

Global-optimale, datengetriebene Applikation längs- und querführender Fahrfunktionen

Zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften (Dr.-Ing.)

von der KIT-Fakultät für
Elektrotechnik und Informationstechnik des
Karlsruher Instituts für Technologie (KIT)

angenommene

DISSERTATION

von

M.Sc. Jacob Langner

geb. am 23.09.1990 in Stralsund

Tag der mündlichen Prüfung:

Hauptreferent:

Korreferent:

18.07.2025

Prof. Dr.-Ing. Eric Sax

Prof. Dr. Stefan Nickel

Kurzfassung

Fahrerassistenzsysteme und hochautomatisierende Fahrfunktionen haben das Potenzial, die Verkehrssicherheit zu erhöhen, indem sie Verkehrsunfälle reduzieren und gleichzeitig den Fahrer entlasten. Zudem tragen sie zur Verbesserung des Fahrkomforts und der Verkehrseffizienz bei.

Diese Dissertation untersucht die Parametrierung solcher Systeme und identifiziert Schwächen im aktuellen Entwicklungsprozess. Ein zentrales Problem ist, dass Tests im realen Straßenverkehr durchgeführt werden müssen und diese aufgrund begrenzter Ressourcen nur Stichproben liefern, die den gesamten Straßenverkehr nicht korrekt repräsentieren. Zudem verhindern sequentiell durchgeführte Tests und eine fehlende Systematik in der Variierung der Parameter die Vergleichbarkeit der Ergebnisse. Weder bestehende Prozesse noch wissenschaftliche Ansätze adressieren diese Herausforderungen ausreichend.

Deshalb wird in dieser Dissertation ein datengetriebener, global-optimaler Prozess für die Parametrierung von Fahrerassistenzsystemen und hochautomatisierenden Fahrfunktionen entwickelt und evaluiert. Dieser nutzt aufgezeichnete Testfahrten aus dem realen Straßenverkehr in einem virtuellen Framework und automatisiert die Suche nach optimalen Parametern mithilfe eines genetischen Optimierungsalgorithmus und objektiver Fahrdynamikmetriken. Mit Hilfe von Gewichtungsfaktoren werden Bias in den Stichproben entfernt und die Repräsentativität bezüglich der Grundgesamtheit hergestellt. Die Evaluation der prototypischen Umsetzung zeigt, wie Verzerrungen in den zugrundeliegenden Daten und regionale Unterschiede die Aussagekraft kleiner Stichproben beeinflussen und dadurch die Parametrierung im realen Straßenverkehr erschweren. Durch Interpretation und Gewichtung während der Datenanalyse wird eine global-optimale Parametrierung

erreicht, die eine Verbesserung des Fahrkomforts und der Fahrdynamikeigenschaften der Systeme ermöglicht.

Abstract

Driver assistance systems and highly automated driving functions offer great potential in the form of a significant reduction in traffic accidents and thus an increase in road safety, while at the same time relieving the driver and creating completely new mobility solutions with increasing automation. They also improve driving comfort and efficiency in traffic.

This phd thesis examines the parameter setting of such systems and identifies potential for improvement in the application process. The challenges in the application in real traffic are systematically researched. It is shown that economically feasible scopes of real world traffic tests always represent random samples with bias in the sense of the entire road traffic and are therefore not representative. In addition, the comparability of the functional quality of sequential road tests is not given. There is no systematic approach to parameter variation. Current application processes and scientific approaches do not sufficiently take these challenges into account.

In this phd thesis, a globally optimized, data-driven application process is developed, prototypically implemented and evaluated. For this purpose, test drives carried out in a virtual application framework are used and optimal parameter sets are automatically searched for using a modified cuckoo search. During the evaluation, the influence of bias and country-specific conditions on small samples of road tests is shown, thereby preventing an optimal application. By means of virtual data analysis and weighting of the samples against a population, a globally optimal application is ensured.

Danksagung

Mein aufrichtiger Dank gilt zunächst meinem Doktorvater Prof. Dr.-Ing. Eric Sax für die Möglichkeit zur Promotion, die fachliche Begleitung über all die Jahre sowie die vielen konstruktiven Diskussionen. Ich danke Dir für die wertvollen Lehren, das ehrliche Feedback und die stete Unterstützung, die den Weg dieser Dissertation entscheidend geprägt haben. Ebenso danke ich Prof. Dr. Stefan Nickel herzlich für die Übernahme des Korreferats und die damit verbundenen Impulse und Diskussionen zur Optimierung, die neue Perspektiven eröffnet und wichtige Anregungen geliefert haben.

Ein großer Dank gilt den Kolleginnen und Kollegen der ADAS- und HAF-Entwicklung bei Porsche, insbesondere Marc Holzäpfel und Kai-Lukas Bauer, sowie meinen beiden Doktorandenkollegen Carl Esselborn und Philip Elspas. Danke für die spannenden Projekte, die inspirierenden Gespräche und die vielfältigen Einblicke, die ich durch Euch gewinnen durfte. Ohne Eure Unterstützung und die gemeinsamen Erfahrungen wäre diese Arbeit in dieser Form nicht möglich gewesen.

Meinen Kolleginnen und Kollegen vom FZI Forschungszentrum Informatik, insbesondere im Bereich ESS, danke ich für die außergewöhnlich bereichernde und kollegiale Zeit. Durch Euch bin ich jeden Tag mit Freude zur Arbeit gekommen - und Ihr habt mir die Energie gegeben, auch in intensiven Phasen der Dissertation dranzubleiben. Ein ganz besonderer Dank geht an Johannes Bach. Lieber Johannes, danke Dir für den „Raketenstart“ am FZI, die Einführungen in das wissenschaftliche und projektspezifische Arbeiten, die gemeinsamen Diskussionen und Veröffentlichungen, langen Abende am FZI und die daraus entstandene

Freundschaft. Mein Dank gilt auch Stefan Otten, der mich ans FZI geholt, mir Vertrauen geschenkt und meine Entwicklung dort wesentlich gefördert hat.

Danken möchte ich auch Lukas Köhrer für die hervorragende Zusammenarbeit im Porsche-Kontext, die unzähligen gemeinsamen Fahrten nach Mönshheim und das Beibringen der Grundlagen der Regelungstechnik sowie Romy Pohl für die gemeinsame Forschung zu Bias in Straßenverkehrstests und die inspirierende Zusammenarbeit. Ein riesiger Dank geht an Carl Philipp Hohl und Max Grobbel für die gemeinsamen Radtouren, die die notwendige Ablenkung und Entspannung geschaffen haben, um dann wieder motiviert an der Dissertation zu arbeiten.

Besonderer Dank gebührt Christian King und Lennart Ries für die gemeinsamen Diss-Urlaube und Schreib-Sessions, den damit verbundenen fachlichen Austausch, das gegenseitige Feedback sowie das unermüdliche Motivieren. Danke, Christian, für Dein wertvolles Feedback zur Struktur und Methodik meiner Ausarbeitungen - und Lennart, für die kreativen Brainstormings, das Story-Building und das Formel-Korrigieren sowie ganz allgemein für die gemeinsame Zeit am FZI und das gegenseitige Freiräume-Schaffen während der Doppelbelastung.

Mein herzlicher Dank gilt zudem meiner Familie und meinen Freunden für ihre stetige Unterstützung, Geduld und den Rückhalt während der gesamten Promotionszeit. Und von ganzem Herzen danke ich meiner Freundin Lara Schweickart für ihre unglaubliche Unterstützung in der Endphase meiner Promotion. Liebe Lara, Danke Dir für die mentale Stärke und Dein Verständnis, für den Zuspruch und das Mitfiebern, das Korrekturlesen, das Üben des Vortrags und die Begleitung durch die gesamte Prüfungsphase! Ohne Dich wäre das alles nicht möglich gewesen.

Berlin, im September 2025

Jacob Langner

Inhaltsverzeichnis

Kurzfassung	i
Abstract	iii
Danksagung	v
1 Einleitung	1
1.1 Motivation	1
1.2 Herausforderungen in der Applikation von ADAS und HAF-Funktionen	7
1.3 Forschungsfragen	9
1.4 Umfeld	10
1.5 Aufbau der Dissertation	10
2 Produktentwicklung - Wertschöpfung in der Automobilindustrie	13
2.1 Das Produktportfolio - Strukturierung und Gruppierung von Marktsegmenten	14
2.2 Fahrzeugplattformen - Wiederverwendung und Skaleneffekte . . .	18
2.3 Der Produktentstehungsprozess - Entwicklung von Fahrzeugmodellen	21
2.4 Funktionsentwicklungsprojekte - Wertschöpfung durch neue Funktionalität	23
2.4.1 Vorgehensmodelle	24
2.4.2 Das V-Modell	26
2.4.3 Fahrerassistenzsysteme	29
2.5 Qualitätssicherung und -prüfung	32

3	Qualifikation und Integration von ADAS und HAF-Funktionen	37
3.1	Integrationsstufen	39
3.2	Qualifikation entlang der Integrationsstufen	40
3.2.1	Fahrzeugtests auf dem Prüfgelände	42
3.2.2	Fahrzeugtests im realen Straßenverkehr	42
3.2.3	Shadow Mode Testing	44
3.3	Applikation auf Gesamtsystemebene	45
3.3.1	Ziel der Applikation	45
3.3.2	Bewertung der Applikation	46
3.3.3	Durchführung der Applikation	49
4	Datengetriebene Entwicklung	53
4.1	Messdaten im Kontext der automobilen Softwareentwicklung . .	54
4.2	Kollektion von Fahrzeugmessdaten	56
4.3	Verwendung und Mehrwert der Fahrzeugmessdaten	59
4.4	Analyse und Strukturierung der aufgezeichneten Fahrzeugtests . .	60
4.4.1	Fahrzeugtests als Stichproben einer Grundgesamtheit . .	61
4.4.2	Bias	62
4.4.3	Repräsentativität einer Stichprobe	63
5	Herausforderungen und Chancen in der Applikation von ADAS und HAF-Funktionen mittels Fahrzeugtests	69
5.1	Herausforderungen in der Applikation von ADAS und HAF . . .	70
5.1.1	Vergleichbarkeit von unterschiedlichen Erprobungsfahrten	70
5.1.2	Repräsentativität von Applikations- und Erprobungsfahrten	74
5.1.3	Prozess- und strukturbedingte Bias in Erprobungsfahrten .	82
5.1.4	Fehlende Systematik in der Parametervariation	87
5.2	Related Work zu den Herausforderungen in der Applikation von ADAS und HAF	89
5.2.1	Simulative Applikation von ADAS und Hochautomatisiertes Fahren (HAF)-Funktionen	89
5.2.2	Nutzerzentrierte Applikation von ADAS	91
5.2.3	Applikation mittels datenbasierter und lernender Verfahren	93
5.2.4	Applikation durch Vehicle in the Loop Verfahren	94
5.3	Diskussion und Potenziale	94

5.4	Anforderungen an die Applikation von ADAS und HAF-Funktionen	96
6	Global-optimale, datengetriebene Applikation von ADAS und HAF-Funktionen	99
6.1	Grundidee der global-optimalen, datengetriebenen Applikation . .	99
6.2	Datenmodell: Daten- und Bedeutungsebene von aufgezeichneten Straßenverkehrstests	102
6.3	Identifikation von Szenarien in Straßenverkehrstests	107
6.3.1	Identifikation von Szenarien mit manöverartigen Fahraufgaben	109
6.3.2	Identifikation von Szenarien mit kontinuierlichen Fahraufgaben	111
6.3.3	Attribution von Fahrtabschnitten	112
6.4	Methodik zur repräsentativen Bewertung von Straßenverkehrstests	115
6.4.1	Definition der Grundgesamtheit	116
6.4.2	Aggregation der Fahrtabschnitte über den Szenario-Attributraum	118
6.4.3	Gewichtung der Subszzenarien	120
6.4.4	Substitution der fehlenden Subszzenarien	122
6.4.5	Aggregation der Subszzenarien	126
6.5	Prozess der global-optimalen, datengetriebenen Applikation . . .	127
6.5.1	Applikationsvorbereitung (A)	128
6.5.2	Applikationsdurchführung (B)	131
6.5.3	Datenaufzeichnung (C)	133
6.5.4	Virtuelle Analyse (D)	133
6.5.5	Parametersätze festlegen oder anpassen (A.2)	142
7	Prototypische Umsetzung am Beispiel einer prädiktiven Längsregelung	143
7.1	Use Case	143
7.2	Datengrundlage	143
7.2.1	Trainings- und Validierungsdatensätze	144
7.3	Analyse- und Bewertungsmethodik	145
7.3.1	Abdeckung des Parameterraums	145

7.3.2	Bewertung der Fahraufgabe Kurvendurchfahrt	145
7.3.3	Aggregation der Fahrtabschnitte im Szenario-Attributraum / Analyse einer einzelnen Applikation	147
7.4	Ablauf der virtuellen Applikation	147
8	Evaluation und Diskussion der prototypischen Umsetzung	153
8.1	Evaluation der Unterteilung des Szenario-Attributraums in Subszenarien und deren gewichtete Aggregation	154
8.2	Bewertung der Repräsentativität einer Applikation	161
8.2.1	Evaluation statistischer Kennzahlen als Indikatoren für die Repräsentativität	161
8.2.2	Evaluation der Subszenario-Abdeckung als Indikator für die Repräsentativität	164
8.3	Evaluation Methodik und Gesamtprozess	166
8.4	Anforderungserfüllung	168
8.5	Erkenntnisgewinn zu den gestellten Forschungsfragen	170
8.5.1	Vergleichbarkeit sequentieller Parameteränderungen	170
8.5.2	Aussagekraft kleiner Stichproben	171
8.5.3	Integration in den Entwicklungsprozess	172
9	Zusammenfassung und Ausblick	175
9.1	Zusammenfassung	175
9.2	Ausblick	177
A	Anhang	179
A.1	Grafiken	179
A.1.1	Variante des Produktentstehungsprozess (PEP) von Porsche Engineering	179
A.1.2	Erprobungsplanung - direkt und adaptiert	179
A.2	Optimierungsalgorithmus	181
A.2.1	Funktionsweise Modified Cuckoo Search	181
A.2.2	Hyperparameter der Modified Cuckoo Search	184
A.3	Kalibrierung und Tuning des virtuellen Applikationsframeworks	184
A.3.1	Diskretisierung des Attributraums	184
A.4	Beschreibung der Umsetzung	191

A.4.1	Datenupload & -aufbereitung	191
A.4.2	Framework zur Analyse und Aggregation von Fahrzeugmessdaten	192
A.5	Weitere Ergebnisse	201
Abbildungsverzeichnis		209
Tabellenverzeichnis		215
Literaturverzeichnis		217
Eigene Veröffentlichungen		233
Betreute Abschlussarbeiten		237

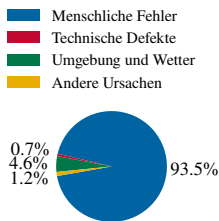
1 Einleitung

1.1 Motivation

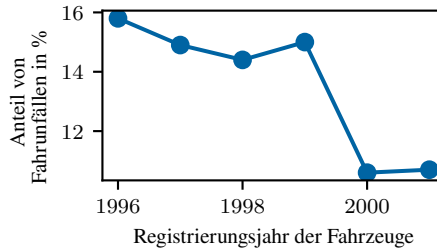
Die Automobilindustrie durchläuft seit den 1970er Jahren einen tiefgreifenden technologischen Wandel vom traditionellen Maschinenbau zur Herstellung und Integration elektrisch-elektronischer Systeme [1]. Der Anteil an elektrisch-elektronischen Systemen im Fahrzeug nimmt dabei seit Jahrzehnten stetig zu [2].

Ein wesentlicher Treiber dieses technologischen Wandels ist die Einführung von Fahrerassistenzsystemen und automatisierten Fahrfunktionen, die Ende der 1970er Jahre mit der Einführung des Anti-Blockier-Systems (ABS) begann und heutzutage eine immer entscheidendere Rolle bei den Kaufentscheidungen der Kunden spielt [3, 4]. Diese Systeme erhöhen die Sicherheit im Straßenverkehr, die Energieeffizienz der Fahrzeuge sowie den Fahrkomfort [5–7]. Gleichzeitig verschärfen sich die regulatorischen Anforderungen hinsichtlich der Sicherheit, Leistungsfähigkeit und Zuverlässigkeit von Fahrerassistenzsystemen und insbesondere automatisierten Fahrfunktionen [8].

Sowohl in den USA als auch in Deutschland werden über 90% der Unfälle im Straßenverkehr durch menschliche Fehler verursacht (vgl. Abbildung 1.1a) [9, 10]. Assistenzsysteme unterstützen den Fahrer bei der Fahraufgabe und haben damit das Potenzial, die Sicherheit im Straßenverkehr zu erhöhen (vgl. Abbildung 1.1b) [5]. Laut einer Studie der Foundation for Traffic Safety (FTS) werden derzeit verfügbare Fahrerassistenzsysteme 37 Mio. Verkehrsunfälle in den USA bis 2050 verhindern [7]. Ein enormes Potenzial, das durch die fortschreitende technologische Entwicklung der Assistenzsysteme ausgebaut werden kann.



(a) Unfallursachen nach GIDAS



(b) Unfallzahlen von Mercedes Benz

Abbildung 1.1: Reduktion von Unfällen durch die Einführung eines serienmäßigen Elektronisches Stabilitätsprogramm (ESP)-Systems in jedem Fahrzeug bei Mercedes-Benz im Jahr 1999 (in Anlehnung an [9])

Hochautomatisierte bzw. autonome Fahrfunktionen stellen die konsequente Weiterentwicklung von Fahrerassistenzsystemen dar. Sie übernehmen die Fahrzeugführung zumindest unter bestimmten Bedingungen und entbinden den Fahrer vollständig von der Fahraufgabe. Die Markteinführung autonomer Fahrzeuge scheint im öffentlichen Diskurs in immer greifbarere Nähe zu rücken. Der tatsächliche Zeitpunkt für die Einführung solcher Systeme ist unter Experten jedoch höchst umstritten. Die Prognosen gehen stark auseinander. So prognostizierte der frühere US-amerikanische Verkehrsminister (eng. US Secretary of Transportation) in 2016 einen landesweiten Einsatz von autonomen Robotaxis für das Jahr 2021 [11]. Andere Experten, wie Chris Urmson, Executive bei Aurora, sind skeptischer und gehen von 30 Jahren und mehr aus bis sich autonome Fahrzeuge im Massenmarkt etabliert haben [12]. Eine im Dezember 2021 veröffentlichte Studie von McKinsey, in der 75 Experten aus der Transport- und Automobilbranche befragt wurden, ergibt ein differenzierteres Bild (vgl. Abbildung 1.2). Autonome Robotaxis im urbanen Bereich werden von den Experten zwischen Ende 2026 und Anfang 2030 erwartet [13]. In den USA gibt es bereits seit 2020 Anbieter von Robotaxi-Services, die allerdings den Betriebsbereich einschränken und auf einzelne Städte mit günstigen Wetter- und Klimabedingungen limitieren. So hat Waymo als erstes Unternehmen im Oktober 2020 seinen Robotaxi-Service in Phoenix, Arizona

gestartet [14]. Seitdem sind sie nur nach San Francisco, Los Angeles und Austin expandiert [15].

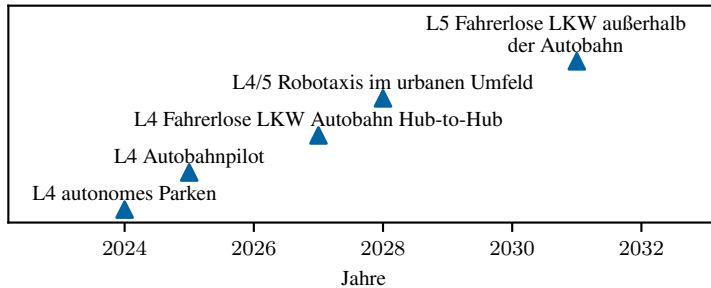


Abbildung 1.2: McKinsey Umfrage: Einschätzung von 75 Experten der Transport- und Automobilindustrie aus dem Jahr 2021 zur kommerziellen Einführung von hochautomatisierten Fahrfunktionen (in Anlehnung an [13])

Das automatisierte Fahren gilt neben der zunehmenden Vernetzung und der Elektrifizierung der Fahrzeuge als ein entscheidender Trend in der Automobilindustrie. Die heutigen Assistenzsysteme spielen dabei eine wichtige Rolle, da sie als Wegbereiter für hochautomatisierte Systeme dienen [16]. Zur Klassifikation automatisierter Fahrsysteme hat die Society of Automotive Engineers (SAE) den Standard J3016 veröffentlicht [17], der sechs Stufen des automatisierten Fahrens definiert (vgl. Tabelle 1.1).

SAE Stufe	Name	Dynamische Fahraufgabe (eng. Dynamic Driving Task)		Rückfallebene für die dynamische Fahraufgabe	Betriebsbereich
		Durchführung	Überwachung		
Der Fahrer übernimmt Teile bzw. die gesamte dynamische Fahraufgabe.					
0	keine Automatisierung	Fahrer	Fahrer	Fahrer	n/a
1	assistiertes Fahren	Fahrer und System	Fahrer	Fahrer	beschränkt
2	teilautomatisiertes Fahren	System	Fahrer	Fahrer	beschränkt
Das automatisierende System übernimmt (während es aktiv ist) die gesamte dynamische Fahraufgabe.					
3	bedingt automatisiertes Fahren	System	System	übernahmebereiter Nutzer	beschränkt
4	hochautomatisiertes Fahren	System	System	System	beschränkt
5	autonomes Fahren	System	System	System	unbeschränkt

Tabelle 1.1: Die 6 Stufen des automatisierten Fahrens laut SAE J3016 [17].

Aktuell auf dem Markt vorhandene Systeme zählen zu den SAE Level 1-3 Systemen, d.h. es handelt sich dabei um assistierende bzw. teil- oder bedingt automatisierende Systeme (vgl. Abbildung 1.3). Der Schritt hin zu SAE Level 3 mit einem umfassenden Betriebsbereich bereitet große Herausforderungen, da der Fahrer als überwachende Instanz entfällt und sich das System selbst über redundante Auslegung überwachen muss. Die Haftung obliegt dann dem System bzw. dem Hersteller des Systems und kann nicht länger auf den Fahrer übertragen werden.

OEM	AD-Level	Betriebsbereich/ Operational Design Domain (ODD)				Markteinführung				Modelle
										
MB	3	Stau auf Autobahn	Autobahn	Stadt	Überland	2023	-	2022	-	S-Klasse, EQS
BMW	3**	Stau auf Autobahn	Autobahn	Stadt	Überland	-	-	2023	-	7er
	2+	Stau auf Autobahn	Autobahn	Stadt	Überland	2022	-	2023	-	7er, 5er
Honda	(3)*	Stau auf Autobahn	Autobahn	Stadt	Überland	-	-	-	2021	Legend
TESLA	2+	Stau auf Autobahn	Autobahn	Stadt	Überland	2022	-	-	-	Model S, X, 3, Y
GM	2+	Stau auf Autobahn	Autobahn	Stadt	Überland	2022	-	-	-	z.B. Lyriq, Silverado
	2+**	Stau auf Autobahn	Autobahn	Stadt	Überland	2023	-	-	-	Celestiq
Ford	2+	Stau auf Autobahn	Autobahn	Stadt	Überland	2022	-	-	-	z.B. Mach-E, F-150
XPENG	(2+)*	Stau auf Autobahn	Autobahn	Stadt	Überland	-	2022	-	-	P5, P7i, G9
BAIC	(2+)*	Stau auf Autobahn	Autobahn	Stadt	Überland	-	2022	-	-	Alpha-S HI

Quelle: CAM | * Stark eingeschränkte Verfügbarkeit, ** Bis Ende 2023 angekündigt

Abbildung 1.3: Serienverfügbarkeit von SAE Level 2+ und SAE Level 3 Systemen (Quelle: [16]).

In den Jahren 2021 und 2023 wurden erste SAE Level 3 Systeme von Automobilherstellern durch das Kraftfahrtbundesamt (KBA) auf deutschen Straßen zugelassen. Der SAE Level 3 *Drive Pilot* von Mercedes Benz wurde am 09. Dezember 2021 nach mehreren Verzögerungen zugelassen [18, 19]. BMW kündigte am 10. November 2023 den SAE Level 3 *Personal Pilot* für den neuen 7er BMW an [20]. Der von Audi bereits im Jahr 2017 angekündigte SAE Level 3 *Staupilot* wurde für den Audi A8 erst verschoben und im Jahr 2020 dann vollständig zurückgezogen [21].

Für die Realisierung hochautomatisierter Fahrfunktionen werden im Fahrzeug für die Planung und Steuerung exakte Informationen über das Fahrzeugumfeld benötigt. Zur Umfeldwahrnehmung wird eine stetig steigende Anzahl und Varianz an Sensoren im Fahrzeug verbaut. In heutigen Fahrzeugen werden Mono- und Stereokameras, Mittel- und Langstrecken-Radarsysteme, Ultraschallsensoren und in der Luxusklasse mit SAE Level 3 Funktionen auch Lidar-Systeme verwendet [22]. Zusätzlich werden Online-Dienste und Backend-Services, wie zum Beispiel elektronische Karten oder aktuelle Verkehrsinformationen, integriert. Die Erweiterung des Funktionsumfangs führt zu einem Anstieg der Fahrsituationen, die durch die Fahrfunktionen durchgeführt und damit durch die Umfeldwahrnehmung erfasst werden müssen. Die Komplexität der Systeme, die in allen Fahrsituationen geeignet reagieren müssen, steigt [23]. Dies zeigt sich unter anderem in der steigenden Anzahl an Steuergeräten und Codezeilen (vgl. Abbildung 1.4) im Fahrzeug. Die redundante Sensorik muss verarbeitet und daraus ein konsistentes Modell der wahrgenommenen Fahrzeugumgebung erstellt werden. Auf Basis dieses Modells müssen immer komplexere Fahraufgaben geplant und durchgeführt werden, wie zum Beispiel die kombinierte Längs- und Querführung des Fahrzeugs.

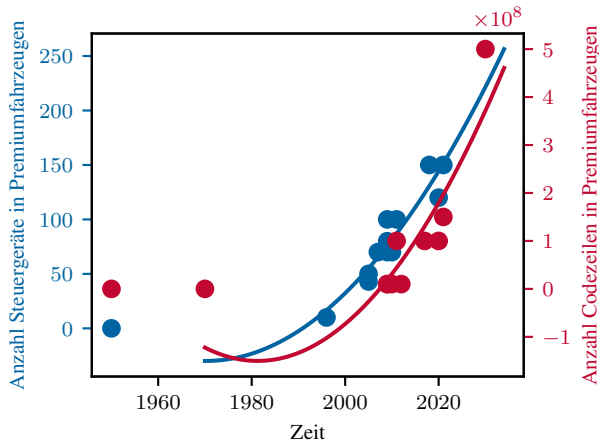


Abbildung 1.4: Entwicklung der Steuergeräteanzahl und Codezeilen in Premiumfahrzeugen über die Zeit (Datenquellen: [?, 1, 24–31]).

Der Anstieg an Situationen, die ein System beherrschen muss, führt zu zahlreichen Herausforderungen während der Entwicklung. Insbesondere beim Testen und der Absicherung stoßen aktuelle Methoden an ihre Grenzen [32]. Auch in der Abstimmung des Systems auf die Umgebung steigen die Aufwände mit der Anzahl der Situationen.

Um Kosten zu sparen und die hohen Aufwände zu rechtfertigen, werden Fahrerassistenzsysteme (eng. Advanced Driver Assistance Systems (ADAS)) und Systeme für das hochautomatisierte Fahren (HAF) nicht nur für ein Fahrzeugmodell entwickelt. Vielmehr wird eine solche Softwarefunktion in verschiedenen Fahrzeugen, Modellen und Baureihen mit unterschiedlichster Bauart, Ausstattung - insbesondere Motorisierung - und damit auch unterschiedlicher Fahrdynamik eingesetzt.

Oberste Priorität bei der Einführung (bedingt-)automatisierender Systeme ist der Nachweis der Sicherheit und korrekten Funktionsweise, gefolgt von der Sicherstellung des markenspezifischen Fahrzeugverhaltens. Insbesondere Premium- und Sportwagenhersteller definieren sich über eindeutige Markeneigenschaften, die Kunden beim Kauf eines Fahrzeugs des Herstellers erwarten. Die Akzeptanz automatisierender Systeme ist eng an die Sicherheit, Funktionalität und das Fahrerlebnis gekoppelt [33].

Um das markenspezifische Fahrverhalten für die unterschiedlichen Fahrzeugmodelle sicherzustellen, werden bei der Entwicklung der Softwarefunktion Parameter eingeführt, die Freiheitsgrade zur Abstimmung der Funktionen mit dem Gesamtsystem und der Fahrzeugumgebung bieten [34]. So kann zum Beispiel der Verzögerungszeitpunkt vor einer scharfen Kurve je nach Fahrzeugmodell, Fahrzeugsegment (bspw. Kleinwagen, Oberklasse oder Sportwagen), gewünschter Fahrdynamik, Fahrkomfort und Sicherheitskriterien unterschiedlich gewählt werden. Mit Hilfe der dafür verwendeten Parameter können diese Fahrzeugcharakteristika verändert werden ohne Änderungen im Softwarecode selbst durchführen zu müssen. Mit steigendem Softwareumfang und steigender Systemkomplexität ist auch die Anzahl der notwendigen Freiheitsgrade gestiegen [35]. In modernen Fahrzeugen gibt es über 5000 Parameter [36, 37]. Der Prozess der Abstimmung des Systems wird in der Automobilindustrie als Applikation bezeichnet.

1.2 Herausforderungen in der Applikation von ADAS und HAF-Funktionen

Systeme müssen mit einem hohen Reifegrad in der tatsächlichen Systemumgebung appliziert werden. Für ADAS und HAF-Funktionen bedeutet dies eine Erprobung in der realen Welt, da nur dort die direkte Erlebbarkeit der Systeme im späteren Einsatzgebiet gewährleistet ist. Dabei applizieren Entwickler und Applikateure das System direkt auf Basis ihres subjektiven Empfindens. Neben der reinen Funktionsfähigkeit des Systems geht es hierbei verstärkt auch um Komforteinstellungen und insbesondere die Sicherstellung gewünschter, markenspezifischer Fahreigenschaften, wie zum Beispiel Beschleunigungs- und Bremsverhalten.

Für die Applikation in der realen Welt existieren wesentliche Herausforderungen:

Funktionsparameter müssen in unterschiedlichen Ausprägungen getestet werden, um das Fahrverhalten der Funktion zu bewerten und dadurch eine optimale Applikation zu ermitteln. Zu jedem Zeitpunkt kann nur eine Kombination an Parameterausprägungen getestet werden. Tests in der Realität sind jedoch nur eingeschränkt wiederholbar. Somit ist ein direkter Vergleich verschiedener Parameterausprägungen nicht möglich. Für Systeme, die nur eine begrenzte Umfeldwahrnehmung benötigen, werden relevante Situationen teilweise reproduzierbar auf dem Prüfgelände herbeigeführt. Mit steigender Umfeldwahrnehmung können jedoch nicht mehr alle Situationen in der realen Welt reproduzierbar getestet werden.

Herausforderung 1: Unterschiedliche Parameterkombinationen müssen sequentiell getestet werden. Da Straßenverkehrstests nicht reproduzierbar sind, ist eine Vergleichbarkeit der Parameterkombinationen nicht gegeben.

Des Weiteren sind die Wirkungsbereiche der Parameter oft nicht auf einzelne Situationen beschränkt, sodass eine Verbesserung in einer Situation durch eine Parameteränderung möglicherweise ungewollte bzw. verschlechternde Auswirkungen auf andere Situationen hat, was den direkten Vergleich der Auswirkungen von Parameteränderungen weiter erschwert.

Herausforderung 2: Parameteränderungen wirken sich auf mehrere Situationen aus. Bei der Bewertung einer Parameterausprägung muss der gesamte Wirkbereich des Parameters berücksichtigt werden.

Eine weitere Herausforderung bei der Applikation in der realen Welt ist, dass das System aus Zeit- und Kostengründen stets nur auf einer kleinen Stichprobe des gesamten späteren Einsatzgebiets appliziert werden kann. Insbesondere bei Funktionen, die viele Informationen aus der Fahrzeugumgebung verarbeiten, müssen regionale Unterschiede berücksichtigt werden. Regionale Unterschiede sind unter anderem unterschiedliche Darstellungen von Verkehrsschildern, unterschiedliche Distanzmaße (Meilen und Kilometer), Rechts- bzw. Linksverkehr, andere Straßenbaunormen und -verordnungen aber auch topographische Gegebenheiten, wie Terrain und Witterung. Für die Applikation muss eine globale Sicht und Abdeckung aller relevanten Variationen gewährleistet werden. Eine lokale Applikation, die nicht alle regionalen Besonderheiten abdeckt, ist nicht ausreichend.

Herausforderung 3: Die Umfänge einer Applikation sind durch Zeit und Kosten limitiert. Dennoch muss eine Repräsentativität der Applikation für das gesamte spätere Einsatzgebiet sichergestellt werden.

In der Automobilindustrie ist in den letzten Jahren ein Wandel zu erkennen. Ergänzend zu weiterhin notwendigen Straßenverkehrstests einzelner Entwickler haben die Fahrzeughersteller größere Fahrzeugflotten in der gesamten Welt etabliert, um Kosten und Zeit zu sparen und mit bezahlten Testfahrern die Fahrzeuge als Gesamtsysteme zu testen. Entwickler und Applikateure einzelner Systeme stellt dies jedoch vor neue Herausforderungen. Sie sitzen nun nicht mehr selbst im Fahrzeug und erleben damit das Verhalten und die Qualität ihrer Softwaresysteme nicht mehr direkt, stattdessen müssen sie große Mengen aufgezeichneter Straßenverkehrstests am Computer analysieren und dadurch das Verhalten ihres Systems bewerten und anpassen.

Herausforderung 4: Die Applikation muss effizienter bzw. skalierbarer werden, um bei gleichbleibenden Entwicklungszyklen die steigenden Funktionsumfänge zu applizieren.

Herausforderung 5: Straßenverkehrstests, die nicht direkt im Fahrzeug erlebt wurden, gewinnen an Bedeutung für die Applikation und müssen herangezogen werden. Dafür ist eine objektive und automatisierte Analyse notwendig.

1.3 Forschungsfragen

In dieser Dissertation sollen Methoden und Werkzeuge für die global-optimale, datengetriebene Applikation teil- und hochautomatisierender Fahrfunktionen zur Längs- und Querverführung von Fahrzeugen entwickelt werden, sodass auch zukünftig ein optimaler Fahrkomfort und die höchstmögliche Sicherheit garantiert werden können. Eine besondere Herausforderung stellt die zunehmende Umfeldwahrnehmung und -interaktion dar, wodurch sich regionale Unterschiede und Besonderheiten zunehmend auf das Fahrverhalten der automatisierenden Fahrfunktionen auswirken. Auf Basis einer Literaturanalyse und unter Berücksichtigung aktueller Trends und Entwicklungen sollen folgende forschungsleitende Fragen beantwortet werden:

Forschungsfrage 1: Vergleichbarkeit sequentieller Parameteränderungen

Wie kann die Vergleichbarkeit von zu unterschiedlichen Zeitpunkten, an unterschiedlichen Orten sowie mit unterschiedlichen Parameterausprägungen durchgeführten Straßenverkehrstests hergestellt werden?

Forschungsfrage 2: Aussagekraft kleiner Stichproben

Wie kann eine repräsentative Aussage über die Qualität der Applikation in Bezug auf das gewünschte Funktionsverhalten auf Basis stichprobenartiger Straßenverkehrstests abgeleitet werden?

Forschungsfrage 3: Integration in den Entwicklungsprozess

Wie und um welche Schritte muss der Applikationsprozess erweitert werden, um eine global-optimale, datengetriebene Applikation zu gewährleisten?

1.4 Umfeld

Diese Arbeit ist am FZI Forschungszentrum Informatik im Rahmen von Forschungsaufträgen der Dr. Ing. h.c. F. Porsche AG entstanden. Für die Erforschung der Methoden und Algorithmen wurden im Rahmen der Forschungsaufträge Fahrzeugmessdaten bereitgestellt. Diese wurden für diese Dissertation verwendet. Aufgrund der Vertraulichkeit können keine Details zu den verwendeten Daten, insbesondere Geodaten, Fahrzeuginformationen und Softwarestände, genannt werden. Die Daten wurden so verändert, dass keine Rückschlüsse auf den tatsächlichen Umfang, Inhalt sowie die Routen der Erprobungsfahrten möglich sind. Die methodischen Grundlagen sind unverändert in dieser Dissertation enthalten.

1.5 Aufbau der Dissertation

Im folgenden Kapitel 2 werden die Grundlagen der automobilen Wertschöpfung in Form der Produktentwicklung und deren Entwicklungsprozesse erörtert,

Daraufhin wird in Kapitel 3 näher auf die Qualifikation und Integration von Fahrerassistenzsystemen und hochautomatisierten Fahrfunktionen eingegangen. Verschiedene Testverfahren mit einem Schwerpunkt auf die Gesamtsystemebene werden erörtert. Insbesondere wird die Applikation von ADAS und HAF-Funktionen näher ausgeführt.

In Kapitel 4 werden das Paradigma der datengetriebenen Entwicklung erläutert und Vorteile für die einzelnen Entwicklungsprozessschritte aufgezeigt. Zudem werden die für diese Dissertation notwendigen statistischen Grundlagen erläutert.

Die Herausforderungen in der Applikation sowie die Potenziale, die sich aus der datengetriebenen Entwicklung ergeben, werden in Kapitel 5 diskutiert. Zusätzlich wird der aktuelle Stand der Technik in der Applikation vorgestellt und Anforderungen an eine global-optimale, datengetriebene Applikation abgeleitet.

In Kapitel 6 wird die in dieser Dissertation erforschte global-optimale, datengetriebene Applikation im Detail erläutert. Dabei werden Schwerpunkte auf das Datenmodell (Abschnitt 6.2), einer Methodik zum Vergleich von zeitlich und räumlich getrennt durchgeführten Straßenverkehrstests (Abschnitt 6.3) und einer Methodik zum repräsentativen Vergleich der Straßenverkehrstests (Abschnitt 6.4) gelegt.

In Kapitel 7 wird die global-optimale, datengetriebene Applikation am Beispiel einer prädiktiven Längsregelung prototypisch umgesetzt.

Abschließend wird die vorgestellte global-optimale, datengetriebene Applikation in Kapitel 8 evaluiert und diskutiert.

Kapitel 9 fasst die Dissertation zusammen und gibt einen Ausblick auf weiterführende Forschungsarbeiten im Kontext der Applikation von Fahrerassistenzsystemen und hochautomatisierenden Fahrfunktionen.

2 Produktentwicklung - Wertschöpfung in der Automobilindustrie

Die Automobilindustrie ist ein Wirtschaftszweig, in dem Produkte (Fahrzeuge) hergestellt werden. Für Produkte, insbesondere Konsumgüter, wird in der Betriebswirtschaftslehre ein Produktlebenszyklus [38] angenommen, der die typischen Phasen, die ein Produkt durchläuft, aus betriebswirtschaftlicher Sicht beschreibt (vgl. Abbildung 2.1). Diesem Modell liegen verschiedene Annahmen zugrunde, zum Beispiel, dass die Lebenszeit eines Produkts begrenzt ist, und, dass ein Produkt nach der Wachstumsphase in eine Sättigungsphase geht, in welcher der Umsatz rückläufig wird. Um dem einbrechenden Umsatz entgegenzuwirken, werden heutzutage regelmäßig neue Produkte auf den Markt gebracht, die dann ihren eigenen Produktlebenszyklus von vorne durchlaufen. Damit ein Unternehmen langfristig wirtschaftlich agieren kann, muss die Einführung neuer Produkte bestmöglich geplant werden. In der Produktstrategie wird langfristig gesteuert, welche Produkte zu welchem Zeitpunkt auf den Markt gebracht werden [1]. Der Markteinführung geht ein mehrjähriger Produktentstehungsprozess voraus (vgl. Abschnitt 2.3).

Definition 1: Lebenszyklusmodell nach ISO 12207

Ein Lebenszyklusmodell beschreibt den Verlauf eines Produkts in Phasen. Die Phasen können sequentiell, überlappend und/oder iterierend sein. Das Lebenszyklusmodell kann sich auf die gesamte Lebensdauer des Produkts von der Konzeption bis zur Entsorgung beziehen oder auf einen projektspezifischen Ausschnitt [39].

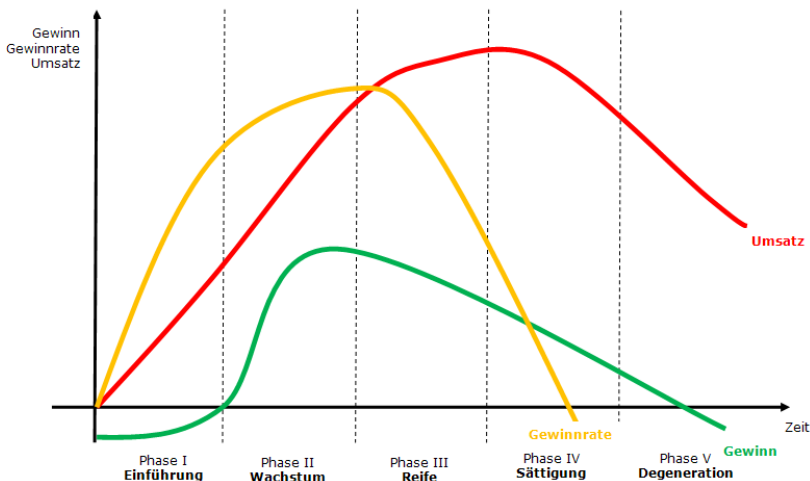


Abbildung 2.1: Umsatzentwicklung eines Produkts über die Phasen des Produktlebenszyklus (Quelle: [38]).

2.1 Das Produktportfolio - Strukturierung und Gruppierung von Marktsegmenten

Der Erfolg einer Produkteinführung hängt maßgeblich von der Nachfrage auf dem Markt ab. Die Kundengruppen besitzen unterschiedliche Ansprüche und Erwartungen an ein Produkt bzw. Fahrzeug, so dass es nicht ein Fahrzeug gibt, das den

kompletten Markt abdeckt. Daher haben die Automobilhersteller Produktportfolios entwickelt, die für jede Kundengruppe ein geeignetes Fahrzeug anbieten sollen.

Definition 2: Produktportfolio

Ein Produktportfolio ist eine vollständige Zusammenstellung aller von einem Unternehmen (oder einer strategischen Geschäftseinheit) geplanten, angebotenen sowie abgekündigten Produkte [40].

Zur Erstellung des Produktportfolios wird der gesamte Markt - also alle potentiellen Fahrzeugkäufer - zunächst anhand ausgewählter Merkmale in homogene Gruppen segmentiert. Diesen Gruppen werden anschließend Fahrzeugmerkmale zugeordnet, die diese als Kaufkriterien herbeiziehen. Über die Jahre ist so eine Sammlung von Fahrzeugsegmenten entstanden, die von den Fahrzeugherstellern mit einzelnen Fahrzeugmodellen adressiert werden. Die Fahrzeugsegmente sind ein Versuch, den vielschichtigen und hochkomplexen Markt in homogene Teilmärkte zu gliedern. Die Segmente sind ständig im Wandel und können je nach Bedarf unterschiedlich stark ausdifferenziert werden. Durch die Sättigung des Automobilmarktes steigt die Anzahl der spezifizierten Segmente stark an (vgl. [41] 1987: 9 Segmente; 1997: 26 Segmente; 2009: 57 Segmente), da die verbliebenen potentiellen Käufer immer gezielter angesprochen und auch kleine Nischenmärkte erschlossen werden sollen. [1]

Definition 3: Fahrzeugsegment

Fahrzeugsegmente stellen eine Klassifikation der Fahrzeugmodelle anhand optischer, technischer und marktorientierter Merkmale dar [42].

Die Europäische Kommission sowie das deutsche KBA haben Fahrzeugsegmente für die Abgrenzung wettbewerbsrechtlicher Marktanteile von Personenkraftwagen definiert [43]. Die Fahrzeugsegmente wurden dabei auf EU- und Bundesebene leicht unterschiedlich definiert (vgl. dazu Tabelle 2.1).

Fahrzeugsegment nach Europäischer Kommission	Fahrzeugsegment nach KBA	Fahrzeugbeispiel
A: Kleinwagen	Minis (Kleinstwagen)	Smart, VW Up!
B: Kleinwagen	Kleinwagen	Fiat Punto, VW Polo
C: Mittelklasse	Kompaktklasse	VW Golf
D: Obere Mittelklasse	Mittelklasse	Volva S-70, VW Passat
E: Oberklasse	Obere Mittelklasse	BMW 5er, Audi A6
F: Luxusklasse	Oberklasse	Mercedes S-Klasse, Audi A8
S: Sportwagen	Sportwagen	Ferrari, Porsche 911
M: Mehrzweckfahrzeuge (oder MPV-Van)	Minivan	Mercedes-Benz B-Klasse
	Großraumvan	VW Sharan
J: Geländewagen (Sport Utility Vehicles - SUV, einschließlich Fahrzeuge mit Allradantrieb)	Sport Utility Vehicle	Suzuki Vitara, BMW X1
	Geländewagen	Suzuki Vitara, Mercedes-Benz G-Klasse

Tabelle 2.1: Fahrzeugsegmente nach Europäischer Kommission [44] und KBA [45]

Für die Segmentierung des Automobilmarktes im Sinne eines Produktportfolios können Merkmale, wie die Fahrzeugklasse (A-E, J für Geländewagen, M für Mehrzweckfahrzeuge, S für Sportwagen), die Aufbauart (Sedan, Kombi, Coupé, Convertible, Roadster, Sports Activity Vehicle (SAV)), regionale Merkmale (z.B. Links- und Rechtslenker, schmalere Fahrzeuge in Japan) oder die Antriebsform (Verbrenner, Elektrisch, Hybrid, Wasserstoff, ..), herangezogen werden.

All diese Merkmale spannen ein Feld auf, welches die Automobilhersteller in die Planung ihrer Produktstrategie einbeziehen und darauf ihre Entscheidungen über

neue Fahrzeugprojekte basieren [1]. Dabei können einzelne Felder des Produktportfolios gezielt adressiert aber auch gezielt nicht adressiert werden. Volumen-Fahrzeughersteller decken ein breites Spektrum des Produktportfolios ab, wohingegen ein Sportwagenhersteller ggf. absichtlich Lücken im Produktportfolio zulässt bzw. nur gezielt einzelne Segmente adressiert (vgl. dazu Abbildung 2.2 Segmentanzahl: Bugatti: 2, Lamborghini: 3, Porsche: 6, Skoda: 11, VW: 17). Die detaillierte Kenntnis über Einflussgrößen und Bestimmungsgründe des Kaufverhaltens der Kunden ist eine wesentliche Voraussetzung für den Erhalt der Wettbewerbsfähigkeit von Automobilherstellern [46].




























































	Fließ- heck	Limou- sine	Kombi	Van	SUV	Coupé	Cabriolet	Roadster	City Van	Pick-Up
E										
D		  				  	 	 		
C		 			 					
B		  	  			 	 	 		
A	  	 	 		 		 			
A0	  	 	 							
A00	 									

Abbildung 2.2: Das Produktportfolio der Fahrzeugmarken des VW-Konzerns 2012 nach Fahrzeugklasse und Aufbauart (in Anlehnung an: [41]).

Die Produktstrategie dient der langfristigen Ausrichtung und Steuerung des Produktportfolios. Unter Berücksichtigung und Antizipation des Marktes wird entschieden, welche Fahrzeuge, Fahrzeugmodelle oder Varianten erfolgversprechend das bestehende Produktportfolio ergänzen können. Zudem werden Weiterentwicklungen bestehender Modelllinien geplant sowie wenig erfolgreiche Modelle und Modelllinien aus dem Produktportfolio entfernt. [1]

Definition 4: Produktstrategie

Die Produktstrategie definiert die Ausrichtung des Produktportfolios auf einen antizipierten, zukünftigen Markt bzw. eine Nachfrage an Fahrzeugen.

2.2 Fahrzeugplattformen - Wiederverwendung und Skaleneffekte

Um ein umfassendes Produktportfolio schnell und mit niedrigen Kosten bedienen zu können, müssen Skaleneffekte ausgenutzt und vorherige Entwicklungen wiederverwendet werden. In Fahrzeugplattformen werden Komponenten gebündelt, die in mehreren Fahrzeugmodellen und -varianten zum Einsatz kommen sollen. Die Bündelung wird dabei so vorgenommen, dass sich die Fahrzeuge einer Plattform optisch stark voneinander unterscheiden können und gleichzeitig viele Komponenten miteinander teilen [1]. Das heißt, dass nicht-differenzierende Komponenten, wie Antriebsstrang, Lenkung und Getriebe, geteilt werden und differenzierende Komponenten, wie Karosserie und Innenraumausstattung, fahrzeugspezifisch gewählt werden.

Definition 5: Plattform

Eine Plattform ist ein geteilter Satz von Komponenten, der über verschiedene Fahrzeuge mitunter auch unterschiedlicher Marken hinweg verwendet wird [1].

Aus einer Fahrzeugplattform können mehrere Modelllinien abgeleitet werden. Modelllinien teilen absichtlich ein gemeinsames Innenraum- und Karosserie-Design mit dem Ziel, einen Wiedererkennungswert zu schaffen und somit die Produktmarke zu stärken. Modelllinien können durch den Entwurf eines Modells und der Ableitung von Derivaten daraus entstehen. Eine andere Möglichkeit ist die Planung aller Varianten im Voraus, um eine hohe Wiederverwendung von

Teilen und Designs zu erreichen und damit das komplette finanzielle Potenzial der Modelllinie auszuschöpfen [1].

Der äußerliche Wiedererkennungswert ist exemplarisch an den Modellen der Porsche 911 Modelllinie zu beobachten (vgl. Abbildung 2.3a). Veränderungen am äußerlichen Erscheinungsbild werden nur geringfügig von Modell zu Modell vorgenommen. Die Porsche 911 Modelllinie besitzt zudem einen eindeutig definierten Erwartungswert der Kunden an Fahrgefühl und Fahrverhalten. Hingegen sind die Modelle der MLB Plattform des Volkswagen Konzerns (vgl. Abbildung 2.3b) wesentlich vielseitiger. Die Plattform wird über mehrere Marken des Konzerns eingesetzt und beinhaltet Fahrzeugmodelle verschiedener Fahrzeugklassen und Aufbauarten (z.B. Sedan, Kombi und SAV).

Definition 6: Modelllinie

Eine Modelllinie ist eine Familie von Fahrzeugen, die neben den nicht sichtbaren Komponenten auch große Anteile der Karosserie und Innenausstattung teilen und somit nicht nur technisch sondern auch optisch zu einer Produktfamilie gehört [1].

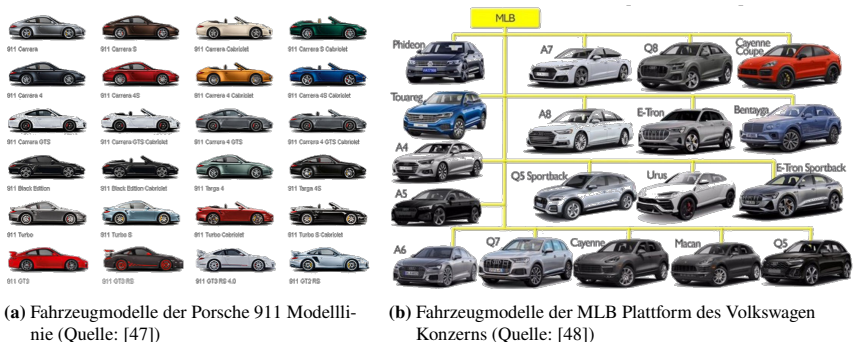


Abbildung 2.3: Äußere Erscheinung von Fahrzeugmodellen einer Modelllinie 2.3a und Fahrzeugmodellen einer Fahrzeugplattform 2.3b

Definition 7: Modell

Ein Fahrzeugmodell bildet die kleinste Einheit der Fahrzeugentwicklung. Fahrzeuge eines Modells bestehen zum überwiegenden Anteil aus den gleichen Bauteilen und unterscheiden sich nur durch Ausstattungsoptionen, die einen geringen Einfluss auf die Erscheinung des Gesamtfahrzeugs besitzen.

Neben den Produktfamilien und Modellen im Produktportfolio eines Automobilherstellers, gibt es zudem Ausstattungsmerkmale sowie länderspezifische Versionen der einzelnen Modelle, die die Vielfalt der Fahrzeuge weiter erhöhen. Weber spricht von drei Versionen: Europe, U.S. und Right Hand Driving [1]. Die Ausstattungsmerkmale können oft frei kombiniert werden und sind damit der Haupttreiber für die Variantenvielfalt im Automobilbereich.

Um diese Vielfalt beherrschbar zu machen, sind aus Sicht von Winner et al. [2] eine Top-Down orientierte Systemarchitektur, eine systematische Funktionsentwicklung sowie ein geeigneter Qualifikationsprozess notwendig. Um ein systematisches und strukturiertes Vorgehen zu gewährleisten, haben die Fahrzeughersteller Referenzprozesse für Fahrzeugentwicklungsprojekte entwickelt.

Definition 8: Prozess nach IEEE Standard Glossary of Software Engineering Terminology

Ein Prozess ist eine Abfolge von Arbeitsschritten, die zu einem bestimmten Zweck ausgeführt werden [49].

Definition 9: Referenzprozess

Ein Referenzprozess ist eine allgemeingültige Prozessvorlage, die ein einheitliches Vorgehen über mehrere gleichartige Vorhaben sicherstellen soll. Referenzprozesse dienen der Wiederverwendung, Standardisierung, Etablierung eines gewohnten Vorgehens sowie der Vergleichbarkeit und einem verbesserten Controlling einzelner Prozess-Instanzen. [50]

Die Abbildung der Arbeitsschritte in Prozessen dient dabei nicht nur der Strukturierung der Arbeitsaufgabe, sondern auch dem Controlling zur Sicherstellung der Einhaltung von Kosten-, Leistungs- und Terminzielen und somit der Wirtschaftlichkeit des Unternehmens. Die Inhalte und Schnittstellen der Arbeitsschritte sind eindeutig definiert.

2.3 Der Produktentstehungsprozess - Entwicklung von Fahrzeugmodellen

Die Gesamtheit der für die Entwicklung eines Fahrzeugmodells notwendigen Prozessschritte von der „strategischen Vision des Fahrzeugs bis zum ersten Serienfahrzeug“ [1] wird als **Produktentstehungsprozess (PEP)** [51] bezeichnet.

Der PEP ist ein Prozess nach dem Vorbild des Stage-Gate-Vorgehensmodells, bei dem die einzelnen Arbeitsschritte in mehrere Hauptphasen gegliedert sind. Die Übergänge zwischen den Phasen werden durch Qualitätsprüfpunkte (eng. Quality Gates) eindeutig definiert. Die jeweiligen Arbeitsschritte der vorherigen Phase müssen abgeschlossen und die Qualitätskriterien erfüllt sein, um in die nächste Phase übergehen zu können. Einzelne Arbeitsergebnisse können durch mehrere, dezentrale Teams erarbeitet werden. Die Quality Gates dienen dann neben dem Review der Ergebnisse und der Controlling-Funktion für das Management zusätzlich der Synchronisation der Arbeitsergebnisse. [52, 53]

Die Auslegung des PEP in der Automobilindustrie ist herstellerspezifisch (vgl. angepasster PEP von Porsche Engineering [54] im Anhang A.1.1) - allen gemein ist die grobe Untergliederung in mehrere, sich inhaltlich stark ähnelnde Hauptphasen. Die Dauer einer Fahrzeugentwicklung im PEP beträgt zwischen 40 und 58 Monaten [55, 56].

Weber [1] generalisiert drei Hauptphasen des PEP (vgl. Abbildung 2.4):

Während der **Initialisierungsphase** wird die Produktvision, die aus der Produktstrategie hervorgeht, konkretisiert, bewertet und anschließend daraus ein konkretes Fahrzeugprojekt abgeleitet. Zusätzlich werden Innovationen aus der Vorentwicklung in das konkrete Fahrzeugprojekt integriert.

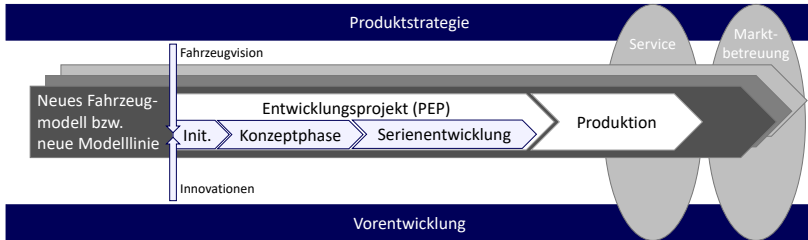


Abbildung 2.4: Start eines neuen PEP für ein Fahrzeugmodell oder eine Modelllinie aus der Produktstrategie mit den Innovationen aus der Vorentwicklung. (in Anlehnung an [1])

Neben der technischen Realisierbarkeit (Geometrie, Design, Beschaffung, Produktion) werden während der Initialisierungsphase des PEP in einem Geschäftsplan (eng. Business Plan) auch die ökonomischen Aspekte betrachtet. Bei guten wirtschaftlichen Aussichten und Gewährleistung der technischen Realisierbarkeit wird das Projekt in die nächste Phase überführt.

In der **Konzeptphase** wird das Fahrzeugdesignkonzept verfeinert. Auf 3D-Modellbasis werden Aspekte, wie Design, Aerodynamik und Agilität bewertet und in einzelnen Prototypen-Fahrzeugen validiert. Des Weiteren wird das Projektkonzept mit Hilfe von Subkonzepten (Produkt-, Produktions-, Beschaffungs-, Marketing-, und Servicekonzepten) quantifiziert. Diese Quantifizierung des Projektkonzepts mündet in einer Zielvereinbarung für das Projekt und beendet damit die Konzeptphase. Verträge mit externen Lieferanten werden unterzeichnet und Projektbudgets freigegeben.

Anschließend beginnt die **Serienentwicklung** mit der Realisierung der einzelnen Bestandteile des Fahrzeugs. Aus dem Produktkonzept werden eindeutige Beschreibungen für die Bauteile und Softwarekomponenten abgeleitet. Diese werden in der Serienentwicklung entweder vom Automobilhersteller selbst oder von

Zulieferern in einzelnen Funktionsentwicklungsprojekten umgesetzt. Die Qualitätssicherung über die einzelnen Entwicklungsprojekte im PEP hinweg wird durch Quality Gates sichergestellt (vgl. Abschnitt 2.5).

2.4 Funktionsentwicklungsprojekte - Wertschöpfung durch neue Funktionalität

In der Funktionsentwicklung muss zwischen Serienentwicklungsprojekten und Vorentwicklungsprojekten bzw. Pilotprojekten unterschieden werden (vgl. Vorentwicklung und Serienentwicklung in Abbildung 2.4). In **Vorentwicklungsprojekten** werden neue Konzepte und Ideen auf Studien- und Prototypenbasis umgesetzt, um deren Realisierbarkeit und Mehrwert für zukünftige Fahrzeuge zu bewerten. Aufgrund der ungewissen Erfolgsaussichten und schlechten Planbarkeit dieser Projekte werden sie nicht einem konkreten Fahrzeugentwicklungsprojekt zugeordnet, sondern verlaufen parallel und kontinuierlich zu den einzelnen Fahrzeugentwicklungsprojekten. Funktionen, die sich in den Pilotprojekten und Machbarkeitsstudien bewährt haben, können in konkrete Fahrzeugentwicklungsprojekte integriert werden. In der **Serienentwicklung** werden vorhandene Funktionen bzw. Funktionsprototypen aus der Vorentwicklung weiterentwickelt und auf einen serientauglichen Reifegrad gebracht [57] (vgl. Abschnitt 2.5). Insbesondere in der Vorentwicklung werden neue Innovationen vorangetrieben, die der Produktdifferenzierung dienen und somit das neue Modell von den Vorgängern und den Modellen der Wettbewerber unterscheiden sollen. Laut Weber werden „90% aller Innovationen durch Elektronik und Software ermöglicht“ [1].

Das Zusammenspiel von Vor- und Serienentwicklung garantiert einen stetigen Innovationshub für neue Fahrzeugprojekte. Allerdings steigt dadurch auch die Anzahl der Funktionen in den Fahrzeugen stetig. In heutigen Fahrzeugen sind zahlreiche Funktionen integriert, die nach Weber [1] entsprechend ihrer funktionalen Zugehörigkeit in die folgenden Fahrzeugdomänen gruppiert werden. Bach [58] erweitert diese Domänen um eine Bordnetz-Domäne zur Abbildung der

zunehmenden Elektrifizierung des Antriebsstrangs und der damit einhergehenden „Batterie-, Starkstrom- und Hochspannungstechnik“:

- Infotainment
- Karosserie (eng. body electronics)
- Chassis und Fahrerunterstützung (eng. driver assistance)
- Antriebsstrang (eng. powertrain)
- Passive Sicherheit (eng. passive safety)
- Bordnetz

Die Kompetenzbereiche in der E/E-Entwicklung sind analog zu den Fahrzeugdomänen aufgeteilt. Sie dienen der Bündelung von Fachkompetenzen sowie der spezialisierten, separierten Entwicklung der einzelnen Domänenbestandteile. Der Fokus in dieser Dissertation liegt auf der Entwicklung und Applikation von ADAS und HAF-Systemen. Folgende Ausführungen beziehen sich daher in erster Linie immer auf die Fahrerunterstützungsdomäne.

2.4.1 Vorgehensmodelle

Um die Komplexität der Fahrzeugentwicklungsprojekte, die durch die unterschiedlichen Arten von Entwicklungsprojekten, der großen Anzahl involvierter Stakeholder und Entwicklungsabteilungen, die hohe Anzahl einzelner Komponenten und Subsysteme im Fahrzeug, den vielen notwendigen Prozessschritten und den hohen Anforderungen an die Sicherheit der Systeme ergibt, ist ein strukturiertes und koordiniertes Vorgehen während der Entwicklung notwendig.

In Vorgehensmodellen werden Referenzprozesse für die Durchführung von Entwicklungen definiert. Sie beschreiben die einzelnen Prozessschritte, die involvierten Rollen, die Arbeitsergebnisse je Prozessschritt und die Schnittstellen zwischen den Schritten.

Definition 10: Vorgehensmodell

Vorgehensmodelle in der Systementwicklung und im Softwareengineering beschreiben Folgen bzw. Bündel von Aktivitäten, die zur Durchführung eines (IT)-Projekts erforderlich sind. Üblich ist eine Gliederung in Phasen, Arbeitsabschnitte und Aktivitätsblöcke, durch die der Entwicklungsprozess in planbare und kontrollierbare Einheiten zerlegt wird (Komplexitätsreduktion und Projektcontrolling). [59]

Im Automotive Systems Engineering (ASE) wird der Prozess der Systementwicklung durch Vorgehensmodelle beschrieben und strukturiert. Das Ziel der Vorgehensmodelle ist es, die Abläufe und Teilschritte in Entwicklungsprozessen zu formalisieren und zu strukturieren, um zum einen ein Projektcontrolling zu ermöglichen und zum anderen die Komplexität der Systementwicklung zu reduzieren [59]. Durch im Voraus festgelegte Phasenziele (Meilensteine) können Abweichungen und Fehler frühzeitig erkannt und beseitigt werden. Es existieren zahlreiche Vorgehensmodelle, die sich in sequentielle, iterative und agile Vorgehensmodelle untergliedern lassen.

Die Eignung der Vorgehensmodelle hängt von der Art, dem Umfang, der Dauer und den Anforderungen des Entwicklungsprojekts ab. Durch das sogenannte „Tailoring“ werden die Referenzmodelle an die Gegebenheiten des Unternehmens, Systems und Projekts angepasst [60]. Für gut strukturierte und starre Projekte, die zu Beginn vollständig durchgeplant werden können und das erwartete Ergebnis bereits bekannt ist, sind sequentielle Vorgehensmodelle, wie das Wasserfallmodell [61] oder das Spiralmodell [62], geeignet. Für weniger stark strukturierte Projekte mit offenem Ergebnis eignen sich dagegen agile oder iterative Vorgehensmodelle, wie Scrum [63] oder Rapid Prototyping [64]. Die agilen Vorgehensmodelle ermöglichen es, während der Entwicklung flexibler auf Änderungen zu reagieren und Anforderungen und Wünsche des Auftraggebers oder Kunden während der Entwicklung noch mit einzubeziehen [65, 66].

2.4.2 Das V-Modell

Das V-Modell ist ein sequentielles Vorgehensmodell, das aus zeitlich aufeinanderfolgenden Phasen besteht. Der Ursprung des V-Modells ist auf Boehm im Jahr 1984 zurückzuführen [67]. Seitdem wurde es mehrfach überholt, z.B. im V-Modell 1992 und V-Modell 1997 [68, 69]. Das V-Modell besteht aus den Phasen System-Anforderungsanalyse, System-Entwurf, Software-Anforderungsanalyse¹, Software-Grobentwurf, Software-Feinentwurf, Software-Implementierung, Software-Integration, System-Integration und Überleitung in die Nutzung [69].

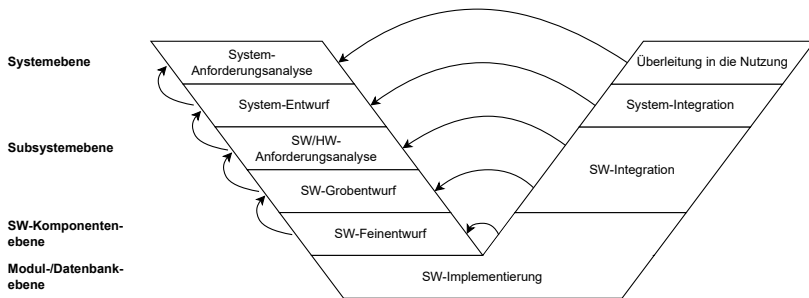


Abbildung 2.5: Das V-Modell 97 (in Anlehnung an [69]).

Die Systementwicklung nach dem V-Modell 97 (vgl. Abbildung 2.5) beginnt mit der **System-Anforderungsanalyse**. Für ein Gesamtfahrzeug ergeben sich die Anforderungen neben legislativen und Sicherheitskriterien auch aus der Positionierung im Produktportfolio und den im Fahrzeugsegment anvisierten Kundenbedürfnissen. Neben einer Gefahren- und Risikoanalyse werden ein technisches Modell des Systems erstellt und der funktionale Umfang des Gesamtsystems definiert. Jedes System trägt damit zur Erreichung der Gesamtanforderungen auf Fahrzeugebene bei.

Während des **System-Entwurfs** wird das System modelliert und in kleinere funktionale Subsysteme zerlegt. Jedes Subsystem wird dabei an eine Anforderung

¹ Die Hardwareentwicklung ist für diese Dissertation nicht relevant und wird nicht betrachtet

geknüpft. Die folgenden Phasen werden für jedes Subsystem durchlaufen. Subsysteme können Hardware oder Software sein.

Für jedes Subsystem werden separat die technischen Anforderungen detailliert. Für Softwaresysteme wird dies während der **Software-Anforderungsanalyse** durchgeführt.

Um eine unabhängige Entwicklung der Softwaresysteme zu gewährleisten, werden in der **Software-Grobentwurfs**-Phase die Schnittstellen zwischen den Subsystemen definiert sowie die Softwarearchitektur festgelegt.

Anschließend wird jedes System während der **Software-Feinentwurfs**-Phase in einzelne Komponenten und Module zerlegt, um eine weitere Komplexitätsreduktion zu erzielen.

Die **Software-Implementierung** ist die Phase, in der die Softwarekomponenten und -module realisiert werden. Hier findet die Implementierungsaufgabe statt. Am Ende der Phase steht ein ausführbarer Softwarecode zum Testen bereit.

In der **Software-Integration** werden die Softwarekomponenten und -module entsprechend ihrer Schnittstellen integriert und getestet. Dazu werden die Anforderungen aus der Entwurfsphase herangezogen. Ziel des Software-Tests ist es, die integrierten Softwarekomponenten auf Fehler zu testen und zu prüfen, ob sie die spezifizierten Anforderungen erfüllen.

Nach erfolgreichem Software-Test werden die Subsysteme, bestehend aus Software und Hardware, in der **System-Integrations**-Phase zu einem System zusammengefügt und auf das spezifizierte Verhalten getestet. Der Fokus der Tests liegt hier auf dem korrekten Zusammenspiel der Subsysteme.

In der letzten Phase, der **Überleitung in die Nutzung**, wird das Gesamtsystem gegen die Nutzeranforderungen getestet. Eine abschließende Prüfung stellt fest, ob die Anforderungen erfüllt sind und das System in die Nutzung übergeleitet werden kann.

Definition 11: System

Ein System ist eine Sammlung interagierender Komponenten, die eine oder mehrere bestimmte Funktionen in einer bestimmten Umgebung erfüllen. [49]

Definition 12: Subsystem

Ein Subsystem ist ein sekundäres oder untergeordnetes System eines größeren Systems. [49]

Definition 13: Software-Komponente

Eine Software-Komponente ist eine minimales Stück Software, für das eine separate Spezifikation existiert. [70]

Definition 14: Software-Modul

Eine Software-Modul ist eine nicht weiter teilbares Stück Software. [70]

Bei strikter Anwendung des V-Modells müsste das fertige Endprodukt - also das zu entwickelnde Fahrzeug - in einem Durchlauf vollständig spezifiziert, entwickelt und anschließend qualifiziert werden. Iterationsschleifen sind nicht vorgesehen. Bei typischen Entwicklungszeiten von 3-5 Jahren für ein Fahrzeugmodell wären Fehler sehr kostspielig und würden die Fertigstellung der Entwicklung potentiell um mehrere Jahre verzögern. In der Praxis werden Fahrzeuge daher nicht in einem V-Modell entwickelt. Stattdessen werden mehrere V-Modell-Iterationen nacheinander durchlaufen. So können Konzeptstudien, Prototypen und Musterstände erzeugt und getestet werden bevor sie anschließend in nachfolgenden Entwicklungszyklen sequenziell weiterentwickelt werden [71]. Zudem sind Rücksprünge in vorherige Entwicklungsschritte realisierbar, falls es in früheren Phasen zu unbemerkten Fehlern gekommen sein sollte [68].

Das V-Modell unterstützt die zielgerichtete und strukturierte Entwicklung von komplexen, sicherheitskritischen Systemen mit zusätzlich verteilter Entwicklungsleistung und mehreren Entwicklungsteams. Es soll sicherstellen, dass nach dem Durchlaufen des Entwicklungsprozesses ein sicheres und den Anforderungen genügendes System entwickelt wurde. Durch die Dekomposition des Systems wird die Komplexität während der Entwicklung reduziert.

2.4.3 Fahrerassistenzsysteme

Zu den Subsystemen eines Fahrzeugs, die während des Produktentstehungsprozesses entwickelt und in das Gesamtsystem integriert werden müssen, zählen auch Softwaresysteme, die den Fahrer bei der Führung des Fahrzeugs unterstützen, sogenannte Fahrerassistenzsysteme. Diese wurden durch die zunehmende Elektrifizierung des Fahrzeugs ermöglicht und beschränkten sich zunächst auf stabilisierende Sicherheitssysteme. Durch die gestiegenen Elektronikumfänge und Sensoreinheiten im Fahrzeug wurden diese Systeme um weitere Sicherheits- und Komfortsysteme mit Umgebungsinteraktion sowie Navigationssysteme erweitert [23] (vgl. Abbildung 2.6).

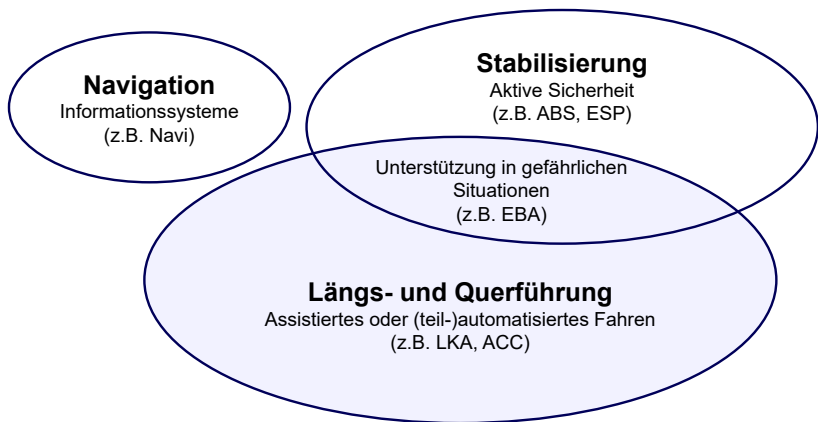


Abbildung 2.6: Einteilung von Fahrerassistenzsystemen in drei Aufgabenklassen (in Anlehnung an [9])

Zu den Stabilisierungssystemen zählen unter anderem das ABS und das Elektronisches Stabilitätsprogramm (ESP), welche dem fahrdynamischen Zustand (z.B. die Gierrate und Querschleunigung) des Fahrzeugs zur automatisierten Stabilisierung in hochdynamischen Situationen verwenden. Stabilisierungssysteme werden auch als Active Safety bezeichnet [23]. Sie besitzen ein großes Potenzial zur Verringerung der Unfallzahlen [72]. Der Kombination aus Emergency Brake Assist (EBA) und Lane Keep Assist (LKA) wird ein theoretisches Sicherheitspotenzial bei einer einhundertprozentigen Marktdurchdringung von 25,1% zugerechnet. Dabei entfallen 17,8% auf das EBA-System und 7,1% auf das LKA-System [73].

Fahrerassistenzsysteme, die dynamische Fahraufgaben (eng. Dynamic Driving Task (DDT), folgend als Fahraufgabe bezeichnet) übernehmen sollen, benötigen zusätzliche Umfeldinformationen. Diese müssen durch die maschinelle Wahrnehmung der Fahrzeugumgebung erlangt werden. Zur Fahraufgabe gehören die laterale und longitudinale Fahrzeugführung, die Überwachung der Fahrzeugumgebung, die Reaktion auf Ereignisse in der Fahrzeugumgebung sowie die Manöverplanung [74]. Zur sicheren Fahrzeugführung ist eine Verständnis der aktuellen Situation sowie der Prädiktion der weiteren Entwicklung der Fahrzeugumgebung notwendig. Für die fahrzeugführenden Fahrerassistenzsysteme wurde eine weitere Klassifikation nach ihrer Automatisierungsstufe eingeführt (vgl. Tabelle 1.1). Fahrzeugführende Fahrerassistenzsysteme werden auch als Advanced Driver Assistance Systems (ADAS) bzw. ab SAE-Level 3 als bedingt- und hochautomatisierte Fahrfunktionen (Hochautomatisiertes Fahren (HAF)-Funktionen) bezeichnet. Um bei steigendem Automatisierungsgrad die Sicherheit im Straßenverkehr zu gewährleisten, nimmt die Umgebungswahrnehmung mit dem Automatisierungsgrad zu. In heutigen Premiumfahrzeugen werden über 2500 Fahrzeugsignale über die Bussysteme ausgetauscht [75]. Einzelne ADAS und HAF-Funktionen verarbeiten Informationen von bis zu 45 Sensoren und damit mehrere hundert Signale am Funktionseingang und berechnen über den zeitlichen Verlauf mehrere Ausgangs- bzw. Steuergrößen (vgl. Abbildung 2.7) [76].

Definition 15: Fahraufgabe

Die Fahraufgabe umfasst alle operativen und taktischen Aufgaben in Echtzeit, die für den Betrieb eines Fahrzeugs im Straßenverkehrs erforderlich sind. [74] Die Fahraufgabe kann zeitlich, räumlich und in Abhängigkeit von den Umgebungsbedingungen eingegrenzt sein.

Definition 16: Fahrfunktion

Ein Subsystem des Fahrzeugs, welches mindestens eine Fahraufgabe realisiert, wird als Fahrfunktion bezeichnet. Insbesondere zählen dazu ADAS und HAF-Funktionen.

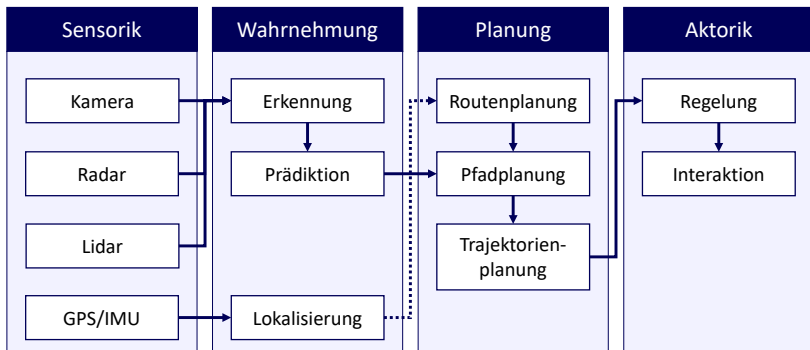


Abbildung 2.7: Schematische Darstellung der Verarbeitungsschritte innerhalb der Wirkkette für ein Fahrerassistenzsystem (in Anlehnung an [23, 77])

Für den Betrieb von Automatisierungssystemen muss deren Betriebsbereich (eng. Operational Design Domain (ODD)) genau spezifiziert werden. Die ODD von ADAS und HAF-Funktionen dient der Einschränkung des Einsatzes auf nachweislich sichere Betriebsbedingungen. Wird die ODD verlassen, muss die Fahrzeugführung rechtzeitig an den Fahrer übergeben werden. [78]

Definition 17: Operational Design Domain (ODD)

Die Operational Design Domain beschreibt das Set an Betriebsbedingungen, für die ein automatisierendes System ausgelegt ist. Zu den Betriebsbedingungen zählen unter anderem Umgebungsbedingungen, geographische Informationen, Tageszeiten sowie das explizite Vorhandensein bzw. die explizite Abwesenheit gewisser Verkehrs- oder Straßenbedingungen. [78, 79]

2.5 Qualitätssicherung und -prüfung

Ziel des ASE ist es, die Entwicklung qualitativ hochwertiger Systeme sicherzustellen. Ein erster Schritt dazu ist das strukturierte und systematische Vorgehen während der Entwicklung, wie es durch die Vorgehensmodelle definiert ist. Im PEP wird der Reifegrad des Systems über Quality Gates - Meilensteine mit fest definierten und objektiv messbaren Qualitätskriterien - sichergestellt. Der Verband der Automobilindustrie (VDA) definiert 8 Reifegradstufen (vgl. Abbildung 2.8) mit dem Ziel einer „Harmonisierung von Inhalten und Abläufen in der Lieferkette“ zur Sicherstellung der „Anlauf-, Anliefer- und Feldqualität des betrachteten Lieferumfangs“ [57]. Zu jeder Reifegradstufe werden Reifegradindikatoren sowie Mess- und Bewertungskriterien für jeden Reifegradindikator definiert. Diese können für ein durchgängiges Reifegradtracking über den gesamten PEP verwendet werden. Der VDA schlägt eine 3-stufige Risikoklassifizierung für Subsysteme und Komponenten vor. Je nach Einstufung können unterschiedliche Maßnahmen bei Zielabweichung eingeleitet werden. [57]

Definition 18: Reifegrad

Der Reifegrad gibt an, wie gut ein System funktional die Anforderungen erfüllt, durch die es spezifiziert wurde. [80]

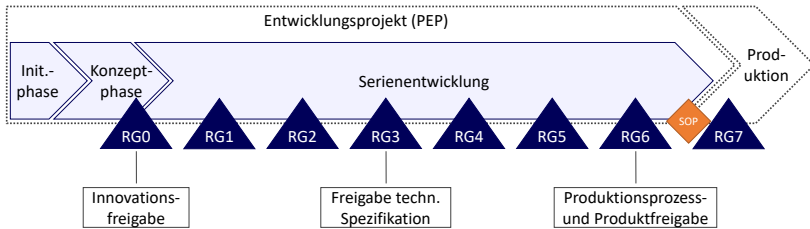


Abbildung 2.8: Die 8 Reifegradstufen der Reifegradbewertung entlang der Phasen des PEP (in Anlehnung an [57]).

Der VDA definiert 8 Reifegradstufen entlang des PEP (vgl. Abbildung 2.9). Während der Konzeptphase werden die Reifegrade RG0-RG2 erreicht. Diese zielen auf die interne Projektorganisation und das Projektmanagement ab. Weiterhin werden in der Konzeptphase Designentscheidungen getroffen, Komponentenlastenhefte erstellt und Leistungsumfänge sowie Schnittstellen für Komponentenentwicklungen festgelegt. Mit RG1 wird für Neuentwicklungen die Funktionalität über A-Muster nachgewiesen. Der in späteren Phasen zu erbringende Nachweis der Produktqualität wird über eine Simulations- und Testplanung festgelegt. Zudem werden erste Lieferantengespräche geführt und die System- und Softwarearchitektur inkl. Schnittstellen definiert, so dass mit Erreichen des RG2 eine interne Serienentwicklung oder externe Beauftragung einzelner Komponenten beginnen kann. Mit RG2 werden zudem die Release- und Erprobungsplanung festgesetzt. [57]

Nach Erreichen des RG2 werden die identifizierten Komponenten unabhängig voneinander entwickelt. Zu jeweils festgelegten Releases werden Soft- und Hardwarestände erzeugt, die die Reifegradstufen RG3-RG6 durchlaufen. Zwischen jedem Release einer Soft- oder Hardwarekomponente wird ein gesamter V-Modell-Zyklus durchlaufen (vgl. Abbildung 2.9). Die Hardware und Mechanik Releases führen zu neuen Musterständen. Während in der Konzeptphase A-Muster erzeugt wurden, entsteht zum ersten Release der Hardware das erste B-Muster (B1.0). Dieses kann über kleinere Releases zu B1.1, B1.2, etc. Mustern bzw. über große Releases zu C- und D-Mustern weiterentwickelt werden. Software-Releases

können entsprechend simulativ oder auf den aktuellsten Musterständen getestet werden. [57]

Die Hardware- und Mechanikentwicklung ist mit Erreichen der D-Musterstände abgeschlossen. D-Muster werden bereits mit Serienwerkzeugen und unter Serienbedingungen entwickelt. Mit Erreichen des Reifegrads RG5 werden für die Hardware- und Mechanikanteile Produktionsprozess- und Produktfreigabeverfahren bestätigt. Die Softwareentwicklung muss zum RG6 abgeschlossen sein, so dass auch die Gesamtsystementwicklung abgeschlossen werden kann. RG7 markiert den finalen Projektabschluss und die Verantwortungsübergabe an die Serie. [57]

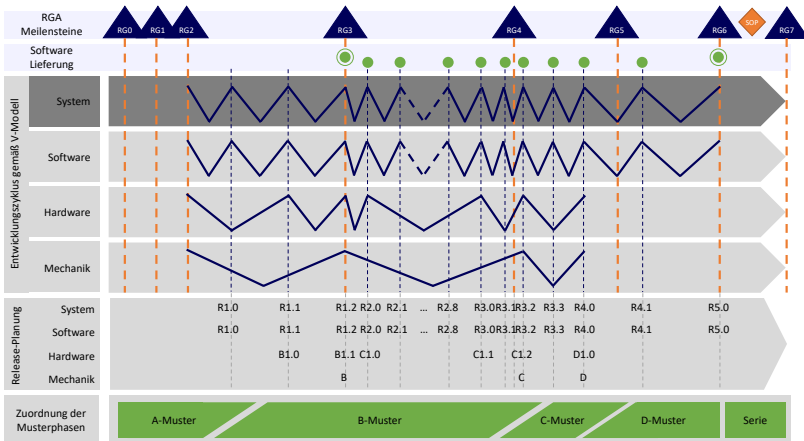


Abbildung 2.9: Systementwicklung entlang der Reifegradstufen im PEP (in Anlehnung an [57]).

Die Reifegradentwicklung des VDA bietet eine Steuerungsmethode für das Projektmanagement und soll Zusatzaufwände und Zeitverzögerungen im Produktentstehungsprozess vermeiden. Schäuffele [81] ergänzt zudem Richtlinien, die der Qualitätssicherung dienen, wie zum Beispiel den Einsatz geschulter, gut ausgebildeter Mitarbeiter oder der Automatisierung manueller, fehlerträchtiger Arbeitsschritte.

Definition 19: Fehler

Ein Fehler bezeichnet die Abweichung eines berechneten, beobachteten oder gemessenen Werts oder Bedingung vom wahren, spezifizierten und theoretisch korrekten Wert bzw. Bedingung. [82]

Während die Qualitätssicherung das Auftreten von Fehlern verhindern soll, liefert sie dennoch keine Garantie für eine fehlerfreie Entwicklung. Daher muss möglichst nach jedem Entwicklungsschritt eine Qualitätsprüfung durchgeführt werden mit dem Ziel, aufgetretene Fehler zeitnah zu identifizieren und zu beheben. Dabei lassen sich Fehler bzw. Unzulänglichkeiten der Spezifikation und Fehler bzw. Unzulänglichkeiten der Umsetzung der Spezifikation, also des Designs und der Implementierung, unterscheiden. Entsprechend sind die Begriffe Verifikation und Validierung wie folgt definiert:

Definition 20: Verifikation

Die Verifikation ist ein Prozessschritt zur Bestimmung der Vollständigkeit und korrekten Spezifikation oder Implementierung der Anforderungen für eine Entwicklungsphase. Am Ende der Entwicklungsphase wird durch die Verifikation geprüft, ob die Ergebnisse den Anforderungen an diese Phase genügen. [82]

Definition 21: Validierung

Die Validierung ist ein Prozessschritt mit zwei Zielen. Zum einen muss der Nachweis erbracht werden, dass die Sicherheitsziele erfüllt und dass die abgeleiteten Sicherheitskonzepte geeignet für den Nachweis der funktionalen Sicherheit des Systems sind. Zum anderen muss der Nachweis erbracht werden, dass die Sicherheitsziele korrekt, vollständig und auf Gesamtsystemebene vollumfänglich erreicht sind. [82]

3 Qualifikation und Integration von ADAS und HAF-Funktionen

Ziel der Qualifikation ist es, die Sicherheit und Funktionsfähigkeit des Gesamtsystems sicherzustellen und damit einen sicheren Gebrauch durch die Anwender sowie Dritte zu gewährleisten. Für diesen Sicherheitsnachweis müssen die Dimensionen **funktionale Sicherheit**, die **Sicherheit der beabsichtigten Funktion**, die **Gebrauchssicherheit**, die **Cyber Security** sowie weitere **ethische, regulatorische und gesetzliche Randbedingungen** betrachtet werden (vgl. Abbildung 3.1). Um ein einheitliches, sicheres und durchgängiges Vorgehen sicherzustellen, wurden mehrere Normen für die Qualifikation und Integration von Fahrzeugen definiert.

Definition 22: Sicherheit

Sicherheit ist die Abwesenheit inakzeptabler Risiken. [82]

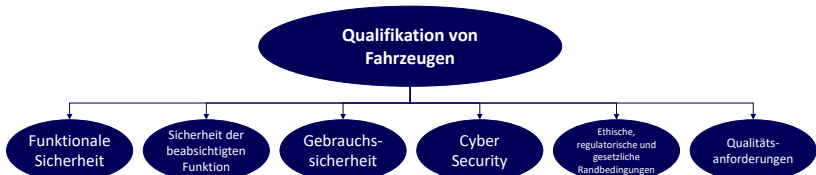


Abbildung 3.1: Die zu betrachtenden Dimensionen zur Qualifikation von Fahrzeugen

Die ISO26262 beschreibt die Anforderungen und Durchführung der Tests zur Sicherstellung der funktionalen Sicherheit des Systems [82]. Nach ISO26262 [82] muss während der Anforderungsanalyse auch eine Risikoanalyse durchgeführt

werden, um daraus Sicherheitsziele für das Produkt abzuleiten. Ziel der Verifikation & Validierung (V&V) ist es den Nachweis zu erbringen, dass diese Sicherheitsziele erfüllt sind. Die V&V mündet in einem Prozess zur Qualifikation und Zulassung des Systems durch staatliche Institutionen, wie dem KBA, der vor der Markteinführung des Systems durchlaufen werden muss. Der dafür notwendige Nachweis der Sicherheit des Systems bestimmt maßgeblich die Ausrichtung und den Umfang des Testens.

Definition 23: Testen

Das Testen ist eine Aktivität bei der ein System oder eine Komponente unter bestimmten Bedingungen ausgeführt wird, die Ergebnisse beobachtet oder aufgezeichnet werden und eine Bewertung über einen Aspekt des Systems oder der Komponente erstellt wird. [83]

In der ISO21448 werden technische Anforderungen an die Sicherheit der beabsichtigten Funktion (eng. Safety of the intended functionality (SOTIF)) formuliert [74]. Ziel der SOTIF ist es, funktionale Unzulänglichkeiten, Systemlimitierungen sowie vorhersehbaren Missbrauch (eng. Foreseeable Misuse), die zu einem potenziell gefährlichen Verhalten und damit erhöhtem Risiko führen können, zu identifizieren und zu vermeiden. Sie beschreibt Maßnahmen zur Risikominimierung und Gefahreneliminierung, die über den PEP hinweg implementiert werden müssen. Dabei wird das Fahrzeug als Gesamtsystem betrachtet und potenzielle Störeinflüsse und Auslöser (eng. Triggering Events) in der Fahrzeugumgebung auf das Fahrzeug bzw. die Fahrzeugführung untersucht. Zur Beschreibung und Strukturierung dieser externen Faktoren wird ein szenarienbasierter Ansatz gewählt. [74]

Definition 24: Szenario

Ein Szenario ist eine Beschreibung der zeitlichen Beziehungen zwischen einer Abfolge von Szenen, wobei einer Situation Ziele und Werte zugeordnet werden können und das Szenario durch Aktionen und Events beeinflusst werden kann. [74]

Definition 25: Szene

Eine Szene ist ein Snapshot der Fahrzeugumgebung, die aus Szenerie, dynamischen Objekten, dem Ego-Fahrzeug und weiteren Akteuren bzw. Beobachtern besteht sowie die Relationen zwischen diesen Entitäten beschreibt. [74]

Definition 26: Situation

Eine Situation betrachtet eine Szene aus der Perspektive eines spezifischen Verkehrsteilnehmers der Szene [84].

Zusätzlich zu den sicherheitsbezogenen Dimensionen ist die Erfüllung der Kunden- und selbstgesteckten Qualitätsanforderungen an das System für eine erfolgreiche Qualifikation entscheidend.

3.1 Integrationsstufen

Für alle Dimensionen der Qualifikation muss nachgewiesen werden, dass die definierten Anforderungen korrekt und vollständig sind (vgl. *Validierung*) sowie die definierten Anforderungen korrekt und vollständig umgesetzt wurden (vgl. *Verifikation*). Dieser Nachweis muss für das Gesamtsystem Fahrzeug erbracht werden. Die durch die Dekomposition im V-Modell (vgl. Abschnitt 2.4.2) entstandenen und getrennt entwickelten Softwaremodule und Hardwarebauteile müssen für den Sicherheitsnachweis des Gesamtsystems stufenweise integriert und die Integrationsstufen jeweils ausgiebig qualifiziert werden. Die Integrationsstufen sind in aufsteigender Reihenfolge Soft- und Hardwarekomponenten, Subsysteme und das gesamte System [68] (vgl. Abbildung 3.2).

Definition 27: Integrationsstufe

Die Integrationsstufe definiert den aktuellen Scope des zu testenden Systems (eng. System Under Test (SUT)), der sich an den durch die Dekomposition im V-Modell eingeführten Ebenen des Gesamtsystems Fahrzeug orientiert.

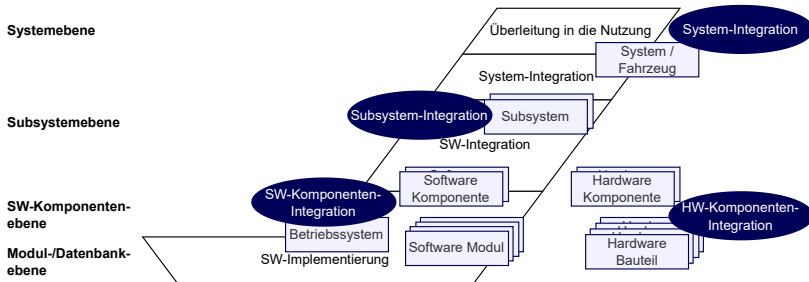


Abbildung 3.2: Artefakte und Integrationsstufen im V-Modell (in Anlehnung an [85])

3.2 Qualifikation entlang der Integrationsstufen

Mit jedem Software- und Hardwarerelease im Reifegradprozess wird für die jeweilige Komponente ein Entwicklungszyklus nach dem V-Modell durchlaufen. Damit einhergehend werden über die Integrationsstufen hinweg von der einzelnen Komponente bis zum Fahrzeug umfangreiche Funktions- und Integrationstests durchgeführt, um Fehler frühzeitig zu finden. Zur Fehlerdetektion werden dabei je nach Integrationsstufe verschiedene Tests durchgeführt. Zur Durchführung dieser Tests wird eine Testumgebung benötigt, die aus dem SUT und einer Testplattform besteht [86].

Definition 28: Testplattform

Eine Testplattform bezeichnet einen physischen und/oder virtuellen Versuchsaufbau zur Durchführung von Tests für ein SUT. Dabei wird die Systemumgebung real aufgebaut oder modelliert.

Verifikations- und Validierungstests können zu unterschiedlichen Reifegradstufen des Systems an unterschiedlichen Testplattformen durchgeführt werden. Diese reichen von einer komplett virtuellen, simulativen Ausführung des SUT bis hin zu Fahrzeugtests unter realen Bedingungen. Softwaretests können statisch oder dynamisch sein. Statische Softwaretests prüfen Dokumente oder Quellcode, während bei dynamischen Tests das SUT durch Stimuli ausgeführt wird. Stimuli sind dabei in der Regel durch Testfälle fest definierte Testdaten. [80]

Beim dynamischen Testen wird das SUT unter variierbaren Bedingungen ausgeführt, um das Verhalten des SUT zu bewerten. Ein Test bzw. ein Testfall geht dabei mit einem erwarteten Ergebnis bzw. Verhalten einher. Bei zustandsbehafteten Systemen muss auch das zeitliche Verhalten betrachtet werden. Bei Systemen, die ihre Umgebung beeinflussen, muss eine Rückkopplung des Systemausgangs über die Systemumwelt auf den Systemeingang des nächsten Zyklus vorgenommen werden. Diese closed-loop Tests können mit verschiedenen Systemrealisierungsgraden durchgeführt werden. Tests dieser Art werden als X-in-the-Loop (XiL)-Tests bezeichnet [87, 88] (vgl. Abbildung 3.3). Abhängig von der Systemrealisierung handelt es sich um Model-in-the-Loop (MIL)-Tests, Software-in-the-Loop (SIL)-Tests, Hardware-in-the-Loop (HIL)-Tests oder Vehicle-in-the-Loop (ViL)-Tests.

Zum Testen des Gesamtsystems können ViL-Tests [89] eingesetzt werden, die insbesondere zum Nachstellen von kritischen Situationen geeignet sind. Dabei muss jedoch neben dem Fahrzeug als SUT selbst das gesamte Fahrzeugumfeld als Testplattform realisiert werden. Daher werden zum Testen des Gesamtsystems vorwiegend **Fahrzeugtests** durchgeführt. Dafür werden Prototypen und Versuchsträger aufgebaut, die einen Systemtest auf speziell ausgerüsteten Prüfgeländen oder unter realen Bedingungen im Straßenverkehr ermöglichen.

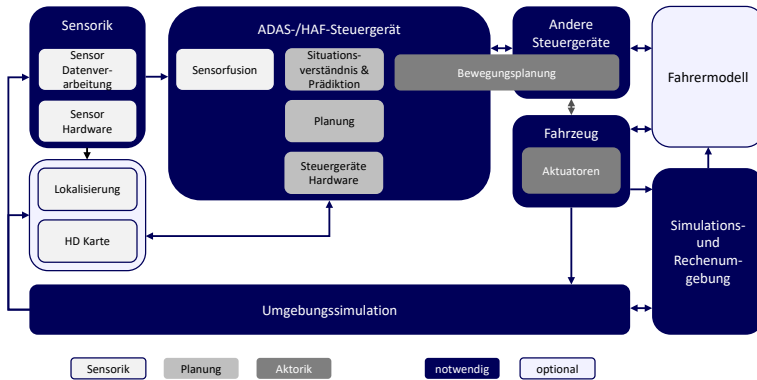


Abbildung 3.3: Das XiL-Prinzip für ADAS und HAF-Systeme. (in Anlehnung an Reisgys [71])

3.2.1 Fahrzeugtests auf dem Prüfgelände

Prüfgelände sind nicht öffentlich zugängliche, begrenzte Straßenabschnitte, die zur Durchführung von Fahrzeugtests angelegt werden. Sie eignen sich besonders für das gezielte Nachstellen einzelner Szenarien, die mit Hilfe von Dummy-Fahrzeugen und unter kontrollierten Bedingungen abseits des realen Straßenverkehrs erprobt werden können. Prüfgeländetests bieten innerhalb einer systemabhängig zu quantifizierenden Unschärfe durch die makroskopisch kontrollierbare Umgebung eine bedingte Reproduzierbarkeit und damit Vergleichbarkeit der durchgeführten Tests.

3.2.2 Fahrzeugtests im realen Straßenverkehr

Zur Durchführung von Fahrzeugtests im realen Straßenverkehr, folgend auch als **Straßenverkehrstests** bezeichnet, muss das gesamte Fahrzeug als SUT real aufgebaut werden. Je nach Reifegradstufe der einzelnen Komponenten geschieht dies als Prototypenfahrzeug, Null-Serie oder auch seriennahes Fahrzeug. Straßenverkehrstests werden in der Regel durch die an der Entwicklung und dem Test des SUT beteiligten Personen durchgeführt. Die Fahrzeugumgebung kann nur bedingt

beeinflusst werden. Sax et al. [68] sprechen daher von Erproben im Sinne des Ausprobierens mit einer vergleichsweise geringen Systematik statt Prüfen bzw. Testen im Sinne gezielter Analysen mit einer systematischen Herangehensweise.

Definition 29: Testfahrt

Eine Testfahrt ist ein Fahrzeugtest im realen Straßenverkehr mit dem Ziel, das SUT im Kontext des Gesamtsystems und späteren Einsatzgebiets zu erproben. Der Testinhalt kann dabei nur indirekt beeinflusst werden und unterliegt dem Zufall. Eine Testfahrt ist zeitlich und räumlich zusammenhängend und damit im Umfang auf wenige Stunden beschränkt.

Definition 30: Erprobung

Eine Erprobung ist ein Fahrzeugtest im realen Straßenverkehr mit dem Ziel, das SUT im Kontext des Gesamtsystems zu erproben und dieses einer hohen Varianz an (unbestimmten) Situationen auszusetzen [80]. Der Testinhalt kann dabei nur indirekt beeinflusst werden und unterliegt dem Zufall. Der Umfang einer Erprobung beträgt mehrere Tage bis wenige Wochen.

Erst im realen Straßenverkehr kann ein umfangreiches Bild der erzielten Sicherheit und Qualität des SUT gewonnen werden, denn beim Testen ist unbedingt das spätere Einsatzgebiet des SUT zu berücksichtigen [80]. Auch die abschließende Bewertung der Fahrer-Fahrzeug-Interaktion sowie der Bedienkonzepte ist erst im realen Straßenverkehr leistbar. Insbesondere die schwer zu fassenden Anforderungen an das erwartete, funktionale Verhalten des Fahrzeugs werden hier mit der impliziten Erwartungshaltung der Entwickler und Fahrverhaltensexperten abgeglichen. Dabei zählt das subjektive Empfinden der Erprobungsfahrer stellvertretend für alle späteren Nutzer. Wahlweise kann mittels Fahrstudien auch die Bewertung durch spätere Fahrzeugkunden mit unterschiedlichen Fahrkenntnissen einbezogen werden.

Die Umfänge der durchgeführten Straßenverkehrstests variiert stark. In einzelnen Testfahrten werden Entwicklungsstände getestet. Für die gesamtheitliche Qualifikation und Freigabe des SUT werden umfangreiche Erprobungen durchgeführt. Darüber hinaus gibt es für besonders wichtige Märkte dauerhaft eingesetzte Fahrzeugflotten (sog. Dauerläufer), die kontinuierlich Funktionen und Fahrzeugmodelle testen. Die Dauerläufer werden durch ihre Einsatzzwecke unterschieden. Für einzelne Domänen werden Domänen-Dauerläufer eingesetzt, um ausführliche Tests der Domänensteuergeräte und -funktionen durchzuführen (z.B. Fahrerassistenzdauerlauf (eng. ADAS-Endurance Run) zum Testen von Fahrerassistenzfunktionen). Für den Gesamtfahrzeugeindruck werden allgemeine Dauerläufer (eng. General Endurance) eingesetzt. Während der Entwicklung gibt es zudem spezielle Entwicklungs-Dauerläufer (eng. During-Development-Endurance-Run). Anlassbezogen können auch umfangreiche Testkampagnen definiert werden (z.B. Wintererprobung mehrerer Modelle und Fahrzeuge unter verschiedenen Witterungsbedingungen). [24]

Guhl et al. beschreiben in [90] zudem verschiedene Aspekte von Straßenverkehrstests. Neben dem Abfahren fest definierter Streckenprofile werden insbesondere auch Langstreckenfahrten zur Alterssimulation der Fahrzeuge, Hochgeschwindigkeitsanteile auf Autobahnen, Schlechtwegestrecken zum Testen von Extremsituationen sowie Nachtfahrten durchgeführt. Für ADAS wird zudem ein besonderer Fokus auf Überland- und Stadtfahrten gelegt.

3.2.3 Shadow Mode Testing

Beim Shadow Mode Testing handelt es sich um Fahrzeugtests im realen Straßenverkehr. Jedoch werden diese nicht aktiv durch Testflotten und -fahrer durchgeführt. Stattdessen wird in Kundenfahrzeugen während der Fahrer oder eine (teil-)automatisierte Fahrfunktion im Produktivbetrieb die Fahrzeugführung übernimmt im Hintergrund eine andere Fahrfunktion ausgeführt. Diese erhält dieselben Sensor- und Buskommunikationsdaten aus dem Fahrzeug wie die Fahrfunktion im Produktivbetrieb erhält bzw. erhalten würde. Dabei greift die Fahrfunktion

im Shadow Mode aber nicht aktiv in die Fahrzeugführung ein. Im Falle einer Abweichung ist daher keine Rückkopplung der Fahrzeugumgebung zum SUT möglich. Es handelt sich um einen open-loop Test. [91–93]

Kommt es zu Abweichungen zwischen dem Fahrer bzw. der aktiven Fahrfunktion und der Funktion im Shadow Mode, wird das Ereignis ins Backend übertragen, wo die Abweichung analysiert und mögliche Ursachen für die Abweichung untersucht werden. [92, 93]

Je nach Größe der Kundenflotte können durch Shadow Mode Testing hunderttausende - im Fall von Tesla mehrere Millionen - Testkilometer innerhalb eines Tages gesammelt werden. Abhängig von der Hardwareausstattung der Fahrzeuge können mehrere Funktionen parallel im Shadow Mode getestet werden. [92, 93]

3.3 Applikation auf Gesamtsystemebene

3.3.1 Ziel der Applikation

Teil der Fahrzeugvision, die in die Initialisierungsphase eines neuen PEP einfließt (vgl. Abschnitt 2.3 und Abbildung 2.4), ist die Definition der gewünschten **charakteristischen Fahrzeugeigenschaften**. Diese sind maßgebend für das Erscheinungsbild aber auch das Fahr-, Komfort- und Sicherheitsgefühl des zu entwickelnden Fahrzeugs. So kann hersteller- sowie segmentabhängig der Fokus auf Sportlichkeit, Komfort oder Luxus gelegt werden. Über Markenleitbilder der Automobilhersteller ist zudem eine klare Erwartungshaltung der Fahrzeugkäufer an ein bestimmtes Fahrgefühl bzw. Fahrverhalten assoziiert (vgl. Tabelle 3.1). Die festgelegten Fahrzeugeigenschaften dienen während der Entwicklung des Systems, der Subsysteme und Komponenten als Maßstab für Design- und Implementierungsentscheidungen (vgl. V-Modell in Abschnitt 2.4.2). [94]

Da die Zielausprägungen für die Fahrzeugeigenschaften auf Fahrzeug- bzw. Gesamtsystemebene definiert werden und verschiedene Subsysteme zu einem Ziel und einzelne Subsysteme zu mehreren Zielen beitragen, gibt es eine Vielzahl von

Marke	Markenleitbild	Assoziierte Fahrzeugeigenschaften
VW	Das Auto	werthaltig; innovativ; verantwortungsvoll
Audi	Vorsprung durch Technik	sportlich; progressiv; hochwertig
Škoda	Simply clever	präzise; funktional; geräumig
Seat	Enjoyneering	designorientiert; jung; dynamisch
Lamborghini	Der ultimative Sportwagen	extrem; kompromisslos; italienisch
Bentley	Driving Luxury Performance	kraftvoll; exquisit; individuell

Tabelle 3.1: Die Markenleitbilder und assoziierten Eigenschaften der Marken des VW-Konzerns (Quelle: [41]).

Wechselwirkungen. Durch die dezentrale Verantwortlichkeit für die Entwicklung der einzelnen Subsysteme ist eine eindeutige Zuordnung von Zuständigkeiten nicht realisierbar. Erst wenn alle Subsysteme entwickelt worden sind, ist eine Abstimmung dieser auf gemeinsame Ziele auf Gesamtsystemebene durchführbar. Dieser Prozess wird als Applikation bezeichnet. [35]

Definition 31: Applikation

Die Applikation ist der Prozess der Abstimmung von Komponenten und Subsystemen im Systemverbund sowie des Systems mit der Systemumgebung zur Erreichung der gewünschten charakteristischen Fahrzeugeigenschaften. In die Software integrierte Parameter werden während der Applikation so eingestellt, dass das gewünschte Systemverhalten erreicht wird.

3.3.2 Bewertung der Applikation

Für die Applikation von assistierenden und automatisierenden Fahrfunktionen spielen zwei Kriterien eine wesentliche Rolle:

Zum einen ist die allgemeine Verfügbarkeit des Systems für den Kundennutzen und damit den Mehrwert der Funktion maßgeblich.

Definition 32: Verfügbarkeit von ADAS- oder HAF-Funktionen

Die Verfügbarkeit beschreibt die Betriebsbereitschaft eines Systems oder einer Komponente [80].

Die Verfügbarkeit einer assistierenden bzw. automatisierenden Fahrfunktion beschreibt demnach die anteilige Zeit, in der die Funktion aufgrund der Umgebungssituation, Umgebungswahrnehmung und der eigenen Betriebsbereitschaft die Fahrzeugführung aktiv übernehmen bzw. aktiv unterstützen kann.

Der zweite Aspekt ist die für die Fahrzeuginsassen empfundene Qualität der Fahrzeugführung während das automatisierende bzw. assistierende System eingeschaltet ist. Die qualitative Bewertung setzt stets voraus, dass das System während der Fahrt bzw. in den zu bewertenden Situationen aktiv die Fahrzeugführung übernommen bzw. assistiert hat. Die Qualität kann unterschiedlich und systemabhängig definiert werden. Sie orientiert sich an den in der Initialisierungsphase des PEP definierten Zielgrößen für die Fahreigenschaften und -charakteristika (vgl. Abschnitt 2.3), die die geweckte Erwartungshaltung des Kunden an das Fahrgefühl beschreiben. Naus et al. definieren beispielsweise die Sicherheit und den Fahrkomfort als die entscheidenden Gütekriterien für ein Adaptive Cruise Control (ACC)-System [95].

Definition 33: Qualität von ADAS- oder HAF-Funktionen

Die Qualität einer assistierenden bzw. automatisierenden Fahrfunktion beschreibt die funktionale Güte der Funktion. Zur Funktionsqualität gehören neben dem spezifizierten Sollverhalten weitere Gütekriterien, wie die empfundene Sicherheit während des aktiven Fahreingriffs, der Fahrkomfort, die Genauigkeit der Manöverausführung, das zeitliche Verhalten sowie die Reaktivität des Systems.

Verfügbarkeit und Fahrgefühl werden maßgeblich durch das Zusammenspiel der einzelnen an der Längs- und Querführung beteiligten Funktionen beeinflusst. Die Abstimmung dieser Funktionen zu einem Gesamtbild findet durch die Applikation

statt. Um die Zielerreichung der festgelegten Fahrdynamik- und Fahrkomforteigenschaften messbar zu machen, bedarf es Methoden und Metriken zur Bewertung des Ist-Zustands und einem Abgleich mit dem gewünschten Zielverhalten. Der Ist-Zustand der Funktionsqualität wird subjektiv und/oder objektiv unter Zuhilfenahme von weichen oder harten Kriterien ermittelt (z.B. hart: mittlere Abweichung von der Spurmitte; weich: angenehme Geschwindigkeit in der Kurvendurchfahrt).

Zur objektiven Bewertung einer Funktion auf Basis aufgezeichneter Testfahrten müssen Bewertungsmetriken definiert werden, die die Beurteilung auf Basis eines in den Busdaten beobachtbaren Fahrzeugverhaltens bewerten. Für die Verfügbarkeit können dazu die Systemstatus-Informationen herangezogen werden. Die Funktionsqualität kann durch die Auswertung der Odometrie-Daten des Fahrzeugs (z.B. Längs- und Querruck sowie Beschleunigungen) oder die beobachtbare Interaktion mit der Fahrzeugumgebung (z.B. Abstand und Zeitlücke zum Vorderfahrzeug) bewertet werden (vgl. dazu auch Abschnitt 5.2.2). Für die Applikation werden ein oder mehrere Bewertungsmetriken eingesetzt. Die Aggregation konkurrierender Zielgrößen (z.B. Sportlichkeit vs. Effizienz) zu einem singulären Fitnesswert stellt dabei eine Herausforderung dar [96, 97]. Pareto-Kurven ermöglichen eine manuelle Abwägung in Konkurrenz stehender Ziele.

Definition 34: Bewertungsmetrik (einer Funktion)

Eine Bewertungsmetrik im Sinne der Funktionsapplikation ist eine Metrik, die eine quantitative Aussage über die Qualität einer Funktion in ihrem aktuellen Zustand - sowohl der Softwareversion als auch des hinterlegten Parametersatzes - aus dem beobachtbaren Fahrzeugverhalten und der Interaktion mit der Fahrzeugumgebung liefert.

Alternativ zur objektivierten Bewertung einer Fahrfunktion durch Metriken, können die eingesetzten Testfahrer mittels geeigneter Messtechnik eine subjektive Bewertung während der Applikationsfahrt abgeben. Dies ist insbesondere für Situationen hilfreich, die nur unzureichend durch die Fahrzeugsensorik wahrgenommen werden können (z.B. Einsehbarkeit von Kurven).

3.3.3 Durchführung der Applikation

Die Applikation beginnt mit den ersten Fahrzeugtests während der Entwicklung und zieht sich über die Reifegradstufen hinweg. In der Praxis findet sie häufig parallel zum Testen und zur Absicherung der Systeme statt. Die Applikation gewinnt mit zunehmendem Reifegrad an Bedeutung, da das Ausmaß und die Häufigkeiten von Änderungen an den Funktionen selbst zurückgehen (vgl. Abbildung 3.4). Die finale Applikation eines Fahrzeugs findet nach dem Function Freeze (RG6) statt, nachdem keine Änderungen am Softwarecode einer Funktion mehr möglich sind.

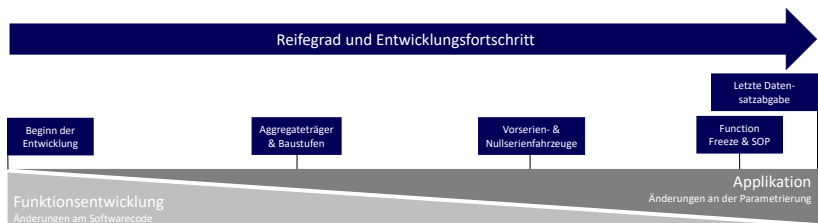


Abbildung 3.4: Verhältnis der Softwareentwicklung zu Applikation mit fortschreitendem Reifegrad der Funktion.

Um auch nach der Entwicklung der Subsysteme Freiheitsgrade zu erhalten, werden im Softwarecode der Subsysteme und Komponenten Parameter eingeführt, die das Funktionsverhalten beeinflussen und über die während der Applikation das Verhalten definiert werden kann. Sie dienen dazu, Änderungen an der Funktionalität des Systems von Änderungen am Softwarecode zu entkoppeln. Es wird eine losgelöste Entwicklung der einzelnen Funktionen vom Gesamtsystem ermöglicht. Durch die eingeführten Parameter sind während der Applikation keine Änderungen am Softwarecode notwendig.

Definition 35: Parameter

Ein Parameter ist eine Variable im Softwarecode, die während der Applikation mit einem festen Wert belegt (bedatet) und in einem Parametersatz festgeschrieben wird.

Definition 36: Parametrierung bzw. Parametersatz

Eine Parametrierung bzw. ein Parametersatz ist die Gesamtheit aller bedachten Parameter für ein konkretes Fahrzeug, Fahrzeugmodell oder Derivat. Ein Parametersatz wird im Fahrzeug hinterlegt, so dass die zugehörigen Funktionen auf die Parameter zugreifen können.

Die Durchführung der Applikation hängt maßgeblich vom Wirkungs- und Einflussbereich der Parameter ab. Insbesondere im ADAS und HAF Kontext interagieren die Funktionen und damit auch ihre Parameter mit der Fahrzeugumgebung. Die Bewertung der Fahrdynamik muss entsprechend den gesamten Regelkreis aus Fahrer bzw. Fahrfunktion, Fahrzeug und Umwelt einbeziehen [94]. Daher findet die Applikation von ADAS auf der Integrationsstufe des Gesamtfahrzeugs (vgl. Abschnitt 3.1) und damit in der realen Welt statt.

Beispielsweise müssen die minimale Zeitlücke vor dem Auslösen einer Notbremsung oder der minimale Abstand auf ein Vorderfahrzeug für ein ACC System abhängig vom jeweiligen Gesamtsystem appliziert werden. Unterschiedliche Bremseigenschaften des Fahrzeugs erfordern hier eine Feinabstimmung zur Sicherstellung eines sicheren und komfortablen Regelverhaltens.

Die starke Interaktion der ADAS und HAF-Funktionen mit der Systemumgebung führt zu einer hohen Anzahl zu applizierender Parameter. SAE Level 2 Systeme besitzen bereits 200-500 Parameter [34]. Die Zahl der Parameter nimmt für bedingt- und hochautomatisierende Fahrfunktionen weiter zu [35]. Zudem können sich Parameter unterschiedlicher Teilfunktionen gegenseitig beeinflussen, falls sie beispielsweise in denselben Umgebungssituationen die Wirkung ihrer jeweiligen Teilfunktion verändern.

Zur Applikation auf Gesamtsystemebene muss das zu applizierende System in einem Prototypenfahrzeug oder Versuchsträger lauffähig integriert werden. Die Applikation wird auf dem Prüfgelände oder im realen Straßenverkehr durchgeführt. ADAS wie EBA und ACC werden auf dem Prüfgelände durch genau spezifizierte Versuche appliziert. In meist klinisch nachgestellten Szenarien werden

dabei einzelne Verhaltenscharakteristika eingestellt, so dass diese den gesetzlichen und sicherheitsrelevanten Rahmenbedingungen (vgl. Abbildung 3.1) entsprechen und zudem ein subjektiv für gut empfundenes Gesamtfahrzeug- bzw. Funktionsverhalten aufzeigen. Diese Art der Applikation gewährleistet die notwendige Vergleichbarkeit unterschiedlicher Parameterwerte, da in zahlreichen direkt aufeinanderfolgenden Tests verschiedene Parametereinstellungen und deren Auswirkungen auf das Systemverhalten beobachtet und damit die beste Applikation gefunden werden können.

Da sich regional oder länderspezifisch die Systemumgebung, bspw. die Darstellung von Verkehrszeichen, die Verkehrsregeln oder Links- und Rechtsverkehr, stark unterscheiden, kann abhängig vom System eine länderspezifische Applikation notwendig sein. Dabei werden verschiedene Parametersätze vorgehalten, die je nach Ländercode eingesetzt werden.

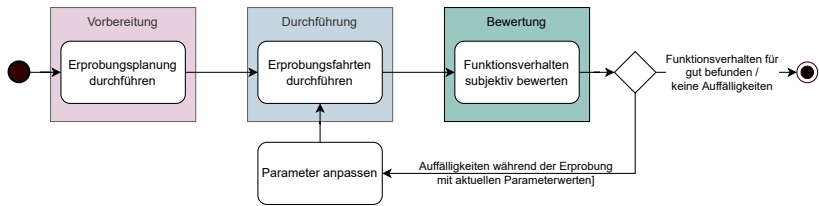


Abbildung 3.5: Für diese Dissertation relevante Teilschritte im Applikationsprozess in der Fahrerunterstützungsdomäne.

Der Applikationsprozess für die Applikation bedingt- und hochautomatisierender Funktionen im realen Straßenverkehr ist in drei Phasen unterteilt (vgl. Abbildung 3.5). Der Applikationsprozess beginnt mit der **Vorbereitungsphase**, in der insbesondere die Erprobungsfahrten geplant und terminiert werden. Zudem wird die Inbetriebnahme der Prototypenfahrzeuge durchgeführt [35]. Anschließend iterieren die **Durchführungs-** und **Bewertungsphase**. In der Durchführungsphase wird die zu applizierende Funktion aktiv im Fahrzeug während einer Erprobungsfahrt eingesetzt. Der Applikateur beobachtet das Funktionsverhalten und bewertet die Funktion subjektiv. Auf Basis dieser Bewertung wird die Entscheidung getroffen,

die Applikation zu beenden oder weitere Straßenverkehrstests durchzuführen und dabei ggf. Änderungen an einzelnen Parametern vorzunehmen.

4 Datengetriebene Entwicklung

Wie in den Abschnitten 3.2.1 und 3.2.2 beschrieben, werden umfangreiche Fahrzeugtests mit Prototypenfahrzeugen zur Verifikation und Validierung durchgeführt. Während das individuelle Erleben des Fahrzeugverhaltens und der subjektive Eindruck einen großen Mehrwert bieten, liegt in den durchgeführten Testfahrten weiteres Potenzial. Durch Aufzeichnung der fahrzeuginternen Buskommunikation und optional weiterer Informationen, wie Wetter- und Verkehrsdaten oder manueller Annotationen der Testfahrer, werden die durchgeführten Tests erfasst und gespeichert. Dies eröffnet neue Möglichkeiten der Datenanalyse und -bewertung, z.B. das Erlangen eines tieferen Funktionsverständnisses, die Rückverfolgung von Fehlern und das Nachvollziehen der Entscheidungsfindung der Algorithmen im Detail.

Darüber hinaus können aufgezeichnete Fahrten aggregiert und gesamtheitlich betrachtet werden. Das Ableiten von Mustern, Häufungen und Verteilungen führt zu datenbasierten Entscheidungsgrundlagen, die nicht nur auf einer Sammlung von Einzelfällen basieren. Dazu zählt auch eine statistische Betrachtung der Systemumgebung, z.B. Auftretenswahrscheinlichkeiten von häufigen, relevanten oder kritischen Szenarien.

Die Verwendung von Daten in Geschäfts- und Entwicklungsprozessen wird über die Automobilentwicklung hinaus als datengetriebene Entwicklung bezeichnet. Der Begriff ist nicht eindeutig definiert, subsummiert aber im Allgemeinen die Verwendung von Daten und Datenanalysen, um fundiertere, objektive Entscheidungen zu treffen. Dazu werden Key Performance Indicator (KPI) oder Objectives and Key Results (OKR) definiert (vgl. Definition 34). [98]

Definition 37: Datengetriebene Entwicklung

Datengetriebene Entwicklung bezeichnet ein Entwicklungsparadigma, bei dem Entscheidungen während der Entwicklung oder Wartung von Produkten oder Software auf Basis von Daten getroffen werden. Die Daten liefern akurates, umsetzbares Feedback für die Entwickler zur Verbesserung der Produkte oder Prozesse. [98, 99]

Zur Umsetzung einer datengetriebenen Entwicklung sind geeignete Datenverarbeitungs- und Datenanalyseverfahren notwendig. Während die Aufzeichnung und Auswertung von Daten - nicht nur aus Prototypenfahrzeugen sondern auch Kundenflotten und Dauerläufern (vgl. Abschnitte 3.2.2 und 3.2.3) - auch für weitere Anwendungsgebiete, wie den After-Sales-Bereich, die Diagnose und Wartung, neue Möglichkeiten eröffnen, beziehen sich die folgenden Betrachtungen auf die Entwicklung von ADAS und HAF-Funktionen.

4.1 Messdaten im Kontext der automobilen Softwareentwicklung

Informationen zwischen Funktionen und Steuergeräten im Fahrzeug werden über Bus-Systeme in Form von Nachrichten ausgetauscht. Diese Nachrichten können mittels geeigneter Messtechnik zur Laufzeit aufgezeichnet und gespeichert werden. Zu jeder Nachricht ist der Sender (die sendende Funktion), ein Zeitstempel und der Nachrichteninhalt bekannt. In einer Nachricht werden in der Regel mehrere Informationen (Signale) codiert versendet.

An den im Fahrzeug verbauten Sensoren, wie Kameras, Radar- und Lidarsensoren, fallen unverarbeitete Sensorrohdaten an. Kameras im Fahrzeug zeichnen Bilddaten bspw. mit 2048x1280 Pixeln und 45 Bildern pro Sekunde¹ auf. Das

¹ BOSCH Multi Purpose Camera [100]

entspricht bei 4 Farbkanälen und einer Farbtiefe von 8 Bits (entspricht 256 Farbnancen pro Farbkanal) einem Speicherbedarf von 471,9 Megabyte pro Sekunde (Mb/s). Radar- und Lidarsensoren liefern hochaufgelöste Punktwolken im dreidimensionalen Raum. Aufgrund des dafür notwendigen hohen Speicherplatzes werden diese Daten nicht dauerhaft erhoben, sondern lediglich in kleinen Studien aufgezeichnet. Sensorrohdaten sind daher im Kontext dieser Dissertation nicht relevant und werden nicht weiter betrachtet.

Neben Sensorrohdaten und Buskommunikationsdaten, die kontinuierlich bzw. zyklisch anfallen, sind auch Metadaten wie z.B. die Version der verwendeten Softwarekomponente (eng. Software Component (SWC)) und der Parametersatz vorhanden und für spätere Analysen relevant. Je nach Verfügbarkeit können zudem subjektive Fahreindrücke und manuell hinzugefügte Informationen als Sprachkommentare oder Textannotationen gesammelt werden.

Definition 38: Fahrzeugmessdaten

Unter dem Begriff Fahrzeugmessdaten werden im Folgenden die Gesamtheit der im Fahrzeug vorhandenen und aufgezeichneten Daten bezeichnet. Fahrzeugmessdaten bestehen aus Buskommunikationsdaten sowie Metadaten. Optional können Sensorrohdaten und Sprach- bzw. Textannotationen hinzugefügt werden.

Je nach eingesetzter Messtechnik werden die Messdaten in unterschiedlichen Formaten erzeugt. Verbreitet ist die Aufzeichnung mittels Automotive Data and Time-Triggered Framework (ADTF) im *.dat*-Format sowie das von der Association for Standardization of Automation and Measuring Systems (ASAM) standardisierte Measurement Data Format (MDF) [101]. Im Wesentlichen gibt es zwei Arten von Dateiformaten, binär-codierte Formate, die nur durch extern bereitgestellte Decodierschemata interpretierbar sind, und kombinierte Dateiformate, die neben den binär-codierten Rohdaten auch die Decodierinformationen zur Interpretation der Buskommunikationsdaten direkt inkludieren. Metadaten über die zeitlichen und geographischen Kennzahlen, das verwendete Fahrzeug, die Steuergerät- und

SWC-Versionen sowie Parametersätze werden je nach Konfiguration der Messtechnik automatisiert erzeugt.

4.2 Kollektion von Fahrzeugmessdaten

Testdaten fallen bei Tests während der Entwicklung und der Qualifikation über die Integrationsstufen hinweg an (vgl. Abschnitt 3.1). Bei Modul-, Komponenten- und Integrationstests können Testdaten aus den unterschiedlichen Testplattformen (z.B. SIL, HIL) aufgezeichnet und gespeichert werden. Während der Integration und Qualifikation in das Gesamtsystem fallen durch Prototypenfahrzeuge, Erprobungen, Dauerläufer und Flottendaten zudem Tausende bis mehrere Millionen Testkilometer im realen Straßenverkehr an. Nach Produktionsbeginn (eng. Start of Production (SOP)) können auch Serienfahrzeuge als Datenkollektoren eingesetzt werden.

Durch die Notwendigkeit spezieller Messtechnik werden die Buskommunikationsdaten nur in Prototypen, Testfahrzeugen und Erprobungsflotten der Automobilhersteller gesammelt. Über die Dauer eines oder mehrerer Fahrzeugprojekte mit jeweils 40 bis 58 Monaten Entwicklungsdauer (vgl. Abschnitt 2.3) entstehen so dennoch Aufzeichnungen über mehrere Millionen Testkilometer. So summieren sich die zurückgelegten Testkilometer für die Freigabe der Mercedes Benz E-Klasse W212 auf 36 Millionen Kilometer [102] und für die darauffolgende Generation W213 auf 12 Millionen Kilometer [103].

Aus Datenschutz- sowie technischen Gründen (z.B. Ausstattung, Speicherplatzbedarf, Übertragungsraten) werden in Kundenflotten nur aggregierte Daten aufgezeichnet und via Internet oder Mobilfunknetz ins Backend der Automobilhersteller gesendet. Im Gegenzug dazu fahren Kundenfahrzeuge in ihrer Gesamtheit in kürzester Zeit mehrere Millionen Kilometer und liefern so wertvolle Gesamteindrücke und Statistiken. Ansätze zur situativen Datenkollektion versuchen hier weitere Potenziale zu erschließen. [92, 104]

Für eine Analyse im Backend eignet sich die auf Effizienz ausgelegte Form der Datenübertragung im Fahrzeug nicht, da die Signalinformationen in verschiedenen Nachrichten codiert und versendet werden. Um den Speicherbedarf einzelner Signale zu minimieren, werden die physikalischen Signalwerte in den Nachrichten mittels signalspezifischer Offsets und Skalierungen in wenigen Bits codiert. Eine Decodierung erfolgt im Backend anhand festgelegter Decodierungsvorschriften - z.B. in Form einer Kommunikationsmatrix, die neben den Offsets und Skalierungsfaktoren zusätzlich die Zykluszeiten sowie Default-, Fehler- und Initialwerte der einzelnen Signale enthält. [24]

Beispiel 1: Die geographische Breite und Länge werden zur Übertragung im Fahrzeug über Bussysteme speichereffizient codiert und müssen zur Interpretation in physikalische Werte decodiert werden (vgl. Abbildung 4.1). Die geographische Breite mit Werten zwischen -90° bis 90° wird wie folgt codiert: 1 Vorzeichen-Bit für die nördliche bzw. südliche geographische Breite und 17 Bits für eine auf 3 Nachkommastellen genaue Darstellung der Dezimalzahl in Grad mit einem Skalierungsfaktor von 1000; die nördliche geographische Breite von $49,013^\circ$ entspräche dann 001011111101110101 (vereinfacht, üblicherweise 6 Nachkommastellen und damit 27 Bits). Analog kann die geographische Länge codiert werden. Aufgrund des größeren Wertebereichs von -180° bis 180° wird ein weiteres Bit benötigt.

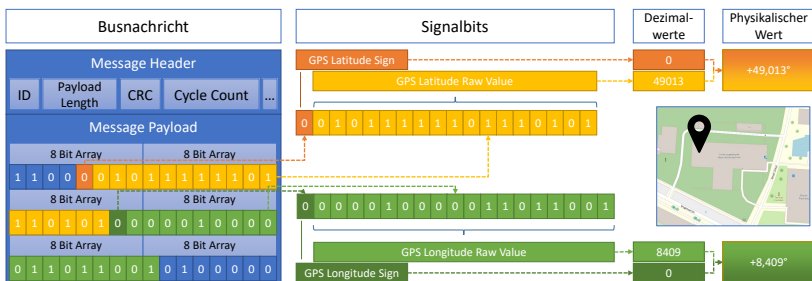


Abbildung 4.1: Schematische Darstellung der Decodierungsschritte für Signale einer Busnachricht am Beispiel der geographischen Länge und Breite.

Für die Analyse ist zudem eine Umwandlung der als Nachrichten codierten Signale zu Zeitreihen-Daten notwendig. Bei Zeitreihen-Daten werden die Signale der verschiedenen Nachrichten auf einen einheitlichen Zeitstempel gebracht, um anschließend mit zeitreihenbasierten Analyseverfahren ausgewertet werden zu können. Dabei wird stets der zuletzt übertragene Signalwert gehalten bis zum Eintreffen einer neuen Nachricht mit einem aktualisierten Signalwert.

Beispiel 2: Busnachrichten eines Odometriesteuergeräts sowie des Navigationssteuergeräts werden zeitversetzt und asynchron auf dem Bus übertragen (vgl. Abbildung 4.2). Die Nachrichten werden auf ein einheitliches Zeitintervall von 20ms abgebildet. Die alten Werte werden bis zum Eintreffen neuer Busnachrichten mit aktualisieren Informationen gehalten.

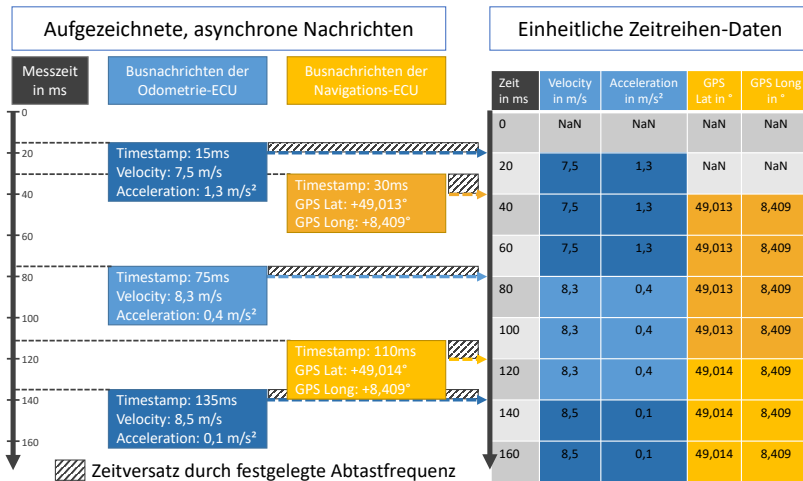


Abbildung 4.2: Schematische Darstellung der Umwandlung zweier Nachrichten in einheitliche Zeitreihen-Daten mit 20ms Abtast-Intervallen.

Fahrzeugmessdaten unterliegen Störeinflüssen und Ungenauigkeiten, beispielsweise bedingt durch Sensorrauschen oder Umgebungseinflüsse. Solche Verzerrungen einzelner Informationen werden durch Signalverarbeitungsmethoden während

der Datenaufbereitung entfernt oder abgemildert. Glättung, Interpolation und Entfernen von Ausreißerwerten sind exemplarische Methoden, um die Datenqualität zu erhöhen. Zusätzlich können weitere Datenquellen, die im Fahrzeug nicht zur Verfügung stehen, hinzugezogen werden, um die Datenqualität weiter zu verbessern.

4.3 Verwendung und Mehrwert der Fahrzeugmessdaten

Fahrzeugmessdaten fallen während der durchgeführten Fahrzeugtests als Nebenprodukt an. Während das Potenzial dieser Daten zunächst stark unterschätzt und die Daten oftmals noch nicht oder nicht vollständig genutzt wurden, haben sich in den vergangenen Jahren die Fahrzeugmessdaten zu einem zentralen Werkzeug in der Entwicklung von ADAS und HAF-Funktionen etabliert. Es wird von datengetriebener Entwicklung (Data-Driven-Development, Big Loop, Data Loop) gesprochen [JL1, 105, 106] (vgl. Definition 37).

Potenzielle Mehrwerte für die Datenauswertung der Fahrzeugkommunikation sind vielseitig und erstrecken sich über alle Entwicklungsphasen im V-Modell (vgl. Abbildung 2.5).² Die für diese Disseration relevanten Potenziale werden im Folgenden kurz erläutert.

Aufgezeichnete Fahrzeugmessdaten werden in aggregierter Form dazu verwendet, ein auf Daten bzw. Fakten basiertes Verständnis des Systems und der späteren Systemumgebung zu erhalten. Dazu werden unter anderem aus den Realdaten typische Einsatzszenarien extrahiert und bei Bedarf abstrahiert. Des Weiteren lassen sich Statistiken, Häufungen und Verteilungen einzelner Signale oder abstrahierter Systeminformationen generieren. [JL1]

² Veröffentlicht in *Data-Driven Development: A Complementing Approach for Automotive Systems Engineering* [JL1] und hier zusammenfassend dargestellt

Die aufgezeichneten Fahrzeugmessdaten dienen zudem als realistische Teststimuli [107] zum Testen der Implementierung mittels open-loop [108] oder closed-loop³ [109] Simulation. Die erstellten simulationsbasierten Tests werden während der Implementierung und Qualifikation verwendet. Zusätzlich werden einmalig durchgeführte Straßenverkehrstests in späteren Iterationszyklen als Input für open- und closed-loop Simulationstests wiederverwendet. [JL1]

Durch die aufgezeichneten Straßenverkehrstests ist eine Variation der Straßentopologien, Wetterbedingungen und Verkehrsszenarien in der Simulation möglich, ohne diese manuell modellieren zu müssen. Zudem sind Fahrten in verschiedenen Regionen und Ländern simulierbar, die die länderspezifischen Besonderheiten enthalten. Durch die Skalierbarkeit der Simulation und die im Vergleich zum Fahrzeugtest geringen Aufwände für einen Softwaretest kann die Testabdeckung bereits vor der vollständigen Integration erhöht werden. [JL1]

4.4 Analyse und Strukturierung der aufgezeichneten Fahrzeugtests

Erprobungen (vgl. Definition 30) liegt eine geringe Systematik zu Grunde. Durch die geringen Einflussmöglichkeiten auf die Fahrzeugumgebung treten Situationen und Szenarien während der Erprobung zufällig und unterschiedlich oft auf. Bei der Analyse der Aufzeichnungen muss daher ein besonderes Augenmerk auf den Inhalt der Daten gelegt werden. Je nach Analyseziel sind zudem die Zielverteilungen unterschiedlich zu bewerten. Während bei Fehlerbetrachtungen jedes Szenario eine potenziell gleichbedeutende Rolle einnimmt, sind bei Qualitätsaspekten häufig vorkommende Szenarien höher zu bewerten, da diese für die Insassen häufiger wahrzunehmen sind. Für beide Analyseziele gilt, dass die Daten nicht ohne Weiteres aggregiert und als repräsentativ für den Analysezweck angenommen werden sollten.

³ z.B. mittels Reactive Replay für längsführende Regelungsfunktionen

Definition 39: Repräsentativität

Repräsentativität liegt dann vor, wenn die Teilmenge ein wirklichkeitsgetreues, verkleinertes Abbild der Grundgesamtheit darstellt, also die gleichen Merkmale aufweist. [110]

4.4.1 Fahrzeugtests als Stichproben einer Grundgesamtheit

Für die Bewertung der Repräsentativität eines Datensatzes muss zunächst die zu repräsentierende Grundgesamtheit definiert werden. Ausgehend davon kann die Art der Repräsentation festgelegt werden - beispielsweise der Häufigkeit entsprechend oder die einmalige Abdeckung aller Szenarien. Im Bereich der Straßenverkehrstests gibt es mehrere Möglichkeiten, die Grundgesamtheit zu definieren. So kann das weltweite Straßennetz als Grundgesamtheit für das spätere Einsatzszenario des Fahrzeugs festgelegt werden. Eine andere Möglichkeit stellt die Reduktion des Straßennetzes auf die von den Kunden gefahrenen Strecken dar, die beispielsweise über Flottendaten ermittelt werden. Weitere Möglichkeiten sind die Einschränkung auf bestimmte Länder oder eine Betrachtung der dynamischen Situationen oder Wetterbedingungen und damit einer gänzlich anderen Definition der Grundgesamtheit. Die Grundgesamtheit muss abhängig vom Analyseziel definiert werden.

Definition 40: Grundgesamtheit

Die statistischen Einheiten, zu denen eine Aussage getroffen werden soll, wird als Grundgesamtheit bezeichnet. [111]

Da es nicht möglich ist, die Grundgesamtheit in den Fahrzeugtests vollständig zu testen, müssen repräsentative Teilmengen davon untersucht und anschließend auf das große Ganze geschlossen werden. Die Teilmengen werden mittels Stichprobenverfahren gebildet. Im Einzelfall gilt es dabei zu prüfen, ob ein Fahrzeugtest per

se eine valide Stichprobe darstellt oder ob durch Datenanalyseverfahren zunächst noch eine repräsentative Stichprobe gebildet werden muss.

Definition 41: Stichprobe

Als Stichprobe wird ein Teil einer Gesamtheit bezeichnet, der nach einem bestimmten Auswahlverfahren zustande gekommen ist. [112]

Definition 42: Zufallsstichprobe

Eine Zufallsstichprobe ist dadurch charakterisiert, dass jedes Element der Grundgesamtheit mit gleicher Wahrscheinlichkeit in die Stichprobe aufgenommen werden kann. [113]

Definition 43: Geschichtete Stichprobe

Eine geschichtete Stichprobe wird hinsichtlich der bekannten, relevanten Einflussgrößen repräsentativ gezogen. D.h. die Verteilung der bekannten Merkmale der Grundgesamtheit wird in der Stichprobe abgebildet. Die Auswahl innerhalb der Schichten erfolgt zufällig. [113]

4.4.2 Bias

Eine Stichprobe kann probabilistisch aus der Grundgesamtheit gezogen werden oder durch nicht probabilistische Verfahren über deskriptive Features bestimmt werden. Probabilistische Verfahren eignen sich bei großen Stichproben. Bei kleinen Stichproben kann die Zufallsverteilung nicht vorausgesetzt werden, daher werden nicht probabilistische Verfahren eingesetzt, die jedoch anfällig gegenüber einem Auswahl Bias sind. Ein systematisches Vorgehen ist zwingend notwendig. [114]

Definition 44: Merkmal

Ein Merkmal ist eine Eigenschaft eines Untersuchungsobjekts, welches dieses genauer beschreibt. Dabei werden unveränderliche Konstanten und Variablen mit verschiedenen Ausprägungen unterschieden. Ein Merkmal kann manifest, also direkt beobachtbar bzw. leicht feststellbar, oder latent, also nicht direkt beobachtbar, sein. [114]

Definition 45: Variable

Die Variable ist ein Merkmal bzw. deskriptives Feature, das verschiedene Ausprägungen annehmen kann. Es wird zwischen stetigen Variablen mit unendlich vielen, fein abgestuften Ausprägungen und diskreten Variablen mit einer kleinen und begrenzten Zahl an Ausprägungen unterschieden. [114]

Definition 46: Bias

Als Bias wird eine systematische Über- oder Unterschätzung eines Schätzers gegenüber dem wahren Wert der Grundgesamtheit bezeichnet. [115]

4.4.3 Repräsentativität einer Stichprobe

Um die Repräsentativität einer vorhandenen Stichprobe S bezüglich der Grundgesamtheit G zu quantifizieren, kann der Stichprobenfehler berechnet werden.

Definition 47: Stichprobenfehler

Der Stichprobenfehler, auch als Standardfehler des Mittelwerts bezeichnet, gibt an, wie sehr der Mittelwert eines Merkmals der Stichprobe vom Mittelwert desselben Merkmals der Grundgesamtheit abweicht. [116]

Bei einer Zufallsstichprobe stellt der Stichprobenfehler die zufällige Abweichung vom Erwartungswert dar und lässt sich mit der Stichprobengröße verringern [115]. Der Stichprobenfehler $\sigma_{\bar{x}}$ ergibt sich aus der Standardabweichung der Grundgesamtheit σ_G sowie der Stichprobengröße n :

$$\sigma_{\bar{x}} = \frac{\sigma_G}{\sqrt{n}} \quad (4.1)$$

In der Regel ist jedoch die Standardabweichung der Grundgesamtheit unbekannt und kann nicht ermittelt werden. Daher kann zur Annäherung der geschätzte Stichprobenfehler $\hat{\sigma}_{\bar{x}}$ mit der Standardabweichung der Stichprobe σ_S berechnet werden.

$$\hat{\sigma}_{\bar{x}} = \frac{\sigma_S}{\sqrt{n}} \quad (4.2)$$

Die Standardabweichung σ beschreibt dabei die durchschnittliche Abweichung vom Erwartungswert. Der Erwartungswert μ ergibt sich aus der Verteilung der Grundgesamtheit. Falls die Verteilung der Grundgesamtheit nicht bekannt ist, muss der arithmetische Mittelwert einer möglichst großen Stichprobe als Erwartungswert verwendet werden.

Für die Berechnung des Stichprobenfehlers einer empirischen Verteilung gegen eine erwartete Verteilung gibt es verschiedene statistische Maße:

Kolmogorow-Smirnow-Test (KS-Test): Der KS-Test vergleicht die empirische Verteilungsfunktion (eng. Empirical Cumulative Distribution Function (ECDF)) mit der kumulativen Verteilungsfunktion (eng. Cumulative Distribution Function (CDF)) der erwarteten Verteilung. [117]

Beim KS-Test werden zwei Hauptwerte berechnet:

Die *Teststatistik (Kolmogorow-Smirnow-Statistik)* misst den maximalen vertikalen Abstand zwischen der ECDF und der CDF der erwarteten Verteilung. Ein

größerer Wert der Teststatistik deutet darauf hin, dass die empirische Verteilung und die erwartete Verteilung weniger gut übereinstimmen.

Der *p-Wert* gibt die Wahrscheinlichkeit an, dass die beobachteten Daten aus der erwarteten Verteilung stammen. Dazu wird eine Nullhypothese aufgestellt, die besagt, dass die beobachteten Daten aus der erwarteten Verteilung stammen. Ein kleiner *p-Wert* deutet darauf hin, dass die beobachteten Daten signifikant von der erwarteten Verteilung abweichen, während ein großer *p-Wert* darauf hindeutet, dass die beobachteten Daten gut mit der erwarteten Verteilung übereinstimmen.

In der Regel wird ein Signifikanzniveau (oft 0,05 oder 0,01) festgelegt. Wenn der *p-Wert* kleiner als dieses Signifikanzniveau ist, wird die Nullhypothese abgelehnt und es wird angenommen, dass die beobachteten Daten nicht aus der erwarteten Verteilung stammen. Andernfalls wird die Nullhypothese beibehalten, was darauf hinweist, dass die beobachteten Daten wahrscheinlich aus der erwarteten Verteilung stammen.

Als Nullhypothese wird postuliert, dass die ECDF der CDF einer Grundgesamtheit entspricht. Die Grundgesamtheit kann bekannt sein oder durch eine angenommene Verteilungsfunktion (z.B. Normalverteilung) approximiert werden.

Der KS-Test beginnt mit der Berechnung der konkreten empirischen Verteilungsfunktion aus einer beobachteten Häufigkeit $S(x_i)$ mit n Beobachtungswerten $x_i, 1 \leq i \leq n$:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(-\infty, x]}(x_i), \quad x \in \mathbb{R} \quad (4.3)$$

Anschließend wird die konkrete empirische Verteilungsfunktion mit der erwarteten Verteilungsfunktion $F_0(x)$ verglichen. Dafür werden eine Null- und Alternativhypothese aufgestellt:

Nullhypothese H_0 : $F_n(x) = F_0(x)$

Alternativhypothese H_1 : $F_n(x) \neq F_0(x)$

Die Teststatistik D ergibt sich aus

$$D_n = \sup_{x \in \mathbb{R}} |F_n(x) - F_0(x)| \quad (4.4)$$

wobei \sup das Supremum bzw. die kleinste obere Schranke ist.

Für ein Signifikanzniveau α wird die Nullhypothese zugunsten der Alternativhypothese abgelehnt, falls d_n das $(1 - \alpha)$ -Quantil übersteigt. Das $(1 - \alpha)$ -Quantil wird für kleine n aus Tabellen abgelesen und für große n numerisch ermittelt [118].

Chi-Quadrat-Test: Der Chi-Quadrat-Test vergleicht die beobachtete Häufigkeit von Ereignissen mit der erwarteten Häufigkeit unter der Annahme einer bestimmten Verteilung. Die Chi-Quadrat-Statistik gibt an, wie gut die beobachteten Daten mit der erwarteten Verteilung übereinstimmen. [117]

Für eine erwartete absolute Häufigkeit n_{0j} wird aus der Wahrscheinlichkeit p_{0j} und der Anzahl der in der Stichprobe vorhandenen Beobachtungen n berechnet: $n_{0j} = p_{0j} * n$. Damit kann für eine vorliegende Stichprobe mit den beobachteten Häufigkeiten N_j die Chi-Quadrat-Statistik berechnet werden:

$$X^2 = \sum_{j=1}^m \frac{(N_j - n_{0j})^2}{n_{0j}} \quad (4.5)$$

Kullback-Leibler-Divergenz (KL-Divergenz): Die KL-Divergenz misst die Ähnlichkeit zwischen zwei Wahrscheinlichkeitsverteilungen, indem sie die Information verfolgt, die verloren geht, wenn eine Verteilung anstelle einer anderen verwendet wird. [119]

Die KL-Divergenz ist für diskrete und stetige Verteilungen unterschiedlich definiert. Für den Fall zweier diskreter Wahrscheinlichkeitsverteilungen P und Q mit $p(x) = P(x)$ und respektive $q(x) = Q(x)$ ist die KL-Divergenz D definiert als

$$D(P\|Q) = \sum_{x \in X} p(x) \log \left(\frac{p(x)}{q(x)} \right) \quad (4.6)$$

Root Mean Squared Error (RMSE): Der RMSE ist ein Maß für die mittlere quadratische Abweichung zwischen den Werten einer empirischen Verteilung und den entsprechenden Werten der erwarteten Verteilung.

Der RMSE ergibt sich aus

$$e = \sqrt{\sum_{i=1}^n e_i^2 / n} \quad (4.7)$$

wobei $e_i = O_i - E_i$ der Fehler für die Beobachtung i ist, die sich aus der Erwartung E_i und Beobachtung O_i berechnet.

Welches Maß am besten geeignet ist, hängt von der Art der Daten und der Fragestellung ab.

5 Herausforderungen und Chancen in der Applikation von ADAS und HAF-Funktionen mittels Fahrzeugtests

Die Funktionsumfänge von heutigen ADAS und HAF-Systemen steigen stetig an (vgl. Abschnitt 2.4.3). Durch die Übernahme von Fahraufgaben in unterschiedlichen Umgebungssituationen erhöht sich die Sicherheit und der Fahrkomfort während der assistierten bzw. automatisierten Fahrt. Allerdings steigen auch die während der Applikation zu berücksichtigten Fahraufgaben und Umgebungssituationen sowie die zu applizierenden Parameter an (vgl. Abschnitt 3.3).

Für die Applikation müssen umfangreiche Straßenverkehrstests durchgeführt werden. In der Aufzeichnung und systematischen Analyse dieser Fahrzeugtests liegen jedoch auch Potenziale, den Herausforderungen in der Applikation von ADAS und HAF-Funktionen zu begegnen (vgl. Abschnitt 4.3).

5.1 Herausforderungen in der Applikation von ADAS und HAF

5.1.1 Vergleichbarkeit von unterschiedlichen Erprobungsfahrten

Während der Applikation werden verschiedene Parameterausprägungen des ADAS bzw. HAF-Systems gegeneinander verglichen und eine optimale Parameterkombination gesucht. Die in Straßenverkehrstests auftretenden Situationen und Bedingungen sind jedoch nicht bzw. der statische Anteil bedingt reproduzierbar, da die Fahrzeugumgebung nicht beeinflusst werden kann und die Umgebungsfaktoren in ihrer jeweiligen Ausprägung einen hochdimensionalen Raum aufspannen (vgl. Abschnitt 3.2.2). Somit ist ein direkter Vergleich unterschiedlicher Parameterausprägungen nicht möglich. Selbst wenn dieselbe Strecke mehrfach direkt hintereinander durchfahren werden würde, ließen sich andere Verkehrsteilnehmer, besondere Kombinationen aus Umgebungssituationen sowie der interne Zustand des SUT nicht exakt reproduzieren. Der Aufwand für das mehrfache Wiederholen einer Situation bzw. eines Szenarios wäre zudem um ein Vielfaches höher, da jeweils zum Aufnahmeort zurückgekehrt, dieser mit ausreichend Vorlauf zum Aufbau der Umgebungsmodelle und Einschwingen der Regelalgorithmen angefahren und die dynamische Fahrzeugumgebung reproduziert werden müssten. Für ein Szenario mit einer Dauer von 10 Sekunden ist mit einem Vorlauf für den Aufbau des Umgebungsmodells inklusive Straßentopologie aus der elektronischen Karte von mindestens 60 Sekunden und mehreren Minuten für das Wenden und Rückkehren in den Ausgangszustand zu rechnen. Jedes einzelne Szenario müsste dabei für alle Parameterkombinationen wiederholt werden, was sowohl für die Anzahl der Szenarien als auch der Parameterkombinationen über den wirtschaftlichen, organisatorischen und technischen Möglichkeiten liegen würde.

Eine Alternative zur wiederholten Durchführung in der Realität ist die Re-Simulation aufgezeichneter Testfahrten¹. Die Anwendbarkeit für die Applikation einer prädiktiven Längsregelung wurde in einem Experiment untersucht:

Dazu wurden 21 Testfahrten in einer closed-loop Simulation eingespielt und die Auswirkung einer einzelnen Parameteränderung auf die Güte der Funktion - in diesem Fall die Prädiktionsgüte für die Kurvengeschwindigkeit - über die Fahrten hinweg bewertet, wobei der Parameter für den Unsicherheitsfaktor für die Kurvenkrümmungen aus der elektronischen Karte und der Kamera variiert wurde (vgl. Abbildung 5.1). Die Analyse zeigt, dass die Auswirkungen der Parameteränderung (Y-Achse) auf die Bewertungsmetrik (Z-Achse) einen geringeren Einfluss haben als die unterschiedlichen Fahrten (X-Achse). Die betrachtete Bewertungsmetrik variiert über die einzelnen Fahrten entlang der X-Achse stärker als über einzelne Parameteränderungen entlang der Y-Achse. Eine Bewertung der Parameteränderungen auf Basis der gewählten Stichprobe von Straßenverkehrstests ist nicht möglich, da die geänderte Funktionsqualität nicht eindeutig auf die Parameteränderungen zurückzuführen ist und die Ergebnisse von der gewählten Stichprobe von Straßenverkehrstests abhängen.

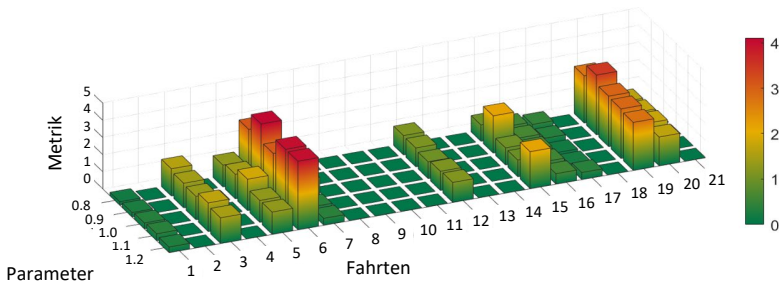


Abbildung 5.1: Auswirkung einer Parameteränderung auf das SUT in verschiedenen Straßenverkehrstests

¹ z.B. Reactive-Replay [58] als closed-loop Simulation für die prädiktive Längsregelung

Um zu untersuchen, ob sich ab einer bestimmten Anzahl an Straßenverkehrstests ein repräsentatives Mittel einstellt, soll das Konvergenzverhalten der Bewertungsmetrik über die Anzahl an Straßenverkehrstests untersucht werden. Dazu wurde ein Datensatz aus 107 Fahrten mit einer Gesamtstrecke von rund 3.750km ausgewählt. Die Fahrten wurden zufällig ausgewählt und decken eine hohe Varianz an Situationen des realen Straßenverkehrs, darunter Stadt- und Überlandfahrten, Fahrten mit geringem und hohem Verkehrsaufkommen, Fahrten mit hoher und niedriger Durchschnittsgeschwindigkeit, etc., ab. Jeder einzelne Straßenverkehrstest wurde für sich mit der bereits im vorangegangenen Beispiel verwendeten Metrik bewertet.

Zur Untersuchung des Konvergenzverhaltens wurden 10.000 zufällige Permutationen ohne Wiederholung der aufgezeichneten Straßenverkehrstests aus dem bewerteten Datensatz gezogen. Das heißt, es wurden 10.000 zufällige Stichproben von Straßenverkehrstests simuliert, indem zunächst jeweils eine zufällige Fahrt ausgewählt und iterativ weitere Fahrten zur Stichprobe hinzugefügt wurden. Der zu jedem Iterationsschritt berechnete Mittelwert über die gefahrene Strecke der aktuellen Stichprobe ist in Abbildung 5.2 dargestellt. Der Mittelwert über alle Fahrten liegt bei 1,8. Eine Annäherung an diesen tatsächlichen Mittelwert - also eine Konvergenz der zufällig gezogenen Stichproben gegen den tatsächlichen Mittelwert - ist nicht zu beobachten. Erst mit Hinzunahme fast aller Fahrten stellt sich ein Wert um 1,8 zuverlässig ein. Die Streuung der Mittelwerte über alle Permutationen ist selbst bei 3.000km bzw. dem Ziehen von über 80% aller Fahrten noch $\pm 0,3$ bzw. $\pm 16,67\%$ und damit zu hoch für einen Rückschluss auf die Applikationsgüte. Die Auswahl eines kleinen Testsets bzw. einer kleinen Stichprobe kann starke Abweichungen in der Funktionsbewertung nach unten oder oben mit sich ziehen.

Das schlechte Konvergenzverhalten ist darauf zurückzuführen, dass die Varianz der Umgebungssituationen in den Straßenverkehrstests sehr hoch ist. Manche - potenziell sehr relevante und kritische Situationen - treten nur in wenigen Fahrten auf. Diese haben bei Hinzunahme bzw. Entfernen aus dem Testdatensatz einen großen Einfluss auf die durchschnittliche Bewertung der Funktion.

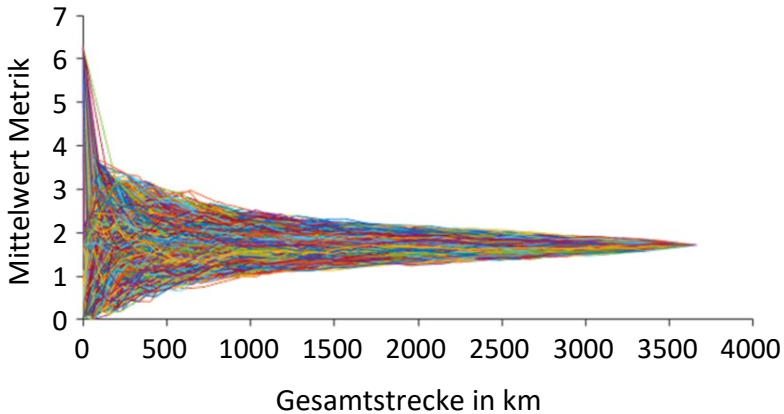


Abbildung 5.2: Permutation von bewerteten Straßenverkehrstests mit einer Gesamtlänge von ca. 3.750km. Eine Annäherung an die tatsächliche Durchschnittsbewertung ist erst bei Hinzunahme fast aller Fahrten gewährleistet.

Für die Applikation der hier untersuchten Funktion und Metrik bedeutet dies, dass für jede einzelne Parameteränderung mindestens 3.750km gefahren werden müssten - wobei auch für diese Strecke nicht abschließend geklärt ist, ob es sich dabei bereits um einen global repräsentativen Mittelwert handelt oder bei Hinzunahme weiterer Testfahrten sich eine noch größere Distanz ergeben würde. Da diese Distanz für jeden Parametersatz während der Applikation gefahren werden müsste, ergibt sich daraus selbst unter der Annahme, dass diese Distanz ausreicht, bei nur drei Parametern mit jeweils fünf Ausprägungen eine Gesamtstrecke von ~468.000km bzw. ~585 Erprobungstagen² für die gesamte Applikation.

Diese Voruntersuchung zeigt, dass ein Vergleich auf der Ebene ganzer Straßenverkehrstests (vgl. Abbildung 5.1) nicht valide ist. Die exemplarisch verwendete, prädiktive Längsregelung ist von äußeren Umgebungseinflüssen abhängig, die sich zwischen einzelnen Fahrten unterscheiden. Für die prädiktive Kurvenregelung ist die Kurvenbeschaffenheit, also z.B. die Kurvenkrümmung, der Kurvenradius und

² unter Annahme einer Durchschnittsgeschwindigkeit von 80km/h, einem Erprobungsfahrzeug und 10h Fahrzeit pro Tag

die zulässige Höchstgeschwindigkeit, aber auch die dynamische Situation, also das Vorhandensein eines potenziell langsameren Vorderfahrzeugs, die Geschwindigkeit des eigenen Fahrzeugs vor der Kurve, sowie die elektronischen Karteninformationen über die Kurve von Relevanz. Die in einem Straßenverkehrstest enthaltene Anzahl von Kurven kann stark variieren. Zudem sind die Anzahl und die Wertebereiche der relevanten Umgebungseinflüsse zu groß, als dass sich ein statistisches Mittel über die Dauer einer einzelnen Fahrt einstellen würde. Ein Vergleich unterschiedlicher Parameterausprägungen bezüglich der Prädiktionsgüte für die Kurvengeschwindigkeit auf Fahrtebene ist nicht möglich. Im Gegenteil kann ein undifferenzierter Vergleich zweier Parameterausprägungen auf Fahrtebene zu einer suboptimalen Parameterauswahl und damit zu einer Verschlechterung der Funktionsqualität führen. Die Vergleichbarkeit von Straßenverkehrstests muss sichergestellt werden.

5.1.2 Repräsentativität von Applikations- und Erprobungsfahrten

Testfahrten und Erprobungen als Grundlage der Applikation sind solange eine valide Lösung, wie sie die innerhalb der ODD auftretenden und für die zu applizierende Funktion relevanten Situationen repräsentativ widerspiegeln. Nur so ist gewährleistet, dass eine auf den Fahrversuchen und Erprobungen basierende Bewertung der Applikation nicht vom Fahrerlebnis abweicht, welches spätere Nutzer der Funktion im Mittel erleben werden.

Definition 48: Repräsentativität von Erprobungsfahrten

Die Repräsentativität einer Menge von Straßenverkehrstests bezeichnet die Validität und Übertragbarkeit einer auf dieser Menge (Stichprobe) hergeleiteten Aussage für die definierte Gesamtheit. Diese kann im Kontext von Straßenverkehrstests das weltweite Straßennetz, eine bestimmte Region oder eine aus Flottendaten gewichtete Teilmenge befahrener Strecken sein.

Charakteristikum	Erprobungsfahrt 1	Erprobungsfahrt 2	Erprobungsfahrt 3	Grundgesamtheit
Umfang	365,32 km	337,35 km	142,26 km	57.316,89 km
Dauer	6,61 h	5,61 h	2,29 h	944,82 h

Tabelle 5.1: Überblick über die drei verwendeten Erprobungsfahrten in der Voruntersuchung zur Repräsentativität.

Die Annahme, dass sich über eine gesamte Erprobung hinweg ein statistisches Mittel einstellt und diese somit ohne eine detaillierte Analyse der einzelnen Situationen für die Grundgesamtheit repräsentativ ist, soll in einer weiteren Voruntersuchung analysiert werden. Dabei ist zu untersuchen, ob Erprobungen bezüglich einzelner Merkmale Bias (vgl. Definition 46) gegenüber der Grundgesamtheit aufweisen.

Dazu werden im Folgenden exemplarisch drei Erprobungsfahrten auf ihre Repräsentativität gegenüber der Grundgesamtheit - in diesem Fall die Menge aller verfügbaren Erprobungsfahrten - untersucht. Zur Untersuchung der Repräsentativität werden die statischen und dynamischen Umgebungsinformationen sowie der Zustand des Ego-Fahrzeugs in der jeweiligen Umgebungssituation einzeln oder in Kombination betrachtet. Für die Voruntersuchung werden exemplarisch die Merkmale *Geschwindigkeit des Ego-Fahrzeugs*, *Abstand zum Vorderfahrzeug*, *geltende Tempolimits* und *Straßenklasse* betrachtet. Die drei ausgewählten Erprobungsfahrten sind in Tabelle 5.1 charakterisiert.

5.1.2.1 Verteilung der Ego-Geschwindigkeit

Abbildung 5.3 zeigt die Verteilung der Geschwindigkeit des Testfahrzeugs im gesamten Erprobungsdatensatz sowie in den drei Erprobungsfahrten. Die Geschwindigkeit des Ego-Fahrzeugs ist in 20km/h Abschnitte (Bins) eingeteilt und über die gefahrene Distanz aggregiert. Geschwindigkeiten über 280km/h sind zusammengefasst. Die Geschwindigkeiten im gesamten Erprobungsdatensatz (vgl.

Abbildung 5.3a) häufen sich im Bereich von 40km/h bis 140km/h. Mit Geschwindigkeiten darüber und darunter wurden nur wenige Kilometer zurückgelegt.

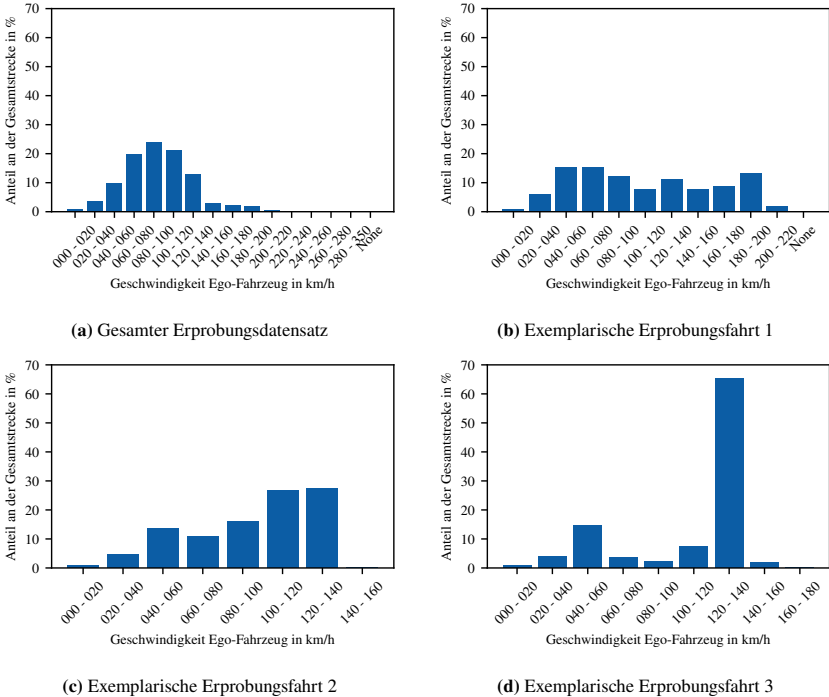


Abbildung 5.3: Verteilung der Geschwindigkeit in km/h für den gesamten Datensatz sowie drei exemplarische Erprobungsfahrten

Die Verteilung der Geschwindigkeiten für die exemplarische Erprobungsfahrt 1 (vgl. Abbildung 5.3b) ist breit gestreut. Geschwindigkeiten zwischen 20km/h und 200km/h kommen häufig und gleichmäßig verteilt vor. Geschwindigkeiten über 220km/h sind nicht enthalten.

In der exemplarischen Erprobungsfahrt 2 sind vorwiegend Geschwindigkeiten im Bereich von 40km/h bis 140km/h enthalten (vgl. Abbildung 5.3c). Der Häufungsbereich entspricht dem gesamten Datensatz allerdings unterscheidet sich die

Verteilung innerhalb der Häufung. Zudem sind Geschwindigkeiten ab 160km/h nicht in der Fahrt enthalten.

Die Geschwindigkeitsverteilung der exemplarischen Erprobungsfahrt 3 (vgl. Abbildung 5.3d) ist anomal. Es wurden fast ausschließlich Geschwindigkeiten im Bereich von 120km/h bis 140km/h gefahren.

Die drei exemplarischen Erprobungsfahrten weisen in den Geschwindigkeitsverteilungen eindeutige Unterschiede auf. Keine der drei Verteilungen entspricht der der Grundgesamtheit. Demzufolge ist keine der drei exemplarischen Erprobungsfahrten bezüglich der Ego-Geschwindigkeit repräsentativ für den gesamten Erprobungsdatensatz.

5.1.2.2 Verteilung des Abstands zum Vorderfahrzeug

Abbildung 5.4 zeigt die Verteilung des Abstands zum Vorderfahrzeug im gesamten Erprobungsdatensatz sowie den drei Erprobungsfahrten. Die maximale Detektionsreichweite der Kamera ist auf 250m begrenzt. Abschnitte, in denen kein Vorderfahrzeug erkannt wurde, sind separat ausgewiesen. Es wird jeweils nur das Vorderfahrzeug mit dem geringsten Abstand zum Ego-Fahrzeug betrachtet. Ein Fahrzeug im Nahbereich schließt in der Realität ein weiteres Fahrzeug davor nicht aus, jedoch ist aufgrund von Verdeckungen die Detektionsrate unzuverlässig. Der Fall wird daher hier nicht betrachtet.

Im gesamten Erprobungsdatensatz (vgl. Abbildung 5.4a) ist ungefähr zu 40% der Strecke kein vorausfahrendes Fahrzeug erkannt worden. Die anderen 60% verteilen sich verstärkt auf den Nahbereich zwischen 20m und 80m. Die Häufigkeit bis zu 250m nimmt stetig ab.

Die Häufigkeitsverteilung des Abstands zum Vorderfahrzeug für die exemplarische Erprobungsfahrt 1 (vgl. Abbildung 5.4b) ähnelt der Verteilung im gesamten Datensatz. Abschnitte ohne Vorderfahrzeug kommen geringfügig seltener vor, dafür häufen sich Fahrzeuge im Nahbereich zwischen 20m und 60m.

Ähnlich verhält sich die Häufigkeitsverteilung für die exemplarische Erprobungsfahrt 2 (vgl. Abbildung 5.4c). Die Häufung im Nahbereich konzentriert sich hier allerdings auf den Bereich von 20m bis 40m.

In der exemplarischen Erprobungsfahrt 3 (vgl. Abbildung 5.4d) sind keine Informationen zum Abstand des Vorderfahrzeugs enthalten. Dies kann zum Beispiel in fehlerhafter Messtechnik oder ausgefallener Sensorik begründet sein.

Auch bezüglich des Vorhandenseins und Abstands einen Vorderfahrzeugs ist keine allgemeingültige einheitliche Verteilung zu erkennen. Vielmehr variiert die Häufung im Nahbereich von Fahrt zu Fahrt. Allerdings ist stets mindestens ein Drittel der Fahrt ohne erkanntes Vorderfahrzeug. Der Fernbereich ist jeweils unterrepräsentiert, was jedoch im Abstand zum dichtesten Fahrzeug begründet sein kann. Zudem kann die Verkehrssituation einen Einfluss darauf haben: So sind bei zählfließendem oder stockendem Verkehr sowie im städtischen Bereich hohe Abstände unwahrscheinlicher.

5.1.2.3 Verteilung der geltenden Tempolimits

Abbildung 5.5 zeigt die Verteilung des jeweils aktuell geltenden Tempolimits in den drei Erprobungsfahrten. Abschnitte ohne Tempolimit sind auf der X-Achse mit 0 dargestellt.

Im gesamten Erprobungsdatensatz (vgl. Abbildung 5.5a) sind Tempolimits enthalten, die jeweils einen Anteil von weniger als 20% an der Gesamtstrecke besitzen. Die längste Streckengültigkeit besitzen die Tempolimits 100, unbegrenzt, 80, 70 und 120.

In der exemplarischen Erprobungsfahrt 1 (vgl. Abbildung 5.5b) waren über 40% der Strecke ohne Tempolimit. Für die restliche Strecke war das Tempolimit von 100 km/h am längsten vorhanden.

Für die Strecke der exemplarischen Erprobungsfahrt 2 (vgl. Abbildung 5.5c) gab es dauerhaft ein Tempolimit. Dabei waren die Beschränkungen auf 110 und 130 km/h mit ~20% bzw. ~35% am häufigsten vertreten.

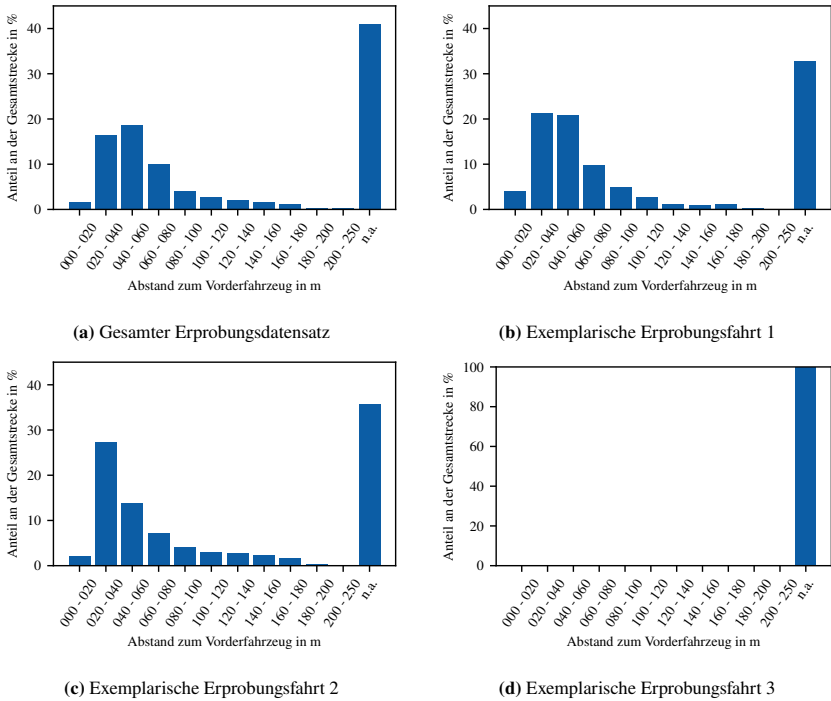


Abbildung 5.4: Verteilung der Abstände zum Vorderfahrzeug in m für den gesamten Datensatz sowie drei exemplarische Erprobungsfahrten

Die exemplarische Erprobungsfahrt 3 (vgl. Abbildung 5.5d) weist auch in der Tempolimitverteilung eine Unregelmäßigkeit auf. Mit einem Tempolimit von 140 für über 70% der Strecke weicht die Verteilung von allen anderen Verteilungen ab.

5.1.2.4 Verteilung der Straßenklasse

Abbildung 5.6 zeigt die Verteilung der Straßenklasse, auf der sich das Testfahrzeug zum Zeitpunkt der Aufnahme befunden hat, in den drei Erprobungsfahrten. Die Straßenklasse wird aus den elektronischen Kartendaten entnommen.

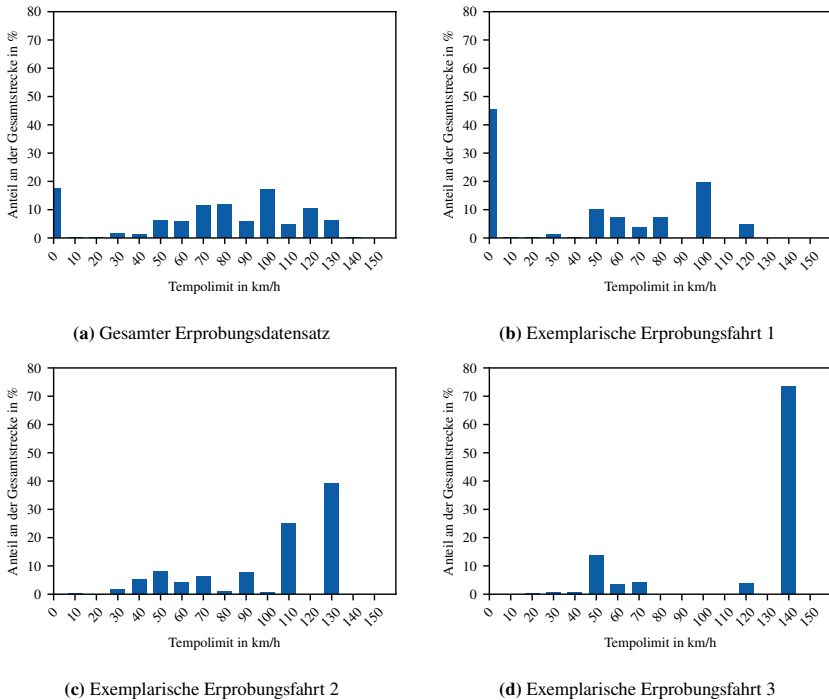


Abbildung 5.5: Verteilung der geltenden Tempolimits in km/h für den gesamten Datensatz sowie drei exemplarische Erprobungsfahrten

Im gesamten Erprobungsdatensatz (vgl. Abbildung 5.6a) sind ca. 50% Autobahn und 25% Bundesstraße enthalten. Den drittgrößten Posten belegt die Landstraße mit rund 15%. Kreis- und Ortsstraßen haben einen geringen Anteil.

In der exemplarischen Erprobungsfahrt 1 (vgl. Abbildung 5.6b) ähnelt die Verteilung der Grundgesamtheit. Kreisstraßen sind nicht enthalten, dafür nimmt der Anteil an Bundes- und Ortsstraßen geringfügig zu.

Die Verteilung in der exemplarischen Erprobungsfahrt 2 (vgl. Abbildung 5.6c) verschiebt sich in Richtung Autobahn und Bundesstraße. Alle anderen Straßenklassen sind entweder unterrepräsentiert oder nicht vorhanden.

Der Autobahnanteil nimmt in der exemplarischen Erprobungsfahrt 3 (vgl. Abbildung 5.6d) nochmal zu und liegt bei ~75%. Die Bundesstraße ist im Vergleich zu den anderen Fahrten stark unterrepräsentiert. Die Verteilung der anderen Straßenklassen entspricht in etwa der Grundgesamtheit.

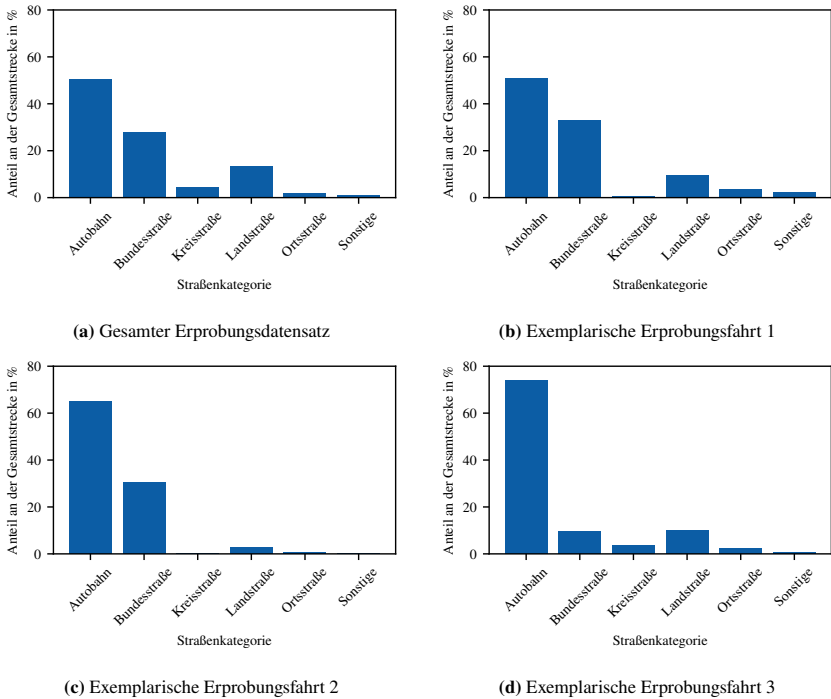


Abbildung 5.6: Verteilung der Straßenklasse für den gesamten Datensatz sowie drei exemplarische Erprobungsfahrten

5.1.2.5 Fazit

Obwohl in der Untersuchung nur wenige Merkmale betrachtet und diese in den Beispielen einzeln und rein empirisch untersucht wurden, sind deutliche Abweichungen zwischen der Grundgesamtheit und einzelnen bzw. allen Erprobungsfahrten zu erkennen. Dies deutet darauf hin, dass von der Repräsentativität einer einzelnen Applikations- oder Erprobungsfahrt nicht ausgegangen werden kann.

5.1.3 Prozess- und strukturbedingte Bias in Erprobungsfahrten

Die Grundlage für die Applikation bilden Straßenverkehrstests, die sich in Testfahrten (vgl. Definition 29) und Erprobungen (vgl. Definition 30) unterscheiden lassen. Die entwicklungsbegleitenden Testfahrten beschränken sich aufgrund ihrer kurzen Dauer regional rund um die Entwicklungsstandorte der Automobilhersteller. Erprobungen besitzen einen größeren Testumfang als die Testfahrten, werden vorab geplant und haben im Bereich der Fahrerassistenzsysteme jeweils einen Fokus auf bestimmte Regionen und Länder, um regionale Besonderheiten in der Wahrnehmung sowie den elektronischen Streckendaten zu erfassen. Die Planung dieser Erprobungen erfolgt in einem eng getakteten Entwicklungsprozess unter hohem Zeitdruck entweder durch die Entwickler und Funktionsverantwortlichen selbst oder durch eine zentrale Koordinationsstelle. Dabei müssen eine Reihe von Randbedingungen berücksichtigt werden, wie z.B. die Verfügbarkeit von Versuchsträgern, die Arbeits- und Urlaubspläne der Entwickler und Applikateure, verfügbare zeitliche und monetäre Budgets und konkurrierende Testziele. Vorhandene Einflussfaktoren³ erzeugen Bias in Erprobungsfahrten (vgl. Tabelle 5.2).

Unter diesen Randbedingungen muss der Testumfang maximiert werden. Bei Erprobungsfahrten liegt der Fokus auf einer hohen Effizienz. In kürzester Zeit

³ Veröffentlicht in *Statistical Consideration of the Representativeness of Open Road Tests for the Validation of Automated Driving Systems* [JL12]

Kategorie	Einflussfaktoren
Beteiligte Personen	Individuelle Fahrercharakteristik Qualifikation (z.B. Fahrsicherheitstraining) Erfahrung und Routine Nutzung von ADAS und HAF-Systemen Experten- bzw. Systemwissen
Routenplanung	Ladesäulen- und Tankstellennetz Nähe zu Entwicklungs- und Unternehmensstandorten Organisatorische Rahmenbedingungen Persönliche Präferenzen Ortskenntnisse
Fahrzeugumgebung	Jahres- und Tageszeit Straßentopographie und -topologie Verkehrsaufkommen Wetter- und Klimaeffekte
Eingesetzte Erprobungsfahrzeuge	Derivatunterschiede Verbaute Komponenten, Funktionen und Systeme Prototypenstatus und Reifegrad
Länder, Regionen und Absatzmärkte	Geltende Verkehrsregeln Länderspezifische Fahrstile Örtliches Verkehrsaufkommen Form und Farbe von Verkehrsschildern Baustile und Bebauung
Projektorganisation	Verfügbare Ressourcen (Zeit, Zeitslots, Anzahl der Fahrzeuge & Applikateure, . . .) Priorisierung der Absatzmärkte Prozessuale Vorgaben

Tabelle 5.2: Einflussfaktoren nach Kategorien aufgeteilt, die Bias in Erprobungsfahrten erzeugen.

sollen möglichst viele Testkilometer in einem oder mehreren Ländern entlang der geplanten Route eingefahren werden. Das systematische Vorgehen und wiederholte Testen in gleichen Situationen (vgl. Definition 26) mit unterschiedlichen Parametern rückt dabei in den Hintergrund. Vielmehr werden Parameter über die Erprobung hinweg verändert. Der sich ändernde wahrnehmbare Fahreindruck kann dabei jedoch auch von der veränderten Situation beeinflusst sein. Ein direkter Rückschluss auf die geänderten Parameter ist nicht sicher gewährleistet.

Definition 49: Fahrtinhalt

Der Fahrtinhalt umfasst alle Situationen (vgl. Definition 26) eines Straßenverkehrstests und beinhaltet insbesondere alle relevanten Umgebungsbedingungen, bspw. die Straßenklassen oder das Höhenprofil. Die relevanten Merkmale bzw. Umgebungsbedingungen einer Situation sind abhängig von der betrachteten Funktion.

Die Durchführung der Erprobungsfahrten unter den aufgeführten Randbedingungen garantiert weder, dass der Fahrtinhalt den von den späteren Kunden gefahrenen Strecken entspricht, noch, dass der Fahrtinhalt einer Auswahl nach dem Zufallsprinzip unterliegt.

Abbildung 5.7 vergleicht die gefahrene Distanz in einem exemplarischen Erprobungsdatensatz innerhalb Deutschlands mit der Länge des deutschen Straßennetzes. In Deutschland sind lediglich 13.155 km der insgesamt 229.601 km überörtlicher Straßen Autobahnen. Im Vergleich dazu ist der relative Anteil im exemplarischen Erprobungsdatensatz mit ~60% wesentlich höher. Die relativen Anteile der Land- und Kreisstraßen sind im Erprobungsdatensatz entsprechend geringer.

Abbildung 5.8 vergleicht die gefahrene Distanz im exemplarischen Erprobungsdatensatz innerhalb Deutschlands mit der Fahrleistung Deutschlands im Jahr 2005⁴.

⁴ Die Straßenverkehrszählungen ab 2010 enthalten nur noch Daten für die Autobahnen und Bundesstraßen.

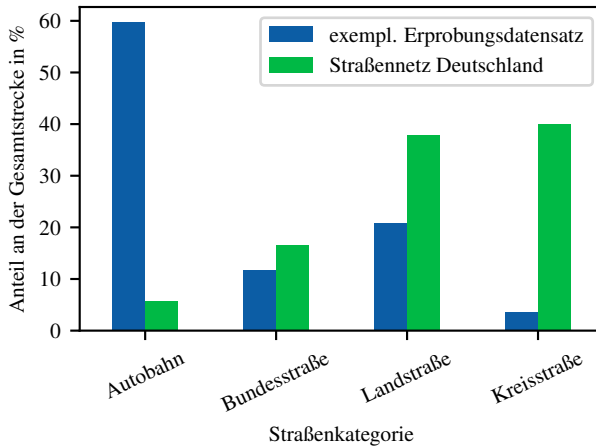


Abbildung 5.7: Vergleich der Verteilung der Straßenkategorien im exemplarischen Erprobungsdatensatz mit der Verteilung im gesamten deutschen Straßennetz als Grundgesamtheit (Quelle für das Straßennetz Deutschlands: BMDV [120])

Während die beiden Verteilungen ähnlicher sind als der Vergleich mit dem gesamten Straßennetz, so gibt es doch Abweichungen. Die Fahrleistung auf Autobahnen liegt 10,9% unter der gefahrenen Distanz im exemplarischen Erprobungsdatensatz. Hingegen ist die Fahrleistung auf Bundesstraßen 13% und auf Kreisstraßen 4,4% höher als die gefahrene Distanz im exemplarischen Erprobungsdatensatz.

Ob die Fahrleistung eines Landes, die vom Fahrzeughersteller selbst erhobene Fahrleistung der Kundenflotte oder das gesamte Straßenverkehrsnetz als Referenz für die Grundgesamtheit geeignet sind, hängt von der Fahrfunktion ab und muss im Einzelfall entschieden werden. In allen Fällen muss darauf geachtet werden, dass die Erprobungsfahrten keine Bias enthalten.

Prozess- und strukturbedingte Bias führen zu einer systematischen Verzerrung in den Daten und müssen in der Auswertung von Erprobungen und allen darauf basierenden Analysen zwingend berücksichtigt werden. Ein Ignorieren dieser Bias verfälscht das Ergebnis und damit die Validität der Datenanalyse.

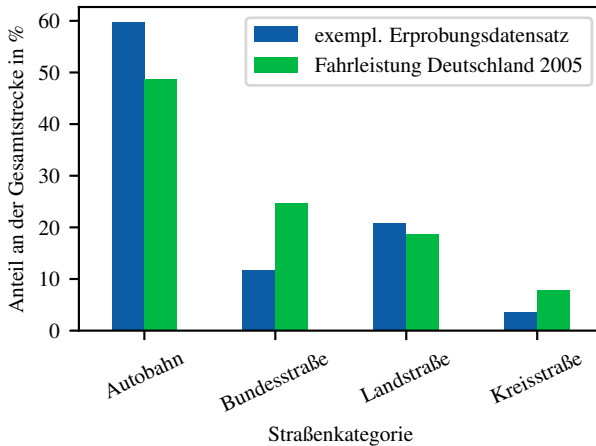


Abbildung 5.8: Vergleich der Verteilung der gefahrenen Distanz im exemplarischen Erprobungsdatensatz mit der Fahrleistung in Deutschland 2005 als Grundgesamtheit (Quelle für die Fahrleistung Deutschlands 2005: ifeu / bast [121, 122])

Abbildung 5.9 zeigt die gefahrene Distanz in einem Erprobungsdatensatz aufgeschlüsselt für das Merkmal *Straßenkategorie*. Bei Betrachtung der Verteilung über alle Länder (Teilgrafik 5.9a) fällt eine Überrepräsentation von Autobahn- und Bundesstraßen-Abschnitten auf. Zwischen einzelnen Ländern zeigen sich zudem Unterschiede in der Verteilung der Straßenklasse. Während in Deutschland (Teilgrafik 5.9b) die Autobahn überrepräsentiert ist, ist es in Norwegen (Teilgrafik 5.9d) die Bundesstraße. Die Verteilung in Schweden (Teilgrafik 5.9c) weist eine verhältnismäßig ausgeglichene Verteilung auf.

Die Verteilung der zurückgelegten Erprobungskilometer über die Straßenklasse und die verschiedenen Länder zeigt, dass es eindeutige Bias in den Erprobungsdatensätzen gibt. Zudem gibt es erhebliche Unterschiede im Umfang der Test- und Erprobungsfahrten in den einzelnen Ländern (vgl. Abbildung 5.10). Diese Unterschiede entsprechen weder der Länge des Straßennetzes noch den von den Kunden gefahrenen Kilometern in den einzelnen Ländern. Die Anzahl an

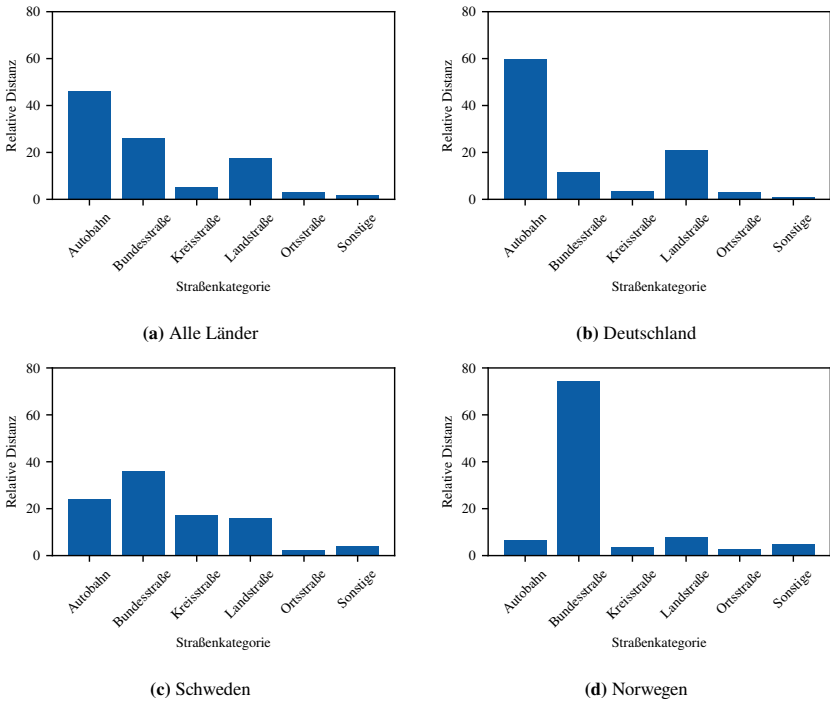


Abbildung 5.9: Verteilung der gefahrenen Distanz nach Straßenkategorie über verschiedene Länder in einem exemplarischen Erprobungsdatensatz

Test- und Erprobungskilometern in Deutschland ist durch entwicklungsbegleitende Testfahrten um den Entwicklungsstandort sowie den Entwicklungsstandort als Ausgangspunkt für Erprobungsfahrten zu begründen.

5.1.4 Fehlende Systematik in der Parametervariation

Die Festlegung der Parameter beginnt mit einem „best guess“ der Entwickler und Applikateure für jeden Parameter. Während der Entwicklung werden für viele Parameter bereits Standard- bzw. Initialwerte definiert. Gegenseitige Abhängigkeiten von Parametern führen dazu, dass Änderungen an einzelnen Parametern neben der

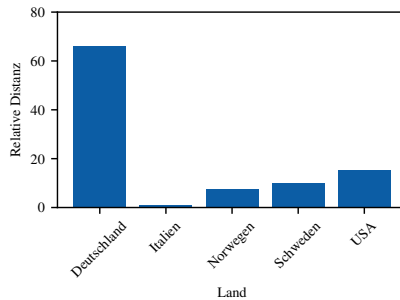


Abbildung 5.10: Verteilung der gefahrenen Distanz auf ausgewählte Länder in einem exemplarischen Erprobungsdatensatz

gewünschten Veränderung des Funktionsverhaltens in einigen Situationen auch nicht vorhersehbare bzw. unerwünschte Veränderungen in der Funktionsweise mit sich führen (vgl. Abschnitt 3.3). Dies fördert eine Verstetigung der Initialwerte der Parameter, die dann während der Applikation nicht mehr systematisch verändert werden. Stattdessen werden Parameter Auffälligkeits-basiert appliziert. D.h. ein Parameter wird nur dann verändert, wenn seine Ausprägung zu einer negativen Auffälligkeit während der Entwicklung oder Applikation führt.

Die Durchführung der Applikation beginnt mit dem aus der Entwicklung hervorgegangenen Parametersatz. Zum einen sind Auffälligkeiten in der Funktionsweise Ursache für Parameteränderungen durch den Applikateur. Dabei wird von der Auffälligkeit auf die die Auffälligkeit verursachende Teilfunktion und die daran beteiligten Parameter geschlossen und diese entsprechend angepasst. Zum anderen werden Parameteränderungen durchgeführt, wenn der Applikateur einen gesamtheitlichen Fahreindruck erhalten hat und mit der Funktionsqualität nicht zufrieden ist. Dabei werden einzelne Parameter verändert, um das Funktionsverhalten an die Erwartungen anzupassen.

In der Gesamtheit entsteht so ein unstrukturiertes und nicht systematisches Vorgehen. Das Funktionsverhalten vor und nach den sequentiellen Parameteränderungen ist nicht direkt vergleichbar. Insbesondere über viele konsekutive Änderungen

nimmt die Vergleichbarkeit drastisch ab. Das Festsetzen der neuen Parameterwerte unterliegt einem „subjektiven Bauchgefühl“.

5.2 Related Work zu den Herausforderungen in der Applikation von ADAS und HAF

5.2.1 Simulative Applikation von ADAS und HAF-Funktionen

Die Modellbildung ermöglicht es, Fahrfunktionen simulativ auszuführen. Die Qualität der Modelle für Fahrzeug, Fahrzeugumgebung und Fahrfunktion beeinflussen den Realitätsgrads und damit die Aussagekraft der Simulation. Die stetige Verbesserung der Modelle wird angestrebt, um die Skalierbarkeit von Simulationen für das Testen von ADAS und HAF-Funktionen (vgl. XiL-Testen in Abschnitt 3.2) nutzbar zu machen sowie zur Verbesserung der Applikation zu führen. Dabei werden vordefinierte Szenarien mit verschiedenen Parameterkombinationen simuliert und der entstandene Fahrindruck mittels Objektivierung verglichen.

In der Fahrerassistenzdomäne gibt es Ansätze, die Applikation simulativ durchzuführen. Aufgrund der Interaktion mit der Systemumgebung muss dabei das gesamte Fahrzeug samt Fahrzeugumgebung modelliert werden. Da die Varianz und der Detailgrad der Modelle durch die manuellen Modellierungsaufwände limitiert sind und die Applikation eine sehr hohe Anforderung an den Realitätsgrad des Gesamtsystems besitzt, wird die Systemumgebung durch Replay aufgezeichneter Realdaten (z.B. [108, 109]) dargestellt. Durch wiederholte Simulationsdurchläufe und Veränderungen einzelner Parameter kann eine Applikation durchgeführt werden.

5.2.1.1 Simulative Applikation einer modellprädiktiven Regelung

Oschlies et al. [123] verwenden zur Applikation einer modellprädiktiven Querführung eine Simulationsumgebung, um die Funktion nicht in der realen Welt parametrieren zu müssen. Durch die Verwendung einer Simulationsumgebung mit Laufzeiten von rund 0.8 Sekunden für einen Streckenabschnitt von 2.5 km ermöglichen sie den Einsatz von Partikelschwarmalgorithmen, die auf Basis einer zu definierenden Anzahl von Agenten den Lösungsraum iterativ nach der besten Lösung absuchen. Das Verfahren ist insbesondere für eine modellprädiktive Regelung (eng. Model Predictive Control (MPC)) geeignet, da diese auf einer Reihe von Stell- und Störgrößen basiert, die zur Optimierung der Regelgröße appliziert werden müssen.

Die Zielfunktion für die Optimierung des MPC leiten Oschlies et al. aus einer Probandenstudie ab. Die subjektiven Bewertungen des Fahreindrucks werden über Objektivnotenmodelle auf objektive Messgrößen abgebildet [124].

5.2.1.2 Simulative Applikation eines ACC-Systems

Pawellek et al. [125] zeigen die Optimierung eines ACC Systems exemplarisch an einer Feldstudie. Sie verwenden die in einer Vorstudie abgeleiteten Objektivnotenmodelle für die Applikation des Systems und validieren das optimierte ACC System in einer Nachstudie. Sie erzielen dabei über alle betrachteten Fälle hinweg eine Verbesserung des subjektiven Systemverhaltens. Pawellek et al. differenzieren bei der Bewertung des Systems nach Fahrmanövern (z.B. Einscheren, Einscherer, Auffahren, etc.) und Fahrumgebung (Stadt, Land, Bundesautobahn). Zusätzlich bewerten sie das System im Normal- und Sportmodus.

5.2.1.3 Simulative Applikation eines SAE Level 3 Systems

Fraikin [35] hat eine virtuelle Applikationsmethodik für automatisierende Fahrfunktionen entwickelt. Der Schwerpunkt liegt dabei auf der Einschränkungen des

Suchraums durch die Identifikation einflussreicher Parameter und Wertebereiche. Ein weiterer Fokus lag auf der Wahl des geeigneten Optimierungsalgorithmus. Für die virtuelle Applikation wurde ein hybrides Fahrzeugmodell basierend auf einem analytischen Einspurmodell sowie künstlicher Intelligenz entwickelt mit dem Fokus auf Effizienz und Rechenzeitoptimierung. Auf Basis dieses Modells wurden die relevanten Bereiche im Suchraum virtuell eingeschränkt, um die Effizienz der Applikation im realen Straßenverkehr zu verbessern. Dabei wurde eine Methodik zur Komplexitätsreduktion des Parameterraums basierend auf der Elementareffektmethode eingesetzt. [35, 126]

Die Methodik wurde prototypisch für ein SAE Level 3 System umgesetzt. Dabei wurde die Optimierung auf Basis von sechs vordefinierten Szenarien auf der Autobahn durchgeführt [35, 127].

5.2.2 Nutzerzentrierte Applikation von ADAS

Bei nutzerzentrierten Ansätzen steht die subjektive Bewertung des Systems durch den späteren Nutzer im Vordergrund. Um die subjektive Bewertung der Nutzer zu quantifizieren, werden Probandenstudien, Experimente oder Feedbackmechanismen eingesetzt. Im Fokus der Wissenschaft steht die Objektivierung des subjektiven Fahreindrucks, um die Skalierbarkeit und Übertragbarkeit der Studien und Experimente zu gewährleisten:

Schick et al. entwickelten ein generisches Vorgehensmodell, um subjektive Eindrücke von späteren Nutzern sowie Experten in objektive KPI für vordefinierte Fahrmanöver zu überführen. Das generische und damit auf andere Fahrmanöver übertragbare Vorgehen wurde am Beispiel einer assistierten Querführung untersucht. Zuerst wurden mittels Expertenworkshops, Benchmark-Tests und Messkampagnen die relevanten Attribute für die Bewertung der Querführung ermittelt und über ein Ebenenmodell verknüpft. Für die Nutzer erlebbare und relevante Fahrzeug- bzw. Fahrfunktionseigenschaften werden auf objektiv messbare Fahrzeugsignale abgebildet. [94, 128]

Bruder et al. [129] untersuchen die allgemeine Aussagekraft menschbezogener Messgrößen in Fahrversuchen. Sie schlussfolgern, dass zur Bewertung von ADAS subjektive Bewertungen, die z.B. durch Fragebögen erfasst werden, und erfasste Fahrzeugdaten geeignet sind. Zu den relevanten Signalgrößen zur Quantifizierung der Fahrzeuglängs- und querverführung zählen laut Bruder et al. die Motordrehzahl, das Motormoment, die Fahrzeuggeschwindigkeit, die Fahrzeugquerbeschleunigung, die Fahrzeuglängsbeschleunigung, die Gierrate, der Lenkradwinkel, die Fahrpedalstellung, der Bremsdruck, der Abstand zum vorausfahrenden Fahrzeug sowie die Abstände zur linken und rechten Begrenzungslinie.

Bauer et al. [34] verwenden den Verlauf des Längsrucks für die Komfortbewertung der Trajektorie. Durch die Quadrierung des Längsrucks werden schnelle, unkomfortable Änderungen der Beschleunigung bestraft.

Holzinger et al. [130] haben eine Methode zur Objektivierung subjektiver Fahreindrücke für ACC und LKA Systeme entwickelt. Sie bewerten dabei nicht nur den Fahrkomfort, sondern auch die empfundene Sicherheit, reale Sicherheit und Fahrdynamik. Aus Referenzfahrten von Fahrverhaltensexperten sowie Fahrzeugkunden unterschiedlichen Alters und mit unterschiedlicher Fahrerfahrung wurden physikalische Parameter zur Beschreibung des Fahrzeugzustands in den bewerteten Situationen extrahiert. Für die Längsdynamik wurden Betriebs- und Fahrverhalten von Motor und Getriebe; für die Querdynamik das Handling, Lenkung und Fahrwerk; und für die Vertikaldynamik der Fahrwerkskomfort ermittelt. Die Studien wurden sowohl aus Fahrer- als auch aus Beifahrersicht durchgeführt. Die abgeleiteten physikalischen Größen werden dann zur Identifikation von Betriebszuständen und Mustern verwendet, die eine Echtzeitbewertung des Systems mit Objektivschulnoten ermöglichen.

Die physikalischen Modelle werden über Korrelations- und Regressionsanalysen [125, 130] aus den bewerteten Messdaten abgeleitet. Holzinger et al. ermitteln für die zu testenden Systeme (ACC und LKA) Haupteinflussgrößen und untersuchen weitere Zusatzgrößen, die einen Einfluss auf die subjektive Bewertung haben könnten. Sie zeigen die Komplexität der Objektivierung auf, die sich aus der Anzahl von Einflussgrößen und den unklaren Abhängigkeiten und Regeln

zwischen den Größen ergibt. So wird beispielsweise das Auffahren auf ein langsames Vorderfahrzeug sensibler bewertet als ein Einscherer, der objektiv aus Systemsicht kritischer zu bewerten ist. Zur Reduktion der Komplexität betrachten Holzinger et al. die Fahraufgaben der Funktion einzeln. Dazu zählen zum Beispiel die Konstantfahrt, Beschleunigung, Verzögerung, Anhalten, Anfahren sowie Ein- und Ausscherer.

Auch Pawellek et al. [125] nutzen Objektivnotenmodelle, um Zielgrößen für eine ACC-Systemauslegung zu quantifizieren. Sie ermitteln beispielsweise die idealen Reaktionszeitpunkte für das ACC im Sport- und Normalmodus. Durch eine Probandenstudie aus Experten und Normalfahrern auf einem Prüfgelände werden Subjektivnoten ermittelt. Es wurden verschiedene Fahrmanöver, darunter Einscheren, Einscherer und das Auffahren auf ein Vorderfahrzeug, in verschiedenen Fahrumgebungen, Stadt, Landstraße und Bundesautobahn, untersucht. Für die unterschiedlichen Ausprägungen der Subjektivnoten wurden physikalische Kennparameter ermittelt, die dann zur Herleitung der Objektivnotenmodelle verwendet wurden. Durch einen Nachversuch wurden die erstellten Modelle validiert.

5.2.3 Applikation mittels datenbasierter und lernender Verfahren

Petersen et al. [JL15] nutzen aufgezeichnete Fahrzeugtests zur Validierung und Bewertung einer Reichweitenmanagementfunktion. Durch eine open-loop Re-Simulation werden zudem verschiedene Parametrierungen getestet und verglichen. Die nachgelagerte, objektive Analyse der Fahrzeugtests bietet darüber hinaus das Potenzial, lernende Verfahren einzusetzen. Denkbar ist beispielsweise das Ableiten von Parameterverbesserungen durch das Aufzeigen von in den bewerteten Fahrzeugmessdaten vorhandenen Korrelationen.

Auch der Ansatz von Oschlies et al. [123] verwendet aufgezeichnete Fahrzeugtests für die Applikation. Für die Applikation des MPC wurde ein Streckenabschnitt auf der Bundesautobahn A39 verwendet. Die Übertragbarkeit auf andere

Streckenabschnitte, Autobahnen oder gar Landstraßen war nicht Gegenstand der Untersuchung.

5.2.4 Applikation durch Vehicle in the Loop Verfahren

In der Forschung gibt es ViL-Ansätze, die das reale Fahrzeug auf dem Prüfgelände einsetzen und so reale Fahrdynamiken erzeugen [131]. Dabei werden nur Teile der Umgebung, wie beispielsweise andere Verkehrsteilnehmer virtuell eingespielt, während die Strecke und das Fahrzeug real vorhanden sind. Auf dem Prüfgelände in Kombination mit dem Einspielen von anderen Verkehrsteilnehmern kann eine Vielzahl von Szenarien reproduzierbar nachgestellt werden. Durch den Ansatz kann auch die Applikation von ADAS und HAF-Funktionen unter realistischen Bedingungen durchgeführt werden.

5.3 Diskussion und Potenziale

Neben der manuellen Applikation ist eine Reihe weiterer Ansätze entstanden, die die Applikation von ADAS und HAF-Systemen verbessern. Die Ansätze müssen dabei nicht alleinstehend angewandt sondern können miteinander kombiniert werden. So ist beispielsweise die datenbasierte Applikation unter Verwendung einer objektiven Bewertung der im Fahrzeug subjektiv wahrgenommenen Qualität des ADAS bzw. der HAF-Funktion denkbar. Die Objektivierung bietet anschließend eine Vergleichbarkeit von Fahrzeugtests mit unterschiedlichen Softwareständen oder Realisierungen von Funktionskonzepten. Zur Ableitung der objektiven Bewertungsmetriken können Ansätze aus der nutzerzentrierten Applikation eingesetzt werden.

Die aktuellen Forschungsansätze (vgl. Abschnitt 5.2.1) fokussieren sich auf die Virtualisierung der Applikation durch den Einsatz von Simulationen und die Verbesserung der Applikation durch Einfangen des subjektiven Fahreindrucks,

insbesondere auch durch spätere Nutzer des ADAS bzw. HAF-Systems. Die simulativen Ansätze ermöglichen eine umfassendere Exploration des Parameterraums in kurzer Zeit und schaffen damit das Potenzial einer optimalen Applikation.

Während die simulativen Ansätze ein großes Potenzial für die Applikation bieten und eine objektive Vergleichbarkeit von Parameteränderungen gewährleisten, sind sie auf einige wenige, vordefinierte Szenarien beschränkt. Die Varianz an Szenarien, die es im realen Straßenverkehr gibt und die für ADAS und HAF-Funktionen entscheidend sind, wird dabei nicht ausreichend abgedeckt. In einigen Ansätzen werden zwar kleine Abschnitte realer Strecken oder einzelne extrahierte Szenarien verwendet, allerdings wird das Potenzial datenbasierter Methoden, das sich durch die Verwendung großer Datenmengen ergibt, noch nicht ausgeschöpft.

Allerdings besteht bei der Verwendung aufgezeichneter Straßenverkehrstests die Gefahr, dass die Parameter auf den verwendeten Datensatz optimiert werden (eng. *overfitting*). Dies führt bei der Verwendung einzelner, nicht repräsentativer Streckenabschnitte oder einem unvollständigen Satz an Szenarien oder Manövern zwar für die jeweilige Auswahl zu einer optimalen Applikation, das Fahrerlebnis kann in der Realität jedoch abweichen, da die Übertragbarkeit auf Strecken und Szenarien, die nicht im Datensatz enthalten waren, nicht gewährleistet ist.

Genau diese Abhängigkeit der für die Applikation verwendeten Strecken oder Szenarien wird in den Ansätzen bisher nicht betrachtet. Nur wenn auch die Fahrzeugumgebung vollständig und hinreichend realistisch für alle relevanten Situationen berücksichtigt wird, kann eine optimale Applikation gefunden werden.

Insbesondere in der datenbasierten Applikation stecken ungenutzte Potenziale, da die während der Entwicklung und Applikation anfallenden Fahrzeugmessdaten noch nicht bzw. nur in kleinen Ausschnitten in den Ansätzen berücksichtigt werden. Für den Einsatz der Fahrzeugmessdaten aus den Test- und Erprobungsfahrten für eine optimale Applikation müssen allerdings die in den Abschnitten 5.1.1, 5.1.2, 5.1.3 und 5.1.4 aufgezeigten Herausforderungen adressiert werden.

5.4 Anforderungen an die Applikation von ADAS und HAF-Funktionen

Ziel der Applikation ist es, das Verhalten der ADAS und HAF-Funktionen im realen Straßenverkehr einzustellen. Dabei ist nicht relevant, wie die Applikation durchgeführt wird, solange die Ergebnisse auf den realen Straßenverkehr übertragbar und dort valide sind. Der gesamte Straßenverkehr kann dabei auf die ODD des Systems eingeschränkt werden.

Anforderung 1: Die Applikation muss im realen Straßenverkehr innerhalb der vorgesehenen ODD durchgeführt oder direkt darauf übertragbar sein.

Ziel der Applikation im Bereich der ADAS und HAF-Systeme ist es, dem Kunden ein optimales Fahrerlebnis innerhalb der sicheren Systemgrenzen zu bieten. Dazu gehört zum einen eine hohe Verfügbarkeit des Systems und zum anderen eine gute Qualität der Funktion auf allen Märkten und für alle Modelle und Derivate, in denen die zu applizierende Funktion zum Einsatz kommt.

Anforderung 2: Innerhalb der ODD muss die Verfügbarkeit und die Qualität der Fahrfunktion maximiert werden.

Ein (globales) Optimum ist ein im statistischen Mittel für alle Situationen bestes Ergebnis. Dabei ist die Gesamtheit der Situationen zu definieren. Diese sind gleichverteilt oder zum Beispiel nach ihrer Häufigkeit in der realen Welt gewichtet. Bei der Applikation - im Gegensatz zur Validierung und Verifikation - spielt die Häufigkeit der einzelnen Situationen eine bedeutende Rolle. Ein suboptimales Funktionsverhalten in häufig auftretenden Fahrsituationen wirkt sich entsprechend wesentlich öfter und damit negativer auf das empfundene Fahrerlebnis aus als ein suboptimales Funktionsverhalten in einer sehr selten auftretenden Situation.

Anforderung 3: Bei der Applikation müssen reale Häufigkeiten von Fahrsituationen berücksichtigt und korrekt abgebildet werden.

Regionale und länderspezifische Besonderheiten durch beispielsweise unterschiedliche Gesetzgebungen, Klima- und Wetterbedingungen oder Straßenbeschaffenheiten müssen bei der Applikation berücksichtigt werden. Während für einzelne Länder und Regionen eigene Datensätze erstellt werden können, soll das Funktions- und Fahrzeugverhalten über alle Länder ähnlich bzw. gleich sein. Insbesondere für Funktionen, die von einer weitreichenden Umgebungswahrnehmung abhängig sind, ist die weltweit variierende Systemumgebung mit in die Applikation einzubeziehen. Neben der rein physischen Umgebungswahrnehmung spielen auch elektronische Karten und deren variierende Datenqualität in einzelnen Regionen eine bedeutende Rolle. Eine global-optimale Applikation muss daher Situationen und deren Häufigkeiten aus allen relevanten Ländern und Regionen berücksichtigen.

Anforderung 4: Die Applikation muss länderspezifische Variationen der Systemumgebung sowie der Datenqualität berücksichtigen.

Anforderung 5: Während die Applikation länderspezifisch sein kann, soll die Funktionsqualität in Bezug auf Fahrdynamik und Fahrkomfort über alle Länder hinweg annähernd identisch sein.

Da bereits der Vergleich zweier aufeinanderfolgender Applikationsfahrten einer hohen Ungenauigkeit und Subjektivität unterliegt (vgl. Abschnitt 5.1.4), ist ein Vorgehen, das auf der Durchführung realer Tests in allen relevanten Märkten durch den bzw. die Applikateure und Funktionsentwickler basiert, nicht sinnvoll. Neben den Kosten für den Transport der Fahrzeuge und Personen sowie dem notwendigen zeitlichen Aufwand, ist der Mehrwert durch den zeitlichen Versatz der einzelnen Applikationsfahrten als gering zu bewerten. Ein subjektiver Vergleich der durchgeführten Applikationsfahrten mit langem zeitlichen Versatz ist durch die Applikateure nicht zu leisten.

Anforderung 6: Eine Vergleichbarkeit der durchgeführten Straßenverkehrstests insbesondere zum Vergleich von Parametervariationen muss sichergestellt werden.

Hinzu kommt, dass die Auswirkungen von Parameteränderungen auf andere Regionen und Länder nicht bewertet werden können. Die Auswirkungen auf bereits getestete Regionen und Länder sind nicht abzusehen, die Auswirkungen auf noch zu testende Länder können erst später beobachtet werden, wobei dann ein Vergleich wiederum erschwert wird. Insgesamt sind mit einer Applikation, die darauf basiert, dass die Funktionsentwickler und Applikateure der jeweiligen Funktion individuelle Realwelttests durchführen, die erzielbaren Testumfänge stark limitiert.

Anforderung 7: Die Funktionsqualität des finalen Parametersatzes einer Applikation muss für alle relevanten Länder bekannt sein.

Die Abhängigkeit von dieser individuellen Bewertung und Applikation durch einige, wenige Applikateure im Fahrzeug muss gelöst werden, um die Testumfänge zu skalieren. Eine Entkopplung von Durchführung und Bewertung ist notwendig.

Anforderung 8: Die Applikation muss skalierbar sein.

Daraus ergibt sich folgendes Begriffsverständnis: Unter global-optimaler, datengetriebener Applikation ist in dieser Dissertation ein datengetriebener Ansatz für die Applikation einer Fahrfunktion zu verstehen, der auf Basis von weltweit dezentral aufgezeichneten Straßenverkehrstests eine möglichst gute Parametrierung unter Berücksichtigung der regions- und umgebungsabhängigen Variationen ermittelt. Dabei ist die Abdeckung und korrekte Abbildung der realen Häufigkeiten der Umgebungssituationen maßgeblich für die Bewertung der Applikation. Datengetrieben bedeutet, dass das Wissen und die Informationen objektiv aus den Daten extrahiert werden und die rein subjektive Bewertung in den Hintergrund rückt.

6 Global-optimale, datengetriebene Applikation von ADAS und HAF-Funktionen

6.1 Grundidee der global-optimalen, datengetriebenen Applikation

Zur Erreichung einer global-optimalen, datengetriebenen Applikation muss die Vergleichbarkeit unterschiedlicher Straßenverkehrstests hergestellt werden (vgl. Anforderung 6). Zudem muss die Gesamtheit der Straßenverkehrstests, die einer Applikation zugrunde liegen, auf ihre Repräsentativität untersucht werden, um den Anspruch des globalen Optimums sicherzustellen (vgl. Anforderung 3). Dafür müssen geeignete Verfahren (Abschnitt 6.3 und Abschnitt 6.4) entwickelt werden. Zudem muss ein strukturierter Prozess (Abschnitt 6.5) etabliert werden, um die notwendige Systematik (vgl. Voruntersuchung in Abschnitt 5.1.4) im Applikationsprozess sicherzustellen.

Die Durchführung der für die Applikation notwendigen Tests kann mit verschiedenen Testplattformen (vgl. Abschnitt 3.2) erfolgen. Die Applikation benötigt für die Bewertung des Fahrerlebnisses eine closed-loop Ausführung sowie einen hohen Realitätsgrad eingesetzter Modelle (Anforderung 1). Der benötigte Realitätsgrad hängt von der zu applizierenden Funktion ab und muss im Einzelfall quantifiziert werden. Zudem müssen Funktionsparameter variiert und die Umgebungsbedingungen abgebildet werden. Die Testumgebung muss skalierbar sein (Anforderung 8).

Software-in-the-Loop-Testmethoden bieten die benötigte Skalierbarkeit. Allerdings ist die Übertragbarkeit in die Realität von der Modellgüte abhängig. Zudem lässt sich die Variation an Umgebungsbedingungen mit den aktuellen Simulationstools nicht ausreichend abbilden. Straßenverkehrstests bieten den höchsten Realitätsgrad und damit die höchste Validität. Durch die Nutzung von Dauerläufern und Kundenflotten mit Shadow Mode (vgl. Abschnitt 3.2.3) kann die Skalierbarkeit von Straßenverkehrstests erhöht werden. Zudem bilden sie bei ausreichendem Umfang die notwendigen Variationen der Umgebungsbedingungen ab. Daher basiert der Ansatz für die global-optimale, datengetriebene Applikation im Kern auf der Analyse real durchgeführter Straßenverkehrstests.

Der Prozess zur global-optimalen, datengetriebenen Applikation (vgl. Abbildung 6.1) basiert auf dem bisherigen Applikationsprozess (vgl. Abbildung 3.5). Die Applikation beginnt mit der Defintion des Applikationsumfangs (Prozessschritt A.1). Die Phasen Vorbereitung (A) und Durchführung (B) werden um eine systematische Parametervariation (Prozessschritte A.2 & B.2) ergänzt, bleiben im großen Ganzen aber unverändert. Hinzu kommt die Phase der Datenaufzeichnung (C), in der die Fahrzeugsignale mit geeigneter Messtechnik aufgezeichnet und anschließend während der virtuellen Analyse (D) offline bewertet werden statt wie bisher eine subjektive Bewertung im Fahrzeug durchzuführen. Die global-optimale, datengetriebene Applikation beruht weiterhin auf der **Durchführung von Straßenverkehrstests bzw. einer Realdaten-basierten (Re-)Simulation**.

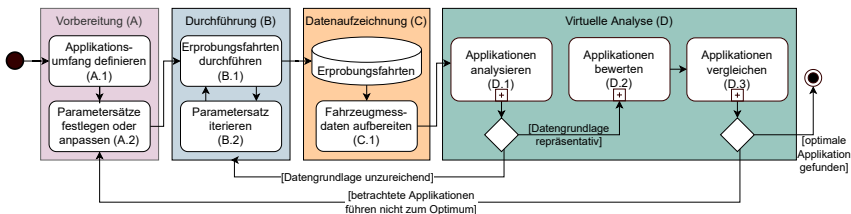


Abbildung 6.1: Prozess der global-optimalen, datengetriebenen Applikation mit vier Prozessphasen (A-D) und den jeweiligen Prozessschritten.

Um länderspezifische Besonderheiten zu erfassen, muss in allen relevanten Ländern appliziert werden (Anforderung 4). Um zusätzlich eine annähernd identische

Funktionsqualität in allen Ländern sicherzustellen (Anforderung 5), müssen die (finalen) Parametervariationen in allen relevanten Ländern getestet werden. Durch einen **iterativen Ansatz** sollen zunächst Parametersätze mit potentiell hoher Funktionsqualität identifiziert werden. Aus Effizienzgründen kann diese Vorauswahl auf kleineren Datensamples bzw. Testfahrten beruhen. Als Startpunkte für die Parametervariation dient das Vorwissen aus der Entwicklung. Anschließend werden auf Basis der systematischen Analyse neue Parametersätze festgelegt (A.2). Während der Durchführung müssen die **Parametersätze im Rahmen von Erprobungsfahrten systematisch variiert** werden ((B.1 & B.2), um die Funktionsqualität in allen relevanten Ländern bewerten zu können (Anforderung 4, 6 und 7).

Straßenverkehrstests sind nur bedingt beeinflussbar. Die auftretenden Situationen können variieren und unterliegen dem Zufall bzw. sind abhängig von der jeweiligen Fahrzeugumgebung, welche potentiell länder- bzw. regionsspezifisch ist. Um dennoch eine Vergleichbarkeit (vgl. Voruntersuchung in Abschnitt 5.1.1) zwischen den Testfahrten mit unterschiedlichen Parametersätzen zu ermöglichen (Anforderung 6), müssen die Fahrten zerteilt und in kleinere Abschnitte aufgetrennt werden. Die Abschnitte in den Fahrten können anschließend in Form von Szenarien (vgl. Definition 24) interpretiert werden (vgl. dazu Abschnitt 6.2 und Abbildung 6.2). Im Rahmen der Analyse der Applikationen (D.1) werden die **Szenarien in den Straßenverkehrstests identifiziert**. Während der Bewertung der Applikationen (D.2) wird eine **Bewertung auf Szenarienebene** durchgeführt, über diese aggregiert und die Funktionsqualität über Szenarien miteinander verglichen. Als Teil der virtuellen Analyse (D) laufen diese Prozessschritte im zentralen Backend ab. Eine vorgelagerte Speicherung und Aufbereitung der Fahrzeugmessdaten (C.1) schaffen die notwendigen Voraussetzungen.

Die Umfänge der Straßenverkehrstests für die Applikation sind aufgrund der organisatorischen Rahmenbedingungen (vgl. Abschnitt 5.1.3) und fest getakteter Entwicklungsprozesse (vgl. Abschnitt 2.3 und 2.5) begrenzt und bilden stets nur eine Teilmenge des gesamten Straßenverkehrs ab. Sie stellen im Bezug auf die Grundgesamtheit Stichproben dar. Um mögliche **Über- bzw. Unterrepräsentationen** (vgl. Voruntersuchung in Abschnitt 5.1.2) - insbesondere auch durch

Bias (vgl. Voruntersuchung in Abschnitt 5.1.3) bedingt - auszugleichen, müssen die Auftretenshäufigkeiten ähnlicher Szenarien erfasst und gegen eine **erwartete Häufigkeitsverteilung in der Grundgesamtheit gewichtet** werden. Die Gewichtung bildet die Grundlage für eine repräsentative Bewertung während der Bewertung einer Applikation, die anschließend den Vergleich der Applikationen (D.3) ermöglicht.

Die global-optimale, datengetriebene Applikation lässt sich durch die Aufzeichnung aller Test- und Erprobungsfahrten während der Integrations- und Qualifikationsphasen des Systems in den Entwicklungsprozess integrieren. Die aufgezeichneten Daten müssen in eine persistente und strukturierte Speicherung in einem Backend überführt werden, in dem die Daten gesammelt und aufbereitet werden. Anschließend stehen diese für Analysen und Vergleiche bereit und ermöglichen somit auch eine virtuelle Bewertung der verwendeten Parametersätze.

Während die Umsetzung einer weltweiten Datenkollektion und -sammlung (Prozessschritte B.2 & C.1) technische und organisatorische Aufwände bedeutet, besteht für die Analyse (D.1), die Bewertung (D.2) und den Vergleich (D.3) der Applikationen Bedarf an geeigneten Verfahren und Methoden. Es fehlt insbesondere an Methoden zur Interpretation der Fahrzeugmessdaten im Backend sowie deren Aggregation und Bewertung für die Applikation. Das Ziel ist, ein strukturiertes und systematisches Vorgehen für die Applikation zu entwickeln und dadurch einen global-optimalen, datengetriebenen Applikationsprozess sicherzustellen.

6.2 Datenmodell: Daten- und Bedeutungsebene von aufgezeichneten Straßenverkehrstests

Die durchgeführten und aufbereiteten Testfahrten liegen in Form von Zeitreihen-Daten vor. Diese enthalten nach der Datenaufbereitung (C.1) alle Informationen, die zu jedem Zeitpunkt der Testfahrt im Fahrzeug zur Verfügung standen. Dennoch

lässt sich aus den Zeitreihen selbst weder direkt auf die erlebten Situationen während der Fahrt noch auf das beobachtete Fahrzeugverhalten schließen.

Ein Vergleich unterschiedlicher Testfahrten darf nicht auf der Datenebene allein erfolgen (vgl. Diskussion in Abschnitt 5.1.1). Für einen Vergleich von Applikationen auf Basis von Testfahrten muss auf der Bedeutungsebene zwischen den während der Testfahrten vom zu applizierenden System durchgeführten Fahraufgaben und den jeweils vorliegenden Umgebungssituationen differenziert werden. Die Kombination aus einer Fahraufgabe und den Umgebungssituationen ergibt im Kontext dieser Dissertation ein Szenario. D.h. ein Szenario enthält immer genau eine vom System ausgeführte Fahraufgabe der zu applizierenden Funktion.

Die Inhalte der Testfahrten müssen durch eine Interpretation der Umgebungssituation und durchgeführten Fahraufgabe auf die Bedeutungsebene transferiert werden. Die Interpretation beschreibt also eine Abbildung der auf der Datenebene vorliegenden Testfahrten bzw. der darin enthaltenen Fahrabschnitte auf Szenarien auf der Bedeutungsebene. Diese Interpretation der Umgebungssituation und durchgeführten Fahraufgabe sowie die Bewertung des Fahrzeugverhaltens geschehen während einer Testfahrt unterbewusst durch die Wahrnehmung des Testfahrers bzw. Applikateurs, müssen aber bei der Analyse von Fahrtaufzeichnungen explizit durchgeführt werden. Die Interpretation muss durch die Kombination der einzelnen aufgezeichneten Signale erfolgen.

Zur Unterscheidung der im Applikationskontext anfallenden, kontinuierlichen Daten von ihrer Bedeutung für die Applikation ist eine eindeutige Trennung notwendig. Es muss zwischen einer Datenebene und einer Bedeutungsebene unterschieden werden (vgl. Abbildung 6.2).

Auf der **Datenebene** existiert für jeden Parametersatz ein Datensatz D , der ein oder mehrere Erprobungen E enthält, die wiederum aus ein oder mehreren Testfahrten T bestehen. Jede Testfahrt besteht nach der Datenaufbereitung aus äquidistant abgetasteten Zeitpunkten f , wobei zu jedem Zeitpunkt f_i die Zeitinformation t_i sowie z weitere Signalinformationen y_r , $1 \leq r \leq z$ mit dem aktuellsten gemessenen Wert vorliegen (vgl. Gleichung 6.1). Dabei gilt für den Laufindex i über

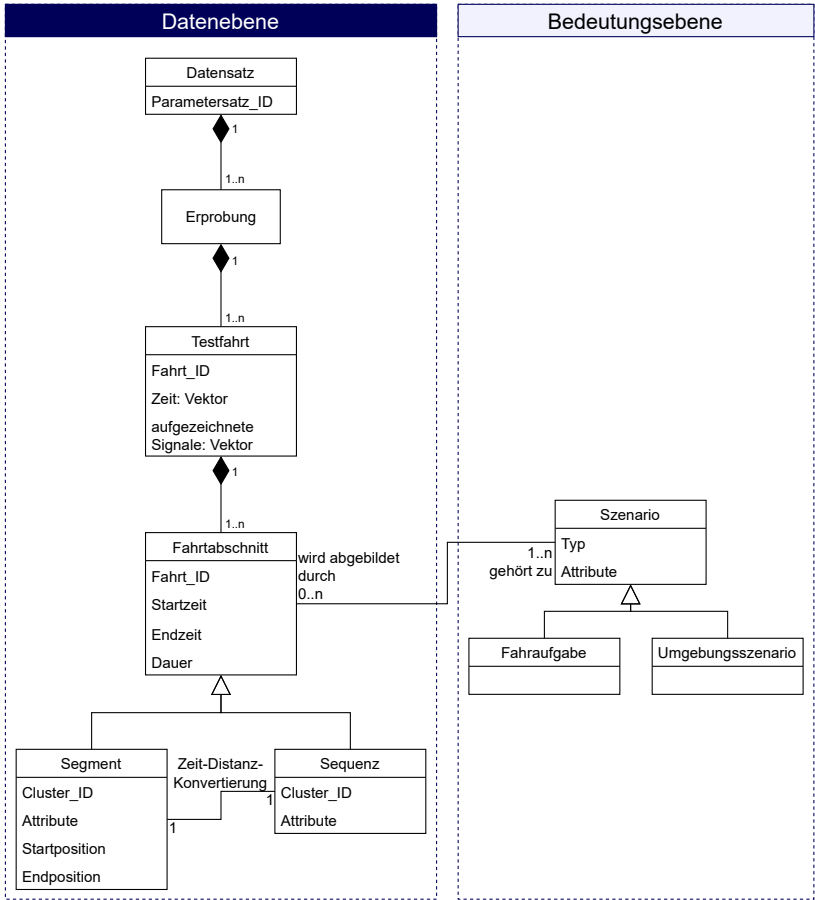


Abbildung 6.2: Klassendiagramm zur Beschreibung der Datenstruktur von Datensätzen und ihrer Bedeutung im Applikationskontext.

die Zeit $0 \leq i \leq \text{dur}(T) * \text{freq}(T)$, wobei $\text{dur}(T)$ die Dauer einer Testfahrt in Sekunden und $\text{freq}(T)$ die Abtastfrequenz der Testfahrt in Hertz angibt.

$$f_i = \begin{pmatrix} t_i \\ y_{1i} \\ \vdots \\ y_{zi} \end{pmatrix}, 0 \leq i \leq \text{dur}(T) * \text{freq}(T) \quad (6.1)$$

Einzelne aufeinanderfolgende Zeitpunkte einer Testfahrt können zu zusammenhängenden Fahrabschnitten F gruppiert werden (vgl. Gleichung 6.2). Jeder Zeitpunkt f des Datensatzes kann keinem, einem oder mehreren Fahrabschnitten zugeordnet sein.

$$F_{ij} = \begin{pmatrix} f_i & \cdots & f_j \end{pmatrix} = \begin{pmatrix} t_i & \cdots & t_j \\ y_{1i} & \cdots & y_{1j} \\ \vdots & \ddots & \vdots \\ y_{zi} & \cdots & y_{zj} \end{pmatrix}, 0 \leq i < j \leq \text{dur}(T) * \text{freq}(T) \quad (6.2)$$

Vereinfachend können die Testfahrten eines Datensatzes aneinandergereiht betrachtet werden, sodass $0 \leq i \leq \text{dur}(D) * \text{freq}(D)$ gilt, wobei $\text{dur}(D)$ die Gesamtdauer der Fahrten des Datensatzes in Sekunden und $\text{freq}(D)$ die einheitliche Abtastfrequenz aller Testfahrten im Datensatz in Hertz angibt. Es darf sich jedoch kein Fahrabschnitt über zwei Testfahrten des Datensatzes erstrecken.

$$f_i = \begin{pmatrix} t_i \\ y_{1i} \\ \vdots \\ y_{zi} \end{pmatrix}, 0 \leq i \leq \text{dur}(D) * \text{freq}(D) \quad (6.3)$$

Der Zeitbezug t eines Fahrabschnittes kann durch Konvertierung auf eine relative Position entlang der gefahrenen Strecke abgebildet werden. Fahrabschnitte mit zeitlicher Referenz werden als Sequenzen und Fahrabschnitte mit Positionsbezug als Segmente bezeichnet.

Die **Bedeutungsebene** steht der Datenebene gegenüber. Sie beschreibt die inhaltliche Interpretation des Fahrtinhalts - also das Verhalten des Ego-Fahrzeugs und anderer Verkehrsteilnehmer sowie die Fahrzeugumgebung. Beim szenarienbasierten Testen wird das Konstrukt des Szenarios (vgl. Definition 24) verwendet, welches sich aus einer Abfolge von Szenen (vgl. Definition 25) zusammensetzt. Die Szenariodefinition und insbesondere die -beschreibungsfom des szenarienbasierten Testens zielt auf eine Handlungsbeschreibung zur Ausführung bzw. Simulation des Szenarios ab und lässt sich nur bedingt auf aufgezeichnete Testfahrten übertragen, da in der Aufzeichnung vorhandene Aktionen und Events nicht beeinflussbar sind und die Ziele des SUT nicht definierbar sind. Das Gedankenkonstrukt des Szenarios als inhaltlich zusammenhängende Abfolge von Szenen bzw. bei aufgezeichneten Testfahrten aus der Perspektive des Ego-Fahrzeugs einer inhaltlich zusammenhängenden Abfolge von Situationen (vgl. Definition 26) soll jedoch verwendet werden.

Ein Fahrabschnitt F kann durch **Interpretation** auf die Bedeutungsebene abgebildet werden. Auf der Bedeutungsebene stellt jeder Zeitpunkt mit allen Signalinformationen f_i eine Situation dar. Ein Szenario besteht aus einer Abfolge von Situationen, wobei eine einzelne Situation wiederholt auftreten oder eine (bestimmte) Abfolge von Situationen vorliegen kann. Durch Algorithmen (vgl. Abschnitt 6.3) wird für jedes Szenario S eine Abbildungsfunktion α_S definiert, die alle Fahrabschnitte des Szenarios im Datensatz identifiziert (vgl. Gleichung 6.4). Dies ermöglicht den semantischen Vergleich unterschiedlicher Fahrabschnitte über die Bedeutungsebene und damit eine tatsächliche Vergleichbarkeit von sequentiellen Straßenverkehrstests sowie die Bewertung unterschiedlicher Parameterausprägungen innerhalb eines Szenarios - also bei gleicher Fahraufgabe und ähnlichen Umgebungssituationen.

Der Datensatz D besteht aus n Fahrabschnitten $F: D = \{F_1, \dots, F_n\}$

$$\alpha_S(D) = \{F_1, \dots, F_m\}; m \leq n \quad (6.4)$$

Ein Fahrabschnitt F eines Datensatzes D kann ein oder mehrere Szenarien repräsentieren. Fahrabschnitte desselben Szenarios sind nicht überlappend, unterschiedliche Szenarien können jedoch gleichzeitig auftreten (vgl. Abbildung 6.3). Damit gilt für jeden Zeitpunkt eines Datensatzes, dass dieser maximal einem Fahrabschnitt in Bezug auf ein Szenario angehören kann:

$$F_{ij}, F_{kl} \in S_1 \implies j < k \quad (6.5)$$

Die Interpretation der Fahrabschnitte als Szenarien mit genau einer zugehörigen Fahraufgabe ermöglicht eine funktionale Bewertung der ausgeführten Fahraufgabe und damit auch der ausführenden Fahrfunktion. Die Bewertungen eines Szenarios werden aggregiert, um eine Gesamtbewertung des Datensatzes bezüglich der Ausführung einer Fahraufgabe abzuleiten. Für die Bewertung müssen die Fahraufgaben aufgrund von Gleichung 6.5 unabhängig voneinander betrachtet werden.

6.3 Identifikation von Szenarien in Straßenverkehrstests

Szenarien (vgl. Definition 24) im Sinne der Applikation einer längs- oder querführenden Fahrfunktion sind die von der Fahrfunktion ausgeführten Fahraufgaben (vgl. Definition 15) in einer bestimmten Abfolge von Umgebungssituationen. Fahraufgaben lassen sich in zwei Arten differenzieren:

Zum einen gibt es manöverartige, zeitlich eindeutig abgrenzbare Fahraufgaben, deren Eigenschaften durch Attribute¹ eindeutig beschrieben werden können. Durch die kurze Dauer von wenigen Sekunden ändert sich die Umgebungssituation von manöverartigen Fahraufgaben nicht. Zum anderen gibt es kontinuierliche, zeitlich nicht eindeutig abgrenzbare Fahraufgaben, währenddessen sich die Umgebungssituation und damit die Ausprägungen der Attribute stark verändern. Eine kontinuierliche Fahraufgabe lässt sich nicht mit einer Ausprägung je Attribut beschreiben. Um kontinuierliche Fahraufgaben dennoch eindeutig zu beschreiben und gegenseitig zu differenzieren, werden diese in den sich veränderten Umgebungssituationen einzeln betrachtet. Beispiele für manöverartige Fahraufgaben sind Spurwechsel, Abbiegevorgänge, Auffahrmanöver, Geschwindigkeitsanpassungen in Kurven sowie Reaktionen auf Ein- und Ausscherer. Kontinuierliche Fahraufgaben sind unter anderem die Spurmittenführung sowie die Folgefahrt (vgl. Abbildung 6.3).

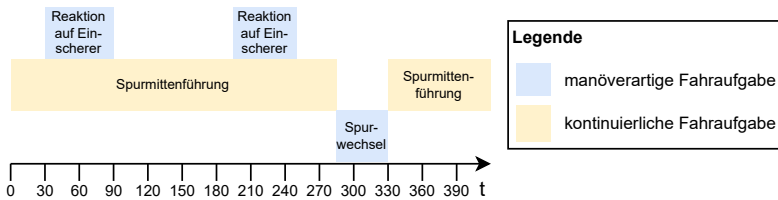


Abbildung 6.3: Schematische Darstellung dreier Fahraufgaben in einem Straßenverkehrstest. Die Reaktion auf Einscherer und der Spurwechsel sind manöverartige Fahraufgaben wohingegen die Spurmittenführung kontinuierlich (außer während des Spurwechsels) aktiv ist.

Für jede Fahraufgabe gibt es spezielle Parameter, die nur während der Durchführung der Fahraufgabe wirken und entsprechend währenddessen appliziert werden müssen. Abhängig von der zu applizierenden Fahraufgabe sind unterschiedliche Fahrtabschnitte der Straßenverkehrstests für die Applikation relevant. Für die Spurmittenführung werden unter anderem Parameter appliziert, die die Geschwindigkeit beeinflussen, in der die Funktion auf eine Veränderung der Spuren

¹ Im Kontext des szenariobasierten Testens wird der Begriff Parameter verwendet. Da dieser im Rahmen der Applikation bereits belegt ist, wird hier von Attributen gesprochen.

reagiert. Ein schnelles Reagieren führt zu einer hohen Treue der Spurmittenführung, wirkt sich aber bei häufigen Spuränderungen negativ auf den Fahrkomfort aus. Beim Spurwechsel werden beispielsweise die maximale Haltezeit eines angeforderten Spurwechsels oder die Größe der Mindestlücke auf der Nebenspur appliziert.

6.3.1 Identifikation von Szenarien mit manöverartigen Fahraufgaben

Es kann explizit nach einzelnen, im Voraus definierten Manövern des Ego-Fahrzeugs oder anderer Verkehrsteilnehmer in den Daten gesucht werden. Dazu wird aus der Fahraufgabe bzw. dem gesuchten Manöver eine Abfolgebeschreibung erstellt, die anschließend in einen Detektor umgewandelt wird, der einzelne Fahrabschnitte aus dem gesamten Straßenverkehrstest extrahiert. Bei den extrahierten Fahrabschnitten handelt es sich um konkrete Ausprägungen der gesuchten Fahraufgabe. Für weiterführende Analysen werden nur die Abschnitte berücksichtigt, in denen die gesuchte Fahraufgabe detektiert wurde.

Zur Detektion von Manövern und Szenarien in Fahrzeugmessdaten gibt es verschiedene Ansätze (vgl. Tabelle 6.1). Im Folgenden werden für die Detektion der Szenarien mit manöverartigen Fahraufgaben (analog zu [JL3]) Umgebungs- bzw. Aktivierungsbedingungen definiert, die als Einstiegspunkt für die Detektion einer Abfolge von Situationen genutzt werden können. Liegen die definierten Situationen mit optionalen Toleranzen für die Übergänge zwischen zwei Situationen vor, kann der Fahrabschnitt zu einem Szenario aggregiert werden (vgl. Abbildung 6.4).

Autor und Quelle	Jahr	Schwerpunkt des Ansatzes	Kommentar
Roesener et al. [132]	2016	Detektion von Spurwechselmanövern, Freifahrt und Einscherern	basiert auf Bus-Kommunikationsdaten und gelabelten Manövern
Elrofai et al. [133]	2016	Detektion von Abbiege- und Spurwechselmanövern des Ego-Fahrzeugs	basiert auf Bus-Kommunikationsdaten und gelabelten Manövern
Barbier et al. [134]	2017	trajektorienbasierter Ansatz zur Detektion von longitudinalen Manövern und Abbiegemanövern	Ansatz nutzt synthetische Daten
Hartjen et al. [135]	2019	Detektion von Manövern des Ego-Fahrzeugs sowie der Interaktion mit anderen Verkehrsteilnehmern und der Infrastruktur	umfassender Manöverkatalog; nutzt Infrastrukturdaten
Erdogan et al. [136]	2019	Detektion von Manövern des Ego- und Vorderfahrzeugs	Ansatz nutzt synthetische Daten zum Training der Klassifikatoren
Hauer et al. [137]	2020	Clustering-Ansatz von Trajektorien um das Ego-Fahrzeug	unüberwacht, wenige Annahmen, Interpretation notwendig
Pfeffer [138]	2020	Detektiert räumliche Manöver anderer Verkehrsteilnehmer relativ zum Ego-Fahrzeug	Ansatz nutzt synthetische Daten; Manöver im Nachgang zu interpretieren
Elsas et al. [JL3]	2020	Detektion konkreter Szenarien; kann Testfall-Logik abbilden	nutzt reguläre Ausdrücke
King et al. [JL6]	2021	Detektion konkreter Szenarien; anschließende Aggregation zu logischen Szenarien	nutzt Infrastrukturdaten
Braun et al. [139]	2022	Detektion und Manöver-basierte Ähnlichkeit von Szenarien	nutzt Drohrendaten

Tabelle 6.1: Überblick über Ansätze zur Manöver- und Szenariendetektion.

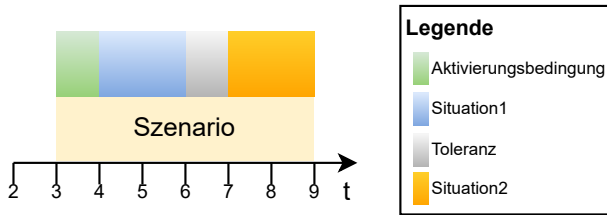


Abbildung 6.4: Schematische Darstellung des Detektionsprinzips für manöverartige Fahraufgaben.

6.3.2 Identifikation von Szenarien mit kontinuierlichen Fahraufgaben

Neben Fahraufgaben, die örtlich und zeitlich eindeutig begrenzt sind, gibt es auch Fahraufgaben, die das Fahrzeug dauerhaft oder mindestens über längere Zeiträume überwachen oder aktiv regeln und deren Verhalten sich nicht eindeutig in Manöver des Ego-Fahrzeugs unterscheiden lässt. Dies kann beispielsweise bei der Spurmittenführung durch ein Automated Lane Keeping System (ALKS) der Fall sein. Um für solche Fahraufgaben eine differenzierte Bewertung zu ermöglichen, muss der Teil der Fahrt, in der die Fahraufgabe durchgeführt wurde, in ihrer Gesamtheit erhalten, jedoch in einzelne, semantisch ähnliche Abschnitte untergliedert werden. Die in Tabelle 6.1 vorgestellten Ansätze erfüllen diese Anforderung nicht. Verbreitete Ansätze wie das Fenstern mittels Sliding Window oder das Segmentieren mit fester Länge führen dazu, dass zusammenhängende Umgebungssituationen zerschnitten oder unpassend zusammengefasst werden und dadurch die Bewertung und Interpretation erschwert wird.

Zur **Segmentierung** der Strecke in Segmente mit dynamischer Länge² werden Attribute, die die relevante Fahrzeugumgebung beschreiben, ausgewählt. Die Umgebungsinformationen liegen dabei entweder als kategoriale oder metrische Werte vor. Die Ausprägungen eines kategorialen Werts lassen sich für ein Segment nicht aggregieren. Daher werden diese zur Segmentierung, also dem Zerschneiden der

² Veröffentlicht in *Logical Scenario Derivation by Clustering Dynamic-Length-Segments Extracted from Real-World-Driving-Data* [JL11]

Testfahrten in Abschnitte, verwendet. Metrische Werte werden als Attribute eines Segments hinzugefügt. Die Aggregation kann zum Beispiel durch Mittelwert- und Medianbildung erfolgen. Bei zu starker Segmentierung (z.B. durch sehr häufige Attributänderungen) können Cluster- und Merge-Strategien oder Methoden zur Entprellung der Attributwerte angewandt werden. Die entstandenen Segmente dynamischer Länge werden anschließend interpretiert und in der weiteren Analyse wie Szenarien behandelt.³

Beispiel 3: Die Segmentierung eines Straßenverkehrstests lässt sich anhand der Attribute Tempolimit, Bauliche Trennung und Straßenklasse darstellen (vgl. Abbildung 6.5). Das erste Segment beginnt bei $t = 0$ und wird durch den Wechsel der Straßenklasse von Autobahn auf Landstraße und dem Tempolimit von 130 km/h auf 100 km/h wird nach $t = 90$ beendet. Das zweite Segment beginnt nahtlos und wird bei $t = 210$ durch den Wechsel der baulichen Trennung beendet. Analog endet das dritte Segment bei $t = 285$ beim erneuten Wechsel des Tempolimits und das vierte Segment beim Ende der Aufzeichnung. Jeder Zeitpunkt des Datensatzes wird einem Segment zugeordnet. Die Segmente können eine unterschiedliche Dauer bzw. Länge aufweisen.

6.3.3 Attribution von Fahrtabschnitten

Zur differenzierten Vergleichbarkeit und repräsentativen Bewertung müssen jedem Abschnitt weitere Attribute zugeordnet werden. Die relevanten Attribute unterscheiden sich je nach Fahraufgabe und zu applizierender Funktion.

³ Einzelnen Teilen der folgenden Abschnitte wurde durch betreute Abschlussarbeiten zugearbeitet [TH3, TH7]

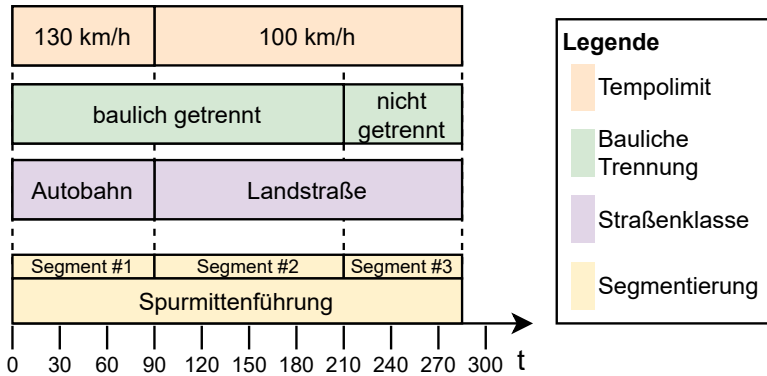


Abbildung 6.5: Schematische Darstellung der Segmentierung von Straßenverkehrstests. Die Ausprägungen der die Fahrzeugumgebung beschreibenden Attribute führen zur Segmentierung der Fahrt. Alle Zeitpunkte werden berücksichtigt.

Beispiel 4: Für die Unterscheidung von *Überholvorgängen* zur Bewertung einer ACC-Funktion ist die Geschwindigkeit des Ego-Fahrzeugs sowie Abstand und Geschwindigkeit der anderen Verkehrsteilnehmer relevant. Beim *Spurhalten* eines LKA hingegen sind Qualität und Anzahl der detektierten Spuren, die Fahrbahnbreite und Wetter- bzw. Witterungseffekte entscheidend.

Aus den aufgezeichneten Fahrzeugdaten sowie den angereicherten Informationen können diese Merkmale für jeden Fahrabschnitt als Attribut hinterlegt werden.

Jedem Szenario S werden durch diese Attribution konkrete Ausprägungen für jedes Attribut A zugewiesen. Für kategoriale Attribute werden Häufigkeiten oder die konkrete Ausprägung des Attributs verwendet. Metrische Attribute werden als statistische Kennzahlen (Mittelwerte, Maxima, Minima, ..) erfasst oder durch Binning diskretisiert.

Über die u Attribute eines Szenarios entsteht ein Attributraum, welcher durch Binning oder Untergliederung der einzelnen Attributwerte strukturiert wird. Jede Attributkombination stellt ein SubszENARIO N des Szenarios dar. Die Menge aller Attributkombinationen bildet sich aus $M_A = A_1 \times \dots \times A_u$.

Damit die Menge der Attributkombinationen und damit die Menge der Subsznarien endlich bleibt, müssen die Wertebereiche beschränkt werden und für die metrischen Attribute eine Diskretisierung erfolgen, sodass jedes Attribut A_v eine endliche Anzahl diskreter Ausprägungen w (Bins oder Kategorien) besitzt (vgl. Gleichung 6.6).

$$A_v = \{a_{vx}\}, 1 \leq x \leq w \quad (6.6)$$

Jedes Element der Menge M_A bildet ein SubszENARIO $N_h, 1 \leq h \leq |M_A|$. Jeder Fahrtabschnitt kann durch seine Attributausprägungen eindeutig einem SubszENARIO N_h zugeordnet werden (vgl. Abbildung 6.6).

$$\forall F \in S, \exists! N \in M_A : F \in N \quad (6.7)$$

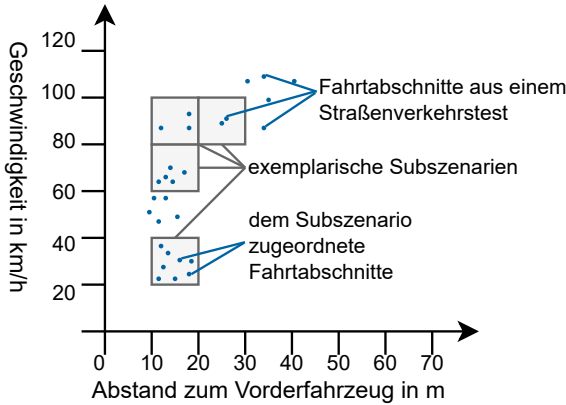


Abbildung 6.6: Schematische Darstellung der Gruppierung von Fahrtabschnitten eines Szenarios durch die Subsznarien im Attributraum.

Jedes Szenario besteht also aus $g = |M_A|$ Subsznarien: $S = \{N_1, \dots, N_g\}$. Die Subsznarien unterscheiden sich in den Attributen $A_v, 1 \leq v \leq u$. Der Datensatz

besteht demnach bezüglich eines Szenarios aus einer Menge von Fahrabschnitten, die durch ihre Attribute Subszenarien zugeordnet werden.

Damit wird die Abbildungsfunktion der Fahrabschnitte bezüglich eines Szenarios (vgl. Gleichung 6.4) erweitert:

$$\alpha'_S(F, A) = N_h \quad (6.8)$$

Durch die Attribution existiert eine eindeutige Abbildung der Fahrabschnitte des Szenarios zu Subszenarien (vgl. Gleichung 6.8).

Beispiel 5: Für ein *Auffahren*-Szenario werden alle Fahrabschnitte, die dem Szenario zugeordnet sind, beispielsweise nach ihren Attributen Abstand zum Vorderfahrzeug sowie der Geschwindigkeit des Ego-Fahrzeugs in Subszenarien unterteilt (vgl. Abbildung 6.6). Bei einem Binning des Abstands von 10m und einer maximalen Distanz von 70m entstehen 7 Bins. Analog wird die Geschwindigkeit in 20km/h Schritten bis zu einer maximalen Geschwindigkeit von 120km/h in 6 Bins unterteilt. Aus dem gewählten Binning ergibt sich die Anzahl Subszenarien von $6 * 7 = 42$ Subszenarien für das *Auffahren*-Szenario. Jeder Fahrabschnitt wird über seine Attributausprägungen eindeutig einem Subszenario zugeordnet. Alle Fahrabschnitte mit einem Abstand zwischen 10m und 20m sowie einer Geschwindigkeit zwischen 20km/h und 40km/h befinden sich im selben Subszenario (vgl. gelb markiertes Subszenario in Abbildung 6.6).

6.4 Methodik zur repräsentativen Bewertung von Straßenverkehrstests

Das weltweite Straßennetz umfasst mehr als 30 Mio. km [140] - die jährliche Fahrleistung allein in den USA betrug im Jahr 2022 3.196.191.000.000 km [141]. Um eine global-optimale Applikation zu ermöglichen, müssen die gesamte ODD der zu applizierenden Funktion (Anforderung 2) sowie alle relevanten Länder und

Regionen (Anforderung 4) berücksichtigt werden. Soll die Funktion im gesamten Straßenverkehr nutzbar sein, muss die Applikation die ODD ohne Einschränkungen vollständig abdecken. Einschränkungen der ODD begrenzen den Umfang der Applikation.

Für eine global-optimale Applikation müssen die identifizierten Fahrabschnitte so aggregiert werden, dass ein allgemeingültiger bzw. repräsentativer Qualitätswert für jedes Szenario abgeleitet werden kann. Aufgrund der Voruntersuchungen zur Repräsentativität (vgl. Abschnitt 5.1.2) und der potentiellen Bias von Erprobungsfahrten (vgl. Abschnitt 5.1.3) kann die Berechnung des Mittelwerts für den Qualitätswert über alle Fahrabschnitte des Szenarios die Erfüllung der Anforderungen nicht garantieren. Vielmehr müssen bei der Aggregation Verzerrungen in den Häufigkeitsverteilungen der Fahrabschnitte des Szenarios durch potenzielle Bias berücksichtigt werden. D.h. überdurchschnittlich häufig aufgetretene Fahrabschnitte müssen weniger stark und unterdurchschnittlich häufig aufgetretene Fahrabschnitte stärker berücksichtigt werden.⁴

6.4.1 Definition der Grundgesamtheit

Um die Gewichtung von Stichproben, wie beispielsweise einzelner Straßenverkehrstests, vorzunehmen, muss zunächst die Grundgesamtheit als Referenz definiert werden. Für die Grundgesamtheit müssen die Häufigkeitsverteilungen der Szenarien und der jeweiligen Subszenarien bekannt sein. Die Grundgesamtheit sollte nach Möglichkeit die gesamte ODD des Systems abbilden, um die realen Auftretenshäufigkeiten aller Szenarien sowie ihrer Attributräume korrekt abzubilden. Für die Festlegung der realen Auftretenshäufigkeiten gibt es unterschiedliche Herangehensweisen (vgl. Tabelle 6.2):

⁴ Einzelnen Teilen der folgenden Abschnitte wurde durch betreute Abschlussarbeiten zugearbeitet [TH10, TH2]

Grundgesamtheit	Gesamtes Streckennetz	Kundenflotte	Gesamtheit aller Erprobungsdaten
Repräsentativität der statischen Umgebung	++	+	--
Repräsentativität der dynamischen Umgebung	--	++	+
Aufwand für die Datenakquise	-	+	++

Tabelle 6.2: Gegenüberstellung potentieller Datenquellen für die Ableitung der Grundgesamtheit.

Für eine Applikation, die insbesondere auf die statische Umgebung abzielt, kann das gesamte Streckennetz als Referenz herangezogen werden. Dabei wird die Häufigkeit der Szenarien durch die Varianz innerhalb des Streckennetzes bestimmt. D.h. zum Beispiel wird die Häufigkeit und Attributsausprägung von Kurven durch die Anzahl gebauter Kurven definiert. Als Datenquelle werden elektronische Kartendaten herangezogen.

Sind für die Applikation dagegen Fahrmanöver bzw. die Interaktion mit anderen Verkehrsteilnehmern von vorrangiger Bedeutung, sind elektronische Kartendaten ungeeignet. Vielmehr müssen aufgezeichnete Straßenverkehrstests aus der Kundenflotte oder den Erprobungsfahrten verwendet werden, in denen die dynamische Fahrzeugumgebung enthalten ist. Im Gegensatz zur Verwendung des Streckennetzes wird bei der Verwendung aufgezeichneter Straßenverkehrstests nicht jeder gebaute Streckenabschnitt genau einmal sondern entsprechend der Anzahl Durchfahrten gezählt. Streckenabschnitte entfallen oder fließen einfach bzw. mehrfach in die Grundgesamtheit ein. Als Datenquelle werden zum Beispiel GPS-Informationen aus Kundenflotten, vorhandene Datensätze, auch von anderen Modellreihen, oder Drohen- und Infrastrukturdatensätze verwendet.

6.4.2 Aggregation der Fahrtabschnitte über den Szenario-Attributraum

Unter Annahme einer objektiven Bewertungsfunktion (rating) $r(F)$ für Fahrtabschnitte eines Szenarios lässt sich das Szenario beispielsweise durch die Bildung des Mittelwerts über alle identifizierten Fahrtabschnitte des Szenarios evaluieren. Die Bewertung R eines Szenarios S ergibt sich aus:

$$R_S = \frac{1}{|F_S|} * \sum_{F \in S} r(F) \quad (6.9)$$

Die Mittelwertbildung des Qualitätswerts aller Fahrtabschnitte eines Szenarios R_S (vgl. Gleichung 6.9) bildet die tatsächlichen Häufigkeiten des Szenarios und seiner Attributausprägungen in der jeweils betrachteten Stichprobe ab. Dies muss aber durch potenzielle Bias nicht den Häufigkeiten der Grundgesamtheit und damit des geplanten Einsatzgebiets der zu applizierenden Funktion entsprechen. Anforderung 3 wird verletzt. Da die Attributausprägung eines Szenarios einen maßgeblichen Effekt auf die Qualität der Funktion hat, muss bei der Aggregation zwischen den Attributausprägungen differenziert werden. Durch den über die Attribution aufgespannten Szenario-Attributraum (vgl. Abbildung 6.7) und die daraus abgeleiteten Subszenarios muss eine differenzierte Bewertung für jedes SubszENARIO durchgeführt werden, um anschließend über alle Subszenarios mit Hilfe von Gewichtungsfaktoren und den aus der Grundgesamtheit bekannten erwarteten Häufigkeitsverteilungen repräsentativ zu aggregieren.

Für alle Fahrtabschnitte F eines Subszenarios N_h kann der mittlere Qualitätswert des Subszenarios R_N über die Mittelwertbildung berechnet werden:

$$R_N = \frac{1}{|F_N|} * \sum_{F \in N} r(F) \quad (6.10)$$

Je nach Anwendungsfall und Applikationsziel sind neben dem Mittelwert weitere Metriken zur Aggregation geeignet. Beispielsweise gewichten quadratische oder

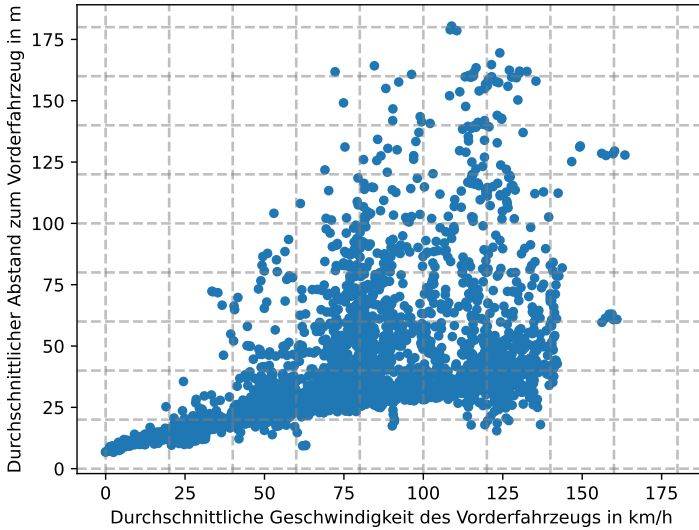


Abbildung 6.7: Häufigkeitsverteilung von Fahrtabschnitten eines Beispieldatensatzes im Attributraum der Geschwindigkeit und des Abstands zum Vorderfahrzeug.

kubische Fehlermetriken (vgl. Root Mean Squared Error in Gleichung 6.11 und Cubic-Root Mean Cubic Absolute Error in Gleichung 6.12) Ausreißer und hohe Abweichungen vom Soll-Verhalten stärker.

$$R_{N_{\text{RMSE}}} = \sqrt{\frac{1}{|F_N|} * \sum_{F \in N} r(F)^2} \quad (6.11)$$

$$R_{N_{\text{CRMCAE}}} = \sqrt[3]{\frac{1}{|F_N|} * \sum_{F \in N} |r(F)|^3} \quad (6.12)$$

6.4.3 Gewichtung der Subszenarien

Bei einer angestrebten Gleichbehandlung aller Subszenarien kann die Bewertung des Szenarios durch Bildung des Mittelwertes über die aggregierten Bewertungen aller Subszenarien berechnet werden:

$$R_S = \frac{1}{|N_S|} * \sum_{N \in S} R_N \quad (6.13)$$

Diese Gleichbehandlung aller Subszenarien genügt jedoch nicht, um Anforderung 3 zu erfüllen. Die korrekte Repräsentation realer Häufigkeiten wird gefordert. Deshalb ist ein Abgleich mit der definierten Grundgesamtheit und eine Gewichtung abweichender Verteilungen im Datensatz erforderlich. Dazu müssen die Auftrenshäufigkeiten aller Subszenarien im aktuellen Datensatz ermittelt werden.

Mit den beiden Häufigkeitsverteilungen H der Subszenarien N des Szenarios S aus der Stichprobe (Erprobung) E sowie der Grundgesamtheit G kann nun eine Gewichtung (weighting) W der unter- bzw. überrepräsentierten Subszenarien vorgenommen werden.

Um aus einer Stichprobe E eine für die Grundgesamtheit G valide Aussage abzuleiten, müssen die Subszenario-Verteilungen in E und G verglichen und gewichtet werden.

$$R_{S_{\text{gewichtet}}} = \sum_{N \in S} R_N * W_N \quad (6.14)$$

mit

$$W_N = \frac{H_{N_G}}{|N_G|} \quad (6.15)$$

Mit der eingeführten Gewichtung der einzelnen Subszenarien N wird der Über- bzw. Unterrepräsentation einzelner Abschnitte in der Stichprobe entgegengewirkt.

Für die Gewichtung der Subszenarien ist die Wahl der Attribute und Attributausprägungen entscheidend, da diese die Größe und Dimensionalität des Attributraums und damit die Anzahl an Subszenarien eines Szenarios definieren. Je mehr Subszenarien entstehen, desto exakter die Repräsentation der Realität. Gleichzeitig nimmt jedoch die Anzahl der Fahrabschnitte je Subszenario ab und verringert damit die Abdeckung und Stichprobengröße in jedem Subszenario.

Bei einer vollfaktoriellen Betrachtung des Attributraums entstehen für ein Szenario mit u Attributen und w Attributausprägungen u^w Subszenarien (vgl. Gleichung 6.16). Bereits bei 10 Attributen mit jeweils 8 Attributausprägungen ergeben sich $8^{10} = 1.073.741.824$ Subszenarien. Eine vollfaktorielle Betrachtung ist damit nicht anwendbar. Stattdessen wird die jeweils paarweise Abdeckung zweier Attribute betrachtet (vgl. Gleichung 6.17). Es ergeben sich für das Beispiel entsprechend $\binom{10}{2} * 8^2 = 2880$ Subszenarien.

$$|N| = g = u^w \quad (6.16)$$

$$|N| = g = \binom{w}{2} * u^2 \quad (6.17)$$

Beispiel 6: Für einen Beispieldatensatz werden Fahrabschnitte mit den Attributen *Geschwindigkeit des Ego-Fahrzeugs*, *Geschwindigkeit des Vorderfahrzeugs*, *Distanz zum Vorderfahrzeug* sowie die *Spurbreite* gebildet. Für jeden Fahrabschnitt sind die Mittelwerte der Attribute berechnet und als Attributkombinationen dargestellt (vgl. Abbildung 6.8). Mittels Binning werden Gruppen für jedes Attribut gebildet, die dann Subszenarien des Szenarios ergeben. Die Anzahl Fahrabschnitte im jeweiligen Subszenario ergibt die Häufigkeit des Subszenarios (vgl. Abbildung 6.9).

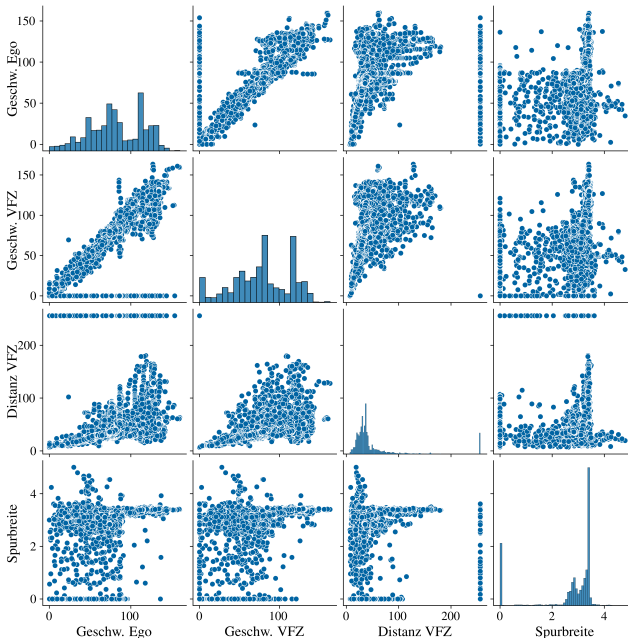


Abbildung 6.8: Paarweise Häufigkeitsverteilung von Fahrtabschnitten eines Beispieldatensatzes für die Attribute Geschwindigkeit des Egofahrzeugs [km/h], Geschwindigkeit des Vorderfahrzeugs (VFZ) [km/h] und Distanz zum Vorderfahrzeug [m] sowie Spurbreite [m]. Auf der Diagonalen sind die eindimensionalen Häufigkeitsverteilungen der Attribute der X-Achse dargestellt.

6.4.4 Substitution der fehlenden Subszzenarien

Die Gewichtung wirkt Unter- und Überrepräsentationen entgegen, funktioniert jedoch nicht, falls für ein SubszENARIO in der Stichprobe kein einziger Fahrtabschnitt vorhanden ist (vgl. Abbildung 6.9). Dieses Nicht-Vorhandensein eines Subszenarios stellt einen Sonderfall dar, da fehlende Werte nicht durch Gewichtung ausgeglichen werden können. Ohne einen einzigen Fahrtabschnitt im SubszENARIO ist es nicht möglich, die Applikationsgüte in dem SubszENARIO zu ermitteln und diese entsprechend gewichtet in die Gesamtbewertung einfließen zu lassen.

Für den Umgang mit fehlenden Subszenarios bieten sich mehrere Möglichkeiten:

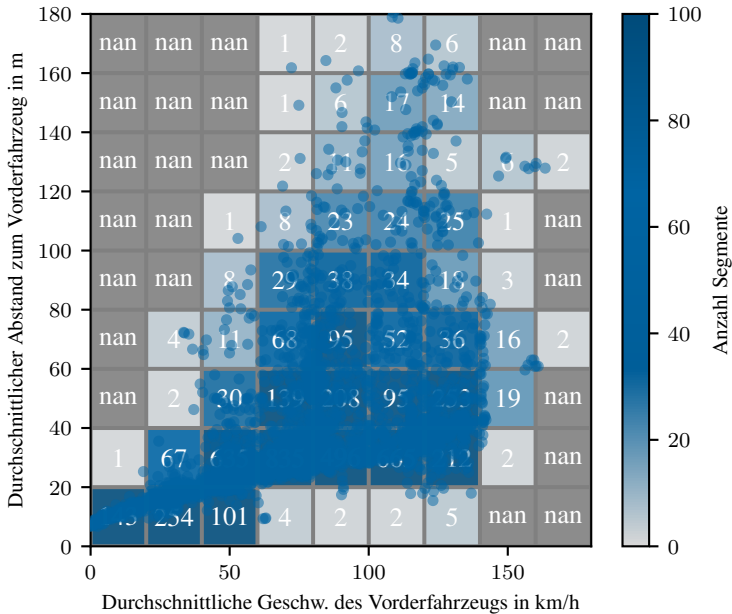


Abbildung 6.9: Unterteilung des Attributraums der Attribute Geschwindigkeit des Vorderfahrzeugs und Distanz zum Vorderfahrzeug in Subszenarien. Dunkel hinterlegte Subszenarien sind im Datensatz nicht vorhanden.

Die Subszenarien befinden sich im mehrdimensionalen Attributraum des Szenarios. Somit ist es möglich zwischen Bewertungen der umliegenden Subszenarien zu **interpolieren** bzw. an den Grenzen des Attributraums zu **extrapolieren**. Dabei muss sichergestellt werden, dass das erwartete Funktionsverhalten und damit die Bewertung der Subszenarien stetig ist und die Inter- bzw. Extrapolationen valide sind.

Ist von einer Stetigkeit nicht auszugehen, müssen fehlende Subszenarien differenziert behandelt werden. Für die Aggregation werden in diesen Fällen **Durchschnitts- oder Strafwerte** angenommen, die für die fehlenden Subszenarien eingesetzt werden. Der Durchschnittswert wird aus dem Mittel aller Subszenarien berechnet. Als

Strafwerte werden künstliche Werte eingefügt, die sich negativ auf die Gesamtbewertung auswirken, um das Fehlen des Subszenarios und damit die unbekannte Funktionsqualität in dem Subszenario negativ in die Bewertung einfließen zu lassen.

Sind Durchschnitts- oder Strafwerte für den Anwendungsfall nicht geeignet, muss die gesamte Applikation als unzureichend bzw. nicht repräsentativ ausgewiesen und zunächst **weitere Testfahrten** durchgeführt werden. Grenzbereiche des Attributraums, die bereits in der Grundgesamtheit selten auftreten, können in Einzelfallentscheidungen als optional eingestuft werden. D.h. das Fehlen von Fahrabschnitten in diesen Subszenarien wirkt sich nicht negativ auf die Gesamtbewertung aus.

Zur Auswahl des geeigneten Substitutionsverfahrens kann ein Entscheidungsprozess durchlaufen werden (vgl. Abbildung 6.10). Zunächst wird die Stetigkeit des Funktionsverhaltens sowie die Relevanz des Subszenarios überprüft. Im Falle eines kritischen Subszenarios oder eines nicht stetigen Funktionsverhaltens wird die Bewertung abgebrochen, um weitere Daten zu akquirieren. Andernfalls wird das fehlende Subszenario substituiert. Dabei wird zwischen Grenzbereich und Kernbereich des Attributraums unterschieden. Subszenarien im Kernbereich müssen durch Strafwerte oder Inter- bzw. Extrapolation ersetzt werden. Subszenarien im Randbereich werden vernachlässigt.

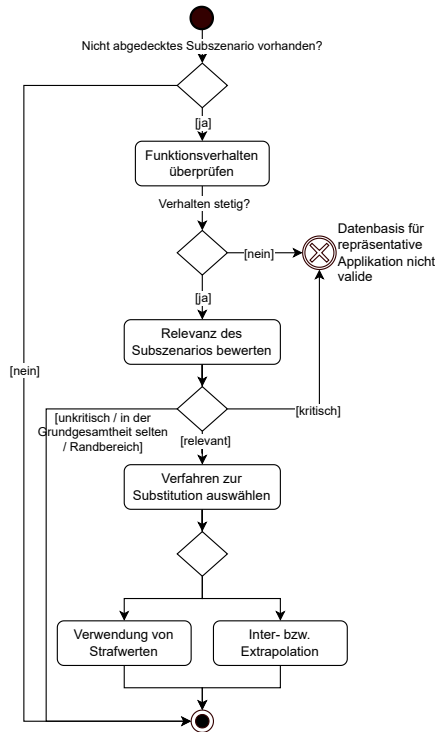


Abbildung 6.10: Entscheidungsprozess für die Auswahl geeigneter Substitutionsverfahren.

Beispiel 7: Für die Geschwindigkeit des Vorderfahrzeugs sowie den Abstand zum Vorderfahrzeug (vgl. Abbildung 6.9) ist zu argumentieren, dass Subszenarios mit einem Abstand > 100m sowie einer Geschwindigkeit > 140 km/h (Bereich oben rechts in der Abbildung) für ein *Auffahren*-Szenario nicht relevant sind. Das erwartete Funktionsverhalten ist stetig und die Attributwerte unkritisch. Eine fehlende Abdeckung ist dementsprechend unkritisch. Hingegen sind die Bereiche mit hohem Abstand und geringer Geschwindigkeit (oben links) bzw. geringem Abstand und hoher Geschwindigkeit (unten rechts) durchaus relevant. Sind diese nicht systemisch ausgeschlossen, müssen sie zwingend in der Applikation berücksichtigt werden. Aufgrund der Kritikalität der Attributkombinationen (hohe Geschwindigkeit bei niedrigem Abstand; niedrige Geschwindigkeit des Vorderfahrzeugs bei potenziell hoher Ego-Geschwindigkeit) können die Subszenarios nicht extrapoliert oder mit Strafwerten belegt werden. Es müssen weitere Daten

6.4.5 Aggregation der Subszenarien

Durch die Bewertung und Gewichtung der einzelnen Subszenarien ergibt sich für jedes Subszenario ein Qualitätswert der Funktionsqualität, ein Gewichtungsfaktor zur Abbildung der Häufigkeit in der Grundgesamtheit sowie durch Multiplikation der beiden Werte der gewichtete Qualitätswert für die Funktionsqualität. Eine gesamtheitliche Qualitätsmetrik der Applikation für das Szenario wird nun durch Aggregation über die Subszenarien abgeleitet. Für die Aggregation können verschiedene Verfahren gewählt werden [TH2]:

Globaler Mittelwert: Bei der globalen Mittelwertbildung wird der Mittelwert über alle Fahrabschnitte eines Szenarios gebildet. Weder die Subszenarien noch die Grundgesamtheit werden berücksichtigt. Die Übertragbarkeit der Applikation wird dadurch nicht sichergestellt.

Subszenario Mittelwert: Für die Berechnung des Subszenario Mittelwerts werden die Fahrabschnitte ihren Subszenarien zugeordnet. Über alle Fahrabschnitte eines Subszenarios wird ein Mittelwert gebildet. Über die Menge aller Subszenarien des Szenarios wird anschließend erneut ein Mittelwert berechnet. Durch die Mittelwertbildung werden alle Bereiche des Szenario-Attributraums gleich gewichtet. Häufig vertretene Subszenarien fallen weniger stark ins Gewicht.

Gewichteter Subszenario Mittelwert: Die Berechnung des gewichteten Subszenario Mittelwerts beginnt analog zur Berechnung des Subszenario Mittelwerts. Jedoch werden bei der Aggregation über die Subszenarien die Gewichtungsfaktoren aus den Häufigkeitsverteilungen der Stichprobe und Grundgesamtheit berechnet und verwendet. Jedes Subszenario fließt entsprechend der relativen Häufigkeit in der Grundgesamtheit in die Bewertung ein. Bei der globalen Gewichtung werden statt der lokalen Häufigkeitsverteilungen die Häufigkeiten in der Grundgesamtheit herangezogen (vgl. Gleichung 6.14). Die Validität der Applikation für die Grundgesamtheit wird sichergestellt, solange für jedes Subszenario eine repräsentative Anzahl an Fahrabschnitten bewertet wurde.

Extrapolation fehlender Subszenarien: Sowohl der globale als auch die Subszenario Mittelwerte können mit der Extrapolation für fehlende Subszenarien

erweitert werden, sodass eine Bewertung der Applikation auch bei lückenhafter Abdeckung des SubszENARIO-Raums vollzogen werden kann. Die Validität der Bewertung muss im Einzelfall betrachtet werden.

Strafwerte für fehlende SubszENARIEN: Analog zur Extrapolation ist die Verwendung von Strafwerten in Kombination mit den Mittelwertbildungen möglich. Durch hohe Strafwerte kann die niedrige Validität der Applikation in der Bewertung berücksichtigt werden. Strafwerte können zudem für in der Grundgesamtheit häufig und selten auftretenden SubszENARIEN abweichend definiert werden.

Zielführend für die global-optimalen, datengetriebene Applikation ist die Anwendung des gewichteten SubszENARIO Mittelwerts, da dieser die globale Häufigkeitsverteilung berücksichtigt und Bias in der lokalen Stichprobe erkennt. Das Substitutionsverfahren muss abhängig von der zu applizierenden Funktion und Fahraufgabe gewählt werden (vgl. Abbildung 6.10). Die Applikation ist dabei bezüglich der verwendeten Diskretisierung, also der gewählten Unterteilung der Szenarien in SubszENARIEN, global-optimal.

6.5 Prozess der global-optimalen, datengetriebenen Applikation

Die repräsentative Bewertung der Applikation mittels Identifikation von Szenarien in den Fahrzeugmessdaten, anschließender Bewertung der für das Szenario relevanten Fahrtabschnitte sowie gewichteter Aggregation über die SubszENARIEN muss in den bestehenden Applikationsprozess integriert werden. Der Prozess für die global-optimalen, datengetriebene Applikation besteht aus vier Phasen mit den notwendigen Prozessschritten (vgl. Abbildung 6.11). Während der Applikationsvorbereitung wird der Applikationsumfang spezifiziert sowie Testumfänge und Parametersätze definiert. Anschließend werden die notwendigen Testfahrten durchgeführt, im Backend zusammengetragen, aufbereitet und analysiert. Abhängig von den Ergebnissen müssen mehrere Iterationsschleifen durchlaufen werden, um weitere Testfahrten durchzuführen oder neue Parametersätze zu betrachten.

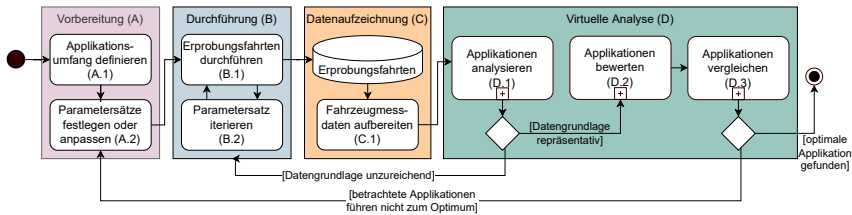


Abbildung 6.11: Prozessschritte der global-optimalen, datengetriebenen Applikation. (Wdh. der Abbildung 6.1)

6.5.1 Applikationsvorbereitung (A)

Im ersten Schritt muss der Applikationsumfang definiert werden. Dieser setzt sich aus der zu applizierenden Funktion und der zu betrachtenden ODD zusammen. Die zu applizierende Funktion kann von einzelnen von der Funktion übernommenen Fahraufgaben bis hin zur kompletten Längs- und Querführung des Fahrzeugs bei SAE Level 3+ Systemen reichen. Von der Funktion bzw. dem System leiten sich zudem die zu betrachtenden Parameter und in Kombination mit der spezifizierten ODD die relevanten Fahraufgaben ab, in denen das zu applizierende (Teil-)System aktiv in die Fahrzeugführung eingreift. Je mehr Fahraufgaben übernommen werden desto größer wird der Applikationsumfang, da mit steigender Anzahl an Fahraufgaben auch die Anzahl der zu applizierenden Funktionsparameter steigt (vgl. Abschnitt 3.3).

Um die Applikation handhabbar zu machen, muss die Applikation der ADAS und HAF-Funktionen in einzelne Fahraufgaben unterteilt werden (vgl. Abbildung 6.12). Dabei wird zwischen manöverartigen und kontinuierlichen Fahraufgaben unterschieden (vgl. Abschnitt 6.3). Einer Fahraufgabe werden anschließend für die Applikation relevante Attribute zugeordnet. Die entstehenden Szenarien können daraufhin unabhängig voneinander und damit parallel appliziert werden. D.h. während derselben Testfahrten können mehrere Szenarien, die voneinander unabhängige Parameter haben, gleichzeitig appliziert werden. Für die Fahraufgaben lässt sich die gesamte ODD des Systems auf einzelne Attributausprägungen der

Systemumgebung reduzieren, falls die zu applizierende Fahraufgabe nur in bestimmten Situationen die Fahrzeugführung übernimmt.

Für alle Attribute muss eine geeignete Diskretisierung festgelegt werden, zum Beispiel durch die Analyse vorhandener Daten. Die Diskretisierung beeinflusst die Anzahl an Subszenarien je Szenario. Dabei muss zwischen feinerer Abtastung des Attributraums durch mehr Subszenarien und robusterer Bewertung jedes Subszenarios durch mehr Fahrtabschnitte pro SubszENARIO abgewogen werden.

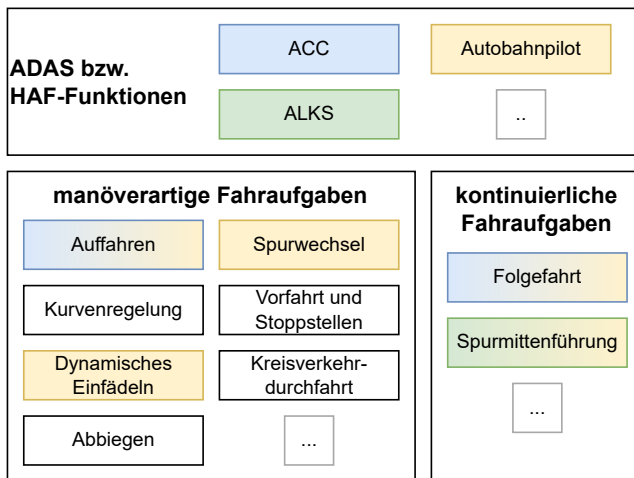


Abbildung 6.12: Dekomposition der Applikation des Gesamtsystems über die Fahraufgaben während der Definition des Applikationsumfangs (vgl. Abbildung 6.11 Prozessschritt A). Die Farbcodierung zeigt relevante Fahraufgaben für die exemplarischen ADAS bzw. HAF-Funktionen

Beispiel 8: Die ODD wird in einem morphologischen Kasten (auch als Zwicky-Box bezeichnet) mit allen relevanten Attributen und deren Ausprägungen dargestellt. Für die kontinuierlichen Fahraufgaben Spurmittenführung (ALKS) und Spurmittenführung (Autobahnпилот) werden die relevanten Attribute und Ausprägungen ausgewählt (vgl. Abbildung 6.13). Über die Kombination der Attributausprägungen ergeben sich die Szenario-Attributräume der beiden Fahraufgaben.

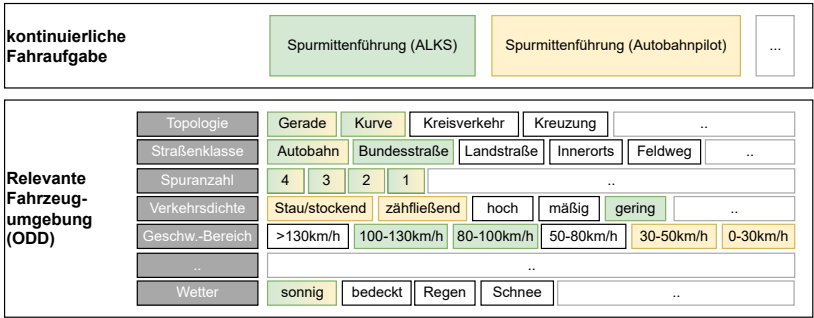


Abbildung 6.13: Festlegung von relevanten Umgebungsinformationen für kontinuierliche Fahraufgaben am Beispiel der Spurmittenführung für ein ALKS und einen Autobahnpiloten.

Zur Definition des Applikationsumfangs gehört auch die Festlegung der Erprobungsumfänge in den zu applizierenden Ländern und Regionen. Die ODD muss durch die Erprobungsumfänge abgedeckt sein. Neben der Routenplanung können auch untere Grenzen für die Beobachtung der zu betrachtenden Szenarien definiert werden. Diese werden für manöverartige Fahraufgaben in Form von Zählern oder bei kontinuierlichen Fahraufgaben kilometerbasiert spezifiziert.

Für die Aufzeichnung der Fahrzeugmessdaten während der Erprobung muss eine entsprechende Messtechnik vorgehalten werden. Zudem müssen Datenformate für die Mess- und Metadaten spezifiziert werden. Darüber hinaus muss sichergestellt werden, dass zu jeder Fahrzeugmessung auch der verwendete Parametersatz p gespeichert wird.

Um eine systematische Variation der Parametersätze sicherzustellen, muss bereits vor der Applikation ein Set zu betrachtender Parametersätze P , also konkreter Parameterausprägungen der relevanten Funktionsparameter, definiert werden, die dann während der Durchführung systematisch und wiederholt iteriert werden können. Die zu betrachtenden Funktionsparameter ergeben sich aus den im Applikationsumfang definierten Fahraufgaben. Anhand der geplanten Erprobungsumfänge kann eine zielführende Anzahl Parametersätze $|P|$ abgeschätzt werden. Zu Beginn

der Applikation müssen die Parametersätze durch Expertenwissen und Erfahrungen aus Tests während der Entwicklung abgeleitet werden. In späteren Iterationen werden neue Parametersätze aus der virtuellen Analyse abgeleitet.

Mit Abschluss der Applikationsvorbereitung ist ein Gerüst für die global-optimale, datengetriebene Applikation definiert (vgl. Abbildung 6.14). Es sind eine oder mehrere für die Applikation relevante Fahraufgaben (blau) definiert. Verschiedene Parametersätze (grau) sind definiert, die jeweils eine Applikation darstellen. Eine Gesamtbewertung der Applikation ergibt sich am Ende des Prozesses aus der Funktionsqualität und -verfügbarkeit in den einzelnen Fahraufgaben. Für jede Fahraufgabe ist ein minimaler Testumfang definiert, welcher aus organisatorischen Vorgaben resultiert oder im späteren Verlauf quantifiziert wird (vgl. Abschnitt 6.5.4.1). Ein oder mehrere Zielregionen bzw. -länder sind definiert, für die die Applikation gültig sein muss.

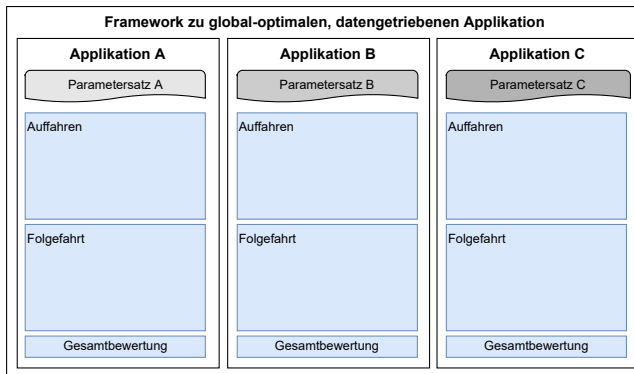


Abbildung 6.14: Framework zur global-optimalen, datengetriebenen Applikation nach Abschluss der Applikationsvorbereitung (vgl. Abbildung 6.11).

6.5.2 Applikationsdurchführung (B)

Hauptaufgabe während der Applikationsdurchführung ist die Realisierung der Straßenverkehrstests im Rahmen der spezifizierten Erprobungen. Dabei gilt es

eine möglichst hohe Varianz der relevanten Umgebungsbedingungen unter den gegebenen Rahmenbedingungen zu erleben. Subjektive Fahrindrücke in Form von Sprach- oder Textannotationen verbessern die Datenqualität. Alle Fahrten müssen mittels Messtechnik aufgezeichnet und anschließend in ein zentrales Datenbackend eingespeist werden.

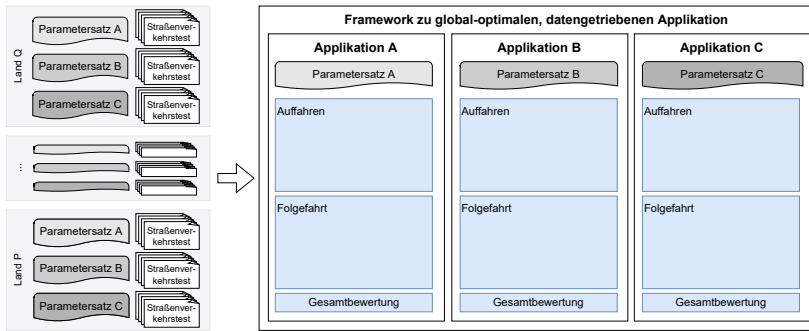


Abbildung 6.15: Die Applikationsdurchführung (vgl. Abbildung 6.11) liefert Fahrzeugmessdaten von Straßenverkehrstests für die global-optimale, datengetriebene Applikation.

Während der Erprobung gilt es die spezifizierten Parametersätze der zu applizierenden Funktion regelmäßig zu iterieren. So wird sichergestellt, dass regionale Gegebenheiten mit allen Parametersätzen erfasst werden, um die Abdeckung für jeden Parametersatz zu maximieren und eine Vergleichbarkeit zwischen den Parametersätzen zu erzielen. Geeignete Prozesse und Festlegungen müssen ein systematisches Vorgehen sicherstellen. Bei Erprobungen mit einem einzelnen Fahrzeug muss die geplante Route in $o * |P|$, $o \geq 1$ Teilrouten zerlegt und jedem Parametersatz P o Teilrouten zugewiesen werden. Bei Erprobungen mit mehreren Fahrzeugen sollten die Parametersätze parallel getestet werden, um viele ähnliche Szenarien mit unterschiedlichen Parametersätzen aufzuzeichnen und somit die Vergleichbarkeit zu verbessern.

Die systematische Strukturierung der erfassten Straßenverkehrstests beinhaltet alle Straßenverkehrstests, die nach verwendetem Parametersatz und Ursprungsland unterteilt sind (vgl. Abbildung 6.15). Jeder Straßenverkehrstest beinhaltet

verschiedene Fahrtabschnitte, in denen jeweils keine, eine oder mehrere Fahraufgaben von der zu applizierenden Funktion durchgeführt wurden. Eine bijektive Zuordnung zwischen Straßenverkehrstest und Fahraufgabe ist daher so nicht möglich.

Aufgrund der nachgelagerten Analyse müssen die Straßenverkehrstests nicht von den Entwicklern und Applikateuren selbst durchgeführt werden, solange mindestens die Buskommunikations- und Metadaten der durchgeführten Fahrten aufgezeichnet werden. Es können also auch Test- und Erprobungsfahrten anderer Funktionen sowie Dauerlaufdaten als Datenquellen herangezogen werden.

6.5.3 Datenaufzeichnung (C)

Die internen Nachrichten der Buskommunikation, die die Fahrzeugsignale enthalten, müssen während der Test- und Erprobungsfahrten aufgezeichnet und für die folgende Identifikation relevanter Fahrtabschnitte aufbereitet werden (vgl. Abschnitt 4.2). Es sind Zeitreihen-Daten mit einer einheitlichen Zeitbasis notwendig. Je nach Anwendungsfall können zusätzliche Datenquellen, wie Wetter- oder Verkehrsinformationen, eingebunden werden. Es ist darauf zu achten, dass die Zuordnung der Buskommunikations- und Metadaten, wie die SWC-Version und der verwendete Parametersatz, durchgängig erhalten bleibt.

Abhängig vom konkreten Anwendungsfall und unter Voraussetzung valider Werkzeuge und Modelle sowie hinreichender Varianz in den Szenarien und Szenario-Ausprägungen können alternativ oder ergänzend Simulationen eingesetzt werden. Die Daten müssen analog bereitgestellt und aufbereitet werden.

6.5.4 Virtuelle Analyse (D)

Die virtuelle Analyse besteht aus den Prozessschritten Applikationen analysieren (D.1), Applikationen bewerten (D.2) sowie Applikationen vergleichen (D.3) und bildet den Kern der global-optimalen, datengetriebenen Applikation (vgl.

Abbildung 6.16). Die ersten beiden Prozessschritte müssen für jede Applikation separat durchlaufen werden, um anschließend die Applikationen miteinander zu vergleichen.

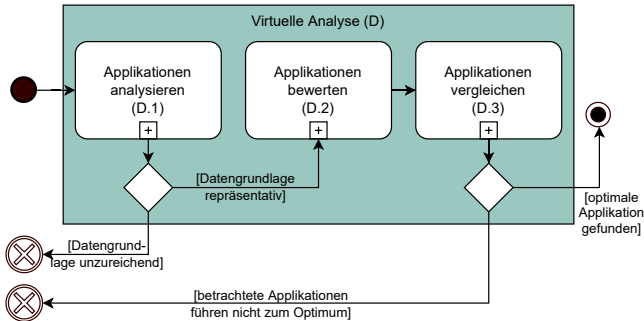


Abbildung 6.16: Prozessschritte der Phase Virtuelle Analyse (vgl. Abbildung 6.11).

6.5.4.1 Applikationen analysieren (D.1)

Während bei der Applikationsdurchführung Daten in Form von Straßenverkehrstests gesammelt wurden, ist eine direkte Bewertung dieser Straßenverkehrstests nicht zielführend (vgl. Abschnitt 6.2). Stattdessen muss zur Gewährleistung der Vergleichbarkeit und Repräsentativität ein differenzierter Analyseansatz gewählt werden.

Ziel der Analyse einer Applikation ist es, die Vergleichbarkeit der unterschiedlichen Straßenverkehrstests sowie eine Zuordnung zu den zu applizierenden Fahraufgaben herzustellen. Dafür müssen zunächst relevante Fahrabschnitte identifiziert und den entsprechenden Szenarien zugeordnet werden. Zudem muss die zugrundeliegende Datenbasis bewertet werden (vgl. Abbildung 6.17).

Um am Ende der Analyse die Bewertbarkeit der verwendeten Parametersätze sicherzustellen, müssen für jeden Parametersatz eigene Datensätze gebildet werden in denen ausschließlich Straßenverkehrstests mit dem jeweiligen Parametersatz

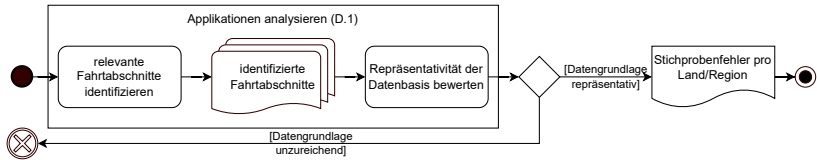


Abbildung 6.17: Prozessschritte des Teilprozesses Applikationen analysieren (vgl. Abbildung 6.16).

enthalten sind. Die folgenden Prozessschritte müssen entsprechend für jeden Datensatz durchlaufen werden.

Relevante Fahrabschnitte identifizieren Die als Ganzes vorliegenden Straßenverkehrstests müssen in relevante, semantisch zusammenhängende Fahrabschnitte zerteilt werden, um die Vergleichbarkeit zwischen den einzelnen Straßenverkehrstests (vgl. dazu Abschnitt 5.1.1 und Anforderung 3) und damit eine differenzierte Bewertung der Fahrfunktion für die einzelnen Parametersätze zu ermöglichen. Um Anforderung 3 zu erfüllen reicht es nicht, die Straßenverkehrstests äquidistant zu zerschneiden. Stattdessen müssen Szenarien extrahiert werden, die die jeweiligen Umgebungsbedingungen im Kontext einer durchgeführten Fahraufgabe beinhalten. So können Fahrabschnitte die im Kontext der Applikation einer Fahraufgabe ähnlich sind über verschiedene Straßenverkehrstests hinweg verglichen werden. Zudem lassen sich die Parametersätze in den jeweils ähnlichen Szenarien bewerten.

Die Identifikation der relevanten Fahrabschnitte muss für alle Straßenverkehrstests über alle Länder hinweg durchgeführt werden (vgl. schematische Darstellung für ein Land in Abbildung 6.18). Durch das Zerschneiden der Straßenverkehrstests in relevante Fahrabschnitte werden diese den Fahraufgaben zugeordnet (vgl. Datenmodell in Abbildung 6.2). Ein Straßenverkehrstest enthält dabei keinen, einen oder mehrere relevante Fahrabschnitte für eine oder mehrere Fahraufgaben (vgl. exemplarisch grün, lila, rot eingefärbte Straßenverkehrstests und daraus extrahierte Fahrabschnitte der Fahraufgaben *Auffahren* und *Folgefahrt*).

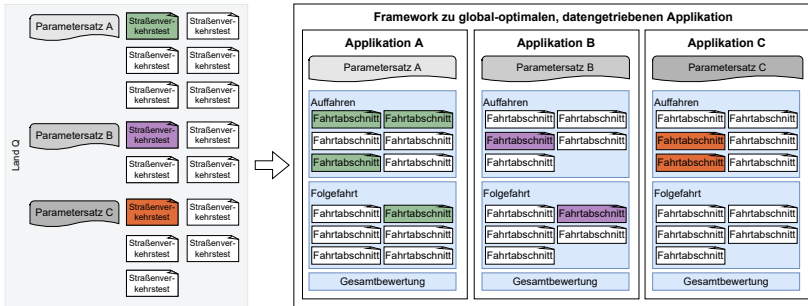


Abbildung 6.18: Framework zur global-optimalen, datengetriebenen Applikation zum Prozessschritt Applikationen analysieren (vgl. Abbildung 6.17).

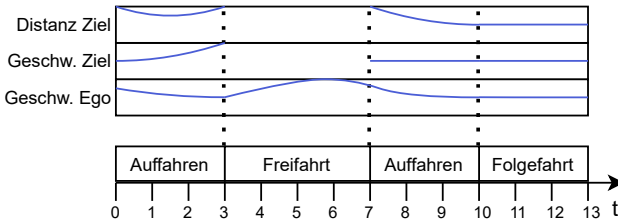


Abbildung 6.19: Detektionsprinzip für Fahrabschnitte eines Straßenverkehrstests am Beispiel der unterschiedlichen Fahraufgaben einer längsregelnden Fahrfunktion im Bezug auf ein Vorderfahrzeug.

Für jede Fahraufgabe werden Umgebungsbedingungen definiert, die diese Fahraufgabe charakterisieren und für die Identifikation relevanter Fahrabschnitte verwendet werden (vgl. exemplarische Identifikation von Fahrabschnitten, die den Fahraufgaben *Auffahren*, *Folgefahrt* und *Freifahrt* zugeordnet werden in Abbildung 6.19). Für die Identifikation wird auf die Buskommunikationssignale für die Geschwindigkeit des Ego-Fahrzeugs sowie der Kamera- bzw. Radar-basierten Objektdetektion des Vorderfahrzeugs in Form der Geschwindigkeit und des Abstands zurückgegriffen.

Dazu ist ein Vorgehen zur Identifikation relevanter, semantisch ähnlicher Fahrabschnitte notwendig (vgl. Abschnitt 6.3).

Repräsentativität der Datenbasis bewerten Nach der Identifikation der Szenarien muss sichergestellt werden, dass die der Analyse zugrundeliegende Datenbasis den Anforderungen genügt. Anforderung 4 verlangt die Applikation in allen relevanten Ländern. Aus Anforderung 2 und Anforderung 3 ergibt sich, dass die gesamte ODD in der Applikation berücksichtigt und über diese die Funktionsverfügbarkeit und -qualität optimiert werden muss.

Um dies zu gewährleisten, muss die Datenbasis hinsichtlich der Abdeckung der ODD bzw. aller Szenarien und ihrer Subszenarien bewertet werden. Unter der Annahme, dass es sich bei den Erprobungsfahrten tatsächlich um zufällige Stichproben handelt, kann der statistische Stichprobenfehler (vgl. Definition 47) herangezogen werden.

Für den Erwartungswert und die Standardabweichung werden die tatsächlichen Verteilungen der einzelnen Szenarien und Subszenarien aus der Grundgesamtheit verwendet und die Abweichung der Stichprobe von den erwarteten Verteilungen ermittelt.

Der berechnete Stichprobenfehler stellt anschließend die Konfidenz der Repräsentativität bzw. Allgemeingültigkeit der auf der Stichprobe abgeleiteten Bewertung für die Grundgesamtheit dar. Aufgrund von länderspezifischen Gegebenheiten sowie den in Abschnitt 5.1.3 aufgezeigten Abweichungen in den Erprobungsumfängen in verschiedenen Ländern und Regionen, soll der Stichprobenfehler $\sigma_{\bar{x}}$ unter Zuhilfenahme der Standardabweichung der länderspezifischen Grundgesamtheit σ_{G_c} (vgl. Gleichung 4.1 und Gleichung 4.2) für jedes Land c einzeln berechnet werden:

$$\sigma_{\bar{x}_c} = \frac{\sigma_{G_c}}{\sqrt{n_c}} \quad (6.18)$$

Der Stichprobenfehler wird maßgeblich vom Umfang - also der gefahrenen Distanz bzw. der Anzahl beobachteter Szenarien - des bewerteten Datensatzes sowie der Abdeckung der unterschiedlichen Szenarien und Subszenarien beeinflusst. Die Berechnung des Stichprobenfehlers je Land stellt sicher, dass vor der Freigabe

der Applikation in einem Land eine Mindestabdeckung der relevanten Szenarien erzielt wurde und somit potentielle länderspezifische Besonderheiten erkannt wurden.

Um Bias (vgl. Abschnitt 5.1.3) zu begegnen, müssen Maßnahmen ergriffen werden, deren Wirksamkeit vom Anwendungsfall abhängig ist. Neben der Erhöhung des Stichprobenumfangs, also der Durchführung weiterer Erprobungsfahrten, können geschichtete Stichproben (vgl. Definition 43) gebildet werden oder aber die Repräsentativität der Stichprobe, zum Beispiel durch den Einsatz unterschiedlicher Fahrzeuge und Testfahrer, erhöht werden.⁵

6.5.4.2 Applikationen bewerten (D.2)

Ziel der Bewertung einer Applikation ist es, eine gesamtheitliche Aussage über die Güte der Funktion für einen konkreten Parametersatz und damit die Güte der Applikation zu ermitteln. Dafür sind mehrere Prozessschritte notwendig (vgl. Abbildung 6.20).

Identifizierte Fahrabschnitte bewerten Jeder einzelne Fahrabschnitt muss bewertet werden. Die Bewertung kann entweder subjektiv (z.B. im Fahrzeug während der Erprobungsfahrt) oder objektiv durch Metriken und KPI's erfolgen. Die Bewertungskriterien für einen Fahrabschnitt sind ggf. von der zugeordneten Fahraufgabe abhängig. Für die Applikation müssen die Verfügbarkeit (vgl. Definition 32) sowie die Qualität (vgl. Definition 33) der Funktion bewertet werden.

Die geeigneten Metriken sind abhängig von der zu applizierenden Fahrfunktion bzw. den übernommenen Fahraufgaben (vgl. Tabelle 6.3). Weitere Metriken sind in der Literatur zu finden (z.B. [142–144]).

⁵ Weitere Maßnahmen zur Bekämpfung von Bias wurden in *Statistical Consideration of the Representativeness of Open Road Tests for the Validation of Automated Driving Systems* veröffentlicht [JL12].

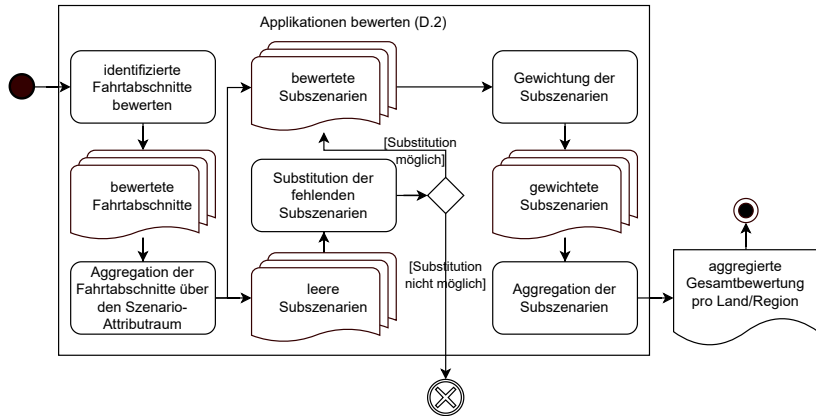


Abbildung 6.20: Prozessschritte des Teilprozesses Applikationen bewerten (vgl. Abbildung 6.16).

Auffahren	Folgefahrt	Spurmittenführung	Spurwechsel
Maximaler Ruck	Varianz des Abstands	Mittlerer Abstand zur Spurmitte	Zeit von Anforderung bis Umsetzung
Minimale Zeitlücke	Mittlere Abweichung von der Soll-Zeitlücke	Schwingungsverhalten	Abstände nach vorne und hinten

Tabelle 6.3: Überblick über ausgewählte Metriken zur Bewertung einzelner Fahraufgaben.

Aggregation der Fahrabschnitte über den Szenario-Attributraum Nachdem die einzelnen Fahrabschnitte bewertet wurden, müssen alle Fahrabschnitte eines Szenarios aggregiert und daraus eine Gesamtaussage für das Szenario abgeleitet werden. Durch den potenziellen Bias einer Stichprobe ist dies jedoch nicht direkt möglich. Zunächst müssen daher der Szenario-Attributraum aufgespannt und die Fahrabschnitte ihren korrespondierenden Subszenarien zugeordnet werden. Für alle Subszenarien wird über Mittelwertbildung (vgl. Gleichung 6.10) oder andere Metriken (vgl. Gleichungen 6.11 und 6.12) eine Subszenariobewertung berechnet.

Substitution der fehlenden Subszenarien Subszenarien ohne zugeordnete Fahrabschnitte können nicht direkt bewertet werden und müssen, falls möglich, substituiert werden. Dazu gibt es verschiedene Verfahren, deren Eignung von der konkreten Fahraufgabe und Applikation abhängt (vgl. Abschnitt 6.4.4).

Gewichtung der Subszenarien Die Repräsentativität der Applikation wird sichergestellt, indem die realen Häufigkeiten der Subszenarien berücksichtigt werden (vgl. Anforderung 3). Die Bildung des Mittelwertes über alle Fahrabschnitte eines Szenarios verletzt die Anforderung 3 (vgl. Voruntersuchung in Abschnitt 5.1.2). Daher kommt die Methodik zur repräsentativen Bewertung von Straßenverkehrstests über die Betrachtung der Häufigkeiten im Szenario-Attributraum zum Einsatz (vgl. Abschnitt 6.4). So werden die realen Häufigkeiten der Subszenarien betrachtet, ungeachtet eines potenziellen Bias in der der Applikation zugrundeliegenden Stichprobe.

Aggregation der Subszenarien Anschließend wird die Bewertung des Szenarios aus den gewichteten Subszenariobewertungen berechnet. Die Szenarien müssen abschließend zu einer Gesamtbewertung der Funktion aggregiert werden.

6.5.4.3 Applikationen vergleichen (D.3)

Für jeden Parametersatz liegt nun ein bewerteter Datensatz und eine aggregierte, länder- bzw. regionsspezifische Gesamtbewertung der Funktionsqualität vor. Neben dieser qualitativen Bewertung gibt es zudem eine Bewertung der Repräsentativität und des Umfangs des Datensatzes in Form des Stichprobenfehlers für jedes Land bzw. jede Region. Um die optimale Applikation zu ermitteln, sind mehrere Prozessschritte notwendig (vgl. Abbildung 6.21).

Applikations-Vergleichs-Tabelle erstellen Für den systematischen Vergleich der Applikationen muss zunächst eine geeignete Struktur geschaffen werden (vgl. Tabelle 6.4). Die Applikations-Vergleichs-Tabelle muss entsprechend für jede

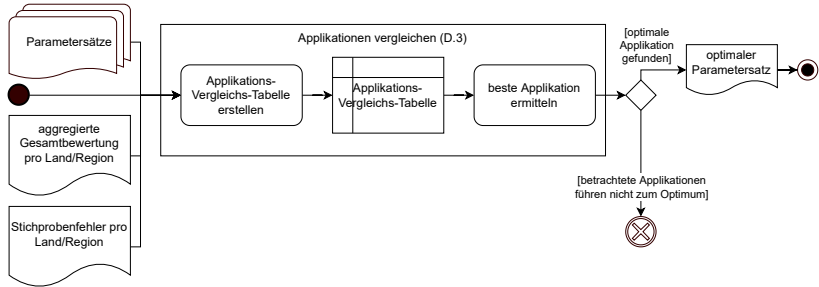


Abbildung 6.21: Prozessschritte und Artefakte des Teilprozesses Applikationen vergleichen (vgl. Abbildung 6.16).

Parametersatz-ID	A_{C1}	R_{C1}	$\sigma_{\bar{x}_{C1}}$	A_{C2}	R_{C2}	$\sigma_{\bar{x}_{C2}}$...	A_{CQ}	R_{CQ}	$\sigma_{\bar{x}_{CQ}}$	A	R	$\sigma_{\bar{x}}$
P_01													
P_02													
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
P_R													

Tabelle 6.4: Schematische Darstellung einer Applikations-Vergleichs-Tabelle.

Applikation die Parametersatz-ID sowie jeweils eine Spalte für die Funktionsverfügbarkeit A , die Funktionsqualität R und den Stichprobenfehler $\sigma_{\bar{x}}$ je Land bzw. Region C aufweisen. Zudem kann je eine aggregierte Spalte über alle Länder bzw. Regionen erstellt werden.

Beste Applikation ermitteln Um Anforderung 2 zu erfüllen, müssen die Verfügbarkeit und Qualität der Fahrfunktion maximiert werden. Die Gewichtung von Verfügbarkeit und Qualität ist abhängig von der Funktion. Je nachdem, ob je Land bzw. Region ein unabhängiger Parametersatz verwendet wird, wird die aggregierte Spalte über alle Länder bzw. Regionen oder die jeweiligen Landes- bzw. Regionsspalten verwenden, um die beste Verfügbarkeit und Qualität zu ermitteln.

Um zusätzlich Anforderung 5 zu erfüllen, darf sich die Funktionsqualität zwischen den Ländern bzw. Regionen nicht unterscheiden. Ein akzeptables Toleranzmaß

muss funktions- und projektabhängig festgelegt werden, beispielsweise als eine maximal tolerierte Standardabweichung. Zudem kann ein absolutes Mindestmaß für die Funktionsqualität definiert werden. Dazu müssen Anforderung 2 und Anforderung 5 berücksichtigt werden.

6.5.5 Parametersätze festlegen oder anpassen (A.2)

Die initialen Parametersätze werden durch Expertenwissen unter Einbindung der Erfahrungen aus der Entwicklung festgelegt. Optional werden weitere Testfahrten zu Festlegung der initialen Parametersätze durchgeführt. Die Anzahl und die Stärke der Variationen der Parametersätze hängt dabei vom geplanten Erprobungsumfang ab. Alternativ können simulative Ansätze Parametersatzvorschläge ermitteln.

Für den Fall, dass durch die Analyse und Auswertung der vorliegenden Applikationsfahrten noch keine optimale Applikation gefunden wurde, müssen die Parametersätze angepasst und neue Erprobungsfahrten durchgeführt werden.

7 Prototypische Umsetzung am Beispiel einer prädiktiven Längsregelung

Zur Evaluation der global-optimalen, datengetriebenen Applikation längs- und querführender Fahrfunktionen wurde eine prototypische Umsetzung am Beispiel einer prädiktiven Längsregelung durchgeführt.

7.1 Use Case

Eine Fahraufgabe prädiktiver, längsregelnder Fahrfunktionen ist die prädiktive Anpassung der Ego-Geschwindigkeit an vorausliegende Kurven in Abhängigkeit der Kurvenkrümmung. Die optimale Geschwindigkeit für jede Kurve hängt dabei von mehreren Attributen der Kurve ab. Die wichtigsten Attribute sind die Krümmung der Kurve und die Steigung, die aus den elektronischen Kartendaten sowie der Kamera ermittelt werden.

7.2 Datengrundlage

Für die Evaluation der global-optimalen, datengetriebenen Applikation wurden mehrere Versuche durchgeführt. Dazu wurden Erprobungen aus vier Ländern ausgewählt (vgl. Tabelle 7.1). Jede Erprobung wurde dabei in zwei disjunkte Datensätze aufgeteilt. Ein Trainingsdatensatz für die Durchführung der virtuellen

Kennzahl	Deutschland	Irland	Italien	Spanien
Länge der gesamten Erprobungsfahrt	7.256,01 km	2.127,88 km	3.048,35 km	4.327,55 km
Länge des Trainingsdatensatzes	2.852,26 km	942,13 km	1.150,03 km	1.474,63 km
Länge des Validierungsdatensatzes	4.403,75 km	1.185,75 km	1.898,32 km	2.852,92 km
Anzahl Fahrten gesamt	144	58	54	46
Anzahl Fahrten im Trainingsdatensatz	30	10	10	10
Anzahl Fahrten im Validierungsdatensatz	114	48	44	36
Dauer der Erprobungsfahrt	95,83 h	32,27 h	40,91 h	52,78 h

Tabelle 7.1: Übersicht über die vier verwendeten Erprobungen.

Applikation sowie ein Validierungsdatensatz zur Evaluation der Ergebnisse des global-optimalen Vorgehens gegenüber anderen Applikationsstrategien.

Die Erprobungsfahrten liegen als aufgezeichnete Messdaten in einem proprietären Format vor. Für die Analyse wurden die relevanten Informationen als .parquet- und .xml-Dateien exportiert.

7.2.1 Trainings- und Validierungsdatensätze

Für jedes Land wurden die 20 bzw. 60 längsten Fahrten zufällig jeweils hälftig in Trainings- und Validierungsdatensätze aufgeteilt. Alle weiteren Fahrten jedes Landes wurden den Validierungsdatensätzen hinzugefügt. Ziel war es eine gleichbleibende Menge an Fahrten für die Applikation für die Länder Irland, Italien

und Spanien zur Verfügung zu stellen. Da für Deutschland ein größerer Datenpool vorhanden ist, wird die Größe des Deutschland-Trainingsdatensatzes auf 30 Fahrten erhöht. Die Vorselektion nach der Länge der Fahrten beeinflusst die Zufälligkeit der Selektion, liegt aber darin begründet, dass die Wahrscheinlichkeit von Fahrthaltungen mit hoher Varianz mit der Fahrtenlänge steigt. Kurze Fahrten sind insbesondere in Deutschland z.B. oft in der Nähe des Entwicklungsstandortes und begünstigen somit potentiellen Bias.

7.3 Analyse- und Bewertungsmethodik

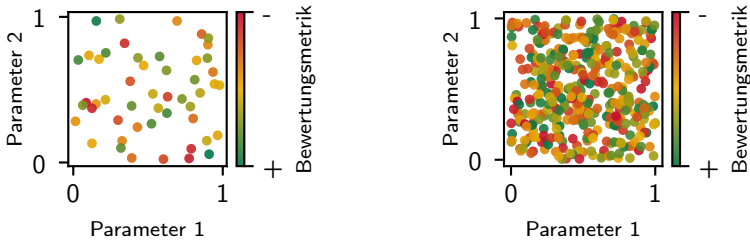
7.3.1 Abdeckung des Parameterraums

Die Funktionsparameter spannen mit ihren Wertebereichen einen mehrdimensionalen Parameterraum auf. Die Abdeckung innerhalb dieses Parameterraums gibt Aufschluss darüber, wie vollständig bzw. umfangreich die Suche nach dem optimalen Parametersatz ist. Zudem lassen sich bei stetigen Funktionen Bereiche identifizieren, in denen wahrscheinlich bzw. unwahrscheinlich optimale Lösungen liegen.

Die Anzahl der Parameter gibt die Dimensionen im Parameterraum vor. Für die Bewertung und zur Visualisierung werden die Parameter paarweise betrachtet. Dabei wird der erste Parameter auf der X-Achse und der zweite Parameter auf der Y-Achse dargestellt. Über die farbliche Kodierung kann zudem die Bewertung der jeweiligen Parameterkombination dargestellt werden (vgl. Abbildung 7.1). Jeder Punkt stellt eine Applikation dar, für die ein aggregiertes Gesamtergebnis berechnet wurde.

7.3.2 Bewertung der Fahraufgabe Kurvendurchfahrt

Für die prototypische Umsetzung wurde der Use Case der prädiktiven Anpassung der Ego-Geschwindigkeit für Kurven gewählt (vgl. Abschnitt 7.1). Zur Bewertung



(a) Geringe Abdeckung mit 50 Datenpunkten

(b) Dichte Abdeckung mit 400 Datenpunkten.

Abbildung 7.1: Visualisierung der paarweisen Parameterraumabdeckung.

einer einzelnen Kurvendurchfahrt werden aufgezeichnete Messsignale verwendet. Die prädizierte Geschwindigkeit für eine Kurve gibt an, mit welcher Geschwindigkeit das prädiktive ACC-System den Kurvenscheitelpunkt durchfahren hätte. Die tatsächliche Geschwindigkeit im Kurvenscheitelpunkt stellt die Referenz dar, mit welcher der Testfahrer die Kurve durchfahren ist. Die Abweichung der beiden Geschwindigkeitswerte stellt für die prototypische Umsetzung die Bewertungsmetrik dar. Die vom Testfahrer gefahrene Geschwindigkeit wird als optimale Referenz verwendet.

$$r(F) = \Delta v = |v_{\text{prädiziert}} - v_{\text{gefahren}}| \quad (7.1)$$

Aus den Messdaten lassen sich weitere Attribute der Kurven extrahieren. Zu jeder Kurve werden die Steigung und die Krümmung extrahiert, um eine Differenzierung der Kurven zu ermöglichen.

7.3.3 Aggregation der Fahrtabschnitte im Szenario-Attributraum / Analyse einer einzelnen Applikation

Die Attribute eines Szenarios spannen einen Attributraum auf, in dem jede Instanz des Szenarios eingeordnet werden kann. Die Attribute haben einen festgelegten Wertebereich. Für jedes Attribut lässt sich der Wertebereich beliebig unterteilen. Dabei kann das Binning durch eine feste Schrittgröße bzw. gewünschte Anzahl an Bins oder einen festgelegten Anteil an Datenpunkten pro Bin berechnet werden. Durch das Binning der Attribute entsteht ein mehrdimensionaler Szenario-Attributraum mit einer definierten Anzahl an Subszenarien (vgl. Abbildung 7.2).

Den entstehenden Subszenarien wird eine ID bzw. Subszenarionummer zugewiesen (vgl. Abbildung 7.3).

Für jedes Subszenario wird ein Qualitätswert (vgl. Abbildung 7.4 obere Zahl in jedem Subszenario) durch Mittelwertbildung der Bewertungen aller Fahrtabschnitte (vgl. Abbildung 7.4 untere Zahl in jedem Subszenario) des Subszenarios berechnet.

Der gesamtheitliche Qualitätswert einer Applikation bezüglich eines Szenarios berechnet sich aus den Qualitätswerten der Subszenarien (vgl. Abbildung 7.5). Die Farbkodierung stellt die Qualitätswerte der Subszenarien grafisch dar. Die Subszenarien werden anschließend miteinander verrechnet. Dazu werden unterschiedliche Verfahren eingesetzt (vgl. Abschnitt 6.4.2).

7.4 Ablauf der virtuellen Applikation

Für die Anwendung des vorgeschlagenen Vorgehens zur global-optimalen, datengetriebenen Applikation bedarf es eines Softwareframeworks für die virtuelle

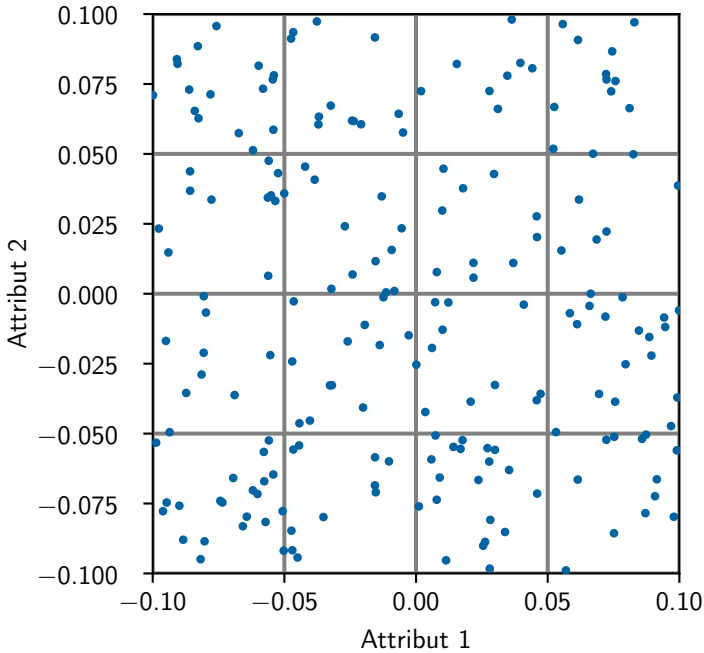


Abbildung 7.2: Visualisierung von Fahrtabschnitten (blau) im Szenario-Attributraum zweier Attribute. Jeder Fahrtabschnitt (Datenpunkt) stellt eine Instanz des Szenarios mit der auf der X- und Y-Achse abzulesenden Attributkombination dar.

Analyse (vgl. Abbildung 6.11) der durchgeführten Applikationsfahrten im Backend. Die notwendigen Prozessschritte, angefangen bei der Analyse der Applikationen über die Bewertung bis hin zum Vergleich der Applikationen, müssen durch geeignete Methoden und Tools abgebildet werden.

Im Rahmen der prototypischen Umsetzung wurde auf aufgezeichnete, vergangene Erprobungsfahrten zurückgegriffen. Zudem stand eine closed-loop Simulationsumgebung (vgl. dazu Bach [58]) zur Verfügung, in der die aufgezeichneten Straßenverkehrstests als Eingangsdaten verwendet und auch unter veränderten

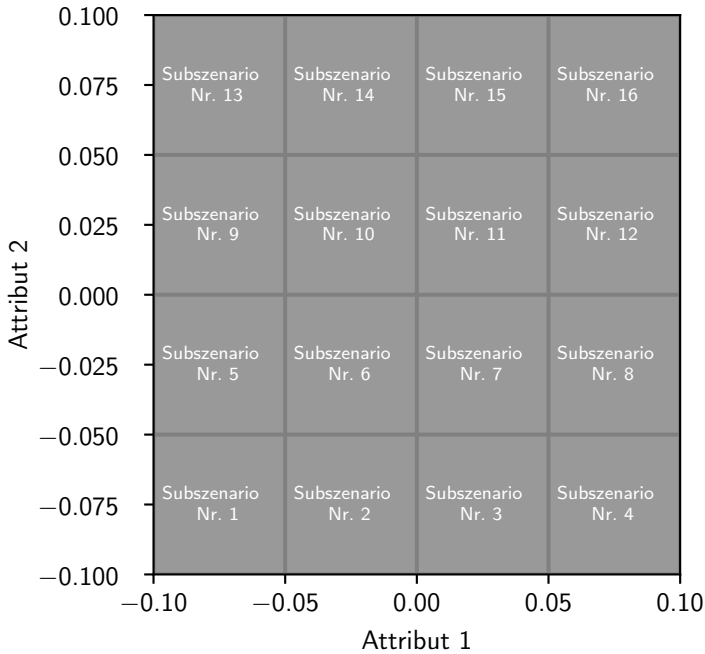


Abbildung 7.3: Visualisierung der Aufteilung des Szenario-Attributraums in Subszenarien. Die Subszenarien werden nummeriert. Die Anzahl der Subszenarien ergibt sich aus der Multiplikation der Anzahl der beiden Attribut-Bins.

Parametern simuliert werden konnten. Da wenig Systemwissen und keine direkten Erfahrungen aus den Erprobungen vorlagen, dafür aber eine erprobte, valide closed-loop Simulation, wurde ein genetischer Optimierungsalgorithmus eingesetzt, um verschiedene Applikationsstrategien umzusetzen. Der Optimierungsalgorithmus wurde so implementiert, dass er ohne Systemwissen heuristisch über mehrere Generationen Lösungen mit unterschiedlichen Parameterkombinationen

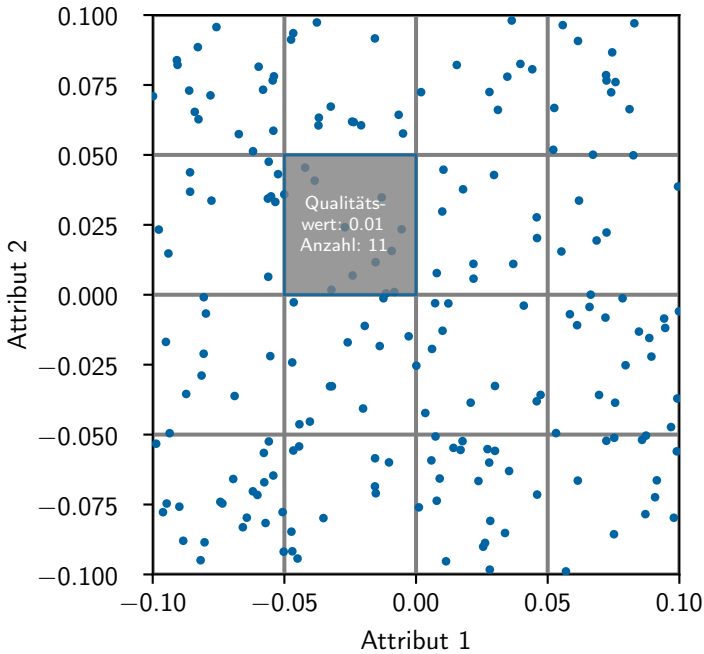


Abbildung 7.4: Visualisierung der Bewertung eines Subszenarios durch die Berechnung des Qualitätswerts aller Fahrtabschnitte des Subszenarios. Der Qualitätswert wird über der Anzahl der Fahrtabschnitte eines Subszenarios dargestellt.

im Parameterraum erzeugen kann¹. Die Parameter werden über eine Konfigurationsdatei gesetzt. Diese wurden unter Verwendung der ausgewählten Datensätze automatisch in der closed-loop Simulation ausgeführt und standen anschließend für die virtuelle Analyse zur Verfügung.

Einzelne Prozessschritte der Prozessphasen Vorbereitung (A), Durchführung (B) und Datenaufzeichnung (C) der global-optimalen, datengetriebenen Applikation

¹ Veröffentlicht in *A Process Reference Model for the Virtual Application of Predictive Control Features* [JL10]

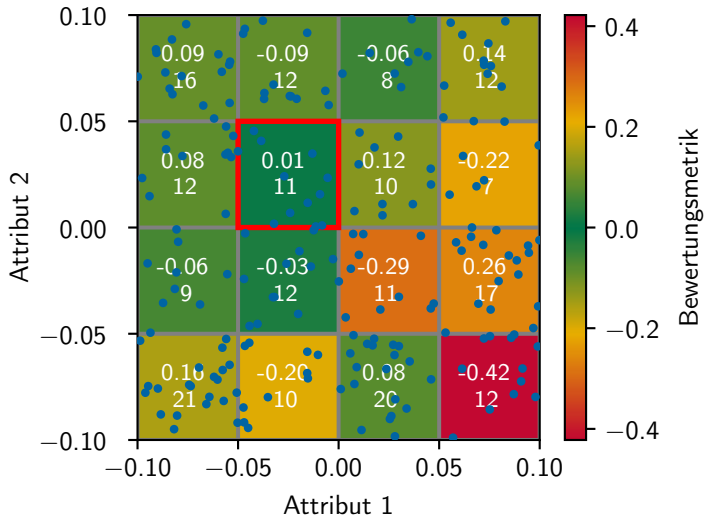


Abbildung 7.5: Vollständige Visualisierung der Bewertung aller Subszenarios über den Mittelwert aller Fahrtabschnitte (blau) jedes Subszenarios. Für jedes SubszENARIO wird die Bewertung (obere Zahl + Farbkodierung) sowie die Anzahl Instanzen (untere Zahl) dargestellt.

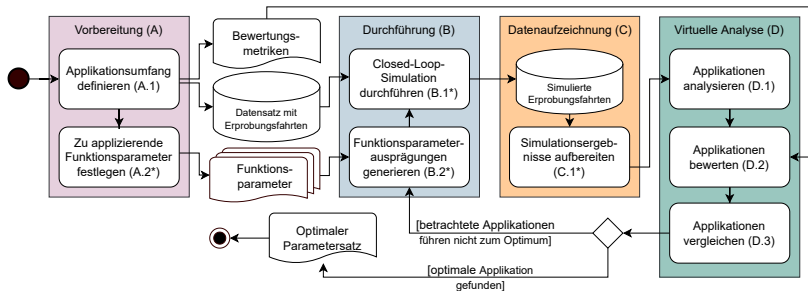


Abbildung 7.6: Adaption des global-optimalen, datengetriebenen Applikationsprozesses auf eine virtuelle, iterative Applikation mittels aufgezeichneter Erprobungsfahrten.

wurden an die virtuelle Applikation angepasst (vgl. Abbildung 7.6 und Abbildung

6.11). Die Anpassung der Funktionsparameter wird in diesem Fall nicht durch Experten während der Applikationsvorbereitung durchgeführt. Stattdessen werden lediglich die relevanten Funktionsparameter und deren Wertebereiche festgelegt (A.2*). Während der iterativen Ausführung werden konkrete Ausprägungen der Funktionsparameter vom Optimierungsalgorithmus generiert (B.2*). Die Durchführung selbst umfasst nun keine Erprobungsfahrten mehr. Stattdessen wird ein Datensatz aufgezeichneter Erprobungsfahrten als Grundlage für eine closed-loop Simulation (B.1*) verwendet. Die simulierten Erprobungsfahrten müssen konvertiert und aufbereitet werden (C.1*). Die virtuelle Analyse bleibt unverändert.

Durch die simulative Ausführung, die lediglich Rechenressourcen anstelle von Personalaufwänden für die Durchführung der Erprobungsfahrten sowie Kosten und Zeit für den Aufbau von Prototypenfahrzeugen benötigt, ist ein stark iterativer Ansatz möglich. Die Optimierungsschleife wird solange durchlaufen, bis ein definiertes Abbruchkriterium (z.B. zeitlich, Abdeckung des Parameterraums oder Schwellwert der Bewertungsmetrik) erreicht wird. Anschließend liegen ein oder mehrere als optimal bewertete Parametersätze vor.

8 Evaluation und Diskussion der prototypischen Umsetzung

Zur Evaluation der global-optimalen, datengetriebenen Applikation längs- und querführender Fahrfunktionen wurden die in Abschnitt 7.2 vorgestellten Applikationsfahrten verwendet. Die Evaluation ist mehrstufig und betrachtet die Kernbestandteile des Konzepts separat.

In Abschnitt 8.1 wird evaluiert, inwieweit mit der global-optimalen, datengetriebenen Applikation bessere, d.h. global-optimale, Ergebnisse erreicht werden.

In Abschnitt 8.2 wird evaluiert, inwieweit die Repräsentativität der verwendeten Straßenverkehrstests quantifiziert und damit eine Aussage über die Grundgesamtheit getroffen werden kann. Dabei wird insbesondere die Stichprobengröße als Quantifizierungsmerkmal untersucht sowie die Gewichtung der Subsznarien evaluiert (vgl. Abschnitt 6.4.3).

Im dritten und letzten Teil der Evaluation in Abschnitt 8.3 wird das Gesamtverfahren betrachtet. Die Ergebnisse eines Applikationsdurchlaufs werden analysiert und die Verwendung im Applikations- bzw. Entwicklungsprozess aufgezeigt.

8.1 Evaluation der Unterteilung des Szenario-Attributraums in Subszenarien und deren gewichtete Aggregation

Im ersten Teil der Evaluation wird der Einfluss der Unterteilung des Szenario-Attributraums in Subszenarien und die anschließende gewichtete Aggregation über die Subszenarien auf die Repräsentativität der Applikation untersucht (vgl. Abschnitte 6.3.3 und 6.4). Der Applikationsprozess wird mehrfach mit unterschiedlichen Aggregationsverfahren für die Aggregation der Subszenarien durchlaufen, um den Mehrwert des global-optimalen Ansatzes zu quantifizieren (vgl. Abschnitt 6.4.5):

Als **Referenzapplikation** wird der *globale Mittelwert* verwendet, der über alle Subszenarien gebildet wird. Dies entspricht dem bisherigen Vorgehen ohne Verwendung der Subszenarien und dient als Referenz für die Evaluation.

In **Applikation SM** wird der *Subszenario-Mittelwert* verwendet, bei dem zunächst die Mittelwerte in den Subszenarien ermittelt und diese dann mit dem Mittelwert über alle Subszenarien verrechnet werden. In diesem Fall wird der Szenario-Attributraum mit allen Subszenarien verwendet.

In **Applikation gSM** wird der *gewichtete Subszenario-Mittelwert* verwendet, bei dem ebenfalls Mittelwerte für die Subszenarien berechnet, diese dann jedoch gewichtet aggregiert werden. Die Gewichte werden dabei aus der Grundgesamtheit, in diesem Fall dem gesamten verfügbaren Datensatz, extrahiert. Jedes Subszenario wird entsprechend seiner Häufigkeit in der Grundgesamtheit in der Bewertung berücksichtigt.

Die Applikation gSM entspricht dem vorgeschlagenen Vorgehen in dieser Dissertation.

Zur Durchführung der drei Applikationen wurde der Spanien-Trainingsdatensatz ausgewählt (vgl. Abschnitt 7.2). Für jede Applikation wurden insgesamt 1300

Parameterkombinationen über 26 Generationen mit jeweils 50 Lösungen erzeugt und auf den Trainingsdaten simuliert.

Die Exploration des Parameterraums muss untersucht werden. Für die Applikation wurden zwei Funktionsparameter verwendet: Ein Parameter für die Gewichtung der Kurvenkrümmung am Kurvenanfang und ein Parameter für die Gewichtung der Kurvenkrümmung am Kurvenende. Beide Parameter spannen einen 2-dimensionalen Raum auf. Für alle drei Applikationen wurde der Parameterraum exploriert mit lokalen Häufungen rund um die besten Lösungen (vgl. Abbildung 8.1). Durch die unterschiedliche Berechnung der Metrik sind die absoluten Metrikwerte über die drei Applikationen nicht vergleichbar und wurden deshalb abstrahiert. Eine Evaluation der jeweils besten Lösung erfolgt separat.

Im Verlauf der Applikation hat sich die jeweils beste Lösung wie in Abbildung 8.2 entwickelt. Aufgrund der unterschiedlichen Aggregationsmethoden sind die absoluten Zahlenwerte nicht vergleichbar. Zu beobachten ist eine sprunghafte Verbesserung über die Generationen hinweg, was an den Wirkprinzipien des gewählten Optimierungsalgorithmus sowie der nicht stetigen Gütefunktion liegt. In der Referenzapplikation wurde keine Verbesserung erzielt. Dies kann zum einen Zufall sein. Zum anderen ist nicht auszuschließen, dass die initialen Parameterausprägungen bereits sehr gut waren und deshalb keine bessere Lösung gefunden werden konnte, da die Aggregation der Kurvenabweichungen in der Referenzapplikation in etwa der Durchführung der Applikation in der realen Welt entspricht und der verwendete initiale Parametersatz das Resultat einer langjährigen Entwicklung mit vielen Test- und Erprobungsfahrten ist.

Die Ergebnisse der jeweils besten Lösung der drei Applikationen im Attributraum der Kurvenkrümmung und Steigung sind in Abbildung 8.3 dargestellt. Zusätzlich ist in Abbildung 8.3a die Ausgangsparametrierung visualisiert. Für den Vergleich der Parametrierungen wurde der Spanien-Validierungsdatensatz (vgl. Tabelle 7.1) verwendet. In diesem Datensatz sind 36 Straßenverkehrstests in Spanien enthalten, die nicht Teil des Trainingsdatensatzes waren, um eine unabhängige Validierung zu gewährleisten.

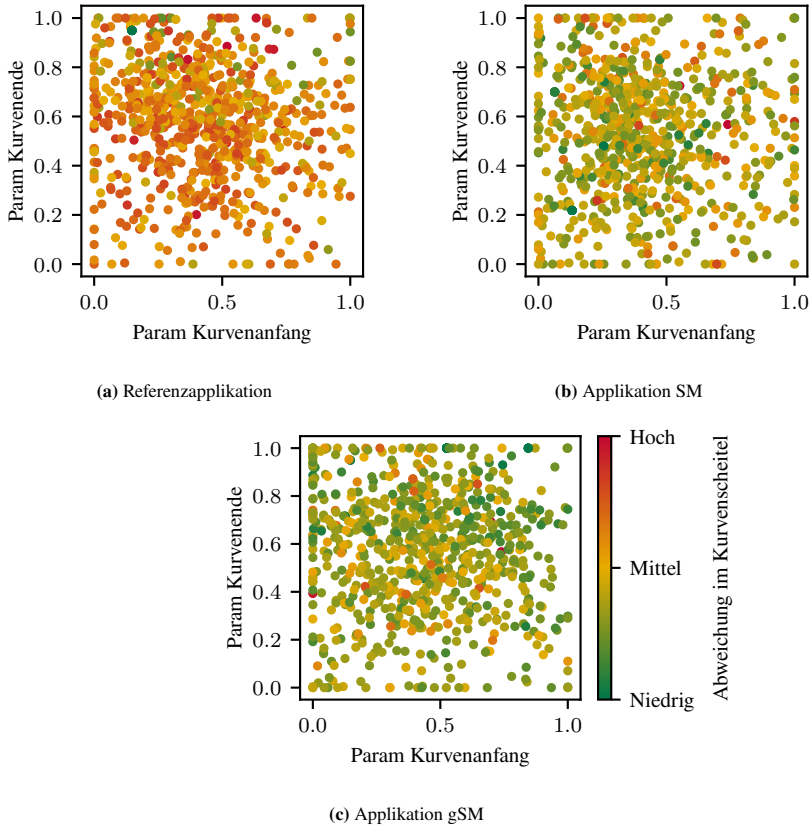


Abbildung 8.1: Vergleich der Exploration des Parameterraums in den drei Applikationsdurchläufen mit dem Spanien-Trainingsdatensatz (vgl. Tabelle 7.1)

Zu beobachten ist, dass in allen vier Fällen die fünf Subszenarien 5, 8, 12, 13 und 16 nicht durch den Spanien-Validierungsdatensatz abgedeckt wurden. In den anderen Subszenarien am Rand ist jeweils eine einstellige Anzahl Fahrtabschnitte vorhanden. Die vier mittleren Subszenarien sind mit 10 bis 33 Fahrtabschnitten vertreten. Dabei variiert die Anzahl der Fahrtabschnitte zwischen den vier Lösungen um bis zu 11 Fahrtabschnitte, was durch die unterschiedlichen Parameterausprägungen und deren Auswirkung auf die Fahrfunktion zu erklären ist.

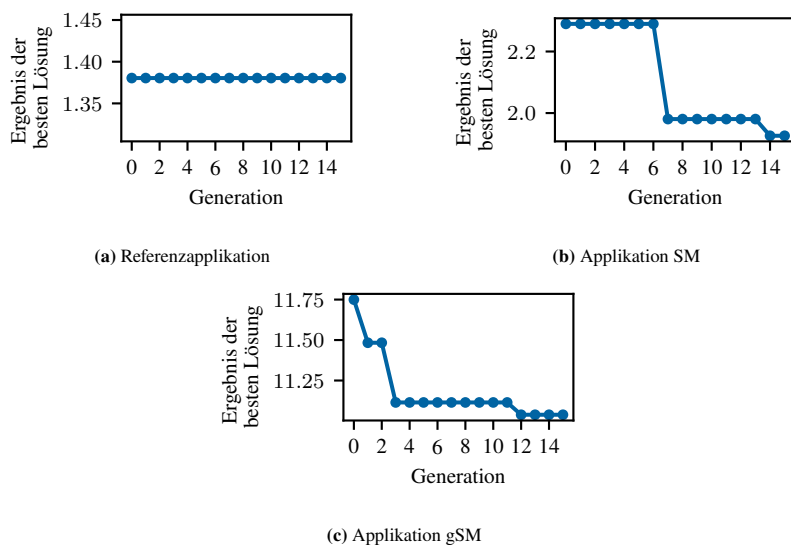


Abbildung 8.2: Entwicklung der besten Lösung über die Generationen in den drei Applikationsdurchläufen mit dem Spanien-Trainingsdatensatz (vgl. Tabelle 7.1)

Aufgrund der geringen Abdeckung des Attributraums mit dem Spanien-Validierungsdatensatz wurde zudem eine Validierung auf dem globalen Validierungsdatensatz durchgeführt. Abbildung 8.4 zeigt dieselbe Darstellung wie für den Spanien-Validierungsdatensatz (vgl. Abbildung 8.3) nun für den globalen Validierungsdatensatz. Alle Subsznarien sind durch mehr Fahrabschnitte repräsentiert und alle Randbereiche sind in den Daten vorhanden.

Visuell sticht sowohl beim Spanien-Validierungsdatensatz (vgl. Abbildung 8.3) als auch beim globalen Validierungsdatensatz (vgl. Abbildung 8.4) keine Lösung als die beste hervor. In Tabelle 8.1 wurden die vier Lösungen daher mit jeweils unterschiedlichen Aggregationsverfahren bewertet. Die jeweils beste Lösung der Applikationsstrategien Referenzapplikation, Applikation SM und Applikation gSM sowie die Ausgangsparametrierung wurden auf beiden Datensätzen simuliert. Dabei wurden jeweils die drei Qualitäts-Metriken globaler Mittelwert,

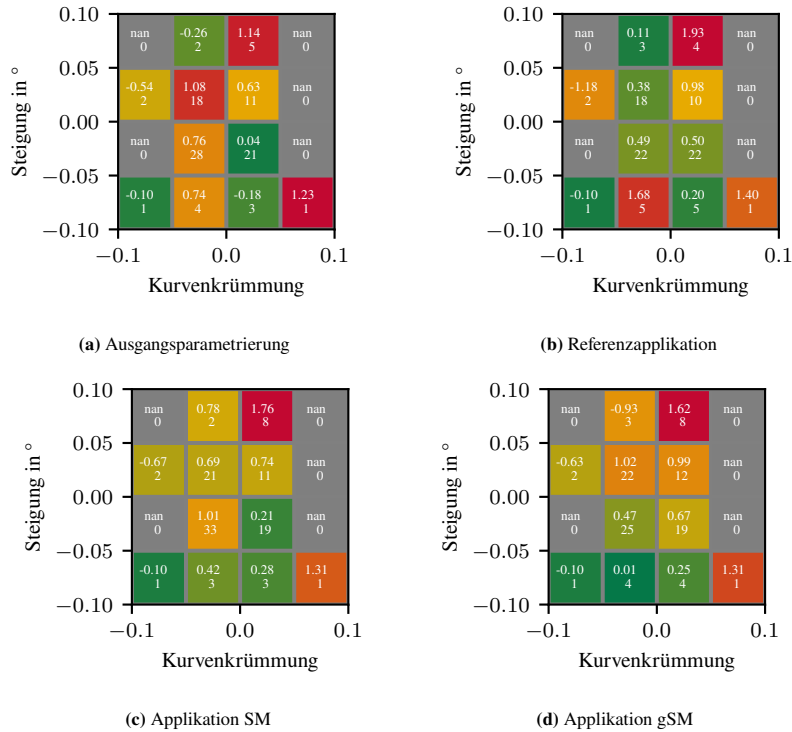


Abbildung 8.3: Mittlere Abweichung im Attributraum der Ausgangsparametrierung sowie der jeweils besten Lösung der drei Applikationsstrategien mit dem Spanien-Validierungsdatensatz (vgl. Tabelle 7.1)

Subszenario-Mittelwert sowie gewichteter Subszenario-Mittelwert für die Validierungsdatensätze berechnet. Während für den Spanien-Validierungsdatensatz die Referenzapplikation in zwei von drei Fällen sowie die Ausgangsparametrierung bei Betrachtung der Subszenario-Mittelwerte die beste Parametrierung hervorgebracht haben, **liefert bei globaler Betrachtung die Applikation gSM, also der gewichtete Subszenario-Mittelwert, das beste Ergebnis**. Sowohl bei Evaluation mit globalem Mittelwert, also dem Mittelwert über alle Kurven im globalen Validierungsdatensatz, sowie dem gewichteten Subszenario-Mittelwert, also der Gewichtung mit der Grundgesamtheit, wurde über die Applikationsstrategie mit

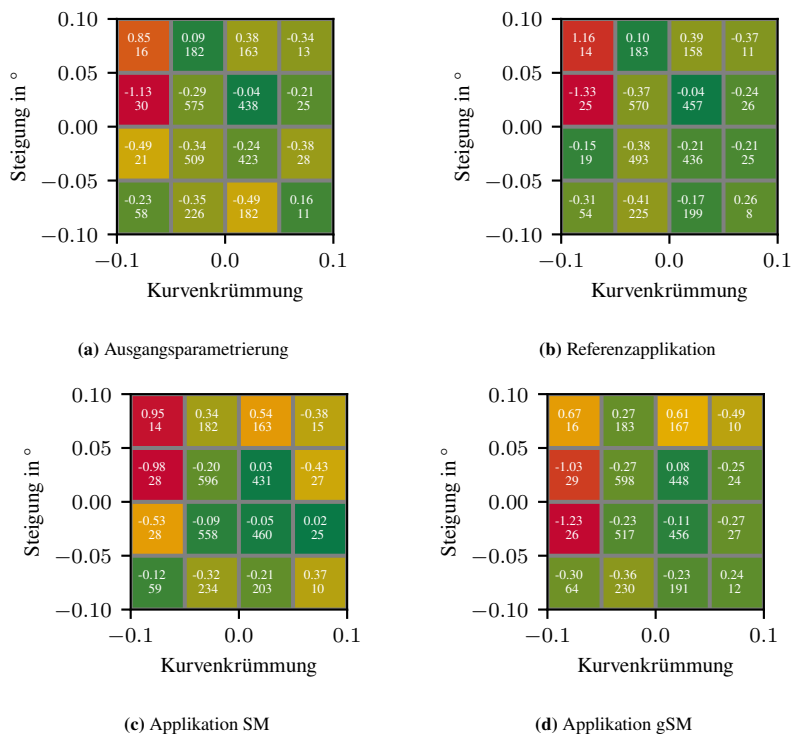


Abbildung 8.4: Mittlere Abweichung im Attributraum der Ausgangsparametrierung sowie der jeweils besten Lösung der drei Applikationsstrategien mit dem globalen Validierungsdatensatz.

gewichteten SubszENARIO-Mittelwerten die beste Parametrierung ermittelt. Lediglich bei der Evaluation bezüglich der SubszENARIO-Mittelwerte führt die Referenzapplikation zu einem besseren Ergebnis.

Grund für die unterschiedliche Güte der Applikationen mittels gewichtetem SubszENARIO-Mittelwert kann sein, dass die Gewichtung aus einem globalen Datensatz entnommen wurde und die Kurven in Spanien dieser Verteilung nicht entsprechen. Zudem können durch den geringeren Umfang des Spanien-Validierungsdatensatzes Kurven nicht repräsentativ verteilt sein. Auffällig ist

zudem, dass die mittleren Abweichungen in Spanien stets positiv, also über der Prädiktion, und im globalen Validierungsdatensatz stets negativ, also unter der Prädiktion, liegen. Dies bestärkt die Aussage, dass die Gesamtheit der Kurven in Spanien nicht normalverteilt ist. Diese Faktoren führen dazu, dass in diesem Fall die Optimierung mit der einfachen Mittelwertbildung ein besseres Ergebnis liefert. Im globalen Validierungsdatensatz sind diese Effekte geringer bzw. nicht mehr vorhanden. **Es zeigt sich, dass bei globaler Betrachtung die Gewichtung der Subsznenarien zu besseren Ergebnissen führt.**

		Validierungsmetrik		
Applikationsstrategie		Globaler Mittelwert	SubszENARIO-Gewichteter Mittelwert	SubszENARIO-Mittelwert
Spanien Validierungs- datensatz	Ausgangsparametrierung	0.23663	0.23824	0.02119
	Referenzapplikation	0.18186	0.29942	0.01907
	Applikation SM	0.48113	0.24125	0.02613
	Applikation gSM	0.38536	0.31824	0.02828
Globaler Validierungs- datensatz	Ausgangsparametrierung	-1.87835	-1.19803	-0.11886
	Referenzapplikation	-1.81546	-0.98817	-0.11450
	Applikation SM	-1.81255	-1.11667	-0.11408
	Applikation gSM	-1.80210	-1.06013	-0.11311

Tabelle 8.1: Ergebnisse der Validierung der jeweils besten Lösung unterschiedlicher Applikationsstrategien auf dem länderspezifischen Spanien-Validierungsdatensatz sowie dem globalen Validierungsdatensatz. Dabei wurden für die Validierung für jede Lösung die Metriken für den globalen Mittelwert, den SubszENARIO-Mittelwert sowie den gewichteten SubszENARIO-Mittelwert berechnet.

8.2 Bewertung der Repräsentativität einer Applikation

8.2.1 Evaluation statistischer Kennzahlen als Indikatoren für die Repräsentativität

Zur Bewertung der Repräsentativität einer Applikation muss der verwendete Datensatz betrachtet werden. Dieser kann als Stichprobe einer Grundgesamtheit interpretiert werden. Für Stichproben lassen sich statistische Kennzahlen, wie der Stichprobenfehler, berechnen, um Rückschlüsse auf die Repräsentativität zu ziehen (vgl. Abschnitt 4.4.3). Für die Berechnung wird die Verteilung der Subszenerarien in der Stichprobe mit der Verteilung der Subszenerarien in der Grundgesamtheit verglichen (vgl. Abschnitt 6.4.3). Für die Evaluation der prototypischen Umsetzung wurde folgender Versuchsaufbau konzipiert:

Die Subszenerario-Verteilung des Szenarios Kurve bzw. der Fahraufgabe Kurvendurchfahrt wird im Attributraum der Krümmung und Steigung betrachtet. Es soll die Verteilung im gesamten Datensatz mit Verteilungen zufällig gezogener kleinerer Stichproben verglichen und mittels statistischer Kennzahlen bewertet werden. Dazu wurde zunächst die Verteilung der Subszenerarien im gesamten Datensatz ermittelt. Ebenfalls wurden zufällig 10x 5 Fahrten, 5x 10 Fahrten, 5x 20 Fahrten, 3x 30 Fahrten und 1x 50 Fahrten aus der Grundgesamtheit gezogen und die Subszenerarioverteilungen in den jeweiligen Datensätzen ermittelt. Mit beiden Verteilungen lassen sich statistische Tests durchführen und eine Aussage über die Ähnlichkeit bzw. Gleichheit der Verteilungen ableiten (vgl. Abschnitt 4.4.3). Die Ergebnisse wurden über alle Datensätze und für die statischen Maße KS-Test (vgl. Gleichung 4.4), Chi-Quadrat-Test (vgl. Gleichung 4.5), KL-Divergenz (vgl. Gleichung 4.6) sowie RMSE (vgl. Gleichung 4.7) berechnet (vgl. Tabelle 8.2). Für den KS-Test und den Chi-Quadrat-Test wurde als Nullhypothese angenommen, dass die jeweilige Stichprobe und die Grundgesamtheit aus derselben Häufigkeitsverteilung stammen.

Datensatz	KS-Statistik	KS-P-Wert	Chi ² -Statistik	Chi ² -P-Wert	KL-Divergenz	RMSE
5 Fahrten #1	0.68750	0.00067	6.52371	0.96955	0.042305	0.018968
5 Fahrten #2	0.56250	0.01120	21.50677	0.12140	0.135323	0.026581
5 Fahrten #3	0.68750	0.00067	12.69212	0.62607	0.581719	0.100219
5 Fahrten #4	0.43750	0.09333	8.33236	0.90977	0.275008	0.043330
5 Fahrten #5	0.75000	0.00012	12.28005	0.65772	0.066352	0.025749
5 Fahrten #6	0.56250	0.01120	10.25326	0.80351	0.220377	0.049931
5 Fahrten #7	0.75000	0.00012	3.27989	0.99930	0.439207	0.068088
5 Fahrten #8	0.50000	0.03500	8.13821	0.91812	0.399441	0.056508
5 Fahrten #9	0.56250	0.01120	32.35287	0.00576	0.824637	0.081329
5 Fahrten #10	0.68750	0.00067	7.44032	0.94426	1.279680	0.140495
10 Fahrten #1	0.43750	0.09333	6.43598	0.97146	0.092119	0.026413
10 Fahrten #2	0.62500	0.00302	27.67481	0.02370	0.121304	0.030372
10 Fahrten #3	0.56250	0.01120	5.42554	0.98787	0.054381	0.021856
10 Fahrten #4	0.68750	0.00067	7.45028	0.94393	0.039257	0.016169
10 Fahrten #5	0.62500	0.00302	14.18040	0.51189	0.088050	0.026542
20 Fahrten #1	0.87500	0.00000	10.51598	0.78609	0.026007	0.015537
20 Fahrten #2	0.68750	0.00067	11.43987	0.72084	0.034787	0.010000
20 Fahrten #3						
20 Fahrten #4	0.62500	0.00302	10.41014	0.79318	0.065612	0.020994
20 Fahrten #5	0.68750	0.00067	13.82520	0.53883	0.072526	0.018528
30 Fahrten #1	0.81250	0.00002	18.42778	0.24084	0.043045	0.018421
30 Fahrten #2	0.75000	0.00012	13.38675	0.57245	0.052602	0.021577
30 Fahrten #3	0.93750	0.00000	9.05611	0.87457	0.024862	0.016738
50 Fahrten #1	0.8125	0.00002	11.887472	0.687526	0.018113	0.011395

Tabelle 8.2: Berechnung verschiedener Statistiken für die Quantifizierung des Stichprobenfehlers in verschiedenen zufällig gezogenen Datensätzen. Für den KS-Test und den Chi²-Test sind die Datensätze mit bestätigter Nullhypothese hervorgehoben. Für die KL-Divergenz sowie den RMSE die jeweils 5 zur Grundgesamtheit ähnlichsten Datensätze.

Für den KS-Test wurde ein Signifikanzniveau von 0.01 festgelegt. Die fettgedruckten Werte in der KS-P-Wert Spalte zeigen die Datensätze, für die die Nullhypothese hält. Auffällig ist, dass dies nur bei Datensätzen mit 5 oder 10 Fahrten der Fall ist (*5 Fahrten #2, #4, #6, #8, #9 und 10 Fahrten #1, #3*). Andere kleine Datensätze (*5 Fahrten #1, #3, #5, #7, #10 und 10 Fahrten #4*) haben dagegen einen sehr kleinen KS-P-Wert, was darauf hindeutet, dass die zur Grundgesamtheit ähnlichen Verteilungen nur durch Zufall getroffen wurden. Wider erwarten zeigt sich jedoch auch bei den größeren Datensätzen nicht, dass die Verteilungen der Grundgesamtheit ähnlich bzw. gleich sind. Dies zeigt sich durch sehr kleine KS-P-Werte und hohe KS-Statistik Werte für *20 Fahrten #1, #2, #5, 30 Fahrten #1, #2, #3 sowie 50 Fahrten #1*.

Beim χ^2 -Test zeigt sich ein ähnliches Bild. Bei einem Signifikanzniveau von 0.9 für den χ^2 -P-Wert konnten ebenfalls nur für Datensätze mit 5 oder 10 Fahrten die Nullhypothese beibehalten werden (*5 Fahrten #1, #4, #7, #8, #10 und 10 Fahrten #1, #3, #4*). Da es jedoch auch hier nicht für alle Datensätze der Fall ist, bestärkt sich die Zufallsannahme. Insbesondere da bei den Datensätzen *5 Fahrten #2, #9 und 10 Fahrten #2* das Signifikanzniveau sehr niedrig ist. Für die Datensätze mit 20, 30 und 50 Fahrten sind die Abweichungen des χ^2 -P-Wertes nach oben und unten weniger ausgeprägt. Bei keinem Datensatz wird das Signifikanzniveau von 0.9 erreicht. Allerdings sinkt es auch nur in einem Fall unter 0.5.

Die KL-Divergenz zeigt ein anderes Ergebnis. Ein niedriger Wert steht für eine hohe Ähnlichkeit des Datensatzes zur Grundgesamtheit. Mit steigender Anzahl an Fahrten sinkt im Mittel die KL-Divergenz (*5 Fahrten: 0.4264; 10 Fahrten: 0.0790; 20 Fahrten: 0.0497; 30 Fahrten: 0.0189; 50 Fahrten 0.0181*) (vgl. Abbildung 8.5). Die Top 5 sind in Tabelle 8.2 hervorgehoben.

Für den RMSE zeigt sich ein noch stärkeres Bild als für die KL-Divergenz. Der mittlere RMSE nimmt über die Anzahl Fahrten ab (*5 Fahrten: 0.0611; 10 Fahrten: 0.0243; 20 Fahrten: 0.0163; 30 Fahrten: 0.0402; 50 Fahrten 0.0114*) (vgl. Abbildung 8.5).

Insgesamt zeigen die statistischen Maße keine eindeutige Tendenz. Für den KS-Test und den χ^2 -Test ist zu schließen, dass für die kleineren Datensätze zufällig

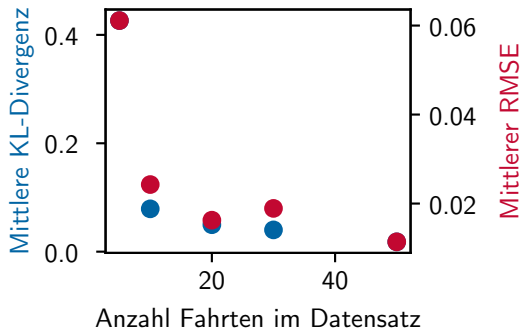


Abbildung 8.5: Mittlere KL-Divergenz (blau) und mittlerer RMSE (rot) über die Anzahl Fahrten in den zufällig gezogenen Datensätzen.

die Verteilung der Grundgesamtheit getroffen wurde. Es lässt sich daraus jedoch ebenfalls nicht schließen, dass die größeren Datensätze zu einer besseren Nachbildung der Stichprobe führen. Für die KL-Divergenz und den RMSE ist eine Tendenz zu erkennen. Größere Datensätze bilden tendenziell die Grundgesamtheit besser ab. **Die vorliegenden Stichprobengrößen sind zu klein, um die Repräsentativität des Datensatzes bezüglich der Grundgesamtheit mittels statistischer Kennzahlen zu differenzieren.**

8.2.2 Evaluation der SubszENARIO-Abdeckung als Indikator für die Repräsentativität

Es bedarf daher eines anderen Indikators für die Repräsentativität eines Datensatzes im Kontext der Applikation. Durch die gewichtete Aggregation der SubszENARIEN spielt das korrekte Verhältnis der SubszENARIEN im Datensatz eine untergeordnete Rolle. **Vielmehr ist die Abdeckung aller SubszENARIEN entscheidend für eine repräsentative Bewertung und damit die Repräsentativität der Applikation.** Daher wurde die Abdeckung aller SubszENARIEN in den jeweiligen Datensätzen als Indikator für die Repräsentativität einer Stichprobe untersucht.

Dazu wurde die relative Abdeckung aller Subszenarien mit mindestens einem Fahrabschnitt sowie mit mindestens 5 Fahrabschnitten berechnet. Die Ergebnisse sind in Anhang A.1.1 Tabelle A.2 ausführlich aufgeführt und in Abbildung 8.6 dargestellt. Es zeigt sich, dass mit zunehmender Anzahl Fahrten die Subszenarioabdeckung zunimmt. Ab 20 bzw. 30 Fahrten sind nahezu alle Subszenarien in den Datensätzen vorhanden. Bei einer Mindestabdeckung von 5 Fahrabschnitten je Subszenario sind immerhin 50% der Subszenarien hinreichend vorhanden.

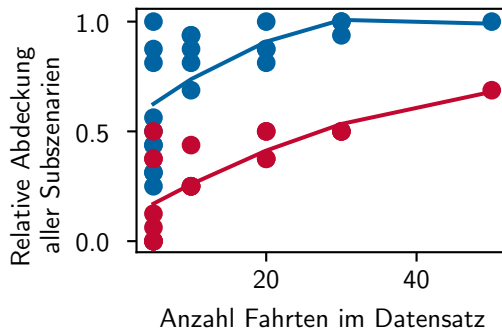


Abbildung 8.6: Relative Abdeckung aller Subszenarien mit >0 (blau) bzw. >5 (rot) Kurven über die Anzahl Fahrten in den zufällig gezogenen Datensätzen.

Die statistischen Kennzahlen zeigen, dass allein die Größe der Datensätze keinen Einfluss auf die Repräsentativität der Stichproben besitzt. Dennoch steigt die Abdeckung der Subszenarien mit der Anzahl Fahrten in der Stichprobe. Darauf aufbauend kann mit Hilfe der in Abschnitt 6.5.4.1 vorgestellten Gewichtung eine repräsentative Bewertung kleiner Stichproben vorgenommen werden, da dadurch nicht mehr die Verteilung der Subszenarien in der Stichprobe entscheiden ist sondern vielmehr das Vorhandensein aller Subszenarien, so dass die Gewichtung vorgenommen werden kann.

Die Abdeckung der Subszenarien in der Stichprobe ist unter Verwendung der gewichteten Bewertung ein valider Indikator für die Repräsentativität der Applikation auf Basis der Stichprobe und bezüglich der Grundgesamtheit.

8.3 Evaluation Methodik und Gesamtprozess

Für die Evaluation der entwickelten Methodik zur global-optimalen, datengetriebenen Applikation soll die Integration in den Entwicklungsprozess betrachtet werden. Die zusätzlichen Prozessschritte, insbesondere die virtuelle Analyse, können nebenläufig zu anderen Entwicklungsaktivitäten stattfinden. Die Ergebnisse der virtuellen Analyse lassen sich anschließend nahtlos in den Applikationsprozess integrieren. Die virtuell ermittelten global-optimalen Parametersätze müssen im Fahrzeug validiert werden. Dabei können Verfügbarkeit und Qualität der Funktion objektiv gegeneinander abgewogen werden. Zudem gibt die SubszENARIOabdeckung eine Möglichkeit zur Abschätzung der Vollständigkeit der Applikation.

Lösungs-ID	Verfügbarkeit	Qualität	Einfache SubszENARIOabdeckung	Fünffache SubszENARIOabdeckung
1	0.380435	-0.058011	62.5%	31.25%
2	0.332130	0.324920	62.5%	31.25%
3	0.354478	0.259917	62.5%	31.25%
4	0.391635	0.564528	62.5%	31.25%
5	0.414449	0.346557	62.5%	37.50%
6	0.375465	0.175903	62.5%	43.75%
7	0.367041	0.406166	62.5%	31.25%
8	0.387597	0.478362	62.5%	31.25%
9	0.386100	0.344748	62.5%	37.50%
10	0.363636	0.477066	62.5%	31.25%

Tabelle 8.3: Auswertungstabelle mit den besten 10 Lösungen aus der Spanien-Applikation. Für jede Applikation wurde die Verfügbarkeit, Qualität sowie die SubszENARIOabdeckung berechnet.

Anhand der Applikation gSM mit dem Spanien-Trainingsdatensatz aus Abschnitt 8.1 soll die Methodik exemplarisch aufgezeigt werden. Dazu werden die 10 Lösungen bzw. Parametersätze mit dem niedrigsten gewichteten SubszENARIO-Mittelwert

ausgewählt und auf dem Spanien-Validierungsdatensatz virtuell simuliert, um Verfügbarkeit, Qualität und die SubszENARIOabdeckung zu ermitteln (vgl. Tabelle 8.3). Aus den Ergebnissen lassen sich Pareto-Darstellungen generieren, die eine manuelle Auswahl der Parametersätze je nach Gewichtung von Qualität und Verfügbarkeit ermöglichen (vgl. Abbildung 8.7).

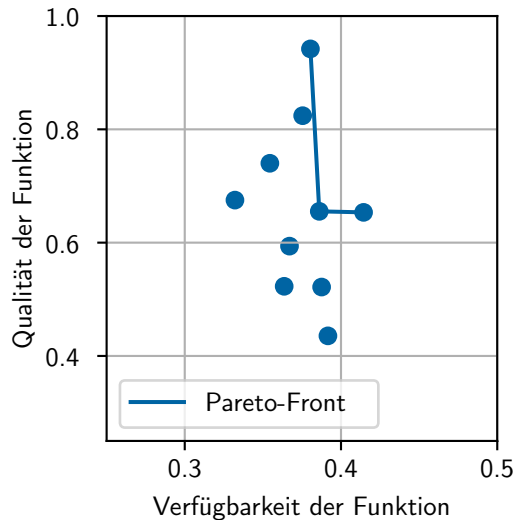


Abbildung 8.7: Darstellung der besten 10 Lösungen als Pareto-Front über Verfügbarkeit und Qualität der Funktion.

Für die global-optimale, datengetriebene Applikation hat sich allgemein gezeigt, dass wesentliche regionale und statistische Unterschiede in den Erprobungsfahrten auftreten. **Eine unstrukturierte und sequentielle Applikation wie bisher kann unter diesen Umständen nicht zu einem global-optimalen Ergebnis führen.** Vielmehr müssen Erprobungsfahrten zur Datensammlung und Validierung der Applikation verwendet werden. Die virtuelle Analyse stellt das Kernstück der Applikation dar, in der die einzelnen Szenarien gesamtheitlich und analytisch bewertet und verglichen werden.

8.4 Anforderungserfüllung

Für die global-optimale, datengetriebene Applikation wurden in Abschnitt 5.4 Anforderungen aufgestellt:

Anforderung 1 fordert die Durchführung der Applikation im realen Straßenverkehr innerhalb der ODD bzw. eine direkte Übertragbarkeit darauf. Durch die Verwendung von aufgezeichneten Straßenverkehrstests sowie der realen Funktionssoftware wird diese Übertragbarkeit im vorgestellten Konzept gewährleistet. Die Anforderung ist erfüllt.

Anforderung 2 fordert die Maximierung der Qualität und Verfügbarkeit der Funktion innerhalb ihrer ODD. Durch die virtuelle Analyse werden Qualität und Verfügbarkeit objektiv ermittelt. Die Paretoauswahl ermöglicht einen Tradeoff zwischen Qualität und Verfügbarkeit manuell zu wählen (vgl. Abschnitt 8.3).

Anforderung 3 fordert, dass Fahrsituationen entsprechend der in der Realität auftretenden Häufigkeiten für die Applikation berücksichtigt werden. Dies wird im vorgestellten Konzept über die Ableitung der Häufigkeiten aus der Grundgesamtheit und der anschließenden Gewichtung kleinerer Stichproben sichergestellt (vgl. Abschnitt 8.2). Die Anforderung wird durch die global-optimale, datengetriebene Applikation erfüllt.

Anforderung 4 fordert die Möglichkeit, länderspezifische Varianten einer Applikation zu erstellen, um regionsspezifische Besonderheiten berücksichtigen zu können. In der virtuellen Analyse können länderspezifische Datensets (vgl. Tabelle 7.1) verwendet werden. Dies kann zum einen für das Finden einer optimalen Applikation für eine bestimmte Region verwendet werden. Zum anderen kann eine gefundene Applikation länderspezifisch validiert werden.

Laut *Anforderung 5* soll die Funktionsqualität über alle Länder hinweg annähernd ähnlich sein. Die Verwendung objektiver Metriken erlaubt eine Vergleichbarkeit über große zeitliche und räumliche Distanzen hinweg. Gleiche Szenarien werden mit denselben Maßstäben bewertet, wodurch eine ähnliche Qualität gezielt ausgewählt werden kann. Alternativ kann ein möglichst diverser Datensatz über

alle Länder hinweg erstellt werden, um eine möglichst optimale Qualität in allen Ländern zu erzielen.

Anforderung 6 fordert eine Vergleichbarkeit der Applikation über die Straßenverkehrstests hinweg. Durch die Einführung der Bedeutungsebene und die Interpretation der Fahrabschnitte als Szenarien sowie deren Unterteilung in Subszenarien ist die geforderte Vergleichbarkeit gewährleistet. Zeitlich und räumlich getrennte Fahrabschnitte können so verglichen werden. Zudem können ähnliche Fahrabschnitte mit unterschiedlichen Parameterausprägungen verglichen werden. Die Anforderung ist damit erfüllt.

Anforderung 7 fordert, dass die finale Qualität der Funktion für alle Länder bekannt sein muss. Dies ist durch die Verwendung von länderspezifischen Validierungsdatensätzen bzw. einem globalen Validierungsdatensatz mit Straßenverkehrstests aus jedem Land sichergestellt.

Anforderung 8 fordert die Skalierbarkeit der Applikation. Durch die virtuelle, automatisierte Analyse ist eine zeitliche und räumliche Skalierung der Applikation gegeben. Zudem können Testfahrer anstelle der Applikateure für die Durchführung der Erprobungsfahrten eingesetzt werden. So lassen sich weltweit Testflotten betreiben, die Daten innerhalb der gesamten ODD einfahren. Unabhängige Funktionen und Fahraufgaben werden parallel appliziert und bieten weitere Skalierbarkeit.

Damit sind alle an die global-optimale, datengetriebene Applikation gestellten Anforderungen durch das vorgestellte Konzept erfüllt.

8.5 Erkenntnisgewinn zu den gestellten Forschungsfragen

8.5.1 Vergleichbarkeit sequentieller Parameteränderungen

Im bisherigen Applikationsprozess ist eine Vergleichbarkeit sequentieller Parameteränderungen nicht gegeben (vgl. Abschnitt 5.1.1). Der Vergleich von Parameteränderungen beruht bisher auf den Erinnerungen der Applikateure und Testfahrer während der Test- und Erprobungsfahrten. Auswirkungen der Parameter sowie deren Änderungen können immer nur stichprobenartig wahrgenommen werden. Dabei ist auch die Wahrnehmung der Fahrzeugumgebung sowie des internen Fahrzeugzustands auf die Wahrnehmungskapazitäten der Applikateure und Testfahrer beschränkt. Insbesondere die Bewertung der Funktionsqualität in einer Fahrsituation ist subjektiven Eindrücken ausgesetzt. Während der Testfahrten, die im Rahmen dieser Dissertation durchgeführt wurden, sind mehrfach Fahrsituationen aufgetreten, die von zwei Applikateuren im selben Fahrzeug sehr unterschiedlich bewertet wurden.

Durch die global-optimale, datengetriebene Applikation werden die genannten Schwachstellen in mehreren Punkten adressiert:

Durch den Einsatz objektiver Metriken ist ein Vergleich zeitlich und räumlicher getrennter Fahrtabschnitte ohne Abweichungen und Ungenauigkeiten möglich. Dadurch kann eine beliebige Anzahl an Fahrtabschnitten unter identischen Kriterien bewertet werden. Eine Veränderung der Metrik auf demselben Fahrtabschnitt ist direkt auf die Parameteränderung zurückzuführen.

Zudem wurde eine Bedeutungsebene geschaffen, die sich Konzepten aus dem szenariobasierten Testen bedient, die eine semantische Unterscheidung und Interpretation der als Ganzes vorliegenden Straßenverkehrstests ermöglicht. Dadurch findet der Vergleich unterschiedlicher Fahrten sowie unterschiedlicher Parametersätze auf semantisch ähnlichen Fahrtabschnitten statt.

Für die Applikation der längs- oder querführenden Fahrfunktion wurde zudem über die ausgeführte Fahrzeugbewegung und die darunterliegenden Fahraufgaben eine Systematik zur Dekomposition der Applikationsaufgabe geschaffen. Darüber hinaus können dieselben Fahraufgaben durch die Attribution weiter differenziert und besser verglichen werden.

Insgesamt werden nun dieselben Fahraufgaben in semantisch ähnlichen Fahrsituationen mittels objektiver Metriken bewertet und können direkt verglichen werden. Dies ermöglicht eine systematische und zielgerichtete Applikation. Die in der prototypischen Umsetzung verwendete Simulationethodik hat jedoch auch Schwächen. So verändert sich durch die Parameteränderung innerhalb der closed-loop Simulation auch das Fahrzeugverhalten. Dadurch kann es dazu kommen, dass dynamische Objekte, deren zeitliches Verhalten positionsbezogen ausgelöst wird, nicht mehr zur selben Zeit auftreten. Die Simulation ist bezüglich der Fahrfunktion weiter valide, jedoch verändert sich die Fahrsituation zwischen den einzelnen Simulationsdurchläufen. Dies erklärt unter anderem auch die variierende Anzahl an Kurven zwischen den Lösungen eines Applikationsdurchlaufs. Ein direkter Vergleich individueller Kurven mit unterschiedlichen Parametern ist nicht möglich, da sich die Auswirkungen der Parameteränderungen über die gesamte Fahrt addieren und demzufolge die Vorbedingungen und der Fahrzeugzustand vor der Kurvendurchfahrt über die unterschiedlichen Simulationsdurchläufe nicht notwendigerweise vergleichbar sind.

Stattdessen wurde im Konzept eine systematische Aggregation vorgesehen, die die Verteilung der Kurven im Attributraum einer jeden Lösung ins Verhältnis einer Grundgesamtheit setzt. Abweichungen in einzelnen Lösungen sowie Unzulänglichkeiten in der Datenbasis wird damit entgegengewirkt.

8.5.2 Aussagekraft kleiner Stichproben

Sowohl in der Voruntersuchung (vgl. Abschnitt 5.1.2) als auch in einem Experiment im Rahmen der Evaluation (vgl. Abschnitt 8.2) wurde gezeigt, dass zufällige, kleine Stichproben keine repräsentative Bewertung einer Applikation erlauben.

Vielmehr wurden Bias in den Daten aufgezeigt, die die Validität des Ergebnisses in Bezug auf die ODD negativ beeinflussen.

Da während der Applikation - sowohl im realen Straßenverkehr als auch in der Simulation - eine Vielzahl von Parameterkombinationen ausprobiert werden muss, um die optimale Applikation zu finden, ist es nicht möglich die Größe der Stichproben so zu erhöhen, dass eine Repräsentativität per se gewährleistet ist. In der Voruntersuchung wurde an einem Beispiel aufgezeigt, dass die notwendigen, zu fahrenden Distanzen nicht wirtschaftlich wären (vgl. Abschnitt 5.1.1).

Im Konzept der global-optimalen, datengetriebenen Applikation wurde deshalb eine Methodik entwickelt, die eine repräsentative Bewertung auch auf kleinen Stichproben ermöglicht, solange alle Subszenarien abgedeckt werden (vgl. Abschnitt 8.2). Zudem wurde am Beispiel des Kurven-Szenarios gezeigt, dass bereits ab 20 Fahrten eine 1-fache Sättigung aller Subszenarien im Attributraum der Krümmung und Steigung erreicht wird. Bei einer Ausweitung des Attributraums auf weitere Attribute muss die Anzahl Fahrten erhöht werden. Um die Anzahl der Attributkombinationen einzugrenzen (Fluch der Dimensionalität), können paarweise Abdeckungen bzw. im Allgemeinen eine t-wise Abdeckung verwendet werden. Die relevanten Attribute sind vom Anwendungsfall abhängig und müssen durch Expertenwissen definiert werden. Die Abdeckung der Subszenarien kann vorab quantifiziert werden. Zudem können die seltenen Randbereiche als optional deklariert werden, um den Fokus der Applikation auf die relevanteren Subszenarien zu richten.

8.5.3 Integration in den Entwicklungsprozess

Das vorgestellte Konzept zur global-optimalen, datengetriebenen Applikation lässt sich in bestehende Entwicklungsprozesse integrieren. Die notwendigen Erprobungsfahrten sind bereits im Entwicklungsprozess verortet und für die Validierung der Systeme zwingend notwendig. Eine Aufzeichnung der Daten bedeutet

zwar einen geringfügigen Mehraufwand, die Daten lassen sich jedoch vielseitig verwenden (vgl. Abschnitt 4.3). Die Kosten für die Speicherung der Daten sind zu vernachlässigen, da es sich bei den benötigten Daten nur um Bus-Kommunikationsdaten im Fahrzeug handelt und nicht um Sensorrohdaten, wie Bild- bzw. Videodaten oder Punktwolken.

Ein Mehraufwand ist für die Implementierung der virtuellen Analyse einzuplanen. Insbesondere sind Algorithmen für die Identifikation der Fahraufgaben sowie Metriken für die Bewertung der Funktionsqualität und -verfügbarkeit zu entwickeln. Die notwendigen methodischen Grundlagen wurden in dieser Dissertation gelegt.

Das systematische Einfahren von Messdaten für einzelne Parametersätze sowie die finale Validierung dieser muss prozessual integriert werden. Dabei lassen sich eingeplante Erprobungs- und Testfahrten nutzen sowie allgemein jegliche Validierungsaufwände, die auch aus anderen Gründen durchgeführt werden müssen. Die systematische Variation der Parametersätze bedeutet wenig Mehraufwand. Um die Repräsentativität zu erhöhen, sollten die in Abschnitt 5.1.3 diskutierten Bias vermieden werden. Dazu ergriffene Maßnahmen fördern zeitgleich den Nutzen der Test- und Erprobungsfahrten für andere Zwecke.

Zusammenfassend lässt sich schließen, dass die Integration in den bestehenden Entwicklungsprozess möglich ist. Die notwendigen Prozessschritte lassen sich mit bestehenden Prozessschritten zusammenfassen bzw. darin integrieren. Für die initiale Implementierung der virtuellen Analyse ist ein Mehraufwand einzuplanen, der jedoch durch die anschließende Automatisierung der Bewertung mehr als ausgeglichen wird. Die zu erwartenden Effizienzgewinne sowie der gesteigerte Qualität der Applikation in Form global-optimaler Ergebnisse stellen ein positives Aufwand-Nutzen-Verhältnis dar.

Der Mehrwert einer global-optimalen, datengetriebenen Applikation liegt im systematischen Vorgehen und der objektivierten, gesamtheitlichen Bewertung. Durch das systematische Vorgehen werden die Parameterkombinationen gezielt variiert

und jeweils auf einer repräsentativeren Datenbasis evaluiert. Die Parameterkombinationen werden nicht mehr sequentiell ausprobiert sondern gezielt in verschiedenen Umgebungssituationen eingesetzt. Die Datenaufzeichnung ermöglicht eine gesamtheitliche Analyse im Nachgang. Dabei führt die objektivierte Bewertung in Kombination mit der Gewichtung der realen Auftretenshäufigkeiten zu repräsentativen Ergebnissen. Einzelne Ausreißer nach unten oder oben, die in Zufallsstichproben auftreten und bei einer subjektiven Bewertung zu stark ins Gewicht fallen können, werden durch die Auftretenshäufigkeiten der Grundgesamtheit relativiert.

Dabei spielt die Wahl der Grundgesamtheit eine entscheidende Rolle. Für die prototypische Umsetzung wurde die gesamte Erprobungsdatenbasis genutzt. Im Produktiveinsatz ist die Verwendung von Kundenflottendaten für das dynamische Fahrverhalten bzw. Kartendaten für statische Umgebungsinformationen vorgesehen.

Um zusätzlich den Aufwand für die Applikation zu verringern und die ressourcentechnisch begrenzten Straßenverkehrstests effizient zu nutzen, ist eine zweistufige Applikation denkbar. In einem der Applikation mittels Straßenverkehrstests vorgelagerten Schritt werden für die zu applizierende Funktion durch simulative Ansätze Parametersatzvorschläge erzeugt. Der Proof-of-Concept dafür wurde in dieser Dissertation erbracht.

Der bzw. die ausgewählten Parametersätze müssen anschließend, wie in Kapitel 6 beschrieben, im realen Straßenverkehr validiert und bei Bedarf angepasst werden. Ein früher Einsatz der virtuellen Applikation erlaubt ein iteratives Vorgehen, so dass mit neu gesammelten Daten weitere Applikationsdurchläufe durchgeführt werden können, um bis zum Reifegrad 7 und damit dem Abschluss der Entwicklung eine optimale Applikation zu erzielen.

9 Zusammenfassung und Ausblick

9.1 Zusammenfassung

Zu Beginn der Dissertation wurde der automobiler Wertschöpfungsprozess vorgestellt. Dabei wurde aufgezeigt, dass die Applikation ein wesentlicher Baustein zur Sicherung der markenspezifischen Fahrcharakteristika der Fahrzeuge ist. Fahrzeugeigenschaften, die Kunden mit den Markenversprechen der Hersteller verbinden, werden auf einzelne Funktionen heruntergebrochen und am Ende der Entwicklung für jedes Fahrzeugmodell appliziert.

Die Applikation findet bisher vorwiegend in der realen Welt mittels Straßenverkehrstests statt. In einer Voruntersuchung wurden Unzulänglichkeiten dieses Vorgehens aufgezeigt. Die Applikation in der realen Welt führt unvermeidlich zu suboptimalen Ergebnissen, da die Vergleichbarkeit von sequentiell durchgeführten Parameteränderungen nicht gegeben ist. Zudem unterliegt die Bewertung durch den Applikateur dem subjektiven Eindruck. Aufgrund limitierter Ressourcen sind die durchgeführten Straßenverkehrstests zudem nicht notwendigerweise repräsentativ.

In der Literatur finden sich Ansätze zur virtuellen Applikation von ADAS und HAF-Systemen. Diese fokussieren sich auf die Objektivierung des subjektiven Fahreindrucks. Für die Umfeldmodellierung werden aufgezeichnete Realdaten verwendet, allerdings nur in kleinem Umfang ohne Betrachtung der Repräsentativität. Zudem wird die Varianz der Fahrzeugumgebung nicht weiter untersucht.

Kern dieser Dissertation ist das Konzept zur global-optimalen, datengetriebenen Applikation. Es wurden zwei methodische Bausteine erarbeitet, die eine repräsentative Applikation auf Basis aufgezeichneter Straßenverkehrstests ermöglichen. Zum einen wurde die Vergleichbarkeit zwischen zeitlich und räumlich getrennten Straßenverkehrstests über die Einführung einer Bedeutungsebene hergestellt. Diese ermöglicht den semantischen Vergleich einzelner Fahrabschnitte. Dazu wurde die Applikation von ADAS und HAF-Funktionen in Fahraufgaben zerlegt. Diese werden anschließend unabhängig voneinander appliziert. Zum anderen wurde durch die Definition einer Grundgesamtheit eine Referenz geschaffen, die es erlaubt auf Basis kleiner Stichproben dennoch repräsentative Analysen abzuleiten.

Das vorgestellte Konzept wurde anschließend prototypisch am Beispiel eines prädiktiven ACC-Systems umgesetzt. Dazu wurden eine Fahraufgabe der Funktion, die prädiktive Kurvenregelung, ausgewählt sowie eine exemplarische Datenbasis vorgestellt. Für unterschiedliche Applikationsstrategien wurden mittels Re-Simulation von aufgezeichneten Erprobungsfahrten Parametersätze simuliert, um anschließend für die virtuelle Analyse bereitzustehen.

Für die Evaluation der methodischen Bausteine wurden verschiedene Experimente durchgeführt. Zum einen wurde die global-optimale, datengetriebene Applikation am Beispiel des Spanien-Erprobungsdatensatzes durchgeführt. Dabei wurden mögliche Aggregationsverfahren gegeneinander evaluiert und der Mehrwert der Gewichtung von Subszenarien bezüglich der Repräsentativität einer Applikation nachgewiesen. Weiterhin wurde die Aussagekraft einer Applikation über den Stichprobenfehler untersucht. Dazu wurden verschieden große Datensätze verwendet und analysiert, inwieweit der Stichprobenfehler Aufschluss auf die Repräsentativität des verwendeten Datensatzes gibt. Nachdem die methodischen Bausteine im Einzelnen evaluiert wurden, wurde der Beitrag zu den gestellten Forschungsfragen aufgezeigt.

Die wesentlichen wissenschaftlichen Erkenntnisse dieser Dissertation lassen sich wie folgt zusammenfassen:

- Die Unzulänglichkeiten bezüglich der Repräsentativität und Vergleichbarkeit bisheriger Applikationsverfahren wurde in Voruntersuchungen aufgezeigt.
- Die Vergleichbarkeit zeitlich und räumlich getrennter Straßenverkehrstests wurde über die Einführung einer Bedeutungsebene hergestellt.
- Durch die Definition einer Grundgesamtheit und dem Vergleich dieser mit den für die Applikation verwendeten Stichproben wurde eine repräsentative Applikation ermöglicht.
- Die Gewichtung der Szenarien und ihrer Subszenarien ermöglicht die Verwendung kleiner Stichproben für die Applikation. Dies eröffnet auch das Potenzial einer virtuellen Applikation mittels Optimierungsalgorithmen.
- Es wurde ein Maß definiert, welches die Repräsentativität einer Stichprobe quantifiziert.
- Zu guter Letzt wurde durch die global-optimale, datengetriebene Applikation eine (Teil-)Automatisierung des Applikationsprozesses ermöglicht.

9.2 Ausblick

Die vorliegende Dissertation hat einen neuen datenbasierten Ansatz für die Applikation längs- und querführender Fahrfunktionen aufgezeigt. Durch das Aufzeichnen von Erprobungsdaten können die Fahrfunktionen virtuell und systematisch analysiert und appliziert werden. Der Applikationsprozess wird gleichzeitig beschleunigt und global optimiert.

Für die Anwendung im industriellen Kontext muss eine Systematik zur Ableitung der Fahraufgaben und deren Attribute entwickelt werden. Dazu ist zu untersuchen, welche Faktoren einen Einfluss auf die Qualität einer Fahrfunktion für eine spezifische Fahraufgabe haben.

Darüber hinaus muss im Kontext der Fahrfunktion untersucht werden, welche Fahraufgaben und deren Parameter sich gegenseitig beeinflussen und welche sich parallel applizieren lassen. So können weitere Effizienzgewinne erzielt werden.

Für die objektive Bewertung von Fahrerassistenzsystemen und hochautomatisierten Fahrfunktionen gibt es zahlreiche Forschungsarbeiten. Weitere Metriken, beispielsweise in Form von Objektivnotenmodellen (vgl. Abschnitt 5.2.2), können integriert werden, um das gewünschte Fahrverhalten noch exakter durch die Metriken abzubilden. Die geeigneten Metriken hängen dabei vom markenspezifischen, gewünschten Fahrverhalten ab.

Zudem könnte der Einsatz künstlicher Intelligenz zur Bewertung der Fahrfunktion untersucht werden. Ein subjektives Labeling durch die Testfahrer könnte einen Datensatz erzeugen, der als Ausgangspunkt für das Training überwachter Lernverfahren verwendet werden kann. Die künstliche Intelligenz hat das Potenzial komplexe Zusammenhänge zwischen Umgebungssituation und erlebtem Fahrverhalten zu erlernen.

In der prototypischen Umsetzung dieser Dissertation wurde bereits gezeigt, dass ein rein virtueller Ansatz für die Applikation möglich ist. In weiterführenden Untersuchungen könnten eine breite Anwendbarkeit für weitere Funktionen und Fahraufgaben betrachtet sowie Kriterien für die Anwendbarkeit von Simulationsumgebungen für die Applikation aufgestellt werden, um die Applikation von längs- und querführenden Fahrfunktionen weiter zu optimieren. Dazu könnte der eingesetzte Optimierungsalgorithmus verfeinert und auf das Optimierungsproblem angepasst werden. Beispielsweise könnte untersucht werden, ob der Einsatz von Tabu-Regionen die Exploration für das gegebene Optimierungsproblem weiter verbessern kann.

A Anhang

A.1 Grafiken

A.1.1 Variante des Produktentstehungsprozess (PEP) von Porsche Engineering

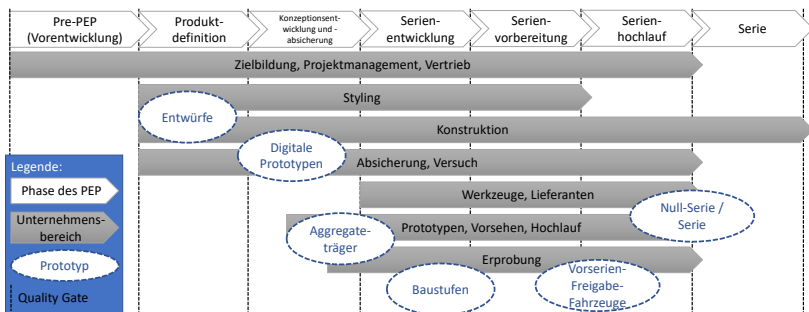


Abbildung A.1: Variante des PEP von Porsche Engineering mit sieben Phasen, den involvierten Unternehmensbereichen sowie erzeugten Prototypen. (in Anlehnung an [54])

A.1.2 Erprobungsplanung - direkt und adaptiert

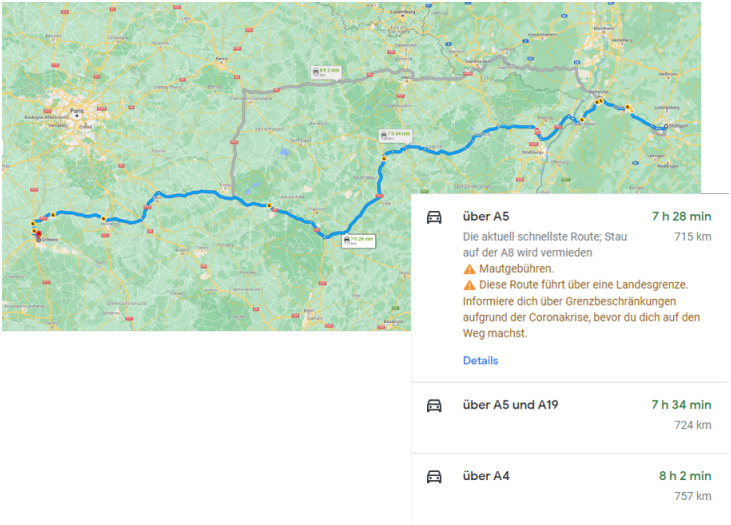


Abbildung A.2: Exemplarische Routenplanung für eine erste Etappe einer Erprobung

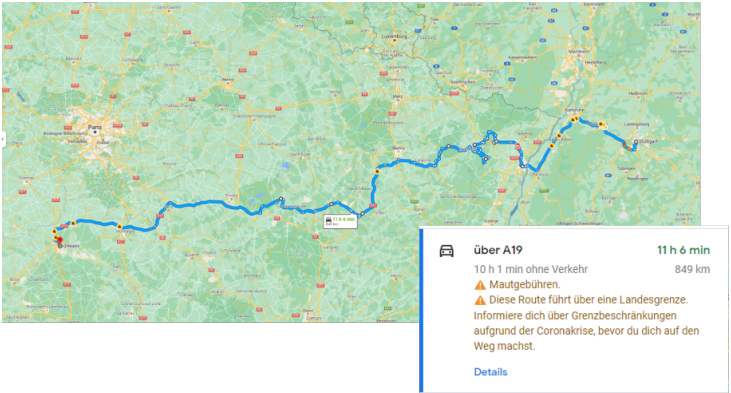


Abbildung A.3: Exemplarische Routenplanung für eine erste Etappe einer Erprobung unter Beachtung potentieller Bias durch die Straßenklasse.

A.2 Optimierungsalgorithmus

A.2.1 Funktionsweise Modified Cuckoo Search

Der Modified Cuckoo Search (MCS) Algorithmus ist ein Optimierungsverfahren, das aus der Natur inspiriert ist. Die Grundlage des Algorithmus beruht auf der Beobachtung des Brutverhaltens einiger Kuckucksarten.

Der MCS Algorithmus gehört zu den genetischen Optimierungsverfahren, die über mehrere Generationen hinweg nach einer optimalen Lösung suchen. Das Grundprinzip lässt sich wie folgt beschreiben:

Zu Beginn wird die Populationsgröße, also die Anzahl an "Kuckucken", festgelegt. In jeder Iteration legt jeder Kuckuck ein Ei", das jeweils einer Ausprägung der Parameterwerte entspricht. In der ersten Iteration wird das Ei in ein zufälliges "Nest", also eine zufällige Position im Suchraum, gelegt. Am Ende einer Iteration werden alle Lösungen bewertet. Auf Basis der besten Lösungen wird die nächste Generation gebildet. Nach dem Evolutionsprinzip wird die beste Lösung beibehalten, die schlechtesten Lösungen verworfen und in der Nähe von guten Lösungen nach besseren Lösungen gesucht. [145]

Der Modified Cuckoo Search with Lévy Flight (MCS-LF) stellt eine Variante des MCS dar, bei dem die Auswahl der Lösungen für die nächste Generation verändert wurde. Dabei wird ein Random Walk Verfahren bei der Auswahl der nächsten Lösung verwendet, wobei die Schrittlänge aus einer Lévy-Verteilung gezogen wird. Diese ist asymmetrisch und besitzt einen schweren Rand (heavy-tailed) wodurch mit hoher Wahrscheinlichkeit ein kleiner Sprung und mit geringer Wahrscheinlichkeit ein großer Sprung durchgeführt wird. Dies soll die Abwägung zwischen Exploration des Suchraums und Ausnutzung bereits gefundener guter Lösungen verbessern und damit die Suche nach der optimalen Lösung beschleunigen. [145, 146]

$$x_{t+1} = x_t + \alpha \cdot \text{Lévy}(\lambda) \quad (\text{A.1})$$

Es gibt verschiedene weitere Modifikationen des MCS-LF, wie beispielsweise die dynamische Anpassung der Lévy-Flug-Parameter zur Anpassung der Suche an die aktuelle Lösungsqualität oder den Einsatz eines AbandonMechanismus, um schlechte Lösungen zu entfernen. [146, 147]

Algorithm 1 Modified Cuckoo Search with Lévy Flight

```

1: Given:  $A$ : maximum Lévy step size,  $\varphi$ : golden ratio
2: Generate an initial population of  $n$  solutions  $X_i, i = 1, 2, \dots, n$ 
3: Evaluate the objective function  $f(x)$  for all initial solutions
4: Set generation counter  $G \leftarrow 1$ 
5: while stopping criterion not met do
6:   Sort all solutions by fitness
7:   for all abandoned solutions  $x_i$  do
8:     Compute Lévy step size  $\beta = \frac{A}{\sqrt{G}}$ 
9:     Generate a new solution  $x_k$  using Lévy flight
10:     $x_i \leftarrow x_k, F_i \leftarrow F_k$ 
11:   end for
12:   for all  $x_i$  in the elite group do
13:     Randomly pick another solution  $x_j$  from the elite group
14:     if  $x_i = x_j$  then
15:       Compute  $\beta = \frac{A}{G^2}$ 
16:       Generate a new solution  $x_k$  using Lévy flight
17:     else
18:       Compute  $d_x = |x_i - x_j|^\varphi$ 
19:       if  $d_x = 0$  then
20:          $x_k \leftarrow \frac{x_i + x_j}{2}$ 
21:       else
22:         Move distance  $d_x$  from the worse of  $x_i$  and  $x_j$  to generate  $x_k$ 
23:       end if
24:     end if
25:     Evaluate fitness  $F_k$  at  $x_k$ 
26:     Randomly select another solution  $x_l$  from the population
27:     if  $F_k < F_l$  then
28:        $x_l \leftarrow x_k, F_l \leftarrow F_k$ 
29:     end if
30:   end for
31:    $G \leftarrow G + 1$ 
32: end while

```

A.2.2 Hyperparameter der Modified Cuckoo Search

Parameter	Wert	Beschreibung
ALG_MCS_DISCARD_RATIO	0,7	Verwurfsrate von Lösungen in jeder Generation
ALG_MCS_STEPSIZE_FACTOR	0,5	Maximale, prozentuale Distanz, die ein Parameter in einem Schritt zurücklegen kann (in Bezug auf $ Max - Min $)
ALG_MCS_NUM_MIN_SOLUTIONS_TO_KEEP	10	Minimale Anzahl an Lösungen zu jedem Zeitpunkt
ALG_MCS_NUM_SOLUTION_TO_DELETE	0	Anzahl Lösungen, die in jeder Generation gelöscht werden sollen
ALG_MCS_ALPHA	0,5	Skalierungsparameter für Lévy-Distanz / Schrittgröße
ALG_MCS_BETA	1,5	Skalierungsparameter für Sigma
ALG_MCS_MY	0,5	Mittelwert um den sich die Normalverteilung verteilt
ALG_MCS_SIGMA_V	0,25	Standardabweichung für Normalverteilung von v

Tabelle A.1: In der C++ Implementierung verwendete Hyperparameter der Modified Cuckoo Search

A.3 Kalibrierung und Tuning des virtuellen Applikationsframeworks

A.3.1 Diskretisierung des Attributraums

Die Diskretisierung des Attributraums in Bins hat einen maßgeblichen Einfluss auf die Qualität der Applikation. Die Optimierung ist stets nur bezüglich der gewählten

Diskretisierung optimal. Entsprechend soll die Wahl des Binnings untersucht werden.

Dafür werden zwei Funktionsparameter mit dem Trainingssplit aus dem Deutschlanddatensatz (vgl. Tabelle 5.1) appliziert. Die Freiheitsgrade der beiden Funktionsparameter belaufen sich auf den Realzahlenbereich zwischen 0 und 1 mit einer Genauigkeit von 6 Nachkommastellen. Für die Bewertung der Applikation wird die relevante Abweichung im Kurvenscheitel in m/s herangezogen.

Für die Applikation wurde ein genetischer Optimierungsalgorithmus eingesetzt, der über 11 Generationen mit jeweils 25 Agenten eine optimale Lösung gesucht hat. Die durch den MCS-LF Algorithmus erreichte Abdeckung des Funktionsparameterraums (vgl. Abbildung A.4) weist vereinzelt Häufungen und an anderen Stellen im Parameterraum Lücken auf. Die jeweilige Bewertung der Applikation schwankt auch zwischen Applikationen, die im Parameterraum benachbart sind.

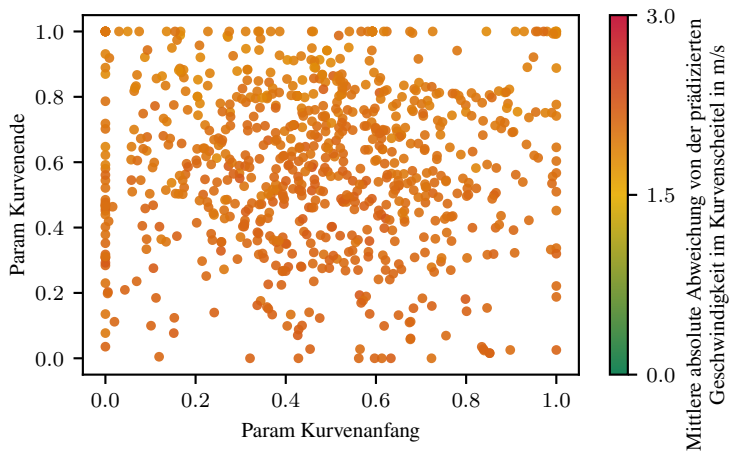


Abbildung A.4: Exploration des Funktionsparameterraums in Versuch 1.

Es wurden zwei Varianten für die Diskretisierung untersucht: Jeweils 8 Unterteilungen für die Attribute Krümmung und Steigung und damit 64 Subszenarien sowie jeweils 4 Unterteilungen je Attribut und damit 16 Subszenarien.

A.3.1.1 Diskretisierung mit jeweils 8 Unterteilungen

Die Analyse ergibt, dass insbesondere die Ränder der betrachteten Attribute und damit die äußeren der 64 Subszenarien im Szenario-Attributraum nur selten in den Daten vorkommen (vgl. Abbildung A.5). Des Weiteren beeinflussen die Funktionsparameter die Resimulation der Daten insofern, dass die Geschwindigkeit des Egofahrzeugs zwischen den einzelnen Durchläufen variiert. Dadurch kann es vorkommen, dass in einigen Durchläufen Kurven beispielsweise hinter anderen Verkehrsteilnehmern durchfahren werden, während in anderen Durchläufen das Vorderfahrzeug erst davor bzw. danach angetroffen wird. Für den Fall, dass das Vorderfahrzeug die Geschwindigkeit in der Kurve beeinflusst, lässt sich die Abweichung zur prädierten Geschwindigkeit nicht mehr auf die Kurve und damit auf die zu applizierenden Funktionsparameter abbilden. Insofern kann die Kurve nicht in die Bewertung aufgenommen werden. Dies erklärt die Schwankungen der Kurvenanzahl und in der Folge - insbesondere bei Schwankungen in den Randbereichen des Szenario-Attributraums - die unterschiedlichen Abdeckungen und Bewertungen (vgl. Abbildung A.5). Exemplarisch ist dies in den Subszenarien 7, 32 und 57 zu sehen. Generation 15 Lösung 1 (vgl. Abbildung A.5b) enthält keine Kurven in den Subszenarien 7, 32 und 57 (rot umrandet). In den anderen Lösungen treten dagegen vereinzelt oder in allen drei Subszenarien Kurven auf (vgl. Abbildungen A.5a, A.5c, A.5d). Einen besonders großen Einfluss hat dies, falls die Kurven, die nicht in allen Lösungen enthalten sind, besonders stark vom Mittelwert abweichen. Subszenario 57 wurde in Lösung 0 und 3 mit einer hohen Abweichung im Mittel bewertet. In den Lösungen 1 und 2 (Abbildungen A.5b und A.5c) ist diese Kurve nicht enthalten. Die möglicherweise schlechte Funktionsqualität in diesem Subszenario wurde nicht ermittelt und kann in der Gesamtbewertung nicht berücksichtigt werden.

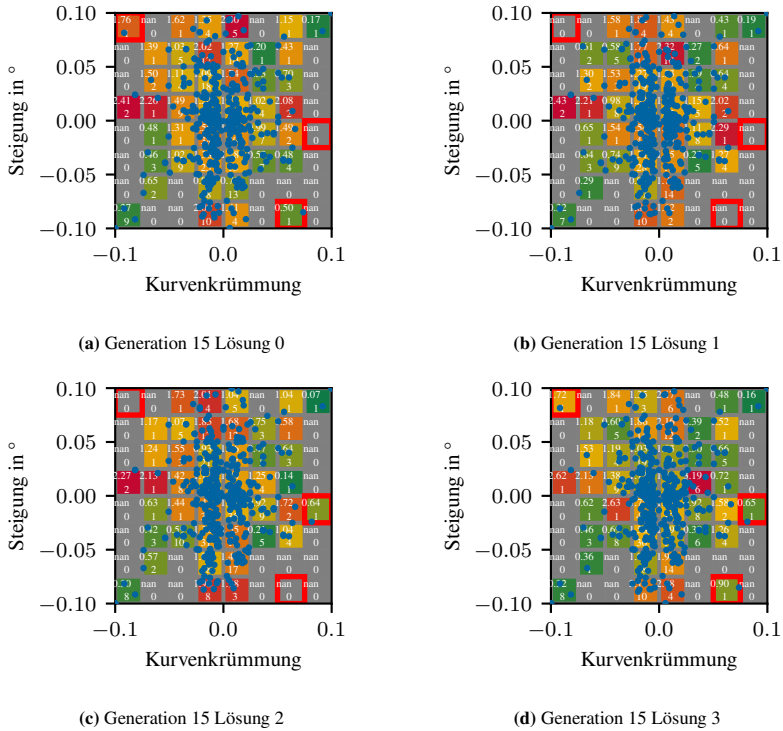


Abbildung A.5: Abdeckung des Szenario-Attributraums in exemplarischen Lösungen des Versuch 1. Nicht vorhandene Subszenarien sind grau hinterlegt und mit NaN bezeichnet. Für die anderen Subszenarien ist die Bewertung des Subszenarios (obere Zahl) und die Anzahl vorhandener Kurven (untere Zahl) dargestellt. Die blauen Punkte zeigen die Datenpunkte der Kurven im Attributraum.

Die in der Analyse aufgefallenen Faktoren haben zur Folge, dass in Applikationen einzelne Kurven unterschiedlich berücksichtigt werden und damit auch die Gesamtzahl berücksichtigter Kurven variiert (vgl. dazu Abbildung A.6). So kann beispielsweise eine Kurve in einer Applikation eine hohe Abweichung von der prädizierten Geschwindigkeit im Kurvenscheitel aufweisen und eine weitere Applikation diese Kurve aufgrund geänderter Funktionsparameter nicht berücksichtigen. Ist diese Kurve nun eine von wenigen Kurven im Subszenario, kann dies

einen starken Einfluss auf die Gesamtbewertung haben. Dadurch sind sprunghafte Veränderungen zwischen benachbarten Applikationen im Funktionsparameter-raum zu erklären.

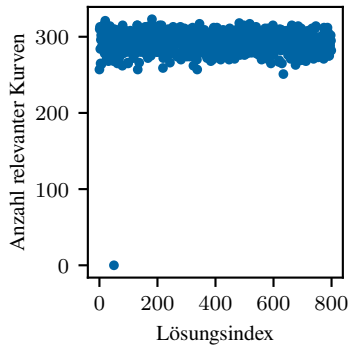


Abbildung A.6: Anzahl der relevanten Kurven für jede Lösung in Versuch 1.

A.3.1.2 Diskretisierung mit jeweils 4 Unterteilungen

Die Diskretisierung mit nur jeweils 4 Unterteilungen je Attribut hat zur Folge, dass bei gleicher Datengrundlage der Attributraum in weniger Subszenarien aufgeteilt wird. Es ergeben sich 16 anstatt 64 Subszenarien.

Die Subszenarien decken damit jeweils den vierfachen Bereich des Attributraums ab. Insbesondere sind in den Subszenarien im Randbereich des Szenario-Attributraums mehr Kurven enthalten, da diese durch die größeren Bins auf jeder Achse doppelt so weit in Mitte des Attributraums reichen, in dem sich sowohl für die Krümmung als auch die Steigung mehr Kurven befinden (vgl. Abbildung A.5 und Abbildung A.7).

Die Abdeckung des Attributraums bzw. der 16 Subszenarien ist in allen Lösungen höher als in der Variante mit 64 Subszenarien. Zudem ist die Anzahl einzelner

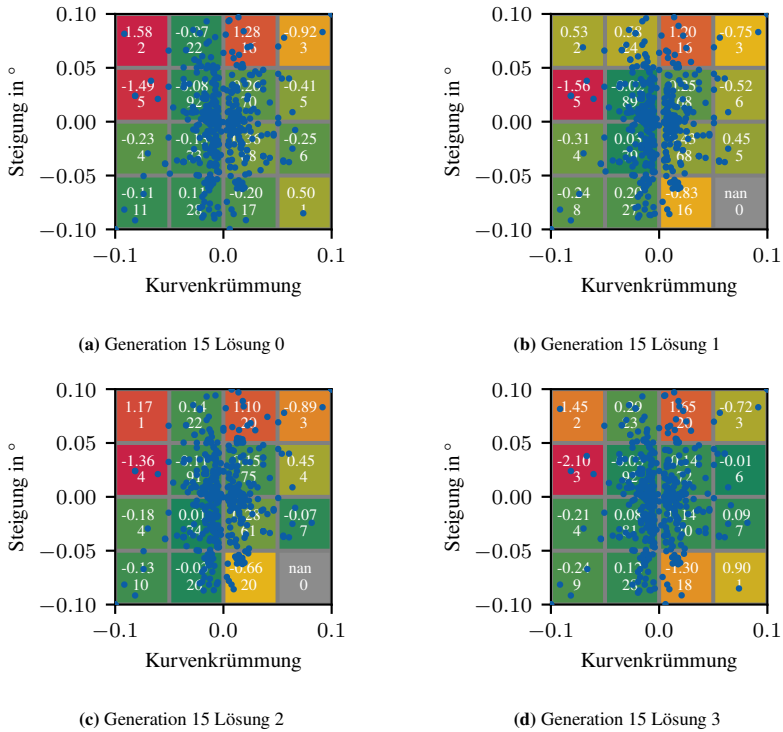


Abbildung A.7: Abdeckung der Subszenarien in exemplarischen Lösungen des Versuch 2

Kurven in den Subszenarien höher, was die Schwankungen bei der Aggregation innerhalb des Subszenarios durch Ausreißer, also einzelne besonders schlecht oder besonders gut bewertete Kurven, reduziert. Zusätzlich verringern sich die Schwankungen bei der Aggregation der einzelnen Subszenarien zu einem Gesamtergebnis durch die geringere Anzahl Subszenarien sowie die höhere Abdeckung der Subszenarien und somit geringere Anzahl fehlender Subszenarien. Die Vergleichbarkeit der Lösungen untereinander wird erhöht.

A.3.1.3 Weitere Diskretisierungsvarianten

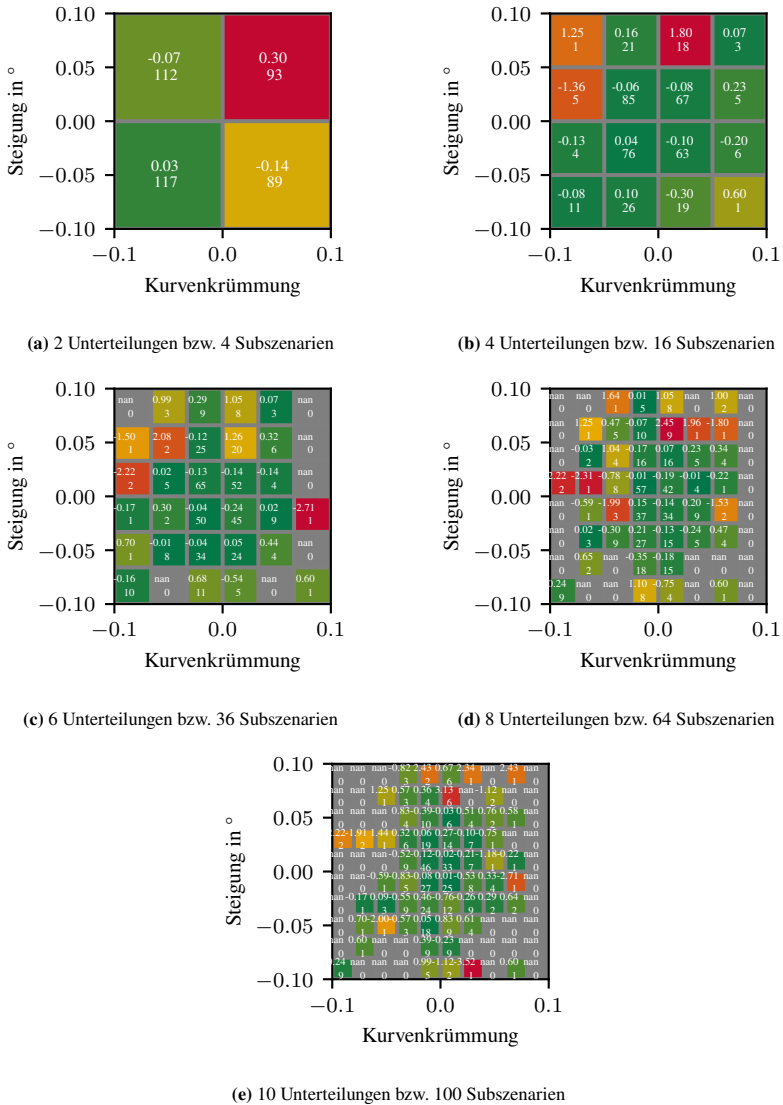


Abbildung A.8: Untersuchung verschiedener Unterteilungen des Attributraums bezüglich der Subsznario-Abdeckung

A.4 Beschreibung der Umsetzung

A.4.1 Datenupload & -aufbereitung

Für die virtuelle Analyse der weltweit durchgeführten Applikationsfahrten wird eine zentrale Datensammlung im Backend benötigt. Dieses Backend kann lokal am Entwicklungsstandort oder in einer selbstgehosteten Cloud betrieben werden. Alternativ kann auch ein externer Cloudanbieter verwendet werden. Für eine zeitnahe Verfügbarkeit der Daten ist ein schneller Up- und Download sowie ein weltweiter Zugang zum Backend von Vorteil. Die Buskommunikationsdaten müssen dann mittels geeigneter Datenkonverter für die Analyse in gängigen Datenanalyse-Tools verfügbar gemacht werden. Dabei eignet sich das parquet-Datenformat, da es mit standardisiert ist und mehrere Programmiersprachen unterstützt werden. Zudem werden die Daten stark komprimiert, wodurch der benötigte Speicherplatz reduziert wird. Konverter-Skripte müssen in der Regel individuell für die jeweiligen Daten entwickelt werden. Dabei kann auf Bibliotheken der Messdaten-Toolhersteller¹ zurückgegriffen werden.

Für den Datenupload und die -aufbereitung gibt es zahlreiche bestehende Tools und Lösungen. Für die Umsetzung der virtuellen Applikation wurden Daten aus einer cloudbasierten Datenmanagementplattform bezogen. Diese ermöglicht einen weltweiten Upload der Daten sowie einen etablierten Dataingest, also standardisierte Prozessschritte zur Konvertierung, Bereinigung und Anreicherung der aufgezeichneten Fahrzeugmessdaten. Auf die in Amazon's *Simple Storage Service*² gespeicherten Daten kann anschließend über virtuelle Maschinen auf einer Amazon Recheninstanz zugegriffen werden. Außerdem ist ein Download für die lokale Verwendung möglich. Die Daten werden im Dataingest in das parquet-Format konvertiert.

¹ bspw. MDF4Lib von Vector Informatik GmbH

² Amazon S3: <https://aws.amazon.com/de/s3/>

A.4.2 Framework zur Analyse und Aggregation von Fahrzeugmessdaten

Aufbauend auf bereitgestellten Zeitreihendaten im parquet-Format kann die Analyse der Fahrzeugmessdaten beginnen. Diese reicht von einfachen Visualisierungen der in den Daten enthaltenen GPS-Positionen bis zu detaillierten Verhaltensanalysen des Fahrzeugs bzw. einzelner Funktionen in bestimmten Situationen. Die dafür notwendigen Methoden sind vielseitig und vom Einzelfall abhängig. Etabliert ist daher die Bereitstellung von Frameworks mit einem über den Einzelfall hinausreichenden Funktionsumfang. Für die virtuelle Applikation wurde dazu ein selbstentwickeltes Framework³ verwendet. Nach der Veröffentlichung aus dem Jahr 2017 wurde das Framework und die dort beschriebenen Methoden über die Zeit weiterentwickelt, in Python transferiert sowie mit einem Analyse-tool fusioniert, das insbesondere die Identifikation von Szenarien mittels regulärer Ausdrücke ermöglicht⁴. Die folgenden Erläuterungen beziehen sich auf die Pythonversion des Frameworks.

Die gesamte Datenanalyse kann mit verschiedenen Tools durchgeführt werden. Für die vorliegende Implementierung wurde Python3⁵ verwendet, da die Programmiersprache Open-Source ist und dadurch sowohl eine weite Verbreitung als auch eine aktive Community mit zahlreichen Funktionsbibliotheken bietet. Für diesen Anwendungsfall besonders geeignet ist die *pandas*⁶-Funktionsbibliothek, die zahlreiche Funktionen für die Analyse von Zeitreihendaten enthält.

³ Veröffentlicht in *Framework for using real driving data in automotive feature development and validation* [JL8]

⁴ Veröffentlicht in *Leveraging regular expressions for flexible scenario detection in recorded driving data* [JL3]

⁵ Python3 verfügbar unter <https://www.python.org/>

⁶ pandas-Funktionsbibliothek verfügbar unter <https://pandas.pydata.org/>

A.4.2.1 Interpretation des Fahrtinhalts

Bei der Analyse von Zeitreihendaten kann jeder einzelne Zeitpunkt bewertet werden. Während einer Erprobungsfahrt werden die Buskommunikationsdaten in ihrer auftretenden Frequenz aufgezeichnet. Diese werden dabei nicht gleichzeitig sondern sequentiell über die Buskommunikationssysteme übertragen. Während der Umwandlung zu Zeitreihendaten mit einem einheitlichen Zeitstempel wird in der Regel die höchste Taktfrequenz verwendet, um einen Informationsverlust zu vermeiden. Da Nachrichten im Fahrzeug mit einer Taktfrequenz von bis 100 Hz übertragen werden, entstehen so große Datenmengen. 100 Hz Taktfrequenz resultiert in 360.000 Datenpunkte pro Stunde, was bei einer achttündigen Erprobungsfahrt 2.880.000 Datenpunkte bedeuten würde. Diese können nicht zielführend ausgewertet werden, da jeglicher Kontext fehlt. Die Datenmenge kann durch die Wahl einer niedrigeren Abtastrate reduziert werden. Dabei wird dann jedoch auch ein Informationsverlust riskiert, beispielsweise von kurzfristigen Statusbitwechseln. Andere Formen der Aggregation der Daten sind daher vorzuziehen. Ein primitiver Ansatz ist die Zusammenfassung einer festen Anzahl von Zeitpunkten - optional mit einem jeweils überlappenden sogenannten Sliding Window.

Für Anwendungen in anderen Domänen, wie beispielsweise die Auswertung von Finanz-, Produktions- oder Nutzerdaten, stellt die Zeit die relevante Dimension der Auswertung und somit eine Aggregation über feste Dauern, wie beispielsweise Tagesverläufe für Aktienkurse oder Quartalsberichte für Geschäftszahlen, ein valides Vorgehen dar. Im Gegensatz dazu müssen Fahrfunktionen über ihren Kontext, also die vorliegende Umgebungssituation, analysiert und aggregiert werden. Eine feste Fensterung könnte sonst beispielsweise ein Überholmanöver oder einen Spurwechsel zerschneiden, was dazu führt, dass diese nicht gesamtheitlich bewertet werden können. Daher bedarf es für die Auswertung von Fahrfunktionen einer semantischen Aggregation der Zeitpunkte einer Erprobungsfahrt.

Für die semantische Interpretation des Fahrtinhalts gibt es zwei grundlegende Möglichkeiten. Zum einen kann die Fahrt stets als Ganzes interpretiert werden. D.h. diese wird jeweils als Referenz für die Gesamtheit herangezogen. Jeder einzelne Zeitstempel tritt in genau einer Sequenz auf. Die Gesamtheit der zerteilten

Sequenzen ergibt wieder die ursprüngliche Fahrt. Alternativ zu einer zeitlichen Betrachtung kann diese Art der Analyse auch auf die Position bzw. Strecke bezogen werden. Zur Unterscheidung wird bei einer zeitlichen Betrachtung von Sequenzen und bei der positionsbezogenen Betrachtung von Segmenten gesprochen. Für die Betrachtung der Fahrzeugumgebung ist die Position in den meisten Fällen die geeignetere Bezugsgröße, da diese geschwindigkeitsbereinigt ist. Fortan wird zur Vereinfachung nur von Segmentierung gesprochen - die Anwendbarkeit auf Sequenzen ist analog. Zeitreihendaten lassen sich über die Geschwindigkeit des Egofahrzeugs in positionsbezogene Daten umwandeln (siehe Gleichungen A.2 bis A.4).

$$\Delta t_i = t_{i+1} - t_i | \forall t_i \in \vec{t} \quad (\text{A.2})$$

$$\Delta pos_i = \Delta t_i * v_{ego_i} | \forall \Delta t_i \in \vec{\Delta t} \quad (\text{A.3})$$

$$pos_i = \sum_{x=0}^i \Delta pos_x \quad (\text{A.4})$$

Eine weitere Möglichkeit ist die Extraktion einzelner Events aus der Gesamtfahrt. Events können sich gegenseitig überlagern (z.B. Spurwechsel und Überholmanöver) und nicht zu jedem Zeitpunkt einer Fahrt muss ein interpretiertes Event vorhanden sein. Beispielsweise sind für die Analyse eines ACC-Systems alle Fahrabschnitte ohne vorausfahrendem Fahrzeug irrelevant und könnten daher vernachlässigt werden.

Der segmentierende Ansatz eignet sich für die Interpretation der statischen Fahrzeugumgebung. Der Event-basierte Ansatz zielt dagegen auf die Erfassung der dynamischen Fahrzeugumgebung sowie des Verhaltens des Ego-Fahrzeugs und damit auch des SUTs ab.

A.4.2.1.1 Event-basierte Interpretation des Fahrthinhalts Events bzw. Manöver lassen sich als Abfolge von Zuständen modellieren. Für ein konkretes Event bzw. Manöver muss zunächst jeweils ein Ausgangszustand z_0 festgelegt

werden. Anschließend werden weitere notwendige Zustände und Zustandsübergänge beschrieben (vereinfacht dargestellt in Algorithmus 2).

Algorithm 2 Definition von Zuständen eines Events

```
function define_state
    z = new State()
    repeat                                ▷ Füge alle relevanten Signale des Zustands hinzu
        z.append(signal, signal_range)
    until user_abort return z
end function
```

Zur Detektion der Events und Manöver werden die Fahrzeugmessdaten sequenziell nach dem Vorhandensein des Ausgangszustands z_0 durchsucht (vereinfacht dargestellt in Algorithmus 3).

Algorithm 3 Detektion von Events in Zeitreihendaten

```
states = []
repeat                                ▷ Definiere alle Zustände des Events
    z = define_state()
    states.append(z)
until user_abort
file = load_testdrive()
t = file.time
iter = 0
sequences = []
while iter ≤  $|t_{end}|$  do
    if  $t[iter] == states[0]$  then
        seq = check_states(iter, states)
        if seq then
            sequences.append(seq)
        end if
        iter ++
    end if
end while
```

Zu jedem gefundenen z_0 wird anschließend die Abfolge der Folgezustände $z_1 - z_n$ mit erlaubten Zustandsübergangszeiten geprüft (vereinfacht dargestellt in Algorithmus 4). Bei Vorhandensein der kompletten Zustandsabfolge wird die gesamte Sequenz von $z_0 - z_n$ als Event bzw. Manöver extrahiert. Diese startet mit dem ersten Zeitstempel in z_0 und endet mit dem letzten Zeitstempel von z_n .

Algorithm 4 Überprüfen der Zustandsabfolge eines Event nach passendem Anfangszustand z_0

```

function check_states(cur_time, states)
    state_ptr = 0
    time_ptr = cur_time ▷ Anfangszustand kann länger als 1 Zeitpunkt sein
    while time_ptr ≤ |tend| do
        if t[time_ptr] ≠ states[state_ptr] then
            if state_ptr == |statesend| then
                sequence = new Sequence(cur_time, time_ptr) return
sequence
            end if
            abort
        else
            while t[time_ptr] == states[state_ptr] || t[time_ptr] ==
states[state_ptr + 1] do ▷ Entweder aktueller Zustand liegt weiterhin vor
oder Übergang zum Folgezustand
                if t[time_ptr] == states[state_ptr + 1] then
                    state_ptr ++
                    continue
                end if
                time_ptr ++
            end while
        end if
    end while
end function

```

A.4.2.1.2 Segment-basierte Interpretation des Fahrtinhalts Für die Segmentierung von Fahrzeugmessdaten wurde ein eigenes Verfahren entwickelt,

welches Segmente mit dynamischer Segmentlänge erzeugt⁷. Durch das Festlegen von segmentierenden Signalen - also Signale deren Änderung das aktuelle Segment beenden und ein neues Segment beginnen - werden semantische und in ihrer Länge flexible Segmente erzeugt, die anschließend durch Clusteringverfahren semantisch nach ihrer Ähnlichkeit gruppiert werden. Zusätzlich zu den segmentierenden Signalen werden weitere Signale als deskriptive Merkmale zu einem Segment angefügt. Dabei bieten sich kategoriale Daten mit Nominal- oder Ordinalskalen für die Segmentierung an. Metrische Daten werden ergänzend über Mittelwerte, Minimal- und Maximalwerte, usw. hinzugefügt (vereinfacht dargestellt in Algorithmus 5). Von dieser Aufteilung kann in begründeten Einzelfällen abgewichen werden. Durch Binning, Monotonieverhalten oder andere Diskretisierungsverfahren werden metrische Daten für die Segmentierung verwendet. Ebenso lassen sich kategoriale Daten beispielsweise über den ersten, letzten oder meist vorkommenden Wert sowie die Wertefolge als Segmentattribute verwenden.

Algorithm 5 Definiere relevante Signale für die Segmentierung

```

function define_relevant_signals
  s = new Signallist()
  repeat      ▷ Füge alle relevanten Signale für die Segmentierung hinzu
    sig = new Signal()
    if is_categorical(sig) then
      s.segmenting_signals.append(sig)
    else
      s.descriptive_signals.append(sig)
    end if
  until user_abort return s
end function
  
```

Anschließend werden die Fahrzeugmessdaten sequentiell durchlaufen und die Segmente gebildet (vereinfacht dargestellt in Algorithmus 6). Je nach Anwendungsfall

⁷ Veröffentlicht in *Logical Scenario Derivation by Clustering Dynamic-Length-Segments Extracted from Real-World-Driving-Data* [JL11]

müssen im Anschluss Merge-Strategien implementiert werden, sollte die Segmentierung zu kleinteilig geworden sein. Dies kann insbesondere bei korrelierten aber leicht zeit- bzw. positionsversetzten Signalen sinnvoll sein, wie beispielsweise dem aktuell ausgewählten Gang und der Geschwindigkeit des Ego-Fahrzeugs. Segmente werden im Postprocessing zusammengefasst.

Algorithm 6 Segmentierung von Fahrzeugmessdaten

```
s = define_relevant_signals()
file = load_testdrive()
t = file.time
iter = 0
t_segmentstart = 0
segments = []
for sig in s.segmenting_signals do
    cur_state.sig == NaN
end for
while iter ≤  $|t_{end}|$  do
    if compare_states(cur_state, file[iter]) == False then
        seg = new Segment(t_segmentstart, iter − 1)
        for sig in s.segmenting_signals do
            seg.sig == cur_state.sig
        end for
        for sig in s.descriptive_signals do
            seg.sig == calc_signal_value(sig, file[t_segmentstart, iter −
1])
        end for
        segments.append(sig)
        t_segmentstart = iter
        cur_state = file[iter]
    end if
    iter ++
end while
```

Je nach Anwendungsfall sind unterschiedliche Signale für die Segmentierung relevant. Das vorgestellte Verfahren erlaubt einen flexiblen Einsatz verschiedener Signale und ist somit auf beliebige Anwendungsfälle anpassbar.

Nachdem die Segmente gebildet wurden, werden die Segmentattribute berechnet. Für die nun bekannte Dauer bzw. Länge der Segmente, werden die metrischen Daten aggregiert und Kennzahlen berechnet.

Algorithm 7 Attribution der extrahierten Segmente

```
function calc_signal_value(signal, segment_data)  
    min_value = min(segment_data)  
    max_value = max(segment_data)  
    avg_value = avg(segment_data) return s  
end function
```

Die extrahierten Segmente werden anschließend mittels Clustering gruppiert, um Segmentgruppen ähnlich zu logischen Szenarien zu erhalten. Dazu muss für jedes Signal - sowohl die segmentierenden als auch die beschreibenden Signale - eine Distanzmetrik zwischen den möglichen Werten definiert werden. Anschließend werden unterschiedliche Clusteringverfahren eingesetzt, um die Segmentgruppen zu bilden.

A.4.2.2 Bewertung von Fahrtsegmenten und Fahrscenarien

Die im vorherigen Schritt extrahierten Segmente oder Events dienen als Grundlage für die Bewertung der Fahrfunktion. Die Bewertung einer Fahrfunktion wird analog event-basiert oder kontinuierlich vorgenommen und hängt vom Anwendungsfall ab. So muss beispielsweise eine automatisierte Spurmittenführung kontinuierlich bewertet werden, indem z.B. die mittlere Abweichung von der Spurmitte über die gefahrene Distanz oder Zeit (vgl. Formel A.5) ermittelt wird. Bei der Bewertung eines Spurwechselassistenten ist hingegen eine Event-basierte Bewertung sinnvoller, da explizit nur die Spurwechsel-Szenarien relevant sind (vgl. Formel A.6 & A.7).

$$\text{lane_deviation}_t = \text{left_line_distance_Y}_t + \text{right_line_distance_Y}_t \quad (\text{A.5})$$

$$\text{lane_change_1} = 1 - \left| \frac{t_{\text{scenario_end}} - t_{\text{scenario_start}}}{t_{\text{desired}}} - 1 \right| \quad (\text{A.6})$$

$$\text{lane_change_2} = 1 - \left| \frac{\max(\overrightarrow{v_ego_y_{\text{scenario}}})}{v_ego_y_{\text{desired}}} - 1 \right| \quad (\text{A.7})$$

Die Bewertung der Fahrfunktionen ist vielseitig durchführbar. Während der Erprobungsfahrt werden Events und Fahrtsegmente durch den Testfahrer bzw. Testbeifahrer manuell - subjektiv - bewertet. Über eine geeignete Toolunterstützung sollten diese Bewertungen automatisiert in das Backend überführt werden. Eine weitere Möglichkeit bietet die Definition von objektivierten Metriken. Diese werden manuell spezifiziert oder aus subjektiv gelabelten Daten abgeleitet. Möglich sind beispielsweise maximal tolerierte Grenzwerte für die Querbeschleunigung oder den Ruck sowie eine maximale Abweichung von der Spurmitte.

Diese objektivierten Metriken lassen sich automatisiert für Events bzw. Segmente berechnen. Die Möglichkeiten der Umsetzung sind vielfältig. Neben einer binären Pass-Fail-Mechanik analog zum Vorgehen bei Testfällen können relative und absolute Skalen verwendet werden. Dabei werden für jedes Segment bzw. Event der Mittelwert, Maximal- oder Minimalwert, Median etc. herangezogen aber auch Standardabweichungen oder Quantile berechnet.

Die Ausführung der Fahraufgabe - längs, quer oder kombiniert - lässt sich nur schwer durch eine einzige Metrik bewerten. Daher ist in vielen Fällen eine Kombination verschiedener Metriken notwendig (vgl. Gleichung A.8). Die Gewichtung hängt vom Einzelfall ab.

$$\text{lane_change} = 0,6 * \text{lane_change_1} + 0,4 * \text{lane_change_2} \quad (\text{A.8})$$

A.4.2.3 Aggregation und Gewichtung bewerteter Fahrtsegmente und Fahrscenarien

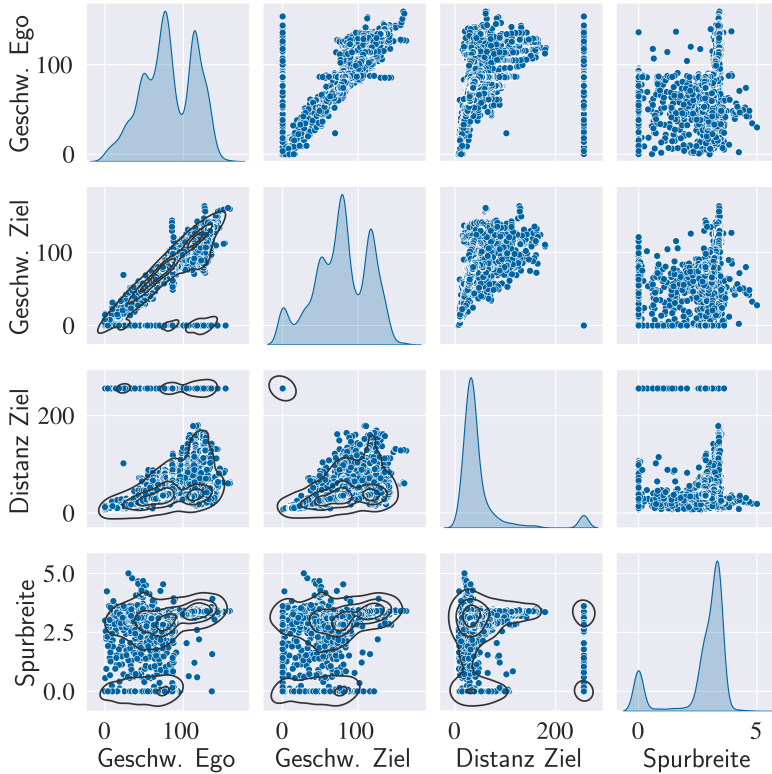


Abbildung A.9: Verteilung nach einzelnen Attributen in der Grundgesamtheit.

A.5 Weitere Ergebnisse

Datensatz	Abdeckung > 0	Abdeckung > 5
5 Fahrten #1	0.8750	0.3750
5 Fahrten #2	0.8125	0.3750
5 Fahrten #3	0.3125	0.0625
5 Fahrten #4	0.5625	0.0000
5 Fahrten #5	1.0000	0.5000
5 Fahrten #6	0.4375	0.1250
5 Fahrten #7	0.2500	0.0000
5 Fahrten #8	0.5000	0.0000
5 Fahrten #9	0.4375	0.0000
5 Fahrten #10	0.3125	0.0000
10 Fahrten #1	0.6875	0.2500
10 Fahrten #2	0.9375	0.2500
10 Fahrten #3	0.8125	0.2500
10 Fahrten #4	0.9375	0.4375
10 Fahrten #5	0.8750	0.2500
20 Fahrten #1	1.0000	0.5000
20 Fahrten #2	0.8750	0.5000
20 Fahrten #3		
20 Fahrten #4	0.8750	0.3750
20 Fahrten #5	0.8125	0.3750
30 Fahrten #1	1.0000	0.5000
30 Fahrten #2	0.9375	0.5000
30 Fahrten #3	1.0000	0.5000
50 Fahrten #1	1.0000	0.6875

Tabelle A.2: Berechnung der relativen Abdeckung aller Subsznarien mit >0 bzw. >5 Kurven über alle zufällig gezogenen Datensätze.

		Validierungsmetrik		
		Globaler Mittelwert	Subszenario Mittelwert	Gewichteter Subszenario Mittelwert
Irland Validierungsdatensatz	Ausgangsparametrierung	-14.26985	-12.32901	-1.13446
	Referenzapplikation	-14.72133	-18.00821	-1.26774
	Applikation SM	-12.45319	-6.53647	-1.07219
	Applikation gSM	-14.32954	-9.83648	-1.43226
Globaler Validierungsdatensatz	Ausgangsparametrierung	-1.87835	-1.19803	-0.11886
	Referenzapplikation	-1.85532	-1.28476	-0.11706
	Applikation SM	-1.89885	-1.00667	-0.11944
	Applikation gSM	-1.81442	-1.01692	-0.11469

Tabelle A.3: Ergebnisse der Validierung der jeweils besten Lösung unterschiedlicher Applikationsstrategien auf dem länderspezifischen Irland-Validierungsdatensatz sowie dem globalen Validierungsdatensatz. Dabei wurden für die Validierung für jede Lösung die Metriken für den globalen Mittelwert, den Subszenario Mittelwert sowie den gewichteten Subszenario Mittelwert berechnet.

Abkürzungsverzeichnis

ABS Anti-Blockier-System

ACC Adaptive Cruise Control

ADAS Advanced Driver Assistance Systems

ADTF Automotive Data and Time-Triggered Framework

ALKS Automated Lane Keeping System

ASAM Association for Standardization of Automation and Measuring Systems

ASE Automotive Systems Engineering

CDF Cumulative Distribution Function

CRMCAE Cubic-Root Mean Cubic Absolute Error

DDT Dynamic Driving Task

EBA Emergency Brake Assist

ECDF Empirical Cumulative Distribution Function

ESP Elektronisches Stabilitätsprogramm

FTS Foundation for Traffic Safety

HAF Hochautomatisiertes Fahren

HIL Hardware-in-the-Loop

KBA Kraftfahrtbundesamt

KL-Divergenz Kullback-Leibler-Divergenz

KPI Key Performance Indicator

KS-Test Kolmogorow-Smirnow-Test

LKA Lane Keep Assist

MCS Modified Cuckoo Search

MCS-LF Modified Cuckoo Search with Lévy Flight

MDF Measurement Data Format

MIL Model-in-the-Loop

MPC Model Predictive Control

ODD Operational Design Domain

OKR Objectives and Key Results

PEP Produktentstehungsprozess

RMSE Root Mean Squared Error

SAE Society of Automotive Engineers

SAV Sports Activity Vehicle

SIL Software-in-the-Loop

SOP Start of Production

SOTIF Safety of the intended functionality

SUT System Under Test

SWC Software Component

V&V Verifikation & Validierung

VDA Verband der Automobilindustrie

ViL Vehicle-in-the-Loop

XiL X-in-the-Loop

Abbildungsverzeichnis

1.1	Reduktion von Unfällen durch die Einführung eines serienmäßigen ESP-Systems in jedem Fahrzeug bei Mercedes-Benz im Jahr 1999 (in Anlehnung an [9])	2
1.2	McKinsey Umfrage: Einschätzung von 75 Experten der Transport- und Automobilindustrie aus dem Jahr 2021 zur kommerziellen Einführung von hochautomatisierten Fahrfunktionen (in Anlehnung an [13])	3
1.3	Serienverfügbarkeit von SAE Level 2+ und SAE Level 3 Systemen (Quelle: [16]).	4
1.4	Entwicklung der Steuergeräteanzahl und Codezeilen in Premiumfahrzeugen über die Zeit (Datenquellen: [?, 1, 24–31]).	5
2.1	Umsatzentwicklung eines Produkts über die Phasen des Produktlebenszyklus (Quelle: [38]).	14
2.2	Das Produktportfolio der Fahrzeugmarken des VW-Konzerns 2012 nach Fahrzeugklasse und Aufbauart (in Anlehnung an: [41]).	17
2.3	Äußere Erscheinung von Fahrzeugmodellen einer Modelllinie 2.3a und Fahrzeugmodellen einer Fahrzeugplattform 2.3b	19
2.4	Start eines neuen PEP für ein Fahrzeugmodell oder eine Modelllinie aus der Produktstrategie mit den Innovationen aus der Vorentwicklung. (in Anlehnung an [1])	22
2.5	Das V-Modell 97 (in Anlehnung an [69]).	26
2.6	Einteilung von Fahrerassistenzsystemen in drei Aufgabenklassen (in Anlehnung an [9])	29
2.7	Schematische Darstellung der Verarbeitungsschritte innerhalb der Wirkkette für ein Fahrerassistenzsystem (in Anlehnung an [23, 77])	31

2.8	Die 8 Reifegradstufen der Reifegradbewertung entlang der Phasen des PEP (in Anlehnung an [57]).	33
2.9	Systementwicklung entlang der Reifegradstufen im PEP (in Anlehnung an [57]).	34
3.1	Die zu betrachtenden Dimensionen zur Qualifikation von Fahrzeugen	37
3.2	Artefakte und Integrationsstufen im V-Modell (in Anlehnung an [85])	40
3.3	Das XiL-Prinzip für ADAS und HAF-Systeme. (in Anlehnung an Reisgys [71])	42
3.4	Verhältnis der Softwareentwicklung zu Applikation mit fortschreitendem Reifegrad der Funktion.	49
3.5	Für diese Dissertation relevante Teilschritte im Applikationsprozess in der Fahrerunterstützungsdomäne.	51
4.1	Schematische Darstellung der Decodierungsschritte für Signale einer Busnachricht am Beispiel der geographischen Länge und Breite.	57
4.2	Schematische Darstellung der Umwandlung zweier Nachrichten in einheitliche Zeitreihen-Daten mit 20ms Abtast-Intervallen.	58
5.1	Auswirkung einer Parameteränderung auf das SUT in verschiedenen Straßenverkehrstests	71
5.2	Permutation von bewerteten Straßenverkehrstests mit einer Gesamtlänge von ca. 3.750km. Eine Annäherung an die tatsächliche Durchschnittsbewertung ist erst bei Hinzunahme fast aller Fahrten gewährleistet.	73
5.3	Verteilung der Geschwindigkeit in km/h für den gesamten Datensatz sowie drei exemplarische Erprobungsfahrten	76
5.4	Verteilung der Abstände zum Vorderfahrzeug in m für den gesamten Datensatz sowie drei exemplarische Erprobungsfahrten	79
5.5	Verteilung der geltenden Tempolimits in km/h für den gesamten Datensatz sowie drei exemplarische Erprobungsfahrten	80
5.6	Verteilung der Straßenklasse für den gesamten Datensatz sowie drei exemplarische Erprobungsfahrten	81

5.7	Vergleich der Verteilung der Straßenkategorien im exemplarischen Erprobungsdatensatz mit der Verteilung im gesamten deutschen Straßennetz als Grundgesamtheit (Quelle für das Straßennetz Deutschlands: BMDV [120])	85
5.8	Vergleich der Verteilung der gefahrenen Distanz im exemplarischen Erprobungsdatensatz mit der Fahrleistung in Deutschland 2005 als Grundgesamtheit (Quelle für die Fahrleistung Deutschlands 2005: ifeu / bast [121, 122])	86
5.9	Verteilung der gefahrenen Distanz nach Straßenkategorie über verschiedene Länder in einem exemplarischen Erprobungsdatensatz .	87
5.10	Verteilung der gefahrenen Distanz auf ausgewählte Länder in einem exemplarischen Erprobungsdatensatz	88
6.1	Prozess der global-optimalen, datengetriebenen Applikation mit vier Prozessphasen (A-D) und den jeweiligen Prozessschritten.	100
6.2	Klassendiagramm zur Beschreibung der Datenstruktur von Datensätzen und ihrer Bedeutung im Applikationskontext.	104
6.3	Schematische Darstellung dreier Fahraufgaben in einem Straßenverkehrstest. Die Reaktion auf Einscherer und der Spurwechsel sind manöverartige Fahraufgaben wohingegen die Spurmittenführung kontinuierlich (außer während des Spurwechsels) aktiv ist.	108
6.4	Schematische Darstellung des Detektionsprinzips für manöverartige Fahraufgaben.	111
6.5	Schematische Darstellung der Segmentierung von Straßenverkehrstests. Die Ausprägungen der die Fahrzeugumgebung beschreibenden Attribute führen zur Segmentierung der Fahrt. Alle Zeitpunkte werden berücksichtigt. . .	113
6.6	Schematische Darstellung der Gruppierung von Fahrtabschnitten eines Szenarios durch die Subsznarien im Attributraum.	114
6.7	Häufigkeitsverteilung von Fahrtabschnitten eines Beispieldatensatzes im Attributraum der Geschwindigkeit und des Abstands zum Vorderfahrzeug.	119

6.8	Paarweise Häufigkeitsverteilung von Fahrtabschnitten eines Beispieldatensatzes für die Attribute Geschwindigkeit des Egofahrzeugs [km/h], Geschwindigkeit des Vorderfahrzeugs (VFZ) [km/h] und Distanz zum Vorderfahrzeug [m] sowie Spurbreite [m]. Auf der Diagonalen sind die eindimensionalen Häufigkeitsverteilungen der Attribute der X-Achse dargestellt.	122
6.9	Unterteilung des Attributraums der Attribute Geschwindigkeit des Vorderfahrzeugs und Distanz zum Vorderfahrzeug in Subszenarien. Dunkel hinterlegte Subszenarien sind im Datensatz nicht vorhanden.	123
6.10	Entscheidungsprozess für die Auswahl geeigneter Substitutionsverfahren.	125
6.11	Prozessschritte der global-optimalen, datengetriebenen Applikation. (Wdh. der Abbildung 6.1)	128
6.12	Dekomposition der Applikation des Gesamtsystems über die Fahraufgaben während der Definition des Applikationsumfangs (vgl. Abbildung 6.11 Prozessschritt A). Die Farbcodierung zeigt relevante Fahraufgaben für die exemplarischen ADAS bzw. HAF-Funktionen	129
6.13	Festlegung von relevanten Umgebungsinformationen für kontinuierliche Fahraufgaben am Beispiel der Spurmittenführung für ein ALKS und einen Autobahnpiloten.	130
6.14	Framework zur global-optimalen, datengetriebenen Applikation nach Abschluss der Applikationsvorbereitung (vgl. Abbildung 6.11).	131
6.15	Die Applikationsdurchführung (vgl. Abbildung 6.11) liefert Fahrzeugmessdaten von Straßenverkehrstests für die global-optimale, datengetriebene Applikation.	132
6.16	Prozessschritte der Phase Virtuelle Analyse (vgl. Abbildung 6.11).	134
6.17	Prozessschritte des Teilprozesses Applikationen analysieren (vgl. Abbildung 6.16).	135
6.18	Framework zur global-optimalen, datengetriebenen Applikation zum Prozessschritt Applikationen analysieren (vgl. Abbildung 6.17).	136
6.19	Detektionsprinzip für Fahrtabschnitte eines Straßenverkehrstests am Beispiel der unterschiedlichen Fahraufgaben einer längsregelnden Fahrfunktion im Bezug auf ein Vorderfahrzeug.	136

6.20	Prozessschritte des Teilprozesses Applikationen bewerten (vgl. Abbildung 6.16).	139
6.21	Prozessschritte und Artefakte des Teilprozesses Applikationen vergleichen (vgl. Abbildung 6.16).	141
7.1	Visualisierung der paarweisen Parameterraumabdeckung.	146
7.2	Visualisierung von Fahrtabschnitten (blau) im Szenario-Attributraum zweier Attribute. Jeder Fahrtabschnitt (Datenpunkt) stellt eine Instanz des Szenarios mit der auf der X- und Y-Achse abzulesenden Attributkombination dar.	148
7.3	Visualisierung der Aufteilung des Szenario-Attributraums in Subszenarios. Die Subszenarios werden nummeriert. Die Anzahl der Subszenarios ergibt sich aus der Multiplikation der Anzahl der beiden Attribut-Bins.	149
7.4	Visualisierung der Bewertung eines Subszenarios durch die Berechnung des Qualitätswerts aller Fahrtabschnitte des Subszenarios. Der Qualitätswert wird über der Anzahl der Fahrtabschnitte eines Subszenarios dargestellt.	150
7.5	Vollständige Visualisierung der Bewertung aller Subszenarios über den Mittelwert aller Fahrtabschnitte (blau) jedes Subszenarios. Für jedes SubszENARIO wird die Bewertung (obere Zahl + Farbkodierung) sowie die Anzahl Instanzen (untere Zahl) dargestellt.	151
7.6	Adaption des global-optimalen, datengetriebenen Applikationsprozesses auf eine virtuelle, iterative Applikation mittels aufgezeichneter Erprobungsfahrten.	151
8.1	Vergleich der Exploration des Parameterraums in den drei Applikationsdurchläufen mit dem Spanien-Trainingsdatensatz (vgl. Tabelle 7.1)	156
8.2	Entwicklung der besten Lösung über die Generationen in den drei Applikationsdurchläufen mit dem Spanien-Trainingsdatensatz (vgl. Tabelle 7.1)	157

8.3	Mittlere Abweichung im Attributraum der Ausgangsparametrierung sowie der jeweils besten Lösung der drei Applikationsstrategien mit dem Spanien-Validierungsdatensatz (vgl. Tabelle 7.1)	158
8.4	Mittlere Abweichung im Attributraum der Ausgangsparametrierung sowie der jeweils besten Lösung der drei Applikationsstrategien mit dem globalen Validierungsdatensatz. .	159
8.5	Mittlere KL-Divergenz (blau) und mittlerer RMSE (rot) über die Anzahl Fahrten in den zufällig gezogenen Datensätzen.	164
8.6	Relative Abdeckung aller Subszzenarien mit >0 (blau) bzw. >5 (rot) Kurven über die Anzahl Fahrten in den zufällig gezogenen Datensätzen.	165
8.7	Darstellung der besten 10 Lösungen als Pareto-Front über Verfügbarkeit und Qualität der Funktion.	167
A.1	Variante des PEP von Porsche Engineering mit sieben Phasen, den involvierten Unternehmensbereichen sowie erzeugten Prototypen. (in Anlehnung an [54])	179
A.2	Exemplarische Routenplanung für eine erste Etappe einer Erprobung .	180
A.3	Exemplarische Routenplanung für eine erste Etappe einer Erprobung unter Beachtung potentieller Bias durch die Straßenklasse. .	180
A.4	Exploration des Funktionsparameterraums in Versuch 1.	185
A.5	Abdeckung des Szenario-Attributraums in exemplarischen Lösungen des Versuch 1. Nicht vorhandene Subszzenarien sind grau hinterlegt und mit NaN bezeichnet. Für die anderen Subszzenarien ist die Bewertung des Subszenarios (obere Zahl) und die Anzahl vorhandener Kurven (untere Zahl) dargestellt. Die blauen Punkte zeigen die Datenpunkte der Kurven im Attributraum.	187
A.6	Anzahl der relevanten Kurven für jede Lösung in Versuch 1.	188
A.7	Abdeckung der Subszzenarien in exemplarischen Lösungen des Versuch 2	189
A.8	Untersuchung verschiedener Unterteilungen des Attributraums bezüglich der SubszENARIO-Abdeckung	190
A.9	Verteilung nach einzelnen Attributen in der Grundgesamtheit.	201

Tabellenverzeichnis

1.1	Die 6 Stufen des automatisierten Fahrens laut SAE J3016 [17].	3
2.1	Fahrzeugsegmente nach Europäischer Kommission [44] und KBA [45]	16
3.1	Die Markenleitbilder und assoziierten Eigenschaften der Marken des VW-Konzerns (Quelle: [41]).	46
5.1	Überblick über die drei verwendeten Erprobungsfahrten in der Voruntersuchung zur Repräsentativität.	75
5.2	Einflussfaktoren nach Kategorien aufgeteilt, die Bias in Erprobungsfahrten erzeugen.	83
6.1	Überblick über Ansätze zur Manöver- und Szenariendetektion.	110
6.2	Gegenüberstellung potentieller Datenquellen für die Ableitung der Grundgesamtheit.	117
6.3	Überblick über ausgewählte Metriken zur Bewertung einzelner Fahraufgaben.	139
6.4	Schematische Darstellung einer Applikations-Vergleichs-Tabelle. . . .	141
7.1	Übersicht über die vier verwendeten Erprobungen.	144
8.1	Ergebnisse der Validierung der jeweils besten Lösung unterschiedlicher Applikationsstrategien auf dem länderspezifischen Spanien-Validierungsdatensatz sowie dem globalen Validierungsdatensatz. Dabei wurden für die Validierung für jede Lösung die Metriken für den globalen Mittelwert, den Subszenario-Mittelwert sowie den gewichteten Subszenario-Mittelwert berechnet.	160

8.2	Berechnung verschiedener Statistiken für die Quantifizierung des Stichprobenfehlers in verschiedenen zufällig gezogenen Datensätzen. Für den KS-Test und den Chi ² -Test sind die Datensätze mit bestätigter Nullhypothese hervorgehoben. Für die KL-Divergenz sowie den RMSE die jeweils 5 zur Grundgesamtheit ähnlichsten Datensätze.	162
8.3	Auswertungstabelle mit den besten 10 Lösungen aus der Spanien-Applikation. Für jede Applikation wurde die Verfügbarkeit, Qualität sowie die SubszENARIOabdeckung berechnet. .	166
A.1	In der C++ Implementierung verwendete Hyperparameter der Modified Cuckoo Search	184
A.2	Berechnung der relativen Abdeckung aller SubszENARIEN mit >0 bzw. >5 Kurven über alle zufällig gezogenen Datensätze.	202
A.3	Ergebnisse der Validierung der jeweils besten Lösung unterschiedlicher Applikationsstrategien auf dem länderspezifischen Irland-Validierungsdatensatz sowie dem globalen Validierungsdatensatz. Dabei wurden für die Validierung für jede Lösung die Metriken für den globalen Mittelwert, den SubszENARIO Mittelwert sowie den gewichteten SubszENARIO Mittelwert berechnet.	203

Literaturverzeichnis

- [1] J. Weber, *Automotive development processes: Processes for successful customer oriented vehicle development*. Springer Science & Business Media, 2009.
- [2] H. Winner, G. Prokop, und M. Maurer, *Automotive Systems Engineering II*. Springer, 2018, Vol. 1.
- [3] E. Seifert, „Sieg der Elektronik: Das ABS-Bremssystem endlich serienreif,” *Arzt U Auto*, Vol. 54, Nr. 10, 1978.
- [4] A. Liu und J. Low, „Where is the L2+ ADAS market taking us?” 2023. [Online]. Verfügbar: <https://www.canalys.com/insights/l2-adas-market>
- [5] M. Aleksa, A. Schaub, I. Erdelean, S. Wittmann, A. Soteropoulos, und A. Fördös, „Impact analysis of Advanced Driver Assistance Systems (ADAS) regarding road safety—computing reduction potentials,” *European Transport Research Review*, Vol. 16, Nr. 1, S. 39, 2024.
- [6] L. Masello, G. Castignani, B. Sheehan, F. Murphy, und K. McDonnell, „On the road safety benefits of advanced driver assistance systems in different driving contexts,” *Transportation research interdisciplinary perspectives*, Vol. 15, S. 100670, 2022.
- [7] R. B. Naumann, L. K. Kreuger, L. Sandt, K. H. Lich, B. Bauchwitz, W. Kumfer, und T. Combs, „Examining the safety benefits of partial vehicle automation technologies in an uncertain future,” 2023.
- [8] Europäisches Parlament, „Verordnung (EU) 2019/2144 des Europäischen Parlaments und des Rates,” *EG*, Vol. 2144, S. 2019, 2019.

- [9] J. Schwarz, „Code of Practice for development, validation and market introduction of ADAS,” in 2. *Tagung Aktive Sicherheit durch Fahrerassistenz*, 2006.
- [10] NHTSA, „Automated Driving Systems 2.0: A Vision for Safety,” Tech. Rep., 2017. [Online]. Verfügbar: https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/13069a-ads2.0_090617_v9a_tag.pdf
- [11] T. Warren, „The future of America ist driverless,” 2016. [Online]. Verfügbar: <https://www.theverge.com/a/verge-2021/secretary-anthony-foxx>
- [12] J. Dujmovic, „Opinion: You will not be traveling in a self-driving car anytime soon. Here’s what the future will look like,” 2021. [Online]. Verfügbar: <https://www.marketwatch.com/story/you-will-not-be-traveling-in-a-self-driving-car-anytime-soon-heres-what-the-future-will-look-like-11623866219>
- [13] K. Heineke, R. Heuss, A. Kelkar, und M. Kellner, „What’s next for autonomous vehicles?” *McKinsey Center for Future Mobility*, S. 1–8, 2021.
- [14] Waymo, „Sharing our safety framework for fully autonomous operations,” Tech. Rep., 2020. [Online]. Verfügbar: <https://waymo.com/blog/2020/10/sharing-our-safety-framework>
- [15] ———, „Waymo One: The future of transportation is here.” Tech. Rep., 2025. [Online]. Verfügbar: <https://waymo.com/waymo-one/>
- [16] L. Block, F. Herrmann, J. Wizl, D. Borrmann, S. Bratzel, und F. Böbber, „Deutschland zum Innovations-standort für das automatisierte und vernetzte Fahren machen,” 2024.
- [17] SAE International, „Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles j3016 202104,” *Society of Automotive Engineers*, Vol. 41, 2021.

- [18] F. Greis, „Staupilot von Mercedes verzögert sich offenbar,” Tech. Rep., 2021. [Online]. Verfügbar: <https://www.golem.de/news/autonomes-fahren-staupilot-von-mercedes-verzoegert-sich-offenbar-2111-160923.html>
- [19] Mercedes-Benz, „Hochautomatisiertes Fahren: Erste international gültige Systemgenehmigung,” Tech. Rep., 2021. [Online]. Verfügbar: <https://group.mercedes-benz.com/innovation/produktinnovation/autonomes-fahren/systemgenehmigung-fuer-hochautomatisiertes-fahren.html>
- [20] BMW Group, „Ab Frühjahr: Hochautomatisiertes Fahren auf Level 3 im neuen BMW 7er.” Tech. Rep., 2023. [Online]. Verfügbar: https://www.bmw.de/de/shop/ls/dp/BMW_Personal_Pilot_de
- [21] H. Schmidt, „Audi steckt beim autonomen Fahren zurück,” 2020. [Online]. Verfügbar: <https://www.nzz.ch/mobilitaet/auto-mobil/autonomes-fahren-stufe-3-audi-verzichtet-im-a8-auf-staupilot-ld.1553933>
- [22] A. Ziebinski, R. Cupek, H. Erdogan, und S. Waechter, „A survey of ADAS technologies for the future perspective of sensor fusion,” in *International Conference on Computational Collective Intelligence*. Springer, 2016, S. 135–146.
- [23] D. Weitzel, H. Winner, C. Peng, S. Geyer, F. Lotz, und M. Sefati, „Absicherungsstrategien für Fahrerassistenzsysteme mit Umfeldwahrnehmung: Bericht zum Forschungsprojekt FE 82.0546/2012,” *Berichte der Bundesanstalt für Straßenwesen Fahrzeugtechnik*, 2014.
- [24] M. Türker und H. Junker, „Der Schlüssel zu effektivem Testing,” *Porsche Engineering Magazin*, Vol. 1, Nr. 1, S. 20–25, 2018.
- [25] M. Berret, „Global Automotive Supplier Study 2018,” *RolandBerger*, 2017.
- [26] J. Ebberg, „Wussten Sie schon, dass...?” Tech. Rep., 2020. [Online]. Verfügbar: <https://www.bosch-presse.de/pressportal/de/de/wussten-sie-schon-dass...-216258.html>

- [27] V. Antinyan, „Revealing the complexity of automotive software,” in *Proceedings of the 28th acm joint meeting on european software engineering conference and symposium on the foundations of software engineering*, 2020, S. 1525–1528.
- [28] T. Häckel, „A Middleware Solution for Open and Dynamic ICT Architectures in Future Cars,” 2018.
- [29] R. N. Charette, „This car runs on code,” *IEEE spectrum*, Vol. 46, Nr. 3, S. 3, 2009.
- [30] C. Buckl, A. Camek, G. Kainz, C. Simon, L. Mercep, H. Stähle, und A. Knoll, „The software car: Building ICT architectures for future electric vehicles,” in *2012 IEEE International Electric Vehicle Conference*. IEEE, 2012, S. 1–8.
- [31] M. Broy, I. H. Kruger, A. Pretschner, und C. Salzmann, „Engineering automotive software,” *Proceedings of the IEEE*, Vol. 95, Nr. 2, S. 356–373, 2007.
- [32] W. Wachenfeld und H. Winner, „Die Freigabe des autonomen Fahrens,” *Autonomes Fahren: technische, rechtliche und gesellschaftliche Aspekte*, S. 439–464, 2015.
- [33] B. Lenz und E. Fraedrich, „Gesellschaftliche und individuelle Akzeptanz des autonomen Fahrens,” *Autonomes Fahren: Technische, rechtliche und gesellschaftliche Aspekte*, S. 639–660, 2015.
- [34] K.-L. Bauer, „Echtzeit-Strategieplanung für vorausschauendes automatisiertes Fahren,” Ph.D. dissertation, Karlsruher Institut für Technologie (KIT), 2020.
- [35] N. Fraikin, *Methodik zur effizienten Applikation automatisierter Fahrfunktionen*. KIT Scientific Publishing, 2024.
- [36] C. Davey, „Automotive software systems complexity: challenges and opportunities,” in *INCOSE International Workshop MBSE Workshop*, 2013.

- [37] D. J. Gebauer, „Ein modellbasiertes, graphisch notiertes, integriertes Verfahren zur Bewertung und zum Vergleich von Elektrik/Elektronik-Architekturen,” Ph.D. dissertation, Karlsruher Institut für Technologie (KIT), 2016.
- [38] R. Olbrich, *Marketing: eine Einführung in die marktorientierte Unternehmensführung*. Springer-Verlag, 2006.
- [39] ISO, *IEC 12207 Systems and software engineering-software life cycle processes*, International Organization for Standardization, 2008.
- [40] G. Schicker und J. Strassl, *Produktportfolio-Management im Zeitalter der Digitalisierung*, Serie Weidener Diskussionspapiere. Ostbayerische Technische Hochschule Amberg-Weiden (OTH), Weiden i.d.OPf., 2019, Nr. 69.
- [41] M. Winterkorn, „Die Bedeutung der Marke in der Automobilindustrie und die Mehrmarkenstrategie des Volkswagen Konzerns,” 2012. [Online]. Verfügbar: <https://docplayer.org/9038526-Die-bedeutung-der-marke-in-der-automobilindustrie-und-die-mehrmarkenstrategie-des-volkswagen-konzerns-prof-dr-martin-winterkorn-16.html>
- [42] Kraftfahrtbundesamt, „Glossar: Segment.” [Online]. Verfügbar: https://www.kba.de/DE/Service/Glossar/glossar_node.html
- [43] —, „Methodische Erläuterungen zu Statistiken über Fahrzeugzulassungen (FZ),” 2018. [Online]. Verfügbar: https://www.kba.de/DE/Statistik/Fahrzeug/fz_methodik/fz_methodische_erlaeueterungen_201810_pdf
- [44] Europäische Kommission, „Kraftfahrzeugbetrieb und -kundendienst in der Europäischen Union; Verordnung (EG) Nr. 1400/2002 der Kommission vom 31. Juli 2002,” 2002. [Online]. Verfügbar: https://ec.europa.eu/competition/sectors/motor_vehicles/legislation/explanatory_brochure_de.pdf
- [45] Kraftfahrtbundesamt, „Fahrzeugzulassungen (FZ) Bestand an Personenkraftwagen nach Segmenten und Modellreihen am 1. Januar 2021 gegenüber 1. Januar 2020,” 2021. [Online]. Verfügbar: https://www.kba.de/DE/Statistik/Fahrzeuge/Bestand/Segmente/segmente_node.html

- [46] M. Unger, *Die Automobil-Kaufentscheidung: ein theoretischer Erklärungsansatz und seine empirische Überprüfung*. Peter Lang International Academic Publishers, 2018.
- [47] C. Rose, „Der Baukasten hinter dem Elfer-Code,” 2011. [Online]. Verfügbar: <https://www.auto-motor-und-sport.de/news/porsche-911-997-modellueberblick-der-baukasten-hinter-dem-elfer-code/>
- [48] Reddit, „One of the main strengths of Volkswagen Group is its ability to save costs through synergies among the brands. The group has currently 17 different platforms for 95 models and still wants to reduce the former with electrification.” 2020. [Online]. Verfügbar: https://www.reddit.com/r/electricvehicles/comments/k3lrfp/one_of_the_main_strengths_of_volkswagen_group_is/
- [49] ISO, *IEEE standard glossary of software engineering terminology (IEEE Std 610.12-1990)*, International Organization for Standardization, 1990.
- [50] C. Hussung, J.-R. Rehse, C. Houy, und P. Fettke, „Entwicklung eines Referenzprozessmodells für Rettungseinsätze der Feuerwehr und Anwendung als Grundlage eines Prozessassistenzsystems.” in *Wirtschaftsinformatik (Zentrale Tracks)*, 2020, S. 522–537.
- [51] J. Feldhusen und K.-H. Grote, „Der Produktentstehungsprozess (PEP),” in *Pahl/Beitz Konstruktionslehre*. Springer, 2013, S. 11–24.
- [52] P. O’Connor, „Implementing a stage-gate process: a multi-company perspective,” *Journal of Product Innovation Management*, Vol. 11, Nr. 3, S. 183–200, 1994.
- [53] R. G. Cooper und E. J. Kleinschmidt, „Stage-gate process for new product success,” *Innovation Management U*, Vol. 3, S. 2001, 2001.
- [54] S. Rudert und J. Trumpfheller, „Vollumf”änglich durchdacht: Der Produktentstehungsprozess,” *Porsche Engineering Magazin*, Vol. 1, Nr. 1, S. 10–13, 2015.

- [55] Volkswagen AG, „Volkswagen richtet Technische Entwicklung neu aus: mehr Tempo bei Produktzyklen und digitalen Angeboten,” Tech. Rep., 2022. [Online]. Verfügbar: <https://www.volkswagen-newsroom.com/de/pressemittellungen/volkswagen-richtet-technische-entwicklung-neu-aus-mehr-tempo-bei-produktzyklen-und-digitalen-angeboten-7768>
- [56] F. Hubik und S. Menzel, „So beschleunigen die Autohersteller die Entwicklung neuer Modelle,” 2022. [Online]. Verfügbar: <https://www.handelsblatt.com/unternehmen/industrie/volkswagen-mercedes-und-co-so-beschleunigen-die-autohersteller-die-entwicklung-neuer-modelle/28268396.html>
- [57] Verband der Automobilindustrie e. V., „Reifegradabsicherung für Neuteile,” VDA, Tech. Rep., 2021.
- [58] J. Bach, „Methoden und Ansätze für die Entwicklung und den Test prädiktiver Fahrzeugregelungsfunktionen,” Ph.D. dissertation, Karlsruher Institut für Technologie (KIT), 2018.
- [59] M. H. Breitner, „Vorgehensmodell,” 2019. [Online]. Verfügbar: <http://www.enzyklopaedie-der-wirtschaftsinformatik.de/lexikon/is-management/Systementwicklung/Vorgehensmodell>
- [60] E. B. Pereira, R. M. Bastos, und T. C. Oliveira, „A systematic approach to process tailoring,” in *2007 International Conference on Systems Engineering and Modeling*. IEEE, 2007, S. 71–78.
- [61] W. W. Royce, „Managing the development of large software systems: concepts and techniques,” in *Proceedings of the 9th international conference on Software Engineering*. IEEE Computer Society Press, 1987, S. 328–338.
- [62] B. W. Boehm, „A spiral model of software development and enhancement,” *Computer*, Nr. 5, S. 61–72, 1988.
- [63] K. Schwaber und M. Beedle, *Agile software development with Scrum*. Prentice Hall Upper Saddle River, 2002, Vol. 1.

- [64] A. Gebhardt, „Rapid Prototyping: Werkzeuge für die schnelle Produktentwicklung,” 1996.
- [65] P. Abrahamsson, O. Salo, J. Ronkainen, und J. Warsta, „Agile Software Development Methods: Review and Analysis,” *VTT*, 2002.
- [66] T. Dybå und T. Dingsøy, „Empirical studies of agile software development: A systematic review,” *Information and software technology*, Vol. 50, Nr. 9-10, S. 833–859, 2008.
- [67] B. W. Boehm, „Verifying and validating software requirements and design specifications,” *IEEE Software*, Vol. 1, Nr. 1, S. 75, 1984.
- [68] E. Sax, *Automatisiertes Testen Eingebetteter Systeme in der Automobilindustrie*. Hanser, 2008.
- [69] W. Dröschel und M. Wiemers, *Das V-Modell 97: der Standard für die Entwicklung von IT-Systemen mit Anleitung für den Praxiseinsatz*. Walter de Gruyter GmbH & Co KG, 2015.
- [70] VDA QMC Working Group 13 / Automotive SIG, „Automotive SPICE Process Assessment / Reference Model,” 2017.
- [71] F. Reisgys, J. Plaum, A. Schwarzhaupt, und E. Sax, „Scenario-based X-in-the-Loop Test for Development of Driving Automation,” in *14. Workshop Fahrerassistenzsysteme und Automatisiertes Fahren*, 2022.
- [72] T. Hummel, M. Kühn, J. Bende, und A. Lang, „Advanced driver assistance systems: an investigation of their potential safety benefits based on an analysis of insurance claims in Germany,” *German Insurance Association Insurers Accident Research, Research Report FS*, Vol. 3, 2011.
- [73] M. Kuehn, T. Hummel, und J. Bende, „Benefit estimation of advanced driver assistance systems for cars derived from real-life accidents,” in *21st International Technical Conference on the Enhanced Safety of Vehicles ESV*, Vol. 15, 2009, S. 18.

- [74] ISO, *ISO/PAS 21448:2019 Road vehicles - Safety of the intended functionality*, International Organization for Standardization, 2019.
- [75] O. Y. Al-Jarrah, C. Maple, M. Dianati, D. Oxtoby, und A. Mouzakitis, „Intrusion detection systems for intra-vehicle networks: A review,” *Ieee Access*, Vol. 7, S. 21 266–21 289, 2019.
- [76] A. Conti, „Fahrassistenzsysteme auf dem Prüfstand: Wie viel Sensorik braucht ein Auto?” *ATZelextronik*, Vol. 18, Nr. 3, S. 62–62, 2023.
- [77] J. Börcsök, „Funktionale Sicherheit,” *Grundzüge sicherheitstechnischer Systeme*, Vol. 2, 2006.
- [78] Europäische Union, „Commission Implementing Regulation (EU) 2022/1426 laying down rules for the application of Regulation (EU) 2019/2144 of the European Parliament and of the Council as regards uniform procedures and technical specifications for the type-approval of the automated driving system (ADS) of fully automated vehicles,” *Online: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF>*, 2022.
- [79] ISO, *ISO/DIS 34501 - Road vehicles — Test scenarios for automated driving systems — Vocabulary*, International Organization for Standardization, 2022.
- [80] S. Schmerler, *Softwaretest in der Praxis: Grundlagen, Methoden und Technologien*. epubli, 2020.
- [81] J. Schäuuffele und T. Zurawka, *Automotive software engineering*. Springer, 2005.
- [82] ISO, *26262: Road vehicles-Functional safety*, International Organization for Standardization, 2011.
- [83] —, *Systems and software engineering-Vocabulary*, International Organization for Standardization, 2010.
- [84] —, *ISO/DIS 34501 - Road vehicles — Terms and definitions of test scenarios for automated driving systems*, International Organization for Standardization, 2022.

- [85] C. King, „Bewertung von Fahrerassistenzsystemen im Umfeld des szenariobasierten Testens,” Ph.D. dissertation, Karlsruher Institut für Technologie, 2023.
- [86] M. Wood, P. Robbel, M. Maass, R. D. Tebbens, M. Meijs, M. Harb, und P. Schlicht, „Safety First for Automated Driving,” Verband der Automobilindustrie e. V. (VDA), Tech. Rep., 2019.
- [87] T. Düser, S. Ott, und A. Albers, „X-in-the-Loop als integrierte Entwicklungsumgebung von komplexen Antriebssystemen,” *Tagung Hardware-in-the-loop-Simulation, Kassel*, Vol. 16, S. 17, 2008.
- [88] H. Shokry und M. Hinchey, „Model-Based Verification of Embedded Software.” *Computer*, Vol. 42, Nr. 4, S. 53–59, 2009.
- [89] T. Bock, K.-H. Siedersberger, M. Zavrel, A. Breu, und M. Maurer, „Simulations-und Testumgebung fuer Fahrerassistenzsysteme/Simulation and test environment for driver assistance systems,” *VDI-Berichte*, Nr. 1900, 2005.
- [90] N. Guhl, „Erprobungsdienstleistungen von Porsche Engineering,” *Porsche Engineering Magazine*, Vol. 1, Nr. 1, 2016.
- [91] J. Golson, „Tesla’s new Autopilot will run in ‘shadow mode’ to prove that it’s safer than human driving,” 2016. [Online]. Verfügbar: <https://www.theverge.com/2016/10/19/13341194/tesla-autopilot-shadow-mode-autonomous-regulations>
- [92] A. Karpathy, „Tesla Autonomy Day,” 2019. [Online]. Verfügbar: <https://www.youtube.com/watch?v=Ucp0TTmvqOE>
- [93] B. Templeton, „Tesla’s ‘Shadow’ Testing Offers A Useful Advantage On The Biggest Problem In Robocars,” 2019. [Online]. Verfügbar: <https://www.forbes.com/sites/bradtempleton/2019/04/29/teslas-shadow-testing-offers-a-useful-advantage-on-the-biggest-problem-in-robocars/?sh=5d2729c3c06c>

- [94] B. Schick, F. Fuhr, M. Höfer, und P. E. Pfeffer, „Eigenschaftsbasierte entwicklung von fahrerassistenzsystemen,“ *ATZ-Automobiltechnische Zeitschrift*, Vol. 121, Nr. 4, S. 70–75, 2019.
- [95] G. Naus, J. Ploeg, M. Van de Molengraft, W. Heemels, und M. Steinbuch, „Design and implementation of parameterized adaptive cruise control: An explicit model predictive control approach,“ *Control Engineering Practice*, Vol. 18, Nr. 8, S. 882–892, 2010.
- [96] F. Lindlar, „Modellbasierter evolutionärer Funktionstest,“ Ph.D. dissertation, Berlin Institute of Technology, 2012.
- [97] F. Schuldt, „Ein Beitrag für den methodischen Test von automatisierten Fahrfunktionen mit Hilfe von virtuellen Umgebungen,“ Ph.D. dissertation, Technische Universität Braunschweig, 2017.
- [98] E. Novoseltseva, „Data Driven Development Elements and Benefits,“ 2021. [Online]. Verfügbar: <https://apiumhub.com/tech-blog-barcelona/data-driven-development-elements-benefits/>
- [99] M. Fenner, „Data-Driven Development,“ 2015. [Online]. Verfügbar: <https://doi.org/10.5438/DHSM-8219>
- [100] Robert Bosch GmbH, „Multi purpose camera,“ Tech. Rep., 2024. [Online]. Verfügbar: https://www.bosch-mobility.com/media/global/solutions/passenger-cars-and-light-commercial-vehicles/driver-assistance-systems/multi-camera-system/multi-purpose-camera/summary_multi-purpose-camera_en.pdf
- [101] ASAM, *ASAM MDF Measurement Data Format*, ASAM e.V., 2019.
- [102] P. Deppe, „Mercedes-Benz Passion: Mercedes-Benz präsentiert in Genf Limousine und Coupe der neuen E-Klasse,“ 2009. [Online]. Verfügbar: <https://blog.mercedes-benz-passion.com/2009/03/mercedes-benz-prasentiert-in-genf-limousine-und-coupe-der-neuen-e-klasse/>

- [103] M. Christiansen, „Mercedes-Benz: In geheimer Mission auf Abnahmefahrt mit der neuen Mercedes E-Klasse W213,” 2015. [Online]. Verfügbar: <https://5komma6.mercedes-benz-passion.com/in-geheimer-mission-auf-abnahmefahrt-mit-der-neuen-mercedes-e-klasse-w213/>
- [104] J. Hofmockel, „Anomalieerkennung in Kommunikationsdaten zur Datenselektion im Fahrzeug,” Ph.D. dissertation, Karlsruher Institut für Technologie (KIT), 2019.
- [105] BMW Group Corporate Communications, „The new BMW Group High Performance D3 platform. Data-Driven Development for Autonomous Driving.” 2019. [Online]. Verfügbar: <https://www.press.bmwgroup.com/global/article/attachment/T0293764EN/427468>
- [106] C. Gillies, P. Wustmann, und D. J. Schaper, „Der große Kreis,” *Porsche Engineering Magazine*, Vol. 1, Nr. 2, 2021.
- [107] J. Bach, K.-L. Bauer, M. Holzäpfel, M. Hillenbrand, und E. Sax, „Control based driving assistant functions test using recorded in field data,” in *7. Tagung Fahrerassistenzsysteme*, 2015. [Online]. Verfügbar: <https://mediatum.ub.tum.de/node?id=1285215>
- [108] S. Fünfgeld, M. Holzäpfel, M. Frey, und F. Gauterin, „Driver state estimation for prediction of vehicle states within control systems,” in *Intelligent Vehicles Symposium (IV), 2016 IEEE*. IEEE, 2016, S. 1386–1391.
- [109] J. Bach, S. Otten, M. Holzäpfel, und E. Sax, „Reactive-Replay approach for verification and validation of closed-loop control systems in early development,” in *SAE Technical Paper 2017-01-1671*, 2017.
- [110] H. Holland und K. Scharnbacher, *Grundlagen der Statistik*. Springer, 1991.
- [111] statista, „Definition Grundgesamtheit,” 2022. [Online]. Verfügbar: <https://de.statista.com/statistik/lexikon/definition/60/grundgesamtheit/>

- [112] Duden, *Duden – Deutsches Universalwörterbuch*, 6. Aufl. Mannheim: Bibliographisches Institut, Okt. 2006.
- [113] K. Burtscher, M. Hartel, R. Raml, und M. Koller, „Stichprobe und Grundgesamtheit,” 2022. [Online]. Verfügbar: https://www.mathe-online.at/lernpfade/hypothesen_testen/?kapitel=1
- [114] N. Döring und J. Bortz, *Forschungsmethoden und Evaluation*. Springer, 2016.
- [115] D. V. Bosch, „Repräsentativität von Stichproben,” 2012. [Online]. Verfügbar: <https://www.marktforschung.de/marktforschung/a/repraesentativitaet-von-stichproben/>
- [116] V. Benning, „Den Standardfehler des Mittelwertes verstehen und berechnen,” 2020. [Online]. Verfügbar: <https://www.scribbr.de/statistik/standardfehler/>
- [117] J. Hedderich und L. Sachs, *Angewandte Statistik*. Springer, 2016.
- [118] L. Sachs, *Angewandte Statistik: Anwendung statistischer Methoden*. Springer, 2013.
- [119] S. Kullback und R. A. Leibler, „On information and sufficiency,” *The annals of mathematical statistics*, Vol. 22, Nr. 1, S. 79–86, 1951.
- [120] BMDV, „Straßennetz (2022),” 2022. [Online]. Verfügbar: <https://bmdv.bund.de/SharedDocs/DE/Artikel/G/infrastruktur-statistik.html>
- [121] T. Kathmann, H. Ziegler, und B. Thomas, „Straßenverkehrszählung 2005: Ergebnisse,” 2007.
- [122] M. Jamet und W. Knörr, „Überleitung der Fahrleistungserhebung nach TREMOD,” 2022.
- [123] H. Oschlies, F. Saust, und S. Schmidt, „Virtuelle Applikation einer modellprädiktiven Regelung am Beispiel einer automatisierten Fahrzeugquerrführung,” in 8. *Tagung Fahrerassistenz*, München, 2017.

- [124] L. Liesner, *Automatisierte Funktionsoptimierung von Adaptive Cruise Control*. Shaker Verlag, 2017.
- [125] T. Pawellek, L. Liesner, R. Henze, und F. Küçükay, „Objektivierungsverfahren für eine kundenoptimale ACC-Systemauslegung,” *ATZ-Automobiltechnische Zeitschrift*, Vol. 118, Nr. 4, S. 74–79, 2016.
- [126] N. Fraikin, K. Funk, M. Frey, und F. Gauterin, „Dimensionality reduction and identification of valid parameter bounds for the efficient calibration of automated driving functions,” *Automotive and Engine Technology*, Vol. 4, S. 75–91, 2019.
- [127] —, „Efficient simulation based calibration of automated driving functions based on sensitivity based optimization,” *IEEE Open Journal of Intelligent Transportation Systems*, Vol. 1, S. 63–79, 2020.
- [128] B. Schick, S. Resch, M. Yamamoto, I. Kushiro, und N. Hagiwara, „Optimization of steering behavior through systematic implementation of customer requirements in technical targets on the basis of quality function deployment,” *Yokohama/Japan: FISITA*, 2006.
- [129] R. Bruder, K. Fuchs, B. Abendroth, und T. Schramm, „„Zur Aussagekraft menschbezogener Messgrößen in Fahrversuchen “,” *Wie realitätsnah lässt sich Fahrerverhalten messen*, S. 1–16, 2009.
- [130] J. Holzinger, P. Schöggel, M. Schrauf, und E. Bogner, „Objektivierung der Fahrbarkeit beim automatisierten Fahren,” *ATZ-Automobiltechnische Zeitschrift*, Vol. 116, Nr. 12, S. 48–53, 2014.
- [131] S. Solmaz, M. Rudigier, M. Mischinger, und J. Reckenzaun, „Hybrid testing: A vehicle-in-the-loop testing method for the development of automated driving functions,” *SAE International Journal of Connected and Automated Vehicles*, Vol. 4, Nr. 12-04-01-0011, S. 133–148, 2021.
- [132] C. Roesener, F. Fahrenkrog, A. Uhlig, und L. Eckstein, „A scenario-based assessment approach for automated driving by using time series classification

- of human-driving behaviour,” in *2016 IEEE 19th international conference on intelligent transportation systems (ITSC)*. IEEE, 2016, S. 1360–1365.
- [133] H. Elrofai, D. Worm, und O. Op den Camp, „Scenario identification for validation of automated driving functions,” in *Advanced Microsystems for Automotive Applications 2016: Smart Systems for the Automobile of the Future*. Springer, 2016, S. 153–163.
- [134] M. Barbier, C. Laugier, O. Simonin, und J. Ibanez-Guzman, „Classification of drivers manoeuvre for road intersection crossing with synthethic and real data,” in *2017 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2017, S. 224–230.
- [135] L. Hartjen, R. Philipp, F. Schuldt, B. Friedrich, und F. Howar, „Classification of driving maneuvers in urban traffic for parametrization of test scenarios,” in *9. Tagung Automatisiertes Fahren*, 2019.
- [136] A. Erdogan, B. Ugranli, E. Adali, A. Sentas, E. Mungan, E. Kaplan, und A. Leitner, „Real-world maneuver extraction for autonomous vehicle validation: A comparative study,” in *2019 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2019, S. 267–272.
- [137] F. Hauer, I. Gerostathopoulos, T. Schmidt, und A. Pretschner, „Clustering traffic scenarios using mental models as little as possible,” in *2020 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2020, S. 1007–1012.
- [138] R. Pfeffer, „Szenariobasierte simulationsgestützte funktionale Absicherung hochautomatisierter Fahrfunktionen durch Nutzung von Realdaten,” Ph.D. dissertation, Karlsruher Institut für Technologie, 2020.
- [139] T. Braun, L. Ries, M. Hesche, S. Otten, und E. Sax, „Maneuver-based Visualization of Similarities between Recorded Traffic Scenarios.” in *DATA*, 2022, S. 236–244.
- [140] Central Intelligence Agency, *The World Factbook 2020*. Central Intelligence Agency, 2020.

- [141] NHTSA, „Traffic Safety Facts Annual Report Tables - National Statistics,” Tech. Rep., 2022. [Online]. Verfügbar: <https://cdan.dot.gov/tsftables/National%20Statistics.pdf>
- [142] P. Junietz, F. Bonakdar, B. Klamann, und H. Winner, „Criticality metric for the safety validation of automated driving using model predictive trajectory optimization,” in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2018, S. 60–65.
- [143] L. Westhofen, C. Neurohr, T. Koopmann, M. Butz, B. Schütt, F. Utesch, B. Neurohr, C. Gutenkunst, und E. Böde, „Criticality metrics for automated driving: A review and suitability analysis of the state of the art,” *Archives of Computational Methods in Engineering*, Vol. 30, Nr. 1, S. 1–35, 2023.
- [144] Y. Lin und M. Althoff, „CommonRoad-CriMe: A Toolbox for Criticality Measures of Autonomous Vehicles,” in *Proc. of the IEEE Intell. Veh. Symp.*, 2023.
- [145] X.-S. Yang und S. Deb, „Cuckoo search via Lévy flights,” in *2009 World congress on nature & biologically inspired computing (NaBIC)*. IEEE, 2009, S. 210–214.
- [146] X.-S. Yang, *Nature-inspired metaheuristic algorithms*. Luniver press, 2010.
- [147] S. Walton, O. Hassan, K. Morgan, und M. Brown, „Modified cuckoo search: a new gradient free optimisation algorithm,” *Chaos, Solitons & Fractals*, Vol. 44, Nr. 9, S. 710–718, 2011.

Eigene Veröffentlichungen

- [JL1] Johannes Bach, Jacob Langner, Stefan Otten, Marc Holzäpfel, and Eric Sax. Data-driven development, a complementing approach for automotive systems engineering. In *2017 IEEE International Symposium on Systems Engineering (ISSE)*, pages 1–6. IEEE, 2017.
- [JL2] Johannes Bach, Jacob Langner, Stefan Otten, Marc Holzäpfel, and Eric Sax. Test scenario selection for system-level verification and validation of geolocation-dependent automotive control systems. In *2017 International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, pages 203–210. IEEE, 2017.
- [JL3] Philip Elspas, Jacob Langner, Michael Aydinbas, Johannes Bach, and Eric Sax. Leveraging regular expressions for flexible scenario detection in recorded driving data. In *2020 IEEE International Symposium on Systems Engineering (ISSE)*, pages 1–8. IEEE, 2020.
- [JL4] Lukas Ewecker, Niklas Wagner, Tim Brühl, Robin Schwager, Tin Stri-bor Sohn, Alexander Engelsberger, Jensun Ravichandran, Hanno Stage, Jacob Langner, and Sascha Saralajew. Detecting oncoming vehicles at night in urban scenarios-an annotation proof-of-concept. In *2024 IEEE Intelligent Vehicles Symposium (IV)*, pages 2117–2124. IEEE, 2024.
- [JL5] Lukas Haas, Arsalan Haider, Rolf Jung, Tim Poguntke, Bernhard Schick, Stefan-Alexander Schneider, Florence Wagner, Thomas Zeh, Martin Benedikt, Eckard Böde, Jacob Langner, et al. Controlling risk for highly automated transportation systems operating in complex open environments: A white paper of the safetrans closing the gap initiative. 2024.

- [JL6] Christian King, Thilo Braun, Constantin Braess, Jacob Langner, and Eric Sax. Capturing the variety of urban logical scenarios from bird-view trajectories. In *VEHITS*, pages 471–480, 2021.
- [JL7] Christian King, Lennart Ries, Jacob Langner, and Eric Sax. A taxonomy and survey on validation approaches for automated driving systems. In *2020 IEEE International Symposium on Systems Engineering (ISSE)*, pages 1–8. IEEE, 2020.
- [JL8] Jacob Langner, Johannes Bach, Stefan Otten, Eric Sax, and Marc Holzäpfel. Framework for using real driving data in automotive feature development and validation. In *8. Tagung Fahrerassistenz*, München, 2017.
- [JL9] Jacob Langner, Johannes Bach, Lennart Ries, Stefan Otten, Marc Holzäpfel, and Eric Sax. Estimating the uniqueness of test scenarios derived from recorded real-world-driving-data using autoencoders. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, pages 1860–1866. IEEE, 2018.
- [JL10] Jacob Langner, Kai-Lukas Bauer, Marc Holzäpfel, and Eric Sax. A process reference model for the virtual application of predictive control features. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, pages 1759–1764. IEEE, 2020.
- [JL11] Jacob Langner, Hannes Grolig, Stefan Otten, Marc Holzäpfel, and Eric Sax. Logical scenario derivation by clustering dynamic-length-segments extracted from real-world-driving-data. In *VEHITS*, pages 458–467, 2019.
- [JL12] Jacob Langner, Romy Pohl, Joshua Ransiek, Philip Elspas, and Eric Sax. Statistical consideration of the representativeness of open road tests for the validation of automated driving systems. In *2023 IEEE International Automated Vehicle Validation Conference (IAVVC)*, pages 1–8. IEEE, 2023.

- [JL13] Jacob Langner, Christian Seiffer, Stefan Otten, Kai-Lukas Bauer, Marc Holzäpfel, and Eric Sax. Qualitative feature assessment for longitudinal and lateral control-features. In *VEHITS*, pages 115–122, 2020.
- [JL14] Hubert Padusinski, Christian Steinhauser, Christian Scherl, Julian Gaal, and Jacob Langner. Point cloud recombination: Systematic real data augmentation using robotic targets for lidar perception validation. *arXiv preprint arXiv:2505.02476*, 2025.
- [JL15] Patrick Petersen, Jacob Langner, Stefan Otten, Eric Sax, Stefan Scheubner, Moritz Vaillant, Sebastian Fünfgeld, and F Porsche. Validation of range estimation for electric vehicles based on recorded real-world driving data. In *19. Internationales Stuttgarter Symposium*, pages 331–344. Springer Vieweg, Wiesbaden, 2019.
- [JL16] Patrick Petersen, Hanno Stage, Jacob Langner, Lennart Ries, Philipp Rigoll, Carl Philipp Hohl, and Eric Sax. Towards a data engineering process in data-driven systems engineering. In *2022 IEEE International Symposium on Systems Engineering (ISSE)*, pages 1–8. IEEE, 2022.
- [JL17] Joshua Ransiek, Johannes Plaum, Jacob Langner, and Eric Sax. Goose: Goal-conditioned reinforcement learning for safety-critical scenario generation. *arXiv preprint arXiv:2406.03870*, 2024.
- [JL18] Lennart Ries, Jacob Langner, Stefan Otten, Johannes Bach, and Eric Sax. A driving scenario representation for scalable real-data analytics with neural networks. In *2019 IEEE Intelligent Vehicles Symposium (IV)*, pages 2215–2222. IEEE, 2019.
- [JL19] Philipp Rigoll, Jacob Langner, Lennart Ries, and Eric Sax. Unveiling objects with sola: An annotation-free image search on the object level for automotive data sets. In *2024 IEEE Intelligent Vehicles Symposium (IV)*, pages 1053–1059. IEEE, 2024.
- [JL20] Philipp Rigoll, Patrick Petersen, Jacob Langner, and Eric Sax. Parameterizable lidar-assisted traffic sign placement for the augmentation of

driving situations with cyclegan. In *International Conference On Systems Engineering*, pages 403–417. Springer, Cham, 2021.

- [JL21] Philipp Rigoll, Patrick Petersen, Jacob Langner, and Eric Sax. Parameterizable lidar-assisted traffic sign placement for the augmentation of driving situations with cyclegan. In *Advances in Systems Engineering: Proceedings of the 28th International Conference on Systems Engineering, ICSEng 2021, December 14-16, Wrocław, Poland 28*, pages 403–417. Springer, 2022.
- [JL22] Hanno Stage, Lukas Ewecker, Jacob Langner, Tin Stribor Sohn, Thomas Villmann, and Eric Sax. Reducing computer vision dataset size via selective sampling. In *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1422–1428. IEEE, 2023.
- [JL23] Hanno Stage, Lennart Ries, Jacob Langner, Stefan Otten, and Eric Sax. Analysis and comparison of datasets by leveraging data distributions in latent spaces. In *Deep Neural Networks and Data for Automated Driving*, pages 107–126. Springer, Cham, 2022.
- [JL24] Christian Steinhauser, Maciej Boncler, Jacob Langner, Steffen Strebel, and Eric Sax. An efficient strategy for testing adas on hil test systems with parallel condition-based assessments. In *VEHITS*, pages 391–399, 2022.
- [JL25] Christian Steinhauser, Joshua Ransiek, Jacob Langner, and Eric Sax. Contextualizing the sim-to-real gap for hardware-in-the-loop testing of highly automated driving functions. In *2024 8th International Conference on System Reliability and Safety (ICSRS)*, pages 404–411. IEEE, 2024.
- [JL26] Christian Steinhauser, Philipp Reis, Hubert Padusinski, Jacob Langner, and Eric Sax. Data quality matters: Quantifying image quality impact on machine learning performance. *arXiv preprint arXiv:2503.22375*, 2025.

Betreute Abschlussarbeiten

- [TH1] Nathalie Brenner. Entwicklung einer modularen testumgebung für fahrzeugregelungsfunktionen unter verwendung aufgezeichneter testdaten und der reactive-replay-methode, 2017.
- [TH2] Markus Glenk. Methodik zur repräsentativen bewertung der virtuellen applikation einer prädiktiven längsregelung, 2020.
- [TH3] Hannes Grolig. Automatisierte generierung von testsets durch semantische interpretation der realfahrten, 2018.
- [TH4] Kevin Huestis. Echtzeit-fähige erzeugung einer global-optimierten trajektorie mittels maschinellern lernen am beispiel ka-raceing, 2018.
- [TH5] Rouven Karp. Ähnlichkeitsanalyse von fahrsszenarien auf basis des visuellen eindrucks, 2020.
- [TH6] Xuesong Li. Scenario-finder to detect similar scenarios within a huge data pool, 2019.
- [TH7] Leonard Nürnberg. Repräsentative streckenabdeckung durch testfahrten auf grundlage von analyse und klassifizierung von kartendaten, 2020.
- [TH8] Philipp Reis. Classification of multivariate time series with machine learning in the automotive industry, 2020.
- [TH9] Lennart Ries. Entwurf und evaluation einer methode zur reduktion von bewertungsaufwänden in der automobilen funktionsentwicklung mittels replicator neural networks, 2017.

- [TH10] Christian Seiffer. Systematische funktions- und applikationsbewertung einer prädiktiven längsregelungsfunktion, 2019.
- [TH11] Shuai Wang. Parameter optimization of an automotive feature using machine learning, 2018.