolation

Uncertain parameterization again results in additional uncertainty of the extrapolation. In the state of the art, the focus is on uncertainty due to the statistic evaluation, while the criticality of each scene that is evaluated is regarded as ground truth. However, arbitrary parametrization that can only be determined with a certain accuracy has an influence on the results. One example is the normalization of reaction time to a criticality value using a cumulative distribution according to equations (5.15) and (5.16). Depending on the parameterization, the trajectory optimization results into different behavior because a smaller criticality caused by the reaction time component leads to later brake reactions. The threshold selection in EVT is another source of uncertainty that is often ignored. Asljung et al.[227] showed that different thresholds lead to very different results. If the working hypothesis that similar parameterization for AD3+ and human drivers can be used, holds true, the metric and extrapolation method could be parameterized using the known human accident rate as calibration. Nevertheless, the true AD3+ accident rate might be over extrapolated when this parameterization is used, if the driving skill of a system is higher or at least more reliable.

The sensitivity analyses showed that the size of the uncertainty due to uncertain parameterization is in the same order of magnitude as the statistical uncertainty and should be considered always when applying EVT based on critical scenes. With more data and a higher mileage, the statistical uncertainty can be reduced. Additional data might also lead to better understanding of TCI parameters. When adapting TCI for AD3+, knowledge about the systems driving capabilities will be helpful parameterizing TCI. Nevertheless, the quantification of uncertainty from uncertain parameters is challenge. Re-assessing the whole dataset is necessary to compare different parameters, which requires high computational effort.

The estimation of the severity is not included and it is questionable, if the severity can be estimated with sufficient certainty. A worst-case assumption based on the kinetic energy is

[227] Asljung, D. et al.: EVT for Vehicle Level Safety Validation (2017).

possible but will overestimate the severity. Hence, the method might be part of the total safety estimation but cannot extrapolate to determine if all MaR requirements will be met.

7.2.3 Developed Metric and Data Evaluation

Developing a metric according to the requirements and design principles revealed two major challenges. First, the description of the driving requirements for the criticality component "precision": While it is obvious that the requirements increase with velocity and decrease with lateral distance, the exact parameterization is challenging. The strength of disturbances is unknown and usually not part of recorded data. Second, the comparison with the driving skill is challenging because the description of driving skill is manifold. Distribution functions of reaction times might be used but it is unknown if the reaction time is dependent from additional parameters, e.g. the complexity of the scene.

Evaluating the metrics on the highD-dataset[228] shows that late overtaking maneuvers are assessed as most critical. The criticality in those cases might be overestimated because the course angle is not part of the data but was interpolated, using several subsequent positions. Nevertheless, the information on the object's position in relation to the driving direction of the ego-vehicle might be compromised. Additionally, it is difficult to detect if an evasion maneuver is initialized, as the steering angle is unknown. This might result in too small reaction times because in reality, the reaction already started. Especially for trucks, the correct representation of the lane change is difficult because the reaction of the trailer is delayed, so it is difficult to observe from above. Further improvement of the course angle estimation might be advised for future work. As TCI includes the remaining reaction time, the decision and planning of the driver would influence the result. When the driver already planned the correct path and is aware that an action will be required, the criticality is reduced. However, the planned path is unknown from data. For human driver, this information is difficult to record, but should be implemented when applying on data from AD3+ systems using information from the path planner.

These results show that the quality of data is crucial. Especially overtaking with relatively small lateral distance results in inaccurate criticality when positions and orientation of vehicles are inaccurate. The difficulty of deriving a course angle from the drone video is a downside of the dataset.

In the data evaluation, a limited set of initial condition was analyzed due to the increasing computational effort. Here, the computation of the metric should be further improved, as the evaluation of the test scenarios show. In order to find the global minimum, more initial conditions or a better pre-selection of candidates is required. A derivative-free optimization method such as the simplex method[229] that is used by matlab's *fminsearch* is an alternative

[228] Krajewski, R. et al.: The highd dataset (2018).

[229] Lagarias, J. C. et al.: Convergence properties of the Nelder-Mead simplex method in low dimensions (1998).

but cannot guarantee to find the minimum. Additionally, it requires a higher computational effort.

To conclude, a method for the top-down development of MiR metrics was designed including requirements, design guidelines and a test and falsification strategy. Based on the design guidelines, the metric TCI was developed that was not falsified by evaluation on the highD-dataset. However, uncertainties from arbitrary parameterization remain. As it is questionable, if the accident severity can be estimated for critical scenes, the extrapolation methods estimates the distance between accidents without categorization of severity.

8 Conclusion and Outlook

Figure 8-1 concludes the overall method discussed in this dissertation. Today's traffic and studies about acceptable risk in other technologies as well as the three principles GAMAB, MEM and ALARP are used to define quantitative macroscopic risk (MaR) requirements for AD3+. To extrapolate MaR based on the microscopic risk (MiR) in single scenes, a new metric called Trajectory Criticality Index (TCI) is proposed.

The developed metric is able to assess human and automated test drives. Using the requirements and design principles, it should be improved further, depending on the available data. For data recorded by AD3+ vehicles, additional information might be used, e.g. information from the path planner to achieve more precise information on reaction time. The analyzed highD-dataset includes over 45.000 vehicle km. Nevertheless, application with EVT has a high variance due to the extremely rare occurrence rate of critical scenes. Hence, the eligibility of TCI as a metric is not falsified and trust in the metric based on the dataset is gained. However, uncertainties due to arbitrary parameterization still exist and should be reduced in future work. Hence, more data should be used to test the metric further and reduce arbitrariness from parametrization of the metric.

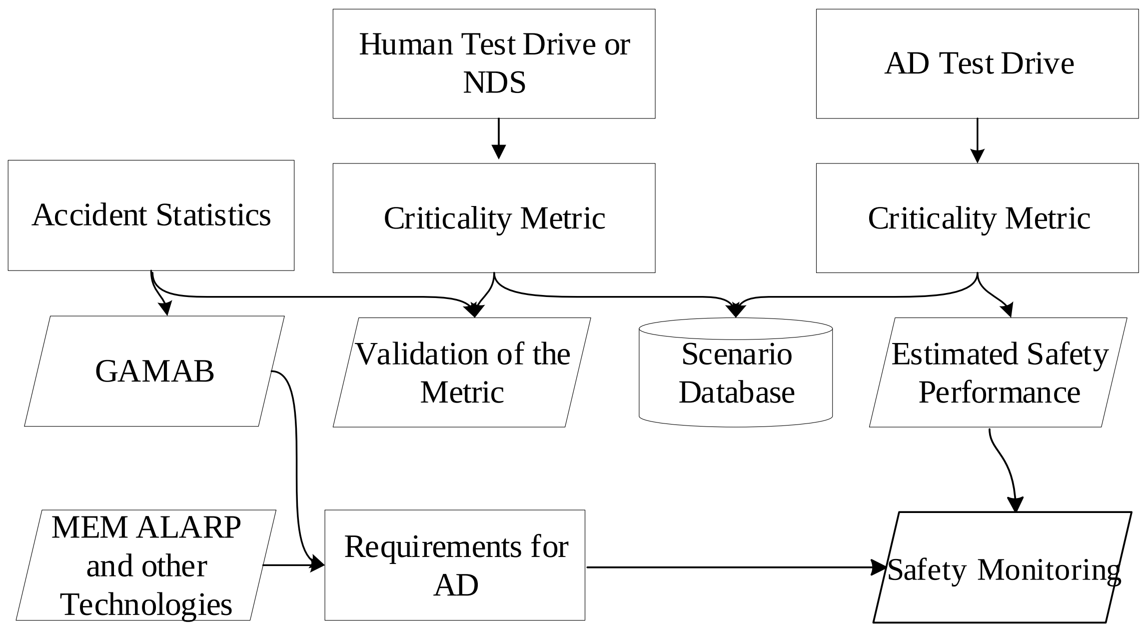

Figure 8-1 Overview of the part of safety validation discussed in this dissertation (modified from Junietz et al.[230])

[230] Junietz, P. et al.: Criticality Metric for the Safety Validation of Automated Driving (2018), p. 61.

Risk extrapolation could support the supervision of MaR requirements during approval or field monitoring in order to provide additional evidence for safety besides the occurring accidents. If the extrapolation suggests a higher risk than accepted, an update for the vehicle could be developed to avoid high-risk observations in the future. The identified critical scenes help improving the driving function. An early identification of system improvement potential is crucial, gaining time for the development, testing and deployment of updates. An identification based on accidents alone, delays the information and might result in legal consequences such as mandatory updates or even prohibition of the system itself. For the online supervision of the fleet, an online calculation of the metric in the AD3+ vehicle is necessary. If computed offline, large amount of data must be saved instead of just saving critical scenes. To achieve real-time applicability, the computational effort must be reduced further. An alternative would be a preselection of scenes with increased criticality without determining the exact value. Data could be saved and post-processed offline.

With knowledge of the system design, the metric should be improved and recalibrated because it cannot be calibrated based on existing data about the average accident distance as it can be done for human drivers. However, the method only extrapolates risks, when the underlying scene that causes the accident risk sometimes results in a critical situation and not an accident. The metric and the method are not eligible to extrapolate the risk for scenes that always lead to an accident. For human driver, this problem does not exist. As far as we know, human drivers are capable handling unknown unknowns that they experience for the first time based on their prior experience. So the method is applicable as long as the data acquisition is able to detect and record the scene correctly. It cannot be assumed that an AD3+ system will handle rare and unknown scenarios in all cases because of lacking knowledge about the system and about the possible environment conditions. As a hypothetical example, a static object on the road that is not detectable by the environment perception system always leads to an accident. Every time, this object is on the road in the path of the AD3+ vehicle, an accident is the consequence, so there are no accident-free critical scenes at least when the sensors recording a dataset do not detect the object as well.

The ultimate goal is to proof that the developed system's safety complies with the derived MaR requirements. The statistical evidence that can be generated from accidents per distance and the occurrence rate of critical scenes increase with additional distance. Dependent on the actual AD3+ systems performance, the requirements from society's and passers-by's viewpoints can be proved statistically. However, the passenger's requirements that are no diluted directly after the introduction cannot be proven because the statistical evidence is not significant enough due to lack of data and the remaining uncertainties from arbitrary parameterization. Knowledge based parametrization can lead to an improved estimation of about one order of magnitude, as the sensitivity analysis showed. Nevertheless, statistical uncertainty remains.

The MaR requirements from passenger's viewpoint cannot be proven at introduction, at least not with the presented extrapolation approach or the state of the art safety approval. As discussed in subchapter 4.5, the passenger might accept the hypothesis of higher safety as long as not proven false. Nevertheless, additional evidence suggesting a sufficiently safe system would increase confidence. The presented MiR approach only contributes towards one of the discussed testing strategies (comp. Figure 2-9). For test approaches other than real world testing (with or without EVT), the uncertainty of the test result cannot be derived. It is unknown how many miles of driving a scenario-based test covers or how uncertainties of model inaccuracies can be described. Nevertheless, all test approaches should be considered together before the market introduction.

To conclude, the following research questions need to be addressed in the future:

- Regarding the method of macroscopic risk estimation:
  - Development of a method or process to gather evidence for safety according to the passenger's requirements
  - Research quantitative uncertainties from test methods
- Regarding the definition of MaR requirements:
  - Which risk will be accepted in a final product and could this be derived before introduction?
  - Do other approaches for the derivation of risks for lower accident severities exist besides reduction proportionally to the risk for fatal accidents?
  - Are there requirements to address the uncertainty of safety in the new technology that cannot be eliminated before introduction?
  - What exactly are the requirements for field observation after introduction?
- Regarding the improvement of MiR metrics:
  - Reduction of uncertainties from arbitrary parameters in the developed MiR metric
  - Improvement of MiR metrics for AD3+using knowledge about the system
  - Improvement of numerical optimization concerning accuracy of the identified minimum compared to the global minimum and computational effort

Especially the quantitative description of uncertainties from other test methods will be a challenge. Otherwise, AD3+ systems must be evaluated in expensive real-world driving tests with uncertain outcome. Alternatively, the uncertainty in the presented approach should be reduced further to reduce the testing effort.

A Trigger Conditions in NDS

A.1 100-Car Study[231]

No.	Trigger Type	Description
1.	Lateral Acceleration	$\lvert a_x \rvert \geq 0.7\ g$
2.	Longitudinal Acceleration	$a_y \leq -0.52\ g$
3.	Event Button	Driver can push a button to mark an event
4.	Forward Time-to-Collision	$\lvert a_x \rvert \geq 0.5\ g$ before or after (± 3 s) around $TTC \leq 4$ s event $TTC \leq 4$ s and $d_{x,front} \leq 30$ m
5.	Rear Time to Collision	• Ignored targets with a speed >44.7 m/s (100 mph) • rear range <15.2 m (50 ft) and the peak longitudinal acceleration of the following vehicle <-0.4 g. • Used a trigger value of two seconds or less
6.	Yaw rate	• The trigger criterion for yaw rate was any set of values that went from neutral (i.e., ~0) yaw rate to +4 degrees/s, oscillated back to -4 degrees/s (or vice versa: -4 to +4), and then returned to neutral within a 3-second time window. • minimum speed of 6.7 m/s (15 mph)

A.2 SHRP2[232]

No.	Type of Triggers	Threshold
1.	Longitudinal Acceleration	• $a_x \leq$ -0,65 g *or* $a_x \geq$ 0,5 g

[231] Dingus, T. A. et al.: The 100-car naturalistic driving study (2006), pp. 331-333.

[232] Hankey, J. M. et al.: Description of the SHRP 2 Naturalistic Database (2016).

3.	Freeway acceleration	• $a_x \leq -0{,}3\ g$; only on freeway
4.	Lateral Acceleration	• $0{,}75\ g \leq a_y \leq -0{,}75\ g$
5.	Swerve	• $\ddot{\psi} \geq 15°/s^2$ and $\ddot{\psi} \leq -15°/s^2$ during a 2 s interval • $v \geq 5$ m/s
6.	Yaw rate	• $\dot{\psi} \geq 8°/s$ and $\dot{\psi} \leq -8°/s$ during a 0.75 s interval • $v \geq 13.4$ m/s
7.	Advanced Safety Systems	• Monitoring activation of systems such as ABS, airbag, ESC and so on
8.	Longitudinal Jerk	• $\dot{a}_x \leq -g$ • $v \geq 5$ m/s
9.	Steering Evasive Maneuver	• $\lvert a_y \rvert > g$ for 0,8 s • $v \geq 5$ m/s

A.3 Critical Scene according to Benmimoun

Benimoun developed heuristics to decide about the criticality of a scene based on driving dynamics[233a], longitudinal distance[233b], driver's reaction[233c] using velocity-dependent thresholds of yaw rate, deceleration, lateral acceleration, TTC and THW.

[233] Benmimoun, M.: Automatisierte Klassifikation von Fahrsituationen (2015).a: p. 106; b: p. 108; c: p. 110

B Information on TCI computation

B.1 Parameters for Optimization

The following parameters were used during optimization:

Parameter	Value
Prediction horizon	2 s
Step size	0.2 s
Optimization Algorithm	interior-point
Core hours for data evaluation	~1400
Lower TCI Limit	10

B.2 Initial Conditions

The following initial conditions are used for the optimization. Six Trajectories are generated based on the current vehicle speed. The duration of the deceleration is adapted at lower velocities, so backward driving is excluded.

In addition to the six trajectories described below, the a posteriori acceleration that were recorded are used as a seventh initial condition.

Full deceleration		medium deceleration		Fast evasion		Medium evasion	
a_x in m/s	a_y in m/s	a_x in m/s	a_y in m/s	a_x in m/s	a_y in m/s	a_x in m/s	a_y in m/s
-10	0	-5	0	0	-5	0	-2.5
…	…	…	…	…	…	…	…
-10	0	-5	0	0	-5	0	-2.5

C Extrapolation

C.1 Equations for EVT

With the number n_{thr} of measurements I_i that exceed a threshold I_{thr}, the mean of the excesses ε over the threshold is given by:

$$\varepsilon = \frac{\sum_{i=1}^{n_{\text{thr}}} I_i - I_{\text{thr}}}{n_{\text{thr}}} \tag{C.1}$$

And the standard deviation ε_σ is given by:

$$\varepsilon_\sigma = \varepsilon / \sqrt{n_{\text{thr}}} \tag{C.2}$$

ε is then calculated for different thresholds $I_{\text{thr}}$and plotted over the thresholds. The resulting curve must be approximately linear until the threshold that is selected and used further.

All measurements over the selected threshold are used to determine the parameters of the distribution function. The following likelihood function is maximized to find the parameters (comp. section 2.2.2.2.3):

$$\ell(\tilde{\sigma}, \xi) = -n\log\tilde{\sigma} - (1 + \frac{1}{\xi}) \sum_{i=1}^{n} \log(1 + \frac{\xi I_i}{\tilde{\sigma}}) \tag{C.3}$$

The 95% confidence intervals of the parmeter pairs are determined maximizing the following formula with the 95% interval of the chi-squared distribution $\chi^2_{1,0.05}$ (comp. section 6.2.3):

$$\ell(\sigma, \xi) \geq \ell(\hat{\sigma}, \hat{\xi}) - 0.5 \cdot \chi^2_{1,0.05} \tag{C.4}$$

The Return Level Plot is given by the following equation, where the input variable *m* describing the number of measurements between two events that can be transformed into the average distance between two events, if its multiplied with the distance of one scenario (420 m in case of the highD-dataset):

$$I_m(m) = I_{\text{thr}} + \frac{\tilde{\sigma}}{\xi} [(m\gamma_{\text{thr}})^{\xi} - 1] \tag{C.5}$$

C.2 Results of Data Evaluation

The following table shows the results of the data evaluation of the highD-dataset. The dataset is divided in 60 parts, each containing about 2000 vehicles:

dataset number	vehicle number	maximal TCI	timestamp in s	vehicle type 1= car 2= truck	Lane Change
2	136	21,23172	1880,16	1	1
2	720	12,38297	9900,44	1	1
3	733	12,75638	9757,8	1	1
3	753	17,07243	10012,2	1	1
3	792	11,51233	10531,28	1	1
4	196	22,65369	2596,84	1	1
4	692	10,73161	9317,36	1	1
4	905	12,94895	12158,24	1	1
5	103	16,03585	1251,72	1	1
5	159	11,95298	1990,48	1	1
5	578	12,75638	7667,16	1	1
6	676	13,48447	9080,44	1	1
6	1006	11,93716	13506,36	1	1
6	1096	14,87581	14692,36	1	1
7	473	17,76028	5793,92	1	1
8	330	19,34367	3931,32	1	1
8	1006	23,11586	12023,24	1	1
9	658	10,92191	8061,72	1	1
9	971	12,75638	11930,8	1	1
9	1015	12,75638	12467,96	1	1
9	1270	12,23689	15726,6	1	1
10	399	17,46344	4934,84	1	1
10	499	10,67332	6107,64	1	1
12	1230	10,37848	21767,08	1	0
12	1322	11,22741	23611,6	1	1
13	1609	28,12887	22995,52	1	1
13	2843	16,51785	40572,76	1	1
16	76	38,0071	1047,48	1	1

dataset number	vehicle number	maximal TCI	timestamp in s	vehicle type 1= car 2= truck	Lane Change
16	153	12,75638	2152,08	1	1
16	630	21,98829	8574,2	1	0
16	643	11,12615	8763,84	1	1
21	115	15,30434	1510,08	1	1
21	124	12,23689	1622,88	1	1
21	947	12,75638	12449,08	1	1
23	433	14,16216	5947,6	1	1
23	1142	11,528	16323	2	1
25	1034	11,12615	29653,28	1	1
26	729	12,23689	20238,76	1	1
26	1856	10,10442	47610,68	1	1
26	2288	16,51785	55873,6	1	1
26	2454	12,23689	59029,12	1	1
29	1576	10,21353	21955	1	1
30	2240	14,78655	32070,12	1	1
31	756	12,75855	10712,64	1	0
32	296	14,87133	4000,72	1	1
32	474	12,30916	6589,64	1	1
32	615	21,23172	8504	1	1
32	1267	12,75638	17723,84	1	1
32	1383	27,47434	19345,24	1	1
33	357	11,97358	5144,76	1	1
33	1583	19,34367	22790	1	1
34	417	17,81569	5810,16	1	0
36	349	11,12615	5001,68	1	0
36	417	19,14605	6047,04	1	1
36	891	25	12872,76	1	1
36	1680	14,62602	24477,84	2	1

dataset number	vehicle number	maximal TCI	timestamp in s	vehicle type 1= car 2= truck	Lane Change
36	2158	16,41325	31466,24	1	1
36	2252	12,75638	32956,2	2	1
36	2320	11,39644	34066,16	2	1
36	2389	12,75638	35275,48	2	1
36	2396	10,37848	35389,32	2	0
36	2531	10,94038	38026,48	1	0
37	104	37,65021	1298,72	1	1
39	1247	10,07419	16552,8	1	1
39	1254	12,332	16651,52	1	1
40	12	10,37848	60,12	1	1
40	687	11,48243	9117,2	1	1
40	1257	19,34367	16793,88	1	1
40	1369	15,60414	18288,44	1	1
40	1598	12,65908	21230,24	1	1
41	852	19,51229	11151,16	1	1
42	769	34,16339	10542,28	1	1
42	2012	13,37678	27203,6	1	1
42	2253	18,19022	30542,28	1	1
43	230	25,96531	2983,84	1	1
43	791	22,05526	10283,24	1	1
43	1321	10,60041	17575,68	1	1
43	2244	10,11919	29979,72	1	1
44	1078	14,09509	14245	1	1
44	2283	14,11744	30524,52	1	1
46	343	17,98727	7386,28	1	1
47	2238	13,394	31331,8	1	1
48	79	17,30619	909,8	1	1
51	219	14,37232	2830,32	1	1

dataset number	vehicle number	maximal TCI	timestamp in s	vehicle type 1= car 2= truck	Lane Change
51	666	17,85901	8920,6	1	1
51	968	15,11027	13015,24	1	1
51	1587	12,75638	21391,56	1	1
51	1703	12,72911	22972,08	1	1
51	2131	19,18285	28852,52	1	1
51	2260	21,17992	30576,16	1	1
54	87	34,64602	1056,96	1	1
54	449	12,75638	5912,88	2	1
54	585	19,94063	7874,44	2	1
55	919	14,52679	12348,08	1	1
55	1388	16,50904	18803,76	1	1
55	1643	14,65173	22226,6	1	1
56	33	15,49139	286,56	1	1
56	107	11,15946	1307,48	1	1
56	1471	23,39749	20012,72	1	1
57	835	11,48243	11165,12	1	1
59	247	17,72402	3100,28	1	1
59	309	11,09642	3935,52	1	1
60	178	28,85808	2291,96	1	1

List of References

Airbus: Commercial Aviation Accidents 1958-2016 (2017)
Airbus: Commercial Aviation Accidents 1958-2016, 2017

Allen, B. L. et al.: Analysis of traffic conflicts and collisions (1978)
Allen, Brian L.; Shin, B. T.; Cooper, Peter J.: Analysis of traffic conflicts and collisions, 1978

Amersbach, C.; Winner, H.: Functional Decomposition (2017)
Amersbach, Christian; Winner, Hermann: Functional Decomposition, in: 8. Tagung Fahrerassistenz, 2017

Amersbach, C.; Winner, H.: Funktionale Dekomposition (2018)
Amersbach, Christian; Winner, Hermann: Funktionale Dekomposition-Ein Beitrag zur Überwindung der Parameterraumexplosion bei der Validation von höher automatisiertem Fahren, in: 12. Workshop Fahrerassistenzsysteme und automatisiertes Fahren, Walting, Germany, 2018

Arrigoni, S. et al.: Safety path planner with collision risk estimation (2016)
Arrigoni, S.; Cheli, F.; Manazza, S. S.; Gottardis, P.; Happee, R.; Arat, M.; Kotiadis, D.: Autonomous vehicle controlled by safety path planner with collision risk estimation coupled with a non-linear MPC, in: The Dynamics of Vehicles on Roads and Tracks: Proceedings of the 24th Symposium of the International Association for Vehicle System Dynamics (IAVSD 2015), Graz, Austria, 17-21 August 2015, 2016

Asljung, D. et al.: Comparing Collision Threat Measures using EVT (2016)
Asljung, Daniel; Nilsson, Jonas; Fredriksson, Jonas: Comparing Collision Threat Measures for Verification of Autonomous Vehicles using Extreme Value Theory, in: IFAC-PapersOnLine (15), Issues 49, pp. 57–62, 2016

Asljung, D. et al.: EVT for Vehicle Level Safety Validation (2017)
Asljung, Daniel; Nilsson, Jonas; Fredriksson, Jonas: Using Extreme Value Theory for Vehicle Level Safety Validation and Implications for Autonomous Vehicles, in: IEEE Transactions on Intelligent Vehicles, 2017

Bagschik, G. et al.: Ontology based scene creation (2018)
Bagschik, Gerrit; Menzel, Till; Maurer, Markus: Ontology based scene creation for the development of automated vehicles, in: 2018 IEEE Intelligent Vehicles Symposium (IV), 2018

Bagschik, G. et al.: Wissensbasierte Szenariengenerierung (2018)
Bagschik, G.; Menzel, T.; Körner, C.; Maurer, M.: Wissensbasierte Szenariengenerierung

für Betriebsszenarien auf deutschen Autobahnen, in: 12. Workshop Fahrerassistenzsysteme und automatisiertes Fahren, 2018

Baum, H. et al.: Volkswirtschaftliche Kosten durch Straßenverkehrsunfälle in Deutschland (2010)
Baum, Herbert; Kranz, Thomas; Westerkamp, Ulrich; für Strassenwesen, Bundesanstalt: Volkswirtschaftliche Kosten durch Straßenverkehrsunfälle in Deutschland, Wirtschaftsverl. NW, Verlag für neue Wiss, 2010

Benmimoun, M.: Automatisierte Klassifikation von Fahrsituationen (2015)
Benmimoun, Mohamed: Automatisierte Klassifikation von Fahrsituationen, fka Forschungsgesellschaft Kraftfahrwesen mbH, 2015

Besch, V.: Produkthaftung für fehlerhafte Arzneimittel (2000)
Besch, Volker: Produkthaftung für fehlerhafte Arzneimittel, 1. Edition, Nomos-Verl.-Ges, Baden-Baden, 2000

Brannstrom, M. et al.: Situation and threat assessment algorithm (2008)
Brannstrom, Mattias; Sjoberg, Jonas; Coelingh, Erik: A situation and threat assessment algorithm for a rear-end collision avoidance system, in: intelligent vehicles symposium, 2008 IEEE, 2008

Broadhurst, A. et al.: Monte carlo road safety reasoning (2005)
Broadhurst, Adrian; Baker, Simon; Kanade, Takeo: Monte carlo road safety reasoning, in: Intelligent Vehicles Symposium, 2005. Proceedings. IEEE, 2005

Büker, M. et al.: Identifikation von Automationsrisiken hochautomatischer Fahrfunktionen (2019)
Büker, Matthias; Kramer, Birte; Böde, Eckard; Maelen, Sebastian V.; Fränzle, Martin: Identifikation von Automationsrisiken hochautomatischer Fahrfunktionen in PEGASUS, in: AAET Automatisiertes und vernetztes Fahren, ITS mobility e. V, 2019

Burckhardt, M.: Reaktionszeiten bei Notbremsvorgängen (1985)
Burckhardt, Manfred: Reaktionszeiten bei Notbremsvorgängen, Verlag TÜV Rheinland, 1985

Coles, S.: An introduction to statistical modeling of extreme values (2001)
Coles, Stuart: An introduction to statistical modeling of extreme values, Springer, London [u.a.], 2001

Crouch, E. A.; Wilson, R.: Risk/benefit analysis (1982)
Crouch, Edmund A. C.; Wilson, Richard: Risk/benefit analysis, in: , 1982

Damerow, F.; Eggert, J.: Predictive risk maps (2014)
Damerow, Florian; Eggert, Julian: Predictive risk maps, in: Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on, 2014

Destatis: Verkehrsunfälle - Fachserie 8 Reihe 7 - 2015 (2015)
Statistisches Bundesamt: Verkehrsunfälle - Fachserie 8 Reihe 7 - 2015, in: , 2015

Destatis: Gesundheit - Todesursachen in Deutschland (2017)
Statistisches Bundesamt: Gesundheit - Todesursachen in Deutschland, 2017

Destatis: Kohortensterbertafeln für Deutschland (2017)
Statistisches Bundesamt: Kohortensterbertafeln für Deutschland, 2017

Destatis: Pressemitteilung Nr. 022 (2017)
Statistisches Bundesamt: Pressemitteilung Nr. 022, 2017

Dieppe, P. et al.: Balancing benefits and harms (2004)
Dieppe, Paul; Bartlett, Christopher; Davey, Peter; Doyal, Lesley; Ebrahim, Shah: Balancing benefits and harms, in: Bmj (7456), Issues 329, pp. 31–34, 2004

Dingus, T. A. et al.: The 100-car naturalistic driving study (2006)
Dingus, Thomas A.; Klauer, Sheila G.; Neale, Vicki L.; Petersen, Andy; Lee, Suzanne E.; Sudweeks, J. D.; Perez, Miguel A.; Hankey, Jonathan; Ramsey, D. J.; Gupta, Santosh; others: The 100-car naturalistic driving study, Phase II-results of the 100-car field experiment, 2006

Döhler, H.; NITSCHE, K.: New mathematical findings on reaction times (2008)
Döhler, H.; NITSCHE, K.: Neue mathematische Erkenntnisse zu Reaktionszeiten bei Notbremsvorgaengen. Teile 1 und 2/New mathematical findings on reaction times during emergency braking. Parts 1 and 2, in: VERKEHRSUNFALL UND FAHRZEUGTECHNIK 9&10, Issues 46, 2008

Donner, E. et al.: RESPONSE 3 - Code of Practice (2007)
Donner, Eckart; Winkle, Thomas; Walz, Rainer; Schwarz, Jürgen: RESPONSE 3 - Code of Practice für die Entwicklung, Validierung und Markteinführung von Fahrerassistenzsystemen, in: VDA Technischer Kongress 2007, Sindelfingen, 2007

Douglas, M.; Wildavsky, A.: How can we know the risks we face? (1982)
Douglas, Mary; Wildavsky, Aaron: How can we know the risks we face? Why risk selection is a social process, in: Risk analysis (2), Issues 2, pp. 49–58, 1982

Eggert, J.: Predictive risk estimation for intelligent adas functions (2014)
Eggert, Julian: Predictive risk estimation for intelligent adas functions, in: 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2014

Eggert, J.: Risk estimation for driving support and behavior planning in intelligent vehicles (2018)
Eggert, Julian: Risk estimation for driving support and behavior planning in intelligent vehicles, in: at-Automatisierungstechnik (2), Issues 66, pp. 119–131, 2018

Eidehall, A.; Petersson, L.: Statistical Threat Assessment for General Road Scenes (2008)
Eidehall, A.; Petersson, L.: Statistical Threat Assessment for General Road Scenes Using Monte Carlo Sampling, in: Intelligent Transportation Systems, IEEE Transactions on (1), Issues 9, pp. 137–147, 2008

Federal Ministry of Transport and Digital Infrastructure: Report of the Ethics Commission (2017)
Federal Ministry of Transport and Digital Infrastructure: Report of the Ethics Commission Automated and Connected Driving, Germany, 2017

Forbes, T. W.: Human factors in highway design, operation and safety problems (1960)
Forbes, Theodore W.: Human factors in highway design, operation and safety problems, in: Human factors (1), Issues 2, pp. 1–8, 1960

Forschungsgesellschaft für Straßen und Verkehrswesen: Anlage von Autobahnen (2009)
Forschungsgesellschaft für Straßen und Verkehrswesen: Richtlinie für die Anlage von Autobahnen (RAA), 2009

Fritzsche, A. F.: Wie sicher leben wir? (1986)
Fritzsche, Andreas F.: Wie sicher leben wir?, Verlag TUV Rheinland, 1986

Gasser, T. M. et al.: Rechtsfolgen zunehmender Fahrzeugautomatisierung (2012)
Gasser, Tom M.; Arzt, Clemens; Ayoubi, Mihiar; Bartels, Arne; Bürkle, Lutz; Eier, Jana; Flemisch, Frank; Häcker, Dirk; Hesse, Tobias; Huber, Werner; Lotz, Christine; Maurer, Markus; Ruth-Schumacher, Simone; Schwarz, Jürgen; Vogt, Wolfgang: Rechtsfolgen zunehmender Fahrzeugautomatisierung, Berichte der Bundesanstalt für Strassenwesen - Fahrzeugtechnik (F), Issues 83, Wirtschaftsverl. NW Verl. für neue Wissenschaft, Bremerhaven, 2012

Geyer, S. et al.: a unified ontology for generating test and use-case catalogues (2013)
Geyer, Sebastian; Baltzer, Marcel; Franz, Benjamin; Hakuli, Stephan; Kauer, Michaela; Kienle, Martin; Meier, Sonja; Weißgerber, Thomas; Bengler, Klaus; Bruder, Ralph: Concept and development of a unified ontology for generating test and use-case catalogues for assisted and automated vehicle guidance, in: IET Intelligent Transport Systems (3), Issues 8, pp. 183–189, 2013

Gibson, S. B.: Risk criteria in hazard analysis (1976)
Gibson, S. B.: Risk criteria in hazard analysis, in: Chemical engineering progress (2), Issues 72, pp. 59–62, 1976

Gordon, T. J. et al.: Multivariate analysis (2013)
Gordon, T. J.; Kostyniuk, L. P.; Green, P. E.; Barnes, M. A.; Blower, D. F.; Bogard, S. E.; Blankespoor, A. D.; LeBlanc, D. J.; Cannon, B. R.; McLaughlin, S. B.: A multivariate analysis of crash and naturalistic driving data in relation to highway factors, in: , 2013

Grunwald, A.: Societal Risk Constellations (2016)
Grunwald, Armin: Societal Risk Constellations for Autonomous Driving. Analysis, Historical Context and Assessment, in: Autonomous Driving, Springer, 2016

Guo, F. et al.: Near Crashes as Crash Surrogate (2010)
Guo, Feng; Klauer, Sheila; Hankey, Jonathan; Dingus, Thomas: Near Crashes as Crash

Surrogate for Naturalistic Driving Studies, in: Transportation Research Record: Journal of the Transportation Research Board, Issues 2147, pp. 66–74, 2010

Haan, L. de**; Ferreira, A.: Extreme value theory (2007)**
Haan, Laurens de; Ferreira, Ana: Extreme value theory, Springer Science & Business Media, 2007

Habenicht, S.: Entwicklung eines manöverbasierten Fahrstreifenwechselassistenten (2012)
Habenicht, Stefan: Entwicklung und Evaluation eines manöverbasierten Fahrstreifenwechselassistenten, VDI Verlag, 2012

Hallerbach, S. et al.: Simulation-based identification of critical scenarios (2018)
Hallerbach, Sven; Xia, Yiqun; Eberle, Ulrich; Koester, Frank: Simulation-based identification of critical scenarios for cooperative and automated vehicles, in: SAE International Journal of Connected and Automated Vehicles 2018-01-1066, Issues1, 2018

Hankey, J. M. et al.: Description of the SHRP 2 Naturalistic Database (2016)
Hankey, Jonathan M.; Perez, Miguel A.; McClafferty, Julie A.: Description of the SHRP 2 Naturalistic Database and the Crash, Near-Crash, and Baseline Data Sets, 2016

Hart, D.: Die Nutzen/Risiko-Abwägung im Arzneimittelrecht (2005)
Hart, Dieter: Die Nutzen/Risiko-Abwägung im Arzneimittelrecht, in: Bundesgesundheitsblatt Gesundheitsforsch. Gesundheitsschutz (2), Issues 48, pp. 204–214, 2005

Hayward, J. C.: TTC (1972)
Hayward, John C.: Near-miss determination through use of a scale of danger, in: Highway Research Record, 1972

Hillenbrand, J. et al.: Situation assessment algorithm (2005)
Hillenbrand, Jörg; Kroschel, Kristian; Schmid, Volker (Eds.) Situation assessment algorithm for a collision prevention assistant, IEEE, 2005

Holder, M. et al.: Measurements revealing Challenges in Radar Sensor Modeling (2018)
Holder, Martin; Rosenberger, Philipp; Winner, Hermann; D'hondt, Thomas; Makkapati, Vamsi P.; Maier, Michael; Schreiber, Helmut; Magosi, Zoltan; Slavik, Zora; Bringmann, Oliver; others: Measurements revealing Challenges in Radar Sensor Modeling for Virtual Validation of Autonomous Driving, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018

Hugemann, W.: Driver reaction times in road traffic (2002)
Hugemann, Wolfgang: Driver reaction times in road traffic, Portoroz, Slovenia, 2002

Hydén, C.: Method for traffic safety evaluation (1987)
Hydén, Christer: The development of a method for traffic safety evaluation: The Swedish traffic conflicts technique
Lund Institute of Technology. Department of Traffic Planning and Engineering, Lund, 1987

Jansson, J.: Collision Avoidance Theory (2005)
Jansson, Jonas: Collision Avoidance Theory
Linköping University Electronic Press, 2005

Jonasson, J. K.; Rootzén, H.: Internal validation of near-crashes in naturalistic driving studies (2014)
Jonasson, Jenny K.; Rootzén, Holger: Internal validation of near-crashes in naturalistic driving studies, in: Accident Analysis & Prevention, Issues 62, pp. 102–109, 2014

Junietz, P. et al.: Metrik zur Bewertung der Kritikalität von Verkehrssituationen und -szenarien (2017)
Junietz, Philipp; Schneider, Jan; Winner, Hermann: Metrik zur Bewertung der Kritikalität von Verkehrssituationen und -szenarien, in: UNI DAS e.V (Ed.): 11. Workshop Fahrerassistenzsysteme, 2017

Junietz, P. et al.: Criticality Metric for the Safety Validation of Automated Driving (2018)
Junietz, P.; Bonakdar, F.; Klamann, B.; Winner, H.: Criticality Metric for the Safety Validation of Automated Driving using Model predictive Trajectory Optimization, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018

Junietz, P. et al.: Evaluation of Safety Validation Approaches (2018)
Junietz, P.; Wachenfeld, W.; Klonecki, K.; Winner, H.: Evaluation of Different Approaches to Address Safety Validation of Automated Driving, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018

Junietz, P. et al.: Gaining Knowledge on Automated Driving's Safety--The Risk-Free VAAFO Tool (2019)
Junietz, Philipp; Wachenfeld, Walther; Schönemann, Valerij; Domhardt, Kai; Tribelhorn, Wadim; Winner, Hermann: Gaining Knowledge on Automated Driving's Safety--The Risk-Free VAAFO Tool, in: Waschl, Harald; Kolmanovsky, Ilya; Willems, Frank (Eds.): Control Strategies for Advanced Driver Assistance Systems and Autonomous Driving Functions : Development, Testing and Verification, Springer International Publishing, Cham, 2019

Junietz, P. et al.: Macroscopic Safety Requirements for Highly Automated Driving (2019)
Junietz, Philipp; Steininger, Udo; Winner, Hermann: Macroscopic Safety Requirements for Highly Automated Driving, in: Transportation Research Record 2673(3), 2019

Käfer, E.: Situationsklassifikation und Bewegungsprognose in Verkehrssituationen (2013)
Käfer, Eugen: Situationsklassifikation und Bewegungsprognose in Verkehrssituationen mit mehreren Fahrzeugen, in: , 2013

Kalra et al.: How Many Miles of Driving Would It Take? (2016)
Kalra; Nidhi; Paddock; M, Susan: Driving to Safety: How Many Miles of Driving Would It

Take to Demonstrate Autonomous Vehicle Reliability?; http://www.rand.org/pubs/research_reports/RR1478.html, 2016, Access 15.04.2016

Kalra, N.; Groves, D. G.: The Enemy of Good (2017)
Kalra, Nidhi; Groves, David G.: The Enemy of Good, in: , 2017

Karlsson, R. et al.: Decision making for collision avoidance application (2004)
Karlsson, Rickard; Jansson, Jonas; Gustafsson, Fredrik: Model-based statistical tracking and decision making for collision avoidance application, in: American Control Conference, 2004. Proceedings of the 2004, 2004

Kinchin, G. H.: Design Criteria, Concepts and Features Important to Safety and Licensing. ANS (1979)
Kinchin, G. H.: Design Criteria, Concepts and Features Important to Safety and Licensing. ANS, in: ENS International Meeting on Fast Reactor Safety Technology, Seattle, Washington (19-23 August, 1979), 1979

Kletz, T. A.: Hazard analysis, its application to risks to the public at large (1978)
Kletz, T. A.: Hazard analysis, its application to risks to the public at large, in: Reliability Engineering (4), Issues 3, pp. 325–338, 1978

Koopman, P. et al.: Credible Autonomy Safety Argumentation (2019)
Koopman, Philip; Kane, Aaron; Black, Jen: Credible Autonomy Safety Argumentation, in: , 2019

Koopman, P.: Edge Cases and Autonomous Vehicle Safety (2019)
Koopman, Philip: Edge Cases and Autonomous Vehicle Safety, Safety-Critical Systems Symposium 2019, Bristol, 2019

Krajewski, R. et al.: The highd dataset (2018)
Krajewski, Robert; Bock, Julian; Kloeker, Laurent; Eckstein, Lutz: The highd dataset, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018

Kühn, M.; Hannawald, L.: Driver Assistance and Road Safety (2016)
Kühn, Matthias; Hannawald, Lars: Driver Assistance and Road Safety, in: Winner, Hermann et al. (Eds.): Handbook of Driver Assistance Systems: Basic Information, Components and Systems for Active Safety and Comfort, Springer International Publishing, Cham, 2016

Kusano, K. D.; Gabler, H. C.: Potential occupant injury reduction (2010)
Kusano, Kristofer D.; Gabler, Hampton C.: Potential occupant injury reduction in pre-crash system equipped vehicles in the striking vehicle of rear-end crashes, in: Annals of Advances in Automotive Medicine/Annual Scientific Conference, 2010

Lagarias, J. C. et al.: Convergence properties of the Nelder-Mead simplex method in low dimensions (1998)
Lagarias, Jeffrey C.; Reeds, James A.; Wright, Margaret H.; Wright, Paul E.: Convergence

properties of the Nelder-Mead simplex method in low dimensions, in: SIAM Journal on optimization (1), Issues 9, pp. 112–147, 1998

Liu, P. et al.: How Safe Is Safe Enough for Self-Driving Vehicles? (2018)
Liu, Peng; Yang, Run; Xu, Zhigang: How Safe Is Safe Enough for Self-Driving Vehicles?, in: Risk analysis, 2018

Mages, M. et al.: Intersection Assistance (2016)
Mages, Mark; Klanner, Felix; Stoff, Alexander: Intersection Assistance, in: Winner, Hermann et al. (Eds.): Handbook of Driver Assistance Systems: Basic Information, Components and Systems for Active Safety and Comfort, Springer International Publishing, Cham, 2016

Mohamed, M.; Saunier, N.: Motion Prediction Methods for Surrogate Safety Analysis (2013)
Mohamed, Mohamed; Saunier, Nicolas: Motion Prediction Methods for Surrogate Safety Analysis, in: Transportation Research Record: Journal of the Transportation Research Board (-1), Issues 2386, pp. 168–178, 2013

Najm, W. G. et al.: Analysis of light vehicle crashes and pre-crash scenarios (2003)
Najm, Wassim G.; Sen, Basav; Smith, John D.; Campbell, B. N.: Analysis of light vehicle crashes and pre-crash scenarios based on the 2000 General Estimates System, 2003

Nörenberg, R.: Effizienter Regressionstest von E/E-Systemen (2012)
Nörenberg, Ralf: Effizienter Regressionstest von E/E-Systemen nach ISO 26262, Issues 3, KIT Scientific Publishing, 2012

Oguchi, T.: Achieving safe road traffic—the experience in Japan (2016)
Oguchi, Takashi: Achieving safe road traffic—the experience in Japan, in: IATSS research (2), Issues 39, pp. 110–116, 2016

PEGASUS Symposium: Approach & Consistency (2017)
PEGASUS Symposium: Approach & Consistency; https://www.pegasusprojekt.de/files/tmpl/PDF-Symposium/01_Approach-and-Consistency.pdf, 2017, Access 29.08.2018

Popper, K. R.: Conjectures and refutations (1969)
Popper, Karl R. *.-1.: Conjectures and refutations, Routledge & Kegan, London, 1969

Risk & Reliability Associates Pty Ltd, Consulting Engineers: Risk and Reliability (2004)
Risk & Reliability Associates Pty Ltd, Consulting Engineers: Risk and Reliability - An Introductory Text, 5. Edition, Risk and Reliability Associates Pty Ltd, 2004

Rizaldi, A. et al.: A Formally Verified Checker of the Safe Distance Traffic Rules (2016)
Rizaldi, Albert; Immler, Fabian; Althoff, Matthias: A Formally Verified Checker of the Safe

Distance Traffic Rules for Autonomous Vehicles, in: NASA Formal Methods Symposium, 2016

Rodemerk, C. et al.: Development of a general criticality criterion (2012)
Rodemerk, Claas; Habenicht, Stefan; Weitzel, Alexander; Winner, Hermann; Schmitt, Thomas: Development of a general criticality criterion for the risk estimation of driving situations and its application to a maneuver-based lane change assistance system, in: Intelligent Vehicles Symposium (IV), 2012 IEEE, 2012

Rogers, E. M.: Diffusion of innovations (1983)
Rogers, Everett M.: Diffusion of innovations, 3. Edition, Free Press, New York, NY, 1983

SAE International Standard J3016: Taxonomy for Automated Driving Systems (2014)
SAE International Standard J3016: Taxonomy and Definitions for Terms related to On-Road Motor Vehicle Automated Driving Systems, SAE International, 2014

Sauerbier, J. et al.: Definition von Szenarien zur Absicherung automatisierter Fahrfunktionen (2019)
Sauerbier, Jan; Bock, Julian; Weber, Hendrik; Eckstein, Lutz: Definition von Szenarien zur Absicherung automatisierter Fahrfunktionen, in: ATZ - Automobiltechnische Zeitschrift (1), Issues 121, pp. 42–45, 2019

Schmidt, C.: Fahrstrategien zur Unfallvermeidung (2014)
Schmidt, Christian: Fahrstrategien zur Unfallvermeidung im Straßenverkehr für Einzel- und Mehrobjektszenarien, Schriftenreihe / Institut für Mess- und Regelungstechnik, Karlsruher Institut für Technologie, Issues 030, KIT Scientific Publishing, Karlsruhe, Baden, 2014

Schoenawa; Stefan: Critical Scenarios for Humand Drivers (2017)
Schoenawa; Stefan: Critical Scenarios for Humand Drivers, PEGASUS Symposium, 2017

Schöner, H.-P.: Challenges and Approaches for Testing of Highly Automated Vehicles (2014)
Schöner, Hans-Peter: Challenges and Approaches for Testing of Highly Automated Vehicles, CESA 3.0 - Congress on Automotive Electronic Systems, Paris, 2014

Schreier, M.: Bayesian environment representation, prediction, and criticality assessment (2016)
Schreier, Matthias: Bayesian environment representation, prediction, and criticality assessment for driver assistance systems, Fortschritt-Berichte VDI. Reihe 12, Verkehrstechnik797 Berichte aus dem Institut für Automatisierungstechnik und Mechatronik der TU Darmstadt, VDI Verlag GmbH, Düsseldorf, 2016

Schuldt, F.: Methodischer Test mit virtueller Umgebung (2017)
Schuldt, Fabian: Ein Beitrag für den methodischen Test von automatisierten Fahrfunktionen mit Hilfe von virtuellen Umgebungen
Technische Universität Carolo-Wilhelmina zu Braunschweig, 2017

Shalev-Shwartz, S. et al.: On a Formal Model of Safe and Scalable Self-driving Cars (2017)
Shalev-Shwartz, Shai; Shammah, Shaked; Shashua, Amnon: On a Formal Model of Safe and Scalable Self-driving Cars, in: arXiv preprint arXiv:1708.06374, 2017

Slovic, P. et al.: Facts and fears: Understanding perceived risk (1980)
Slovic, Paul; Fischhoff, Baruch; Lichtenstein, Sarah: Facts and fears: Understanding perceived risk, in: Societal risk assessment, Springer, 1980

Slovic, P.: The perception of risk (2011)
Slovic, Paul: The perception of risk, Earthscan, London [u.a.], 2011

Songchitruksa, P.; Tarko, A. P.: The EVT approach to safety estimation (2006)
Songchitruksa, Praprut; Tarko, Andrew P.: The extreme value theory approach to safety estimation, in: Accident Analysis & Prevention (4), Issues 38, pp. 811–822, 2006

Stählin, U.: Eingriffsentscheidung für ein Fahrerassistenzsystem zur Unfallvermeidung (2008)
Stählin, Ulrich: Eingriffsentscheidung für ein Fahrerassistenzsystem zur Unfallvermeidung, VDI-Verl., Düsseldorf, 2008

Starr, C.: Social benefit versus technological risk (1969)
Starr, Chauncey: Social benefit versus technological risk, in: Science, pp. 1232–1238, 1969

Starr, C.: Benefit-cost relationships in socio-technical systems (1971)
Starr, Chauncey: Benefit-cost relationships in socio-technical systems, in: Environmental aspects of nuclear power stations, 1971

Steininger, U. et al.: Validation of Assisted and Automated Driving (2016)
Steininger, U.; Schöner, Hans-Peter; Schiementz, Mark; Mazzega, Jens: Validation of Assisted and Automated Driving, crash.tech, Munich, 2016

Steininger, U.; Wech, L.: Wie sicher ist sicher genug? Sicherheit und Risiko zwischen Wunsch und Wirklichkeit (2013)
Steininger, U.; Wech, L.: Wie sicher ist sicher genug? Sicherheit und Risiko zwischen Wunsch und Wirklichkeit, in: VDI-Berichte (2204), 2013

Stipancic, J. et al.: Vehicle manoeuvers as surrogate safety measures (2018)
Stipancic, Joshua; Miranda-Moreno, Luis; Saunier, Nicolas: Vehicle manoeuvers as surrogate safety measures, in: Accident Analysis & Prevention, Issues 115, pp. 160–169, 2018

Stumper, D. et al.: Towards Characterization of Driving Situations (2016)
Stumper, Daniel; Knapp, Andreas; Pohl, Martin; Dietmayer, Klaus: Towards Characterization of Driving Situations via Episode-Generating Polynomials, in: Advanced Microsystems for Automotive Applications 2016, Springer, 2016

Stumper, D.; Dietmayer, K.: Towards Criticality Characterization of Situational Space (2018)
Stumper, Daniel; Dietmayer, Klaus: Towards Criticality Characterization of Situational

Space, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018

Tingvall, C.; Haworth, N.: Vision Zero (2000)
Tingvall, Claes; Haworth, Narelle: Vision Zero, in: 6th ITE International Conference Road Safety & Traffic Enforcement: Beyond, 2000

Toledo, T.; Zohar, D.: Modeling duration of lane changes (2007)
Toledo, Tomer; Zohar, David: Modeling duration of lane changes, in: Transportation Research Record (1), Issues 1999, pp. 71–78, 2007

U.S. Department of Transportation Federal Highway Administration: Fatality Rate per VMT (2013)
U.S. Department of Transportation Federal Highway Administration: FATALITY RATE PER 100 MILLION ANNUAL VMT, 2013

U.S. Department of Transportation NHTSA: Comparison of Fatality Rates (2016)
U.S. Department of Transportation NHTSA: Comparison of 2013 VMT Fatality Rates in U.S. States and in High-Income Countries, 2016

U.S. Department of Transportation NHTSA: A Vision for Safety 2.0 (2017)
U.S. Department of Transportation NHTSA: A Vision for Safety 2.0, 2017

U.S. Department of Transportation NHTSA: Fatality Analysis Reporting System (2017)
U.S. Department of Transportation NHTSA: Fatality Analysis Reporting System, 2017

Ulbrich, S. et al.: Defining and Substantiating the Terms Scene, Situation, and Scenario
Ulbrich, Simon; Menzel, Till; Reschka, Andreas; Schuldt, Fabian; Maurer, Markus: Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving, in: 2015 IEEE 18th International Conference on Intelligent Transportation Systems - (ITSC 2015)

Ulbrich, S.; Maurer, M.: Towards tactical lane change behavior planning for automated vehicles (2015)
Ulbrich, Simon; Maurer, Markus: Towards tactical lane change behavior planning for automated vehicles, in: Intelligent Transportation Systems (ITSC), 2015 IEEE 18th International Conference on, 2015

VDA: VDA 702 Situationskatalog E-Parameter nach ISO 26262-3 (2015)
VERBAND DER AUTOMOBILINDUSTRIE E. V.: VDA 702 Situationskatalog E-Parameter nach ISO 26262-3, 2015

Wachenfeld, W. et al.: The worst-time-to-collision metric for situation identification (2016)
Wachenfeld, Walther; Junietz, Philipp; Wenzel, Raphael; Winner, Hermann: The worst-

time-to-collision metric for situation identification, in: 2016 IEEE Intelligent Vehicles Symposium (IV), 2016

Wachenfeld, W. et al.: Safety Assurance based on an Objective Identification of Scenarios (2016)
Wachenfeld, Walther; Junietz, Philipp; Winner, Hermann; Themann, Philipp; Pütz, Andreas: Safety Assurance based on an Objective Identification of Scenarios, TRB Vehicle Automation Workshop 2016, San Francisco, 2016

Wachenfeld, W.; Winner, H.: Virtual Assessment of Automation in Field Operation (2015)
Wachenfeld, Walther; Winner, Hermann: Virtual Assessment of Automation in Field Operation. A New Runtime Validation Method, in: UNI DAS e.V (Ed.): 10. Workshop Fahrerassistenzsysteme, 2015

Wachenfeld, W.; Winner, H.: The Release of Autonomous Vehicles (2016)
Wachenfeld, Walther; Winner, Hermann: The Release of Autonomous Vehicles, in: Maurer, Markus et al. (Eds.): Autonomous Driving: Technical, Legal and Social Aspects, Springer Berlin Heidelberg, Berlin, Heidelberg, 2016

Wachenfeld, W. H.: Dissertation, How Stochastic can Help to Introduce AD (2017)
Wachenfeld, Walther H. K.: How Stochastic can Help to Introduce Automated Driving, Dissertation
Technische Universität Darmstadt, Darmstadt, 2017

Webb, G. A.; McLean, A. S.: Insignificant levels of dose (1977)
Webb, G. A.; McLean, A. S.: Insignificant levels of dose, 1977

Weitzel, D. A.: Kontrollierbarkeit nicht situationsgerechter Reaktionen (2013)
Weitzel, Dirk A.: Objektive Bewertung der Kontrollierbarkeit nicht situationsgerechter Reaktionen umfeldsensorbasierter Fahrerassistenzsysteme, VDI-Verlag, 2013

Winner, H. et al.: Maße für den Sicherheitsgewinn von Fahrerassistenzsystemen (2013)
Winner, Hermann; Geyer, Sebastian; Sefati, Mohsen: Maße für den Sicherheitsgewinn von Fahrerassistenzsystemen, in: Winner, Hermann; Bruder, Ralph (Eds.): Maßstäbe des sicheren Fahrens. 6. Darmstädter Kolloquium Mensch + Fahrzeug, Ergonomia Verlag, Stuttgart, 2013

Winner, H. et al.: Validation and Introduction of Automated Driving (2018)
Winner, Hermann; Wachenfeld, Walther; Junietz, Phillipp: Validation and Introduction of Automated Driving, in: Winner, Hermann; Prokop, Günther; Maurer, Markus (Eds.): Automotive Systems Engineering II, Springer International Publishing, Cham, 2018

Winner, H.; Wachenfeld, W.: Risk-limited Introduction of Automated Driving, (2017)
Winner, Hermann; Wachenfeld, Walther: Risk-limited Introduction of Automated Driving,, TECHROAD (Keynote), Illertissen, 2017

Winner, H.; Weitzel, A.: Die Freigabefalle des autonomen Fahrens (2011)
Winner, Hermann; Weitzel, Alexander: Die Freigabefalle des autonomen Fahrens, in: Winner, Hermann; Bruder, Ralph (Eds.): Mensch und Fahrzeug, tuprints, Darmstadt, 2011

Wolff, C.: Grundlegendes zum Bremsvorgang (2017)
Wolff, Claus: Grundlegendes zum Bremsvorgang, in: Bremsenhandbuch, Springer, 2017

Yi, B. et al.: Real time integrated vehicle dynamics control and trajectory planning (2016)
Yi, Boliang; Gottschling, Stefan; Ferdinand, Jens; Simm, Norbert; Bonarens, Frank; Stiller, Christoph: Real time integrated vehicle dynamics control and trajectory planning with MPC for critical maneuvers, in: Intelligent Vehicles Symposium (IV), 2016 IEEE, 2016

Yi, B.: Integrated planning and control for collision avoidance systems (2018)
Yi, Boliang: Integrated planning and control for collision avoidance systems, Schriftenreihe / Institut für Mess- und Regelungstechnik, Karlsruher Institut für Technologie (KIT), Issues 039, KIT Scientific Publishing, Karlsruhe, 2018

Zahar, E. G.: Falsifiability (2007)
Zahar, Elie G.: Falsifiability, in: Keuth, Herbert (Ed.): Karl Popper - Logik der Forschung, Akad.-Verl., Berlin, 2007

Zhao, D. et al.: Evaluation of lane-change scenarios based on importance sampling techniques (2017)
Zhao, Ding; Lam, Henry; Peng, Huei; Bao, Shan; LeBlanc, David J.; Nobukawa, Kazutoshi; Pan, Christopher S.: Accelerated evaluation of automated vehicles safety in lane-change scenarios based on importance sampling techniques, in: IEEE Transactions on intelligent transportation systems (3), Issues 18, pp. 595–607, 2017

Own Publications

Wachenfeld, Walther, Junietz, Philipp, Wenzel, Raphael, Winner, Hermann (2016) The worst-time-to-collision metric for scene identification. In 2016 IEEE Intelligent Vehicles Symposium (IV), 2016, pp. 729–734.

Wachenfeld, Walther; Junietz, Philipp; Winner, Hermann; Themann, Philipp; Pütz, Andreas (2016): Safety Assurance Based on an Objective Identification of Scenarios. One Approach of the PEGASUS – Project. TRB-AUVSI-Automated Vehicle Symposium. San Francisco, 20.07.2016.

Junietz, Philipp, Schneider, Jan, Winner, Hermann (2017): Metrik zur Bewertung der Kritikalität von Verkehrssceneen und -szenarien. In 11. Workshop Fahrerassistenzsysteme, Walting, 2017.

Junietz, Philipp, Winner, Hermann (2017): Identifying Critical Traffic Scenarios for Test Case Derivation of Automated Driving, WKM Symposium, Braunschweig, 20.05.2017.

Junietz, Philipp; Wachenfeld, Walther; Klonecki, Kamil; Winner, Hermann (2018): Evaluation of Different Approaches to Address Safety Validation of Automated Driving, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018

Junietz, Philipp; Bonakdar, Farid; Klamann, Björn; Winner, Hermann (2018): Criticality Metric for the Safety Validation of Automated Driving using Model Predictive Trajectory Optimization, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018

Klonecki, Kamil; Wachenfeld, Walther; Junietz, Philipp; Winner, Hermann (2018): Bewertung verschiedener Ansätze zur Verifikation und Validierung des automatisierten Fahrens – Die Eignung von Ansätzen für die Argumentation von Sicherheit. VDI-Tagung, Wolfsburg, 2018.

Junietz, Philipp, Wachenfeld, Walther, Schönemann, Valerij,Domhardt, Kai, Tribelhorn, Wadim, Winner, Hermann (2019): Gaining Knowledge on Automated Driving's Safety--The Risk-Free VAAFO Tool. In Control Strategies for Advanced Driver Assistance Systems and Autonomous Driving Functions: Development, Testing and Verification, H. Waschl, I. Kolmanovsky and F. Willems, eds. Springer International Publishing, Cham, 2019, pp. 47–65.

Junietz, Philipp; Steininger, Udo; Winner, Hermann (2019): Macroscopic Safety Requirements for Highly Automated Driving, in: Transportation Research Record, 2019.

Supervised Theses

Choung, Zihoung: Entwicklung einer Simulationsumgebung zum automatisierten Valet Parken in IPG CarMaker, Bachelorthesis No.: 1250/15.

Wenzel, Raphael: Implementierung und Bewertung des Konzepts „Minimale Zeit bis zur Kollision“ zur Kritikalitätsbewertung der Fahrt im Straßenverkehr, Bachelorthesis No.: 1251/15.

Münch, Manuel: Bewertung von Systemarchitekturen für das automatisierte Valet-Parken, Masterthesis No.: 600/15.

Schönemann, Valerij: Entwicklung und Implementierung der Funktion des automatisierten Valet-Parkens in einem Versuchsträger, Masterthesis No.: 602/15.

Stobert, Arthur: Vergleich und Bewertung von Verfahren zur Berechnung möglicher Ausweichtrajektorien im longitudinalen Verkehr, Bachelorthesis No.: 1270/16.

Prjapati, Nabin: Entwicklung und Anwendung einer objektiven Bewertungsmethode zur manuellen Bewertung der Kritikalität von Verkehrsszenarien Bachelorthesis No.: 1284/16.

Schneider, Jan: Entwicklung und Anwendung von Metriken zur Identifikation kritischer Szenarien für die Verwendung im Freigabeprozess automatisierter Fahrzeuge, Masterthesis No.: 616/16.

Haas, Julia: Anwendung und Analyse von Kritikalitätsmetriken in Autobahn Szenarien, Bachelorthesis No.: 1294/17.

Konradt, Lukas: Berechnung und Kritikalitätsbewertung von Fahrstreifenwechselmanövern auf Autobahnen, Masterthesis No.: 639/17.

Schmidt-Wolf, Melanie: Anwendung und Vergleich verschiedener statistischer Methoden sowie maschineller Lernverfahren zur Kalibrierung von Kritikalitätsmetriken, Masterthesis No.: 641/17.

Bulut, Serdar: Definition and Quantification of the Complexity Experienced by Autonomous Vehicles in their Environment and Driving Task, Masterthesis No.: 706/18.

Lehmann, Valentin: Ermittlung von Risikoanforderungen für automatisierte Fahrzeuge in verschiedenen Ländern und Anwendung als Randbedingungen für die risiko-limitierte Einführung, Bachelorthesis No.: 1327/18.

Wang, Qin: Weiterentwicklung einer auf numerischer Optimierung basierender Kritikalitätsmetrik durch einen Kritikalitätstensor, Masterthesis No.: 708/18.

Wu, Jiaming: Data Analysis of Recorded Autobahn Data using Criticality Metrics, Bachelorthesis No.: 1334/18.

Lebenslauf

Persönliche Daten

Name:	Philipp Matthias Junietz
Geburtsdatum:	04.03.1989
Geburtsort:	Münster

Schulzeit

2001 – 2008	Gymnasium „In der Wüste“, Osnabrück

Zivildienst

2008 – 2009	Heilpädagogischer Sonderkindergarten, Osnabrück

Studium

2009 – 2015	Studium Allgemeiner Maschinenbau Technische Universität Darmstadt, Abschluss M. Sc.
2013 – 2015	Master Automotive Engineering Tongji University, Abschluss mit Double Degree

Berufstätigkeit

2015 – 2019	Wissenschaftlicher Mitarbeiter am Fachgebiet Fahrzeugtechnik der Technischen Universität Darmstadt