

# Methode zur Bewertung der Risikobilanz autonomer Fahrzeuge aus Kundensicht

Yannik Potdevin\*    Dirk Nowotka  
Christian-Albrechts-Universität zu Kiel  
{ypo, dn}@informatik.uni-kiel.de

28. Juni 2021

## Kurzfassung

Wir stellen eine Methode zur Ermittlung der Risikobilanz autonomer Fahrzeuge im Vergleich zu durchschnittlicher menschlicher Fahrleistung vor, welche für Kunden (insbesondere Fahrzeughalter und Fahrzeugzulasser) anwendbar ist. Mithilfe unserer Methode werden Systemtests zur Prüfung von autonomen Fahrzeugen entwickelt und die zugehörigen Testergebnisse bewertet. Im Rahmen von zwei Fallstudien erproben wir unsere Methode und stellen die Ergebnisse vor.

## 1 Einleitung

Die bisherige Praxis auf öffentlichen Teststrecken für autonomes Fahren in Deutschland besteht darin, dass im selbstfahrenden Fahrzeug stets ein Begleiter (auch Steward oder Operator genannt) anwesend ist, um nötigenfalls in riskanten Situationen eingreifen zu können (dies ergibt sich aus den Paragraphen 1a und 1b des Straßenverkehrsgesetzes). Allerdings passierte am 28. Mai 2021 ein Gesetzentwurf den Bundesrat, der autonome Fahrzeuge der Stufe 4 bei Erfüllung einiger Zulassungskriterien grundsätzlich im öffentlichen Straßenverkehr gestattet.<sup>1</sup> Außerdem begann Waymo am Ende des letzten Jahres in Phoenix, Arizona den Betrieb autonomer Fahrzeuge ohne Begleiter im öffentlichen Straßenverkehr.<sup>2</sup>

Der Fahrzeughalter und der Begleiter (Fahrzeugführer) haften für Personen- und Sachschäden, welche beim Betrieb des Fahrzeugs entstehen können. Ist die Ursache für obige Schäden ein Produktfehler des Fahrzeugs, haftet der Hersteller [1].

Langfristig sollen autonome Fahrzeuge ohne Begleiter auskommen, damit für diesen keine Personalkosten anfallen und ihr Betrieb im Rahmen von Transportdienstleistungen wirtschaftlich ist. Hieraus ergeben sich zwei Fragen:

1. Aus Sicht des Zulassers: Wie ist zu prüfen und zu bewerten, dass ein autonomes Fahrzeug für die Zulassung zum Straßenverkehr geeignet ist?
2. Aus Sicht des Halters: Wie ist das verschobene Haftungsrisiko einzuschätzen, welches sich durch den Wegfall des Fahrzeugführers vermutlich zum Halter und Hersteller hin verschieben wird (so vermuten es Stöber, Möller und Pieronczyk [1])? Wie sicher ist das Fahrzeug tatsächlich im Betrieb?

---

\*Gefördert durch das BMVI-Projekt *Entwicklung und Evaluierung eines ÖPNV-on-demand-Angebots mit autonomen Fahrzeugen im öffentlichen Personennahverkehr in ländlichen Regionen (NAF-Bus)*, Förderkennzeichen 16AVF2018H.

<sup>1</sup><https://www.bmvi.de/SharedDocs/DE/Artikel/DG/gesetz-zum-autonomen-fahren.html> (Zugriff: 17. Juni 2021).

<sup>2</sup><https://www.reuters.com/article/us-waymo-autonomous-phoenix-idUSKBN26T2Y3> (Zugriff: 27. April 2021).

Aus beiden Fragen ergibt sich der Bedarf einer für Kunden (neben Haltern und Zulassern ist diese Arbeit zumindest indirekt auch für Versicherer relevant; an Hersteller richtet sich diese Arbeit explizit nicht) anwendbare Methode, das von autonomen Fahrzeugen ausgehende Risiko zu prüfen und zu bewerten. Eine naive Methode ist es, mit dem Fahrzeug (inklusive Begleiter) eine gewisse Zahl von Kilometern im öffentlichen Verkehr zurückzulegen und die dabei auftretenden Sach- und Personenschäden aufzuzeichnen. Die Anzahl der Schädigungen dividiert durch die Fahrleistung ergibt eine Fehlerrate. Ist diese Fehlerrate gering genug, wäre dies ein Argument für den Einsatz des autonomen Fahrzeugs. Was jedoch ist genug? Im Bericht der Ethikkommission [2] besagt die zweite Regel, dass die Zulassung von automatisierten Systemen nur vertretbar sei, wenn sie im Vergleich zu menschlichen Fahrleistungen zumindest eine Verminderung von Schäden im Sinne einer positiven Risikobilanz verspreche. Eine notwendige Bedingung ist daher, dass die Schadensrate des autonomen Fahrzeugs geringer ist als die eines durchschnittlichen menschlichen Verkehrsteilnehmers.

Eine Ausprägung von Schäden ist das tödliche Verunglücken von Fahrzeugbenutzern im Straßenverkehr in Deutschland. Bäumer et al. schätzen die Verunglücktenrate<sup>3</sup> in Deutschland im Jahr 2014 auf durchschnittlich 3 Personen pro  $10^9$  Fahrzeugkilometer (vergleiche [3, Abschnitt 4.4]). Dies interpretieren wir als das durchschnittliche Risiko menschlicher Fahrleistung (beschränkt auf das tödliche Verunglücken). Um mit einem Konfidenzbereich von 95 % statistisch signifikant nachzuweisen, dass von einem autonomen Fahrzeug höchstens ein Risiko von 0,3 tödlich verunglückten pro  $10^9$  Fahrzeugkilometer ausgeht (das autonome Fahrzeug wäre nach diesem Maß mindestens 90 % sicherer), müsste dieses Fahrzeug circa  $10^{10}$  km im typischen Straßenverkehr zurücklegen, ohne dabei einen einzigen Menschen tödlich zu verletzen (gemäß Kalra und Paddock [4, Gleichung (2)]). Unter der realistischen Annahme, dass während des Nachweises doch Menschen tödlich verunglücken würden, erhöhte sich die zu erbringende Fahrleistung noch einmal deutlich. Selbst wenn 100 baugleiche Fahrzeuge 24 Stunden pro Tag und 365 Tage im Jahr mit einer Durchschnittsgeschwindigkeit von  $30 \text{ km h}^{-1}$  die Fahrleistung unter sich aufteilen, benötigten sie rund 380 Jahre für diese Aufgabe. Ein Sicherheitsnachweis auf diese Art ist unpraktikabel.

In dieser Arbeit schlagen wir daher als praktikable Alternative vor, riskante Verkehrssituationen bewusst herbeizuführen – statt aktiv mit (Test)Fahrzeugen Fahrleistung zu erbringen und auf das Eintreten solcher Situationen zu hoffen – und das Verhalten des autonomen Fahrzeugs dabei zu beobachten und zu bewerten. Werden diese Situationen wie ein Test gestaltet und anschließend vom Fahrzeug bestanden, so erhöht dies das Vertrauen in die Sicherheit des Fahrzeugs. Schlagen die Tests jedoch fehl, ist dies ein Indiz für ein erhöhtes Risiko.

Die Hauptbeiträge dieser Arbeit sind:

1. Entwicklung einer Methodik zur Erstellung von Systemtests relevant für Halter und Zulasser von autonomen Fahrzeugen
2. Entwicklung einer Methodik zur Bewertung der Resultate obiger Systemtests
3. Angabe von Systemtests, die mittels der obigen Methodik entstanden
4. Angabe von Testergebnissen obiger Systemtests im Rahmen von 2 Fallstudien, einschließlich ihrer Bewertung

## 2 Methodik

### 2.1 Zweck und Abgrenzung

Die Rollen in der Zielgruppe unserer Methodik sind Fahrzeugzulasser und Fahrzeughalter. Ihr Interesse ist es, das spezifisch von autonomen Fahrzeugen (im Vergleich zu Fahrzeugen, die von Menschen geführt werden) ausgehende Risiko zu beurteilen. Ihre Fähigkeiten beschränken sich darauf, das autonome Fahrzeug als *Black Box* zu betrachten, das heißt sie haben nicht die Fähigkeiten und Kapazitäten des Herstellers, einzelne Komponenten und deren Implementierung

---

<sup>3</sup>Anzahl verunglückter Kfz-Benutzer pro Kfz-Inlandsfahrleistung.

zu betrachten. Dazu passend stellen wir in dieser Arbeit Systemtests vor, welche das Verhalten des Steuerungsprogramms<sup>4</sup> von autonomen Fahrzeugen hinsichtlich ihrer Verkehrssicherheit prüfen. Es handelt sich um Systemtests, weil sie für Zulasser und Halter ohne Kenntnis und Zugriff auf innere Komponenten durchführbar sind. Im Fokus der Tests steht nur das Steuerungsprogramm, denn wir nehmen an, dass das Verkehrsverhalten eines autonomen Fahrzeugs primär durch dessen Steuerungsprogramm geprägt wird. Andere sicherheitsrelevante Aspekte von autonomen Fahrzeugen, insbesondere jene, die sie mit nichtautonomen Fahrzeugen teilen, betrachten wir in dieser Arbeit nicht.

## 2.2 Skizze

In der Einleitung sprachen wir vom Herbeiführen riskanter Verkehrssituationen. Welche genau sind die für autonome Fahrzeuge riskanten Verkehrssituationen? Uns ist keine erschöpfende Aufzählung solcher Situationen bekannt. Jedoch kennen wir eine Möglichkeit, eine Reihe mindestens zu prüfender Situationen herzuleiten. Basierend auf Regel 2 des Berichts der Ethikkommission [2] ergibt sich nämlich: Die Risikobilanz autonomer Fahrzeuge muss für jene Verkehrssituationen, die für menschliche Fahrer riskant sind, im Vergleich zu menschlicher Fahrleistung strikt positiv ausfallen. Wäre dem nicht so, das heißt wäre die Risikobilanz autonomer Fahrzeuge im Vergleich zu menschlicher Fahrleistung bestenfalls neutral, dann folgt, dass autonome Fahrzeuge im Vergleich zu menschlicher Fahrleistung keine positive Gesamtrisikobilanz aufweisen. Dies wäre ein Widerspruch zu Regel 2 des Ethikkommissionsberichts [2] und stünde daher der Einführung von autonomen Fahrzeugen in den öffentlichen Verkehr entgegen. Das heißt: Es muss nachgewiesen werden, dass autonome Fahrzeuge in für die menschliche Fahrleistung riskanten Verkehrssituationen eine strikt positive Risikobilanz aufweisen. Andere Risikofaktoren, beispielsweise nur für autonome Fahrzeuge relevante Risikofaktoren, müssen zwar aus Sicht von Zulassern und Haltern betrachtet werden, sind jedoch nicht Teil dieser Arbeit.

Was für menschliche Fahrzeugführer riskant ist, leiten wir aus der Straßenverkehrsunfallstatistik [5] ab. Darin werden alle polizeilich erfassten Unfälle mit Personenschaden aufgeführt. Neben der Zahl und der Schwere der Schädigungen wird auch notiert, welche Ursachen den jeweiligen Unfälle zugrunde lagen. Für die Einordnung stehen 79 verschiedenen Ursachen (siehe [5, S. 13 ff.]) zur Verfügung. Wir gruppieren die verzeichneten Verkehrsunfälle nach gleicher Ursache und betrachten pro Gruppe, wie sich die Schwere der Personenschäden, das heißt die Zahl der leichten Verletzungen, schweren Verletzungen und tödlichen Verletzungen, verteilt. Daraus folgern wir, mit welcher Schwere Unfallfolgen einer solchen Gruppe durchschnittlich einhergehen. Die durchschnittliche, also gewissermaßen die zu erwartende Schwere von Unfällen einer bestimmten Ursache, ist eine wichtige Komponente, um anhand der Risikobewertung nach ISO 26262 [6, Teil 3] eine Risiko-Kategorisierung der menschlichen Fahrleistung bei Unfällen mit bestimmten Unfallursachen vorzunehmen. Zusammengefasst: Mithilfe der ISO-26262-Risikoanalyse und der Straßenverkehrsunfallstatistik können wir Unfallszenarien mit bekannten Ursachen in Risiko-Kategorien einteilen.

Diese Unfallszenarien mit ihren zugewiesenen Risiko-Kategorien sind der Vergleichsmaßstab den wir nutzen, um die Risikobilanz von autonomen Fahrzeugen im Vergleich zu menschlicher Fahrleistung zu bewerten. Besteht ein autonomes Fahrzeug ein Szenario, das bei menschlicher Fahrleistung mit einem hohen Risiko einhergeht, ist dies ein positiver Beitrag zur Gesamtbilanz. Versagt es in einem Szenario, das bei menschlicher Fahrleistung mit einem vernachlässigbaren Risiko korrespondiert, so ergibt sich hier ein negativer Beitrag zur Gesamtbilanz. Auf die Behandlung in den anderen Fällen gehen wir in Abschnitt 2.6 ein.

Existieren unter den Einzelbeiträgen zur Risikobilanz sowohl positive, als auch negative, so ist im Einzelfall zu prüfen, ob für jeden negativen Beitrag ein geeigneter positiver Beitrag zur Aufrechnung der zu erwartenden Personenschäden vorhanden ist. Dies halten wir im Rahmen der 9. Regel des Ethikkommissionsberichts [2] für vertretbar. Bleiben nach diesen Erwägungen negative, nicht aufrechenbare Beiträge bestehen, so sehen wir die notwendige Bedingung für die Einführung von autonomen Fahrzeugen (Regel 2) als nicht erfüllt an. Könnten jedoch alle

---

<sup>4</sup>Hierunter verstehen wir jenes System (beziehungsweise Systeme), welches die Fahrzeugumgebung inklusive anderer Verkehrsteilnehmer und der Verkehrsinfrastruktur erfasst, verarbeitet, interpretiert, daraus Schlüsse für die Fahrzeugführung (Längs- und Querführung) zieht, und diese dann veranlasst. Siehe auch Tabelle 1.

negativen Beiträge verrechnet werden, beziehungsweise gab es von vornherein keine, so betrachten wir Regel 2 nicht als verletzt. Ob diese Regel auch als erfüllt betrachtet werden kann, ist nicht allein auf Basis unserer Methode zu entscheiden. Eine positive Risikobilanz nach unserer Methode ist lediglich ein Indiz für die Erfüllung von Regel 2 und bedarf weiterer Argumente.

### 2.3 Risikoanalyse nach ISO 26262

Die Risikoanalyse nach ISO 26262 [6] beginnt damit, dass der Untersuchungsgegenstand (*item* [6, Teil 1, Def. 1.69]) definiert und abgegrenzt wird. In diesem Fall ist der Untersuchungsgegenstand einerseits das Steuerungsprogramm eines modellhaften autonomen Fahrzeugs (siehe Tabelle 1) und andererseits im Vergleich dazu die durchschnittliche menschliche Fahrleistung (diese definieren wir nicht explizit, sondern setzen auf ein implizites Verständnis).

Tabelle 1: Definition des Untersuchungsgegenstandes »Steuerungsprogramm eines modellhaften autonomen Fahrzeugs« (SAE-Level 4–5) gemäß ISO 26262.

Zweck	Sicherer und komfortabler Transport von Personen (Fahrgästen) von einem Ort zu einem anderen Ort im öffentlichen Straßenverkehr.
Funktionen	Fahrzeugsteuerung, Geschwindigkeitsregelung Kurz-, mittel- und langfristige Routenplanung Wahrnehmung und Verständnis der näheren Fahrzeugumgebung, insbesondere des Geländes, der Verkehrsinfrastruktur und anderer Verkehrsteilnehmer
Anforderungen	Einhalten der Verkehrsregeln Für Fahrgäste und andere Verkehrsteilnehmer sichere und komfortable Fahrweise
Abhängigkeiten und Interaktionen mit der Umwelt	Koordination mit anderen Verkehrsteilnehmern und Verkehrsleitinfrastruktur Grundsätzlich sind alle Umweltbedingungen zulässig, mit Ausnahme etwa von starkem Regen, starkem Schneefall, Eis oder Schnee auf der Strecke, starkem Wind oder Sturm, Staub oder Nebel oder Dampf auf der Strecke und extremen Temperaturen (je nach Fahrzeug entfallen manche Bedingungen hiervon, beziehungsweise kommen neue Bedingungen hinzu).
Abhängigkeiten und Interaktionen mit anderen <i>Items</i>	Sensoren, welche die nähere Umgebung in hinreichendem Umfang, Qualität und Geschwindigkeit erfassen

Ist der Untersuchungsgegenstand definiert, gilt es das von ihm ausgehende Risiko (*risk* [6, Teil 1, Def. 1.99]) zu beurteilen. Dazu fehlen uns jedoch noch einige Grundbegriffe des Risikomanagements nach ISO-26262:

- Gefährdung (*hazard* [6, Teil 1, Def. 1.57]): Eine potentielle Schadensquelle, ausgelöst durch fehlerhaftes Verhalten des Untersuchungsgegenstands.
- Betriebssituation (*operational situation* [6, Teil 1, Def. 1.83]): Ein Szenario, welches während des Betriebs des Fahrzeugs eintreten kann.
- Gefährliches Ereignis (*hazardous event* [6, Teil 1, Def. 1.59]): Eine Kombination aus einer Gefährdung und einer Betriebssituation.

- Schweregrad (*severity* [6, Teil 1, Def. 1.120]): Ein Schätzwert für das Ausmaß an Personenschäden, die bei einem gefährlichen Ereignis entstehen können. Er wird in eine der vier Klassen S0–S3 eingeteilt.
- Häufigkeit (*exposure* [6, Teil 1, Def. 1.37]): Die Wahrscheinlichkeit dafür, dass sich das Fahrzeug während seines Betriebs in jener Betriebsituation befindet, die Grundlage für ein gefährliches Ereignis ist. Sie wird in einer der fünf Klassen E0–E4 eingeteilt.
- Beherrschbarkeit (*controllability* [6, Teil 1, Def. 1.19]): Die Fähigkeit, eines gefährlichen Ereignisses zugeordnete Personenschäden durch zeitnahe Reaktionen involvierter Personen abzuwenden. Sie wird eine der vier Klassen C0–C3 eingeteilt.
- ASIL (*Automotive-Safety-Integrity-Level* [6, Teil 1, Def. 1.6]): Eine vier<sup>5</sup> Level umfassende Skala, aus der sich notwendigerweise zu erfüllende Anforderungen der ISO 26262 und anzuwendende Sicherheitsmaßnahmen ergeben, um inakzeptables Restrisiko zu vermeiden. D ist das strengste und A das am wenigsten strenge Level.

Die ISO 26262 definiert Risiko als die Kombination aus der Schwere eines Personenschadens (*harm* [6, Teil 1, Def. 1.56]) und der Eintrittswahrscheinlichkeit dieses Personenschadens. Das Problem der Ermittlung des vom Untersuchungsgegenstand (als Ganzes) ausgehenden Risikos wird in Teilprobleme zerlegt. Mithilfe von Risikoanalysetechniken wird versucht, eine möglichst vollständige Auflistung der zum Untersuchungsgegenstand gehörigen gefährlichen Ereignisse zu ermitteln. Zu jedem gefährlichen Ereignis werden die drei Parameter Schweregrad, Häufigkeit und Beherrschbarkeit ermittelt. Ergibt sich, dass mindestens einer der Parameter den jeweiligen Nullwert annimmt, bewertet die ISO 26262 das Risiko als vernachlässigbar (kein ASIL vergeben). Anderenfalls ist das Risiko nicht vernachlässigbar und ein ASIL wird vergeben (siehe Tabelle 4 der ISO 26262 [6, Teil 3]). Das ASIL QM steht dabei für ein akzeptables Risiko, das keiner gesonderten<sup>6</sup> Reduktion bedarf. Das ASIL A hingegen kennzeichnet inakzeptables Risiko und fordert geringe Risikoreduktionsmaßnahmen (RRM) ein. Das ASIL D kennzeichnet ebenfalls inakzeptables Risiko und fordert hohe RRM ein. Die ASIL B und C stellen Zwischenstufen dar. Erst wenn alle geforderten RRM aller gefährlichen Ereignisse umgesetzt werden, gilt das von den gefährlichen Ereignissen ausgehende Restrisiko als akzeptabel. Aus der erfolgreichen Behandlung der Teilprobleme ergibt sich insgesamt, dass das vom ganzen Untersuchungsgegenstand ausgehende Restrisiko als akzeptabel einzustufen ist.

## 2.4 Straßenverkehrsunfalldaten und gefährliche Ereignisse

Die Straßenverkehrsunfallstatistik beinhaltet Angaben zu Unfällen, Beteiligten, Fahrzeugen, Verunglückten und Unfallursachen. Ihr Zweck ist es, bundesweit vergleichbare Daten zum Straßenverkehrsunfallgeschehen zu erfassen [7].

Der Jahresbericht zu den Verkehrsunfällen des Jahres 2014 [5, S. 13 ff.] listet in seinen methodischen Erläuterungen 79 Unfallursachen auf.<sup>7</sup> Der für uns relevante Teil umfasst Fehler der Fahrzeugführer (die Nummern 8–49) und allgemeine Unfallursachen (die Nummern 70–89). Beispiele für Fehler der Fahrzeugführer sind etwa ungenügender Sicherheitsabstand (Nr. 14), Fehler beim Überholtwerden (Nr. 23) und falsches Verhalten gegenüber Fußgängern an Fußgängerüberwegen (Nr. 38). Für die allgemeinen Unfallursachen sind etwa Glätte oder Schlüpfrigkeit der Fahrbahn aufgrund von Laub und Lehm (Nr. 74), mangelhafte Beleuchtung der Straße (Nr. 78) und Wild auf der Fahrbahn (Nr. 86) Beispiele.

Die Unfallursachen der Straßenverkehrsunfallstatistik verwenden wir als Grundlage um durch Brainstorming gefährliche Ereignisse zu ermitteln, die zu einem Unfall eben dieser Ursache führen können. Jede so entstandene Situation muss vier Kriterien erfüllen.

<sup>5</sup>In dieser Arbeit zählen wir QM als fünftes Level mit.

<sup>6</sup>Das heißt keine Maßnahmen, die über gewöhnliches Qualitätsmanagement hinausgehen.

<sup>7</sup>Wir nutzen die deutsche Unfallstatistik, weil wir uns mit der Zulassung und Haltung von autonomen Fahrzeugen im öffentlichen Straßenverkehr in Deutschland beschäftigen. Das Bezugsjahr ist 2014, weil die Zahl der Fahrerassistenzsysteme im Fahrzeugbestand jenen Jahres noch gering ist (siehe auch [https://www.adac.de/infotestrat/technik-und-zubehoer/fahrerassistenzsysteme/fakten\\_fuer\\_experten/default.aspx](https://www.adac.de/infotestrat/technik-und-zubehoer/fahrerassistenzsysteme/fakten_fuer_experten/default.aspx) [Zugriff: 18. Juni 2021]) und so ein prototypischer Vergleich von autonomer Fahrleistung mit (nahezu) rein menschlicher Fahrleistung ermöglicht wird.

- Vergleichbarkeit: Das gefährliche Ereignis gilt für das Steuerungsprogramm eines modellhaften autonomen Fahrzeugs, als auch für menschliche Fahrleistung.
- Eignung: Im Rahmen der beiden Fallstudien (siehe Abschnitt 3) ist das gefährliche Ereignis für die konkreten Fahrzeugmodelle EZ10 von Easymile und ARMA von Navya anwendbar (hinsichtlich des Funktionsumfangs).
- Durchführbarkeit: Aus den Situationen lassen sich auf dem Gelände des GreenTec Campus und auf der Insel Sylt (siehe Fallstudien) durchführbare Testfälle ableiten.
- Einverständnis: Die Betreiber der Fahrzeuge (GreenTec Campus und die Sylter Verkehrsbetriebe) sind mit der Durchführung von aus den Situationen abgeleiteten Testfällen einverstanden.

In Tabelle 2 findet sich das Ergebnis unseres Brainstormings.

Die für uns relevanten gefährlichen Ereignisse sind nun bekannt. Der nächste Schritt ist es, die zugehörigen ASIL zu bestimmen. Hierfür benötigen wir drei Informationen:

- Beherrschbarkeit des gefährlichen Ereignisses
- Schweregrad des gefährlichen Ereignisses
- Häufigkeit der zugrundeliegenden Betriebssituation (des gefährlichen Ereignisses)

Die Beherrschbarkeitsklasse setzen wir pauschal und pessimistisch auf C3, weil ein Fahrgast eines autonomen Fahrzeugs des SAE-Automatisierungsgrades [8] 4 oder 5 (wie bei unserem Modellfahrzeug der Fall) explizit von dessen Steuerung und Überwachung entbunden ist.

Den Schweregrad leiten wir aus der Straßenverkehrsunfallstatistik (StVU) ab. Zu einem gefährlichen Ereignis kennen wir die motivierende Unfallursache. Durch eine Analyse der StVU ist uns die Zahl der in Unfällen dieser Ursache leichtverletzten, schwerverletzten und getöteten Personen bekannt.<sup>8</sup> Die relativen Anteile an unverletzten, leichtverletzten, schwerverletzten und getöteten Personen zu einer Unfallursache interpretieren wir als die Wahrscheinlichkeit, bei einem Unfall dieser Ursache entsprechend zu verunglücken (für die detaillierte Berechnung der relativen Häufigkeiten siehe Abschnitt 2.5). Liegt die Wahrscheinlichkeit bei einem Unfall einer bestimmten Ursache zu sterben bei mindestens 10 %, weisen wir gemäß ISO 26262 [6, Teil 3, Anhang B] der Unfallursache, und damit auch dem abgeleiteten gefährlichen Ereignis, den Schweregrad S3 zu. Ist die Wahrscheinlichkeit getötet zu werden geringer als 10 %, aber die Wahrscheinlichkeit für eine schwere Verletzung mindestens 10 %, ergibt sich der Schweregrad S2. Ist der Schweregrad weder S3 noch S2 und liegt die Wahrscheinlichkeit einer leichten Verletzung bei mindestens 10 %, folgt der Schweregrad S1. An anderen Fällen (also bei einer Wahrscheinlichkeit von mindestens 90 % unverletzt zu bleiben) erhält das gefährliche Ereignis den Schweregrad S0.

Für die Ermittlung der Häufigkeit einer Betriebssituation, welche einem gefährlichen Ereignis zugrunde liegt, bedienen wir uns wiederum der ISO 26262 [6, Teil 3, Anhang B], denn dort sind bereits zahlreiche Betriebssituationen samt Häufigkeitsklasse aufgeführt. Für Betriebssituationen aus Tabelle 2, die nicht in der ISO 26262 aufgeführt sind, suchen wir Betriebssituationen heraus, welche unserer Einschätzung nach ähnlich häufig vorkommen und übernehmen anschließend deren Häufigkeitsklasse. Dies ist zum Beispiel bei den Betriebssituationen zu den Unfallursachen 27 und 38 der Fall.

Sind die drei oben erwähnten Parameter bestimmt (wobei der Grad der Beherrschbarkeit mit C3 bereits feststeht), erfolgt mithilfe von Tabelle 3 die Ermittlung des entsprechenden ASIL.

---

<sup>8</sup>Dabei gilt eine durch einen Unfall verletzte Personen als getötet, falls sie innerhalb von 30 Tagen an den Folgen des Unfalls stirbt; sie gilt als schwerverletzt, falls sie unmittelbar zur stationären Behandlung (mindestens 24 Stunden) in ein Krankenhaus aufgenommen wird und als leichtverletzt gilt sie in den übrigen Fällen.

Tabelle 2: Aus den Unfallursachen der Straßenverkehrsunfallstatistik abgeleitete Betriebssituationen und gefährliche Situationen, die für beide Untersuchungsgegenstände (Steuerungsprogramm und menschliche Fahrleistung) anwendbar sind. Der Einfachheit halber erfolgt die Beschreibung aus Sicht des Steuerungsprogramms. UNr. entspricht Unfallursachennummer, E<sub>n</sub> entspricht Häufigkeitsklasse und S<sub>n</sub> entspricht Schweregrad.

UNr.	Betriebssituation	E <sub>n</sub>	Gefährliches Ereignis	S <sub>n</sub>	ASIL
(14, 15)	Dem Fahrzeug fährt ein anderer Verkehrsteilnehmer unmittelbar voraus.	E4	Die Steuerungssoftware stellt einen ungenügenden Sicherheitsabstand zum Vorausfahrenden ein. Verringert der Vorausfahrende plötzlich seine Geschwindigkeit stark, könnte es zu einem Zusammenstoß mit ihm kommen.	S1	B
15*	Dem Fahrzeug fährt ein anderer Verkehrsteilnehmer unmittelbar hinterher.	E4	Die Steuerungssoftware veranlasst ohne zwingenden Grund eine starke Bremsung. Es könnte zu einem Zusammenstoß mit dem Hinterherfahrenden kommen.	S1	B
23	Das Fahrzeug wird (von einem anderen Verkehrsteilnehmer) überholt.	E3	Beim Überholtwerden veranlasst die Steuerungssoftware ein fehlerhaftes Manöver (d. h. sie weicht vom Grundsatz »Spur und Geschwindigkeit halten« in potentiell schädlicher Weise ab). Es könnte zu einem Zusammenstoß mit dem Überholenden kommen. Außerdem könnte das fehlerhafte Manöver dazu führen, dass der Überholende mit Dritten oder Hindernissen kollidiert.	S2	B
27	Das Fahrzeug befährt eine Kreuzung oder Einmündung mit der Regelung »rechts vor links«.	E4	Die Steuerungssoftware missachtet die Regel »rechts vor links«, missachtet also die Gebote in §8 StVO. Es könnte zu einem Zusammenstoß mit dem vorrangigen Verkehrsteilnehmer kommen.	S1	B
35	Das Fahrzeug biegt nach links oder rechts ab.	E4	Die Steuerungssoftware veranlasst ein fehlerhaftes Manöver beim Abbiegen nach links oder nach rechts (beispielsweise missachtet sie die Gebote in §9 StVO). Es könnte zu einem Zusammenstoß mit anderen Verkehrsteilnehmern oder Hindernissen kommen.	S1	B
36	Das Fahrzeug wendet oder fährt rückwärts.	E4	Die Steuerungssoftware veranlasst ein fehlerhaftes Manöver beim Wenden oder Rückwärtsfahren (beispielsweise missachtet sie die Gebote in §9 StVO). Es könnte zu einem Zusammenstoß mit anderen Verkehrsteilnehmern oder Hindernissen kommen.	S1	B
37	Das Fahrzeug fährt in den fließenden Verkehr ein.	E4	Die Steuerungssoftware veranlasst ein fehlerhaftes Manöver beim Einfahren in den fließenden Verkehr (beispielsweise verstößt es gegen die Gebote in §10 StVO). Es könnte zu einem Zusammenstoß mit anderen Verkehrsteilnehmern oder Hindernissen kommen.	S1	B
38	Das Fahrzeug ist im Begriff einen Fußgängerüberweg zu kreuzen, an/ auf dem sich ein Fußgänger befindet.	E4	Die Steuerungssoftware veranlasst falsches Verhalten gegenüber Fußgängern an Fußgängerüberwegen. Es könnte zu einem Zusammenstoß mit Fußgängern kommen.	S2	C
39	Das Fahrzeug ist im Begriff eine Fußgängerfurt zu kreuzen, an/auf der sich ein Fußgänger befindet.	E4	Die Steuerungssoftware veranlasst falsches Verhalten gegenüber Fußgängern an Fußgängerfurten. Es könnte zu einem Zusammenstoß mit Fußgängern kommen.	S2	C
74	Die Fahrbahn ist beeinträchtigt durch andere Einflüsse als Öl, Fahrbahnbenutzerverschmutzungen, Schnee, Eis oder Regen; zum Beispiel durch Laub, angeschwemmten Lehm, etc.	E2	Die Steuerungssoftware passt das Fahrverhalten nicht entsprechend der Fahrbahnbeeinträchtigung an (etwa Anpassung der Geschwindigkeit, Anpassung der Abstände, etc.). Unnötig langsames Fahren oder unnötig starkes Abbremsen ist kein angepasstes Fahrverhalten. Es könnte zu einem Zusammenstoß mit anderen Verkehrsteilnehmern oder Hindernissen kommen.	S2	A
85	Das Fahrzeug begegnet einer nicht, oder nur unzureichend gesicherten Arbeitsstelle auf der Fahrbahn.	E1	Wird die Arbeitsstelle nicht rechtzeitig als solche erkannt, könnte es zu einem Zusammenstoß mit Teilern der Arbeitsstelle oder mit dort arbeitenden Personen kommen.	S2	QM

Tabelle 3: Angabe der ISO 26262 ASIL, nach Festlegung der Beherrschbarkeit auf Klasse C3, in Abhängigkeit des Schweregrades (Zeilen) und der Häufigkeit (Spalten).

	E1	E2	E3	E4
S1	QM	QM	A	B
S2	QM	A	B	C
S3	A	B	C	D

## 2.5 Ermittlung des Schweregrades der gefährlichen Ereignisse

Den Schweregrad von gefährlichen Ereignissen leiten wir aus der StVU<sup>9</sup> ab. Die gefährlichen Ereignisse ergeben sich jeweils aus einer Unfallursache der StVU. Zu jedem polizeilich erfassten Unfall (dies ist die Erhebungseinheit der StVU) werden, soweit durch Polizeibeamte vor Ort feststellbar, Unfallursachen aufgezeichnet. Für den Unfall als solchen können bis zu zwei allgemeine Unfallursachen (Nr. 70–89) erfasst werden. Hinzu kommen pro Beteiligten<sup>10</sup> bis zu drei personenbezogene Fehlverhalten (Nr. 1–69). Weiterhin wird pro Unfall die Anzahl der getöteten, der schwerverletzten und der leichtverletzten Personen notiert. Pro Unfallbeteiligten wird zusätzlich dessen Unfallfolgen (also ob er un-, leicht-, schwerverletzt ist oder getötet wurde), die Zahl der Verletzten, beziehungsweise getöteten Mitfahrer und die Zahl der Benutzer des Fahrzeugs des Beteiligten verzeichnet. Anschaulich ergeben sich Unfalldaten, wie sie in den Tabellen 4 und 5 zu sehen sind.

Tabelle 4: Fiktiver Auszug aus der StVU mit den für uns relevanten und zu einem Unfall gehörigen Merkmalen. UID identifiziert eindeutig einen Unfall. U Ursache{1, 2} kennzeichnen die primäre und die sekundäre allgemeine Unfallursache. UGetoe, USchwer und ULeicht stehen jeweils für die Anzahl der Getöteten, der Schwerverletzten und der Leichtverletzten.

UID	U Ursache1	U Ursache2	UGetoe	USchwer	ULeicht
...	...	...	...	...	...
10	83	84	2	3	5
...	...	...	...	...	...

Tabelle 5: Fiktiver Auszug aus der StVU mit den für uns relevanten und zu den beteiligten eines Unfalls gehörigen Merkmalen. UID identifiziert eindeutig einen Unfall. BID identifiziert pro Unfall eindeutig einen Beteiligten. BUrsache{1, 2, 3} kennzeichnen primäres, sekundäres und tertiäres personenbezogenes Fehlverhalten. BUnfallf steht für die Unfallfolge, die der Beteiligte selbst erlitt, BMitf für die Zahl der Verletzten oder getöteten Mitfahrer und BFzBenutz für die Zahl der Fahrzeugbenutzer des Fahrzeugs des Beteiligten.

UID	BID	BUrsache1	BUrsache2	BUrsache3	BUnfallf	BMitf	BFzBenutz
...	...	...	...	...	...	...	...
10	1	12	1	–	1	3	4
10	2	13	–	–	1	4	7
...	...	...	...	...	...	...	...

Um nun den Schweregrad eines gefährlichen Ereignisses zu ermitteln, gehen wir grob<sup>11</sup> wie folgt vor:

<sup>9</sup>Quelle: FDZ der Statistischen Ämter des Bundes und der Länder, Straßenverkehrsunfälle, 2014, eigene Berechnungen.

<sup>10</sup>Auszug aus den methodischen Erläuterung zu der StVU [5, S. 12]: »Als Beteiligte an einem Straßenverkehrsunfall werden alle Fahrzeugführer oder Fußgänger erfasst, die selbst - oder deren Fahrzeug - Schäden erlitten oder hervorgeufen haben. Verunglückte Mitfahrer zählen somit nicht zu den Unfallbeteiligten.«

<sup>11</sup>Die Details können wir der Implementierung des Analyseprogramms entnehmen, welches wir dem zuständigen Forschungszentrum zur Ausführung übermittelten, siehe <https://github.com/ypotdevin/severity-unfallstatistik>.



1. Wir schränken die Daten auf Unfälle ein, bei denen der Hauptverursacher ein Fahrzeug eines Typs führt, der vermutlich in der Zukunft vollautomatisierbar ist.
2. Wir berechnen pro Unfall eine untere Schranke für die Zahl der durch diesen Unfall gefährdeten Verkehrsteilnehmer.
3. Pro Unfallursache  $u$  der 79 Unfallursachen:
  - (a) Wir wählen genau die Unfälle aus, welche  $u$  als eine Ursache nennen.
  - (b) Unter jenen Unfällen summieren wir jeweils die Zahlen der Leichtverletzten ( $L_u$ ), der Schwerverletzten ( $S_u$ ), der Getöteten ( $Get_u$ ) und der Gefährdeten ( $Gef_u$ ) untereinander auf und berechnen anschließend die Quotienten  $\frac{L_u}{Gef_u}$ ,  $\frac{S_u}{Gef_u}$  und  $\frac{Get_u}{Gef_u}$ .
  - (c) Wir interpretieren die oben berechneten Quotienten als die Wahrscheinlichkeit, bei einem Unfall der Ursache  $u$  eine der jeweiligen Unfallfolgen zu erleiden und weisen dann, wie in Abschnitt 2.4 beschrieben,  $u$  den passenden Schweregrad zu.
  - (d) Den für  $u$  bestimmten Schweregrad weisen wir allen auf  $u$  basierenden gefährlichen Ereignissen zu.

Die in Nummer 1 getroffene Einschränkung auf »vermutlich zukünftig automatisierbare Fahrzeuge« begründen wir damit, dass wir nur jene Unfälle betrachten wollen, die so prinzipiell auch für autonome Fahrzeuge relevant, beziehungsweise diesen zuzurechnen sind. Unfälle, die durch Verkehrsteilnehmer einer anderen Art der Verkehrsbeteiligung verursacht wurden, sollen die errechneten Kennzahlen nicht verfälschen. Im Detail verstehen wir unter »vermutlich zukünftig automatisierbar« alle Kraftfahrzeuge<sup>12</sup> (Art der Verkehrsbeteiligung mit Schlüsselnummern 1–59, siehe [5, S. 18 f.]), nicht jedoch motorisierte Zweiräder (Schlüsselnummern 1–15). Letztere schließen wir aus, weil wir zurzeit keinen Markt für autonome motorisierte Zweiräder sehen, da unserer Einschätzung nach die Gründe für das Fahren eines Motorrads hauptsächlich emotionaler Natur (Fahrspaß etc.) sind, Menschen daher Motorräder selbst fahren wollen. Ebenso schließen wir, aus offensichtlichen Gründen, Fußgänger und sonstige Fahrzeuge aus (Schlüsselnummern 71–93). Zu Schienenfahrzeugen (Schlüsselnummern 61, 62) sei jedoch erwähnt: Eine Vollautomatisierbarkeit dieser Fahrzeuge halten wir für durchaus realistisch. Dennoch betrachten wir Schienenfahrzeuge nicht in dieser Arbeit, weil sie eine andere Methodik zur Risikobewertung erfordern.

Die ISO 26262 fordert bei der Ermittlung des Schweregrades eines gefährlichen Ereignisses die Auswirkungen auf jede potentiell gefährdete Person zu betrachten (vergleiche [6, Klausel 7.4.3.2, Notiz 1]). Es ist demnach erforderlich anzugeben, welcher Personenkreis als durch einen Unfall gefährdet gilt. Allerdings ist dabei nur auf Personenschäden zu achten – Unfälle mit nur Sachschäden brauchen nicht betrachtet zu werden. Die folgenden drei Regeln nutzen wir, um aus der StVU die durch einen Unfall gefährdeten Personen zu ermitteln. Erstens: Alle im Rahmen eines Unfalls verletzten Personen (inklusive der Getöteten) gelten als gefährdet. Dies schließt insbesondere verletzte Fußgänger ein. Zweitens: Ist mindestens ein Benutzer<sup>13</sup> eines Fahrzeugs verletzt, gelten alle Benutzer des Fahrzeugs ebenfalls als gefährdet. Drittens: Alle Benutzer eines Fahrzeugs, welches mit einem weiteren Fahrzeug zusammenstößt und verunfallt, gelten als gefährdet.

Da unverletzte Fußgänger nicht in der StVU erfasst werden, sind sie nicht Teil des obigen Regelwerks. Außerdem lassen es unsere Ressourcen nicht zu, die 230466 für uns relevanten Unfälle im Einzelfall zu beurteilen, weshalb die Regeln ausschließlich allgemein formuliert sind. Die Zahl der tatsächlich gefährdeten Personen wird damit systematisch unterschätzt (dies ist der Grund für die Formulierung »untere Schranke« in Nummer 2), was jedoch hinnehmbar ist, da so die in Nummer 3 berechneten Wahrscheinlichkeiten überschätzt werden (vergleiche [6, Klausel 7.4.3.1, Notiz]).

<sup>12</sup>Hierunter fallen insbesondere schwere Fahrzeuge wie Kraftomnibusse oder Lastkraftwagen. Da die ISO 26262 seit ihrer Überarbeitung im Jahr 2018 auch für solche Fahrzeuge anwendbar ist, halten wir unsere Methodik für kompatibel.

<sup>13</sup>Verletzte und unverletzte Personen in und auf dem Fahrzeug, vergleiche [5, S. 12].

## 2.6 Testerstellung und -bewertung

Wie gefährliche Ereignisse zu bestimmen und zu bewerten sind, wurde in den Abschnitten 2.3–2.5 beschrieben. Die nächsten und letzten Schritte sind das Generieren von Testfällen (auf Basis der gefährlichen Ereignisse), ihre Durchführung, die Bewertung der Einzeltestergebnisse und die Zusammenfassung der Einzeltestergebnisse zu einem Gesamtergebnis.

Da gefährliche Ereignisse Situationen beschreiben, die auf Ebene des Fahrzeugs zu beobachten sind (im Vergleich zur Ebene der Komponenten), prüfen wir das Fahrzeugverhalten mittels Systemtests. Die Tests sind so konzipiert, dass sie das jeweilige gefährliche Ereignis realistisch abbilden und in Kernpunkten vergleichbar sind. Jedoch erlauben wir explizit einige Freiheitsgrade um einerseits den Testaufwand zu reduzieren und um andererseits der Dynamik und Unschärfe des Verkehrsgeschehens gerecht zu werden (siehe auch Abschnitt 3.2 für weitere Erläuterungen). Beispiele für solche Freiheitsgrade sind etwa der Verzicht auf die genaue Vorgabe des Wetters (sofern es sich nicht extrem ist); der Form und Beschaffenheit des in manchen Tests benötigten Fahrrads, Rollers und Personenkraftwagens; der Größe und Form von Testpersonen oder der detaillierten Umgebung (etwa Straßenbelag und -beschaffenheit).

Eine Übersicht der Testfälle ist in Tabelle 6 zu finden.

Nun beschäftigen wir uns mit der Bewertung der Testergebnisse. Abhängig von dem ASIL des Testfalls (welches sich aus dem gefährlichen Ereignis ergibt), bewerten wir das Bestehen (beziehungsweise das Fehlschlagen) eines Tests unterschiedlich. Unsere Motivation dafür ist, ein ASIL als Aufforderung zur Risikoreduktion zu betrachten. Daraus schließen wir: Ist die menschliche Fahrleistung mit einem hohen ASIL bewertet, geht von ihr ein zu hohes Risiko aus. Ist sie mit dem ASIL QM bewertet, ist das Risiko akzeptabel. Bezüglich des Vergleichs von durchschnittlicher menschlicher Fahrleistung und der Fahrleistung autonomer Fahrzeuge legen wir fest: Besteht das autonome Fahrzeug einen Test mit dem ASIL D, das heißt ein Szenario welches mit maximaler Risikoreduktionsaufforderung für die durchschnittliche menschliche Fahrleistung verknüpft ist, gilt dies als Risikobilanzbeitrag zugunsten (positiv) des autonomen Fahrzeugs. Fällt das autonome Fahrzeug bei einem ASIL-QM-Test durch, also bei einem Szenario ohne Risikoreduktionsaufforderung, ist der Risikobilanzbeitrag zulasten (negativ) des autonomen Fahrzeugs. Als neutrale Beiträge definieren wir das Bestehen eines ASIL-QM-Testfalls und das Fehlschlagen eines ASIL-D-Testfalls.

Weniger eindeutig ist die Bewertung in den übrigen sechs Fällen. Wird ein Testfall mit ASIL-C statt ASIL-D bewertet, ist die Risikoreduktionsanforderung eine Größenordnung geringer. Dies gilt auch für die Abstufungen von ASIL-C auf ASIL-B und ASIL-B auf ASIL-A. Das von der durchschnittlichen menschlichen Fahrleistung ausgehende Risiko nimmt folglich mit jeder Stufe deutlich ab. Ein Vergleich, der nur auf der Prüfung des Bestehens oder Nichtbestehens basiert, erfasst diese Abstufungen nicht genügend. Wir berücksichtigen diesen Umstand, indem wir in solchen Fällen den Risikobilanzbeitrag nicht mehr punktgenau beziffern, sondern einen Ermessensspielraum schaffen, der im Einzelfall zu betrachten ist. Dies haben wir in Abbildung 1 mithilfe von breiten Intervallen dargestellt. Besteht beispielsweise das autonome Fahrzeug einen ASIL-C-Test, schätzen wir dies weder als negativ, noch als nur neutral ein. Ob dies jedoch als eindeutig positiv, oder nur als bedingt positiv zu werten ist, hängt vom Einzelfall ab. Deshalb bewerten wir so einen Fall als eher positiv bis positiv. Für das Bestehen eines ASIL-B-Tests schließen wir lediglich einen negativen Beitrag aus und lassen offen, ob oder wie positiv dies ist. Beim Bestehen eines ASIL-A-Tests schließen wir zusätzlich auch noch einen eindeutig positiven Bilanzbeitrag aus, weil wir hier davon ausgehen, dass das von der durchschnittlichen menschlichen Fahrleistung ausgehende Risiko nicht mehr ausreichen hoch ist, um dem autonomen Fahrzeug einen deutlichen Vorsprung im Risikovergleich zuzusprechen. Für das Fehlschlagen bei Tests mit ASIL-Bewertung A–C führen Symmetrieüberlegungen und obige Argumente zu den in Abbildung 1 aufgeführten Bewertungen.

Zum Schluss stellen wir uns die Frage, wie aus den einzelnen Risikobilanzbeiträgen eine Gesamtbilanz, beziehungsweise ein Gesamtbild zu destillieren ist. Eine naheliegende Idee ist, ausgewählte Risikobeiträge paarweise miteinander zu verrechnen; etwa das Bestehen eines ASIL-D-Tests mit dem Fehlschlagen eines ASIL-QM-Tests. Eigentlich ist das Ziel der Ermittlung der Risikobilanz herauszufinden, ob autonome Fahrzeuge im Vergleich zur durchschnittlichen menschlichen Fahrleistung insgesamt weniger Personenschäden verursachen. Das heißt

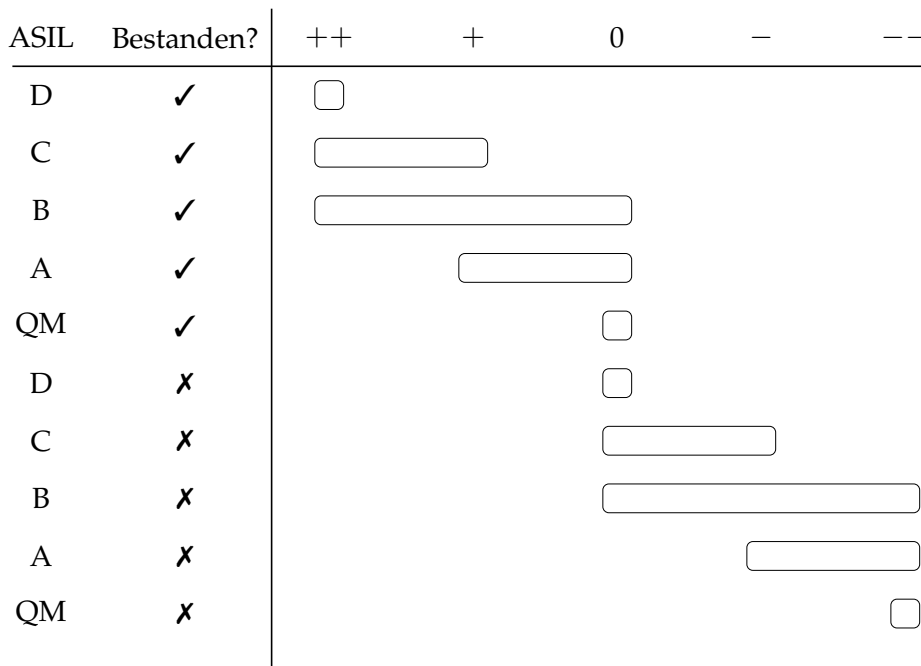


Abbildung 1: Beiträge zur Risikobilanz in Abhängigkeit davon, wie ein Test des autonomen Fahrzeugs ausgeht und welches ASIL dem Test zugeordnet ist. Dabei ist (++) ein positiver Beitrag, (+) ein eher positiver Beitrag, (0) ein neutraler Beitrag, (-) ein eher negativer Beitrag und (--) ein negativer Beitrag. Die breiten Intervalle deuten, im Vergleich zu den Punkten, den Bewertungsspielraum an.

um eine paarweise Aufrechnung zu rechtfertigen, müsste bekannt sein, in welchem Umfang Personenschäden durch das Bestehen des einen Testfalls vermieden<sup>14</sup> werden und auch in welchem Umfang zusätzliche Personenschäden durch das Fehlschlagen eines anderen Testfalls zu erwarten sind. Unsere Methodik bietet für keinen der beiden Fälle geeignete Maße. Daher geben wir hier keine Prinzipien für die paarweise Aufrechnung an und verweisen allgemein auf die Einzelfallbetrachtung.

Um dennoch ein grobes Gesamtbild zu erhalten, welches uns ermöglicht zumindest in eindeutigen Fällen zu entscheiden, ob die notwendige Bedingung der positiven Risikobilanz vom untersuchten Fahrzeug erfüllt wird, schlagen wir folgendes vor: Alle Risikobilanzbeiträge werden gruppiert, gezählt und in einem Histogramm dargestellt (wie beispielhaft in Abbildung 2 gezeigt). Machen die negativen Beiträge mindestens die Hälfte aller Beiträge aus, so gehen wir davon aus, dass die Gesamtbilanz negativ ist. Ist die Zahl der negativen Beiträge null, oder vernachlässigbar<sup>15</sup> gering, halten wir die Gesamtbilanz für positiv und damit auch die notwendige Bedingung für erfüllt. In den übrigen Fällen bleibt uns vorerst nur die Einzelfallentscheidung.

### 3 Fallstudien

#### 3.1 Testumgebung und Testobjekte

Aus den gefährlichen Situationen in Tabelle 2 leiten wir Testfälle ab, die wir am Beispiel eines Fahrzeugs des Modells EZ10<sup>16</sup> von Easymile SAS und auch am Beispiel eines Fahrzeugs des Modells ARMA<sup>17</sup> von Navya SA ausführen. Bei beiden Fahrzeugen handelt es sich um autonome Kleinbusse. Sie verfügen über GPS-Antennen, LiDAR-Sensoren und Odometer und

<sup>14</sup>Im Vergleich zur durchschnittlichen menschlichen Fahrleistung.

<sup>15</sup>Was als vernachlässigbar gilt, haben die Anwender der Methode zu entscheiden.

<sup>16</sup>Modellgeneration 2; Software zum Testzeitpunkt im Mai und September 2020: Voyager 7.

<sup>17</sup>Softwareversion 4.11

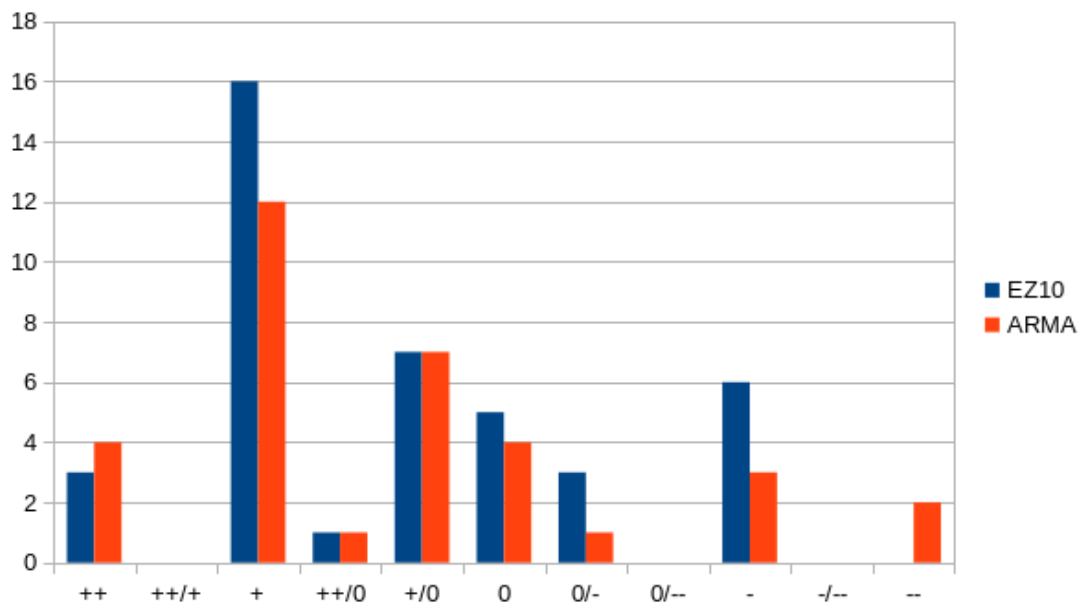


Abbildung 2: Die nach gleicher Art zusammengezählten Einzelbeiträge zur jeweiligen Risikobilanz des EZ10 und des ARMA.

Nutzen diese zur Lokalisierung, Erkennung der Umgebung (insbesondere Hindernisse) und Erkennung anderer Verkehrsteilnehmer.

Beide Fahrzeuge fahren nur auf bekannten Strecken autonom (SAE-Level 4), das heißt nur auf Strecken zu denen ein den besonderen Anforderungen der Fahrzeuge genügendes virtuelles Abbild<sup>18</sup> existiert. Ein solches Abbild dient den Fahrzeugen als Referenz, metaphorisch gesprochen als »virtuelle Schiene«, auf der sie fortan entlang fahren. Die Konsequenz ist, dass beide Fahrzeuge (zum Testzeitpunkt) nicht zu dynamischen Ausweichmanövern fähig sind, falls sich ein Hindernis auf der »virtuellen Schiene« befindet. Beide blieben stattdessen stehen.

Die vom EZ10 autonom befahrbare Strecke befindet sich auf dem Privatgelände der Green-TEC Campus GmbH, die zum ARMA gehörige Strecke befindet sich in Keitum auf Sylt (öffentlicher Straßenverkehr). Für die Durchführung der Tests wählen wir jeweils geeignete Streckenabschnitte und Bereiche aus, etwa lange und gerade Teilstücke, Kreuzungen und freie Flächen.

### 3.2 Testszenerien

Die in Tabelle 6 aufgeführten Testszenerien erläutern wir kurz in der folgenden Aufzählung. Die Nummern der Aufzählung stimmen mit der Nummer der Tests (TNr.) überein. Der Grund für die lückenhafte und unsortierte Nummerierung: Die Testnummer ist eine laufende Nummer, sobald ein Test konzipiert ist, erhält er die niedrigste freie Nummer. Es besteht keine Korrelation zur Unfallursachenummer (UNr.). Da manche Tests konzipiert wurden, letztendlich jedoch nicht im Rahmen der Fallstudien durchgeführt wurden, bestehen Lücken in der Aufzählung. Die nun folgende Ordnung orientiert sich an der Ordnung in Tabelle 6.

- 2 Bei diesem Szenario bewegt sich das Fahrzeug auf einen anderen Verkehrsteilnehmer (Pkw oder Fahrrad / Elektroroller) zu, der sich 25 %, 50 % oder 75 % langsamer als das Fahrzeug bewegt. Alle 6 Kombinationen werden ausgetestet.
- 16 Ein anderer Verkehrsteilnehmer begibt sich (sei es durch Anhalten oder durch Eindringen von der Seite) mit einem Abstand von  $x$  Metern vor das Fahrzeug, wobei  $x$  die Länge des

<sup>18</sup>Im Detail bedeutet dies: Die Strecke, auf der das jeweilige Fahrzeug fahren soll, wird mit einem Spezialfahrzeug des Herstellers oder dem Fahrzeug selbst manuell befahren, während gleichzeitig die Signale der erwähnten Sensoren aufgezeichnet werden. Anschließend wird die so aufgezeichnete Karte manuell nachbearbeitet und auf das Fahrzeug aufgespielt.

Anhaltewegs für einen menschlichen Fahrer mit 1 s Reaktionszeit unter Idealbedingungen ist (konkret sind dies ca. 6,5 m). So kann der Operator noch eingreifen, sollte das Fahrzeug nicht vorab reagieren.

- 17 Laub und Müll (2 Varianten) wird einen Meter vor dem Fahrzeug aufgewirbelt. Ein menschlicher Fahrer würde deshalb keine starke Bremsung einleiten.
- 10 Das Fahrzeug wird von verschiedenen Verkehrsteilnehmern (Pkw oder Fahrrad / E-Roller) in verschiedenen Geschwindigkeiten ( $50 \text{ km h}^{-1}$ ,  $30 \text{ km h}^{-1}$  und  $20 \text{ km h}^{-1}$ ) mit verschiedenen seitlichen Abständen (2 m, 1,5 m, 1 m und 0,5 m) überholt. Nicht jede Kombination wird geprüft (höchstens 7). Beim Pkw nur die schnelleren Geschwindigkeiten und beim Fahrrad / E-Roller nur die kürzeren Abstände.
- 11 Ein Pkw und ein Fahrrad, beziehungsweise E-Roller, befinden sich links vor dem Fahrzeug (wie am Ende eines Überholvorganges). Sie scheren mit  $x + 2 \text{ m}$ ,  $x$ ,  $x - 2 \text{ m}$  und  $x - 4 \text{ m}$  in die Spur des Fahrzeugs ein. Dabei ist  $x$  der Anhalteweg unter Idealbedingungen mit 1 s Reaktionszeit bei einer Geschwindigkeit von rund  $15 \text{ km h}^{-1}$  (hier 6,5 m). Alle Kombinationen, bis auf die Kombination Pkw und  $x - 4 \text{ m}$ , sollten getestet werden.
- 4 Das Fahrzeug biegt ein einer Kreuzung rechts ab (nach Freigabe durch Operator). Hinter der Kurve steht ein Pkw oder eine Person. Einmal stehen die Hindernisse 2 m bis 3 m hinter der Mündung, sodass der Abbiegevorgang nicht abgeschlossen werden kann. Ein anderes Mal stehen sie 5 m bis 6 m hinter der Mündung. Alle vier Kombinationen werden geprüft.
- 18 Das Fahrzeug steht an einem fixierten Ausgangspunkt. Vor, beziehungsweise hinter dem Fahrzeug wird ein Bereich eingezeichnet, wo an bestimmten Stellen (siehe die orangefarbenen Kreuze in Abbildung 3) Hindernisse platziert werden (immer nur eines zurzeit). In einer Variante ist ein Leitkegel das Hindernis, in einer anderen eine Person. In einer weiteren Variante geht eine Person auf den Halbkreisen hin und her – sowohl vor, als auch hinter dem Fahrzeug.
- 13 Dieses und das folgende Szenario betrifft das Abfahren von einer Haltestellen in Verbindung mit dem Einfahren in den fließenden Verkehr. Hier steht die Interaktion mit aussteigenden Fahrgästen im Fokus. Ein Fußgänger steigt aus dem Fahrzeug aus, geht dicht am Fahrzeug vorbei (in Fahrtrichtung), begibt sich vor das Fahrzeug und überquert die Straße. Es werden drei Varianten geprüft in denen sich der Fußgänger mal zügiger und sicher, und mal gemächlicher und unsicher bewegt.
- 14 Im Vergleich zu Nummer 13 steht in diesem Szenario die Interaktion mit dem fließenden Verkehr im Vordergrund. Genau dann wenn das Fahrzeug (etwa von einer Haltestelle) in den fließenden Verkehr einfährt, überholt bereits ein anderer Verkehrsteilnehmer (Pkw oder Fahrrad / E-Roller) das Fahrzeug.
- 7 Das Fahrzeug stößt einmal beim Linksabbiegen und auch beim Rechtsabbiegen auf eine Fußgängerfurt. Beide Fälle werden in zwei Varianten geprüft – je eine pro Seite von welcher der Fußgänger die Furt überquert. Die Begegnung des Fahrzeugs und des Fußgängers wird so koordiniert, dass sie sich (nacheinander) an fünf definierten Punkten (je mit einem Meter Abstand zueinander) auf der Fußgängerfurt treffen. Der mittlere der fünf Punkte ist der, wo die Fahrspur des Fahrzeugs auf die Laufspur des Fußgängers trifft.
- 15 Auf der Fahrspur des Fahrzeugs wird ein Bereich von ca.  $2 \text{ m}^2$  auf verschiedene Weisen verunreinigt. Dafür werden Lehm, Laub und Müll verwendet. Dabei werden die Verunreinigungen einmal flach ausgebreitet und ein anderes Mal werden Häufchen gebildet. Außerdem soll das Fahrzeug einmal in voller Fahrt darüber hinweg fahren und einmal mit den Verunreinigungen vor sich anfahren. Acht Varianten aus diesen Kombinationsmöglichkeiten werden geprüft.



35	4	Erkennt das Fahrzeug während eines Abbiegevorgangs, trotz Freigabe des Operators, ein Hindernis und reagiert entsprechend?	Je nach Entfernung des Hindernisses stoppt das Fahrzeug rechtzeitig in angemessener Intensität, beziehungsweise verweigert von vorne herein die Anfahrt.
36	18	Das Fahrzeug steht. Verschiedene stillstehende und sich bewegende Objekte werden in verschiedenen Bereichen und verschiedenen Entfernungen vor und hinter dem Fahrzeug platziert. Erkennt das Fahrzeug alle Objekte und reagiert angemessen auf diese?	In allen Fällen erkennt das Fahrzeug die Objekte und verweigert die Anfahrt, bzw. kommt nach Anfahrt rechtzeitig wieder zum Stehen.
37	13	Ein Fußgänger steigt an einer Bushaltestelle aus dem Fahrzeug aus. Die Türen werden geschlossen und der Fußgänger bewegt sich eng am Fahrzeug nach vorn, betritt den Bereich vor dem Fahrzeug und quert von dort, mit Aufenthalt zur Verkehrsprüfung, die Straße. Wie verhält sich das Fahrzeug währenddessen?	Das Fahrzeug bleibt, solange der Fußgänger in unmittelbarer Umgebung (insbesondere in der Spur des Fahrzeugs) ist, stehen und lässt den Fußgänger ungehindert gewähren. Erst wenn der Fußgänger diesen Bereich verlässt, möge das Fahrzeug anfahren.
37	14	Wie verhält sich das Fahrzeug beim Überholtwerden, als es von einer Haltestelle abfährt?	Das Fahrzeug bleibt stehen oder fährt mit mäßiger Geschwindigkeit an. Es wahrt ausreichend seitlichen Abstand zum Überholenden – weder behindert noch gefährdet es den Überholenden.
39	7	Wie verhält sich das Fahrzeug gegenüber Fußgängern auf Fußgängerfurten?	Würde eine Kollision des Fahrzeugs mit dem Fußgänger drohen, sollte das Fahrzeug Bremsmanöver einleiten, um einen Abstand zum Fußgänger in Länge des Anhaltewegs, mindestens jedoch in Länge des Gefahrenanhaltewegs zu erzwingen. Droht hingegen keine Kollision ist lediglich darauf zu achten, dass der Fußgänger nicht behindert oder belästigt wird.
74	15	Wie verhält sich das Fahrzeug bei Verschmutzung der Straße durch Laub, Lehm und Müll (Fokus auf Verunreinigung der Oberfläche)?	Das Fahrzeug wird nicht beeinträchtigt und fährt so weiter, als wäre die Verunreinigung nicht vorhanden.

Der öffentliche Straßenverkehr ist reich an Variationen, von der Landschaft, über das Straßennetz und die Verkehrsinfrastruktur, bis zu den Verkehrsteilnehmern. Eine vollständige und präzise Modellierung ist mit heutigen Methoden nicht durchführbar. Daher ist unrealistisch anzunehmen, das Testobjekt in allen Facetten des öffentlichen Straßenverkehrs prüfen zu können. In solchen Fällen ist es üblich, die Zahl der Tests auf kritische Fälle und Grenzfälle zu reduzieren, um mit überschaubarem Aufwand dennoch zu heuristisch guten Ergebnissen zu kommen. Aus diesen Überlegungen folgern wir, insbesondere in Anbetracht der begrenzten Ressourcen der Zielgruppe unserer Methode, dass es genügt aus den gefährlichen Situationen lediglich prototypische Testfälle abzuleiten, die zusätzlich einigen Spielraum bei der Durchführung zulassen – letzteres reflektiert noch einmal die Unschärfe und Variation des öffentlichen Straßenverkehrs.

So ist für uns nur wesentlich, dass der Prüfgegenstand zu der gefährlichen Situation passt und auch tatsächlich überprüfbar ist, dass also ein Bestehen oder ein Fehlschlagen auch entscheidbar ist. Sind bei den Tests Hindernisse, Personen oder Fahrzeuge involviert, halten wir die genauen Ausmaße, die Form, das Gewicht, das Modell und ähnliche Eigenschaften nicht für wesentlich, solange diese Subjekte und Objekte im öffentlichen Straßenverkehr prinzipiell anzutreffen sind. Ein autonomes Fahrzeug der Stufe 5 muss ohnehin mit allen realistischen Varianten der Subjekte und Objekte im Straßenverkehr zurechtkommen und ein Fahrzeug der Stufe 4 muss zumindest erkennen, dass es sich im Zweifelsfall außerhalb seines definierten Einsatzbereichs befindet. Das heißt es gibt keinen Grund spezielle Eigenschaften der Subjekte und Objekte vorzuziehen. Wir akzeptieren daher explizit die Subjekte und Objekte so, wie sie zum Test vorliegen.

Beispiele für weitere Sachverhalte, denen wir uns aus analogen Gründen hingeben, sind:

- das Wetter, sofern der Einsatz der Fahrzeuge bei der konkreten Ausprägung (laut Hersteller) zulässig ist
- die Uhrzeit, beziehungsweise die Tageszeit, und damit verbunden auch die Helligkeit der Umgebung (aus praktischen Gründen testeten wir tagsüber zwischen 9 Uhr und 18 Uhr), weil die konkret vorliegende Fahrzeugtechnik keinen Unterschied zwischen Tag und Nacht erkennt<sup>19</sup>
- die genaue bauliche Realisierung einer Straße, einer Kreuzung oder einer verkehrstechnischen Einrichtung, sofern sie prinzipiell so auch im öffentlichen Straßenverkehr anzufinden wäre, beziehungsweise zulässig ist
- Bauweisen von Bushaltestellen (ebenerdig, erhöht, mit oder ohne Häuschen, ...)

### 3.3 Beobachtungen

Die wesentlichen während der Tests gemachten Beobachtungen sind in Tabelle 8 aufgelistet. Auf eine detaillierte Beschreibung verzichten wir an dieser Stelle, weil die Methode im Vordergrund steht – nicht das Abschneiden der Fahrzeuge.

### 3.4 Bewertung

In Tabelle 8 sehen wir die einzelnen, während der Testfälle ermittelten Risikobilanzbeiträge der Fahrzeuge (zusammengefasst sind sie auch in Abbildung 2 zu finden). Die Verteilung der einzelnen Risikobilanzbeiträge ist sowohl beim EZ10 von Easymile, also auch beim ARMA von Navya, gemäß Abschnitt 2.6 nicht eindeutig. Daher bewerten wir beide Fälle individuell.

Die deutliche Mehrheit (32 von 41) der Risikobilanzbeiträge des EZ10 ist nicht negativ, 19 davon sogar strikt positiv. Lediglich 6 von 41 sind strikt negativ und bei den übrigen drei Beiträgen besteht noch Unsicherheit (die drei  $[0, -]$  Beiträge). Bei den strikt negativen Beiträgen zeigt sich jedoch eine große Schwäche des EZ10: Das Fahrzeug besaß zum Zeitpunkt des Tests nicht die Fähigkeit beim Detektieren eines Objekts mit seinen Sensoren die Natur des Objekts einzuschätzen. Ob es nun ein Verkehrsteilnehmer, Müll, Laub, ein Insekt oder auch nur ein Grashalm ist – gerät es zu nahe ans Fahrzeug, wird ein der Nähe dementsprechend starkes Bremsmanöver ausgelöst. Da diese Schwäche systematisch ist, wiegen unserer Ansicht nach die positiven Beiträge die negativen nicht auf. Wir kommen zu dem Schluss, dass die notwendige Bedingung einer positiven Gesamtrisikobilanz durch den EZ10, zumindest mit der damaligen Version der Steuerungssoftware, nicht erfüllt wird. Wie zu erwarten, ist der EZ10 damit (noch) nicht bereit für die allgemeine Zulassung für den öffentlichen Straßenverkehr.

Auch beim ARMA ist die Mehrheit (28 von 34) der Risikobilanzbeiträge nicht negativ, 16 davon zweifelsfrei positiv. Fünf Beiträge sind strikt negativ, zwei davon besonders deutlich. Jene beiden rühren daher, dass das Fahrzeug in manchen Konstellationen große Schwierigkeiten damit hat, durch seine Sensoren einen Leitkegel zu erkennen. Selbst wenn es einen solchen erkennt, und kurz vor dem Kontakt mit ihm stoppt, nimmt es oftmals unmittelbar danach die Fahrt wieder auf und fährt den Leitkegel an oder überfährt ihn. Die Tatsache, dass das Fahrzeug solch nicht zu vernachlässigende Objekte des Straßenverkehrs nicht zuverlässig erkennt, gepaart damit, dass das Fahrzeug nicht erkennt, dass gerade eine Kollision stattfindet, lässt uns zu dem Schluss kommen, dass die mehrheitlich positiven Beiträge die wenigen, jedoch stark ausgeprägten negativen Beiträge nicht aufwiegen. Daher können wir auch hier die notwendige Bedingung einer positiven Gesamtrisikobilanz durch den ARMA mit der damaligen Version der Steuerungssoftware nicht als erfüllt betrachten. Nicht von dieser Einschätzung berührt wird der prototypische Einsatz unter besonderen Sicherheitsvorkehrungen, wie etwa stark begrenzter Geschwindigkeit und dem Einsatz eines Operators.

<sup>19</sup>Bei anderen Fahrzeugen, etwa welchen die Kamerasysteme zur Umgebungserkennung einsetzen, ist es hingegen wesentlich, zu welcher Tageszeit und Helligkeit getestet wird.



Tabelle 8: Die wichtigsten Beobachtungen aus dem Testbetrieb der Fahrzeuge EZ10 von Easy-mile auf dem GreenTEC Campus und dem ARMA von Navya auf Sylt, aufgeschlüsselt nach Testszenario (TNr.), vergleiche auch Tabelle 6. TE steht für Testergebnis und RB für Risikobizanzbeitrag.

TNr.	Fhz.	Beobachtungen	TE	RB
2	EZ10	Im Rahmen der Durchführung anderer Tests zeigte sich, dass das Fahrzeug stets einen Mindestabstand von ca. 10 m zum vorderen Verkehrsteilnehmer einstellt (etwa durch Verringerung der Geschwindigkeit). Um am Testtag Zeit einzusparen und Redundanz zu vermeiden, deklarierten wir die 6 Testvarianten als Bestanden.	6 · ✓	6 · +
2	ARMA	Gleiche Annahme wie beim EZ10.	6 · ✓	6 · +
16	EZ10	Aus anderen Tests ergab sich, dass das Fahrzeug nicht zulässt, den Sicherheitsabstand nach vorn auf unter 10 m zu reduzieren (Ziel waren 6,5 m). Das Test war daher nicht so wie ursprünglich spezifiziert durchführbar. Stattdessen begab sich eine Person ca. 6,5 m vor dem Fahrzeug auf dessen Fahrbahn und löste so eine Vollbremsung aus. Das Fahrzeug kam rechtzeitig zum Stillstand.	✓	[++, 0]
16	ARMA	Ähnliche Ergebnisse wie beim EZ10.	✓	[++, 0]
17	EZ10	Dicht (ca. 1 m) vor das Fahrzeug geworfener Müll und Laub, sowie dicht vor dem Fahrzeug aufgewirbelter Müll, lösten eine Vollbremsung aus.	2 · ✗	2 · -
17	ARMA	Ähnliche Ergebnisse wie beim EZ10.	2 · ✗	- , [0, -]
10	EZ10	Wurde das Fahrzeug mit 2 m Seitenabstand von einem Pkw überholt, änderte es sein Fahrverhalten nicht. Wurde es mit 1 m bis 1,5 m Abstand überholt, unabhängig davon ob mit Pkw oder E-Roller, verringerte es seine Geschwindigkeit mäßig stark. Beim Überholen mit nur 0,5 m Abstand, löste das Fahrzeug eine Vollbremsung aus.	2 · ✓, 5 · ✗	2 · [+ , 0], 0, 3 · [0, -], -
10	ARMA	Vor Ort erwies sich ein Überholmanöver mit einem Pkw und den geforderten Abständen als nicht durchführbar. Daher entfielen diese Tests. Durchgeführt wurden nur jene mit dem E-Roller. Beim Überholen mit 1,5 m bis 2 m Abstand verringert das Fahrzeug kurzzeitig leicht seine Geschwindigkeit (was wir noch als akzeptabel bewerten); bei kleineren Abständen führte das Fahrzeug eine starke Bremsung, beziehungsweise eine Vollbremsung durch.	✓, 2 · ✗	0, 2 · -
11	EZ10	Vor Ort stellte sich heraus, dass von den sieben geplanten Varianten nur die ersten drei durchführbar sind (die Abstandsermittlung zwischen Fahrzeug und E-Roller war nicht zuverlässig möglich). Die ersten beiden wurden tatsächlich durchgeführt (das Fahrzeug löste in beiden Fällen eine Vollbremsung aus, d. h. es bestand die Tests nicht). Den dritten Test führten wir aus Zeitgründen nicht mehr durch, weil wir annahmen, das Verhalten bliebe gleich.	3 · ✗	3 · -
11	ARMA	Im Nachhinein erwiesen sich die meisten Abstände als zu kurz bemessen (deutlich kürzer als bei den Tests des EZ10) und daher nicht aussagekräftig.	k. A.	k. A.
4	EZ10	Der Pkw und die Person werden in allen vier Varianten erkannt und das Fahrzeug kam rechtzeitig zum Stillstand.	4 · ✓	4 · +
4	ARMA	Ähnliche Ergebnisse wie beim EZ10.	4 · ✓	4 · +
18	EZ10	Das Fahrzeug erkannte das Hindernis und die Person jeweils sicher und verweigert die Anfahrt, sofern dies höchstens 2 m entfernt liegt. War das Objekt weiter weg, nahm das Fahrzeug die Fahrt auf, bis es auf 2 m herankam und blieb dann stehen.	5 · ✓	5 · +
18	ARMA	In den 4 Varianten in denen es galt eine Person zu erkennen (stillstehend oder auch in Bewegung), schnitt das Fahrzeug insgesamt gut ab und erkannte die Personen rechtzeitig (kam zum Teil jedoch sehr nahe heran). Bei der Erkennung eines Leitkegels hingegen kam es vorwärts- und rückwärtsfahrend einige Male zu Zusammenstößen (zum Teil fuhr das Fahrzeug sogar trotz physischen Kontakts weiter).	4 · ✓, 2 · ✗	2 · +, 2 · [+ , 0], 2 · --
13	EZ10	Im Kern wurde das Test in allen Varianten vom Fahrzeug bestanden – auch wenn die Durchführung vor Ort versehentlich unwesentlich von der Spezifikation abwich.	3 · ✓	3 · [+ , 0]
13	ARMA	Ähnliche Ergebnisse wie beim EZ10.	3 · ✓	3 · [+ , 0]

14	EZ10	Wurde das Fahrzeug mit 1,5 m bis 2 m überholt, hatte dies keine Auswirkungen auf dessen Fahrweise. Wurde es nur mit geringem Seitenabstand überholt, verringerte es die Geschwindigkeit, oder kam teilweise sogar kurz zum Stillstand.	2 · ✓	2 · [+ , 0]
14	ARMA	Aufgrund von starken Platzbeschränkungen vor Ort und dementsprechend deutlich anderer Durchführung ist der Test nicht mit dem des EZ10 vergleichbar. Das Fahrzeug behinderte weder den überholenden Pkw, noch den E-Roller. Es fuhr an, kam aber zum Stillstand, bevor die anderen Verkehrsteilnehmer bedrängte oder gefährdete.	2 · ✓	2 · [+ , 0]
7	EZ10	Stand der Fußgänger im Weg, kam das Fahrzeug stets mit 1,5 m bis 2 m Abstand zum Stillstand. Meistens geschah dies auch durch wohl dosiertes Bremsen. In einem Fall fiel das Bremsmanöver jedoch sehr stark aus, sodass Passagiere leicht gefährdet wurden – das Abstand zum Fußgänger betrug jedoch immer noch 1 m bis 1,5 m. Stand der Fußgänger nicht im Weg, fuhr das Fahrzeug an ihm vorbei.	4 · ✓	3 · ++ , +
7	ARMA	Das Fahrzeug erkannte den Fußgänger gut und kam stets rechtzeitig mittels eines angemessenen Bremsmanövers zum Stillstand – wenige Male kam es dem Fußgänger jedoch auf ca. 0,5 m nahe, was als belästigend (jedoch nicht gefährdend) empfunden werden könnte.	4 · ✓	4 · ++
15	EZ10	Durch ruhig auf dem Boden befindliche Verunreinigungen wie Laub, Lehm und Müll wurde das Fahrzeug nicht beeinträchtigt (für aufgewirbeltes Laub vergleiche mit Testfall Nummer 17).	4 · ✓	4 · 0
15	ARMA	Durch ruhig auf dem Boden befindliche Verunreinigungen wie Lehm und Grünschnitt (Laub war am Testort nicht verfügbar) wurde das Fahrzeug nicht beeinträchtigt. Tests mit Müll konnten, wegen starken Windes, nicht durchgeführt werden.	3 · ✓	3 · 0

## 4 Evaluation

Ein Vorteil unserer Methode ist, dass sie auf Systemstests basiert. Es ist daher kein Wissen über den Aufbau und die Implementierung einzelner Komponenten nötig. Auch sind spezielle Testgeräte und -aufbauten unnötig. So ist unsere Methode leicht aus Kundensicht anwendbar. Dem entgegen steht, dass gewisse Fehler so vermutlich nicht aufgedeckt werden. Beispielsweise könnte ein Fehler in der Standby-Komponente eines redundant ausgelegten Systems während eines Systemtests unentdeckt bleiben, weil stets die Primärkomponente aktiv ist, ein Komponententest der Standbykomponente hätte den Fehler jedoch erkennen können.

Der Umfang, die Komplexität und die Anzahl der Systemtests ist allein durch den Anwender bestimmt – wir und auch Dritte erlegen keine Pflichten zur Durchführung und zur Dokumentation auf. Die Folge ist ein schlanker Prüfprozess. Dem Anwender obliegt es selbst einzuschätzen, ob die selbst festgelegte Testabdeckung den eigenen Ansprüchen genügt. Besonders bedacht werden sollte, ob ein Testszenario einmalig oder mehrmalig durchlaufen werden müsse. Wird ein Test lediglich einmalig durchlaufen bleibt die Frage, ob das Testergebnis nur auf die Fähigkeiten des Fahrzeugs zurückzuführen ist, und nicht etwa durch Störfaktoren, oder gar durch Zufall beeinflusst wurde.

Je nach ASIL des gefährlichen Ereignisses nimmt unsere Methode bereits einen Teil der Interpretation ab und fordert nur in uneindeutigen Fällen die Betrachtung des Einzelfalls. Leider treten die uneindeutigen Fälle, jene mit dem ASIL B, siehe auch noch einmal Abbildung 1, recht häufig auf – eindeutige Fälle wie jene mit ASIL D oder QM kommen dagegen selten vor (siehe Tabelle 10 im Anhang). Außerdem ist anzunehmen, dass die Verteilung der einzelnen Risikobilanzbeiträge selten so eindeutig ausfällt, dass ohne Einzelfallbetrachtung eine Aussage zur Gesamtbilanz getroffen werden kann.

Werden Gefährliche Situationen, wie bei uns, so formuliert, dass sie allgemein für autonome Fahrzeuge anwendbar sind und nicht von einem konkreten Modell abhängen, sind sie leicht auf neue Anwendungsfälle übertragbar. Der Aufbau auf bereits vorhandenem Wissen würde so erleichtert.

Zuletzt gehen wir auf die Anforderungen an ein Testkonzept ein, die Wachenfeld und Winner [9, Abschnitt 21.3] stellen. Diese sind (1) Repräsentativität/Validität, (2) Variierbarkeit, (3) Beobachtbarkeit, (4) Ökonomie, (5) Reproduzierbarkeit, (6) Frühzeitigkeit und (7) Sicherheit.

Die Anforderungen 2 und 4 erfüllt unsere Methode per Konstruktion. Anforderung 1 ist nur zum Teil erfüllt. Einerseits testen wir das Fahrzeug als ganzes auf Testfeldern oder in der Praxis im öffentlichen Straßenverkehr. Andererseits überlassen wir es explizit dem Anwender, ein Testszenario in mehreren Varianten, beispielsweise mit unterschiedlichen Umweltbedingungen, durchzuführen. Das die Anforderungen 3 und 7 erfüllt werden liegt vollständig in der Hand der Anwender. Die Anforderungen 5 und 6 sind aus Kundensicht nicht relevant (diese richten sich primär an Hersteller).

## 5 Verwandte Arbeiten

Um die Steuerungssoftware eines autonomen Fahrzeugs zu testen, bieten sich auf der einen Seite Komponententests an, welche die Komponente »Steuerungsprogramm« isoliert überprüfen (etwa durch Softwaretests). Auf der anderen Seite bieten sich Systemtests an, welche die Steuerungssoftware im Zusammenspiel mit dem Gesamtsystem Fahrzeug überprüfen [10]. Systemtests, auf Testfeldern oder in der Praxis, genießen im Allgemeinen eine höhere Validität gegenüber Komponententests [11], sind jedoch aufwendiger in ihrer Umsetzung. Eine Zwischenform stellen Simulationen dar (*software in the loop*), welche insbesondere in der Fahrzeugentwicklung eingesetzt werden [12, 13]. Eine Übersicht hierzu bieten Huang u. a. [14] und Alvarez u. a. [15].

Bei der Konzeption von Testfällen für ein Black-Box-System (ein System über dessen inneren Aufbau und Implementierung dem Tester keine Kenntnisse vorliegen) ist eine Option, risikobasiert vorzugehen [6]. Eine andere Option ist es, in großer Zahl verschiedene Parameter einer Simulation zu variieren, um negative Auswirkungen auf das Testobjekt zu ermitteln [16]. Weiter geht der Ansatz von Mullins, Stankiewicz und Gupta [17], indem sie ein Gütemaß für das Fahrverhalten eines autonomen Fahrzeugs definieren und dann zielgerichtet nach Punkten suchen, in deren Nähe sich der Gütewert deutlich verschlechtert.

Die Unfallstatistik nutzen wir als Referenz für kritische Verkehrssituationen, um gefährliche Ereignisse zu erkennen und damit einen Rahmen für die Konzeption von Testfällen zu schaffen. Die Unfallstatistik basiert auf polizeilich erfassten Verkehrsunfällen. Gambi, Huynh und Fraser [18] verwenden die Unfallberichte direkt, um kritische Verkehrssituationen (in Simulationen) nachzustellen. Zuletzt ist es nützlich, primär aus Sicht von Fahrzeugentwicklern, die in früheren Feldtests aufgetretenen kritischen Verkehrssituationen wiederzuverwerten [11, 19].

## 6 Ausblick

In unserer Arbeit bewerten wir die Risikobilanz von autonomer Fahrleistung mit nahezu rein menschlicher Fahrleistung (StVU 2014). Dieser Vergleich ist prototypisch und entspricht so dem Wortlaut der Regel 2 des Berichts der Ethikkommission. Er berücksichtigt jedoch nicht die Risikoreduktion, welche bereits jetzt auf vorhandene Fahrerassistenzsysteme – das heißt Teilautomatisierung – zurückzuführen ist. Bei der Einführung fortgeschrittener autonomer Fahrfunktionen, bis hin zum vollautonomen Fahrzeug, sollte deren Risikobilanz im Vergleich zur menschlichen Fahrleistung inklusive gängiger Fahrerassistenzsysteme aufgestellt werden. Damit würde sich das Vergleichsobjekt menschliche Fahrleistung stetig verändern. Wir vermuten, dass hiervon ausgehende Risiko nimmt im Laufe der Zeit stetig ab, was zur Folge hätte, dass die Sicherheitsanforderungen an zuzulassende (teil)autonome Fahrfunktionen stetig steigen würden (ansonsten wäre deren Risikobilanz nicht mehr positiv). Unsere Methode ist in der Lage diese Entwicklung abzubilden, indem die ASIL-Bewertung der gefährlichen Situationen stets mit aktuellen Unfallstatistiken durchgeführt wird.

## 7 Zusammenfassung

In dieser Arbeit stellten wir eine Methode vor, mit der die Risikobilanz von autonomen Fahrzeugen im Vergleich zur durchschnittlichen menschlichen Fahrleistung ermittelt werden kann. Die Methode beinhaltet das Erstellen von Systemtests. Im Rahmen von zwei Fallstudien, ein autonomes Fahrzeug auf einem Testfeld des GreenTec Campus und ein anderes autonomes

Fahrzeug im öffentlichen Straßenverkehr in Keitum auf Sylt, stellten wir unserer Methodik folgend Systemtests und entsprechende Testergebnisse vor, um die Risikobilanz beider Fahrzeuge zu ermitteln.

## Literatur

- [1] Michael Stöber, Annelie Möller und Marc-Christian Pieronczyk. „Haftungsrechtliche Probleme des autonomen Fahrens (3 Teile)“. In: *Verkehr und Technik* 5, 6 & 7 (2019).
- [2] Ethik-Kommission. *Automatisiertes und Vernetztes Fahren*. Bericht. Berlin, DE: Bundesministerium für Verkehr und digitale Infrastruktur, Juni 2017.
- [3] Marcus Bäumer u. a. *Fahrleistungserhebung 2014 - Inlandsfahrleistung und Unfallrisiko*. Berichte der Bundesanstalt für Straßenwesen Verkehrstechnik. Bremen: Fachverlag NW, 2017. ISBN: 978-3-95606-337-4.
- [4] Nidhi Kalra und Susan M. Paddock. „Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?“ In: *Transportation Research Part A: Policy and Practice* 94 (2016), S. 182–193. ISSN: 0965-8564. DOI: [10.1016/j.tra.2016.09.010](https://doi.org/10.1016/j.tra.2016.09.010). URL: <http://www.sciencedirect.com/science/article/pii/S0965856416302129>.
- [5] Statistisches Bundesamt (Destatis), Hrsg. *Fachserie 8 Reihe 7 (2016): Verkehr 2014. Verkehrsunfälle*. URL: [https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEHeft\\_derivate\\_00018327/2080700147004\\_Korr15022016.pdf](https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEHeft_derivate_00018327/2080700147004_Korr15022016.pdf).
- [6] ISO 26262-3:2011(E). *Road vehicles – Functional safety – Part 3: Concept Phase*. Geneva, CH, Nov. 2011.
- [7] Statistisches Bundesamt (Destatis), Hrsg. *Qualitätsbericht (2017): Statistik der Straßenverkehrsunfälle*. URL: [https://www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Verkehrsunfaelle/strassenverkehrsunfaelle.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Verkehrsunfaelle/strassenverkehrsunfaelle.pdf?__blob=publicationFile).
- [8] On-Road Automated Driving (ORAD) committee. *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. Juni 2018. DOI: [10.4271/J3016\\_201806](https://doi.org/10.4271/J3016_201806).
- [9] Walther Wachenfeld und Hermann Winner. „Die Freigabe des autonomen Fahrens“. In: *Autonomes Fahren*. Hrsg. von Markus Maurer u. a. Springer Vieweg, Berlin, Heidelberg, 2015, S. 439–464.
- [10] Matthew Schwall u. a. „Waymo Public Road Safety Performance Data“. In: *CoRR* abs/2011.00038 (2020). arXiv: [2011.00038](https://arxiv.org/abs/2011.00038). URL: <https://arxiv.org/abs/2011.00038>.
- [11] Lutz Eckstein und Adrian Zlocki. „Safety potential of ADAS–combined methods for an effective evaluation“. In: *23rd International Technical Conference on the Enhanced Safety of Vehicles (ESV) Seoul, South Korea*. 2013, S. 15–25.
- [12] Shuo Feng u. a. „Intelligent driving intelligence test for autonomous vehicles with naturalistic and adversarial environment“. In: *Nature communications* 12.1 (2021), S. 1–14.
- [13] Cumhur Erkan Tuncali u. a. „Simulation-based Adversarial Test Generation for Autonomous Vehicles with Machine Learning Components“. In: *2018 IEEE Intelligent Vehicles Symposium, IV 2018, Changshu, Suzhou, China, June 26-30, 2018*. IEEE, 2018, S. 1555–1562. DOI: [10.1109/IVS.2018.8500421](https://doi.org/10.1109/IVS.2018.8500421). URL: <https://doi.org/10.1109/IVS.2018.8500421>.
- [14] Wuling Huang u. a. „Autonomous vehicles testing methods review“. In: *19th IEEE International Conference on Intelligent Transportation Systems, ITSC 2016, Rio de Janeiro, Brazil, November 1-4, 2016*. IEEE, 2016, S. 163–168. DOI: [10.1109/ITSC.2016.7795548](https://doi.org/10.1109/ITSC.2016.7795548). URL: <https://doi.org/10.1109/ITSC.2016.7795548>.

- [15] Stephanie Alvarez u. a. „Prospective effectiveness assessment of ADAS and active safety systems via virtual simulation: a review of the current practices“. In: *25th International Technical Conference on the Enhanced Safety of Vehicles (ESV) National Highway Traffic Safety Administration*. 2017.
- [16] Saurabh Jha u. a. „AVFI: Fault Injection for Autonomous Vehicles“. In: *48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops, DSN Workshops 2018, Luxembourg, June 25-28, 2018*. IEEE Computer Society, 2018, S. 55–56. DOI: [10.1109/DSN-W.2018.00027](https://doi.org/10.1109/DSN-W.2018.00027). URL: <http://doi.ieeecomputersociety.org/10.1109/DSN-W.2018.00027>.
- [17] Galen E. Mullins, Paul G. Stankiewicz und Satyandra K. Gupta. „Automated generation of diverse and challenging scenarios for test and evaluation of autonomous vehicles“. In: *2017 IEEE International Conference on Robotics and Automation, ICRA 2017, Singapore, Singapore, May 29 - June 3, 2017*. IEEE, 2017, S. 1443–1450. DOI: [10.1109/ICRA.2017.7989173](https://doi.org/10.1109/ICRA.2017.7989173). URL: <https://doi.org/10.1109/ICRA.2017.7989173>.
- [18] Alessio Gambi, Tri Huynh und Gordon Fraser. „Generating effective test cases for self-driving cars from police reports“. In: *Proceedings of the ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, ESEC/SIGSOFT FSE 2019, Tallinn, Estonia, August 26-30, 2019*. Hrsg. von Marlon Dumas u. a. ACM, 2019, S. 257–267. DOI: [10.1145/3338906.3338942](https://doi.org/10.1145/3338906.3338942). URL: <https://doi.org/10.1145/3338906.3338942>.
- [19] Olaf J Gietelink u. a. „Pre-crash system validation with PRESCAN and VEHIL“. In: *IEEE Intelligent Vehicles Symposium, 2004*. IEEE. 2004, S. 913–918.

## A Zusatzmaterial

Tabelle 10: Alle Unfallursachenkategorien der Straßenverkehrsunfallstatistik von 2014, bewertet mit den Unfallzahlen jenen Jahres. Wo uns möglich, nennen wir bei sich aus der Unfallursache ergebenden Betriebssituationen eine Häufigkeitsklasse. UNr. entspricht Unfallursachenummer,  $En$  entspricht Häufigkeitsklasse und  $Sn$  entspricht Schweregrad.

UNr.	$Sn$	$En$	ASIL
1	2		
2	2		
3	2		
4	2		
10	1		
11	2		
12	2		
13	2		
14	1	4	B
15	1	4	B
16	1		
17	2		
18	1		
19	2		
20	1		
21	2		
22	1		
23	2	3	B
24	1		
25	1		
26	1	4	B
27	1	4	B
28	1	4	B
29	1		
30	2		
31	1		
32	1		
33	0		
35	1	4	B
36	1	4	B
37	1	4	B
38	2	4	C
39	2	4	C
40	2		
41	2		
42	2		
43	1	4	B
44	2	4	C
45	1	4	B
46	2	4	C
47	2	4	C
48	1	4	B
49	1		
50	2		
51	2		
52	1		
53	2		

54	2		
55	2		
60	2		
61	2		
62	2		
63	2		
64	2		
65	2		
66	2		
67	2		
68	2		
69	2		
70	2		
71	2		
72	2	2	A
73	2	3	B
74	2	2	A
75	2		
76	2		
77	2		
78	2		
80	2		
81	2		
82	1		
83	2		
84	2		
85	2	1	QM
86	2		
87	2		
88	2		
89	1		

---