
Konzept zu statistisch repräsentativen Verkehrsbeobachtungen

Fachveröffentlichung der
Bundesanstalt für
Straßenwesen

Konzept zu statistisch repräsentativen Verkehrsbeobachtungen

von

Peter Hosten, Alexander Tenbrock, Julian Bock, Adrian Zlock
fka GmbH, Aachen

Impressum

Fachveröffentlichung zu Forschungsprojekt: 82.0735
Konzept zu statistisch repräsentativen Verkehrsbeobachtungen

Fachbetreuung:
Maximilian Grabowski, Marcus Wisch

Referat:
Automatisiertes Fahren, Passive Fahrzeugsicherheit, Biomechanik

Herausgeber:
Bundesanstalt für Straßenwesen
Brüderstraße 53, D-51427 Bergisch Gladbach
Telefon: (0 22 04) 43 - 0

<https://doi.org/10.60850 fv-f-82.0735>

Bergisch Gladbach, Juli 2024

Zu diesem Forschungsprojekt werden nur die Kurzfassung und der Kurzbericht veröffentlicht. Die Langfassung des Schlussberichts kann auf Anfrage an verlag@bast.de zur Verfügung gestellt werden.

Es wird darauf hingewiesen, dass die unter dem Namen der Verfasser veröffentlichten Berichte nicht in jedem Fall die Ansicht des Herausgebers wiedergeben. Nachdruck und photomechanische Wiedergabe, auch auszugsweise, nur mit Genehmigung der Bundesanstalt für Straßenwesen, Stabsstelle Presse und Kommunikation.

Kurzfassung

Konzept zu statistisch repräsentativen Verkehrsbeobachtungen

Ziel des vorliegenden Forschungsvorhabens ist die Entwicklung eines Konzeptes, das eine statistisch repräsentative Erhebung von Verkehrsszenarien in verschiedenen Verkehrsräumen erlaubt. Die dadurch zu gewinnende Datengrundlage soll der Bewertung von automatisierten Kraftfahrzeugsteuerungen dienen.

Ausgangspunkt und Orientierung für das methodische Vorgehen ist in diesem Forschungsvorhaben das methodische Rahmenkonzept für Verhaltensbeobachtung im fließenden Verkehr (Hautzinger, Pfeiffer & Schmidt, 2012). Durch Adaptation und Anwendung der Entwicklungsschritte dieses Rahmenkonzeptes konnte ein Erfahrungswissen gewonnen werden, auf dessen Grundlage ein Konzept für einen Stichprobenplan erstellt worden ist.

Dem Rahmenkonzept folgend wurde zunächst der Untersuchungsgegenstand abgrenzt. Nach Vorstellung verschiedener Szenarienkonzepte und ihrer Beschreibungsgrößen erfolgte die Beschreibung eines allgemeinen Modells der Grundgesamtheit. Dieses Modell berücksichtigt die Abhängigkeit der Grundgesamtheit von Kontextfaktoren als Einflussgrößen auf den Verkehr. Darüber hinaus sieht das Modell der Grundgesamtheit eine Partitionierung des Untersuchungsgebietes und Untersuchungszeitraums in Untersuchungseinzelemente vor, in denen die Untersuchungseinheiten – in vorliegendem Fall die Szenarien – mittels der verschiedenen Formen der Erhebung beobachtet werden können (vgl. (Bäumer et al., 2017b, S. 25)). Diese Partitionierung hat zum Ziel, den Einfluss der Kontextfaktoren auf die Ausprägung und Häufigkeit der Szenarien innerhalb eines Untersuchungseinzelementes zu minimieren und damit eine Modellierung als von diesen Kontextfaktoren unabhängigen Zufallsprozess zu ermöglichen. Die nächsten beiden Entwicklungsschritte des Rahmenkonzeptes betreffen die Auswahl der Beobachtungsorte und -zeiten. Im Rahmen dieses Vorhabens werden diese beiden Aspekte verallgemeinert bei der Auswahl der Beobachtungseinheiten behandelt. Allgemein kann zwischen stationären Beobachtungseinheiten (z. B. Beobachtung mittels Drohne) und instationären Beobachtungseinheiten (z. B. Beobachtung mittels fahrzeuggebundener Sensorik) unterschieden werden. Im Gegensatz zur stationären Beobachtungseinheit wird bei einer instationären Beobachtungseinheit bedingt durch den bewegten Beobachter zu verschiedenen Zeitpunkten ein anderer Streckenabschnitt des Untersuchungsgebietes beobachtet. Bei der Auswahl der Beobachtungseinheiten muss demnach bei einer stationären Beobachtungseinheit ein Ort und ein Zeitpunkt der Beobachtung definiert werden, während bei einer instationären Beobachtungseinheit darüber hinaus auch die Bewegung des Beobachters definiert werden muss. Da die Auswahl des Beobachtungsortes bei stationären und instationären Beobachtungseinheiten bei der Stichprobenplanung ein wesentlicher Aspekt ist, wurde ein Schwerpunkt auf die Entwicklung einer Methode zur Zufallsauswahl von Beobachtungsorten gelegt. Ein weiterer Entwicklungsschritt betrifft die Bereitstellung technischer Hilfsmittel und die Beobachterschulung. In diesem Zusammenhang wurden verschiedenen Formen der Erhebung detailliert betrachtet und gegenübergestellt. Bei der Betrachtung des letzten Entwicklungsschrittes des Rahmenkonzeptes wurde ein Hochrechnungsverfahren entwickelt, welches die Tatsache berücksichtigt, dass im Regelfall nur ein Teil eines Untersuchungseinzelementes beobachtet werden kann und somit eine Extrapolation der Beobachtungen auf das gesamte Untersuchungseinzelement notwendig ist. Dazu wurde das Auftreten der Szenarien als Zufallsprozess modelliert. Dies erlaubt die Schätzung einer Wahrscheinlichkeit für das Erfassen eines Szenarios innerhalb eines Untersuchungseinzelementes mittels einer spezifischen Form der Erhebung. Diese Erfassungswahrscheinlichkeit kann dann genutzt werden, um eine Extrapolation der Beobachtungen durchzuführen. Für die weitere Untersuchung wurde das Hochrechnungsverfahren auf reale Verkehrsdaten angewendet und als Untersuchungseinheit das Fahrstreifenwechselszenario gewählt.

Basierend auf den obengenannten Arbeiten wurde ein Konzept für einen Stichprobenplan erstellt, das auf eine möglichst repräsentative Erhebung in den Verkehrsräumen Autobahn, Überland und Innerorts abzielt. In diesem Zusammenhang wird darauf hingewiesen, dass die Beschreibung der Fahraufgabe mittels beobachteter Verkehrsszenarien zwangsläufig aufgrund des Zufallscharakters des Auftretens der Verkehrsszenarien unvollständig ist. Aus praktischer Sicht ist es daher notwendig, den Grad der Vollständigkeit der Beschreibung der Fahraufgabe auf eine effiziente Weise zu erhöhen. Ein möglicher Ansatz kann darin bestehen, den durch die empirischen Daten aufgespannten Szenarienraum basierend auf zusätzlichem Wissen anzureichern. Dieser Ansatz wurde bereits im PEGASUS Projekt bei Einführung der logischen Szenarien verfolgt (PEGASUS, 2019).

Abstract

Concept for Statistically Representative Traffic Observations

The objective for this research project is the development of a concept that facilitates a statistically representative observation of various traffic domains. The dataset to be obtained in this way is intended to be used for the assessment of automated driving functions.

Starting point and orientation for the methodical approach in this research project is the methodological framework for observational surveys in moving traffic (Hautzinger et al., 2012). The experience-based knowledge gained in the adaptation and application of the development steps of the framework forms the basis for the conception of a sampling plan.

First, the object of study has been delimited according to the framework. After the introduction of various scenario concepts and their describing parameters, a general model for the population of scenarios is presented. This model incorporates the dependence of the population on context factors influencing the traffic. Furthermore, the model requires the partitioning of the area and period of investigation into investigation elements (cf. (Bäumer et al., 2017b, S. 25)). Within these, the unit under investigation – scenarios in this case – can be observed using different survey techniques. The aim of the partitioning is to minimize the influence of the context factor on the characteristic and the frequency of the scenarios and thus to enable the modelling as a stochastic process independent of these context factors. The next steps of the framework concern the selection of the observation location and times. In the context of this project, these aspects are handled in a generalized way when selecting observation units. These can be differentiated into stationary observation units (e.g. observation using drones) and instationary units (e.g. using a measurement vehicle). Unlike stationary observation units, in the case of instationary observation units a moving observer is covering different sections of the road at different points in time. The selection of the observation units must therefore consist of the definition of the time and location of the observation, while for instationary observation units the movement of the observer has to be defined as well. Because the selection of the observation location is a core aspect when setting up a sampling plan, the focus has been set on the development of a random sampling method for observation locations. A further step is related to the provision of technical resources. In this context, different forms of traffic survey techniques are presented and compared. In the final step of the framework an extrapolation method has been developed. This method considers the fact, that in many cases only a part of an individual investigation element can be observed and thus an extrapolation of the observed samples onto the complete investigation element is required. To this end, the occurrence of scenarios has been modelled as a stochastic process. This facilitates the estimation of the probability to capture a scenario within an individual investigation element by means of a specific survey technique. This probability can be used for the extrapolation of the observation. Finally, this method has been applied to real-world traffic data with a lane change scenario chosen as unit under investigation.

Based on the work described above, a concept has been drafted aiming at a representative survey in the traffic domains federal highways as well as rural and urban areas. It is pointed out that the description of the driving task using observed scenarios is inherently incomplete because of the stochastic nature of their occurrences. From a practical point of view, it is therefore necessary to increase the completeness of the description of the driving task in an efficient way. A possible starting point can be the enrichment of the scenario space partially covered by empirical data using additional knowledge. This approach has already been used at the introduction of logical scenarios (PEGASUS, 2019).

Kurzbericht zum BAST Forschungsprojekt FE 82.0735/2019 „Konzept zu statistisch repräsentativen Verkehrsbeobachtungen“

1 Aufgabenstellung

Der Absicherungs- bzw. Freigabeprozess für hochautomatisierte Fahrsysteme ist eine der größten aktuellen Herausforderungen auf dem Weg zum Serienbetrieb dieser Systeme. Eine der Kernproblemstellungen ist dabei die große Menge und Vielfalt an Situationen im realen Straßenverkehr, für die ein sicherer Betrieb nachgewiesen werden muss. Die klassische Methodik des „Freifahrens“, also dem Fahren einer sehr großen Gesamtdistanz mit dem aktiven Testsystem, ist für das hochautomatisierte Fahren nicht mehr wirtschaftlich und praktisch umsetzbar, da nach Berechnungen eine Distanz von mehreren Milliarden Kilometern mit dem Zielsystem bewältigt werden müsste (Wachenfeld & Winner, 2015, S. 458).

Es wird also eine Möglichkeit zur Raffung benötigt, um einen Freigabeprozess in einem realistischen Zeitraum durchführen zu können. Dazu wird auf Szenarien zurückgegriffen, in welche die Gesamtheit der Fahraufgabe eingeteilt wird. Während die Systematik zur Beschreibung relevanter Verkehrsszenarien in verschiedenen Forschungsprojekten derzeit entwickelt wird, fehlt ein Konzept, das deren repräsentative Erhebung in bestimmten Verkehrsräumen definiert. Die repräsentativen Szenarien sind jedoch unerlässlich, um die automatisierte Kraftfahrzeugsteuerung hinsichtlich einer positiven Risikobilanz zu bewerten.

Ziel des vorliegenden Forschungsvorhabens ist die Entwicklung eines Konzeptes, das eine statistisch repräsentative Erhebung von Verkehrsszenarien in verschiedenen Verkehrsräumen erlaubt. Die dadurch zu gewinnende Datengrundlage soll der Bewertung von automatisierten Kraftfahrzeugsteuerungen dienen. Folgende Anforderungen können in diesem Zusammenhang formuliert werden. Eine auf Basis des zu erstellenden Konzeptes erhobene Stichprobe von Verkehrsszenarien kann verwendet werden, um statistisch abgesicherte Aussagen über die tatsächliche Häufigkeit einzelner Szenarien im gesamten Verkehrsgeschehen bzw. innerhalb des betrachteten Verkehrsraums zu machen. Der Stichprobenplan berücksichtigt verschiedene Formen der Erhebung: fahrzeuggebundene Erhebung aus dem Verkehr, infrastrukturegebundene Beobachtungen sowie Beobachtungen aus der Luft. Die mit der Art der Erhebung einhergehenden Einschränkungen sollen berücksichtigt werden – vgl. Bild 1-1.

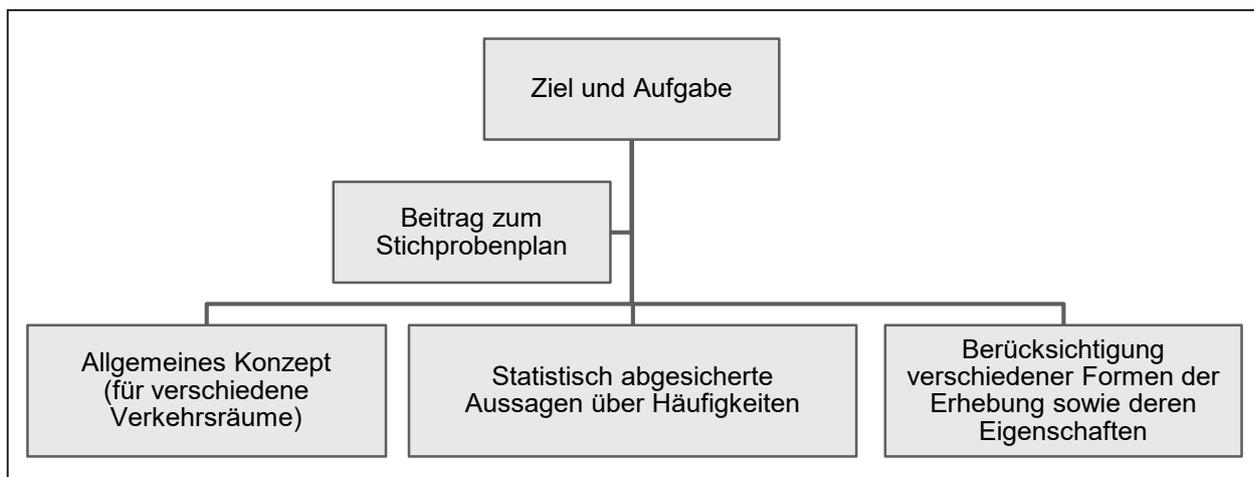


Bild 1-1: Ziel und Aufgabe des Forschungsvorhabens

2 Methodisches Vorgehen

Ausgangspunkt und Orientierung für das methodische Vorgehen ist in diesem Forschungsvorhaben das methodische Rahmenkonzept für Verhaltensbeobachtung im fließenden Verkehr (Hautzinger, Pfeiffer & Schmidt, 2012). Dieses Rahmenkonzept gliedert sich in sechs Entwicklungsschritte – vgl. Bild 2-1.

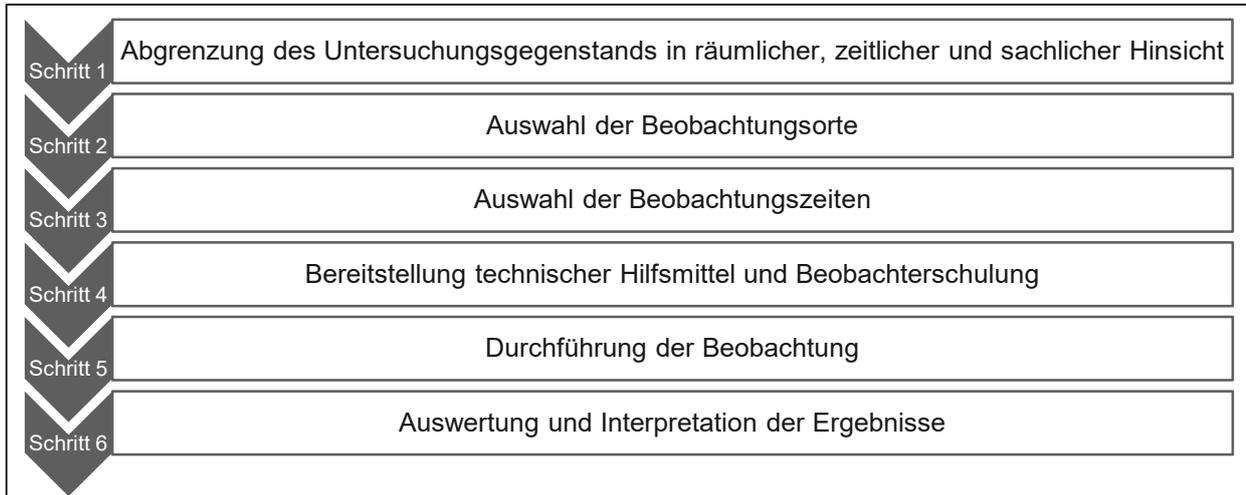


Bild 2-1: Entwicklungsschritte basierend auf dem methodischen Rahmenkonzept für Verhaltensbeobachtung im fließenden Verkehr nach (Hautzinger et al., 2012).

Im ersten Entwicklungsschritt ist eine Abgrenzung des Untersuchungsgegenstandes in räumlicher, zeitlicher und sachlicher Hinsicht vorzunehmen. Daher erfolgte zunächst die Konzeption eines allgemeinen Modells der Grundgesamtheit der Szenarien. Dieses Modell berücksichtigt die Abhängigkeit der Grundgesamtheit von Kontextfaktoren. Kontextfaktoren stellen statische, makroskopische oder verkehrsexterne Einflussgrößen auf das Verkehrsgeschehen dar. Beispiele für erstere Kategorie sind der Straßenverlauf in Form von Steigungen und Kurven, vorausgehende Unfälle, die Verkehrsstärke und Staus sind Faktoren makroskopischer Art. Die Witterungsbedingungen und der witterungsbedingte Straßenzustand beeinflussen als verkehrsexterne Größen ebenfalls das Verkehrsgeschehen und damit die Verkehrsszenarien. Für eine differenziertere Einordnung der Kontextfaktoren wird auf das 6-Ebenen-Modell verwiesen (Scholtes et al., 2021). Darüber hinaus sieht das Modell der Grundgesamtheit eine Partitionierung des Untersuchungsgebietes und Untersuchungszeitraums in Untersuchungseinzelemente vor, in denen die Untersuchungseinheiten – in vorliegendem Fall die Szenarien – mittels der verschiedenen Formen der Erhebung beobachtet werden können. Diese Partitionierung hat zum Ziel, den Einfluss der Kontextfaktoren auf die Ausprägung und Häufigkeit der Szenarien innerhalb eines Untersuchungseinzelementes zu minimieren – vgl. (Bäumer et al., 2017) – und damit eine Modellierung als von diesen Kontextfaktoren unabhängigen Zufallsprozess zu ermöglichen.

Der zweite und dritte Entwicklungsschritt betrifft die Auswahl der Beobachtungsorte und –zeiten. Im Rahmen dieses Vorhabens werden diese beiden Aspekte verallgemeinert bei der Auswahl der Beobachtungseinheiten behandelt. Allgemein kann zwischen stationären Beobachtungseinheiten (z. B. Beobachtung mittels Drohne) und instationären Beobachtungseinheiten (z. B. Beobachtung mittels fahrzeuggebundener Sensorik) unterschieden werden. Im Gegensatz zur stationären Beobachtungseinheit wird bei einer instationären Beobachtungseinheit bedingt durch den bewegten Beobachter zu verschiedenen Zeitpunkten ein anderer Streckenabschnitt des Untersuchungsgebietes beobachtet. Bei der Auswahl der Beobachtungseinheiten muss demnach bei einer stationären Beobachtungseinheit ein Ort und ein Zeitpunkt der Beobachtung definiert werden, während bei einer instationären Beobachtungseinheit darüber hinaus auch die Bewegung des Beobachters definiert werden muss. Aus stichprobentheoretischer Sicht ist eine uneingeschränkt zufällige Auswahl der Auswahlseinheiten notwendig, um statistisch abgesicherte Aussagen treffen zu können. Da die Auswahl des Beobachtungsortes bei stationären und instationären Beobachtungseinheiten bei der Stichprobenplanung ein wesentlicher Aspekt ist, wurde ein Schwerpunkt auf die Entwicklung einer Methode zur Zufallsauswahl von Beobachtungsorten im Verkehrsraum Autobahn gelegt.

Dazu wurde basierend auf BISStra Daten unter Verwendung von kubischen Splines ein kontinuierliches Modell des Streckennetzes parametrisiert, das eine Zufallsauswahl des Beobachtungsortes als auch eine Schätzung des Hilfsmerkmals Kurvenradius an den ausgewählten Beobachtungsorten ermöglicht. Exemplarisch für weitere Kontextfaktoren wurde basierend auf einer Klasseneinteilung des Hilfsmerkmal Kurvenradius untersucht, mit welcher relativen Häufigkeit eine Klasse des Hilfsmerkmals in einer Zufallsstichprobe enthalten ist. Die Untersuchung zeigt auf, dass neben der für die Klasseneinteilung gewählten Quantisierung der Stichprobenumfang – in diesem Fall die Anzahl der Beobachtungsorte – maßgeblich die Auswahlwahrscheinlichkeit bestimmter Klassen des Hilfsmerkmals bestimmt.

Der vierte Entwicklungsschritt betrifft die Bereitstellung technischer Hilfsmittel und die Beobachterschulung. In diesem Vorhaben wurden daher verschiedene Erhebungsverfahren betrachtet und gegenübergestellt. Dazu wurden zunächst einige in diesem Zusammenhang relevante Erhebungsgrößen erläutert, woran sich eine Vorstellung der zur Erfassung dieser Größen relevanten Erhebungsmethoden und –verfahren anschloss. Um eine Wissensgrundlage für die Auswahl eines spezifischen Erhebungsverfahrens zu schaffen, erfolgt danach eine Abschätzung der Eignung der Erhebungsverfahren hinsichtlich der Erfassung der dynamischen Objekte verschiedener Verkehrsszenarien auf Basis eines exemplarischen Szenarienkataloges.

Der fünfte Entwicklungsschritt umfasst die Durchführung der Beobachtung. Die Datenerhebung erfolgte im Rahmen vorheriger Aktivitäten. Für dieses Forschungsvorhaben wurde auf den Drohnendatensatz highD zurückgegriffen (Krajewski, Bock, Kloeker & Eckstein, 2018).

Der sechste Entwicklungsschritt betrifft die Auswertung und die Interpretation der Ergebnisse. Die Hochrechnung muss als Teil der Auswertung das Erfassungspotential der verschiedenen Formen der Erhebung berücksichtigen. Allen Erhebungsmethoden gemein ist, dass sie nur einen Teil der vorhandenen relevanten Szenarien erfassen können – also hinsichtlich ihres Erfassungspotentials limitiert sind. Daher wurde ein Hochrechnungsverfahren entwickelt, welches die Tatsache berücksichtigt, dass im Regelfall nur ein Teil eines Untersuchungseinzelelementes beobachtet werden kann und somit eine Extrapolation der Beobachtungen auf das gesamte Untersuchungseinzelelement notwendig ist. Um eine Aussage über die nicht beobachtete Raumzeit gewinnen zu können, müssen Annahmen über das Auftreten der relevanten Szenarien innerhalb eines Untersuchungseinzelelementes getroffen werden. Dazu wurde das Auftreten der Szenarien als Zufallsprozess modelliert. Dies erlaubt die Schätzung einer Wahrscheinlichkeit für das Erfassen eines Szenarios innerhalb eines Untersuchungseinzelelementes mittels einer spezifischen Form der Erhebung. Diese Erfassungswahrscheinlichkeit kann dann genutzt werden, um eine Extrapolation der Beobachtungen durchzuführen. Darüber hinaus erlaubt die Schätzung der Erfassungswahrscheinlichkeit in Kombination mit der Annahme über die Auftretenswahrscheinlichkeit eines Szenarios innerhalb eines Untersuchungseinzelelementes eine Schätzung der Beobachtungswahrscheinlichkeit seltener Ereignisse. Dieser Zusammenhang kann genutzt werden, um eine untere Schranke für den Stichprobenumfang – in diesem Fall die Anzahl der zu beobachtenden Untersuchungseinzelelemente – abzuschätzen. Für die weitere Untersuchung wurde das Hochrechnungsverfahren auf reale Verkehrsdaten angewendet und als Untersuchungseinheit das Fahrstreifenwechselszenario gewählt. Als Datenquelle diente der highD Drohnendatensatz, welcher Fahrzeugtrajektorien basierend auf einer naturalistischen Verkehrsbeobachtung beinhaltet. Des Weiteren wurden zwei verschiedene Definitionen des Fahrstreifenwechselszenarios zur Detektion verwendet, die auf unterschiedlichen charakteristischen Merkmalen basieren: der Lateralposition und der Lateralgeschwindigkeit. Zunächst wurde die Hypothese geprüft, dass die Verkehrsszenarien bzgl. ihrer örtlichen Lage gleichverteilt auftreten. Dazu wurden verschiedene statistische Anpassungstest verwendet. Im Anschluss erfolgte die eigentliche Hochrechnung zur Abschätzung einer Häufigkeitsdichte basierend auf einer Gewichtung mit den reziproken Erfassungswahrscheinlichkeiten.

3 Untersuchungsergebnisse

Die durchgeführten Untersuchungen zeigen, dass die Gleichverteilungsannahme bzgl. der örtlichen Lage innerhalb des Beobachtungsfensters der Drohne für einen Großteil der Fahrstreifenwechselszenarien basierend auf der Definition über die Lateralposition durch die statistischen Anpassungstests gestützt werden kann. Für Fahrstreifenwechselszenarien basierend auf einer Definition über die Lateralgeschwindigkeit kann eine leichte Abweichung von der Gleichverteilung an den Rändern des Beobachtungsfensters beobachtet werden.

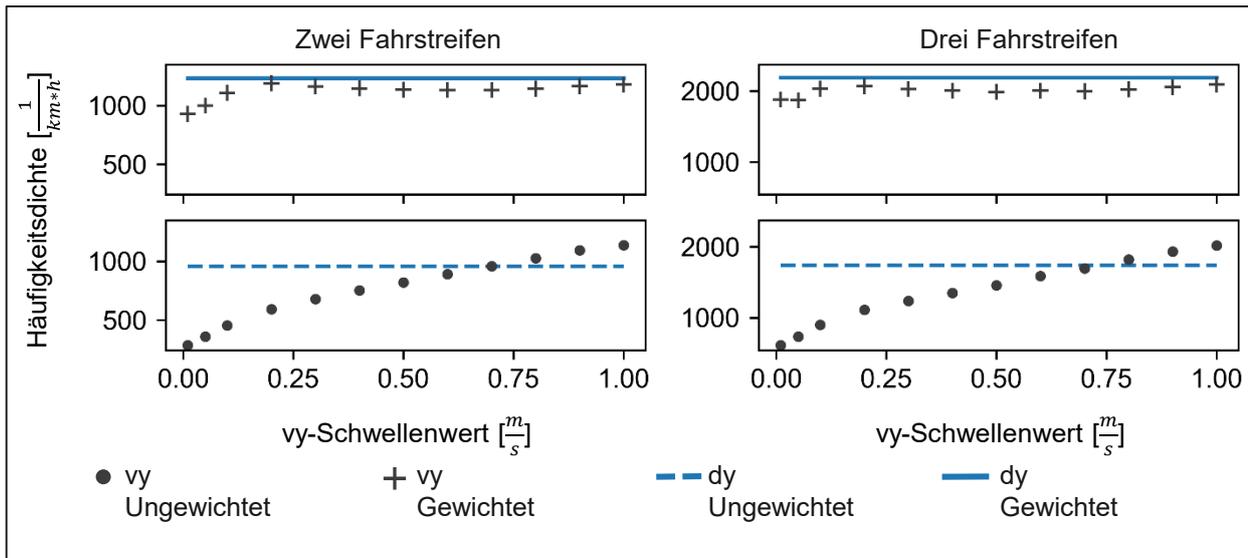


Bild 3-1: Einfluss des Lateralgeschwindigkeits-Schwellenwertes auf die mittlere Häufigkeitsdichte der Fahrstreifenwechselszenarien (ungewichtet und gewichtet)

Des Weiteren wurde der Einfluss der Gewichtung der Beobachtungen zur Abschätzung einer Häufigkeitsdichte untersucht. Dazu wurden Hochrechnungen für Fahrstreifenwechselszenarien mit verschiedenen Schwellenwerten für die Lateralgeschwindigkeit (v_y) durchgeführt und mit der Hochrechnung für Fahrstreifenwechselszenarien basierend auf der Lateralposition (dy) verglichen. Die mittleren Häufigkeitsdichten als Resultat dieser Hochrechnungen jeweils basierend auf einer Gewichtung mit ,1' (ungewichtet) und einer Gewichtung basierend auf den reziproken Erfassungswahrscheinlichkeiten (gewichtet) sind in Bild 3-1 dargestellt. Es wird deutlich, dass die Gewichtung basierend auf den reziproken Erfassungswahrscheinlichkeiten für verschiedene Schwellenwerte für die Lateralgeschwindigkeit zu ähnlichen mittleren Häufigkeitsdichten führen. Diese Beobachtung kann als Plausibilisierung der vorgestellten Methode für das betrachtete Fallbeispiel verstanden werden.

4 Folgerung für die Praxis

Für die Bewertung einer automatisierten Fahrfunktion spielt die Vollständigkeit mit der die Fahraufgabe mittels beobachteter Verkehrsszenarien beschrieben werden kann eine wichtige Rolle. Wie in diesem Forschungsvorhaben herausgearbeitet wurde, hängt die Beobachtungswahrscheinlichkeit eines Verkehrsszenarios innerhalb der Stichprobe neben der Erfassungswahrscheinlichkeit maßgeblich von der Auftretenswahrscheinlichkeit innerhalb eines Untersuchungselementes ab. Die Beschreibung der Fahraufgabe mittels beobachteter Verkehrsszenarien ist somit zwangsläufig aufgrund des Zufallscharakters des Auftretens der Verkehrsszenarien unvollständig. Aus praktischer Sicht ist es daher notwendig, den Grad der Vollständigkeit der Beschreibung der Fahraufgabe auf eine effiziente Weise zu erhöhen. Ein möglicher Ansatz kann darin bestehen, den durch die empirischen Daten aufgespannten Szenarienraum basierend auf zusätzlichem Wissen anzureichern. Dieser Ansatz wurde bereits im PEGASUS Projekt bei Einführung der logischen Szenarien verfolgt (PEGASUS, 2019).

5 Literaturverzeichnis

- Bäumer, M., Hautzinger, H., Pfeiffer, M., Stock, W., Lenz, B., Kuhnimhof, T. et al. (2017). *Fahrleistungserhebung 2014 - Inländerfahrleistung* (Berichte der Bundesanstalt für Straßenwesen V, Verkehrstechnik, Heft 290). Bremen: Fachverlag NW.
- Hautzinger, H., Pfeiffer, M. & Schmidt, J. (2012). *Entwicklung eines methodischen Rahmenkonzepts für Verhaltensbeobachtung im fließenden Verkehr*. Zugriff am 31.10.2022. Verfügbar unter: <https://bast.opus.hbz-nrw.de/opus45-bast/files/530/m227b.pdf>
- Krajewski, R., Bock, J., Kloeker, L. & Eckstein, L. (2018). *The highD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems*. Zugriff am 31.10.2022. Verfügbar unter: <https://arxiv.org/pdf/1810.05642>
- PEGASUS. (2019). *PEGASUS Method: An Overview*. Zugriff am 31.10.2022. Verfügbar unter: <https://www.pegasusprojekt.de/de/>
- Scholtes, M., Westhofen, L., Turner, L. R., Lotto, K., Schuldes, M., Weber, H. et al. (2021). 6-Layer Model for a Structured Description and Categorization of Urban Traffic and Environment. *IEEE Access*, 9, 59131–59147. <https://doi.org/10.1109/ACCESS.2021.3072739>
- Wachenfeld, W. & Winner, H. (2015). Die Freigabe des autonomen Fahrens. In M. Maurer (Hrsg.), *Autonomes Fahren. Technische, rechtliche und gesellschaftliche Aspekte* (Springer Open, S. 439–464). Berlin: Springer Vieweg. https://doi.org/10.1007/978-3-662-45854-9_21

Summary

BASSt research project FE 82.0735/2019 “Concept for statistically representative traffic observations”

1 Problem Definition

The safety validation and release process for highly automated driving systems is one of the major current challenges on the way to market introduction of these systems. In particular, the problem needs to be tackled that for the large number and variety of situations in road traffic a safe operation of these systems needs to be proven. The conventional method of driving a very large total distance with the active test system is economically and practically infeasible, since a total distance of several billion test kilometres must be driven according to estimates (Wachenfeld & Winner, 2015, p. 458).

Therefore, a reduction of the testing effort is required in order to enable the release process within a realistic period of time. For this purpose, scenarios are used, into which the entire driving task is divided. While the systematics for describing relevant traffic scenarios is currently being developed in various research projects, a concept is missing that defines their representative survey in specific traffic areas. However, representative scenarios are essential to evaluate the automated driving systems with regard to a positive risk balance.

The objective of this research project is the development of a general concept that allows a statistically representative survey of driving scenarios in various traffic areas. The dataset to be obtained in this way is intended to be used for the evaluation of automated driving functions. In this context, the following requirements can be formulated. A sample of scenarios derived by the sampling plan can be used to make statistically secure arguments about the actual frequency of occurrence of scenarios in the overall traffic or within the traffic area under consideration. The sampling plan considers different kind of techniques for traffic observation: Vehicle-based observations, infrastructure-based observations and aerial observations. The limitations associated with the type of survey should be taken into account – compare Figure 1-1.

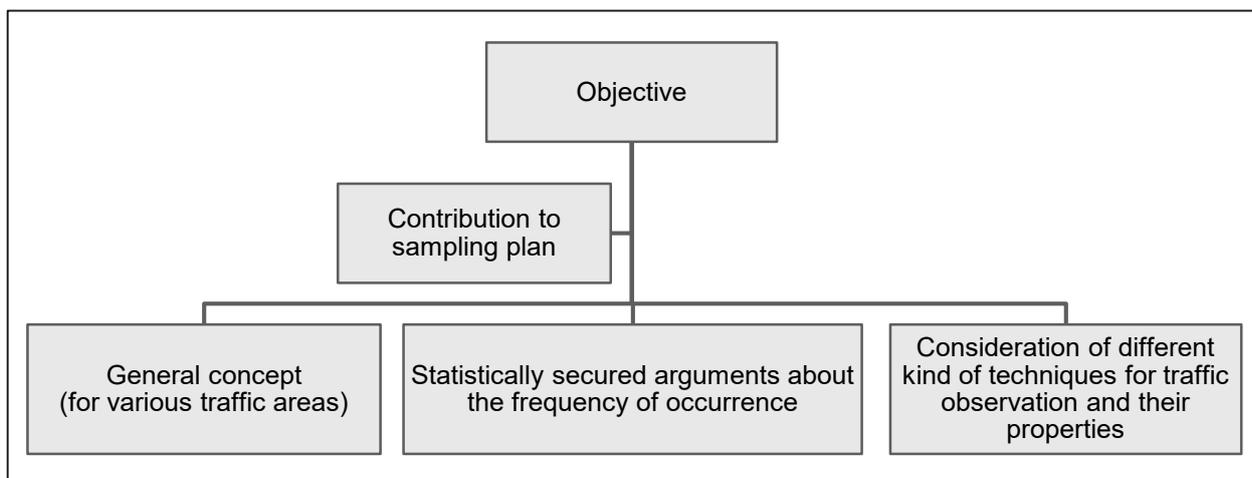


Figure 1-1: Research project objective

2 Methodical Approach

Starting point and orientation for the methodical approach in this research project is the methodological framework for observational surveys in moving traffic (Hautzinger, Pfeiffer & Schmidt, 2012). This framework proposes six development steps – compare Figure 2-1.

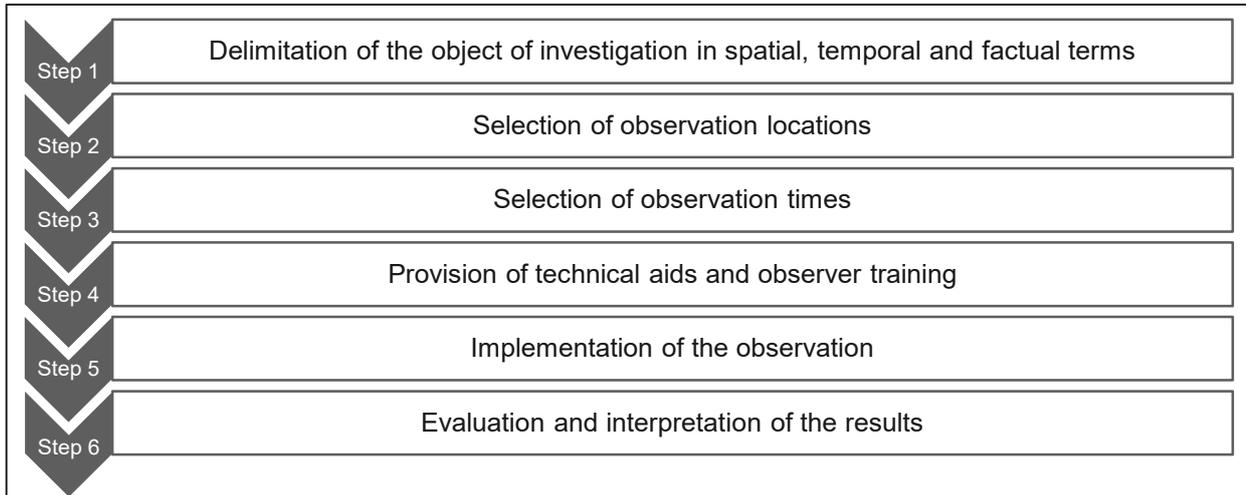


Figure 2-1: Development steps of the methodological framework for observational surveys in moving traffic – cf. (Hautzinger et al., 2012)

In the first development step the object of investigation is delimited in spatial, temporal and factual terms. Hence, a general model for the population of scenarios has been designed. The model takes into account the dependence of the population on context factors. Context factors represent static, macroscopic, or traffic-external variables. Examples of the first category are the road course in the form of inclines and curves, preceding accidents and traffic jams are factors of macroscopic type. The weather conditions and the weather-related road conditions also influence the traffic situation and thus the scenarios as external context factors. For a more differentiated structuring it is referred to the 6-Layer Model (Scholtes et al., 2021). Next to the dependence on context factors, the general model of the population requires a partitioning of the investigation area and period into investigation elements, in which the scenarios can be observed using different kind of techniques for traffic observation. The aim of the partitioning is to minimize the influence of the context factors on the characteristics and frequency of the scenarios within the investigation elements – compare (Bäumer et al., 2017) – and thus to enable a modelling as a random process independent of these context factors.

The second and third development step is related to the selection of observation locations and times. In the scope of this project, these aspects are considered in a generalized manner when selecting the observation units. In general, a distinction can be made between stationary observation units (e.g. observation using drones) and non-stationary observation units (e.g. observation using on-board vehicle sensors). In contrast to the stationary observation unit, a different section of the investigation area is observed in a non-stationary observation unit due to the moving observer at different points in time. When selecting the observation units, a location and time of observation must be defined for a stationary observation unit, while the movement of the observer must also be defined for a non-stationary observation unit. From the perspective of sampling theory, an unrestricted random sampling of the selection units is necessary in order to be able to make statistically secure arguments. Since the selection of the observation locations for stationary and non-stationary observation units is an essential aspect of the sampling plan, a focus was placed on the development of a method for the random selection of observation locations on federal motorways.

For this purpose, a continuous model of the road network was parameterized based on BISStra data using cubic splines, which allows a random selection of the observation location as well as an estimation of the auxiliary variable curve radius at the selected observation locations. As an example for further context factors, it has been investigated by means of a binning of the auxiliary variable curve radius, with which relative frequency a class of the auxiliary variable is contained in a random sample. The investigation revealed that, in addition to the quantization chosen for the binning, the sample size – in this case the number of observation locations – essentially determines the selection probability of certain classes of the auxiliary variable.

The fourth development step is related to the provision of technical aids and the observer training. Therefore, various survey techniques were characterized and compared. To this end, some relevant survey variables were first introduced, followed by a presentation of the survey methods and procedures relevant to the observation of these variables. In order to create a knowledge base for the selection of a specific survey procedure, a rough assessment of the suitability of the survey procedures with regard to the observation of the dynamic objects of different traffic scenarios has been performed based on an exemplary scenario catalogue.

The fifth development step involves the implementation of the observation. The data collection took place in previous activities. In this research project, the drone dataset highD has been used (Krajewski, Bock, Kloeker & Eckstein, 2018).

The sixth development step concerns the evaluation and interpretation of the results. As part of the evaluation, the extrapolation from the observed scenarios must take into account the potential of the various techniques of traffic observation. All survey techniques have in common, that they can only capture a part of the existing scenarios. For that reason, an extrapolation method has been developed that takes into account the fact that normally only a part of an investigation element can be observed and that an extrapolation of the observations to the entire investigation element is therefore necessary. In order to be able to make a statement about the unobserved space-time, assumptions about the occurrence of the scenarios within an investigation element must be made. For this purpose, the occurrence of the scenarios has been modelled as a random process. This allows to estimate a probability of observing a scenario within an investigation element using a specific survey technique. This observation probability can then be used for extrapolation. Moreover, the estimation of the observation probability in combination with an assumption about the probability of occurrence of a scenario within an investigation element allows to estimate the probability of observing rare events in a sample. This relationship can be used to estimate a lower bound for the sample size – in this case the number of investigation elements to be observed. For the further investigation, the extrapolation method was applied to real traffic data and the lane change scenario was selected as scenario to be investigated. For this purpose, the drone dataset highD has been used, which contains vehicle trajectories based on a naturalistic traffic observation using drones. Furthermore, two different definitions for the lane change scenario have been applied, which are based on different characteristic features: the lateral position and the lateral speed. In the first step of the investigation, the hypothesis has been tested, that the scenarios occur uniformly distributed with regard to their longitudinal position. To this end, various statistical tests have been used. In second step of the investigation, the extrapolation of the observation is performed by means of estimating the frequency density based on a weighting with the reciprocal observation probabilities.

3 Investigation Results

The conducted investigations show that the assumption regarding the uniformly distributed occurrence of the longitudinal position within the observation window of the drone can be supported by the performed goodness-of-fit tests for a large part of the lane change scenarios based on the definition of the lateral position. For lane change scenarios based on a lateral velocity definition, a slight deviation from the uniform distribution can be observed at the borders of the observation window.

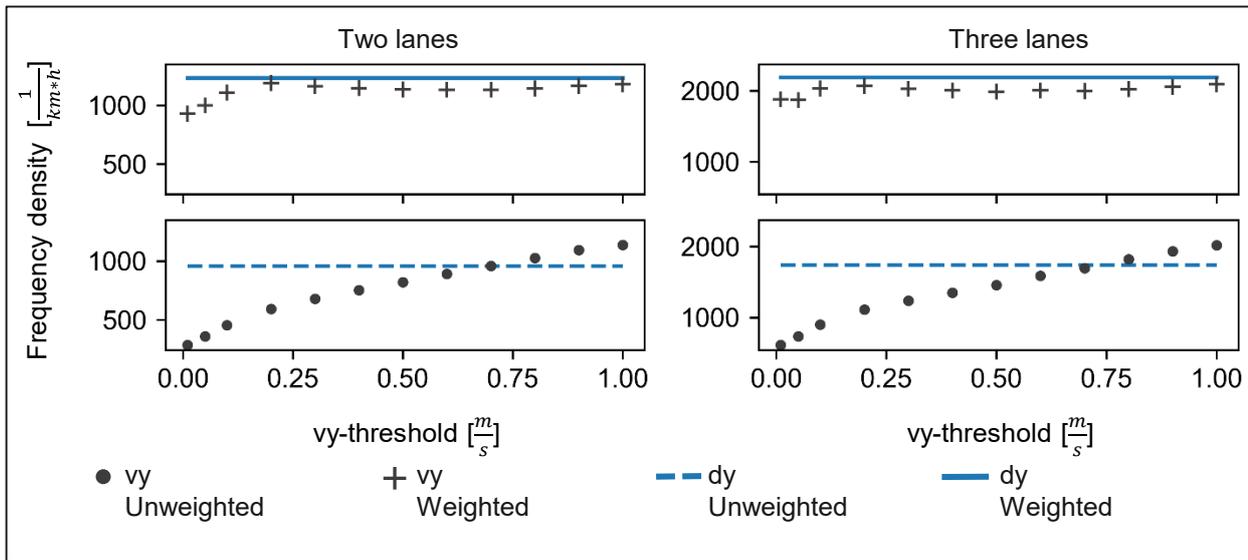


Figure 3-1: Influence of the lateral velocity threshold on the mean frequency density of the lane change scenarios (unweighted and weighted)

Furthermore, the influence of the weighting of the observations for the estimation of a frequency density was examined. For this purpose, extrapolations for lane change scenarios with different threshold values for the lateral velocity (v_y) were performed and compared with the extrapolation for lane change scenarios based on the lateral position (d_y). The mean frequency densities as a result of these extrapolations, each based on a weighting of '1' (unweighted) and a weighting based on the reciprocal observation probabilities (weighted) are shown in Figure 3-1. It becomes clear that the weighting based on the reciprocal observation probabilities for different threshold values of the lateral velocity results in similar mean frequency densities. This can be understood as a plausibility check of the presented method for the considered example.

4 Conclusion

The completeness with which the driving task can be described by the observed scenarios is important for the evaluation of an automated driving function. In this research project it could be shown, that the probability of observing a scenario within a random sample depends not only on the observation probability but also on the probability of occurrence within an investigation element. The description of the driving task by means of observed scenarios is therefore inevitably incomplete due to the random character of the occurrence of the scenarios. From a practical point of view, it is therefore necessary to increase the degree of completeness of the description of the driving task in an efficient way. A possible approach can be the enrichment of the scenario space spanned by the empirical data based on additional knowledge. This approach was already pursued in the PEGASUS project at the introduction of logical scenarios (PEGASUS, 2019).

5 Literature

- Bäumer, M., Hautzinger, H., Pfeiffer, M., Stock, W., Lenz, B., Kuhnimhof, T. G. et al. (2017). *Fahrleistungserhebung 2014 - Inlandsfahrleistung und Unfallrisiko* (Berichte der Bundesanstalt für Straßenwesen V, Verkehrstechnik, Heft 291). Bremen: Fachverlag NW.
- Hautzinger, H., Pfeiffer, M. & Schmidt, J. (2012). *Entwicklung eines methodischen Rahmenkonzepts für Verhaltensbeobachtung im fließenden Verkehr*. Access at 31.10.2022. Available under: <https://bast.opus.hbz-nrw.de/opus45-bast/files/530/m227b.pdf>
- Krajewski, R., Bock, J., Kloeker, L. & Eckstein, L. (2018). *The highD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems*. Access at 31.10.2022. Available under: <https://arxiv.org/pdf/1810.05642>
- PEGASUS. (2019). *PEGASUS Method: An Overview*. Access at 31.10.2022. Available under: <https://www.pegasusprojekt.de/de/>
- Scholtes, M., Westhofen, L., Turner, L. R., Lotto, K., Schuldes, M., Weber, H. et al. (2021). 6-Layer Model for a Structured Description and Categorization of Urban Traffic and Environment. *IEEE Access*, 9, 59131–59147. <https://doi.org/10.1109/ACCESS.2021.3072739>
- Wachenfeld, W. & Winner, H. (2015). Die Freigabe des autonomen Fahrens. In M. Maurer (Hrsg.), *Autonomes Fahren. Technische, rechtliche und gesellschaftliche Aspekte* (Springer Open, S. 439–464). Berlin: Springer Vieweg. https://doi.org/10.1007/978-3-662-45854-9_21