
Otto-von-Guericke Universität Magdeburg



Fakultät für Informatik
Institut für Wissens- und Sprachverarbeitung

Master Thesis

**Schätzen verdeckter Fahrereigenschaften auf Basis des
Fahrverhaltens**

Autor:
Pascal Held

5. April 2011

VOLKSWAGEN
AKTIENGESELLSCHAFT

Außenbetreuer:

Dr. Mark Gonter
Integrierte Sicherheit und Licht
Volkswagen AG
Berliner Ring 2
38440 Wolfsburg

Colin Bauer
Carneq GmbH
Carnotstr. 4
10587 Berlin

Betreuer:

Prof. Dr. Rudolf Kruse
Fakultät für Informatik
Otto-von-Guericke Universität
Universitätsplatz 2
39106 Magdeburg, Germany

Matthias Steinbrecher
Fakultät für Informatik
Otto-von-Guericke Universität
Universitätsplatz 2
39106 Magdeburg, Germany

Held, Pascal:

Schätzen verdeckter Fahreigenschaften auf Basis des Fahrverhaltens

Master Thesis, Otto-von-Guericke Universität

Magdeburg, 2011.

Veröffentlichungen über den Inhalt der Arbeit sind nur mit schriftlicher Genehmigung der Volkswagen AG zugelassen.

Die Ergebnisse, Meinungen und Schlüsse dieser Masterarbeit sind nicht notwendigerweise die der Volkswagen AG.

Diese Masterarbeit wäre ohne die Unterstützung einiger Personen nicht möglich gewesen. Aus diesem Grund möchte ich mich bei allen bedanken, die mich mit Rat und Tat unterstützt haben.

Ein ganz besonderer Dank geht an meinen Betreuer Colin Bauer, der mir immer Feedback und weitere Anregungen für meine Arbeit gegeben hat. Auch war er jederzeit und überall für mich verfügbar.

Auch bei dem Bereich integrierte Sicherheit der Konzernforschung der Volkswagen AG möchte ich mich für die angenehme Arbeitsumgebung und dem Interesse an meiner Arbeit bedanken. Ohne dieses Arbeitsumfeld wäre diese Arbeit nicht möglich gewesen.

Des Weiteren möchte ich mich bei Till Freiburg für seine Vorarbeit und die Einführung in das Thema bedanken. Die Ergebnisse seiner Arbeit und seine Erfahrungen haben mir sehr weitergeholfen und meine Tätigkeit vereinfacht.

Mein Dank geht auch nach Magdeburg an die Arbeitsgruppe Computational Intelligence der FIN. Insbesondere an meinen Betreuer Matthias Steinbrecher, der immer konstruktives Feedback zu meiner Arbeit gab und an Christian Möwes, der mich immer mal wieder mit interessanter Literatur versorgt hat.

Ein Großteil meiner Arbeit bestand in der Durchführung einer Probandenstudie. In diesem Kontext möchte ich mich bei Clements Grunert für seine Hilfe bei der Vorbereitung der Studie und natürlich bei allen Probanden, die freiwillig zwei Stunden Zeit dafür geopfert haben, bedanken.

Vielen Dank euch allen!

Abstract

In modern vehicles the incorporation of assistance systems is constantly increasing. Their goals are, among other things, to reduce accident consequences or completely avoid them altogether. The greatest challenge these systems face is the decision when to intervene. An early intervention can increase the efficiency of the system, but this results, however, in drivers more frequently perceiving the activation as unnecessary. Given a better knowledge about the driver the systems could be specifically adapted to him, potentially increasing the systems' acceptance. Such an estimation of the driver is the goal of this work. The purpose is to estimate hidden driver properties underlying his driving behavior. This estimate is based on the basic assumption that drivers with similar properties behave similarly concerning certain quantities. With the aid of an evolutionary algorithm these quantities as well as their weights are determined. In a study with 26 subjects three hidden target properties were examined and estimated. Between the predicted values and given measurements a correlation of approx. 0.7 was reached. In order to exclude the overlearning of the model, an additional optimization was carried out using random values. In this case only a considerably smaller correlation of 0.36 could be achieved. As a consequence it can be observed that the basic assumption applies at least to the three examined target properties and that these can be estimated based on driving behavior.

Zusammenfassung

In modernen Fahrzeugen werden immer mehr Fahrassistenzsysteme eingesetzt. Diese dienen unter anderem dazu Unfallfolgen zu vermindern oder bestenfalls Unfälle komplett zu vermeiden. Die größte Herausforderung an diese Systeme ist die Entscheidung, wann diese eingreifen. Ein frühes Eingreifen kann den Wirkungsgrad erhöhen, führt aber dazu, dass Fahrer dies häufiger als Falschauslösung wahrnehmen. Durch ein besseres Wissen über den Fahrer könnten die Systeme direkt auf den Fahrer angepasst werden, was dessen Akzeptanz gegenüber der Systeme erhöhen könnte. Eine solche Einschätzung des Fahrers soll mit Hilfe dieser Arbeit ermöglicht werden. Ziel ist es verdeckte Fahrereigenschaften auf Basis seines Fahrverhaltens zu schätzen. Diese Schätzung basiert auf der Grundannahme, dass Fahrer mit ähnlichen Eigenschaften sich bezüglich gewisser Größen ähnlich verhalten. Mit Hilfe eines evolutionären Algorithmus werden diese Größen, sowie deren Gewichtung bestimmt. In einer Studie mit 26 Probanden wurden drei verdeckte Zielgrößen untersucht und geschätzt. Zwischen den vorhergesagten Werten und den gegebenen Messwerten wurde dabei eine Korrelation von ca. 0,7 erreicht. Um ein Überlernen des Modells auszuschließen, wurde zusätzlich eine Optimierung auf Zufallswerte durchgeführt. Dabei konnte nur eine wesentlich geringere Korrelation von etwa 0,36 erreicht werden. Insgesamt kann festgehalten werden, dass die Grundannahme zumindest für die drei untersuchten Zielgrößen zutrifft und dass sich diese Werte auf Basis von Fahrverhalten schätzen lassen.

Inhaltsverzeichnis

| | |
|--|------------|
| Zusammenfassung | I |
| Abbildungsverzeichnis | VII |
| Tabellenverzeichnis | IX |
| Abkürzungsverzeichnis | XI |
| 1 Einführung | 1 |
| 1.1 Hintergrund | 1 |
| 1.2 Motivation | 2 |
| 1.3 Ziele | 2 |
| 1.4 Vorgehen | 3 |
| 2 Datenbasis | 5 |
| 2.1 Fahrerleistungsdatenbank | 5 |
| 2.1.1 Beschreibung der Strecke | 5 |
| 2.1.2 Beschreibung der Fahrten | 7 |
| 2.2 Simulatorstudie des DLR | 8 |
| 2.2.1 Beschreibung der Szenarien | 8 |
| 2.2.2 Aufbau der Datenbank | 9 |
| 2.3 Zusatzstudie | 9 |
| 2.3.1 Versuchsaufbau | 10 |

| | | |
|----------|--|-----------|
| 2.3.2 | Strecke / Szenarien | 12 |
| 2.3.3 | Probanden | 14 |
| 3 | Untersuchung des Fahrverhaltens | 17 |
| 3.1 | Beschreibung der Fahrsituationen | 17 |
| 3.1.1 | Globale Betrachtung | 18 |
| 3.1.2 | Haltevorgang | 18 |
| 3.1.3 | Fahren | 19 |
| 3.1.4 | Freie Vorgänge - Gebundene Vorgänge | 20 |
| 3.2 | Ermittlung von Kenngrößen | 20 |
| 3.2.1 | Auswahl der Kenngrößen | 22 |
| 3.3 | Parameter und Kenngrößen | 23 |
| 3.3.1 | Abstand zum voraus fahrenden Fahrzeug | 23 |
| 3.3.2 | Beschleunigung | 25 |
| 3.3.3 | Fahrpedalwert | 26 |
| 3.3.4 | Geschwindigkeit | 28 |
| 4 | Schätzung des verdeckten Fahrerparameters | 31 |
| 4.1 | Aufzeichnen von Messwerten und Extraktion der Kenngrößen | 32 |
| 4.2 | Normalisieren der Daten | 33 |
| 4.3 | Bestimmung ähnlicher Fahrten | 34 |
| 4.4 | Schätzen des Zielwertes | 37 |
| 4.4.1 | Beispiel für eine Zielwertschätzung | 37 |
| 5 | Bestimmung relevanter Kenngrößen | 41 |
| 5.1 | Übersicht Evolutionärer Algorithmus | 41 |
| 5.2 | Kodierung der Individuen | 42 |
| 5.3 | Startpopulation | 42 |

| | | |
|----------|--|-----------|
| 5.4 | Fitnessfunktion | 43 |
| 5.5 | Auswahlverfahren | 44 |
| 5.5.1 | Elitismus | 44 |
| 5.5.2 | Glücksradauswahl | 44 |
| 5.6 | Mutation | 48 |
| 5.7 | Crossover | 51 |
| 6 | Anwendung und Auswertung | 53 |
| 6.1 | Simulatorstudie | 54 |
| 6.1.1 | Datenbasis | 54 |
| 6.1.2 | Anwendung | 55 |
| 6.2 | Realfahrten | 57 |
| 6.2.1 | Datenbasis | 58 |
| 6.2.2 | Anwendung auf Zielgröße A | 59 |
| 6.2.3 | Anwendung auf Zielgröße B | 60 |
| 6.2.4 | Anwendung auf Zielgröße C | 60 |
| 6.2.5 | Kontrolluntersuchung | 63 |
| 6.3 | Auswertung der Ergebnisse | 64 |
| 7 | Schlussfolgerungen und Abschluss | 65 |
| 7.1 | Zusammenfassung der Ergebnisse | 65 |
| 7.2 | Vergleich mit Erwartungen und Zielen | 66 |
| 7.3 | Einschränkungen und Anwendbarkeit | 67 |
| 7.4 | Erweiterungsmöglichkeiten | 67 |
| 7.5 | Einsatzmöglichkeiten und Nutzen | 68 |
| | Literaturverzeichnis | 71 |
| | Selbstständigkeitserklärung | 73 |

Abbildungsverzeichnis

| | | |
|-----|--|----|
| 2.1 | Darstellung der Teststrecke | 6 |
| 2.2 | Streckenplan des Stadtszenarios | 8 |
| 2.3 | Frontkamera im Versuchsfahrzeug | 10 |
| 2.4 | Ausgabe des ASR300 | 11 |
| 2.5 | komplette Teststrecke | 13 |
| 2.6 | Geschlecht der Probanden | 15 |
| 2.7 | Alter der Probanden | 15 |
| 2.8 | jährliche Fahrleistung der Probanden | 16 |
| 2.9 | Anteil vereiste Fahrbahn | 16 |
| 3.1 | exemplarische Darstellung eines Haltevorgangs | 18 |
| 3.2 | Verteilung der TTC-Messwerte | 24 |
| 3.3 | Verteilung der Zeitlücke-Messwerte | 25 |
| 3.4 | Verteilung der Längs- und Quereschleunigung | 25 |
| 3.5 | Zusammenhang zwischen Längs- und Quereschleunigung | 27 |
| 3.6 | Verteilung der Fahrpedalwerte | 28 |
| 3.7 | Verteilung der Geschwindigkeit | 29 |
| 3.8 | Verteilung der Geschwindigkeitsdifferenz | 29 |
| 4.1 | Ablauf der Fahrerparameterschätzung | 32 |
| 4.2 | Ablauf der Kenngrößenermittlung | 33 |
| 4.3 | Umrechnung Kenngröße in Prozentrang | 34 |

| | | |
|------|---|----|
| 4.4 | Darstellung von Referenzfahrten für zwei Kenngrößen | 35 |
| 4.5 | Darstellung von Referenzfahrten für zwei Kenngrößen, wobei eine skaliert ist | 36 |
| 4.6 | Darstellung des gegebenen (blau) und des geschätzten (rot) Zielwertes | 39 |
| 5.1 | Beispielkonfiguration eines Individuums | 42 |
| 5.2 | Erzeugung der Startpopulation | 43 |
| 5.3 | Entwicklung der Fitness bei Standardglücksradauswahl | 45 |
| 5.4 | Entwicklung der Fitness bei Glücksradauswahl exponentiell skaliert . | 46 |
| 5.5 | Entwicklung der Fitness bei Glücksradauswahl dynamisch skaliert . . | 47 |
| 5.6 | Entwicklung der Fitness mit unterschiedlichen Mutationsintensitäten | 49 |
| 5.7 | Entwicklung der Fitness mit unterschiedlichen Mutationswahrscheinlichkeiten | 50 |
| 5.8 | Uniform Crossover | 51 |
| 6.1 | Beispielgrafik zum Zustand des evolutionären Algorithmus | 53 |
| 6.2 | Generation 1 - Simulatorstudie | 56 |
| 6.3 | Generation 20 - Simulatorstudie | 56 |
| 6.4 | Generation 100 - Simulatorstudie | 57 |
| 6.5 | Generation 1 - Probandenstudie - Kenngröße A | 59 |
| 6.6 | Generation 100 - Probandenstudie - Kenngröße A | 60 |
| 6.7 | Generation 1 - Probandenstudie - Kenngröße B | 61 |
| 6.8 | Generation 100 - Probandenstudie - Kenngröße B | 61 |
| 6.9 | Generation 1 - Probandenstudie - Kenngröße C | 62 |
| 6.10 | Generation 100 - Probandenstudie - Kenngröße C | 62 |
| 6.11 | Generation 1 - Probandenstudie - Zufallswerte | 63 |
| 6.12 | Generation 100 - Probandenstudie - Zufallswerte | 64 |

Tabellenverzeichnis

| | | |
|-----|---|----|
| 2.1 | Aufteilung der Teststrecke nach Straßentyp | 6 |
| 2.2 | Geschwindigkeitsbegrenzungen der Teststrecke der FLDB | 6 |
| 2.3 | Geschwindigkeitsbegrenzungen des Stadtszenarios | 14 |
| 2.4 | Geschwindigkeitsbegrenzungen des Landstraßenszenarios | 14 |
| 4.1 | Zielwerte für Beispielschätzung | 38 |
| 4.2 | Zielwerte und geschätzte Werte für Beispielschätzung | 38 |
| 6.1 | Zielwerte der Simulatorstudie | 55 |
| 6.2 | Zielgrößen der Probandenfahrten | 58 |
| 6.3 | Ergebnisse der Untersuchung | 64 |

Abkürzungsverzeichnis

| | |
|------|--|
| ACC | Adaptive Cruise Control |
| ADTF | Automotive Data and Time-triggered Framework |
| CAN | Controller Area Network |
| DLR | Deutsches Zentrum für Luft- und Raumfahrt |
| FLDB | Fahrerleistungsdatenbank |
| TTC | Time to Collision |

Kapitel 1

Einführung

In heutigen Fahrzeugen kommen immer mehr Fahrassistenzsysteme zum Einsatz. Einige von diesen können durch den Fahrer konfiguriert werden, wie beispielsweise der gewünschte Abstand bei einer Abstandsregelautomatik. Andere hingegen sind fest konfiguriert.

Wenn das Fahrzeug gewisse Rahmenparameter zum Verhalten des Fahrers kennen würde, könnten einige dieser Systeme automatisch (vor-)konfiguriert werden, wodurch sich die Akzeptanz für den Fahrer verbessern könnte.

1.1 Hintergrund

Bei vielen Assistenzsystemen gibt es eine wichtige Frage. In welchem Moment soll das System eingreifen und den Fahrer unterstützen?

Prinzipiell kann durch ein früheres Eingreifen des Systems dessen Wirkung verbessert werden. Es gibt beispielsweise Systeme, die ein Auffahren auf ein vorausfahrendes Fahrzeug vermeiden (Collision Avoidance) oder dessen Folgen zumindest verringern (Collision Mitigation). Sie greifen in Notfallsituationen aktiv in die Fahrdynamik des Fahrzeugs ein, um entsprechende Wirkungen zu erhalten.

Ein früheres Eingreifen solcher Systeme führt zu einem früheren Beginn des Bremsvorgangs. Dies ermöglicht es die Kollision zu verhindern oder reduziert die Kollisionsgeschwindigkeit, was zu verminderten Unfallfolgen führt. Auf der anderen Seite hat ein frühes Eingreifen auch Nachteile für den Fahrer. So kann ein zu frühes Bremsen den Fahrer aber an einem erfolgreichen Ausweichen hindern und somit die Gesamtsituation kritischer gestalten.

Es gibt Fahrer die beispielsweise ein sehr nahes Auffahren noch als unkritisch einstufen. Diese würden dann einen frühen Eingriff als Fehlalarm interpretieren. Treten solche Situationen häufig auf, neigt der Fahrer in der Regel dazu den Fahrassistenten zu deaktivieren und in seinem nächsten Fahrzeug nicht mehr zu bestellen.

Um diesem Problem entgegen zu wirken gibt es im Moment zwei Strategien. Komfortsysteme, wie die Abstandsregelautomatik, lassen sich manuell konfigurieren. Sicherheitssysteme, die hingegen nur in Gefahrensituationen aktiv eingreifen, werden konfiguriert, sodass sie nur auslösen, wenn der Fahrer nicht mehr anders reagieren kann. Im Beispiel der Unfallvermeidung wäre das zu dem Zeitpunkt, in dem es dem Fahrer nicht mehr möglich ist, durch Bremsen und Ausweichen einen Unfall zu verhindern.

Mit einem genaueren Wissen über den Fahrer könnten Alternativlösungen, wie das Ausweichen, bereits früh ausgeschlossen werden. Wenn der Fahrer beispielsweise nicht die Fähigkeit oder Erfahrung oder den Mut besitzt, dieses Manöver durchzuführen. Dieses Wissen lässt sich jedoch meistens nicht direkt vom Fahrverhalten ablesen und werden deswegen verdeckte Eigenschaften genannt.

1.2 Motivation

Durch eine Charakterisierung des Fahrzeugführers während der Fahrt ist es möglich das Fahrzeug näher an den Fahrer anzupassen. Dazu gehören sowohl das Parametrisieren von Fahrassistenzsystemen, als auch das Vorkonfigurieren von Komfortsystemen, wie beispielsweise die Auswahl der Härte des Fahrwerkes.

Besonders Sicherheitssysteme lassen sich auf diese Weise vielleicht so konfigurieren, dass sie etwas früher greifen, ohne dabei gefühlt mehr Fehlauflösungen zu verursachen. Dies würde zur erhöhten Verkehrssicherheit und zu einer besseren Akzeptanz dieser Systeme beitragen.

Im Bereich der Komfortfunktionen ist eine adaptive Anpassung an den Fahrer natürlich auch vorteilhaft. Passt sich das Fahrzeug an den Fahrer an, so könnte das Wohlbefinden des Fahrers steigen.

1.3 Ziele

In dieser Arbeit soll daher ein System entstehen, welches verdeckte Parameter des Fahrers schätzt. Mögliche Werte für einen solchen Parameter sind Erfahrungen im Straßenverkehr, Fähigkeiten im Umgang mit Fahrzeugen, aktuelle Zustände des Fahrers oder Einstellungen des Fahrers bezüglich Regeln und Beschränkungen. Das System soll die Schätzungen während der Fahrt auf Basis des gemessenen Fahrverhaltens kontinuierlich durchführen.

Diese Arbeit basiert dabei auf einer wichtigen Annahme:

Fahrer mit ähnlichen Eigenschaften, verhalten sich bezüglich bestimmten Teilen des Gesamtfahrverhaltens ähnlich.

Diese Annahme ist in sehr vielen Bereichen sicherlich gegeben. So reagiert ein unaufmerksamer Fahrer später auf Gefahrensituationen und ein sportlicher Fahrer wird stärkere Beschleunigungen im Fahrzeug verursachen.

Unter dieser Annahme lassen sich verdeckten Parameter durch den Vergleich mit einer entsprechenden Referenzdatenbank ermitteln. Ein wichtiger Punkt für den Abgleich sind die Vergleichskriterien, sowie deren Gewichtung. In dieser Arbeit wird ein Weg beschrieben, wie diese mit Hilfe eines evolutionären Algorithmus bestimmt werden können.

1.4 Vorgehen

Grundlegend für den Vergleich von Fahrten sind Kenngrößen, die eine spezielle Fahrt beschreiben. Zu Beginn der Arbeit werden hierfür verschiedene Parameter, wie Geschwindigkeit, Beschleunigung oder Abstand zum Vordermann, vorgestellt. Aus diesen Parametern werden Kenngrößen, wie Mittelwert oder Varianz, ermittelt, welche einen Vergleich untereinander ermöglichen. Mit Hilfe dieser werden anschließend die Referenzfahrten analysiert und beschrieben.

Anschließend wird auf das Modell zur Schätzung der verdeckten Fahrerparameter eingegangen. Die Schätzung basiert grundlegend auf der Ähnlichkeit zu Referenzfahrten. Die Ähnlichkeit ist wiederum Abhängig von gewichteten Kenngrößen.

Zur Bestimmung der benötigten Kenngrößen und deren Gewichtung, wird ein evolutionärer Algorithmus vorgestellt. Dieser optimiert die Gewichtung der einzelnen Kenngrößen, so dass ein vorgegebener Zielwert möglichst gut geschätzt werden kann.

Kapitel 2

Datenbasis

Als Grundlage für diese Arbeit dienen aufgezeichnete Fahrten aus zwei großen Quellen. Zum einen wurden 2005 einige Realfahrten sehr detailliert in der sogenannten Fahrerleistungsdatenbank (FLDB) aufgezeichnet. Zum anderen wurde durch das Deutsche Zentrum für Luft- und Raumfahrt DLR eine Simulatorstudie durchgeführt, bei der Probanden ein Stadt- und ein Landstraßenszenario durchfuhren.

Die Fahrerleistungsdatenbank diene in erster Linie dazu die verschiedenen Fahrparameter während normalen Realfahrten zu analysieren. Die Daten der Simulatorstudie wurden zur Entwicklung des später beschriebenen Algorithmus benutzt, da diese Daten unter sehr kontrollierten Bedingungen aufgezeichnet wurden.

Zusätzlich wurde noch eine weitere Studie durchgeführt, bei der die Probanden verschiedene Szenarien im Straßenverkehr durchfuhren. Diese Daten dienen als Grundlage zur Validierung des Algorithmus.

Im Folgenden werden diese drei Datenquellen näher beschrieben.

2.1 Fahrerleistungsdatenbank

Die Datenbank enthält 54 durchgeführte Fahrten. Zu den einzelnen Testfahrten sind die Fahrzeugsensordaten, und GPS-Daten aufgezeichnet. Diese Datenbank bildet ein normales Fahrverhalten im durchschnittlichen Straßenverkehr ab.

2.1.1 Beschreibung der Strecke

Alle Fahrer sind die Strecke aus Abbildung 2.1 in beide Richtungen gefahren. Insgesamt ist die Strecke etwa 108 *km* lang. Sie stellt einen Mix aus Autobahn, Bundes- und Landstraßen sowie Ortsdurchfahrten dar. Die Aufteilung der Strecke in die verschiedenen Abschnittstypen ist Tabelle 2.1 zu entnehmen. Eine Übersicht über die Verteilung der Geschwindigkeitsbegrenzungen ist in Tabelle 2.2 gezeigt.

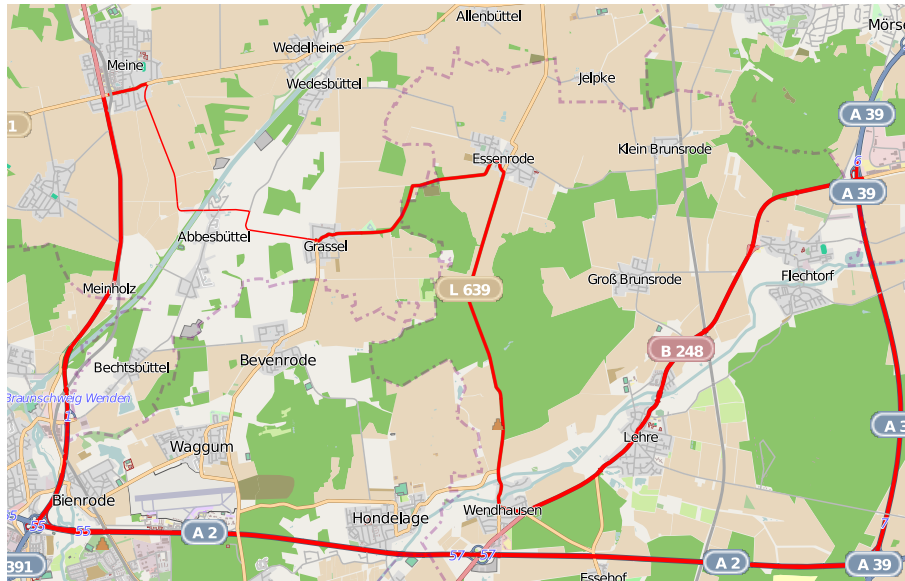


Abbildung 2.1: Darstellung der Teststrecke (entnommen aus der FLDB, ©Karte: OpenStreetMap und Mitwirkende, CC-BY-SA)

| Straßentyp | Länge | Anteil | Geschwindigkeiten |
|------------|---------|--------|--------------------------------|
| innerorts | 13,3 km | 12,3 % | 50 |
| außerorts | 46,7 km | 43,1 % | 60, 80, 100 |
| Autobahn | 48,3 km | 44,6 % | 80, 100, 120, keine Begrenzung |

Tabelle 2.1: Aufteilung der Teststrecke nach Straßentyp

| Geschwindigkeit | Länge | Anteil |
|------------------|---------|--------|
| 50 | 13,3 km | 12,3 % |
| 60 | 8,4 km | 7,7 % |
| 80 | 29,3 km | 27,1 % |
| 100 | 19,6 km | 18,1 % |
| 120 | 28,0 km | 25,8 % |
| keine Begrenzung | 8,9 km | 9,0 % |

Tabelle 2.2: Aufteilung von verschiedenen Geschwindigkeitsbegrenzungen der Teststrecke

Der Stadtanteil der Strecke besteht fast nur aus Ortsdurchfahrten auf der Hauptstraße. Somit lassen sich typische Fahrmanöver der Stadt, wie häufiges Anfahren oder Abbiegen unter Beachtung von Vorfahrt anderer Verkehrsteilnehmer, unzureichend gut analysieren. Auch auf der Autobahn kann es zu Ungenauigkeiten bezüglich der Geschwindigkeitsbegrenzungen kommen, da häufig variable Begrenzungen auf der A2 aktiv sind.

Ablage in der FLDB

Die Teststrecke wurde für die FLDB genau vermessen und mit in der Datenbank gespeichert. Die gesamte Strecke ist in Streckenabschnitte aufgeteilt, welche Straßenstücke zwischen zwei Kreuzungen entsprechen. Zu den Abschnitten sind jeweils die Fahrbahnbegrenzungsmarkierungen als einzelne Linienpunkte in Geokoordinaten gespeichert. Zusätzlich wurden noch alle Verkehrszeichen mit den entsprechenden Koordinaten in die Datenbank aufgenommen.

2.1.2 Beschreibung der Fahrten

Die 54 Testfahrten wurden von insgesamt elf Fahrern durchgeführt. Zu jeder Testfahrt wurde eine Vielzahl von Daten aufgezeichnet. Dazu gehören zum Beispiel die Nachrichten des CAN-Busses, die Informationen aus dem Laserscanner und die Positionen des Ortungssystems.

Hier ein kleiner Ausschnitt aus den gespeicherten Messwerten:

CAN-Bus Zeitstempel, Geschwindigkeit, Gierrate, Querbefleunigung, Lenkradwinkel, Lenkradwinkelgeschwindigkeit, Längsbefleunigung, Fahrpedalwert, Bremslichtschalter, ABS, ESP, Gierrgeschwindigkeit, Bremsdruck, Blinker, Kilometerstand und Motordrehzahl

Laserscanner Objektliste der erkannten Objekte (Objekt-ID, Testfahrt-ID, Zeitstempel, Sensor-schmutzig-Flag, Regenflag, aktuelle Kurvenkrümmung, Objektklassifizierung, Anzahl der Messpunkte und einige weitere Werte, die aber entweder für diese Arbeit nicht relevant sind, oder nicht vollständig vorhanden sind),

Liste der Messpunkte (Punkt-ID, Testfahrt-ID, Objekt-ID, Zeitstempel, Typ, Koordinaten relativ zum Auto, Geokoordinaten sowie Informationen zur Bewegung des Punktes, die aber leider häufig nicht korrekt gespeichert sind)

GPS-Daten Testfahrt-ID, Zeitstempel, Geokoordinaten in Länge und Breite, sowie als Rechts- und Hochwert, Höhe über n.N., Geschwindigkeit, Orientierung, sowie einige für diese Arbeit irrelevante Informationen

2.2 Simulatorstudie des DLR

In einer Simulatorstudie durch das DLR fuhren 21 Probanden jeweils ein Stadt-, sowie ein Landstraßenszenario.

2.2.1 Beschreibung der Szenarien

Die Studie besteht aus zwei Teilszenarien, die unabhängig von einander zu fahren sind. Dabei handelt es sich um ein Stadtszenario, in dem typische Stadtsituationen, wie z. B. Anhalten an einer Ampel simuliert werden und um eine Landstraßensimulation, in der speziell Überholvorgänge betrachtet werden.

Stadtszenario

In dem Stadtszenario fuhren die Probanden die in Abbildung 2.2 gezeigte Strecke durch eine virtuelle Stadt. Dabei wurde besonders auf Ampelsituationen und Abbiegevorgänge eingegangen.



Abbildung 2.2: Streckenplan des Stadtszenarios (aus der Dokumentation der Studie)

An den Ampeln werden insgesamt sechs verschiedene Situationen simuliert. Es gibt drei verschiedene Aufgabentypen, bei denen jeweils ein oder kein Fahrzeug vorausfährt. Bei der einfachsten Aufgabe handelt es sich um Ampeln, die die gesamte Zeit auf grün stehen. Bei den leichten Aufgaben schalten die Ampeln vier Sekunden bevor das Egofahrzeug die Ampel erreicht auf gelb und eine Sekunde vorher auf rot. Zusätzlich gibt es noch schwere Aufgaben, bei denen die Ampeln 2,5 Sekunden vorher auf gelb und eine Sekunde vorher auf roten schalten.

Am Ende des Szenarios wird eine Unfallsituation provoziert, indem ein vorausfahrendes Fahrzeug sehr stark verzögert.

In dem gesamten Stadtgebiet herrscht durchschnittlicher Verkehr und eine Richtgeschwindigkeit von 50 km/h .

Landstraßenszenario

Im Landstraßenszenario sollten die Probanden auf einer kurvigen Landstraße fahren. Auf der gesamten Strecke war eine Geschwindigkeitsbegrenzung von 80 km/h vorgeschrieben. Bei zu starken überschreiten dieser Geschwindigkeit wurde der Proband darauf hingewiesen. Während der Fahrt stießen die Probanden immer wieder auf Fahrzeuge, die mit recht geringer Geschwindigkeit auf der eigenen Spur fahren. Zusätzlich gab es regelmäßig Gegenverkehr, so dass einige der langsamen Fahrzeuge nicht direkt überholt werden konnten.

2.2.2 Aufbau der Datenbank

Für jede Simulatorfahrt und für jedes Teilszenario existieren zwei Tabellen, in denen die Messwerte getrennt für das Egofahrzeug und die Werte für alle anderen Verkehrsteilnehmer gespeichert sind. Bei den Stadtszenarien gibt es zusätzlich noch eine Tabelle in der die aktuellen Zustände der Ampeln abgelegt sind.

2.3 Zusatzstudie

In einer ergänzenden Studie sollten zusätzliche Daten gesammelt werden, da die Strecken der FLDB sehr stark durch Autobahnen und Landstraßen geprägt sind. Aus diesem Grund wurde eine Strecke gewählt, bei der die Probanden einen wesentlich größeren Anteil an Stadtbereichen absolvieren mussten.

Die Gesamtstrecke ist in drei Szenarien unterteilt, so dass die Probanden je ca. $\frac{1}{3}$ Autobahn, Landstraße und Stadtverkehr fahren. Insgesamt wurden 26 Probanden untersucht.

2.3.1 Versuchsaufbau

Die Versuchsfahrten wurden mit einem Versuchsträger durchgeführt, der mit zusätzlichen Sensoren und Aufzeichnungssystemen ausgestattet wurde. Diese sollen hier kurz vorgestellt werden.

Aufzeichnung der CAN-Nachrichten

Wie bereits in der FLDB wurden auch während dieser Versuche die Nachrichten des Antriebs-CAN aufgezeichnet. Jedoch sind diese noch nicht gefiltert, so dass aus Sicht des Fahrzeugs die komplette Fahrt, inklusive aller ausgelösten Sicherheitssysteme, nachvollzogen werden kann.

Frontkamera

In dem Fahrzeug ist an der Windschutzscheibe eine Kamera montiert, die während der Fahrt ein Bild in Fahrtrichtung aufzeichnet. Das aufgezeichnete Video wird durch den Messrechner mit den aufgezeichneten CAN-Nachrichten synchronisiert, so dass sich die komplette Fahrt am Rechner rekonstruieren lässt. Durch das aufgezeichnete Video lassen sich Situationen, in denen auffällige Messwerte aufgezeichnet wurden, besser aufklären. In Abbildung 2.3 ist ein Bild der Kamera während der Fahrt zu sehen.



Abbildung 2.3: Frontkamera im Versuchsfahrzeug

Radar und Laserscanner

Das Versuchsfahrzeug ist mit einem Radarsystem (ASR300 von Continental) und mit einem Laserscanner (IDIS 2 der Firma Hella) ausgestattet. Beide Systeme sind nach vorne gerichtet und können Objekte und Hindernisse vor dem Fahrzeug bis zu einer Entfernung von 200 Metern erkennen und verfolgen.

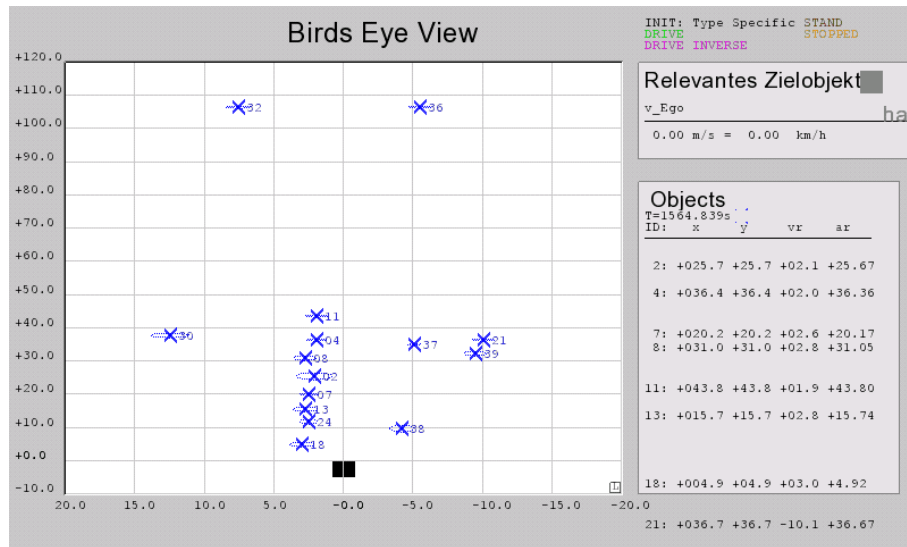


Abbildung 2.4: Ausgabe des ASR300

In Abbildung 2.4 sind die Positionsinformationen der Objekte des ASR300 zu sehen. Abbildung 2.3 zeigt die gleiche Situation aus der Frontkamera. Es ist zu erkennen, dass sowohl stehende Objekte (Leitplanke, Objekte 2, 4, 7, 8, 11, 13, 18 und 24), als auch fahrende Objekte auf der eigenen (vorausfahrende Fahrzeuge, Objekte 37 und 38) und auf der Gegenseite (LKW, Objekt 30) erkannt werden. Solange sich ein Objekt im Sichtbereich des Sensors befindet wird diesem eine Identifikationsnummer zugewiesen mit dessen Hilfe die Bewegung des Objekt analysiert werden kann.

Mit Hilfe dieser Sensorsysteme kann der Abstand, die Geschwindigkeit und Bewegungsrichtung zum vorausfahrenden Fahrzeugs bestimmt werden. Auch diese Informationen werden mit Hilfe von CAN-Nachrichten an den Messrechner im Fahrzeug übertragen.

Digitale Karte

Neben den technischen Systemen ist im Fahrzeug auch ein Informationssystem eingerichtet. Auf dem Fahrzeugrechner wurde eine digitale Karte installiert. Mit Hilfe eines GPS-Systems lässt sich so, ähnlich wie bei einem Navigationsgerät, die aktuelle Position des Fahrzeugs bestimmen. Über die digitale Karte lassen sich so verschiedene Informationen über die gerade befahrene Straße ermitteln. In dieser Arbeit ist hauptsächlich die erlaubte Geschwindigkeit relevant. Es werden jedoch auch Informationen, wie Kurvigkeit, Abstand zur nächsten Kreuzung und einige Umfeldinformationen, verfügbar gemacht.

Fragebogen

Im Anschluss an die durchgeführten Versuchsfahrten wurden die Probanden gebeten einen Fragebogen zu beantworten. Dieser beinhaltete Fragen zur Person, zum allgemeinen Fahrverhalten und zu der durchgeführten Fahrt. Aus diesen Fragen sind, die im letzten Abschnitt der Arbeit genutzt, Zielgrößen abgeleitet worden.

2.3.2 Strecke / Szenarien

Als Teststrecke wurde eine Route geplant die insgesamt ca. 105 km lang ist. Durchschnittlich beträgt die Fahrzeit etwa eine Stunde und 40 Minuten. Die Gesamtroute ist in drei Abschnitte bzw. Szenarien unterteilt, die direkt nacheinander durchfahren wurden.

In Abbildung 2.5 ist die komplette Versuchsstrecke dargestellt. Im Vergleich zu der FLDB unterscheidet sich diese Strecke hauptsächlich im Stadtbereich. So wurde ein großes Teilstück quer durch Braunschweig hinzugefügt um auch häufiger Anhalte- oder Abbiegevorgänge analysieren zu können. Der Landstraßenabschnitt gleicht größtenteils der FLDB um auf diesem Abschnitt einen Vergleich zwischen diesen Fahrten und den Fahrten der FLDB zu gewährleisten. Die A2 wurde in dieser Studie nicht befahren, da dort dynamische Geschwindigkeitsbegrenzungen eingerichtet werden können. Dies würde eine Vergleichbarkeit zwischen den einzelnen Fahrten erschweren.

Autobahnabschnitt

Im ersten Drittel der Fahrt wurde hauptsächlich das Verhalten auf der Autobahn untersucht. Die Teststrecke führte ca. 44 km über die A39. Die Autobahn ist auf der gesamten Strecke zweispurig. In den meisten Bereichen gibt es keine Geschwindigkeitsbegrenzungen. Während der Durchführung der Studie war im südlichen Bereich der Autobahn eine Baustelle eingerichtet.

Interessante Größen, die in diesem Abschnitt gemessen wurden waren das Auffahr- und Überholverhalten des Fahrers, sowie der Umgang mit Geschwindigkeitsbegrenzungen oder die favorisierte Geschwindigkeit.

Stadtverkehr

Die Fahrt in der Stadt stellt das zweite Drittel der Gesamtfahrt dar. Im Stadtverkehr wurden die Probanden mit verschiedensten Verkehrssituationen konfrontiert. Dazu gehören unter anderem das Verhalten an Ampeln oder während des Abbiegens. Zusätzlich gibt es Bereiche mit unterschiedlichen Geschwindigkeitsbegrenzungen. So ist in Wohngebiete ein anderes Verhalten als auf Hauptstraßen zu erwarten. In Tabelle 2.3

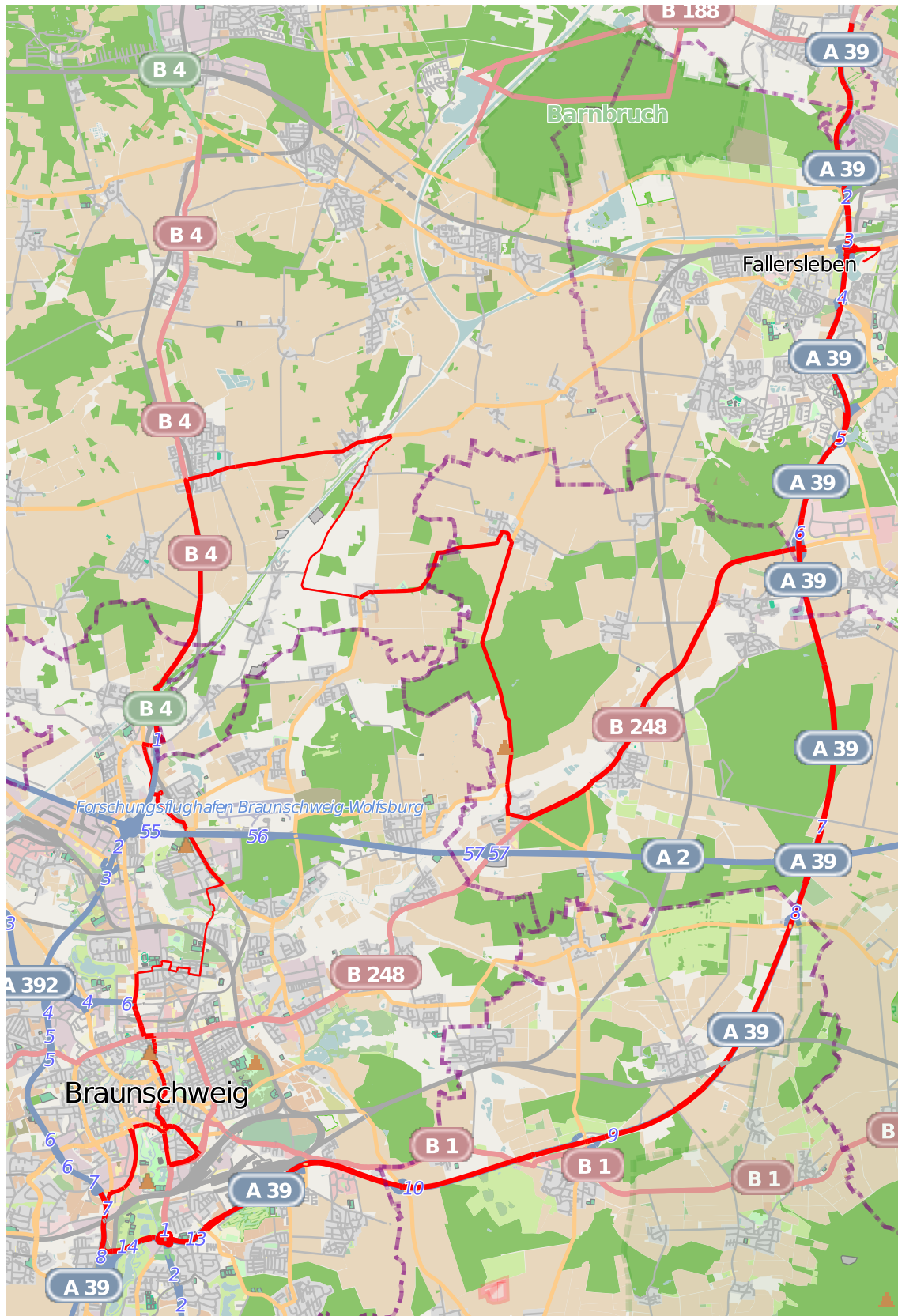


Abbildung 2.5: komplette Teststrecke (©Karte: OpenStreetMap und Mitwirkende, CC-BY-SA)

wird eine Übersicht über den Anteil der verschiedenen Geschwindigkeitsbegrenzungen innerhalb des Stadtszenarios gegeben.

| Geschwindigkeit | Länge | Anteil |
|-----------------|---------|--------|
| 30 | 1,9 km | 12,4 % |
| 50 | 13,4 km | 87,6 % |

Tabelle 2.3: Aufteilung von verschiedenen Geschwindigkeitsbegrenzungen des Stadtszenarios

Landstraße

Im letzten Abschnitt der Teststrecke fuhren die Probanden hauptsächlich über Land- und Bundesstraßen. Jedoch gibt es auch einen Anteil Autobahn und Ortsdurchfahrten. Eine genaue Aufteilung der Geschwindigkeitsbegrenzung ist der Tabelle 2.4 zu entnehmen.

| Geschwindigkeit | Länge | Anteil |
|------------------|---------|--------|
| 30 | 0,3 km | 0,7 % |
| 50 | 7,6 km | 18, % |
| 60 | 2,2 km | 5,3 % |
| 70 | 0,4 km | 1,0 % |
| 80 | 8,3 km | 19,9 % |
| 100 | 20,9 km | 50,0 % |
| 120 | 2,1 km | 5,0 % |
| keine Begrenzung | 1,5 km | 3,6 % |

Tabelle 2.4: Aufteilung von verschiedenen Geschwindigkeitsbegrenzungen des Landstraßenszenarios

Der Landstraßenteil deckt sich größtenteils mit dem der FLDB. Jedoch endet dieser Abschnitt wieder am Mobile Life Campus der Volkswagen AG.

2.3.3 Probanden

Für die Studie standen nur eine recht homogene Gruppe von Personen zur Verfügung. Dadurch ist die durchgeführte Studie nicht repräsentativ für alle Fahrer, kann aber trotzdem für eine prototypische Umsetzung angewendet werden. Für einen produktiven Einsatz sollte eine repräsentative Normstichprobe aufgenommen werden, was aber im Rahmen dieser Arbeit nicht möglich war, da dafür mehrere hundert Fahrten aufgezeichnet werden müssen.

Insgesamt haben 26 Probanden an der Studie teilgenommen, von denen sechs weiblich waren, siehe Abbildung 2.6. Die meisten der Probanden waren unter 30 Jahren alt. In Abbildung 2.7 ist eine detaillierte Darstellung vom Alter zu sehen.

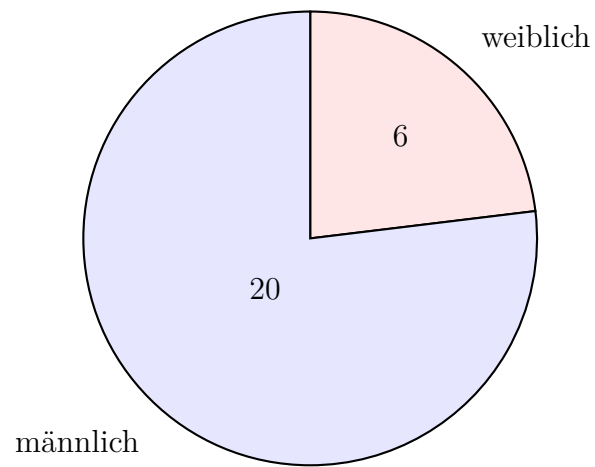


Abbildung 2.6: Geschlecht der Probanden

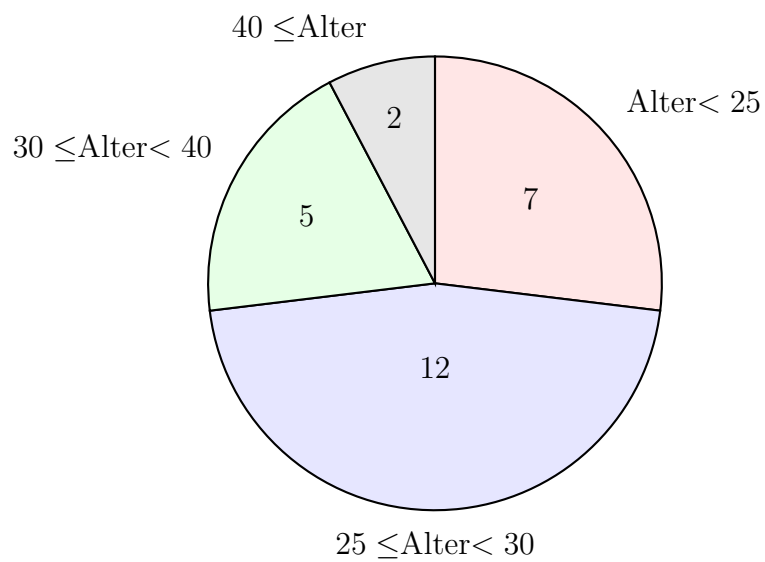


Abbildung 2.7: Alter der Probanden

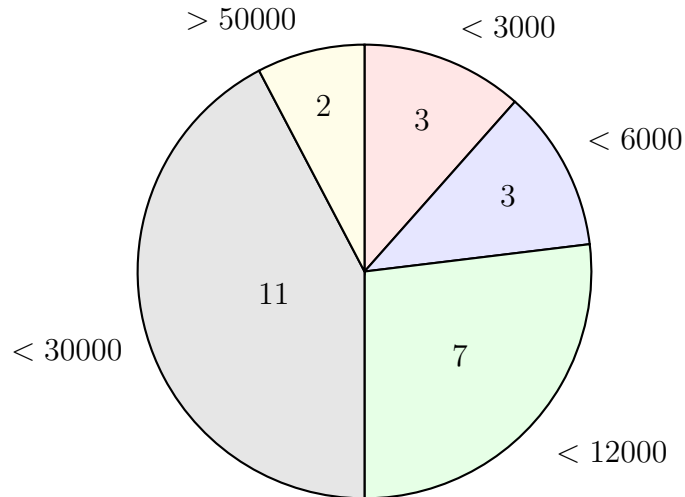


Abbildung 2.8: jährliche Fahrleistung der Probanden

Wie zu erwarten, unterscheiden sich die Fahrleistungen der Probanden sehr stark. Die meisten liegen bei einem Wert zwischen 12000 km/Jahr und 30000 km/Jahr . In Abbildung 2.8 ist die genaue Verteilung zu sehen. Die Einteilung erfolgte dabei analog zu der Einteilung des DLR in der Simulatorstudie. Es gab keine Probanden, die zwischen 30000 km/Jahr und 50000 km/Jahr pro Jahr fahren.

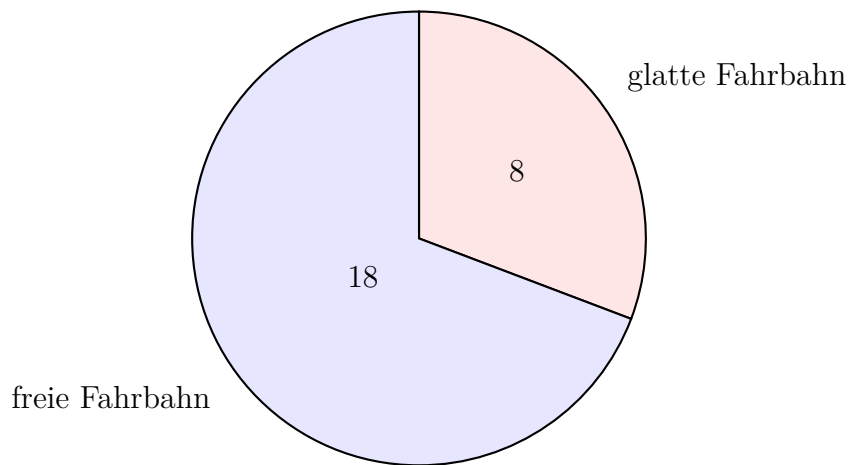


Abbildung 2.9: Anteil vereiste Fahrbahn

Da die Studie in den Wintermonaten durchgeführt wurde, mussten einige Probanden witterungsbedingt auf Eis und Schnee fahren. Insgesamt betraf dies 8 der 26 Probanden, siehe Abbildung 2.9. Bei der Betrachtung der Zielgrößen muss darauf geachtet werden, in wie weit das Fahren auf glatter Fahrbahn das Fahrverhalten beeinflusst.

Kapitel 3

Untersuchung des Fahrverhaltens

Grundlage der Zielgrößenschätzung soll eine Auswertung des Fahrverhaltens während der Fahrt sein. Zu diesem Zweck ist es nötig, dass das Fahrverhalten hinsichtlich verschiedener Parameter näher beschrieben wird.

In dieser Arbeit werden insbesondere das Fahren und die Haltevorgänge näher betrachtet. Dabei wird jeweils zwischen freien und gebundenen Manövern unterschieden. Frei bedeutet, dass es kein vorausfahrendes Fahrzeug auf der Fahrspur gibt. Bei gebundenen Manövern befindet sich der Fahrer hinter einem vorausfahrenden Fahrzeug.

Zu den genannten Fahrsituationen werden gewisse Parameter bestimmt, welche das Verhalten des Fahrers gut beschreiben. Auf folgende Parameter soll dabei in Abschnitt 3.3 näher eingegangen werden, da diese nach Freiburg [Fre10] regelmäßig auftreten und leicht bestimmbar sind:

- Abstand zum vorausfahrenden Fahrzeug
- Beschleunigung
- Fahrpedalwert
- Geschwindigkeit
- Querschleunigung
- Zeit bis zur Kollision
- Zeitlücke

3.1 Beschreibung der Fahrsituationen

In dieser Arbeit soll nicht nur globales Fahrverhalten des Fahrers betrachtet werden, sondern auch auf häufig auftretende Fahrsituationen eingegangen werden. Dabei werden insbesondere das Fahren und das Verhalten bei Haltevorgängen betrachtet. Da diese zusätzlich noch in freie und gebundene Vorgänge unterteilt werden, ergeben sich folgende Konstellationen:

- globales Verhalten
- freie Haltevorgänge
- gebundene Haltevorgänge
- freies Fahren
- gebundenes Fahren

3.1.1 Globale Betrachtung

Neben expliziten Fahrsituationen sollen auch Parameter für die gesamte Fahrt, wie in [Fre10] beschrieben, betrachtet werden. Bei der Parameterauswahl durch den evolutionären Algorithmus können so auch Modelle entstehen, für die keine Extraktion von Fahrsituationen nötig ist.

3.1.2 Haltevorgang

Besonders bei der Fahrt in Städten oder Ortschaften kommt es häufig zu Haltevorgängen. Diese treten beispielsweise an Ampeln oder Kreuzung auf.

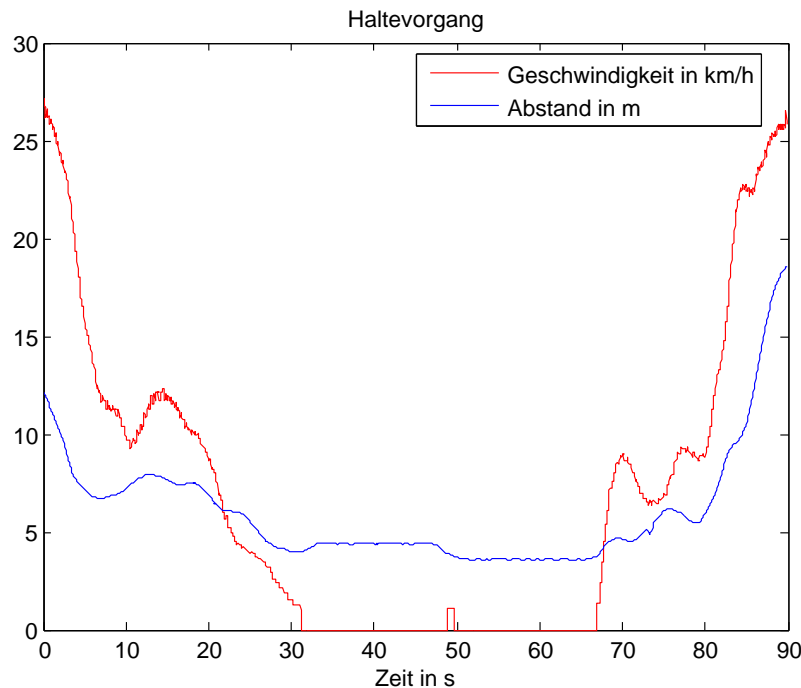


Abbildung 3.1: exemplarische Darstellung eines Haltevorgangs

In den Betrachtungen dieser Arbeit gehören zu einem Haltevorgang drei Phasen. Diese sind das Anhalten, sprich das Reduzieren der Geschwindigkeit bis das Fahrzeug steht, der Stand an sich und das Anfahren, bei dem der Fahrer wieder auf die allgemeine Fahrgeschwindigkeit beschleunigt. Diese Phasen lassen sich sehr gut in Abbildung 3.1 erkennen.

Anhalten

Damit ein Haltevorgang detektiert werden kann, ist es nötig, gewisse Bedingungen zu definieren, die genau dieses Manöver beschreiben. In dieser Arbeit wird dabei das Unterschreiten einer gewissen Geschwindigkeit als Startbedingung für das Anhalten genutzt. Diese wird auf 25 km/h festgelegt, da diese Geschwindigkeit in der Regel bei einer normalen Fahrt überschritten wird. Ausnahmen stellen hier insbesondere verkehrsberuhigte Bereiche dar. Diese führen hier entsprechend zu Störungen und werden deswegen aus der Betrachtung ausgeschlossen. Das Anhalten ist beendet, sobald das Fahrzeug steht, sprich eine Geschwindigkeit von 0 km/h erreicht hat.

In der kontinuierlichen Analyse der Fahrt kann es vorkommen, dass die Startbedingung für das Anhalten erreicht wird, es aber tatsächlich nicht zu einem Stillstand kommt. Um dies zu beheben, wird in der Betrachtung nur die Zeit ab dem letzten Eintreten dieser Bedingung vor dem Stoppen bewertet. Somit können Situationen in denen der Fahrer aus anderen Gründen langsamer wird von der Auswertung ausgeschlossen werden.

Anfahren

Nachdem der Fahrer gehalten hat, wird er in der Regel weiterfahren und somit seine Geschwindigkeit erhöhen bis er seine bevorzugte Reisegeschwindigkeit hat. Auch für das Anfahren muss eine Endbedingung definiert werden. Analog zum Anhalten wird hier das Erreichen der Geschwindigkeit von 25 km/h festgelegt.

Während eines Haltevorganges kann es vorkommen, dass der Fahrer kurz anfährt, ohne aber die Endbedingung zu erreichen. Dies kann beispielsweise dann auftreten, wenn er an einer Ampel auf das voraus fahrende Fahrzeug aufrückt, nachdem bei einer kurzen Grünphase nur einige Autos die Ampel passieren konnten. Um diese Stördaten aus der Analyse heraus zu filtern, werden die Fahrdaten während kurzer Fahrphasen innerhalb eines Haltevorgangs ignoriert.

Stand

Als Stand gilt die gesamte Zeit in der das Fahrzeug während eines Haltevorgangs steht, sprich eine Geschwindigkeit von 0 km/h hat. Sollte es während einer Standphase zu kurzen Fahrsituationen kommen, wie gerade beschrieben, werden diese entsprechen nicht mit betrachtet.

3.1.3 Fahren

Als Fahren wird in dieser Arbeit die gesamte Zeit zwischen Haltevorgängen betrachtet. Durch die Randbedingungen der Haltevorgänge besitzt das Fahrzeug beim Fahren eine Geschwindigkeit von über 25 km/h . Allerdings zählen auch Bereiche, in denen

das Fahrzeug nur temporär diese Geschwindigkeit unterschreitet ohne dabei zum Stillstand zu kommen, zum Fahren.

3.1.4 Freie Vorgänge - Gebundene Vorgänge

Ein weiteres Merkmal für die hier vorgestellten Fahrsituationen ist der voraus fahrende Verkehr. So ist es möglich, dass der Fahrer quasi alleine auf der Straße ist und frei fahren kann, ohne dass er einem anderen Fahrzeug hinterher fährt. Eine Unterscheidung ist dahingehend sinnvoll, da der Fahrer bei freier Fahrt die Geschwindigkeit frei regeln kann, diese bei der gebundenen Fahrt jedoch vom vorausfahrenden Fahrzeug begrenzt ist. Dafür lässt sich bei der gebundenen Fahrt sehr gut der Abstand zum Vordermann ermitteln, welcher bei der freien Fahrt wenig Sinn ergibt.

Zur Beschreibung der aktuellen Situation wird mit Hilfe von Abstandsmesssystemen, wie Laserscanner, Radar- oder Ultraschallsensoren, der Abstand zum vorausfahrenden Fahrzeug ermittelt. Erkennt das System kein Objekt in der Fahrspur oder ist das erkannte Objekt zu weit entfernt, so wird davon ausgegangen, dass es sich um eine freie Fahrt handelt. Andernfalls wird eine gebundene Fahrt angenommen.

3.2 Ermittlung von Kenngrößen

Während einer Fahrt fallen sehr viele Messwerte an, die nicht direkt miteinander vergleichbar sind. Aus diesem Grund ist es notwendig, dass aus den einzelnen Messreihen wenige Kenngrößen ermittelt werden, welche sich einfach vergleichen lassen. In Frage dafür kommen alle charakteristischen Werte der deskriptiven Statistik. Diese sind nach Steinbrecher [Ste10] zum Beispiel:

Mode x^*

Der Mode ist bei einer diskreten Menge von Messwerten die Ausprägung des Attributes mit dem häufigsten Vorkommen.

Median \tilde{x}

Der Median beschreibt den mittleren Wert aller Werte der Zufallsstichprobe. Ist x eine sortierte Liste der Stichprobe mit n Elementen, so ist der Median wie folgt definiert.

$$\tilde{x} = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & n \text{ ist ungerade} \\ \frac{1}{2} \left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)} \right) & n \text{ ist gerade} \end{cases}$$

Arithmetisches Mittel \bar{x}

Das arithmetische Mittel beschreibt den Mittelwert aus n Elementen, so wie er allgemein bekannt ist. Es wird folgendermaßen berechnet:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Da es sich bei Messwerten im Fahrzeug nur um eine Stichprobe des Fahrverhaltens handelt, sollte eher folgender erwartungstreuer Schätzer verwendet werden:

$$\bar{x} = \frac{1}{n+1} \sum_{i=1}^n x_i$$

Range R

Der Range beschreibt den Umfang des Wertebereiches.

$$R = x_{max} - x_{min}$$

 p -Quantile

Das p -Quantil beschreibt den Wert unter dem p -Prozent aller Messwerte liegen. Das 50 %-Quantil entspricht genau dem Median.

Varianz v^2 bzw. s^2

Die Varianz beschreibt die durchschnittliche quadratische Abweichung zum Mittelwert und ist definiert als:

$$v^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Auch bei der Varianz sollte eher ein erwartungstreuer Schätzer verwendet werden, da es sich um eine Schätzung basierend auf einer Stichprobe handelt. Folgender Schätzer kann dafür verwendet werden:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Standardabweichung s

Die Standardabweichung ist die Quadratwurzel der Varianz.

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

3.2.1 Auswahl der Kenngrößen

Im Fahrzeug sollen die Kenngrößen zu den einzelnen Fahrparametern kontinuierlich ermittelt werden. Da die Messwerte ca. zehn Mal pro Sekunde anfallen eignen sich besonders Kenngrößen, die sich beim Eintreffen von neuen Messwerten leicht aktualisieren lassen, ohne dass alle bisherigen Messwerte gespeichert werden müssen.

Das arithmetische Mittel lässt direkt aus der Summe sowie der Anzahl der bisherigen Messwerte berechnen. Beide Werte lassen sich leicht aktualisieren. Auch Minimal- und Maximalwerte lassen sich aus den Werten, sowie den eintreffenden Messwerten leicht neu bestimmen, ohne dass alle Messwerte bekannt sein müssen.

Somit sind das arithmetische Mittel, sowie der Range besonders geeignet. Mit Hilfe des Verschiebungssatzes [TSWH09, S. 33] lässt sich die Varianz so umformen, dass auch sie berechnet werden kann, ohne dass alle Datensätze gespeichert werden müssen.

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right)$$

Ein wenig schlechter eignet sich der Mode, da für ihn zwar nicht alle Messwerte, aber zumindest alle möglichen Messwerte mit einem entsprechenden Zähler, zwischengespeichert werden müssen.

Der Median, sowie beliebige andere p -Quantile lassen sich nur mit Hilfe aller Messwerte bestimmen. Aus diesem Grund werden sie in dieser Arbeit nicht weiter betrachtet, auch wenn der Median nach Steinbrecher [Ste10] stabiler gegen Ausreißer und Messfehler ist.

Auch der Range wird in dieser Arbeit keine Anwendung finden, da er auf Extremwerten der einzelnen Messreihen basiert. Nach Freiburg [Fre10] treten diese jedoch zum Teil erst sehr spät und in unregelmäßigen Abständen auf. Auch können diese Extremwerte durch einzelne Messfehler und Ausreißer gestört werden.

Eine weitere Kenngröße die leicht kontinuierlich zu bestimmen ist, ist der Anteil aller Messwerte, die eine bestimmte Bedingung erfüllen. Eine solche Bedingung könnte zum Beispiel ein bestimmter Schwellwert sein, der überschritten sein muss.

3.3 Beschreibung der Parameter mit den zugehörigen Kenngrößen

Zur Ermittlung der Zielgrößen werden verschiedene Fahrparameter extrahiert. Da es zu jedem Parameter sehr viele Messwerte gibt, werden diese durch charakteristische Kenngrößen beschrieben. Sie ermöglichen einen einfacheren Vergleich von verschiedenen Fahrten.

Es sollen nun folgende Parameter näher betrachtet werden, da diese das Fahrverhalten gut beschreiben.

- Abstand zum vorausfahrenden Fahrzeug (absoluter Abstand, Time to Collision (TTC), Zeitlücke)
- Beschleunigung (Längsbeschleunigung, Querschleunigung)
- Fahrpedalwert
- Geschwindigkeit (allgemein, bevorzugte Richtgeschwindigkeit, Differenz zur erlaubten Geschwindigkeit)

3.3.1 Abstand zum voraus fahrenden Fahrzeug

Ein wichtiger Parameter zur Beurteilung des Fahrverhaltens ist der Abstand, den der Fahrer zu dem voraus fahrenden Fahrzeug hält. Dabei sollte jedoch auch der relative Abstand in Bezug auf die eigene bzw. die relative Geschwindigkeit zum voraus fahrenden Fahrzeug betrachtet werden. Dabei handelt es sich zum Einen um die Zeitlücke, die der Quotient aus Entfernung und eigener Geschwindigkeit ist und zum Anderen die Zeit bis zur Kollision, welche den Geschwindigkeitsunterschied zum Vordermann einbezieht.

absoluter Abstand

Der absolute Abstand ist während der Fahrt wenig aussagekräftig, da Fahrer bei höheren Geschwindigkeiten meistens größere Abstände einhalten. Anders ist dies jedoch im Stand, da dort alle Geschwindigkeiten 0 km/h betragen.

In dieser Arbeit wird der absolute Abstand bei der Analyse von Haltevorgängen betrachtet. Dabei soll für jeden gebundenen Haltevorgang der minimale Abstand ermittelt werden, den der Fahrer beim Anhalten lässt.

Neben dieser direkten Auswertung ist der absolute Abstand auch die Grundlage zur Bestimmung der TTC und der Zeitlücke. Er wird in der Regel mit Hilfe eines Entfernungsmessers, wie zum Beispiel eines Laserscanners, Radar- oder Ultraschallsensors, ermittelt.

TTC

Die Zeit bis zur Kollision entspricht nach Winner [Win09, S. 527] dem Quotienten aus absolutem Abstand d und der relativen Geschwindigkeit $v_{rel.}$ zum vorausfahrenden Fahrzeug. Diese ist die Differenzgeschwindigkeit zwischen vorausfahrendem und dem eigenen Fahrzeug.

$$TTC = \frac{d}{v_{rel.}}$$

Nach Abendroth und Bruder [AB09, S. 13] kann davon ausgegangen werden, dass die TTC die Aktionen des Fahrers bestimmt.

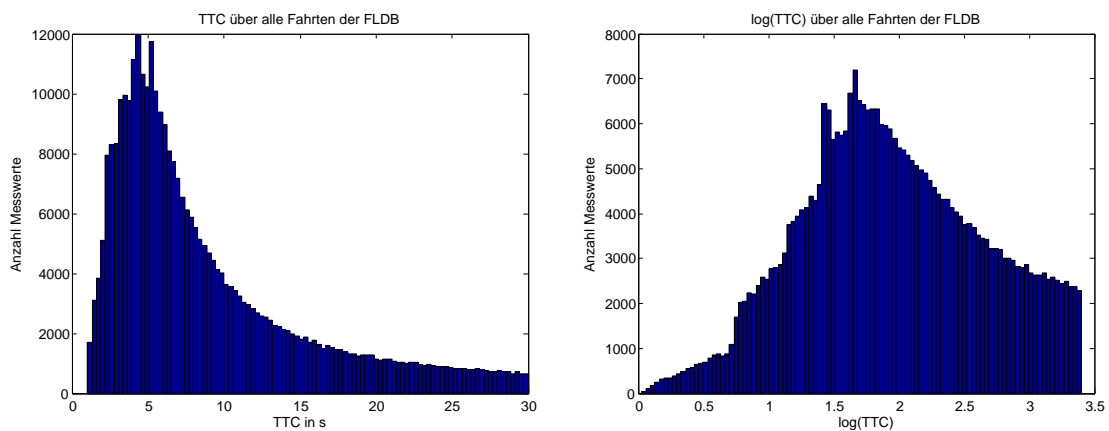


Abbildung 3.2: Verteilung der TTC-Messwerte über alle Fahrten der FLDB, links unverändert, rechts logarithmiert

In Abbildung 3.2 ist die Verteilung der TTC über alle Fahrten der FLDB dargestellt. Auf der rechten Seite sind die Werte vorher logarithmiert worden. Es ist zu erkennen, dass die logarithmierten Werte näherungsweise eine Dreiecksverteilung darstellen. Eine solche fast symmetrische Verteilung lässt sich gut mit Hilfe von arithmetischem Mittel und Varianz beschreiben.

Zeitlücke

Die Zeitlücke entspricht nach Winner et. al [WDS09, S. 479] dem zeitlichen Abstand τ zum vorausfahrenden Fahrzeug.

$$\tau = \frac{d}{v}$$

Wie bei der TTC ist auch bei der Zeitlücke in Abbildung 3.3 zu erkennen, dass die logarithmierten Werte wesentlich symmetrischer sind. Die logarithmierte Zeitlücke entspricht fast einer Normalverteilung, wodurch sich diese besonders gut mit Hilfe der Standardabweichung und dem arithmetischen Mittel beschreiben lässt.

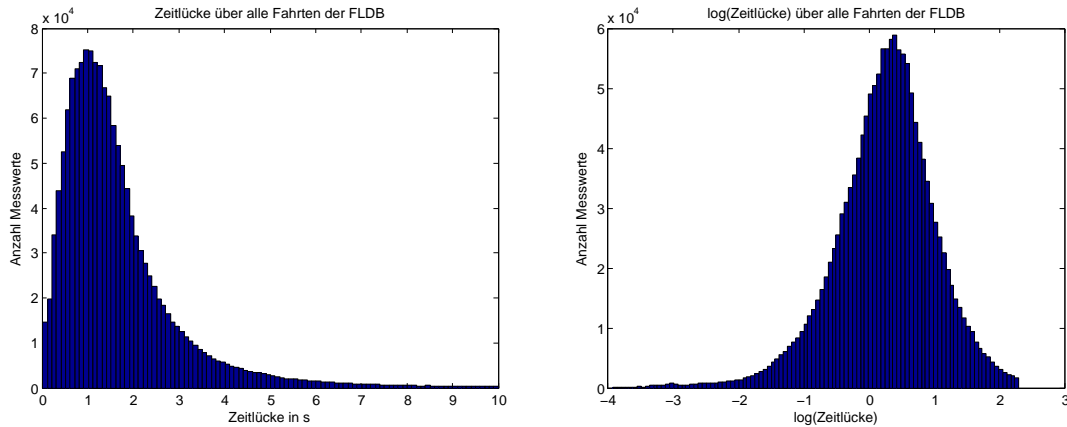


Abbildung 3.3: Verteilung der Zeitlücke-Messwerte über alle Fahrten der FLDB, links unverändert, rechts logarithmiert

3.3.2 Beschleunigung

Die Beschleunigung unterteilt sich in zwei Arten. Diese sind die Längsbeschleunigung, welche beim allgemeinen Beschleunigen bzw. beim Abbremsen auftreten, und die Querbeschleunigung, die direkt proportional zur Zentrifugalkraft innerhalb von Kurven ist.

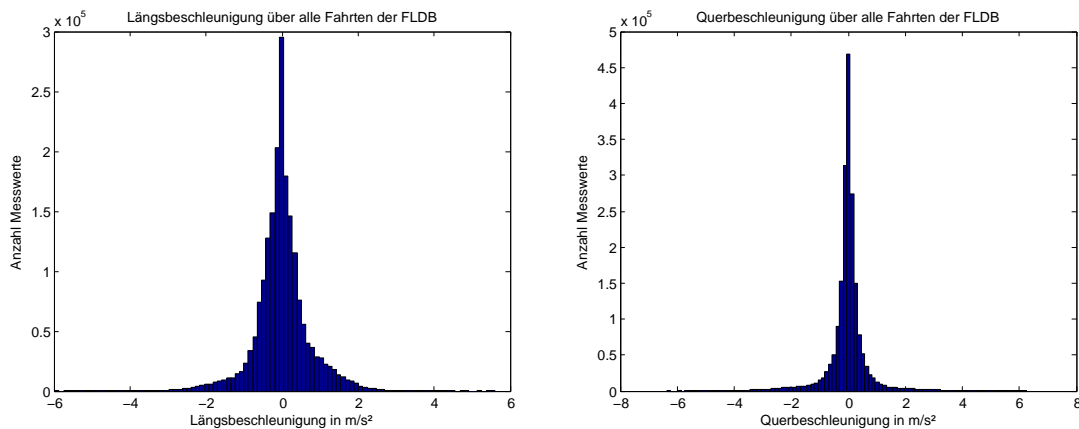


Abbildung 3.4: Verteilung der Längs- (links) und Querbeschleunigung (rechts) über alle Fahrten der FLDB

Längsbeschleunigung

Längsbeschleunigung tritt beim Fahrzeug bei zwei verschiedenen Fahrsituationen auf. Beim Beschleunigen, sprich bei der Erhöhung der Fahrzeuggeschwindigkeit, tritt eine positive Längsbeschleunigung auf. Beim Bremsen hingegen wird eine negative Längsbeschleunigung aufgebaut. Der positive und negative Teil der Längsbeschleunigung werden einzeln betrachtet, da Fahrer unterschiedliches Beschleunigungs- und Bremsverhalten aufweisen können.

In Abbildung 3.4 ist auf der linken Seite die Verteilung der Längsbeschleunigung zu sehen. Betrachtet man nur den positiven, bzw. nur den Betrag der negativen Beschleunigung, so ist diese annähernd exponentiell mit dem Erwartungswert $\frac{1}{\lambda}$ verteilt. Zur Schätzung dieser Verteilung reicht somit ein λ welches sich aus der Varianz leicht bestimmen lässt [TSWH09, S. 187]. Aus diesem Grund wird als Kenngröße für die Längsbeschleunigung die Varianz, jeweils für positive und negative Werte getrennt, genutzt.

$$s^2 = \frac{1}{\lambda}$$

Querbeschleunigung

Unter der Querbeschleunigung wird die Beschleunigung verstanden, die senkrecht zur Fahrtrichtung wirkt. Je enger die Kurve bzw. je höher die Geschwindigkeit, desto größer ist die daraus resultierende Querbeschleunigung. Auch bei der Querbeschleunigung sollte positiver und negativer Anteil separat betrachtet werden, da für den Fahrer Links- und Rechtskurven unterschiedlich stark wirken. Eine Verteilung der Querbeschleunigungswerte ist in Abbildung 3.4 dargestellt.

Die Bestimmung der Kenngrößen für die Querbeschleunigung verläuft analog der Längsbeschleunigung.

Kombination aus Längs- und Querbeschleunigung

Während des Fahrens wirken oft Längs- und Querbeschleunigung gleichzeitig auf das Fahrzeug. Abbildung 3.5 skizziert das Beschleunigungsverhalten verschiedener Fahrertypen. In x-Richtung ist dabei die Quer- und in y-Richtung die Längsbeschleunigung abgetragen. Die Kreise beschreiben die maximale Beschleunigung, mit der das Fahrzeug Haftung auf der Fahrbahn bei unterschiedlichen Witterungsbedingungen hat (Kammischer Kreis). Die eingezeichneten Hüllkurven entsprechen 85%-Perzentile der Gesamtbeschleunigung [Don09]. Diese Kurven lassen sich entsprechend auch für andere Perzentile wie den Median bestimmen.

Zur Charakterisierung des Fahrers ist es nun möglich kontinuierlich zu bestimmen, zu welchem Anteil sich der Fahrer unterhalb des durchschnittlichen Medians der Referenzdaten befindet. So kann die Kombination aus Quer- und Längsbeschleunigung in dieser Kenngröße beschrieben werden.

3.3.3 Fahrpedalwert

Das Fahrpedal ist eines der direkten Bedienelemente, mit dem der Fahrer sein Fahrzeug steuern kann. Dank des elektrischen Fahrpedals lässt sich die Neigung des

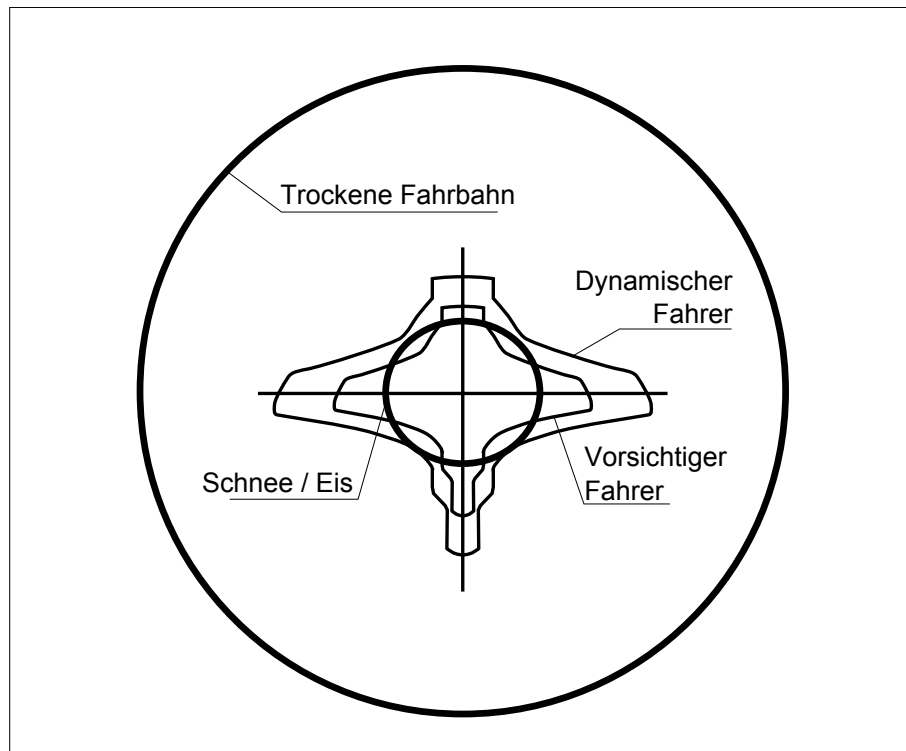


Abbildung 3.5: Zusammenhang zwischen Längs- und Querbeschleunigung [Don09, S. 21]

Fahrpedals einfach auslesen. Die Werte entsprechen dabei dem prozentualen Anteil des durchgetretenen Fahrpedalweges.

In Abbildung 3.6 ist eine Verteilung des Fahrpedalwertes über alle Fahrten der FLDB zu sehen. Zu erkennen sind zwei besonders auffällige Balken am Anfang und am Ende des Wertebereiches. Während einer sehr großen Zeit des Fahrens betätigen die Fahrer das Fahrpedal gar nicht und rollen somit oder bremsen ihr Fahrzeug. Bei einigen Fahrern tauchen viele Messwerte am Ende des Wertebereiches auf, der durch ein vollständiges Durchtreten des Fahrpedals erzeugt wird.

Die Messwerte bei denen das Fahrpedal nicht betätigt wurde, werden für die weitere Betrachtung nicht berücksichtigt, da sie keine Aussage über den Fahrer zulassen. Das vollständige Durchtreten hingegen ist ein Indikator für das Fahrverhalten und wird deswegen speziell betrachtet.

Im rechten Teil der Abbildung 3.6 ist der Fahrpedalwert von diesen zwei besonderen Häufungen bereinigt. Es ist zu erkennen, dass die übrigen Werte recht symmetrisch und scheinbar normal verteilt sind. Aus diesem Grund wird für die Beschreibung des Fahrpedalverhaltens wieder die Standardabweichung sowie das arithmetische Mittel genutzt. Zusätzlich wird noch der Anteil der Messwerte bestimmt, die über dem Wert 95 liegen.

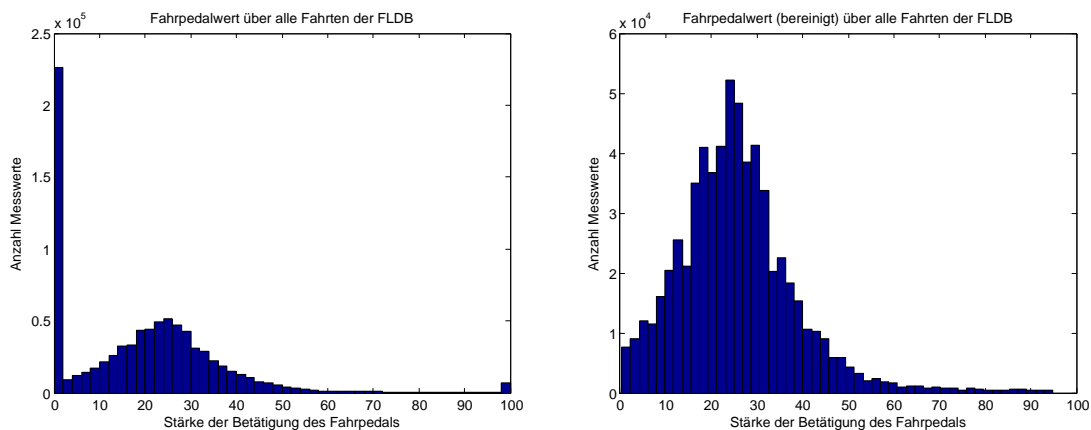


Abbildung 3.6: Fahrpedalwert-Verteilung über alle Fahrten der FLDB, links unverändert, rechts bereinigt

3.3.4 Geschwindigkeit

Das Geschwindigkeitsverhalten ist eine direkte Ausdrucksweise des Fahrverhaltens. Wie jedoch beim Abstand ist die aktuelle Geschwindigkeit direkt nur wenig aussagekräftig. So gibt es innerorts zum Beispiel Straßen, auf denen nur 30 km/h erlaubt sind oder Autobahnen, wo es keine Begrenzungen gibt. Zur Charakterisierung der Fahrer wird in dieser Arbeit insbesondere die Differenz zur erlaubten Geschwindigkeit, sowie die gefahrene Geschwindigkeit, wenn es keine Geschwindigkeitslimits gibt, betrachtet.

Bei der gebundenen Fahrt kann der Fahrer jedoch nicht seine bevorzugte Geschwindigkeit fahren, sondern wird durch das voraus fahrende Fahrzeug begrenzt. Aus diesem Grund wird bei der Analyse der Geschwindigkeit vor allem auf die freie Fahrt eingegangen.

bevorzugte Richtgeschwindigkeit

Als bevorzugte Richtgeschwindigkeit wird in dieser Arbeit die Geschwindigkeit bezeichnet, mit der ein Fahrer fährt, wenn ihn keine Fahrbedingungen beschränken. Dazu zählen insbesondere voraus fahrender Verkehr und allgemeine Geschwindigkeitsbegrenzungen auf der Straße. Somit werden nur Fahrabschnitte auf Autobahnen oder Kraftfahrstraßen betrachtet, bei denen es keine Geschwindigkeitsbeschränkungen gibt.

In Abbildung 3.7 ist die Verteilung der Geschwindigkeit unter den o. g. Bedingungen der Fahrt 500 der FLDB zu sehen.

Die Werte sind in der Regel nicht normal verteilt. Dennoch ist das Arithmetische Mittel und die Varianz charakteristisch für gewisses Fahrverhalten.

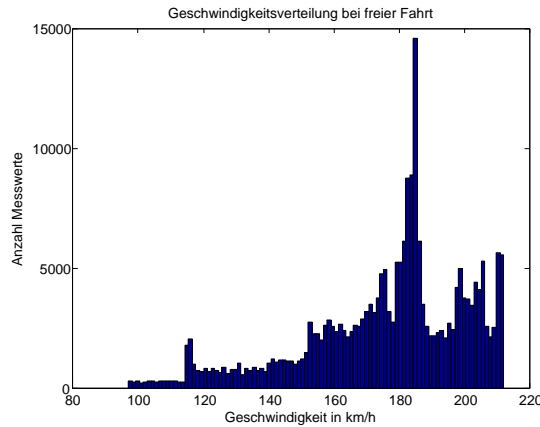


Abbildung 3.7: Verteilung der Geschwindigkeit bei freier Fahrt und ohne Geschwindigkeitsbegrenzung der Fahrt 500 der FLDB

Differenz zur erlaubten Geschwindigkeit

In Abschnitten, in denen es nicht erlaubt ist beliebige Geschwindigkeiten zu fahren, wird einen weiteren Rückschluss auf das Fahrverhalten ermöglicht. Auch hier werden nur Abschnitte betrachtet, in denen der Fahrer nicht hinter einem anderen Fahrzeug her fährt, da dieses eine physische Geschwindigkeitsbegrenzung darstellt.

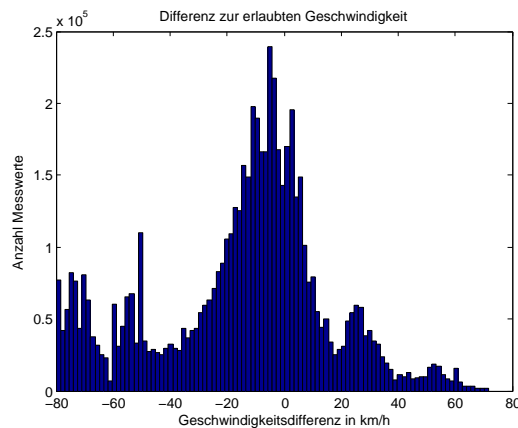


Abbildung 3.8: Verteilung der Geschwindigkeitsdifferenz zur erlaubten Geschwindigkeit bei freier Fahrt der Fahrt 500 der FLDB

In Abbildung 3.8 ist die Verteilung der Differenz zur erlaubten Geschwindigkeit für die Fahrt 500 der FLDB abgebildet. Zu erkennen ist, dass der Fahrer die erlaubte Geschwindigkeit zum Teil sehr stark übertreten hat.

Für diesen Parameter lassen sich mehrere interessante Kenngrößen bestimmen. Prinzipiell ist auch hier wieder das arithmetische Mittel, sowie die Varianz interessant. Zusätzlich lässt sich der Anteil der Geschwindigkeitsüberschreitung und die maximale Geschwindigkeitsüberschreitung bestimmen. Wie jedoch in Abschnitt 3.2.1 beschrieben, sind Extremwerte eher schlecht geeignet.

Kapitel 4

Schätzung des verdeckten Fahrerparameters

Verdeckte Fahrerparameter lassen sich nicht, oder nur sehr schwer, direkt im Fahrzeug bestimmen. Aus diesem Grund sollen die Werte mit Hilfe von Referenzdaten geschätzt werden. Dabei wird folgende Grundannahme getroffen:

Fahrer mit ähnlichen verdeckten Parametern verhalten sich während der Fahrt ähnlich. Somit werden auch die ermittelten Kenngrößen ähnliche Ausprägungen haben.

Im Folgenden wird der spezielle Fahrerparameter, der ermittelt werden soll, Zielwert genannt. Abbildung 4.1 zeigt eine grobe Übersicht über den Ablauf der Schätzung des Zielwertes. Zunächst werden während der Fahrt Messwerte aufgezeichnet und in einzelne Kenngrößen überführt. Diese Werte werden anschließend normalisiert, so dass sie sich untereinander besser vergleichen lassen. Mit Hilfe der normalisierten Kenngrößen werden nun ähnliche Fahrten ermittelt mit denen dann der Zielwert geschätzt wird.

Dieser Ansatz entspricht prinzipiell einem Fallbasierten Schließen (case based reasoning), wie durch Kolodner [Kol92] beschrieben. Jedoch ist die Lernphase des Systems auf das Auswerten der Referenzdatenbank beschränkt, da im produktiven Einsatz im Fahrzeug die Zielgrößen nicht mehr direkt bestimmt und somit der ermittelte Wert nicht überprüft werden kann.

Grundsätzlich sollte die Schätzung von gewissen Zielgrößen auch mit anderen Lernverfahren möglich sein. Ein Beispiel dafür sind Neuronale Netze, wie durch Borgelt et. al [BKKN03] beschrieben. Jedoch bestehen Neuronale Netze hauptsächlich aus Gewichtungen und Schwellwerten, die sehr schwer interpretiert werden können. Ein System, welches auf der Ähnlichkeit zu anderen Fahrern besteht, ist im Vergleich dazu wesentlich einfacher zu verstehen.

Eine andere Möglichkeit wäre ein Fuzzy Controller, welcher auf entsprechenden Regeln basiert. Diese Fuzzy Regeln sind im Allgemeinen recht leicht zu verstehen.

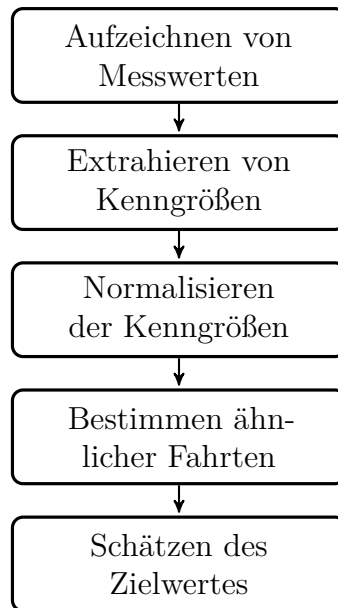


Abbildung 4.1: Ablauf der Fahrerparameterschätzung

Eine dieser Regeln könnte zum Beispiel wie folgt aussehen:

starke Beschleunigungen \wedge hohe Querbeschleunigungen \Rightarrow Fahrer fährt sportlich

Dieser Ansatz wird jedoch in dieser Arbeit nicht verfolgt, da sich solche Regeln vielleicht nicht immer finden lassen und ein Ansatz auf Ähnlichkeit vielversprechender scheint.

4.1 Aufzeichnen von Messwerten und Extraktion der Kenngrößen

Heutzutage wird in modernen Fahrzeugen ein CAN-Bus eingesetzt, mit dessen Hilfe die einzelnen Sensoren und Steuersysteme miteinander kommunizieren. Die über diesen Bus übertragenen Messwerte, wie z.B. Geschwindigkeit, Abstand und andere, lassen sich mit Hilfe eines prototypischen Steuergerätes auslesen und auswerten. Über ein solches System können auch die in Abschnitt 3.3 beschriebenen Kenngrößen ermittelt werden.

In Abbildung 4.2 sind die nötigen Schritte zur Ermittlung der Kenngrößen dargestellt. Je nach Art der Kenngröße werden dafür, wie in Abschnitt 3.2 beschrieben, unterschiedliche mathematische Funktionen genutzt.

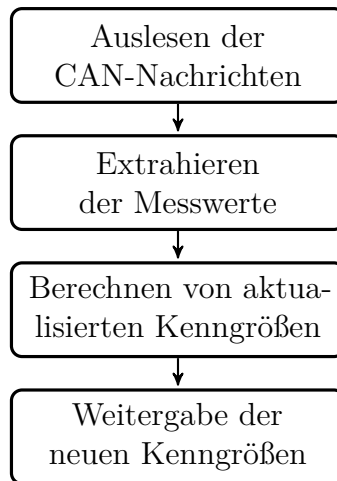


Abbildung 4.2: Ablauf der Kenngrößenermittlung

4.2 Normalisieren der Daten

In dem hier vorgestellten Modell sollen verschiedene Daten miteinander verglichen und zur Bestimmung von Ähnlichkeiten zwischen Fahrten genutzt werden. Jedoch haben die einzelnen Fahrparameter unterschiedliche Wertebereiche und Verteilungen. So gibt es Geschwindigkeitswerte beispielsweise zwischen 0 km/h und 250 km/h , und Beschleunigungswerte im Bereich von -10 m/s^2 bis 10 m/s^2 . Genauso verschieden, wie die möglichen Werte sind auch die Verteilungen, wie Abschnitt 3.3 zu entnehmen ist.

Wünschenswert wären Kenngrößen, die alle denselben Wertebereich besitzen und möglichst gleichverteilt sind. Dabei ist eine Gleichverteilung vorteilhaft, da so Abstände im gesamten Wertebereich gleich aussagekräftig sind.

Aus diesem Grund werden die ermittelten Kenngrößen in Prozentränge umgerechnet. Diese können auf Basis der zur Verfügung stehenden Referenzdaten aus der FLDB ermittelt werden.

Zur Berechnung des Prozentrangs werden folgende Größen benötigt:

PR_x der Prozentrang für den Rohwert x
 $cumf_x$ Anteil der Probanden, die einen kleineren Rohwert als x besitzen
 f_x Anteil der Probanden mit dem Rohwert x
 N Stichprobengröße

Der Prozentrang errechnet sich wie folgt:

$$PR_x = 100\% \cdot \frac{cumf_x + f_x/2}{N}$$

In Abbildung 4.3 ist eine solche Umrechnung zu sehen. In der FLDB sind recht wenige Fahrten aufgezeichnet und somit nicht alle Prozentränge belegt. Deshalb

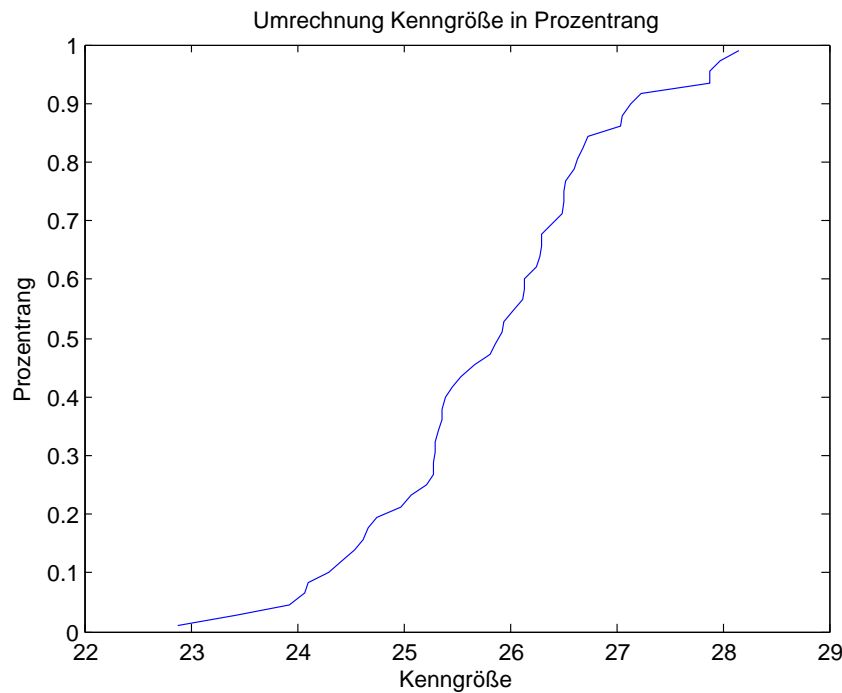


Abbildung 4.3: Umrechnung Kenngröße in Prozentrang

werden Werte, die zwischen den einzelnen Referenzwerten liegen, interpoliert. Dazu wird in dieser Arbeit die lineare Interpolation genutzt. Werte, die unterhalb des Wertebereiches der Referenzdaten liegen, werden dem Prozentrang 0 und Werte die oberhalb liegen dem Prozentrang 1 zugeordnet.

4.3 Bestimmung ähnlicher Fahrten

Nachdem die zu untersuchende Fahrt durch die entsprechenden Parameter charakterisiert ist, werden im nächsten Schritt ähnliche Fahrten ermittelt. In Abbildung 4.4 ist eine Verteilung von Referenzfahrten für zwei Kenngrößen dargestellt. Das blaue Kreuz entspricht der aktuellen Fahrt, welche geschätzt werden soll.

Gewichtung der Koordinaten

In den bisherigen Abschnitten wurden eine Vielzahl von Parametern mit entsprechenden Kenngrößen beschrieben. Jedoch beschreiben nicht alle den Zielwert in gleicher Weise. So könnte beispielsweise das Fahrpedal ein Indiz für die Sportlichkeit des Fahrers und die Differenz zur erlaubten Geschwindigkeit ein Indiz für die persönliche Stellung des Fahrers zu Gesetzen und Vorschriften sein.

Aus diesem Grund werden die verschiedenen Kenngrößen mit einer entsprechenden Gewichtung versehen, die von der gewünschten Zielgröße abhängig ist. Ein Verfahren,

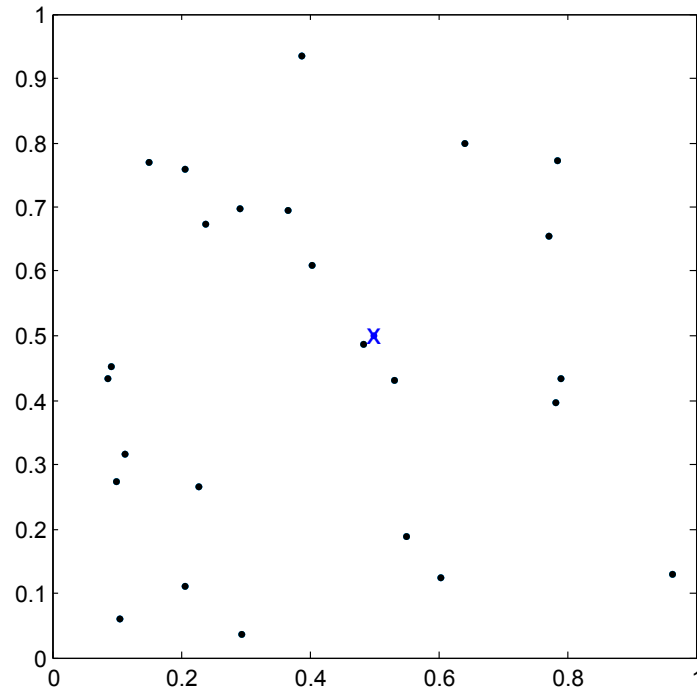


Abbildung 4.4: Darstellung von Referenzfahrten für zwei Kenngrößen

mit dem eine solche Gewichtung bestimmt werden kann, wird im nächsten Kapitel vorgestellt.

Die Gewichtungen wirken sich als Verzerrungen und Projektionen auf den Suchraum aus. Eine Gewichtung von 0 bedeutet, dass diese Kenngröße nicht zur Ähnlichkeitsbestimmung genutzt wird. Sie projiziert den Suchraum von n Dimensionen auf $n - 1$ Dimensionen, wobei die entsprechende Dimension für die Kenngröße entfällt.

In Abbildung 4.5 sind dieselben Daten dargestellt wie in Abbildung 4.4, wobei jedoch die vertikale Dimension auf 0,3 skaliert ist.

Ähnlichkeitsmaß

Damit die Ähnlichkeit zwischen zwei Fahrten x und y ermittelt werden kann, ist es notwendig ein Abstands- bzw. Ähnlichkeitsmaß zu definieren. In dieser Arbeit soll ein Ähnlichkeitsmaß s genutzt werden, welches auf dem euklidische Abstand d basiert.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$s(x, y) = 1 - d(x, y) = 1 - \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

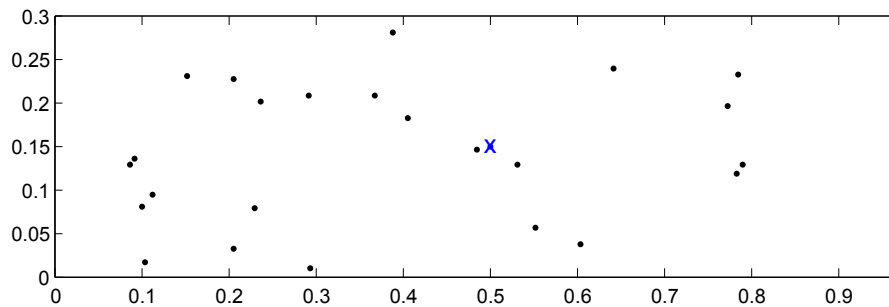


Abbildung 4.5: Darstellung von Referenzfahrten für zwei Kenngrößen, wobei eine skaliert ist

Durch diese Konstruktion erfüllt das hier beschriebene Ähnlichkeitsmaß folgende notwendige Bedingungen:

$$\begin{aligned} s(x, y) &= s(y, x) \\ s(x, y) &\leq s(x, x) \\ s(x, x) &= 1 \end{aligned}$$

Normalisieren der Gewichtungen

Die Größe des Suchraums wird durch zwei Faktoren beeinflusst. Zum Einen durch die Verteilung der jeweiligen Gewichtungen k_i und zum Anderen durch die Anzahl der Kenngrößen n die für die Bestimmung der Ähnlichkeit relevant sind. Deshalb wird die Größe des Suchraums normalisiert indem die Gewichtungen entsprechend angepasst werden. Ziel ist es, dass bei einer Ähnlichkeit von 0 beide Messpunkte an zwei gegenüberliegenden Ecken des Suchraums zu finden sind. Da der euklidische Abstand Grundlage der Ähnlichkeit ist, besitzt die Hauptdiagonale die Länge l . Werden nun die einzelnen Gewichtungen durch diese Länge geteilt, besitzt die Hauptdiagonale genau die Länge 1.

$$l = \sqrt{\sum_{i=1}^n k_i^2}$$

$$\hat{k}_i = \frac{k_i}{l} = \frac{k_i}{\sqrt{\sum_{i=1}^n k_i^2}}$$

Durch diese Transformation ist der maximale Abstand zwischen zwei Punkten 1 und somit ist folgende Bedingung erfüllt.

$$s(x, y) \in [0, 1]$$

4.4 Schätzen des Zielwertes

Das Schätzen des Zielwertes basiert auf dem gewichteten Mittel der Zielwerte z_i ähnlicher Fahrten. Dafür wird die zu untersuchende Fahrt mit allen Fahrten aus der Referenzdatenbank verglichen. Zu jeder Fahrt wird die Ähnlichkeit s_i an Hand des beschriebenen Ähnlichkeitsmaßes bestimmt. Die drei Fahrten, die am ähnlichsten sind dienen als Grundlage der Zielwertschätzung. Als Gewichtung für die Berechnung des Mittels wird die Ähnlichkeit benutzt. Der geschätzte Zielwert \hat{z} berechnet sich wie folgt:

$$\hat{z} = \frac{\sum_{i=1}^3 z_i \cdot s_i}{\sum_{i=1}^3 s_i}$$

4.4.1 Beispiel für eine Zielwertschätzung

Im Folgenden wird das beschriebene Schätzverfahren an einem kleinen Beispiel gezeigt.

In Tabelle 4.1 sind 10 Zufallswerte als sortierte Beispielzielwerte z_i angegeben. Diese stellen nun unsere Referenzdaten dar. Ferner wird angenommen, dass die Auswertung der Kenngrößen folgende Ähnlichkeit zwischen den einzelnen Fahrten bestimmt hat.

$$s(i, j) = \frac{1}{2^{|i-j|}}$$

Die Schätzung erfolgt nun nach der o. g. Formel. Dabei werden alle Referenzdaten, also auch der eigentlich zu schätzende Wert, mitgenutzt. Bei der Anwendung im normalen Einsatz sind diese Werte natürlich unbekannt. Die \hat{z}_i , die sich für die Schätzung ergeben sind in Tabelle 4.2 dargestellt.

| i | z_i |
|-----|--------|
| 1 | 0,0986 |
| 2 | 0,1420 |
| 3 | 0,1683 |
| 4 | 0,1962 |
| 5 | 0,2176 |
| 6 | 0,2510 |
| 7 | 0,3164 |
| 8 | 0,3175 |
| 9 | 0,7032 |
| 10 | 0,8929 |

Tabelle 4.1: Zielwerte für Beispielschätzung

| i | z_i | \hat{z}_i |
|-----|--------|-------------|
| 1 | 0,0986 | 0,1210 |
| 2 | 0,1420 | 0,1377 |
| 3 | 0,1683 | 0,1687 |
| 4 | 0,1962 | 0,1946 |
| 5 | 0,2176 | 0,2206 |
| 6 | 0,2510 | 0,2590 |
| 7 | 0,3164 | 0,3003 |
| 8 | 0,3175 | 0,4137 |
| 9 | 0,7032 | 0,6542 |
| 10 | 0,8929 | 0,7565 |

Tabelle 4.2: Zielwerte und geschätzte Werte für Beispielschätzung

In Abbildung 4.6 sind die gegebenen Zielwerte z_i in blau und die die geschätzten Zielwerte \hat{z}_i in rot dargestellt. Die Linien zwischen den Punkten dienen nur der besseren Wahrnehmung der Punkte und sollen keine Interpolation verdeutlichen. Zu erkennen ist, dass die Schätzung an Hand der beschriebenen Methode in diesem Fall recht gut funktioniert. Lediglich am Rand des Wertebereiches kommt es zu leichten Verzerrungen, da dort nicht größere und kleinere Referenzwerte vorhanden sind.

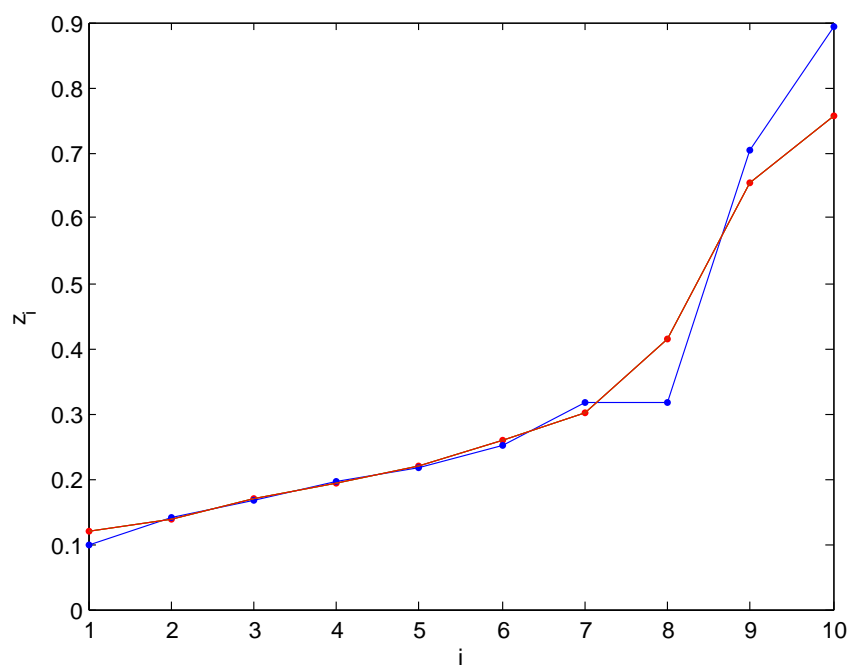


Abbildung 4.6: Darstellung des gegebenen (blau) und des geschätzten (rot) Zielwertes

Kapitel 5

Bestimmung relevanter Kenngrößen

In dem vorherigen Kapitel wurde ein Algorithmus vorgestellt, mit dessen Hilfe verdeckte Zielgrößen geschätzt werden können. Ein wichtiger Parameter für den Algorithmus ist die Auswahl und Gewichtung der einzelnen Kenngrößen, welche die Grundlage der Ähnlichkeitsbestimmung sind. In diesem Kapitel soll nun ein Verfahren vorgestellt werden, mit dessen Hilfe sich die relevanten Kenngrößen einschließlich ihrer Gewichtung bestimmen lassen.

Für die Bestimmung der Auswahl und der Gewichtung wird in dieser Arbeit ein evolutionären Algorithmus genutzt, welcher an Hand von Referenzdaten relevante Kenngrößen lernt. Auf diesen Algorithmus wird nun im Folgenden näher eingegangen.

5.1 Übersicht Evolutionärer Algorithmus

Ein evolutionärer Algorithmus basiert darauf, dass es eine Population von Individuen gibt, die sich im Laufe von Generationen entwickeln. Zur Initialisierung werden 50 zufällige Individuen erschaffen.

Aus den Individuen einer Generation entstehen 50 Individuen für die folgende Generation. Dabei werden zehn Individuen, darunter die drei besten der Generation, direkt ohne Veränderung übernommen. Die restlichen Individuen entstehen jeweils zur Hälfte durch Mutation einzelner oder dem Crossover von jeweils zwei Individuen. Die Auswahl der Eltern für die Mutation bzw. das Crossover geschieht zufällig auf Basis der Fitness der Individuen.

Dadurch, dass bessere Individuen höhere Chancen haben, sich zu vermehren und zu verbessern, werden im Laufe der Generationen immer bessere Individuen, also Gewichtungsverteilungen für die Kenngrößen, entstehen bis diese irgendwann gegen eine Lösung konvergieren. Sollte innerhalb von 50 Generationen kein neuer bester Kandidat gefunden werden, so wird die Suche abgebrochen und der beste bisher

gefundene Kandidat als Lösung ausgegeben. Da sich die meisten Verbesserungen innerhalb der ersten 100 Generationen ergeben, wird in der Auswertung immer nur auf diese eingegangen.

Hier nochmal die relevanten Eckdaten in einer Kurzzusammenfassung:

| | |
|----------------------------|---|
| Anfangspopulation: | 50 zufällige Individuen |
| Zwischenpopulation: | 50 Individuen |
| keine Veränderung: | 10 zufällige Individuen (darunter besten 3) |
| Mutation: | 20 Individuen |
| Crossover: | 20 Individuen aus je zwei Eltern |
| Endbedingung: | 50 Iteration ohne Verbesserung |

Die nachfolgenden Betrachtungen beziehen sich auf die Optimierung der Gewichtung von 41 Kenngrößen. Diese ergeben sich aus der Kombination der Kenngrößen mit den einzelnen Fahrparametern und Fahrsituationen des letzten Kapitels.

5.2 Kodierung der Individuen

Jedes Individuum stellt eine mögliche Konfiguration für die Gewichtungsverteilung der einzelnen Kenngrößen dar. Aus dieser Tatsache lässt sich direkt eine Kodierung der Individuen ableiten. Jedes Gen wird als nicht negative reelle Zahl dargestellt und entspricht der Gewichtung einer Kenngröße. Ein Beispiel für eine solche Darstellung ist in Abbildung 5.1 für einige ausgewählte Kenngrößen zu sehen.

| | | |
|--|-------|------|
| $\bar{x}_{\text{Geschwindigkeit}}$ | 0,3 | 0,2 |
| $s_{\text{Fahrpedal}}^2$ | 0,9 | 0,6 |
| $\bar{x}_{\text{Zeitlücke (freie Fahrt)}}$ | 0,075 | 0,05 |
| $\min_{\text{Abstand beim Anhalten}}$ | 0,225 | 0,15 |

Abbildung 5.1: Beispielkonfiguration eines Individuums, links Kenngrößen, Mitte Gewichtung, rechts normalisiert

Um jeder Gewichtungskonfiguration genau einem Individuum und umgekehrt zuzuordnen, müssen die einzelnen Individuen normalisiert werden. Zu diesem Zweck werden alle Gene g_i eines Individuums so skaliert, dass sie in der Summe eins ergeben.

$$\hat{g}_i = \frac{g_i}{\sum_{j=1}^N g_j}$$

5.3 Startpopulation

Wie bereits erwähnt werden für die Startpopulation 50 Individuen zufällig erzeugt. Dabei wird jedes Gen eines Individuums mit Hilfe einer standardnormalverteilten

Zufallsgröße Z wie folgt berechnet:

$$g_i = \frac{1 + Z}{2}$$

Diese Werte entsprechen noch nicht den Bedingungen wie sie in Abschnitt 5.2 beschrieben sind. So kann es durch die normalverteilten Zufallswerte zu negativen Werten kommen. Diese werden auf 0 gesetzt. Das bedeutet, dass die entsprechende Kenngröße in diesem Individuum nicht betrachtet wird. Anschließend werden die Individuen normalisiert. In Abbildung 5.2 ist dieser Ablauf noch mal dargestellt.

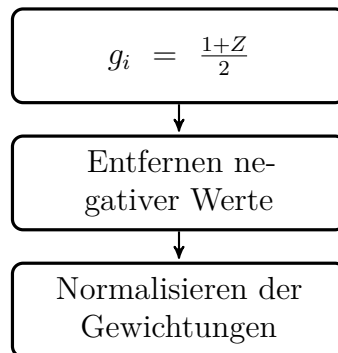


Abbildung 5.2: Erzeugung der Startpopulation

5.4 Fitnessfunktion

Einzelne Individuen sollen mit Hilfe einer Fitnessfunktion bewertet werden. Ziel dieser Bewertung ist es den einzelnen Individuen eine Güte zuzuschreiben. Mit Hilfe dieser Zuordnung lässt sich entscheiden, welches das beste Individuum der Generation ist. Somit stellt die Fitness der Individuen eine Basis für die Auswahl in die nächste Generation dar.

Zur Bestimmung der Fitness wird jedes Individuum einem Test unterzogen. In diesem wird jede Referenzfahrt mit Hilfe der anderen Referenzfahrten geschätzt. Für diese Schätzung kommt der Algorithmus aus dem vorherigen Kapitel zum Einsatz. Als Gewichtungskonfiguration dient das zu untersuchende Individuum. Über die Referenzdaten lassen sich so nun für jede Fahrt ein gemessener bzw. vorgegebener Zielwert und ein geschätzter Zielwert bestimmen. Zwischen diesen beiden Datenreihen wird nun die Korrelation bestimmt und der Korrelationskoeffizient k dient als Grundlage der Bestimmung der Fitness f .

Je näher dieser Wert an 1 ist, desto stärker liegen die geschätzten und die gemessenen Werte auf einer Gerade. Ein Wert von 0 bedeutet, dass kein linearer Zusammenhang zwischen den Werten vorhanden ist. Eine negative Korrelation sagt aus, dass je größer der eine, desto kleiner der andere Wert ist. Ziel ist es eine Korrelation von möglichst 1 zu erreichen um eine gute Schätzung zu gewährleisten. Da negative Werte kontraproduktiv sind und die Auswahlwahrscheinlichkeit proportional zur Fitness ist, wird für eine negative Korrelation die Fitness auf 0 gesetzt.

Durch die vielen möglichen Kenngrößen kann es vorkommen, dass sehr viele dieser ausgewählt werden und somit ein sehr komplexes Modell entsteht. Dabei besteht jedoch die Gefahr, dass das System die Referenzdaten zu sehr lernt und somit quasi „überlernt“ ist. In diesem Fall kann das System dann die Referenzdaten perfekt bestimmen, ist jedoch für die Schätzung anderer Daten weniger geeignet. Um dies zu verhindern wird in die Berechnung der Fitness ein Strafterm eingebaut, welcher die Fitness um ein Prozent je genutzten Parameter senkt. Dadurch werden einfachere Modelle bevorzugt.

In dieser Arbeit wird für die Bestimmung der Fitness somit folgende Funktion genutzt:

$$f = \underbrace{k}_{\text{Korrelation}} \cdot \underbrace{\left(1 - \frac{1}{100}n\right)}_{\text{Strafterm}}$$

5.5 Auswahlverfahren

Ein wichtiger Schritt bei einem evolutionären Algorithmus ist die Auswahl der Kandidaten für die nächste Generation. In dieser Arbeit werden zwei Verfahren kombiniert genutzt. So kommt eine einfache Glücksradauswahl in Kombination mit Elitismus zum Einsatz.

5.5.1 Elitismus

Wurden durch den evolutionären Algorithmus bereits in sehr frühen Phasen gute Individuen erzeugt, so sollen diese natürlich nicht verworfen werden. Aus diesem Grund werden jeweils die drei besten Individuen einer Generation unverändert in die nächste Generation übernommen. Durch dieses Verfahren ist sichergestellt, dass die im gesamten Verlauf beste Lösung bis zum Ende überlebt.

5.5.2 Glücksradauswahl

Die Glücksradauswahl stellt ein sehr einfaches Verfahren zur Auswahl der verwendeten Individuen dar. Dabei werden Individuen mit einer Wahrscheinlichkeit proportional zu ihrer Fitness ausgewählt. Die Wahrscheinlichkeit für ein Individuum i aus der Population der Größe N ausgewählt zu werden beträgt:

$$p_i = \frac{f_i}{\sum_{j=1}^N f_j}$$

Besonders in der Startphase der Evolution kann es häufig zu Individuen kommen, die eine Fitness von 0 besitzen. Solche Individuen würden nach diesem Verfahren

eine Wahrscheinlichkeit von 0 besitzen ausgewählt zu werden und somit nicht am Auswahlprozess teilnehmen. Im schlimmsten Fall könnte es sein, dass die gesamte Population eine Fitness von 0 besitzt und somit kein Individuum für die nächste Generation ausgewählt werden könnte. Um diesen Missstand zu beheben wird jedem Individuum für die Auswahl eine Mindestfitness von 0,01 zugeordnet. Somit hat jedes Individuum eine gewisse, wenn auch sehr kleine, Chance in die nächste Generation übernommen zu werden.

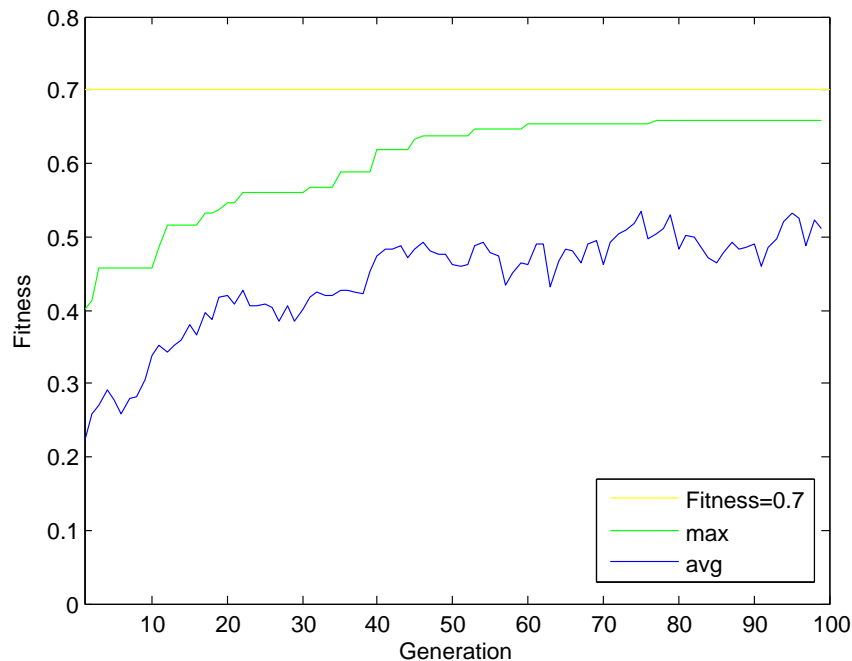


Abbildung 5.3: Entwicklung der Fitness bei Standardglücksradauswahl

In Abbildung 5.3 ist die Entwicklung der Fitness der Population während der ersten 100 Generationen dargestellt. Dabei greifen alle Bestandteile des evolutionären Algorithmus. Das heißt, dass Elitismus, Glücksradauswahl, sowie Mutation und Crossover verwendet werden. Dies trifft auch auf alle andere Variationen des Algorithmus zu, die im Folgenden gezeigt werden. Wie leicht zu erkennen, steigt die maximale Fitness am Anfang recht kontinuierlich an. Circa ab der 50. Generation gibt es aber kaum noch Steigerungen. Ähnlich sieht es bei der durchschnittlichen Fitness der Individuen aus. Anfangs ist eine recht kontinuierliche Steigerung zu erkennen. Die durchschnittliche Fitness pegelt sich nach ca. 40 Generationen zwischen 0,4 und 0,5 ein.

Gillies führt 1985 in seiner Arbeit [Gil85], das von ihm so genannte „Power Scaling“ ein. Diese skaliert die Fitness für den Auswahlprozess exponentiell in der Form:

$$f^* = f^k;$$

Diese Skalierung verändert nur die Wahrscheinlichkeit für die Glücksradauswahl, nicht jedoch die betrachtete Fitness zur Einschätzung der Population.

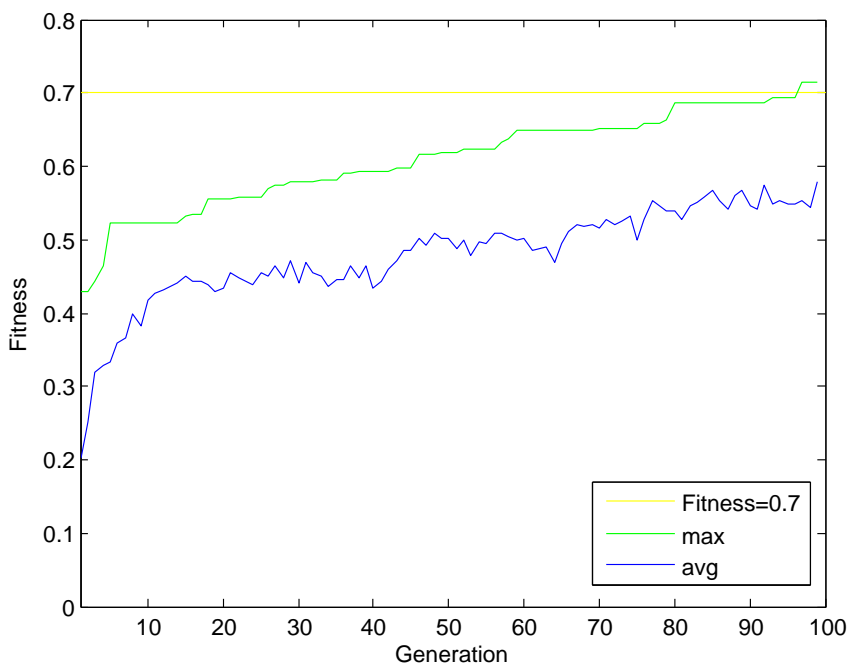


Abbildung 5.4: Entwicklung der Fitness bei Glücksradauswahl exponentiell skaliert mit $k = 1,005$

Nach Gerdes et. al. [GKK04] ergeben k knapp über 1 (bei Gillies für $k = 1,005$) recht gute Werte. In Abbildung 5.4 ist die Fitness der ersten 100 Generationen für die Glücksradauswahl mit $k = 1,005$ exponentiell skalierten Fitness gezeigt. Wie zu erkennen ist, steigt hier die maximale Fitness etwas stärker und kontinuierlicher an. Auch die durchschnittliche Fitness der Population stagniert hier nicht auf einem mittleren Wert, sondern steigt während der gesamten ersten 100 Generationen.

Neben der statischen exponentielle Skalierung werden in [GKK04] auch dynamische Skalierungen vorgestellt. Dabei verändert sich das k über die Zeit, so dass zu Beginn sehr kleine k den Einfluss der Fitness sehr niedrig gestalten. Im Laufe der Generationen wird durch steigendes k der Einfluss der Fitness immer stärker und somit der Selektionsdruck höher.

In dieser Arbeit soll eine dynamische Skalierung zum Einsatz kommen. Da jedoch keine Aussage über die Anzahl der Generationen gemacht wird, es aber Beschränkungen für die erwarteten Zielwerte gibt, soll das k nicht von der Generationsnummer, sondern von der Fitness der Population abhängig gemacht werden. Dazu wird k wie folgt bestimmt.

$$k = c \cdot \max(f)$$

In Abbildung 5.5 sind die Entwicklungen der Population für verschiedene Werte für c dargestellt. Die besten Resultate ergeben sich $c = 3$ und $c = 4$. Bei $c = 5$ ist der Einfluss der Fitness bereits so hoch, dass dort sehr schnell lokale Maxima angestrebt

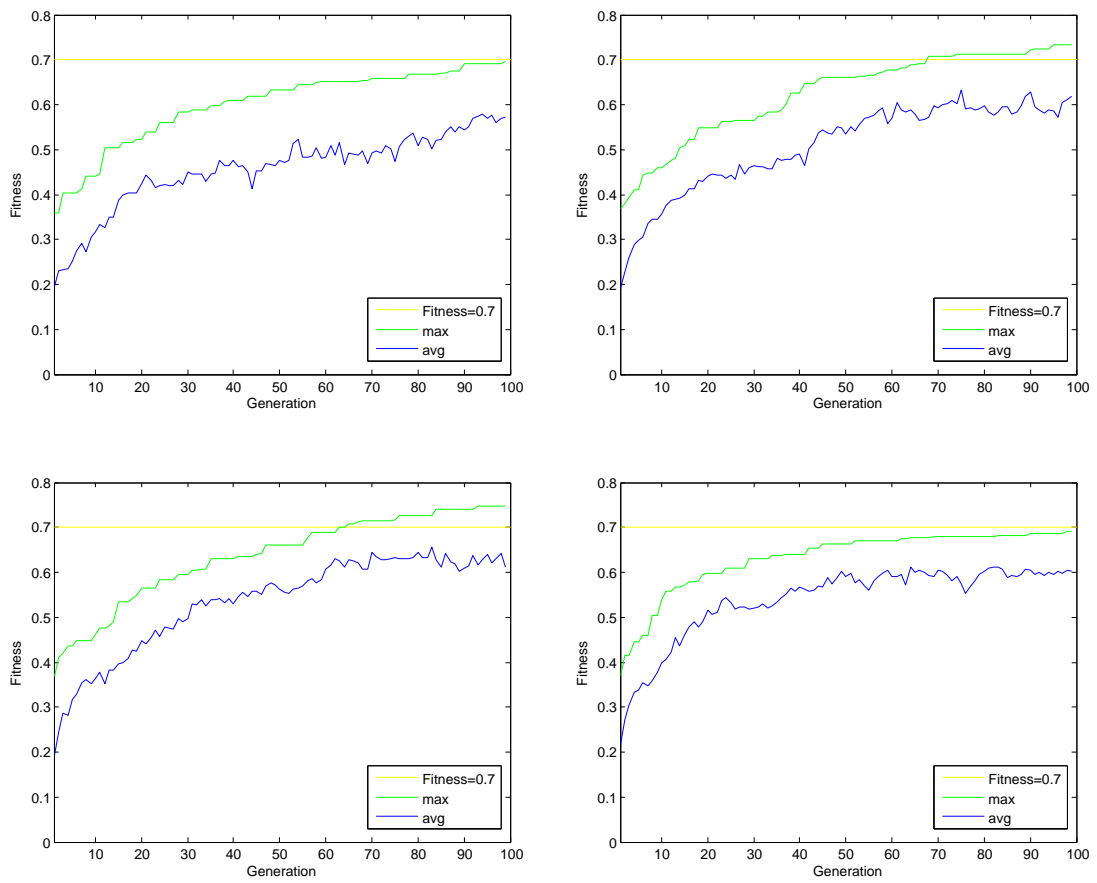


Abbildung 5.5: Entwicklung der Fitness bei Glücksradauswahl mit dynamischer Skalierung für verschiedene c (oben links $c = 2$, oben rechts $c = 3$, unten links $c = 4$, unten rechts $c = 5$)

werden und der gesamte Suchraum nicht gut durchforstet wird. Dadurch lassen sich hier nur kleinere Maximalwerte erreichen.

In dieser Arbeit wird für c der Werte 3 festgesetzt, da dieser auf der einen Seite eine recht gute Entwicklung der Population und auf der anderen Seite aber eine breiter Durchforstung des Suchraumes zulässt als $c = 4$.

5.6 Mutation

Mutation eines Individuums bedeutet, dass einzelne Gene zufällig verändert werden. Dabei hat die Mutation im evolutionären Algorithmus zwei wichtige Aufgaben. Zum Einen sollen Individuen näher an ein Optimum in der nahen Umgebung gebracht werden, so dass ein lokales Optimum erreicht werden kann. Zum Anderen, soll die Mutation jedoch auch so starke Veränderungen ermöglichen, dass Individuen lokale Optima überwinden und somit auch ein anderes, vielleicht besseres lokales Optimum erreichen können.

In dieser Arbeit wird für die Mutation eines Gens eine standardnormalverteilte Zufallsgröße X verwendet. Dieser Wert wird mit einem entsprechenden Faktor k zu dem bisherigen Wert des Gens addiert.

$$g_i^* = g_i + k \cdot X$$

Die Normalverteilung lässt sowohl sehr kleine Veränderungen zu, die zu einem lokalen Optimum führen, als auch sehr große Veränderungen, mit denen andere Gebiete des Suchraums erreicht werden können. Nach der Modifizierung wird das Individuum wieder normalisiert.

In Abbildung 5.6 sind die Entwicklungen für verschiedene Werte für k dargestellt. Für $k = 0,1$ ist zu sehen, dass das beste Individuum kontinuierlich besser wird. Jedoch ist diese Verbesserung im Vergleich zu höheren Werten sehr langsam. Dies spricht dafür, dass insbesondere lokale Optima angestrebt werden. Bei hohen Werten ($k = 0,4$) ist zu erkennen, dass sehr schnell gute Individuen erreicht werden. Durch die starke Mutation wird jedoch sehr schnell von einem lokalen Maximum zum nächsten gesprungen. Dadurch wird dieses jedoch kaum erreicht. Die besten Ergebnisse liefern die Werte $k = 0,2$ und $k = 0,3$. Bei diesen Werten ist ein schnelles und kontinuierliches Wachstum zu erkennen. Da für ein $k = 0,2$ besser lokale Optima erreicht werden können, wird dieser Wert im weiteren Verlauf der Arbeit genutzt.

Ein weiterer wichtiger Parameter für die Mutation ist die Wahrscheinlichkeit p , mit der ein Gen mutiert wird. Nach Gerdes et. al [GKK04] wird häufig eine Mutationswahrscheinlichkeit von $p = \frac{1}{n}$, für n gleich Anzahl der Gene pro Individuum, benutzt. Auch Mühlenbein gibt in seiner Arbeit [Müh92] eine theoretische Begründung für diesen Wert.

In Abbildung 5.7 sind Verläufe für verschiedene Mutationswahrscheinlichkeiten

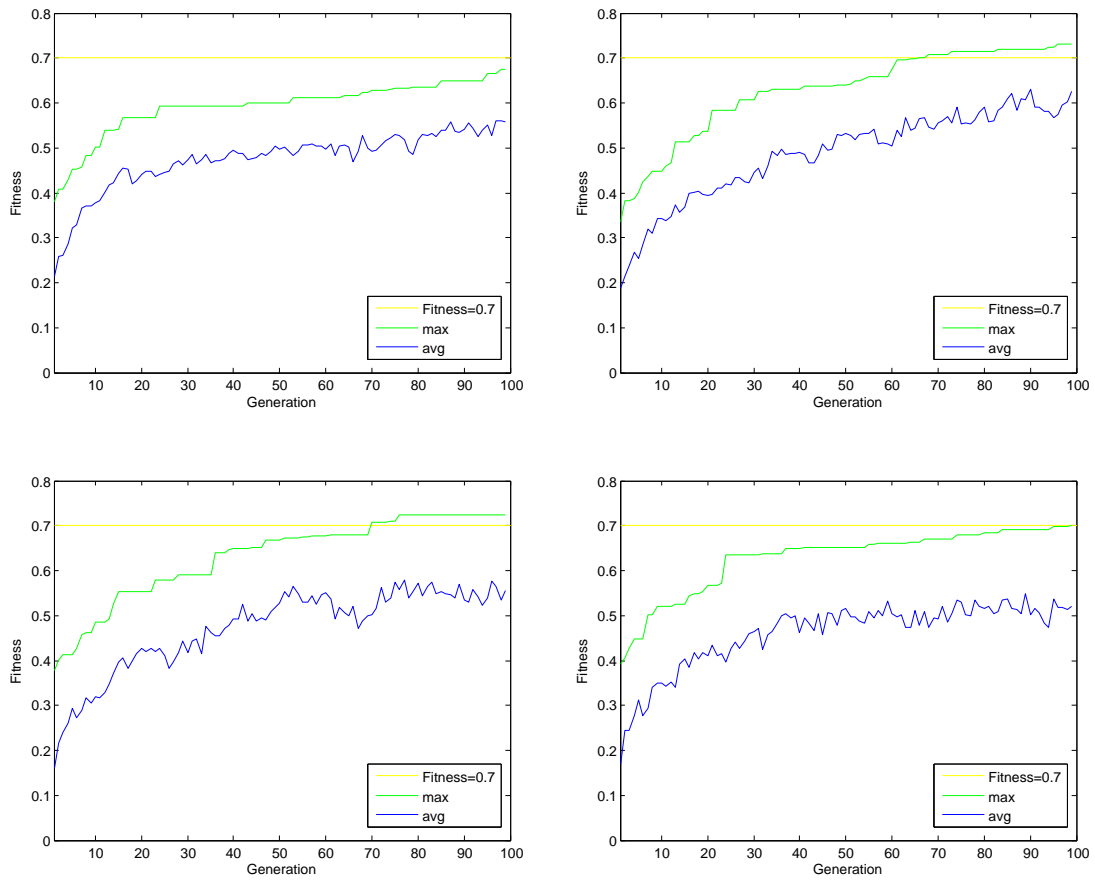


Abbildung 5.6: Entwicklung der Fitness mit unterschiedlichen Mutationsintensitäten (oben links $k = 0,1$, oben rechts $k = 0,2$, unten links $k = 0,3$, unten rechts $k = 0,4$)

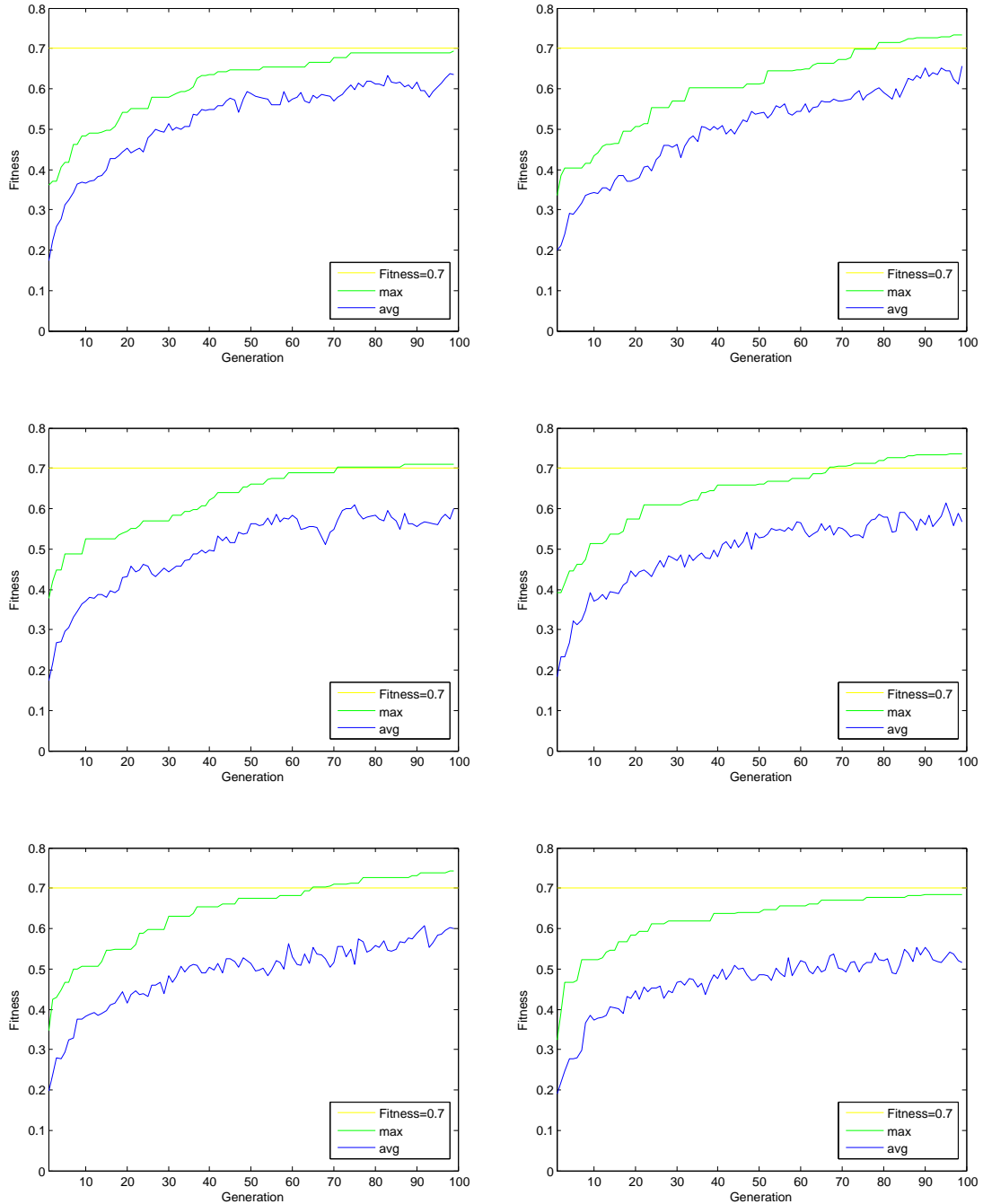


Abbildung 5.7: Entwicklung der Fitness mit unterschiedlichen Mutationswahrscheinlichkeiten (oben links $p = \frac{1}{n}$, oben rechts $p = \frac{2}{n}$, mitte links $p = \frac{3}{n}$, mitte rechts $p = \frac{4}{n}$, unten links $p = \frac{5}{n}$, unten rechts $p = \frac{6}{n}$)

gezeigt. Je näher der Durchschnitt der Population am Maximum liegt, desto weniger stark wird der Suchraum durchforstet, da davon ausgegangen werden kann, dass sich sehr viele Individuen im Bereich lokaler Maxima befinden. Ein frühes Konvergieren der maximalen Fitness ist ein Anzeichen dafür, dass die Individuen in lokalen Maxima festhängen und somit keine besseren lokalen Maxima mehr erreichen.

Bei kleineren Werten $p \leq \frac{2}{n}$ wird der Suchraum recht schlecht durchsucht. Große Werte z. B. $p = \frac{6}{n}$ springen recht stark im Suchraum, so dass gute Werte eher zufällig getroffen als gezielt verbessert werden. Bei noch größeren Mutationswahrscheinlichkeiten ist dieser Effekt noch stärker.

Trotz der guten Erfahrungen, die in anderen Arbeiten mit dem Wert $p = \frac{1}{n}$ gemacht wurden, wird in dieser Arbeit ein Wert von $p = \frac{4}{n}$ festgelegt, da dieser sowohl eine gute Durchsuchung des Raums, als auch ein Erreichen von lokalen Optima, ermöglicht und wesentlich bessere Ergebnisse liefert als $p = \frac{1}{n}$.

5.7 Crossover

Das Crossover ist ein Operator, der zwei oder mehr Individuen aus einer Population kreuzt. Dadurch wird der biologische Schritt der Paarung aus der Evolution nachempfunden. Mit Hilfe eines Crossovers lassen sich Ausschnitte von verschiedenen Individuen mit einander verbinden. So lassen sich durch Kombination von Genen guter Individuen vielleicht noch bessere erzeugen.

In dieser Arbeit wird das Uniform Crossover verwendet, welches von Syswerda [Sys89] vorgestellt wurde.

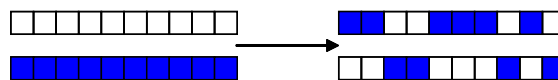


Abbildung 5.8: Uniform Crossover

Grundlage des Verfahrens sind zwei Elternindividuen, aus denen durch Kreuzung zwei Kinderindividuen entstehen. Dabei wird für jedes Gen zufällig bestimmt, an welches Kindindividuum es weitergegeben wird. In Abbildung 5.8 ist dies schematisch dargestellt.

Kapitel 6

Anwendung auf Realdaten und Auswertung der Ergebnisse

In den letzten Kapiteln wurde beschrieben, wie gewünschte Zielgrößen an Hand von Fahrparametern geschätzt werden können. In diesem Kapitel soll das Verfahren nun angewandt werden. Im ersten Schritt erfolgt eine Anwendung auf die Daten der Simulatorstudie, da mit deren Hilfe der Algorithmus entwickelt wurde. Im zweiten Schritt erfolgt eine Anwendung auf die Daten der Realfahrten aus der selbst durchgeführten Studie

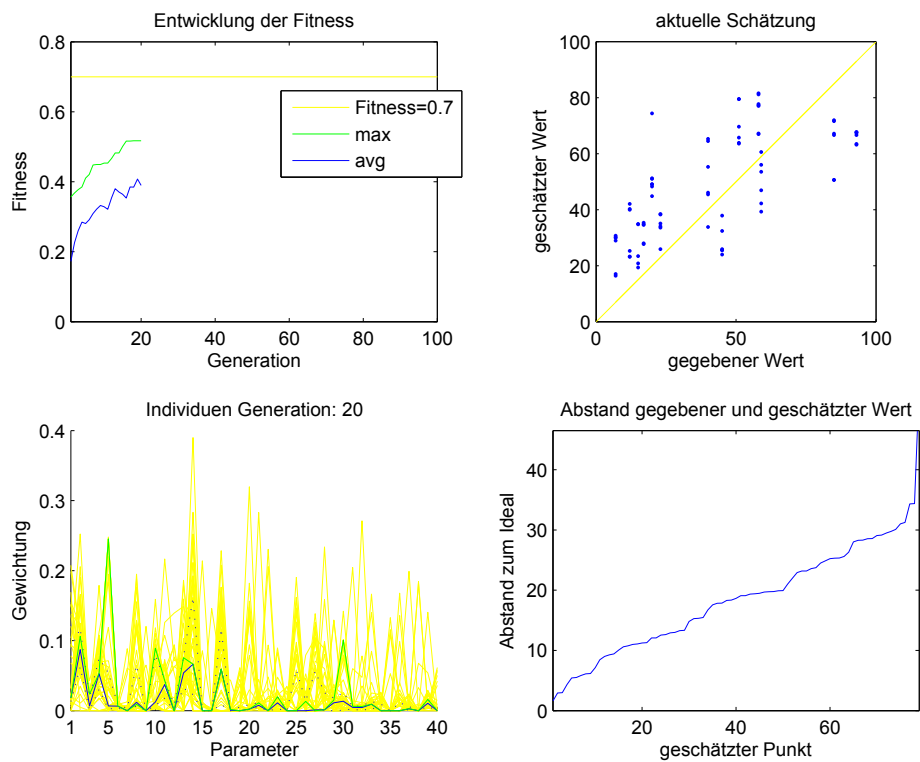


Abbildung 6.1: Beispielgrafik zum Zustand des evolutionären Algorithmus

Im Folgenden werden einige Abbildungen, wie Abbildung 6.1, gezeigt die den Zustand des Evolutionären Algorithmus zu gewissen Zeiten darstellt. An dieser Stelle soll eine

kleine Einführung in die Abbildungen gegeben werden. Die Abbildungen bestehen aus vier Diagrammen.

Oben links wird die Entwicklung der Population im Laufe der Zeit bis zur jeweiligen Generation gezeigt. Diese sind dieselben Diagramme, die auch bereits in dem vorherigen Kapitel Verwendung fanden. Alle anderen Diagramme beziehen sich explizit auf die aktuelle Generation.

Unten links befindet sich eine Darstellung der einzelnen Individuen. Auf der x-Achse sind die einzelnen Kenngrößen und auf der y-Achse die einzelnen Gewichtungen abgetragen. Jedes Individuum entspricht einer der gelben Linien im Diagramm. Durch diese Linien ist zu sehen, wie stark der Suchraum durchsucht wird. Die grüne Linie ist eine Darstellung des besten Individuums der Generation und zeigt somit die aktuelle empfohlene Konfiguration. Die blauen gepunkteten Linien entsprechen dem 25 %- und 75 %-Quantil. Somit befinden sich 50 % aller Individuen zwischen diesen Linien und je 25 % der Individuen darüber bzw. darunter. Die Durchgezogene Linie entspricht dem Median.

Oben rechts ist auf Basis des besten Individuums der Generation eine Schätzung der Zielwerte dargestellt. Für jede Fahrt sind mehrere Messzeitpunkte eingetragen, wodurch mehrere Punkte mit den selben gegebenen Werten entstehen. Die gelbe Linie ist die Identität. Im optimalen Fall liegen alle geschätzten Punkte auf dieser Linie.

Es kann passieren, dass viele Messpunkte auf dem gleichen Punkt im gerade beschriebenen Diagramm liegen. Besonders wenn die Punkte sich der gelben Linie annähern kann so leicht der Eindruck entstehen, dass gar nicht so viele Punkte gut liegen, da sie direkt hintereinander liegen. Im letzten Diagramm unten rechts wurde deswegen der Abstand der Punkte zur gelben Linie geordnet dargestellt. So entsteht eine charakteristische Kurve. Je flacher diese Kurve ist, desto besser ist die Schätzung. Bei der jeweils 100sten Generation ist zusätzlich in rot die Kurve der ersten Generation eingeblendet.

6.1 Simulatorstudie

Während der Entwicklung des Algorithmus wurde mit den Daten der Simulatorstudie gearbeitet, um eine Vergleichbarkeit mit den Ergebnissen von Freiburg [Fre10] zu ermöglichen. Aus diesem Grund soll diese hier zuerst betrachtet werden.

6.1.1 Datenbasis

In der Simulatorstudie wurden 21 Fahrten betrachtet, für die eine Zielgröße festgelegt wurde. In Tabelle 6.1 sind für die einzelnen Fahrer die Zielgrößen aufgezeigt.

Einige Fahrer-IDs sind nicht belegt, da die Fahrten entweder abgebrochen wurden,

| ID | Zielgröße |
|----|-----------|
| 13 | 2 |
| 16 | 51 |
| 18 | 12 |
| 19 | 45 |
| 20 | 7 |
| 21 | 85 |
| 22 | 58 |
| 23 | 59 |
| 24 | 93 |
| 25 | 15 |
| 26 | 17 |
| 27 | 23 |
| 28 | 4 |
| 30 | 9 |
| 31 | 84 |
| 32 | 46 |
| 35 | 51 |
| 36 | 29 |
| 37 | 12 |
| 38 | 12 |
| 40 | 2 |

Tabelle 6.1: Zielwerte der Simulatorstudie

keine Zielgrößen bestimmt werden konnten, oder die Fahrt aus anderen Gründen nicht genutzt werden kann.

6.1.2 Anwendung

In der Simulatorstudie wurde eine Zielgröße betrachtet. Diese hat einen Wertebereich zwischen 0 und 100.

$$\text{Zielgröße} \in [0, 100]$$

In Abbildung 6.2 ist die Startpopulation zu sehen. Es ist zu erkennen, dass die Individuen noch sehr weit im Suchraum verstreut sind. Auch bei der Schätzung lässt sich erkennen, dass die Werte eher geraten als gut geschätzt sind.

In der 20ten Generation sind bereits erste Tendenzen zu erkennen, siehe Abbildung 6.3, die einen erfolgreichen Schätzer erwarten lassen. Die Individuen bewegen sich schon näher im Umfeld einige Parameter, andere Parameter sind schon sehr stark ausgeschlossen und werden nur wenig untersucht. In der Schätzung ist bereits ein grober Trend zu sehen, der zeigt, dass die geschätzten Werte in Richtung der vorgegebenen Werte wandern.

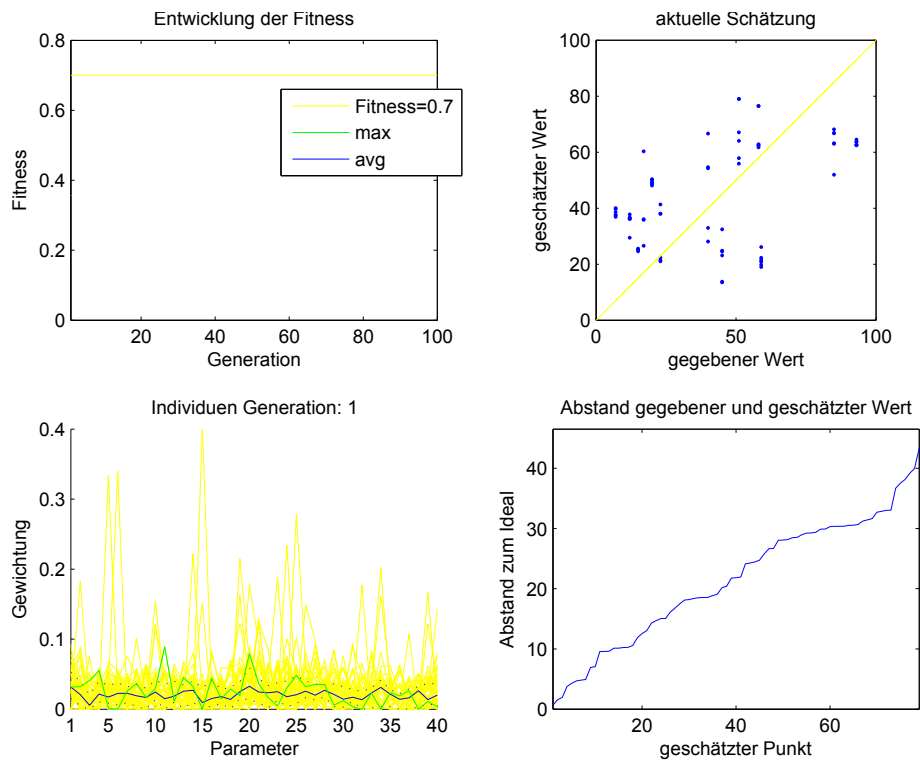


Abbildung 6.2: Generation 1 - Simulatorstudie

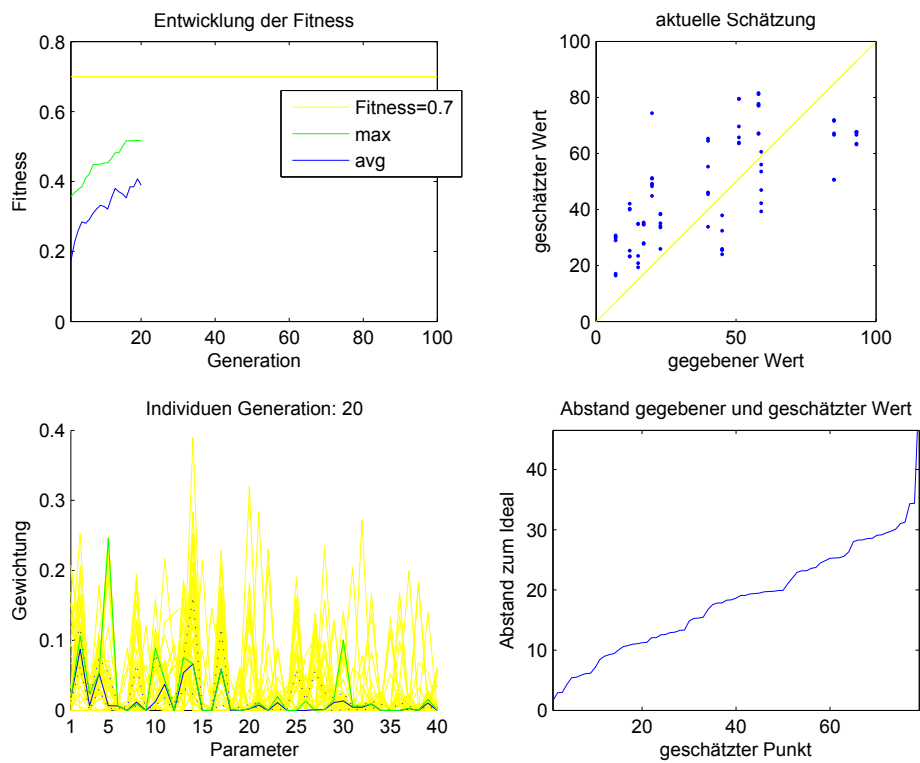


Abbildung 6.3: Generation 20 - Simulatorstudie

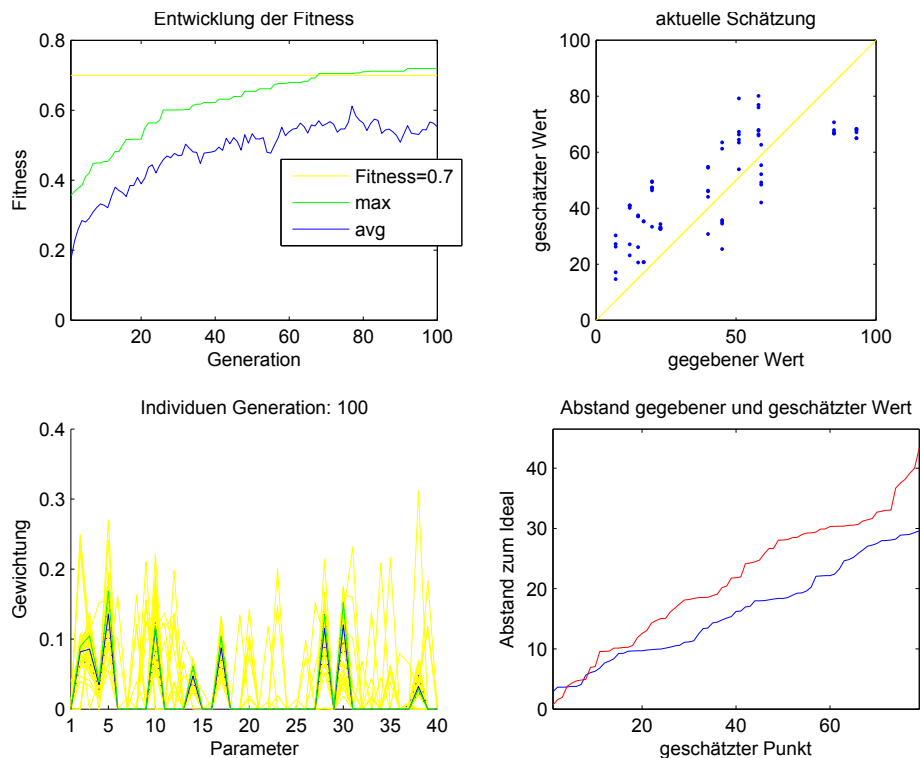


Abbildung 6.4: Generation 100 - Simulatorstudie

Nach der 100sten Generation wurde der Algorithmus hier abgebrochen, da nur noch kleinere Verbesserungen erwartet werden. Es ist zu erkennen, dass der Suchraum nun sehr stark eingegrenzt ist und sich fast alle Individuen in der Umgebung des besten Kandidaten befinden. In der Schätzung ist nun eine starke Korrelation zwischen geschätztem und gegebenem Wert zu erkennen. Auch die Kurve im Diagramm unten rechts ist wesentlich flacher als bei den früheren Generationen. Wie Abbildung 6.4 zu entnehmen wurde eine Fitness von ca. 0,72 mit Hilfe von 10 Parametern erreicht. Die Korrelation kann aus der Fitness bestimmt werden, wenn der abgezogene Strafterm wieder eingerechnet wird. Dies ergibt eine Korrelation von ca. 0,8.

$$\text{Korrelation} = \frac{\text{Fitness}}{1 - \frac{\#\text{Parameter}}{100}}$$

Auf Basis dieses Schätzers lassen sich somit, im Rahmen der Daten der Simulatorstudie, gute Vorhersagen treffen.

6.2 Realfahrten

Wie in Abschnitt 6.1 gezeigt, funktioniert das hier beschriebene Verfahren recht gut zur Schätzung von Fahrparametern. Im folgenden Abschnitt soll dieses Verfahren nun auf Realfahrten aus der Zusatzstudie angewendet werden.

6.2.1 Datenbasis

Als Grundlage dienen Fahrten von 26 Probanden. Jeder dieser Probanden ist drei Szenarien gefahren, die jeweils ca. eine halbe Stunde lang waren. Eine genaue Beschreibung dieser Szenarien ist in Abschnitt 2.3.2 gegeben.

| Code | IDs | Zielgrößen | | | Schnee/Eis |
|------|------------|------------|----|----|------------|
| | | A | B | C | |
| A0AX | 55, 56, 57 | 3 | 49 | 25 | nein |
| A5H7 | 1, 2, 3 | 20 | 41 | 40 | nein |
| A9B9 | 4, 5, 6 | 5 | 54 | 25 | ja |
| B5M7 | 7, 8, 9 | 5 | 33 | 40 | nein |
| B8G9 | 61, 62, 63 | 3 | 52 | 40 | nein |
| C7U9 | 10, 11, 12 | 12 | 47 | 30 | nein |
| D0I8 | 13, 14, 15 | 6 | 50 | 20 | ja |
| E2H0 | 16, 17, 18 | 11 | 59 | 25 | nein |
| E2S2 | 19, 20, 21 | 3 | 34 | 30 | ja |
| G2J5 | 64, 65, 66 | 13 | 50 | 25 | nein |
| G3C7 | 22, 23, 24 | 18 | 43 | 35 | ja |
| G4M7 | 25, 26, 27 | 81 | 49 | 45 | nein |
| H0K0 | 73, 74, 75 | 1 | 60 | 15 | ja |
| H3S7 | 28, 29, 30 | 10 | 40 | 25 | nein |
| H7B0 | 58, 59, 60 | 10 | 54 | 40 | nein |
| H7L1 | 67, 68, 69 | 23 | 57 | 40 | nein |
| I6E8 | 31, 32, 33 | 5 | 48 | 15 | nein |
| J0A6 | 34, 35, 36 | 2 | 57 | 35 | ja |
| J3P4 | 70, 71, 72 | 2 | 52 | 20 | nein |
| K0C7 | 76, 77, 78 | 7 | 44 | 40 | nein |
| K6S9 | 37, 38, 39 | 12 | 54 | 25 | ja |
| MXS5 | 40, 41, 42 | 10 | 44 | 35 | nein |
| P0B4 | 43, 44, 45 | 74 | 50 | 40 | nein |
| R1D4 | 52, 53, 54 | 4 | 53 | 25 | nein |
| U7A2 | 46, 47, 48 | 30 | 50 | 40 | nein |
| U9E1 | 49, 50, 51 | 25 | 48 | 40 | ja |

Tabelle 6.2: Zielgrößen der Probandenfahrten

Tabelle 6.2 zeigt eine Übersicht über die 26 Probanden, einschließlich der Werte für die zu untersuchenden Zielgrößen. Die acht Fahrten, die auf glatter Fahrbahn durchgeführt wurden, sind entsprechend gekennzeichnet. Der Probandencode setzt sich aus Name und Geburtsjahr der Eltern zusammen. Damit ist eine Nachträgliche Zuordnung von Versuchsperson zu Testdaten möglich. Eine Zuordnung in die andere Richtung ist jedoch nicht möglich.

Die Untersuchung ist auf drei erhobene Zielgrößen durchgeführt worden. Diese werden im Folgenden einzeln behandelt. Zusätzlich wurde mit dem Algorithmus eine Optimierung auf Zufallswerte durchgeführt.

6.2.2 Anwendung auf Zielgröße A

Die Zielgröße A war die primäre Zielgröße, die mit diesem Verfahren bestimmt werden sollte. Der mögliche Wertebereich liegt zwischen 0 und 100. Die Probanden der Studie wiesen hauptsächlich Werte zwischen 0 und 30 auf. Lediglich zwei Probanden zeigten sehr hohe Werte. Durch diese Tatsache fehlten besonders in den oberen Wertebereichen Referenzdaten, so dass Schätzungen in diesem Bereich erwartungsgemäß recht schlecht ausfallen werden. Für die Untersuchung dieser Zielgröße wurden die Fahrten bei glatter Fahrbahn nicht mit berücksichtigt, da sie durch glatte Fahrbahn sehr stark beeinflusst wurde.

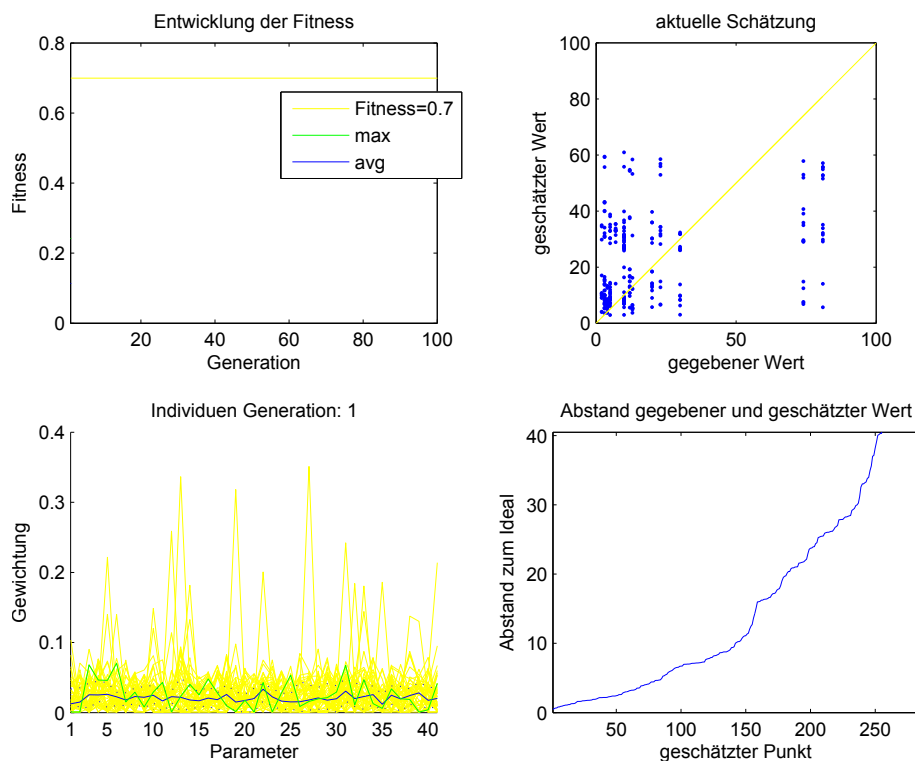


Abbildung 6.5: Generation 1 - Probandenstudie - Kenngröße A

Abbildung 6.5 zeigt die Startpopulation. Zu erkennen ist eine sehr breite Streuung der geschätzten Werte. Besonders im kleinen Wertebereich sind bereits einige Punkte gut geschätzt, was die Anfangsfitness von ca. 0,22 erklärt.

Nach 100 Generationen ist in Abbildung 6.6 zu sehen, dass die Schätzung wesentlich besser als in der Startpopulation aussieht. Besonders im kleinen Wertebereich wurden gute Ergebnisse erreicht. Auch die Schätzwerte für die Probanden im hohen Wertebereich ist besser geworden, auch wenn diese auf Grund der fehlenden Referenzdaten nicht ganz die Zielwerte erreichen kann.

Insgesamt wurde eine Fitness von 0,62 mit Hilfe von 13 Parametern erreicht. Die entspricht einer Korrelation zwischen geschätzten und gemessenen Wert von ca. 0,71.

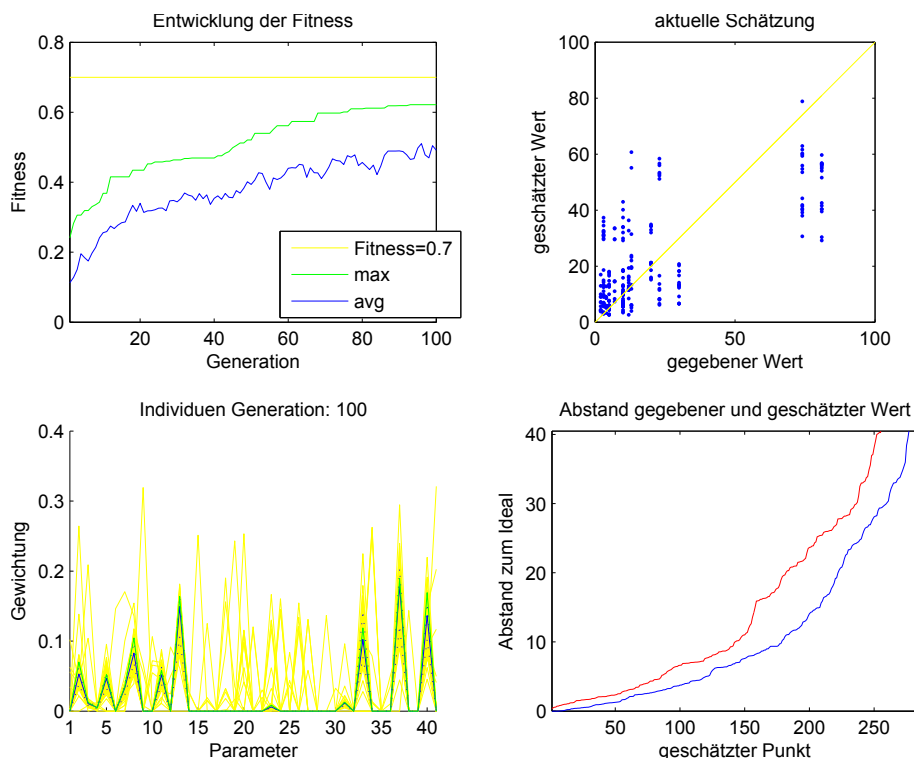


Abbildung 6.6: Generation 100 - Probandenstudie - Kenngröße A

6.2.3 Anwendung auf Zielgröße B

Für die zweite Kenngröße sind prinzipiell alle positiven Werte möglich. Jedoch wurden durch die Probanden nur Werte im Bereich zwischen 33 und 60 erreicht. Die Werte sind in diesem Wertebereich recht gleichmäßig verteilt.

In der Startpopulation, siehe Abbildung 6.7 sind die Schätzwerte noch recht gleichmäßig verteilt und eine Korrelation ist nicht zu erkennen.

Nach 100 Generationen, siehe Abbildung 6.8 ist eine Korrelation zu erkennen. Auch in dem Abstandsdiagramm ist die dargestellte Kurve um einiges flacher als in Abbildung 6.7.

Es wurde eine Fitness von etwa 0,67 erreicht. Um diesen Wert zu erreichen wurden zwölf Parameter ausgewählt. Dies entspricht einer Korrelation von ca 0,76.

6.2.4 Anwendung auf Zielgröße C

Die letzte untersuchte Zielgröße hat einen Wertebereich zwischen 10 und 50. Während der Studie wurden die Zielwerte jedoch nur mit einer Genauigkeit von 5 bestimmt. Die Randbereiche des Wertebereichs wurde durch die Probanden nicht erreicht.

Wieder sind in der Startpopulation, siehe Abbildung 6.9, die Schätzwerte noch recht zufällig verteilt. Allerdings ist bereits zu erkennen, dass kleinere Werte eher mit kleineren und größere eher mit größeren geschätzt werden.

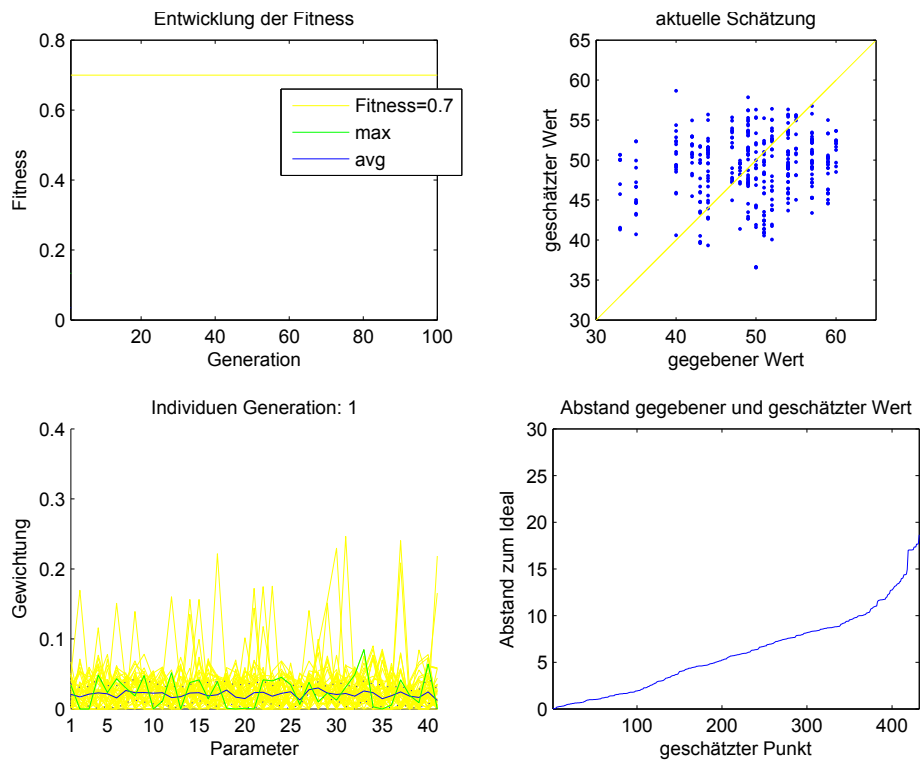


Abbildung 6.7: Generation 1 - Probandenstudie - Kenngröße B

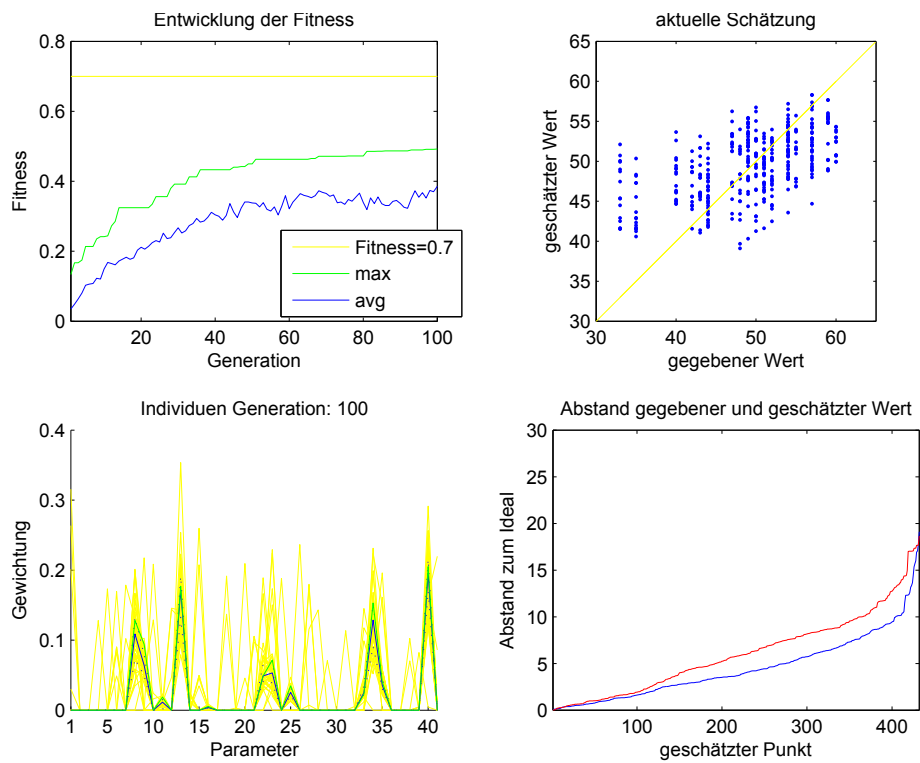


Abbildung 6.8: Generation 100 - Probandenstudie - Kenngröße B

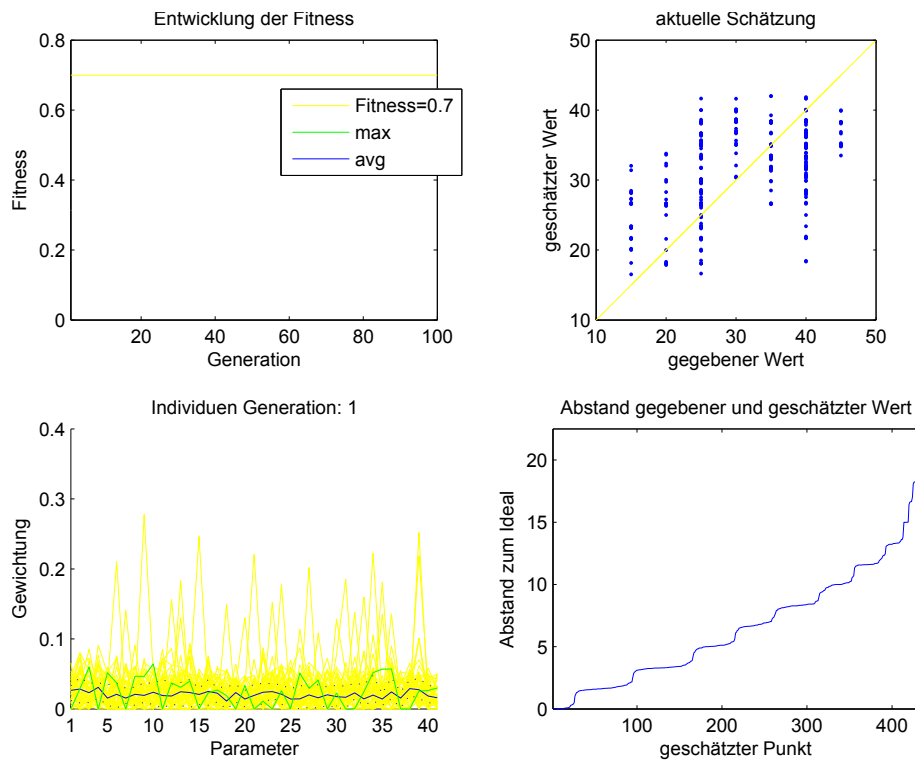


Abbildung 6.9: Generation 1 - Probandenstudie - Kenngröße C

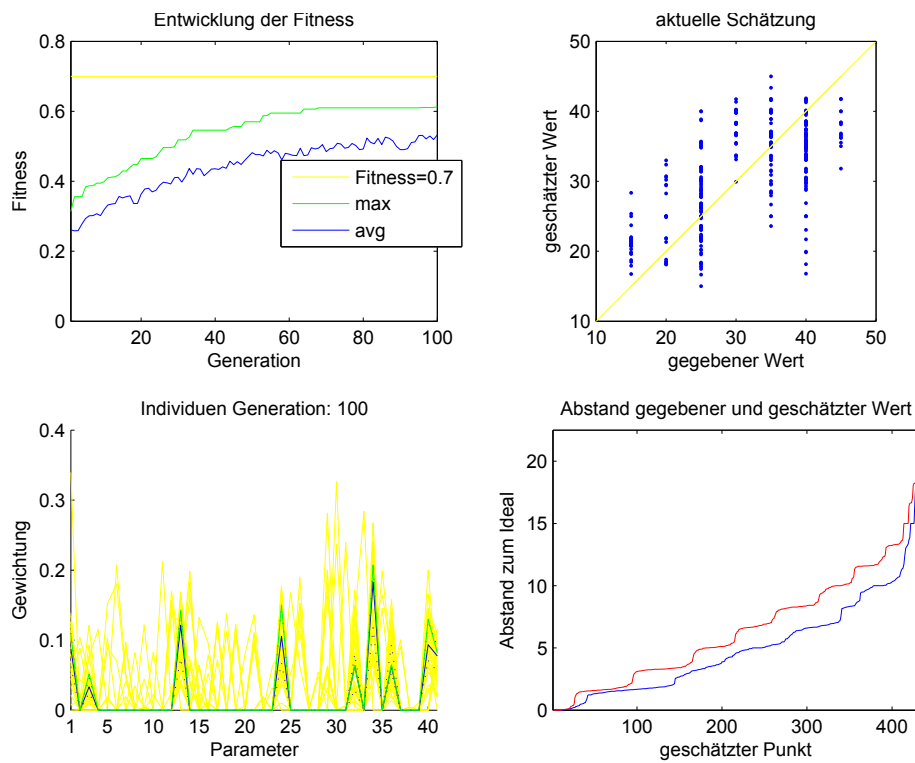


Abbildung 6.10: Generation 100 - Probandenstudie - Kenngröße C

Auch für diese Kenngröße ist nach 100 Generationen wieder eine gute Verbesserung zu erkennen. Zum Einen ist die Abstandskurve, siehe Abbildung 6.10, wesentlich flacher, zum Anderen sind die Schätzwerte näher an die gelbe Linie gerückt.

Insgesamt wurde eine Fitness von ca 0,62 mit Hilfe von neun Parametern erreicht. Dies entspricht einer Korrelation von ca. 0,67.

6.2.5 Kontrolluntersuchung

Um zu zeigen, dass die bisher erreichten Korrelationen nicht nur Zufall, bzw. systematisch durch die Modellgröße bedingt sind, wird eine weitere Zielgröße betrachtet, die jedoch nichts mit den Eigenschaften des Fahrers zu tun hat, sondern Zufallswerte darstellen. Diese Werte liegen gleichverteilt im Bereich zwischen 0 und 100.

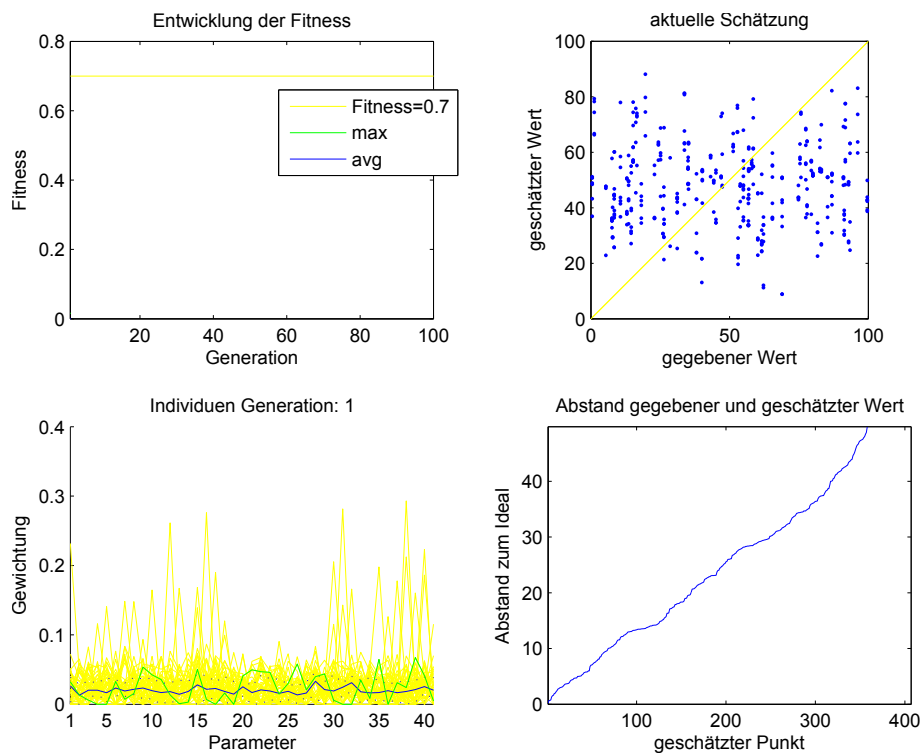


Abbildung 6.11: Generation 1 - Probandenstudie - Zufallswerte

In Abbildung 6.11 ist zu sehen, dass die Zufallswerte gleichmäßig über den Wertebereich verteilt sind. Auch die Schätzwerte sind in der Startpopulation zufällig, so dass keinerlei Korrelation zu erkennen ist.

Nach 100 Generationen konnte zwar ähnliche Verbesserung der Schätzung erreicht werden, siehe Abbildung 6.12. Jedoch zeigt weder das Diagramm mit den Schätzwerten, noch die Abstandskurve eine Korrelation.

Nach 100 Generationen wurde lediglich eine Fitness von ca. 0,3 erreicht werden. Bei den Zielgrößen A bis C wurde dieser Wert zum Teil schon mit der Startpopulation erreicht. Bei den benutzen 16 Parametern entspricht dies einer Korrelation von ca. 0,36.

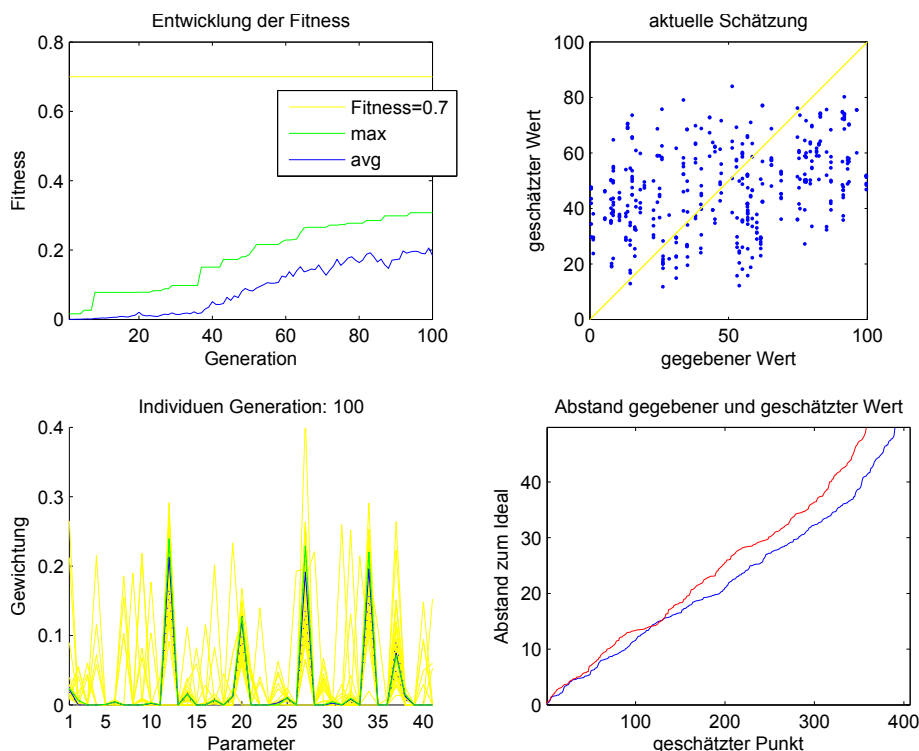


Abbildung 6.12: Generation 100 - Probandenstudie - Zufallswerte

6.3 Auswertung der Ergebnisse

Der vorgestellte Algorithmus wurde für Werte der Simulatorstudie, drei Zielgrößen bezüglich der Realfahrten und für eine Zielgröße aus Zufallswerten angewendet. In Tabelle 6.3 sind die Ergebnisse der einzelnen Untersuchungen noch einmal zusammengefasst.

| Untersuchung | max. Fitness | Anz. Parameter | Korrelation |
|-----------------|--------------|----------------|-------------|
| Simulatorstudie | 0,72 | 10 | 0,8 |
| Zielgröße A | 0,62 | 13 | 0,71 |
| Zielgröße B | 0,67 | 12 | 0,76 |
| Zielgröße C | 0,62 | 9 | 0,67 |
| Zufallswerte | 0,3 | 16 | 0,36 |

Tabelle 6.3: Ergebnisse der Untersuchung

Es wurde gezeigt, dass auf Basis von ähnlichen Fahrten Fahrer bezüglich gewisser Zielgrößen eingeschätzt werden können. Die Ergebnisse beruhen nicht nur auf einem überlernen des Systems, welches sich genau auf die Referenzdatenbank anpasst. Andernfalls hätte das System auch die Zufallswerte lernen können, jedoch sind die Ergebnisse für die Zufallswerte wesentlich schlechter.

Kapitel 7

Schlussfolgerungen und Abschluss

In den vorangegangenen Kapiteln wurde gezeigt, wie mit Hilfe von kontinuierlich anfallenden Fahrparametern Rückschlüsse auf verschiedene Fahrereigenschaften gewonnen werden können. In diesem Kapitel sollen die Ergebnisse noch einmal zusammengefasst werden. Anschließend erfolgt ein Ausblick, wie das entstandene System noch eingesetzt bzw. erweitert werden kann.

7.1 Zusammenfassung der Ergebnisse

Im ersten Schritt der Arbeit wurde das Fahrverhalten von verschiedenen Fahrern an Hand von früher aufgezeichneten Realfahrten analysiert. Dabei wurde gezeigt, wie sich aus kontinuierlich anfallenden Fahrparametern, wie z. B. aktuelle Geschwindigkeit, Abstand zum Vordermann oder Beschleunigung charakteristische Kenngrößen ermitteln lassen. Mit Hilfe dieser Kenngrößen kann das bisherige Fahrverhalten der Fahrer während der Fahrt beschrieben werden.

Im zweiten Schritt wurde mit dessen Hilfe ein Algorithmus entwickelt, der verdeckte Fahrereigenschaften schätzen kann. Dieser ruht auf der Grundannahme, dass Fahrer mit ähnlichen Eigenschaften sich in gewisser Weise ähnlich verhalten. Aus dieser Annahme abgeleitet basiert die Schätzung auf der Ähnlichkeit der zu untersuchenden Fahrt zu Fahrten aus einer Referenzdatenbank. Die Auswahl und Gewichtung der zu betrachtenden Kenngrößen wurden mit Hilfe eines evolutionären Algorithmus getroffen und optimiert. Während der Entwicklung wurde dabei auf eine Zielgröße optimiert. Für diese Untersuchung stellten die Fahrten der Simulatorstudie die Referenzdatenbank dar.

Es wurde eine eigene Studie durchgeführt, um den Algorithmus für verschiedene Zielwerte an Realfahrten zu testen. Für diese Studie sind 26 Probanden jeweils eine Strecke von ca 100 km zwischen Wolfsburg und Braunschweig gefahren. Diese Strecke wurde in drei verschiedene Szenarien unterteilt, die jeweils ein unterschiedliches charakteristisches Profil hatten. Im ersten Szenario ging es primär um das Verhalten auf Autobahnen. Das zweite Szenario war eine Stadtfahrt in Braunschweig. Das

letzte Szenario war eine Kombination aus Landstraße, einigen Ortsdurchfahrten und einem kleinen Stück Autobahn.

Die Anwendung des evolutionären Algorithmus auf die Fahrten der Simulatorstudie und die selbst eingefahrenen Realfahrten zeigten gute Ergebnisse. In der Simulatorstudie konnte für die untersuchte Zielgröße eine Korrelation zwischen geschätzten und gegebenen Werten von 0,8 erreichen. Die Untersuchung der Realfahrten zeigte nicht ganz so gute Ergebnisse. Es wurde für die drei Zielgrößen Korrelationen von 0,67, 0,71 und 0,76 erreicht. Ein Grund für diese etwas schlechteren Ergebnisse könnte sein, dass die Simulatorstudie unter Laborbedingungen und die Realfahrten als Feldversuch durchgeführt wurden. Dadurch ist bei den Realfahrten mit wesentlich mehr Störeinflüssen zu rechnen.

Um ein Überlernen des Modells auszuschließen, wurden die Realfahrten hinsichtlich einer Zielgröße, die lediglich aus Zufallswerten bestand, optimiert. Das entstandene Modell konnte bei der Schätzung der Zielgröße lediglich eine Korrelation von 0,36 erreichen. Dieser Wert ist wesentlich schlechter als die Korrelationen auf tatsächliche Fahreigenschaften. Damit kann davon ausgegangen werden, dass nicht systembedingt immer solch hohen Korrelationen erreicht werden können.

7.2 Vergleich der Ergebnisse mit den Erwartungen und Zielen

Im Rahmen dieser Arbeit sollte ein Verfahren entstehen, mit dem verdeckte Fahreigenschaften kontinuierlich während der Fahrt geschätzt werden können.

Es wurde ein Modell entwickelt, welches zumindest für die untersuchten Zielgrößen bei den untersuchten Probanden recht gut die Fahreigenschaften schätzen kann. Mit Hilfe des evolutionären Algorithmus ist es möglich das Modell, durch die benutzten Gewichtungen der einzelnen Kenngrößen, so zu konfigurieren, dass prinzipiell jede numerische Zielgröße geschätzt werden kann.

Alle Parameter sind statistische Kennwerte, die eine gewisse Größe beschreiben. Die benötigten Messwerte stehen in modernen Fahrzeugen über den Fahrzeug-CAN zur Verfügung. Es wurde insbesondere auch darauf geachtet, dass die Ermittlung der Kennwerte kontinuierlich möglich ist, ohne dass alle bisherigen Messwerte gespeichert werden müssen.

Da für die untersuchten Zielgrößen eine recht hohe Korrelation zwischen geschätzten und gegebenen Werten erreicht wurde, kann davon ausgegangen werden, dass die Grundannahme auf der die gesamte Arbeit basiert, zumindest für die gewählten Zielgrößen gehalten werden kann.

Das entstandene Modell, einschließlich der ermittelten Konfigurationen für die einzelnen Zielgrößen, wurden mit Hilfe von ADTF in einem Testfahrzeug umgesetzt. ADTF ist eine Plattform, mit deren Hilfe Prototypen von Fahrassistenzsystemen im Fahrzeug getestet werden können. Das System im Fahrzeug ist in der Lage die

Messwerte kontinuierlich auszuwerten und somit kontinuierlich eine Schätzung der Fahrereigenschaft zu geben. Die genaue Umsetzung des Systems in ADTF ist nicht Bestandteil dieser Arbeit.

Insgesamt wurde das anfangs gesetzte Ziel erreicht.

7.3 Einschränkungen der Ergebnisse und deren Anwendbarkeit

Eine Einschränkung basiert auf den Referenzdaten. Diese wurden von lediglich 26 Personen eingefahren, die zusätzlich auch noch recht homogen verteilt waren. Dies ist auch der Grund dafür, dass die Verteilung der Zielwerte unter den Probanden, insbesondere für Zielgröße A, nicht sehr gleichmäßig ist. Für einen produktiven Einsatz im Fahrzeug ist es essentiell, dass eine Normstichprobe aufgezeichnet wird. Durch eine entsprechend größere Referenzdatenbank sollte es auch möglich sein die Probanden besser einzuschätzen, da davon auszugehen ist, dass dann auch ähnlichere Fahrer gefunden werden können.

Eine zweite Einschränkung ist systembedingt gegeben. Durch die Interpolation zwischen ähnlichen Fahrern ist eine Schätzung nur für numerische Werte sinnvoll umsetzbar. Geordnete Klassen lassen sich evtl. noch auf numerische Werte zurückführen. Geschätzte Werte könnten durch entsprechende Rundung den Klassen zugeordnet werden. Prinzipiell ist eine Einordnung in nicht geordnete Klassen, wie männlich oder weiblich, jedoch nicht sinnvoll möglich. Für eine solche Einordnung wäre beispielsweise, der Einsatz von Bayes-Klassifikatoren, siehe auch [BH03], besser geeignet.

7.4 Erweiterungsmöglichkeiten

Wie bei jeder Arbeit gibt es auch in dieser Arbeit eine Menge Erweiterungsmöglichkeiten. Einige von diesen sollen hier gezeigt werden.

Wie bereits erwähnt wäre ein erster Schritt in Richtung einer produktiven Anwendung das Aufzeichnen einer Normstichprobe mit mehreren hundert Referenzfahrten. Diese ermöglicht eine genauere und sicherere Schätzung der gewünschten Zielgrößen.

Des Weiteren könnte das Modell ausgebaut werden. Dies ist in verschiedene Richtungen möglich. Eine Möglichkeit ist die Untersuchung weiterer Fahrparameter, beispielsweise der Lenkwinkel, der Bremsdruck, der Seitenabstand zu anderen Fahrzeugen oder Hindernissen oder der benutzte Drehzahlbereich. Einige dieser Größen lassen sich bereits aus den CAN-Nachrichten des Fahrzeugs entnehmen, andere, wie den Seitenabstand, müssten erst durch weitere Sensoren verfügbar gemacht werden.

Außerdem wäre sicherlich die Untersuchung weiterer Fahrsituationen sinnvoll. So könnten Überholvorgänge interessant sein. Diese lassen sich dahingehend unter-

scheiden, ob Fahrzeuge entgegen kommen können oder nicht. Auch das Verhalten bei Abbiegevorgängen oder beim Einparken lässt vielleicht Rückschlüsse auf den Fahrer zu. Beim Betrachten des Straßenverkehrs fallen sicherlich noch viele weitere Situationen auf, die sich entsprechend analysieren lassen.

Zur Beschreibung des Fahrers könnten auch Größen oder Verhaltensmuster angewendet werden, die nicht direkt aus dem Fahrverhalten resultieren. Kandidaten hierfür wären zum Beispiel das Radio oder die Heizung. Welchen Sender hört der Fahrer? Schaltet er häufig um? Bevorzugt er hohe oder niedrige Temperaturen? Wird häufig gelüftet? Mit Hilfe der Spiegel- und Sitzposition lassen sich vielleicht Aussagen über die Größe und Statur des Fahrers treffen. All diese Informationen könnten vielleicht genutzt werden um weitere Rückschlüsse auf das Fahrverhalten zu ziehen.

Ein weiterer Punkt der noch nicht betrachtet wurde ist der Einsatz von Fahrassistenzsystemen. Diese Betrachtung ist hinsichtlich zweier Punkte interessant. Zum Einen: Nutzt der Fahrer Assistenzsysteme wie Adaptive Cruise Control (ACC) oder einen Spurhalteassistenten. Andererseits ist es auch interessant, wie diese Assistenten die Fahrparameter beeinflussen. So könnte beispielsweise mit einer Geschwindigkeitsregelautomatik beschleunigt werden, ohne dass der Fahrer das Fahrpedal betätigen muss. In der durchgeführten Studie wurden die Probanden dazu angehalten weder einen Spurhalteassistenten noch eine solche Geschwindigkeitsregelautomatik zu nutzen.

Eine naheliegende Erweiterung der Arbeit wäre natürlich die Anwendung des Algorithmus auf weitere Fragestellungen. So könnte man versuchen alle möglichen numerischen Werte des Fahrers zu schätzen.

Auch systematisch sind Änderungen oder Erweiterungen an dem beschriebenen Modell denkbar. So könnte beispielsweise nicht die komplette bisherige Fahrt zur Schätzung genutzt werden, sondern nur die letzten Minuten. Dadurch ließen sich Größen schätzen, die sich während der Fahrt ändern. Beispiele hierfür wären z. B. die Müdigkeit oder das aktuelle Stresslevel. Auch ist es möglich den evolutionäre Algorithmus durch andere Lernverfahren zu ersetzen. Wie bereits genannt, könnte unter anderem ein Bayes-Klassifikator genutzt werden um das Geschlecht des Fahrers zu schätzen.

7.5 Einsatzmöglichkeiten und Nutzen

Der primäre Einsatzzweck ist die Parametrisierung von Fahrassistenzsystemen. Komfortsysteme lassen sich bereits auf Werte des Fahrers vorkonfigurieren, so dass bei einer Abstandsautomatik beispielsweise ein fahrertypischer Abstand voreingestellt werden kann. Sicherheitssysteme lassen sich dadurch so konfigurieren, dass sie etwas früher auslösen, ohne, dass für den Fahrer gefühlt mehr Fehlauflösungen vorkommen.

Auch in der Fahrermodellierung, siehe auch [BGR10], können die geschätzten Eigenschaften genutzt werden, um den Fahrzeugführer und sein Verhalten näher zu beschreiben.

Abschließend kann gesagt werden, dass mit dieser Arbeit prinzipiell gezeigt wurde, dass sich Fahreigenschaften auf Grund des Fahrverhaltens schätzen lassen. Das entstandene System bietet viel Potenzial für Erweiterungen und kann sinnvoll bei der Parametrisierung von Fahrerassistenzsystemen eingesetzt werden.

Literaturverzeichnis

- [AB09] *Kapitel Die Leistungsfähigkeit des Menschen für die Fahrzeugführung.* In: ABENDROTH, Bettina ; BRUDER, Ralph: *Handbuch Fahrerassistenzsysteme.* 2009, S. 4–14
- [BGR10] BAUER, Colin ; GONTER, Mark ; ROJAS, Raul: Fahrerspezifische Analyse des Fahrverhaltens zur Parametrierung aktiver Sicherheitssysteme. In: *Sicherheit durch Fahrerassistenz.* München, April 2010
- [BH03] BERTHOLD, M. ; HAND, D.J.: *Intelligent data analysis: an introduction.* Springer Verlag, 2003. – ISBN 3540430601
- [BKKN03] BORGELT, Christian ; KLAWONN, Frank ; KRUSE, Rudolf ; NAUCK, Detlef: *Neuro-Fuzzy-Systeme: Von den Grundlagen künstlicher neuronaler Netze zur Kopplung mit Fuzzy-Systemen.* Vieweg+ Teubner Verlag, 2003. – ISBN 3528252650
- [Don09] *Kapitel Fahrerverhaltensmodelle.* In: [WH09], S. 15–23
- [Fre10] FREIBURG, Till: *Entwicklung eines fahrzeuginternen Verfahrens zur Ermittlung der Risikobereitschaft des Fahrers,* RWTH Aachen, Diplomarbeit, 2010
- [Gil85] GILLIES, A.M.: *Machine Learning Procedures for Generating Image Domain Feature Detectors.* Michigan, University of Michigan, PhD Thesis, 1985
- [GKK04] GERDES, Ingrid ; KLAWONN, Frank ; KRUSE, Rudolf: *Evolutionäre Algorithmen.* Wiesbaden : Vieweg, 2004
- [Kol92] KOLODNER, Janet L.: An introduction to case-based reasoning. In: *Artificial Intelligence Review* 6 (1992), 3-34. <http://dx.doi.org/10.1007/BF00155578>. – ISSN 0269–2821. – 10.1007/BF00155578
- [Müh92] MÜHLENBEIN, H.: How genetic algorithms really work: Mutation and hillclimbing. In: MÄNNER, R. (Hrsg.) ; MANDERICK, B. (Hrsg.): *Foundations of Genetic Algorithms Parallel Problem Solving from Nature.* North-Holland, Amsterdam, 1992, S. 15–26
- [Ste10] STEINBRECHER, Matthias: *Intelligent Data Analysis.* 2010. – Vorlesungsfolien

- [Sys89] SYSWERDA, Gilbert: Uniform Crossover in Genetic Algorithms. In: SCHAFFER, J. D. (Hrsg.): *ICGA*, Morgan Kaufmann, 1989. – ISBN 1-55860-066-3, S. 2–9
- [TSWH09] TOUTENBURG, H. ; SCHOMAKER, M. ; WISSMANN, M. ; HEUMANN, C.: *Arbeitsbuch zur deskriptiven und induktiven Statistik*. Springer, 2009. – ISBN 3540890351
- [WDS09] *Kapitel Adaptive Cruise Control*. In: [WH09], S. 478–521
- [WH09] WINNER, Hermann ; HAKULI, Stephan ; WOLF, Gabriele (Hrsg.): *Handbuch Fahrerassistenzsysteme*. Wiesbaden : Vieweg+Teubner, 2009
- [Win09] *Kapitel Frontalkollisionsschutzsysteme*. In: [WH09], S. 522–542

Selbstständigkeitserklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Magdeburg, den 5. April 2011

Pascal Held