Technische Universität München
TUM School of Engineering and Design

Development of Driver Behavior Models for Simulation-Based Safety Assessment of Automated Vehicles in Urban Traffic

Michael Alexander Harth

Vollständiger Abdruck der von der TUM School of Engineering and Design der Technischen Universität München zur Erlangung des akademischen Grades eines *Doktors der Ingenieurwissenschaften (Dr.-Ing.)* genehmigten Dissertation

Vorsitz: Prof. Dr. phil. Klaus Bengler

Prüfer der Dissertation:

1. Prof. Dr.-Ing. Klaus Bogenberger
2. Prof. Dr.-Ing. Wolfram Remlinger

Die Dissertation wurde am 22.01.2025 bei der Technischen Universität München eingereicht und durch die *TUM School of Engineering and Design* am 11.06.2025 angenommen.

Acknowledgement

First and foremost, I would like to express my gratitude to my supervisor Prof. Dr. Klaus Bogenberger for giving me the opportunity to be an external doctoral candidate at his chair at UniBw and TUM. I was always grateful for his advice, which helped me move forward every time we met. I highly appreciate his always friendly attitude and the very pleasant interaction during the entire dissertation time.

Next, I want to thank Prof. Dr. Wolfram Remlinger, who serves as the second supervisor of this dissertation. I am very grateful to Wolfram for always being available as my mentor for intensive technical discussions that helped me a lot during my whole time at AUDI AG. Even beyond the technical discussions, the relationship with Wolfram has always been exceptionally respectful and friendly.

I would like to express my sincere thanks to Dr. Ronald Kates, who has been an immense support throughout the whole dissertation. It is indescribable to have had someone with such technical and linguistic competence as a contact person. I thank Ron very much for always being available when I needed his advice.

My special thanks go to Lutz Morich, who gave me the opportunity to start my dissertation time in a wonderful team and in an inspiring location at AUDI AG. Lutz has always been a valuable advisor for me. I will always have fond memories of the time I was able to spend at the research project SAVe:. My further thanks go to the other PhD candidates in my team: Benedikt Schwab, Christoph Stadler, Dr. Francesco Montanari and especially Marcel Langer, who has been an incredibly valuable sparring partner for so many discussions (while we drank half a cup of coffee) and a very good friend. I would like to thank everyone involved in the SAVe: research project for the very pleasant and good cooperation.

Moreover, I want to thank all the other PhD students at UniBw and TUM with whom I greatly appreciated the technical as well as informal exchange, especially in the research group I had the opportunity to be part of. Furthermore, I would like to thank all the students who supported this dissertation with a thesis or an internship.

Especially during the hardest time, when I had to finish my dissertation in addition to my daily work, my two bosses at CARIAD SE were a very strong support. Therefore, my sincere thanks go to Stefan Menken and Dr. Michael Dobmann, who were always patient and gave me the time I needed to finish my dissertation.

Finally, I am so grateful for the support my family and friends have given me over all the years. My warmest thanks go to my mother Dr. Ursula Harth, who has been the best proofreader I can imagine. In the same way, I thank my father Dr. Uwe Harth and my sister Dr. Julia Harth for their endless moral support whenever I needed it. Lastly, I cannot thank my beloved partner Leonie Roß enough for her never-ending support and patience. Thank you so much!

Abstract

The great complexity of traffic poses a major challenge for the introduction of automated vehicles. Manufacturers must perform extensive tests to prove that an automated vehicle covers all driving scenarios within the scope of the system. Since such tests cannot be carried out in reality at a reasonable cost, virtual tests offer a way out. However, the validity of testing places high demands on the quality of the static and dynamic virtual environment. In order to be able to test automated vehicles validly in a traffic system, both the global and local traffic conditions and the individual driving behavior of human drivers must be modeled precisely. However, driver models currently used in traffic simulation cannot represent driver behavior in sufficient detail.

In this thesis, a data-driven longitudinal driver behavior model is developed for the virtual testing of automated vehicles in an urban traffic system. Since virtual tests have a special focus on exceptional traffic situations, the model must be able to generate such critical situations. These situations are often caused by human failure, which results from human error-prone information perception and processing. Therefore, the developed driver model takes into account psychological processes of the human driver and includes human factors such as reaction time and anticipation. The architecture of the model is divided into perception, information processing and action. For the reproduction of the individual behavior of different driver types, the model is trained with separate data sets of different driver groups.

In order to be able to use the driver behavior model in a virtual test with automated vehicles, the traffic system of the German city of Ingolstadt is modeled as a test environment using the traffic simulation software SUMO. For this purpose, the traffic flows and traffic lights are simulated based on real data. Both automated vehicles and the developed driver model are integrated into the traffic model. In an exemplary test, the automated vehicles are each equipped with an algorithm. The algorithm is used to recognize the different driver types and to adapt the driving behavior when certain types are recognized. In the last part of the thesis, the influence of different recognition methods on traffic safety and efficiency is analyzed.

Kurzfassung

Eine große Herausforderung für die Einführung automatisierter Fahrzeuge stellt die große Komplexität des Straßenverkehrs dar. Hersteller müssen durch umfangreiche Tests nachweisen, dass ein automatisiertes Fahrzeug alle im Systemumfang liegenden Fahrszenarien abdeckt. Da derartige Tests in der Realität nicht mit vertretbarem Aufwand durchführbar sind, bieten virtuelle Tests einen Ausweg. Allerdings stellt eine valide Absicherung hohe Anforderungen an die Qualität der statischen und dynamischen virtuellen Umgebung. Um automatisierte Fahrzeuge valide in einem Verkehrssystem testen zu können, müssen sowohl die globalen und lokalen Verkehrszustände sowie das individuelle Fahrverhalten menschlicher Fahrer präzise modelliert werden. Derzeit in der Verkehrssimulation zum Einsatz kommende Fahrermodelle können das Fahrerverhalten jedoch nicht detailliert genug abbilden.

Im Rahmen dieser Arbeit wird ein datengetriebenes longitudinales Fahrerverhaltensmodell entwickelt, das für den virtuellen Test automatisierter Fahrzeuge in einem urbanen Verkehrssystem dienen soll. Da in virtuellen Tests ein besonderer Fokus auf außergewöhnlichen Verkehrssituationen liegt, muss das Modell in der Lage sein, solche meist kritischen Situationen zu erzeugen. Diese Situationen sind häufig auf menschliches Versagen zurückzuführen, das aus der menschlichen fehlerbehafteten Informationsaufnahme und -verarbeitung resultiert. Daher berücksichtigt das entwickelte Fahrermodell psychologische Prozesse des menschlichen Fahrers und schließt menschliche Faktoren wie Reaktionszeit und Antizipation mit ein. Die Architektur des Modells gliedert sich in die Wahrnehmung, die Informationsverarbeitung und die Handlung. Um das individuelle Verhalten verschiedener Fahrertypen abbilden zu können, wird das Modell mit separaten Datensätzen verschiedener Fahrergruppen trainiert.

Um das Modell in einem virtuellen Test mit automatisierten Fahrzeugen einsetzen zu können, wird als Testumgebung das Verkehrssystem der deutschen Stadt Ingolstadt mit der Software SUMO modelliert. Dafür werden die Verkehrsströme sowie die Ampelschaltungen auf Basis realer Daten nachgebildet. In dieses Verkehrsmodell werden sowohl automatisierte Fahrzeuge als auch das entwickelte Fahrermodell integriert. In einem beispielhaften Test werden die automatisierten Fahrzeuge jeweils mit einem Algorithmus ausgestattet. Der Algorithmus dient dazu, die verschiedenen Fahrertypen wiederzuerkennen und bei Erkennung bestimmter Typen das Fahrverhalten anzupassen. Im letzten Teil der Arbeit wird der Einfluss verschiedener Erkennungsverfahren auf die Verkehrssicherheit und -effizienz analysiert.

Contents

1. Introduction

Automated driving represents a transformative shift in the future of transportation, promising enhanced safety, increased mobility, and reduced traffic congestion. As the technology rapidly evolves, the need for comprehensive evaluation frameworks becomes increasingly critical to ensure that these systems can operate safely and effectively in complex, real-world environments. As part of this evaluation framework, simulative evaluation places high demands on the validity of human driver behavior models. This thesis investigates a driver behavior model which includes significant characteristics of real human drivers, called Human Factors (HFs). The developed driver model is embedded into a virtual traffic system in order to control human-driven vehicles in the surrounding traffic in a virtual test of automated vehicles.

1.1 Research Context

The vision of accident-free and efficient traffic due to the introduction of automated vehicles has fascinated both academia and industry as well as the general public for decades. The current advances of Artificial Intelligence (AI) methods, e.g., ChatGPT [OPENAI, 2023], which are attracting extensive media attention, suggest that a takeover of vehicle driving by AI-based algorithms is possible in the near future. Nevertheless, the market-ready developments in the field of automated driving in recent years show that excluding the driver from the monitoring and driving task (temporarily or even completely) represents a major challenge. Currently available assistance systems are primarily aimed at use within a limited functional scope in highway traffic. However, urban traffic, which is characterized by a large number of different road users, especially Vulnerable Road Users (VRU), as well as by crossing traffic, places much higher demands on the functional scope of an automated system [CAMPBELL et al., 2010; MOONEY AND JOHNSON, 2014].

An important motivation for the introduction of automated systems into traffic is the goal of improved traffic safety. In the past, the mandatory introduction of numerous driver assistance systems, e.g., ABS in 2004 and ESC in 2014 in the European Union (EU), has resulted in a significant increase in traffic safety. The number of traffic fatalities in the EU could be more than halved between 2001 and 2022 [EUROPEAN COMMISSION, 2017, 2022]. In Germany, the number of accidents and the number of people injured and killed in road traffic are stagnating, although the number of vehicles has increased steadily over the years [DEUTSCHES ZENTRUM FÜR LUFT- UND RAUMFAHRT E.V. (DLR) AND DEUTSCHES INSTITUT FÜR WIRTSCHAFTSFORSCHUNG (DIW), 2022; STATISTISCHES BUNDESAMT DESTATIS, 2023b].

Automated driving is expected to have a similar, if not greater, impact on traffic safety. Due to a permanent observation of the vehicle environment by sensors as well as the significantly lower delay between an observation and the subsequent reaction, there is the potential to detect critical situations early or even to avoid them entirely. However, several factors can also lead to the emergence of new types of accidents [GASSER et al., 2012]. For example, disturbed perception, e.g., due to weather conditions, or technical errors in the interpretation of a situation by an automated vehicle can lead to misjudgment and thus to unsafe behavior. Furthermore, people may overestimate the functional scope of vehicle automation, especially at low automation levels, and use systems in situations for which they were not designed. For both cases, prominent examples exist in the press that have led to accidents [SIDDIQUI AND MERRILL, 2023; WAKABAYASHI, 2018]. Such events can contribute to both a distorted public impression of the technology and a deterioration of manufacturers' reputations. However, these events show the public the system limits of vehicle automation, which in turn can lead to a more conscious usage.

In addition to the social challenges, economic and legal issues must also be clarified. One crucial question, for example, is who is liable in the event of damage during automated driving: the vehicle owner or the manufacturer. A major challenge is furthermore posed by technical issues relating to the development, testing and safeguarding of automated systems. Validation of driving functions purely on real roads is not a viable alternative in terms of time and cost due to the enormous testing effort involved [WACHENFELD AND WINNER, 2015]. Instead, a comprehensive and differentiated test concept consisting of real driving tests as well as virtual testing must be applied [BRUNNER et al., 2019; ECKSTEIN AND ZLOCKI, 2013]. However, this test concept places high demands on the validity of the simulation. It must be ensured that simulation and reality deliver equivalent results under the same boundary conditions. To meet this requirement, the virtual test environment as well as all road users involved must exhibit realistic behavior. In the near future, automated vehicles will operate in a traffic environment where they will have to interact with human-driven vehicles [WAGNER, 2015]. Therefore, modeling human driver behavior in the surrounding traffic plays an important role. In order to be able to take human characteristics into account in a model, the consecutive processes of perception, processing of information towards a decision must be realistically modeled. This is the only way to ensure that errors that occur at a certain step in the process chain that can potentially cause a critical situation can be recreated.

Within the steps of human cognitive processes, differences arise from person to person (inter-driver-specific), which can manifest in different types of reactions in the same initial situation. Furthermore, there are even differences for the same driver (intra-driver-specific) under the same dynamic boundary conditions, e.g., depending on the constitutional state. These characteristics of human drivers must be taken into account in a realistic model in order to be able to reproduce their effects on driver behavior and, in particular, errors in the cognitive and behavioral process chain.

1.2 Research Objective and Methodology

The main objective of this thesis is to develop a cognitive driver model for the longitudinal behavior of human drivers in urban traffic. The purpose of the development is to use the model in a virtual test in the surrounding traffic of automated vehicles in an urban traffic model.

Currently, the majority of virtual tests of automated driving systems are carried out using a scenario-based approach. Like the script of a movie, scenarios describe the initial situation and the sequence of actions of all traffic participants in a traffic situation [BAGSCHIK et al., 2018; SCHULDT, 2017]. These scenarios are created, for example, using expert knowledge or by reconstructing previously experienced, recorded situations. Scenarios can then be simulated in large numbers with variations in order to test the functionality of an automated driving system. In particular, the scenario-based test approach is used to test critical situations that pose a challenge for an automated vehicle. The scenario-based approach is contrasted with the concept of the virtual endurance run [KRAJZEWICZ et al., 2013; J. OLSTAM AND ELYASI-POUR, 2013; SEMRAU et al., 2016]. In a virtual endurance run, one or more automated vehicles are integrated into a virtual traffic system. In this type of simulation, automated vehicles are mainly exposed to everyday, non-hazardous traffic situations. However, critical situations that an automated vehicle has to cope with can also occur during a virtual endurance run. In contrast to the predefined situations in the scenario-based approach, these situations arise randomly from the traffic situation. In addition to testing automated driving systems in everyday traffic situations, the endurance run approach can lead to the discovery of new situations that pose a challenge for the systems.

However, the prerequisite for valid testing in a virtual traffic system is that all traffic participants which interact with automated vehicles are reflected in realistic models. Apart from VRUs, human-controlled vehicles in particular must behave realistically in the surrounding traffic of automated vehicles. In traffic simulation, established human driver models known as *"engineering"* models [SAIFUZZAMAN AND ZHENG, 2014] such as the Intelligent Driver Model (IDM) [TREIBER et al., 2000] and the Krauss model [KRAUß, 1998] idealize human driver behavior. These types of models assume perfect perception of ground truth, ideal immediate processing of information, and an accident-free implementation of the driver's decision. However, real human driver behavior is characterized by imperfect perception, delayed information processing, and non-ideal action execution that depends on various factors affecting the driver and his/her environment. In addition, human driver models should address the inter- and intra-driver-specific differences mentioned above.

From this identified research gap, the need for modeling a new driver model that takes HFs into account is derived. The following research questions arise for the modeling of human driving behavior:

RQ 1.1: What methods can be used to model human driver behavior considering HFs?

RQ 1.2: How can the individuality of individual and different drivers be captured and replicated in driver behavior modeling?

In the context of the two research questions, it is also investigated how the calibration and training process must be designed so that the developed model reproduces human driver behavior as realistically as possible. In addition, it is considered which traffic situations are suitable for investigating the influence of the explicit consideration of HFs.

For performance validation, the developed driver model is compared with existing established driver models. The reference for the comparison is human driver behavior recorded in reality. For the application of the driver model in a virtual endurance run, the modeling of a traffic system is required. The traffic model can either be synthetic or correspond to a real-world example. The advantage of a real traffic model lies in its comparability with tests carried out in reality (see Chapter 4). Such a traffic system corresponding to the real road network and traffic of the German city Ingolstadt is set up in this thesis. The developed driver behavior model is integrated into this traffic system and simulated in mixed traffic conditions with automated vehicles. The following research questions arise for the application of the model:

RQ 2.1: Can the developed driver behavior model, which takes into account HFs in its architecture, reproduce human driver behavior more realistically than existing driver behavior models?

RQ 2.2: Can the developed driver model reproduce the behavior of specific driver types so that the type can be recognized by an automated vehicle?

RQ 2.3: Does the adaptation of driving behavior of automated vehicles to specific recognized driver types impact traffic safety and efficiency?

1.3 Outline of the Thesis

In this thesis, an AI-based driver model for longitudinal driver behavior in urban traffic is developed. Figure 1.1 shows the structure of the thesis.

Chapter 2 describes the current state of the art on the topics of virtual testing of automated vehicles, human driver behavior and its modeling, and modeling of automated vehicles in simulation. Based on the state of the art, Section 2.6 outlines the research gap which directly corresponds to the research questions in Section 1.2. Chapter 3 describes the driver behavior model developed within this thesis. The chapter discusses the structure of the model, which includes HFs, the training procedure, and the performance analysis compared to existing models. Chapter 4 presents the traffic model of Ingolstadt used for the simulations performed in Chapter 5, a framework is created to integrate the developed longitudinal driver behavior model into the traffic model of Ingolstadt in order to investigate a potential use case for the driver behavior model. First, interactions between human drivers in the traffic model are analyzed to validate the developed driver model in comparison with real data. Then, mixed traffic with automated vehicles is simulated. Different risk mitigation strategies for the automated vehicles interacting with different human driver types are compared and evaluated. Finally, Chapter 6 discusses and summarizes the findings obtained in this thesis, answers the research questions and provides an outlook for future research fields.

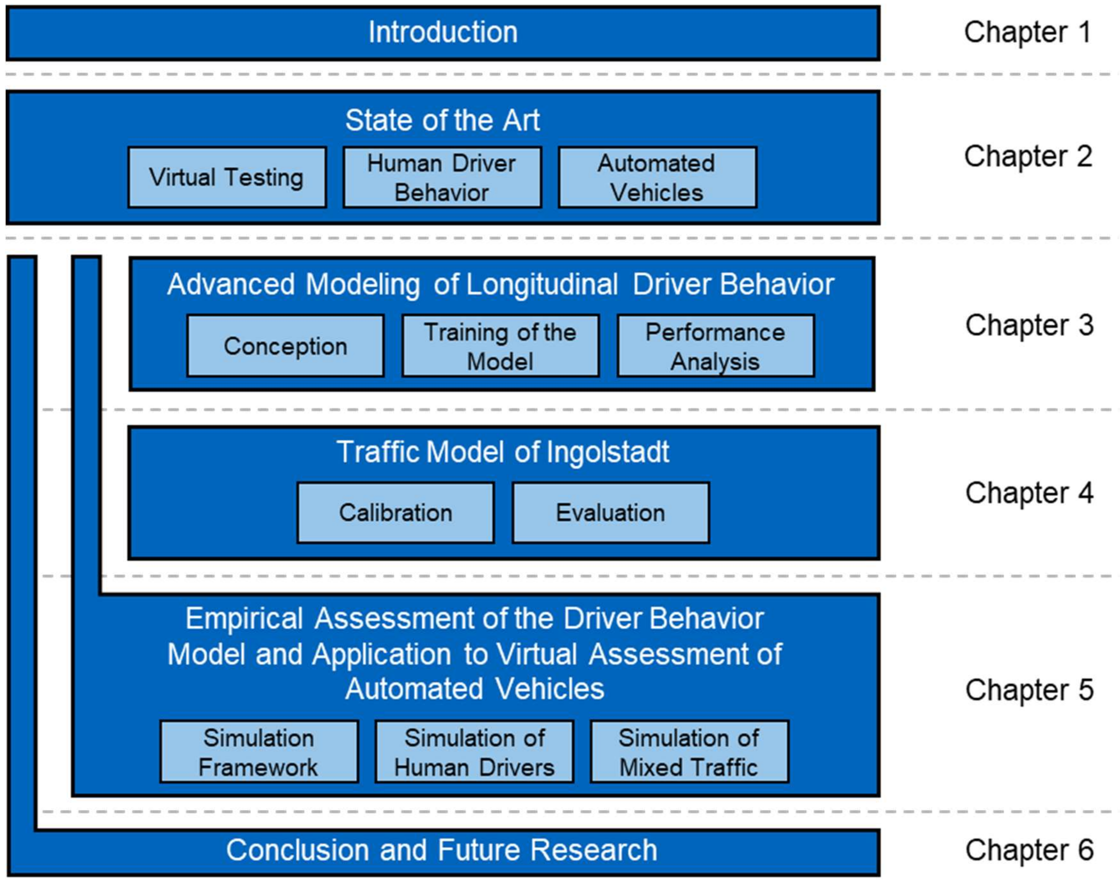

Figure 1.1: Structure of the thesis.

2. State of the Art

While research and development in vehicle technology has been significantly influenced by classic mechanical engineering for decades, an interdisciplinary approach is now required in the field of automated driving. The connectivity between traffic participants is leading to a new perspective that is moving away from the individual vehicle perspective and towards the consideration of entire traffic systems. With increasing digitalization of vehicle development and testing, information and simulation technology is further growing in importance. In virtual tests of automated vehicles, the representation of realistic traffic participants in the surrounding traffic is essential. For the modeling of human traffic participants, the consideration of human psychology is inevitable. Chapter 2 summarizes the current state of the art for testing automated vehicles and modeling human driver behavior.

Section 2.1 introduces the topic of virtual testing of automated vehicles. Based on Section 2.1, Section 2.2 identifies requirements regarding a valid simulation environment. Section 2.3 gives an overview of the psychological processes of a human driver. In Section 2.4, existing driver models for longitudinal driver behavior are presented. It also discusses how HFs can be incorporated into a driver model. Section 2.5 presents the state of the art regarding the modeling of automated vehicles in simulation. Finally, Section 2.6 outlines the research gap derived from the state of the art.

2.1 Testing and Validation of Automated Vehicles

The automation of road traffic holds the great potential of a sustainable improvement of traffic safety and efficiency. Equipped with sensor technology to monitor the entire vehicle environment as well as real-time situation analysis and response, automated vehicles can potentially avoid a large number of critical situations and accidents caused by human error [GASSER et al., 2012]. Furthermore, a forward-looking driving style and the combination of several vehicles into a group (platooning) lead to an improved driver experience from an ecological and comfort perspective as well as to an increased traffic flow.

However, the introduction of automated systems is accompanied by a considerable effort associated with testing and validation of driving functions during development. Based on the statistically determined distance of two fatal accidents and the annual driving performance of all drivers in Germany, WACHENFELD AND WINNER [2015] estimate the required scope of field tests for a comprehensive evaluation of the driving performance of automated vehicles. They conclude that a test distance of 2.1 billion km is required to verify the safety of an automated driving system. Since the verification process requires a partial or even complete repetition of the driven distance for each system update, operational tests in the field cannot be performed with a reasonable time and economic effort.

Research has therefore been directed towards deriving test procedures with optimal levels of abstraction for tests with different scopes. ECKSTEIN AND ZLOCKI [2013] derive test procedures in four levels of abstraction, which distinguish the included number of real elements, i.e. driver, vehicle and environment. The first level is based solely on simulation and does not contain any real-world elements. In contrast, the second and third level cover driving simulator tests as well as controlled field tests, e.g., on test tracks, incorporating either just a driver or a driver and a vehicle into the test procedure. The highest validity in this scheme for the evaluation of automated driving systems would be achieved by field operational tests or naturalistic driving studies combining all referred elements and therefore forming the last abstraction level.

In order to reduce effort, the majority of testing has to be shifted to lower abstraction levels regarding the model of ECKSTEIN AND ZLOCKI [2013]. Figure 2.1 shows different abstraction levels of testing of automated driving systems.

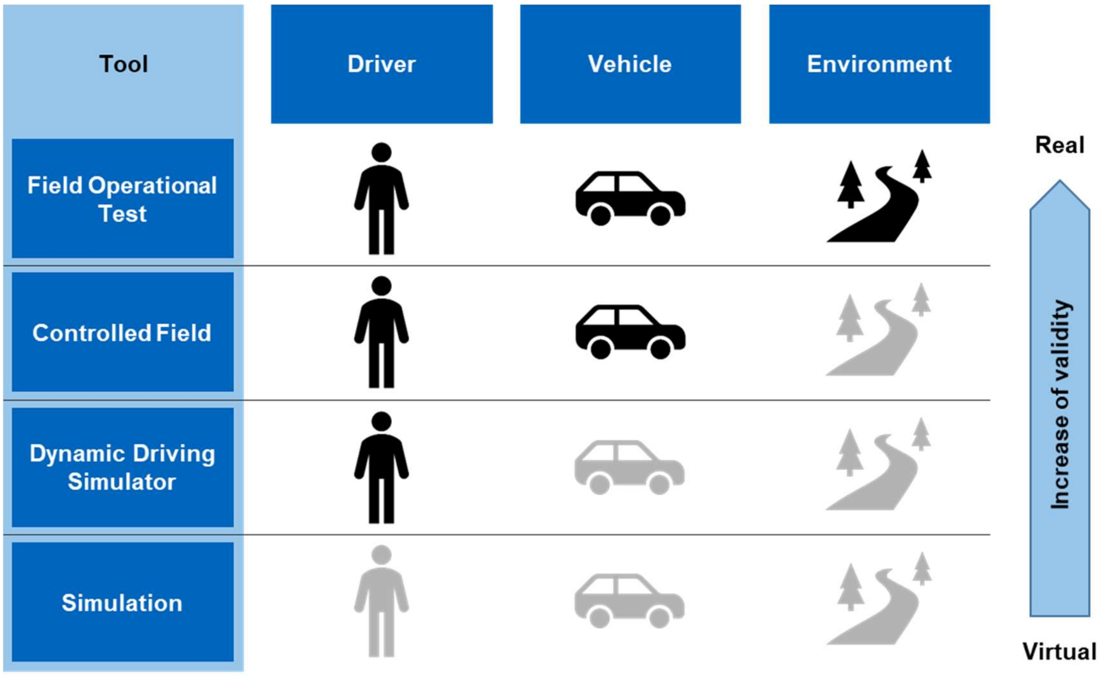

Figure 2.1: Abstraction levels of testing of automated driving systems [ECKSTEIN AND ZLOCKI, 2013].

Especially simulation-based approaches for the evaluation of automated driving systems provide a solution for efficient and replicable testing of a multitude of potentially dangerous traffic situations. The main challenge arising from evaluation of automated driving systems by simulation is to assess and prove the validity of the applied simulation models [SCHULDT, 2017].

2.2 Requirements Towards Simulation

For the development of automated driving systems, there are several available submicroscopic vehicle simulation tools that take the driving dynamics of all traffic participants into account and simulate specific vehicle components, e.g., sensors, as well as vehicle software, e.g., an automated driving function. These simulation tools are mostly applied using scenario-based methods [BAGSCHIK et al., 2018; SCHULDT, 2017]. A scenario fully defines a concrete situation in which a vehicle under test (VuT), e.g., an automated driving system, has to perform. These scenarios are simulated with varying parameter input for performance assessment of an automated driving system, especially in critical situations. The main groundwork required for scenario-based testing is the predefinition of important scenarios as well as the selection of the most relevant scenarios for a given system.

While certain scenarios require detailed examination in scenario-based simulation, it has been suggested to begin testing of an automated driving system in a realistic but randomized traffic environment. This approach provides the advantage of evaluating a system in everyday situations over a large timespan, thus laying the groundwork for further evaluation of critical situations in scenario-based tests. While submicroscopic simulation tools excel in providing an accurate representation of one system under test, they are not designed to provide a realistic traffic surrounding in an extensive road network [SEMRAU, 2018]. However, in a simulation-based evaluation of an automated driving system, realistically behaving traffic participants play a key role for valid simulation results. Therefore, there have been several efforts to combine submicroscopic vehicle simulation with microscopic traffic simulation to take advantage of the potentials of both simulation types [KRAJZEWICZ et al., 2013; J. OLSTAM AND ELYASI-POUR, 2013; SEMRAU et al., 2016]. Although microscopic traffic simulations focus on modeling the movement of all individual traffic participants, they have usually been implemented with the main purpose of investigating traffic flow. In an endurance run of an automated driving system, coupled traffic simulation is required to provide an environment that enables a detailed examination of the microscopic interaction between individual traffic participants. This is an entirely different focus and a new field of application for traffic simulation.

In order to provide a VuT with a realistic traffic environment in the context of a virtual test, the surrounding virtual human road users must also behave realistically. The underlying behavioral models must emulate essential human processes in the reception and processing of information from the environment as well as in decision-making, since these play an essential role in the broad spectrum of possible human behaviors. In particular, human error behavior and resulting critical situations are elementary for considering the traffic safety of automated driving. However, such modeling requires a comprehensive picture of the environment surrounding a road user and its relevant influences. These influences range from interaction with other road users and static influences of the road to buildings, vegetation and weather conditions. Since the consideration of these influences is beyond the scope of a microscopic traffic simulation, submicroscopic simulation is used for this purpose. There are efforts to enrich the virtual environment with a highly accurate database concerning the road space as well as the immediate surroundings, which provides the necessary information about the static environment [SCHWAB et al., 2020]. Furthermore, infrastructure data, such as real traffic light programs, can be included in the simulation. From this set of potential influences, the relevant ones have to be identified and provided to a behavioral model that implements a realistic response to a given situation [HARTH, AMJAD, et al., 2022].

2.3 The Human Driver

A valid virtual test of automated vehicles requires not only an exact representation of the static environment of a VuT but also, in particular, a realistic modeling of the movements of the surrounding road users. Especially, the interaction with the surrounding motorized, manually controlled individual traffic plays an important role. In order to realistically reproduce human driver behavior in the simulation, the process how a driver perceives a situation, derives and processes information, makes a decision and performs an action must be explicitly modeled. From a psychological perspective, the human driving task can be divided into three levels: Perception, information processing, and action. Various HFs, e.g., reaction time, can be assigned to the individual levels or to the overall process. In the following, the levels, their psychological modeling as well as related HFs are explained in more detail.

2.3.1 Perception

While driving, the driver is exposed to a large number of stimuli that are relevant for the execution of the driving task. The majority of the information necessary for the driving task is perceived via the visual sense [ROCKWELL, 1972]. Therefore, the eyes represent the major source for the reception of information. In addition to perceiving spatial depth and size, the eyes also serve to perceive color, objects, and motion. When receiving information, the eyes perform three major tasks. To perform the first task, adaptation, the sensitivity of the eye is adjusted to the prevailing luminance. This is done by decreasing the pupil diameter in strong light conditions and in turn increasing it in weak light conditions. The second task of the eyes is called accommodation. It describes the process of adjusting the focal length of the lens of the eye to allow sharp vision at different visual distances. The third essential task is called fixation. Fixation describes the alignment of the eyes with a visual object so that both visual axes are convergent. In the eye, fixation of an object is performed at the point of sharpest vision on the retina, called the fovea centralis. A sufficiently sharp image results only within a cone of vision with an aperture angle of approx. 2° - 3° [BUBB et al., 2015].

In addition to the visual reception of information, acoustic, haptic, kinesthetic, and vestibular perception also play a role in vehicle guidance. Essential characteristics of acoustic information reception are adaptation, auditory pattern recognition, which is mainly applied in speech comprehension, and acoustic spatial orientation, which is enabled by binaural hearing. Adaptation is the adjustment of the auditory system to different noise levels. Acoustic stimuli relevant for the driver reach him/her both from the vehicle environment, e.g., from sirens of an emergency vehicle, and from the vehicle interior, e.g., in the form of a warning pre-crash system. For the haptic reception of information, receptors in and under the skin are used, which enable a sensation of pressure, touch and vibration. The kinesthetic perception system senses movements of joints and stretches of muscles of the driver. Accelerations and speeds can be detected via the vestibular perception system. Finally, the organ of equilibrium located in the inner ear, as part of the vestibular system, provides information for maintaining balance as well as orientation in space [SCHMIDT AND LANG, 2013].

The perception of stimuli is limited by the perception thresholds. There are two different types of perception thresholds. The minimum intensity of a stimulus that is still detectable for humans is called absolute threshold. In contrast, the differential perceptual threshold defines the just noticeable difference (JND) of the intensity of two stimuli that can be detected by humans as a deviation. According to WEBER's law, this minimal difference is proportional to the original stimulus [BAIRD AND NOMA, 1978; SCHMIDT AND LANG, 2013; WEBER, 1834]. Figure 2.2 visualizes the described dependency.

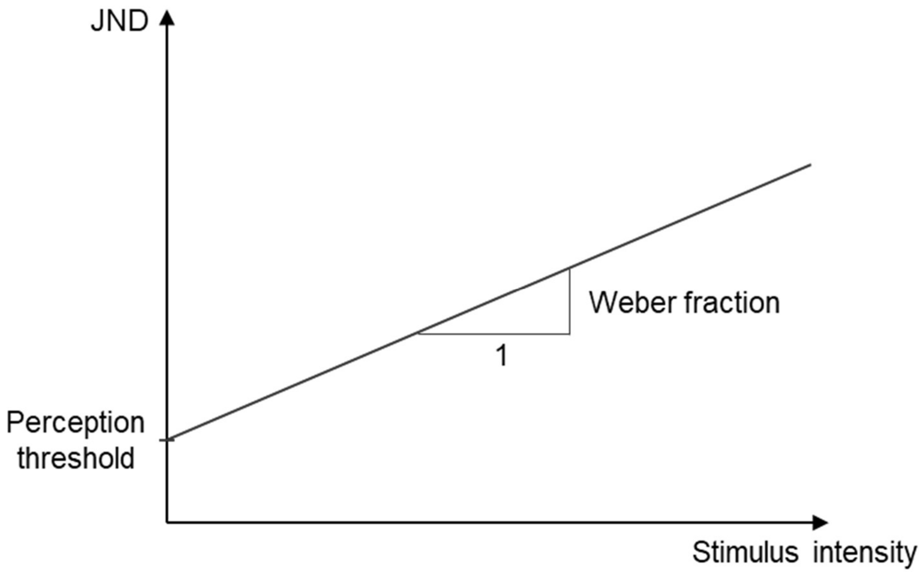

Figure 2.2: WEBER's Law according to WEBER [1834]. Figure according to NASH et al. [2016].

In addition to perceptual thresholds, the limited ability to estimate distances, (relative) velocities, and accelerations represents another potential source of error in the subsequent processing and evaluation of information [CALDERONE AND KAISER, 1989; LOOMIS et al., 1992; RUNESON, 1974; WATAMANIUK AND HEINEN, 2003]. The human driver compensates for this imperfect perception by using learned behavior and anticipation at the information processing level discussed in the following Section 2.3.2.

2.3.2 Information Processing

The information received at the perception level must be interpreted and evaluated by the driver in order to subsequently make a decision for an action. The processing of visual information is a two-stage process. The preattentive phase is characterized by the initial detection of edges, movements, and the orientation of structures without the driver consciously directing attention to the perceived objects. Subsequently, the stimuli perceived in the preattentive phase are transmitted by the eyes via the nerves to the brain, where they are interpreted. This interpretation process in the attentive, i.e. consciously attentive, phase of information processing additionally uses the driver's knowledge and previous experience to be able to grasp and evaluate a prevailing situation.

The SEEV model describes the extent to which visual attention is directed towards specific areas of interest, e.g., a front vehicle or a display in the interior. According to the SEEV model, the factors influencing visual attention can be divided into four areas [WICKENS et al., 2015]. Figure 2.3 shows an overview of the model.

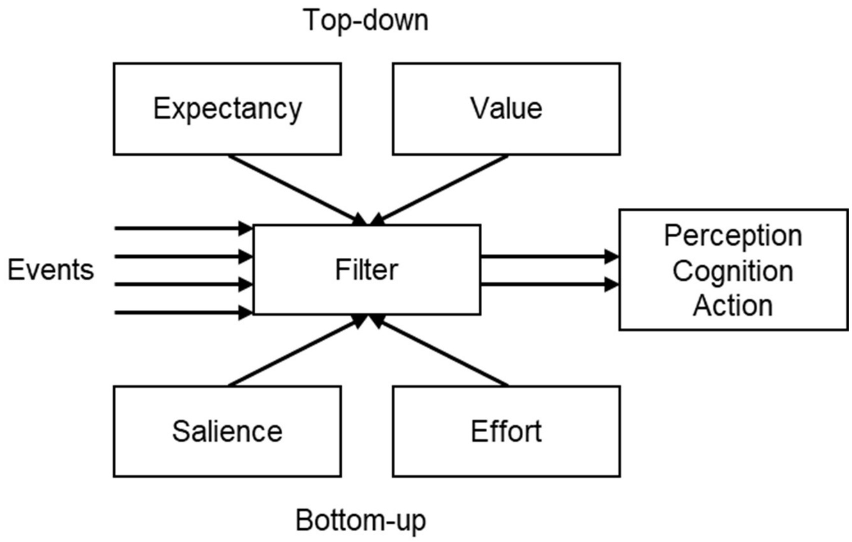

Figure 2.3: SEEV model according to WICKENS et al. [2015].

With the help of the four influencing factors *salience*, *effort*, *expectation*, and *value*, the driver accesses the stimuli he/she receives in terms of their relevance. The allocation of attention depends proportionally on each of the four influencing factors. *Salience* describes the extent to which a received stimulus, e.g., illuminated brake lights, stands out from the surrounding stimuli and thus attracts attention. Attention attraction involves a certain amount of *effort*, which is expressed by a movement of the eyes, the head, or the upper body. WICKENS et al. [2015] describe a higher reluctance to visually cover distances, if a rotation of the head or upper body is required. A driver's *expectation* determines what the driver perceives or overlooks and how he/she behaves when decision time is short [SCHLAG AND HEGER, 2002]. In addition, the *expectation* of an impending event has a significant influence on the corresponding reaction time [HINRICHS AND KRAINZ, 1970]. Finally, the *value* of a signal determines its relevance to the driver. For example, a vehicle ahead that is decelerating sharply attracts more attention than a pedestrian walking on the side of the road. While *salience* and *effort* depend on the physical properties of the environment (bottom-up), *expectancy* and *value* are each influenced by knowledge-based effects (top-down).

The attention available to a driver can be conceived as a limited cognitive resource that can be selectively allocated to different reference points. As soon as a task requires a higher level of attention, the adequate execution of parallel tasks is no longer ensured due to limited cognitive load capacity [KAHNEMAN, 1973]. In particular, distraction by tasks unrelated to the actual driving task can lead to driving errors that can potentially cause critical situations [YOUNG AND REGAN, 2007].

The conscious execution of the driving task requires the driver to have a sufficient understanding of the current driving situation. According to ENDSLEY [1995], situational awareness (SA) can be divided into three levels. At the first level, the driver perceives relevant elements and their characteristics within the dynamic environment. By merging all information received at the first level, the driver gains an understanding of the driving situation. Based on this understanding, the relevance of objects and events in the situation is evaluated on the second level. The first two levels then serve as the basis for predicting the future spatial and temporal evolution of the situation at the third level [ENDSLEY, 1995].

2.3.3 Execution of Action

From the information recorded and processed by the driver, he/she makes decisions about the future driving state. These decisions are executed at the level of action via motor movements of the hand-arm system and the foot-leg system. The physical workload of the driver is comparatively low compared to the cognitive workload caused by the processes of information acquisition and processing. In addition, there is a large number of technical assistance systems, e.g., power steering and brake boosters, which support the driver in performing actions.

The process from the reception of information to processing and execution is not immediate, but takes a certain amount of time, the Perception-Reaction-Time (PRT). According to GREEN [2000], the PRT can be divided into a sequence of three successive temporal sequences.

1. Mental processing time describes the time span from the perception of a signal to the decision. Furthermore, it can be divided into the following three time periods:

    a. Sensation defines the amount of time that elapses before the driver has detected an object on the road. PRT decreases with higher signal intensity, a signal in the foveal field of view, and better visibility conditions.

b. Perception is the time in which the meaning of perceived objects is recognized. Surprising signals, uncertainties in recognition, and response to multiple signals contribute to an increase in PRT.

c. Response selection and programming is the time that elapses from a possible decision of a reaction to its mental planning and design. The time increases when there is a choice between several possible responses.

2. Movement time is the time in which the driver performs the muscle movement required for the reaction. While complex movements increase movement time, increased arousal level and extensive practical driving experience decrease it.

3. The device response time describes the time between the driver's reaction and the corresponding response by the vehicle.

Figure 2.4 illustrates the composition of the PRT.

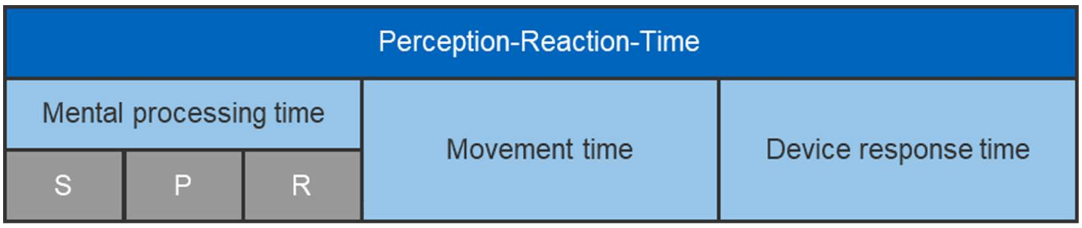

Figure 2.4: Composition of PRT; Labels: S: Sensation, P: Perception, R: Response selection and programming. For the purpose of simplicity, the time spans are chosen to be of equal duration in the figure, but in reality they can be of different durations.

As described, unexpected, surprising stimuli contribute to an increase in PRT. In contrast, expected, anticipated stimuli decrease the duration of the PRT. In expected situations, anticipation helps to (partially) compensate for reaction time. Only when a surprising event occurs, the actual PRT emerges, so that potentially more critical situations can result [Green, 2000].

2.3.4 Individuality of Human Drivers

The characteristics of HFs, such as PRT duration, not only vary from driver to driver (inter-driver), but also for a single driver (intra-driver). A distinction is made between two types of influences on these characteristics: individual differences and situational factors [Hamdar, 2012].

Individual differences represent permanent attributes of the driver and vehicle. Drivers are characterized by demographic factors, such as age and gender, socioeconomic factors, such as wealth and education, and psychological characteristics, such as aggressiveness, impatience, and risk-taking. Vehicle characteristics, such as vehicle size and power, also contribute to a driver's individual driving style. Situational factors represent circumstances that are variable in time and space. A distinction is made between factors related to the individual driver and factors related to the vehicle environment. Situation-related individual factors describe temporary distractions, impairments due to stress, fatigue, or alcohol, and characteristics of the trip, such as the trip purpose and duration. The environment of the vehicle is influenced by time of day and weather conditions such as the position of the sun, rain or slippery roads. Furthermore, the characteristics of the surrounding traffic system, such as road type, number of lanes, speed limit, and traffic volume play a significant role as influences on human drivers and their behavior [HAMDAR, 2012].

The individual influences on the driver cannot be considered separately, but are in part mutually dependent. In their entirety, they determine the mental load of a driver, which can be considered a limited capacity [KAHNEMAN, 1973]. The capacity depends, among other things, on the situation-related individual attributes of the driver. For example, fatigue, low arousal, and alcohol consumption have a detrimental effect on mental load capacity. Reaching or exceeding the limit of this capacity has the effect of overtaxing the driver, which, if the primary driving task is neglected, can lead to misjudgment or overlooking another object [BROOKHUIS AND DE WAARD, 2010; PRECHT et al., 2017].

Parts of the individual and situational influences on the driver cannot be measured or can only be measured with considerable effort, depending on the data collection method. For example, in studies, the driver and his eye movements can be recorded to gain knowledge about his condition as well as gaze control, e.g., [BROOKHUIS AND DE WAARD, 2010; PALINKO et al., 2010]. Collecting data representative of a driver population that captures both the permanent individual characteristics of the driver and vehicle as well as the situational characteristics of the driver and vehicle environment requires a high effort. However, even if the internal factors remain hidden due to the way of data collection, their manifestations can be captured in the form of physically measurable driver behavior. For example, drivers show an individual acceleration and speed profile. Similarities and differences can be identified between different drivers when comparing driver behavior. Drivers with similar characteristics can therefore be classified into specific groups. Depending on the situational factors, it is also possible to classify a single driver into different groups.

CONSTANTINESCU et al. [2010] group drivers into five different categories of aggressiveness from "non-aggressive" to "very aggressive" based on GPS traces. To do this, they determine parameters that characterize the driving behavior from the trajectories and perform a Principal Component Analysis (PCA) in combination with a hierarchical clustering procedure to group the drivers. However, no parameters are used to draw meaningful conclusions about interactions with other road users, e.g., time headway. JOHNSON AND TRIVEDI [2011] use information recorded via smartphone sensors on accelerations, position, driver gestures, and maneuver recognition to distinguish between non-aggressive and aggressive driver behavior.

DÖRR et al. [2014] use fuzzy logic to classify driver behavior into three clusters based on parameters describing the driving state: comfortable, normal, and sporty. The validation of the algorithm is done simulatively by detecting the respective driver type of a vehicle, which is calibrated by a driver model with a parameter set specific to each driver type. Finally, CHEN et al. [2020] use a k-means clustering procedure to classify a driver population into three clusters timid, normal, and aggressive based on parameters such as maximum acceleration and average time gap obtained for each driver from a variety of trajectory data.

2.4 Existing Models of Human Driver Behavior

In order to validate an automated vehicle in a virtual test in a realistic environment, the driver-vehicle units must exhibit realistic behavior in the surrounding traffic. For this purpose, it is necessary that the underlying driver models take into account the human characteristics described in Section 2.3 with regard to individually different perception, information processing as well as action execution. These three steps are also found in the three-level model according to RASMUSSEN [1983] which is related to the three-level hierarchy according to DONGES [1982] to model the human driving task. Figure 2.5 shows the combination of both models. The elements associated with perception are highlighted in light blue, those associated with information processing in dark blue, and those associated with action execution in orange.

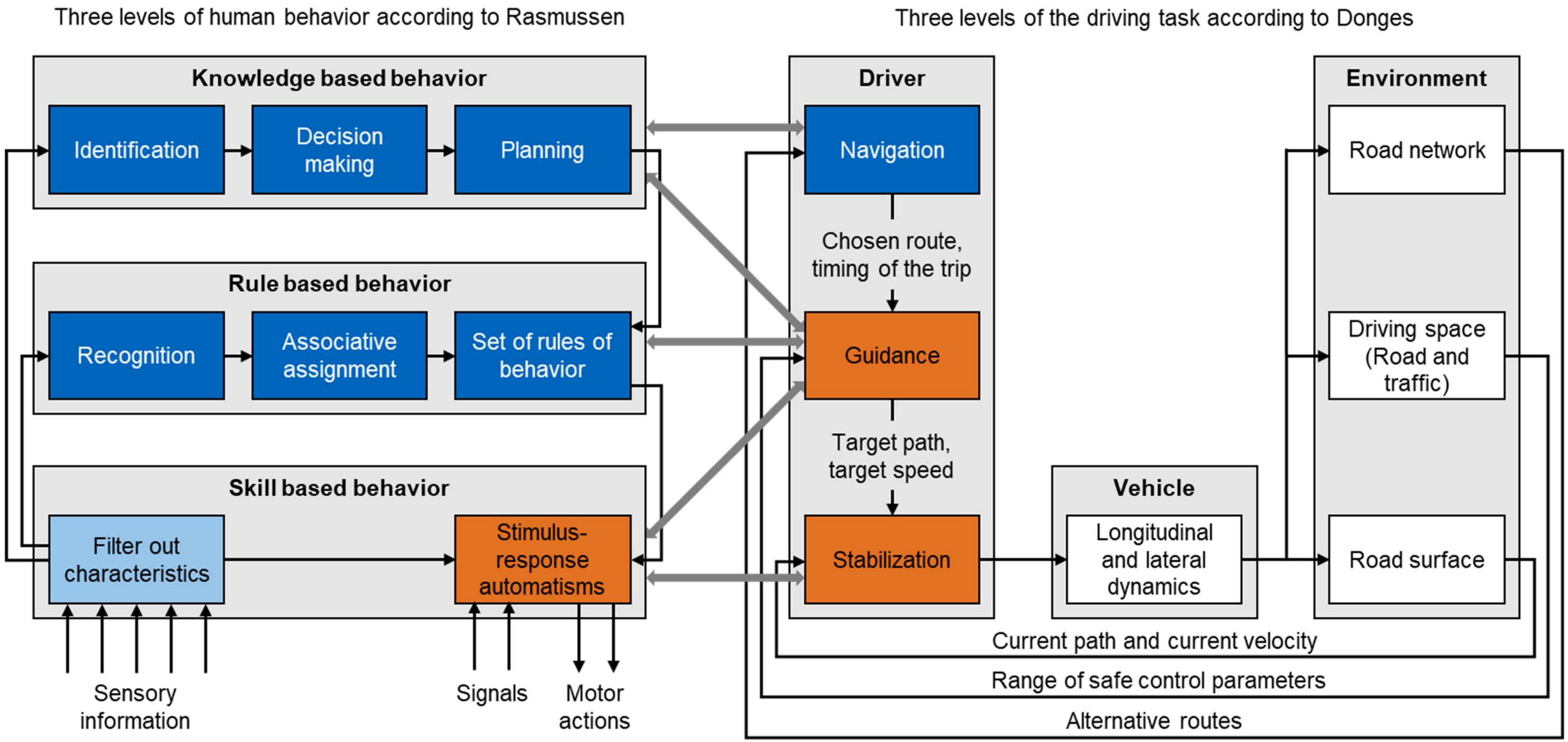

Figure 2.5: Three-level model for human behavior according to RASMUSSEN [1983] in combination with the three-level model of the driving task according to DONGES [1982].

Depending on the intensity of the cognitive demand, RASMUSSEN distinguishes three levels for goal-directed activities of humans (left side of Figure 2.5):

1. Knowledge-based behavior comes into play in complex situations that catch people unprepared and require previously untrained behavior. On the basis of their previous knowledge, people must evaluate the situation and weigh up alternative actions in order to cope with the situation. New action types can also expand the person's capabilities through a learning effect.

2. In known situations that have already occurred frequently, humans use rule-based behavior. After recognizing the situation, the human falls back on an alternative action that is known and proven to him.

3. The skill-based level is characterized by the execution of trained actions without conscious control by the human being. The skill-based behavior thus corresponds to routine action sequences.

The three levels of behavior according to RASMUSSEN interact with the three levels of the driving task according to DONGES. DONGES differentiates the driving task (right side of Figure 2.5) as follows:

1. At the navigation level, the driver makes a decision on a suitable route and estimates the expected travel time. Due to traffic-related disruptions, it may be necessary to change the choice of route during the journey. The navigation level can be assigned to different behavior levels, depending on whether the driver is moving in a traffic area that is familiar or unfamiliar to him/her. The execution of navigation in an unknown traffic space requires a conscious planning of the route and thus a knowledge-based behavior. In a familiar traffic area, no activities are required of the driver. The navigation task is already fulfilled before the start of the journey.

2. The guidance level comprises the adjustment of target variables such as target speed and target course as well as anticipatory measures to keep future deviations from the optimal course as low as possible. Depending on the experience of the driver and the criticality of a given situation, the guidance level can be assigned to one or more behavioral levels. Inexperienced drivers, on the one hand, perform parts of the driving task at the knowledge-based level and develop behavioral rules and the ability to perform unconscious actions only with increasing routine. In the case of experienced drivers, on the other hand, skill-based behavior dominates in non-critical situations. Only unexpected situations with increasing criticality demand rule-based or even knowledge-based behavior from the experienced driver.

3. Finally, the driver corrects deviations of the actual variables from the target variables at the stabilization level. The driver resorts to skill-based behavior for stabilization.

The stabilizing deviations detected by the driver are converted into longitudinal and lateral dynamic motion via the vehicle's actuators. The response behavior of the vehicle and the environment is fed back to the driver in a closed control loop for continuous adjustment of the target variables.

The simulatively usable implementation of the abstract modeling of the human driving task according to RASMUSSEN and DONGES requires a mathematically or computer-technically realizable implementation of the described human processes. The implementation of widespread driver behavior models, cf. [KRAUß, 1998; TREIBER et al., 2000; WIEDEMANN, 1974], often aims at modeling the processes within the guidance level. At the same time, navigation is decoupled from the model and stabilization is neglected due to the assumption of an ideal conversion of target to actual variables in the normal range of driving dynamics. In the following, existing models are presented that describe human longitudinal driver behavior, especially in car-following situations. Since the longitudinal behavior is in the foreground in this thesis, lateral dynamic models will not be discussed here. Implementations as well as overviews of common lane change models can be found in [BEN-AKIVA et al., 2006; GIPPS, 1986; KESTING et al., 2007; TOLEDO et al., 2003].

2.4.1 Longitudinal Driver Behavior Models

Longitudinal dynamic driver behavior can mainly be divided into two types: free driving without a leading vehicle and car-following behind a leading vehicle. Models describing longitudinal dynamic behavior differ conceptually mainly for the car-following, while the free driving results in a drive at the driver's desired speed. The beginnings of modeling longitudinal dynamic driver behavior go back to REUSCHEL [1950] and PIPES [1953], both in the 1950s. Since then, a large number of driver models have been developed that reproduce human behavior using various approaches.

According to OLSTAM AND TAPANI [2004], car-following models can be classified according to their internal logic. The class of *Gazis-Herman-Rothery* (GHR) models [GAZIS et al., 1961] describes the movement of a following vehicle as a direct response to a movement of the leading vehicle. There is a proportional relationship of the speed of the following vehicle, the relative speed, and the distance to the leading vehicle to the acceleration of the following vehicle. A well-known representative of the GHR models is the Intelligent Driver Model (IDM) [TREIBER et al., 2000]. It describes the acceleration a_n of the following vehicle by

$$a_n = a_{max,n}\left[1 - \left(\frac{v_n}{v_{des,n}}\right)^{\delta} - \left(\frac{s^*(v_n, \Delta v_n)}{s_n}\right)^2\right], \tag{1}$$

where $a_{max,n}$ is the maximum desired acceleration, v_n is the current speed, $v_{des,n}$ is the desired speed, Δv_n is the relative speed with respect to a leading vehicle, s_n is the distance to the leading vehicle, and δ is a tuning factor chosen by default to be 4. The parameter s^* denotes the desired gap and is determined by

$$s^*(v_n, \Delta v_n) = s_{0,n} + T_n v_n + \frac{v_n \Delta v_n}{2\sqrt{a_{max,n}\, b_{des,n}}}. \tag{2}$$

In (2), $b_{des,n}$ describes the maximum desired deceleration, $s_{0,n}$ denotes the standstill gap, and T_n is the desired time gap.

The *safety distance* or *collision avoidance* class of models always maintains a safe distance from the leading vehicle to ensure collision-free driving. Some of the models maintain a sufficiently large gap so that even severe deceleration of the leading vehicle does not lead to a collision. Well-known representatives of this model class are the Gipps model [GIPPS, 1981] and the Krauss model [KRAUß, 1998]. In contrast to IDM, the Krauss model directly considers velocity instead of acceleration. The desired velocity $v_{des,n}$ at time step t is derived by a comparison of three different velocity values according to

$$v_{des,n}(t) = \min\left(v_{max,n},\ v_n(t) + a_{max,n}(v_n)\Delta t,\ v_{safe,n}(t)\right). \tag{3}$$

v_n is the current velocity, $v_{max,n}$ states the driver's maximum desired velocity, $v_{safe,n}$ is the maximum safe velocity, $a_{max,n}$ the maximum acceleration and Δt the time step size. The current value of the maximum safe velocity $v_{safe,n}$ is calculated as

$$v_{safe,n}(t) = v_{n-1}(t) + \frac{s_n(t) - v_{n-1}(t)\,\tau_{k,n}}{\frac{v_n(t) + v_{n-1}(t)}{2b_{des,n}} + \tau_{k,n}}, \tag{4}$$

where v_{n-1} is the velocity of the leading vehicle, s_n the current gap, $\tau_{k,n}$ the reaction time and $b_{des,n}$ the desired maximum deceleration. $\tau_{k,n}$ influences the desired time headway. Finally, the velocity for the next time step t + Δt is given by

$$v_n(t + \Delta t) = \max\left(0, v_{des,n}(t) - \eta\right), \tag{5}$$

where η is introduced as a random perturbation factor ($\eta > 0$) in order to reproduce imperfect driving. The Krauss model is the default car-following model in the microscopic simulation software SUMO [LOPEZ et al., 2018].

Optimal velocity models (OVM) [BANDO et al., 1998] form another class of models that determine an optimal safe velocity for the following vehicle, which depends on the distance to the vehicle ahead. The acceleration resulting from the model a_n is determined from the difference between the current speed and the optimal speed. One model used in this work from the group of OVMs is a modified full velocity difference (FVD) model [ZHANG et al., 2019], which is designed for use at signalized intersections. In addition to the distance and relative speed to a leading vehicle using the parameter θ, the traffic signal is also taken into account for determining the desired acceleration.

$$\theta = \begin{cases} 0, & \text{the traffic light signal is red.} \\ 1, & \text{the traffic light signal is green.} \end{cases} \tag{6}$$

Depending on θ, the driver's desired acceleration a_n is defined by

$$a_n = \begin{cases} \kappa_r \left(V_r(\Delta x_n - p_n) - v_n\right) + \lambda_r\, \Delta v_n + \xi & \text{if } \theta = 0, \\ \kappa_g \left(V_g(\Delta x_n - q_n) - v_n\right) + \lambda_g\, \Delta v_n + \xi & \text{if } \theta = 1, \end{cases} \tag{7}$$

in which κ_r, κ_g, λ_r, and λ_g are sensitivity parameters that are dependent on red (r) or green (g) light, respectively. Δv_n describes the relative velocity and ξ is a randomly chosen perturbation term. The parameter p_n is calculated by $\eta \cdot v_n$, where η is a number greater than 0. In addition, q_n ensures smaller distances to the leading vehicle during green light. q_n is determined from ω / v_n, where ω is a number greater than 0. V_r and V_g are optimal velocity functions, both of which are determined using the following equation:

$$V(\Delta x_n) = V_1 + V_2 \tanh(C_1(\Delta x_n - l) - C_2), \tag{8}$$

in which the optimal velocity is determined dependent on the traffic light signal based on the distance to the front vehicle Δx_n, the constants V_1, V_2, C_1, and C_2, and the vehicle length l.

In contrast to the models presented so far, *psycho-physical* or *action point* models take into account perceptual thresholds or action points of the human being in order to change driver behavior. According to this model class, drivers can only react to changes in distance or relative speed to the leading vehicle as soon as they exceed a threshold value. Well-known examples of psycho-physical driver models are the models of WIEDEMANN [1974] and FRITZSCHE [1994].

Beyond the presented approaches for modeling longitudinal driver behavior, *data-driven methods* have been increasingly developed in recent years. In particular, recurrent neural networks (RNNs) have been shown to process temporal sequences of input data. It has further been shown that RNNs are capable of reproducing asymmetric human behavior, such as hysteresis, intensity differences between acceleration and deceleration, and discrete acceleration and deceleration behavior [HUANG et al., 2018]. MORTON et al. [2017] compare two Long Short-Term Memory (LSTM) networks, as representatives of RNNs, with two feedforward Gaussian mixture networks as well as with the IDM model in order to assess the performance in reproducing driver behavior. The training/calibration of each of the five models is based on the NGSIM data set for Interstate 80 [HALKIAS AND COLYAR, 2006]. Both the feedforward models and the LSTM models consist of two hidden layers with 128 nodes each. As a result, LSTM networks are able to replicate characteristics of the human driving behavior while feedforward networks show equal qualitative results. However, the feedforward networks show a deficiency in detecting and following trends in driver behavior, e.g., in acceleration. Furthermore, LSTM produces smoother trajectories. Especially for the similarity of generated and real car-following behavior, the machine learning based models outperform IDM [MORTON et al., 2017]. In addition to the sources mentioned above, there are several others that use LSTM to model human driving behavior, e.g., [FAN et al., 2019; WANG et al., 2018; ZHOU et al., 2017].

Table 2.1 summarizes the introduced model classes, their characteristics and the given model examples:

Model class	Characteristics	Examples
Gazis-Herman-Rothery (GHR) models	Stimulus-response behavior	GHR model [GAZIS et al., 1961] IDM [TREIBER et al., 2000]
Safety distance or collision avoidance models	Ensure collision-free driving by always keeping a safe distance to the leading vehicle	[GIPPS, 1981] [KRAUß, 1998]
Optimal velocity model (OVM)	Determination of an optimal safe velocity to safely follow the leading vehicle	OVM [BANDO et al., 1998] Modified FVD [ZHANG et al., 2019]
Psycho-physical or action point models	Driver changes his/her behavior with respect to thresholds or action points	[WIEDEMANN, 1974] [FRITZSCHE, 1994]
Data-driven models (RNN)	Consideration of temporal sequences for decision making Ability to imitate human driving characteristics	LSTM and feedforward Gaussian mixture networks [MORTON et al., 2017] LSTM [HUANG et al., 2018].

Table 2.1: Overview of longitudinal driver model classes, their main characteristics and associated examples.

2.4.2 Human Factors in Driver Behavior Models

Classical, so-called "engineering" [SAIFUZZAMAN AND ZHENG, 2014], driver behavior models, which are primarily used in traffic flow modeling, represent human behavior in an idealized manner. The models use *current* ground truth inputs such as distance and relative speed to the leading vehicle to determine *current* acceleration or speed. However, important HFs related to the levels of processing information as well as the individuality of human drivers (cf. Section 2.3), such as reaction time and anticipation, are not or insufficiently considered. Therefore, there are efforts in current research to extend engineering driver behavior models with HFs.

TREIBER et al. [2006] add a Human Driver Model (HDM) to the IDM that accounts for inaccurate perception as well as reaction time and temporal and spatial anticipation across multiple vehicles. The model's erroneous perception is modeled using a Wiener process. The model estimates the motion of its own vehicle by assuming constant acceleration over a prediction horizon. In contrast, movements of vehicles ahead in their own lane are anticipated using a constant velocity model. Considering only PRT without anticipatory effects, the IDM shows unstable behavior in the traffic system. However, the destabilizing effects caused by the introduction of PRT can be counterbalanced with the help of temporal and spatial anticipation [TREIBER et al., 2006]. LINDORFER et al. [2018] additionally extend the HDM by variable reaction times depending on the driving state (car-following, free driving, stationary), by inattention due to driver distraction, and by driving errors and irregularities.

VAN LINT AND CALVERT [2018] implement the psychological concepts of dynamic situational awareness (SA) [ENDSLEY, 1995] and of driver task demand (TD) [KAHNEMAN, 1973; PRECHT et al., 2017] (cf. Section 2.3.2) into a variant of the IDM [SCHAKEL et al., 2012]. TD describes the superposition of loads acting on the driver due to the exhibition of the driving task as well as secondary tasks. VAN LINT AND CALVERT [2018] determine SA of the modeled driver based on current TD. SA, in turn, has a significant influence on the HFs PRT, perceptual errors, as well as the nature of the driver's response. By varying the maximum task capacity, i.e., the load capacity of each driver, a heterogeneous driver population with different skill levels can be modeled [VAN LINT AND CALVERT, 2018]. In a more advanced work, CALVERT et al. [2020] extend the TD and SA enriched model with the concept of anticipation reliance (AR). AR enables drivers to cope with a higher cognitive load by anticipating the dynamic vehicle environment. CALVERT et al. do not implement anticipation by explicitly estimating the movements of all surrounding vehicles. Instead, AR reduces TD in the model, resulting in a positive effect on HFs.

CHEN et al. [2020] use the IDM and the Gipps model [GIPPS, 1981] and draw on the concept of long- and short-term driver-specific characteristics. Long-term characteristics describe characteristics intrinsically associated with the driver type, such as an aggressive driving style. Short-term characteristics, in contrast, are due to external stimuli. After temporally limited changes in driving behavior, the driver returns to his/her long-term behavior. The long- and short-term influences are implemented as additional parameters in IDM and the Gipps model. Using a k-means clustering, a population of drivers selected from trajectory data from the NGSIM I-80 data set [HALKIAS AND COLYAR, 2006] is divided into the three driver types timid, normal, and aggressive based on parameters such as acceleration and time gap. Based on the separation, the model parameters of the two studied models are calibrated for each driver type [CHEN et al., 2020].

According to the so-called τ-theory by LEE [1976] the coordination of visual-motor tasks is based on changes of projections on the human retina. Thus, human drivers can perceive changes in the longitudinal motion of a preceding object in the driver's field of view as angular changes of projections on the retina. Figure 2.6 shows the model of the human eye moving at a time t with a velocity v(t) exclusively along the z-axis.

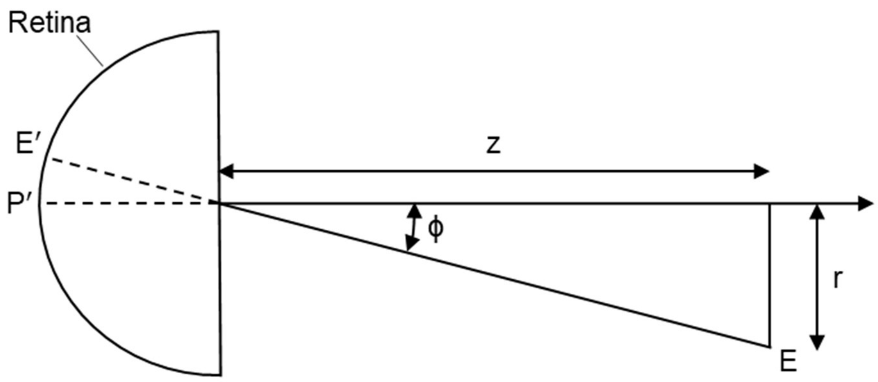

Figure 2.6: Projection of a longitudinally moving object E on the human retina.

The angular deviation ϕ of the object E can be expressed by the ratio of distance z and lateral constant deviation r for small angles:

$$\frac{z(t)}{r} = \frac{1}{\phi(t)}. \tag{9}$$

The derivative of equation (9) with respect to the time leads to equation (10), in which v is the visually perceived relative velocity and v_ϕ the angular velocity of the angle ϕ.

$$-\frac{v(t)}{r} = \frac{v_\phi(t)}{\phi(t)^2} \tag{10}$$

Merging (9) and (10) yields the Time-To-Collision (TTC) as the optical variable τ. TTC defines the time remaining until the collision.

$$TTC(t) = -\frac{z(t)}{v(t)} = \frac{\phi(t)}{v_\phi(t)} = \tau(t) \tag{11}$$

The driver therefore perceives the angular deviation and its rate of change in relation to a surface element and uses this to determine the time remaining before a collision with it. The driver uses this information to initiate a deceleration and prevent a collision. In addition to τ, the driver perceives the time derivative of τ, which is given by equation (12), in which the variable a represents the relative acceleration.

$$\dot{\tau}(t) = -1 + \frac{z(t) \cdot a(t)}{v(t)^2} \tag{12}$$

The driver uses $\dot{\tau}$ to regulate the deceleration so that he/she avoids a collision as smoothly and as controlled as possible. A constant value of $\dot{\tau}$ = -0.5 implies that the driver brakes with a constant deceleration so that he/she comes to a stop directly behind a stationary vehicle or a vehicle that is also braking constantly. A collision can therefore only be avoided if the driver regulates the deceleration in such a way that a value of $\dot{\tau} \geq -0.5$ is set. Situations in which $\dot{\tau}$ has values smaller than -0.5 consequently result in a collision. In a collision situation without relative acceleration between the involved vehicles, a value of $\dot{\tau} = -1$ would result.

The findings of the $\dot{\tau}$-theory are taken up by MAI et al. [2015] and are used in a cognitive driver model. In addition, further refinements are made by YILMAZ AND WARREN [1995] in the optimal control of $\dot{\tau}$ into the modeling of longitudinal driver behavior.

OSSEN et al. [2006] compare different engineering models with respect to their ability to model the driver behavior of different driver types. In order to model different driver types, it is not sufficient to adapt the parameter set of a single model to the respective driver type. Instead, it is suggested to use different models instead of a single model in order to better cover the broad spectrum of different driver types. In this way, the optimal model can be selected individually for each driver type and its parameter set can be calibrated.

In addition to engineering models, HFs are also being added to data-driven models. As described earlier, RNNs are able to reproduce human characteristics while driving [HUANG et al., 2018] (cf. Section 2.4.1). MA AND QU [2020] use a sequence-to-sequence approach based on LSTM to perform anticipation of subsequent behavior over a time horizon. The model is shown to outperform a simple LSTM as well as an IDM in replicating real-world trajectories. Furthermore, human heterogeneous driving behavior as well as hysteresis in the following behavior are shown in the simulation of the model.

2.4.3 Summary of Driver Behavior Models

Table 2.2 summarizes all mentioned driver behavior models categorized by the driver model type – engineering models and data-driven models – and the inclusion of HFs. The column “Remarks” contains additional information regarding each model and, if available, a list of the considered HFs. All models of each category are chronologically ordered. As data-driven models implicitly learn HFs like imperfect driving, all data-driven approaches have been sorted into the group that considers HFs.

	Engineering models		Data-driven models	
	Model/Authors	Remarks	Model/Authors	Remarks
Without consideration of HFs	[REUSCHEL, 1950]	First car-following models		
	[PIPES, 1953]			
	GHR model [GAZIS et al., 1961]	Direct stimulus-response behavior		
	[GIPPS, 1981]	Collision avoidance model that keeps a safe gap		
	[KRAUß, 1998]	Collision avoidance model; standard model in SUMO		
	OVM [BANDO et al., 1998]	Optimal velocity model		
	IDM [TREIBER et al., 2000]	GHR model; widely used in science		
	Modified FVD model [ZHANG et al., 2019]	OVM designed for use at signalized intersections		
With consideration of HFs	[WIEDEMANN, 1974]	Psycho-physical models with perception thresholds	[MORTON et al., 2017]	LSTM models with short-term memory
	[LEE, 1976]	Psycho-physical model based on perceived TTC	[ZHOU et al., 2017]	
	[FRITZSCHE, 1994]	Like [WIEDEMANN, 1974]	[WANG et al., 2018]	
	[YILMAZ AND WARREN, 1995]	Refinement of model by [LEE, 1976]	[HUANG et al., 2018]	LSTM model replicating human driving characteristics
	HDM [TREIBER et al., 2000]	Extended IDM with disturbed perception, anticipation, PRT	[FAN et al., 2019]	Autoencoder and LSTM models with short-term memory
	[MAI et al., 2015]	Psycho-physical model based on [LEE, 1976]	[MA AND QU, 2020]	Sequence-to-sequence LSTM model with anticipation
	[LINDORFER et al., 2018]	Extended HDM with variable PRT, inattention and driving errors	[HARTH, ALI, et al., 2021]	CNN-LSTM model with short-term memory, PRT
	[VAN LINT AND CALVERT, 2018] and [CALVERT et al., 2020]	Extended IDM with SA,TD, AR, PRT, anticipation	[HARTH, AMJAD, et al., 2022]	LSTM model with short-term memory, disturbed perception, PRT, anticipation, individual driver characteristics
	[CHEN et al., 2020]	Extended IDM with driver-specific characteristics		

Table 2.2: Summary of engineering and data-driven driver behavior models with and without HFs.

2.4.4 Data Analysis for the Calibration of Driver Behavior Models

In addition to the conceptual structure of a driver behavior model, the data used to calibrate or train models play an essential role. The size, quality and nature of the data set selected for calibration have a decisive influence on the quality of the calibrated driver behavior. Nowadays, there exists a large number of different data sets that address human driver behavior. A fundamental distinction is made in the way the data are collected. On the one hand, data can be obtained via the sensors in vehicles, so-called Floating Car Data (FCD) [HARTH, ALI, et al., 2021; KESTING AND TREIBER, 2009]. On the other hand, cameras from stationary elevated positions or drones are used to obtain trajectory data from a top view [BOCK et al., 2019; COLYAR AND HALKIAS, 2007; HALKIAS AND COLYAR, 2006; ROBERT KRAJEWSKI, JULIAN BOCK, 2018]. In the case of FCD, the level of detail of the data set is differentiated more precisely. If the FCD data set contains not only GPS positions but also sensor information regarding surrounding vehicles, it is referred to as extended Floating Car Data (xFCD). Table 2.3 compares the suitability of the two selected data types, xFCD and trajectory data from a stationary camera, which are available in this thesis for calibrating driver behavior models.

	xFCD	Trajectory data
Effort to record data	-	++
Effort to process data	-	-
Available quantity	-	+
Driver variance	-	+
Spatial coverage	++	-
Temporal coverage	o	o
Data quality	+	-

Table 2.3: Evaluation of the xFCD and trajectory data sources with respect to various criteria. Scale: --, -, o, +, ++.

For the recording of xFCD, a comparatively high effort is required, since vehicles equipped with sensors and a data logger as well as the corresponding drivers have to be available. In contrast, there is a low effort for recording the trajectory data, since only the installation of a camera is required. The subsequent data processing to enable driver model calibration after recording is a time-consuming process for both types of data. An advantage of the trajectory data from the camera images is the high number of observed vehicles with different driver types in a short period of time. In comparison, the xFCD data set is limited to a few different drivers. In contrast, xFCD have a high coverage of the road network, which allows a higher generalizability of the data and thus a broader local applicability of the calibrated driver model. The temporal coverage is similar for both data sets. Due to the limited resolution of the camera image, trajectory data show higher geometric deviations, especially for points further away from the camera, as well as oscillations during detection, which have to be smoothed, e.g., by a Kalman filter. In addition, occlusions due to the perspective in which the camera captures the traffic complicate the correct detection of all objects. In comparison, the recorded xFCD are processed and prepared for vehicle functionality by vehicle algorithms, such as driver assistance systems. For example, each detected vehicle in the surrounding traffic is assigned a unique identification number that persists even after temporary occlusion. Accordingly, the higher data quality in the comparison with the trajectory data can be attributed to xFCD. All in all, both data sets show advantages and disadvantages. While xFCD stand out in quality and special coverage, the recording effort is very high. In contrast, the trajectory data set comes with a low effort to record and a high coverage of different drivers, but with limited spatial capability and less data quality.

2.5 Automated Vehicles in Simulation

In order to virtually investigate the impact of automated vehicles on traffic safety and efficiency, it is necessary to precisely model the behavior of the systems. In recent literature, engineering models are often used for this purpose [KIM AND HEASLIP, 2023; LU et al., 2019; PORFYRI et al., 2018]. Their parameters are adjusted to the assumed characteristics of automated vehicles. The main assumption for the characteristics of automated vehicles is often a lower reaction time compared to the human driver, which allows a smaller feasible time gap to the leading vehicle during the following maneuver [WAGNER, 2015]. Furthermore, automated vehicles which are controlled by an algorithm move more homogeneously than manually controlled vehicles. Another common assumption is higher compliance with the law regarding traffic rules like the applicable speed limit. In addition to the characteristics of the movement of automated vehicles, the communication of vehicles with other vehicles (Car2Car) or the infrastructure (Car2X) is considered in simulation of automated vehicles. In this thesis, however, collaborative effects on traffic safety and efficiency resulting from Car2Car or Car2X communication are only considered marginally.

In a simulation experiment, WAGNER [2015] investigates the traffic efficiency when automated systems are introduced into an urban traffic system. The automated vehicles are modeled with a Krauss model whose desired time gap is set to 0.5 s. The study assumes an ideal smooth and uniform driving style. In contrast, human drivers are modeled with a time gap of 1 s and a volatile driving behavior. It is shown that automated vehicles increase the efficiency of the transportation system by saving between 5 % and 80 % of lost time, depending on the time of day. LU et al. [2019] also use different parameter sets of the Krauss model to distinguish automated and human-controlled vehicles. In addition to the desired time headway, which is set to 0.6 s for automated vehicles and 0.9 s for manually controlled vehicles, the minimum standstill gap is reduced for the automated systems and the acceleration capability is slightly increased. In both a generic grid network and a simulation of a real-world section of an urban network, there is a 16 % to 23 % increase in capacity with increasing penetration of automated vehicles. Since an automated system is similar to an Adaptive Cruise Control (ACC) in the longitudinal direction, it can be used to model the longitudinal behavior of an automated vehicle. PORFYRI et al. [2018] compare the influence of different proportions of ACC-equipped vehicles in traffic. While the Krauss model with a reaction time of 0.7 s is chosen for manually controlled vehicles, the ACC controlled vehicles have a reaction time of only 0.1 s. The desired time headway of the automated vehicles is drawn from distributed values between 1.1 s and 1.6 s while the time gap of human drivers is set to 1.64 s. As the penetration of automated vehicles increases, PORFYRI et al. determine an increase in capacity of more than 20 %. RÖSENER [2020] models an SAE level 3 Motorway-Chauffeur and an SAE level 4 Urban Robot-Taxi and investigates the impact of the introduction of both functions on traffic safety. In contrast to the previously presented work, a larger time gap of 1.8 s is chosen for modeling both types of automated vehicles compared to the human driver with 1.05 s. It can be shown by a scenario-based approach that the introduction of automated systems promises a significant gain in traffic safety. The introduction of the Urban Robot-Taxi leads to a reduction of about 50 % in longitudinal urban conflicts.

2.6 Research Gap

As described in Section 2.4 in the state of the art, a large number of driver behavior models already exists and is used in various automotive and traffic engineering applications. In particular, the need for realistic driver behavior models for virtual testing of automated vehicles places high demands on the validity of the models used. The human driving task represents a complex heterogeneous process that must be reflected in a driver behavior model in its comprehensiveness. However, the so-called "engineering" models [SAIFUZZAMAN AND ZHENG, 2014] which are commonly used in traffic simulation idealize human longitudinal driver behavior. It is significant that the *current* desired acceleration is determined based on *current* ground truth input data, such as distance and relative speed. However, a real human driver creates a perceptual model of ground truth that can contain inaccuracies or even missing information; moreover, human drivers require a processing time to respond to perceived stimuli in the driving environment.

Human driving behavior is further characterized by various influences that act on the driver as individual differences or situational factors [HAMDAR, 2012] (cf. Section 2.3.4). The individual differences lead to specific driver behavior of different driver types (inter-driver). In addition, situational factors are reflected in varying driver behavior of an individual driver in similar situations (intra-driver). Both types of individual expressions of driving behavior can have an impact on traffic safety and efficiency. Therefore, it is important to consider individualities in the driver behavior model which is developed in this thesis.

Therefore, the challenge arises to find a suitable method that is able to cover the complexity of the human driving task, especially HFs, in order to represent human driver behavior more accurately than currently existing models. In addition to the actual method, the type of calibration or training should also be selected in such a way that the model can reproduce human driver behavior realistically. Furthermore, traffic scenarios must be investigated and chosen for calibration in which HFs play a significant role. For example, highway traffic is characterized by prolonged periods without significant changes in driving dynamics. In contrast, urban traffic in particular shows a significantly increased number of dynamic maneuvers influenced by HFs due to a variety of external influences. This identified research gap to find and calibrate an appropriate method which is able to consider HFs is addressed by **RQ 1.1** and **RQ 1.2** (cf. Section 1.2).

Existing driver behavior models, as described, often rely on the distance and relative speed to the leading vehicle to calculate a response behavior. However, in order to reproduce human behavior realistically, the developed driver behavior model should process those parameters that humans are able to perceive and use to perform the driving task. Furthermore, Section 2.3 shows various characteristics such as reaction time and anticipation of the human driver, which are often not or only insufficiently represented in existing engineering models. However, the representation of these HFs is essential in order to be able to capture human error behavior and the heterogeneity of driver behavior. Only due to the non-ideal characteristics of the human driver, errors can occur, which can lead to potentially critical situations, especially when the driver is surprised. This research gap raises the question of whether the developed driver model that takes HFs into account can represent human driver behavior more realistically than engineering models – despite or rather because it is impaired by HFs. This research gap is reflected by **RQ 2.1**.

Finally, the developed driver model is intended to control human-driven vehicles in the surrounding traffic of automated vehicles in a virtual endurance run. Therefore, when developing the driver behavior model, it must already be considered that it can be integrated into simulation. After successful integration, the applicability of the developed driver behavior model in a virtual test of automated vehicles in a traffic simulation must be investigated. The focus of this investigation is on whether the behavior of different driver types occurs concisely in the simulated traffic. Accordingly, the behavior of different driver types will be recognized by the automated vehicles and assessed in terms of criticality. Based on this assessment, the automated vehicles shall adapt their driving behavior in order to mitigate the potential criticality of car-following situations behind certain driver types. **RQ 2.2** addresses if the different driver types can be reliably recognized by an automated vehicle. Based on the driver type recognition, **RQ 2.3** discusses if the adaptation of the driving behavior of automated vehicles to specific recognized driver types impacts traffic safety and efficiency.

3. Advanced Modeling of Longitudinal Driver Behavior

As discussed above, in order to model human driving behavior in a broad spectrum of complex traffic situations, a data-driven approach must be combined with a realistic model of human perception and human response to stimuli in the driving environment. Based on the analysis of the state of the art as summarized in Chapter 2, a longitudinal driver behavior model designed for urban traffic is developed in this thesis. The model addresses the layers of guidance and stabilization in Figure 2.5 shown on the right-hand side. Section 3.1 discusses which type of modeling is appropriate for the driver model considering the state of the art. Section 3.2 presents the structure of the model, which is based on the psychological processes of humans, in detail. The data required for training the model are presented in Section 3.3. In addition to intra-driver-specific characteristics, such as reaction time, the modeling also includes inter-driver differences. The classification of drivers into four different driver types and their modeling is described in Section 3.4. Section 3.5 outlines the type of training process and the loss function of the model. Finally, in Section 3.6, the ability to reproduce real trajectories is evaluated on the basis of a comparison with existing engineering models.

Some aspects of the data, methodology and results presented in this chapter have been previously reported in [HARTH, ALI, et al., 2021] and [HARTH, AMJAD, et al., 2022].

3.1 Selection of an Appropriate Method for the Driver Behavior Model

Human driver behavior represents a complex process characterized by HFs and individual characteristics of individual drivers. A model that covers the entire process or parts of it must be able to explicitly reproduce the characteristics that arise from HFs. As already discussed in Section 1.2 and Section 2.4.1, conventional engineering models such as IDM and the Krauss model often idealize human driver behavior. In particular, HFs associated with the different levels of perception, information processing and action are only partially considered or not considered at all. Therefore, as discussed in Section 2.4.2, there are efforts to enrich classical engineering models by the explicit modeling of HFs in order to cover effects, which are reflected in the human driver behavior and result from HFs.

However, all these approaches combine the description of basic human driver behavior as a function of parameterized algebraic equations to guarantee mostly collision-free driving. This results in a driver response behavior limited to the logic of the equations, which does not cover the broad spectrum of possible heterogeneous behavioral possibilities in a given situation. Furthermore, engineering models are limited in their equations to parameters that mostly refer exclusively to the driver's own vehicle or the leading vehicle. Beyond these parameters, however, there are other important influences of the static and dynamic environment that play a major role in the driver's response behavior, depending on the current driving situation.

This thesis aims to model the human driver behavior in urban areas for the virtual tests of automated vehicles. Urban traffic scenarios are characterized in particular by the presence of traffic lights, which have a significant influence on the movements of individual vehicles. The approach to a signalized intersection represents a highly dynamic traffic situation, since the driver has to react to traffic light state changes as well as to movements of the surrounding traffic at the same time. These traffic light state transitions cause a change in the dynamic behavior of the traffic situation. This creates a multitude of different possible traffic scenarios that can potentially challenge the driver. Especially in unexpected situations, HFs, such as reaction time, play an important role, as they can potentially increase the criticality of a situation as a disruptive factor. The advantage of this scenario over others, such as the investigation of aggressive lane changes, is the locally limited observability of surprising situations for the driver. Based on the mentioned challenges of urban traffic, the developed model will be evaluated exemplarily in the urban scenario of approaching a traffic light controlled intersection. Engineering models, however, are designed particularly for highway situations on open roadways and can reproduce most human behavior in these scenarios well [TREIBER et al., 2000]. In contrast, urban scenarios place much more complex demands on driver behavior modeling because of the multitude of influences on the driver.

Data-driven methods have emerged as a possible approach to meet the outlined challenges to modeling (cf. Sections 2.4.1 and 2.4.2). These methods enable the learning of complex relationships from a large number of influencing parameters. The number of parameters can be arbitrarily extended depending on the availability of data. For modeling human driver behavior, the use of RNNs and especially of LSTM networks has proven particularly useful. Due to their ability to process temporal sequences, an essential property of humans to determine future behavior based on a time series of stimuli stored in short-term memory is covered. As discussed in Section 2.4.1, RNNs are able to reproduce several human characteristics in driving behavior. MORTON et al. [2017] have pointed out that without explicitly considering other HFs, particularly reaction time, LSTM relies heavily on current acceleration trends to predict future driving state. However, this behavior does not correspond to real-world driver behavior, which lacks responses to immediate stimuli due to reaction time. Therefore, HFs resulting from the levels of perception, information processing, as well as action must be explicitly modeled and their manifestations on driver behavior must be included in the training process of LSTM. As described in Section 2.4.2, LSTM outperforms feedforward neural networks in reproducing real-world trajectories.

Based on the current state of the art, LSTM is chosen as the method for the developed driver behavior model in this thesis. In particular, the ability to generate a response behavior based on a time sequence rather than a single point in time qualifies the method for use in driver behavior modelling.

3.2 Conceptual Design of the Driver Behavior Model

The model developed in this thesis is based on the psychological three-stage concept of perception, information processing and action presented in Section 2.3. The model must cover driving conditions without a leading vehicle (free driving) as well as with a vehicle in front (car-following). Figure 3.1 shows the architecture of the model as a closed-loop process including the HFs implemented in the respective modules.

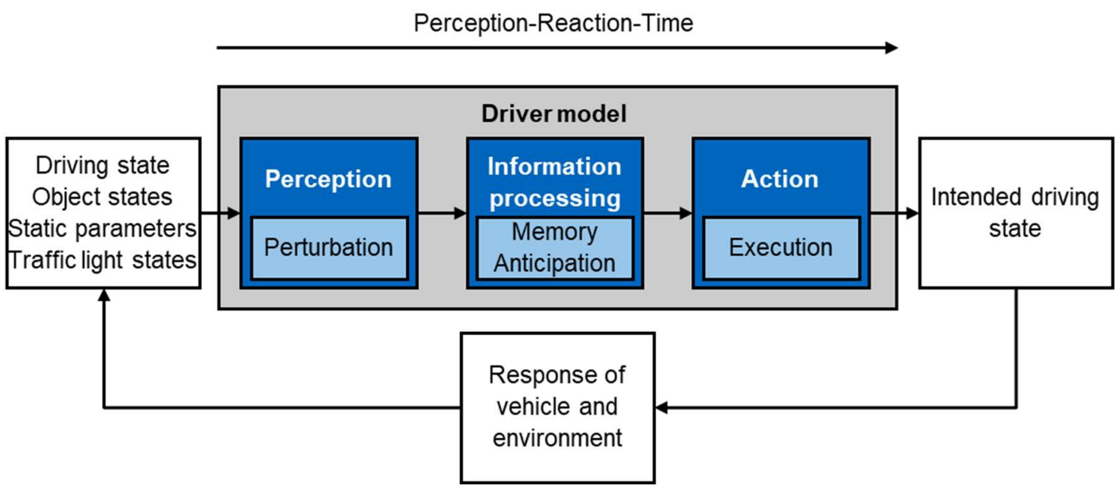

Figure 3.1: Architecture of the developed driver model considering the three layers perception, information processing and action.

The model uses its own current dynamic driving state and information about the driving state of an object in front. In addition, parameters about the static environment and dynamic information about a traffic light ahead are taken into account. This information is captured by the perception module. To simulate human inaccurate perception, the exact variables are perturbed in this module. The perturbed data is passed to the information processing module. There, a temporal sequence of the data is stored in a short-term memory. Based on the information stored in the short-term memory about the observed state from the own vehicle as well as the environment, the temporal and spatial parameters are projected into the near future. In the third module, action, the temporal sequences of perceived and anticipated information are combined and a decision regarding the desired driving state is made and executed. The entire process described, from perception to information processing to action execution, does not occur instantaneously but is subject to a certain period of time, the Perception-Reaction-Time (PRT) (cf. Section 2.3.3). Thus, the output variables are delayed compared to the input variables. The resulting action by the model causes a changed traffic state, which is returned to the model in a closed loop as input information for the next time step. The architecture of the model shown is identical for both driving states, free driving and car-following. Only the input parameters differ between the two variants: while the car-following model additionally considers parameters concerning the leading vehicle, the parameters of the free driving model are limited to information about the own driving state, the static environment and the traffic light signals.

The implementation of HFs must take into account the capabilities of LSTM. Parts of the HFs, such as a heterogeneous fluctuating speed profile, are implicitly learned by LSTM from the data. In addition, LSTM has a short-term memory that allows to respond to a temporal sequence of input signals. In contrast, the response of engineering models is usually based on only a single historical time step. However, incorporating other HFs into LSTM, such as PRT, requires adaptations to adequately represent these HFs. In the following Sections 3.2.1 - 3.2.3, the individual modules of the architecture and the implementation considering the capabilities of LSTM are presented in more detail.

3.2.1 Perception Module and Input Parameters

The concrete selection of the input parameters of the developed driver model is inspired by the influences perceivable by a real human driver. Figure 3.2 represents the module of perception and the perceived parameters. In the following, the individual input parameters are presented in more detail and their selection is justified.

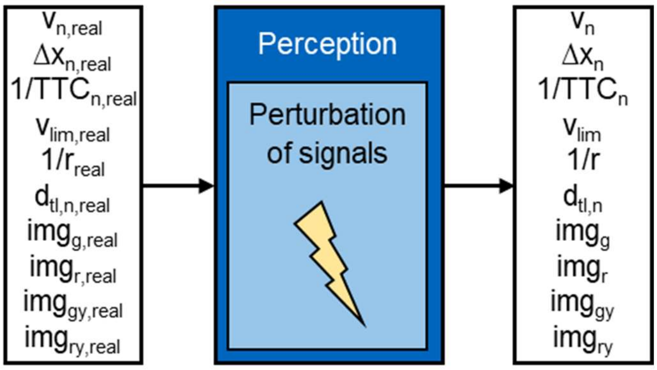

Figure 3.2: Perception module of the developed driver model.

The driver of a vehicle n estimates the prevailing speed $v_{n,real}$ and corrects the estimate by sporadically reading the speedometer. Speed is a crucial parameter for evaluating the current driving condition and is therefore selected as an input parameter of LSTM. Studies also indicate that humans have a relatively poor ability to estimate acceleration [CALDERONE AND KAISER, 1989; WATAMANIUK AND HEINEN, 2003]. Therefore, the introduction of acceleration as an additional input parameter is deliberately omitted. Instead, the estimation of acceleration is implicitly covered by the ability of LSTM to process temporal sequences of velocity. In the car-following driving condition, two additional parameters are considered that characterize the interaction with the leading vehicle. The first parameter is the net distance $\Delta x_{n,real}$ to the front vehicle, which is determined visually by the driver. In addition, the relative motion of both vehicles plays an essential role in evaluating the criticality of a situation. An important parameter for this evaluation is represented by the time-to-collision $TTC_{n,real}$, which is calculated from the quotient of distance $\Delta x_{n,real}$ and relative speed $\Delta v_{n,real}$. LEE's [1976] $\dot{\tau}$-theory (cf. Section 2.4.2) describes that human drivers perceive differences of longitudinal movements as angular deviations of projections on the retina. $TTC_{n,real}$ can therefore be regarded as an optical variable τ, which takes into account the visually perceived projection angle and its temporal change. In order to limit the range of values as well as to avoid the discontinuity at relative velocities in the range around 0 m/s and thus to create a better precondition for the learning process of LSTM, the inverse $1/TTC_{n,real}$ is considered as input parameter instead of $TTC_{n,real}$.

In addition to the influences of moving road users, the static environment also has an effect on the driver. The currently prevailing speed limit $v_{lim,real}$ sets a guide value for the selected speed of the driver and is therefore selected as an additional input parameter. Furthermore, the course of the lane has an influence on the driving behavior. A current or preceding more curved lane causes the driver to adjust the selected speed. The current curvature $1/r_{real}$ of the road thus serves the driver model as a parameter describing the geometry of the static environment. r_{real} is the current radius of the road.

Traffic in urban areas is significantly influenced by signal control through traffic lights. For the driver model, the traffic light state of the preceding traffic light is used as an input. Instead of processing the state as one variable, the individual traffic light states green $img_{g,real}$, red $img_{r,real}$ and the phase transitions from green to yellow to red $img_{gy,real}$ and from red to yellow to green $img_{ry,real}$ are explicitly split. It is obvious that the prevailing traffic light state becomes more relevant with decreasing distance to the traffic light. The distance to the stop line $d_{tl,n,real}$ represents the last input parameter of the driver model.

Apart from the speed limit and the signal status of the traffic light, all parameters visually perceived by the driver are subject to driver- as well as situation-dependent inaccuracies. To model the perturbation of perfect perception, a perturbation variable is introduced to be applied to the visually perceived information. The disturbance variable is drawn for each driver from an assumed lognormal distribution with mean 0.03 and standard deviation 0.05. The shape of the distribution causes drivers to evaluate situations slightly too critically on average, but occasionally to underestimate criticality. The input variables to the model are multiplied or divided by the confounding variable in a way that affects criticality in a consistent manner. For example, higher criticality implies higher speed but lower TTC.

3.2.2 Information Processing Module

The information processing module receives the data output from the perception module. Figure 3.3 shows the information processing module of the driver model.

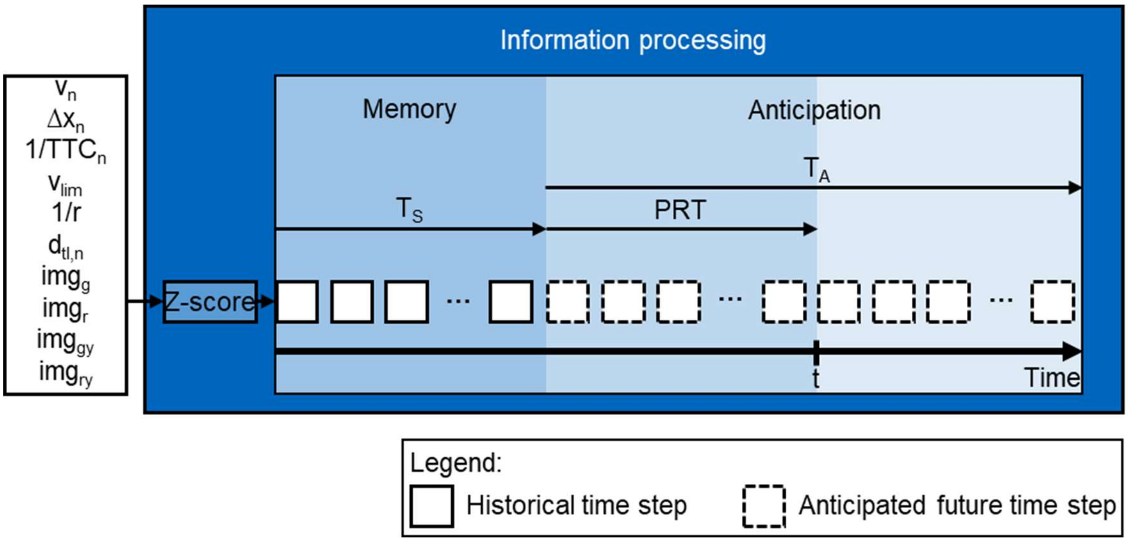

Figure 3.3: Information processing module of the developed driver model.

To facilitate the learning process of LSTM, all input parameters of information processing except traffic light states and curvature are first prepared so that LSTM can process them optimally. Since all signals are scaled differently, they are standardized for training LSTM using the Z-score.

$$z = \frac{x - \mu}{\sigma} \tag{13}$$

In equation (13), μ and σ correspond to the mean value and standard deviation of the original data x.

Information processing includes short-term memory, which stores a time sequence of perturbed input data with a duration of T_S = 1 s. WHEELER et al. [2016] show that despite providing temporally sequenced inputs, the response behavior of LSTM is mainly based on current acceleration trends. This behavior produces correct driver behavior in the majority of cases, as humans can use anticipatory effects to bridge reaction time in non-critical situations. However, only in a few surprising situations the "true" reaction time is revealed. However, LSTM does not learn it implicitly. Therefore, the input data is delayed by the PRT.

However, a real human driver does not plan future decisions solely on past stimuli from short-term memory. Instead, an important human ability is the projection of the last perceived state into the near future. This anticipation allows the driver to react to and mitigate potentially critical situations ahead of time. In addition, anticipation allows the driver to partially compensate for the PRT. LSTM has the potential to project a prevailing situation into the near future, as MA AND QU [2020] show. However, the data set available for this thesis is mainly characterized by comparatively short observation periods of real drivers. Therefore, LSTM is not assumed to be able to learn human anticipation sufficiently well. Common models for anticipation assume a constant acceleration or speed for the vehicle ahead and the vehicle following. For example, TREIBER et al. [2006] use a constant-velocity heuristic to predict the driving state of the following vehicle and a constant-acceleration heuristic for the state of the leading vehicle in the near future. However, TREIBER et al. [2006] study only highway scenarios, which are characterized by lower dynamics than usual in urban traffic. Since highly dynamic situations in the approach to urban signalized intersections are considered in this thesis, a constant acceleration is assumed for both vehicles over a time horizon of $T_A = 2$ s into the future. In particular, the constant-velocity heuristic does not adequately describe an approach to a red traffic light or a queue.

In addition to the vehicle-related dynamic variables, the state of the preceding traffic light is anticipated. A sustained green or red signal is assumed to be constantly green or red, respectively, over the entire time horizon of 2 s. Phase changes erroneously not predicted by this assumption therefore represent a surprise for the driver. Actual phase changes from green to yellow and from red to yellow detected by the driver are predicted with a duration of 3 s and 1 s, respectively. The duration of the phase transitions is subjected to the same perturbation variable, analogous to the perception module, in order to emulate the estimation inaccuracy.

In addition to temporal anticipation, the spatial state is projected into the future. Using path planning and the current velocity, the anticipated path over the anticipated time horizon is determined. Based on this path, the values for the speed limit and curvature are determined. Particularly in front of tight curves, e.g., turns, the driver can adjust the speed at an early stage.

The generated temporal and spatial anticipated input sequence is added to the historical input data and provided to the model.

3.2.3 Action Module

In the action module, the time sequence of the input parameters from short-term memory is combined with the subsequent anticipated time sequence. Taking into account the PRT, a reaction, expressed by the desired acceleration, for the upcoming time step is determined on the basis of the two time sequences. Figure 3.4 schematically shows the internal processing of the historical and anticipated data over time.

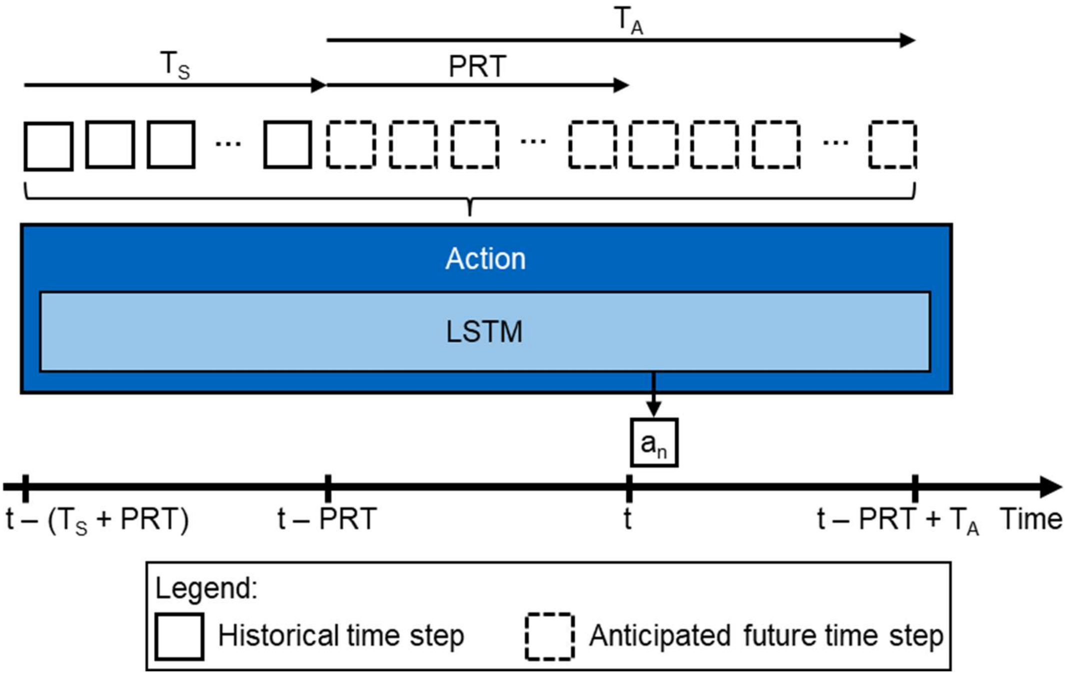

Figure 3.4: Action module of the developed driver model.

Figure 3.4 describes the procedure of processing the model input data over the time t shown in the lower part of the figure. The figure is basically divided into three time ranges from $t - (T_S + PRT)$ to $t - PRT$, from $t - PRT$ to t, and from t to $t - PRT + T_A$. The first time range describes the range of the sequence of historical data. These data are divided into individual time steps, which are schematically represented as boxes with solid frame lines. The two subsequent time intervals from $t - PRT$ to $t - PRT + T_A$ cover the range of anticipated data represented by the boxes with dashed lines. The overlap of the anticipated time horizon T_A and PRT reveals the compensatory effect of anticipation over reaction time. As long as the driver is moving in non-critical situations, anticipation bridges the information missing due to reaction time. However, as soon as the driver is confronted with a surprise, such as the phase transition of a traffic light from green to yellow, the anticipation is wrong. Therefore, surprises reveal the actual reaction time of the driver, who cannot react to a surprising stimulus until he/she has consciously perceived it.

The core of the action module represents a neural network consisting of two LSTM layers with 256 units each, a dense layer with 256 units and an output layer with 1 node. The training of the model is performed using the optimizer Adam. The optimization parameters are initialized with learning rate = 0.001, beta1 = 0.9, beta2 = 0.999, epsilon = 1e-07 and decay = 0. The architecture of LSTM allows short-term storage of temporal sequences of input variables in a short-term memory and their subsequent processing. LSTM is particularly suitable for emulating human driver behavior, as it has been proven to be able to reproduce human characteristics [HUANG et al., 2018]. In the action module, the combined time sequences with sequence length $T_S + T_A$ serve as the input sequence for the LSTM network. A desired acceleration a_n for the current time step t results from the internal processing. This output acceleration is delayed with respect to the quantities consciously perceived by the driver due to the PRT. The model, as described in the figure, is used for both driving states free flowing as well as car-following. However, the quantities Δx_n and $1/TTC_n$, which refer to the leading vehicle, are omitted for the free flowing case. For both driving conditions, different models are trained independently of each other with the respective input variables.

3.2.4 Composition of the Model

Figure 3.5 shows the overall composition of the modules described in Sections 3.2.1 to 3.2.3.

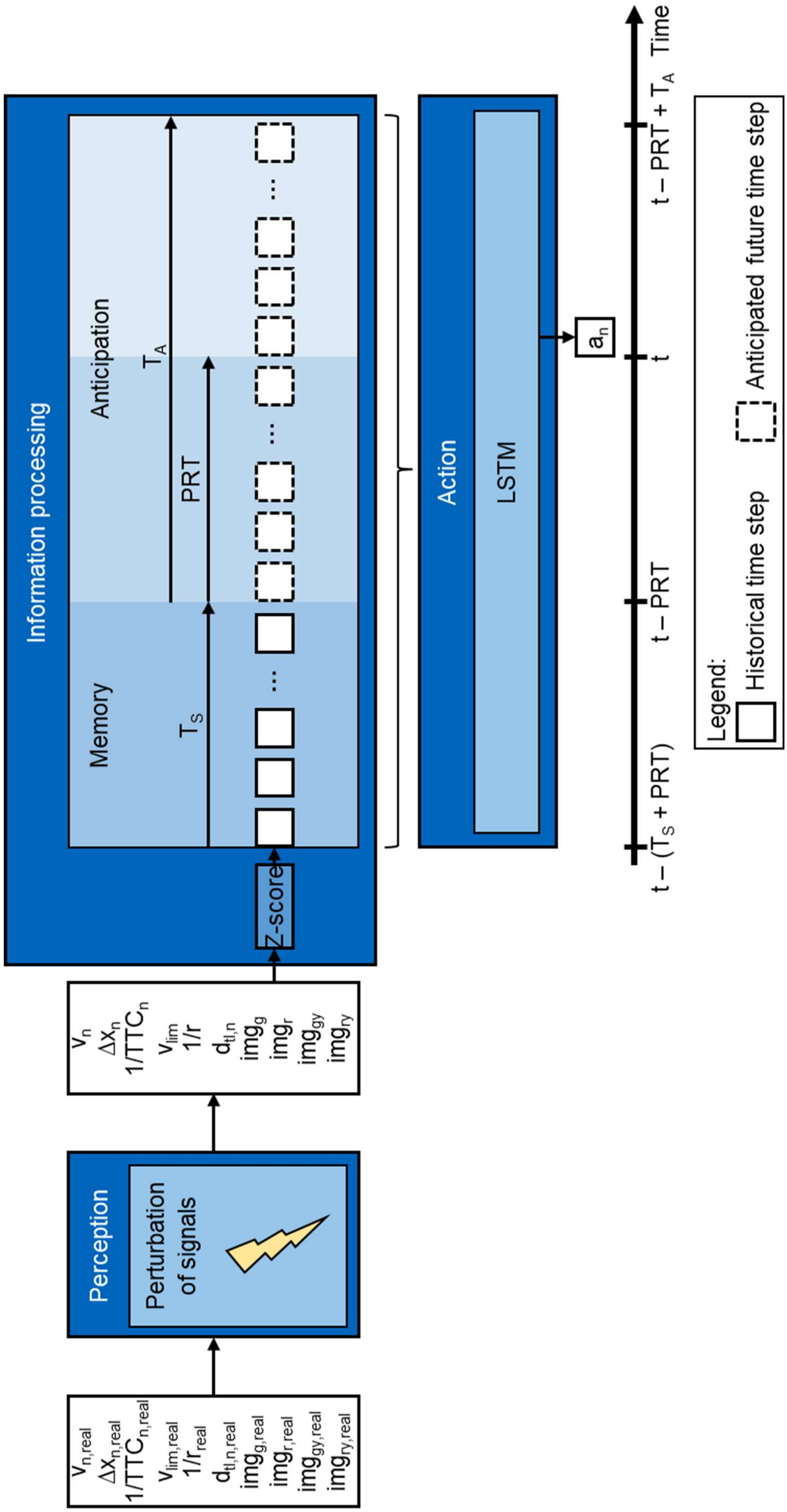

Figure 3.5: Composition of the three modules of the developed driver model.

3.3 Data Analysis and Preparation

In the following, the data sources available for training and evaluation are presented in detail. Furthermore, the suitability of the data sources is analyzed with regard to their applicability for the training of LSTM. Finally, the special situation of the phase change of the traffic light from green to yellow with the associated dilemma situation for the driver is investigated. In this dilemma situation, the driver has to decide to drive through or to stop. Based on this analysis, a decision model is developed.

3.3.1 Available Data

For the training of the developed driver model, the data comes from two different sources. On the one hand, xFCD from vehicles of AUDI AG are used. The term xFCD is used to describe FCD that include not only the dynamic variables of the driver's own vehicle but also the driving conditions of surrounding vehicles, such as distance and relative speed. For the recording of xFCD, vehicles of AUDI AG had a data logger with a recording frequency of 25 Hz. The vehicles were equipped with a series sensor set consisting of a camera and radar. The recorded xFCD data set includes 610 approaches to a variety of signalized intersections in the urban area of the German city of Ingolstadt. The trajectories are each truncated at a distance of 150 m from the stop line at the nearest traffic light. Furthermore, records of historical traffic light conditions are available for 75 traffic lights in Ingolstadt (cf. Section 4.2). The information about which signal groups of the traffic lights were relevant at the times of the recorded xFCD approaches is determined using map-matching methods. For this purpose, the xFCD are matched to the road network of the traffic simulation in SUMO (cf. Section 4.1) before they are broken down into a large number of individual approaches, using a procedure that considers the logical connections of the map. As a result, the position data is unambiguously assigned to the geometrically overlapping connections within intersections. In addition, the available signal groups are assigned to the respective incoming lanes in front of an intersection and the connecting lanes within an intersection using a reference map according to the method which is described in Section 4.1. By superimposing the results of the matching of xFCD and signal groups, the relevant signal group for each approach can be determined. The procedure enables an unambiguous assignment of a trajectory to a signal group as well as a maneuver to be executed over an intersection. Using the time stamp in xFCD and the historical traffic light data, the individual switching states of the relevant signal groups can be assigned to the vehicle data. Figure 3.6 shows the speed profile of an exemplary selected approach to a traffic light from the xFCD.

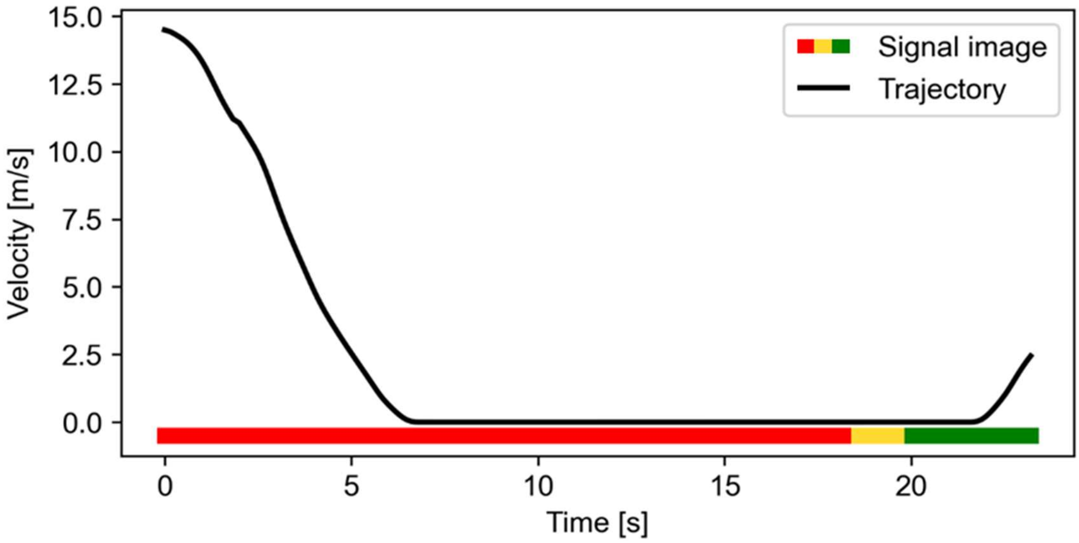

Figure 3.6: Exemplary trajectory towards a traffic light obtained from xFCD.

In the figure, the speed of a vehicle is plotted against time. The signal of the traffic light is shown as a narrow bar in the lower part of the figure. The color of the bar symbolizes the respective state of the traffic light. Already at the beginning of the approach, the traffic light shows a red signal. The driver decelerates from approx. 14.5 m/s to standstill. After approx. 18 s, the state of the traffic light changes to yellow and then to green. The driver now continues his/her journey.

In addition to the data from xFCD, trajectory data recorded by a camera at a single signalized urban four-way intersection in Ingolstadt on two days is used. In total, the data set comprises 20860 trajectories. Figure 3.8 shows the scene recorded by the camera. The camera has a resolution of 1920 x 1080 and is installed in an elevated position above the intersection. It provides a video stream with a frequency of 30 frames per second. In order to identify the objects, their position as well as their object category in the camera image, an image recognition method based on Mask R-CNN [HE et al., 2020] is used. A geometric transformation is applied to convert the camera image data in pixels into georeferenced position data of the objects. In the following, the position data is smoothed using a Kalman filter. This is necessary because oscillations occur when the objects are detected. The accuracy of the georeferenced position was determined to be 0.57 m on average by performing a large number of measurement runs with a vehicle equipped with high-precision sensor technology. The velocity and acceleration of all objects are calculated from the position data via the derivatives with respect to the time. Figure 3.7 shows the same traffic situation as

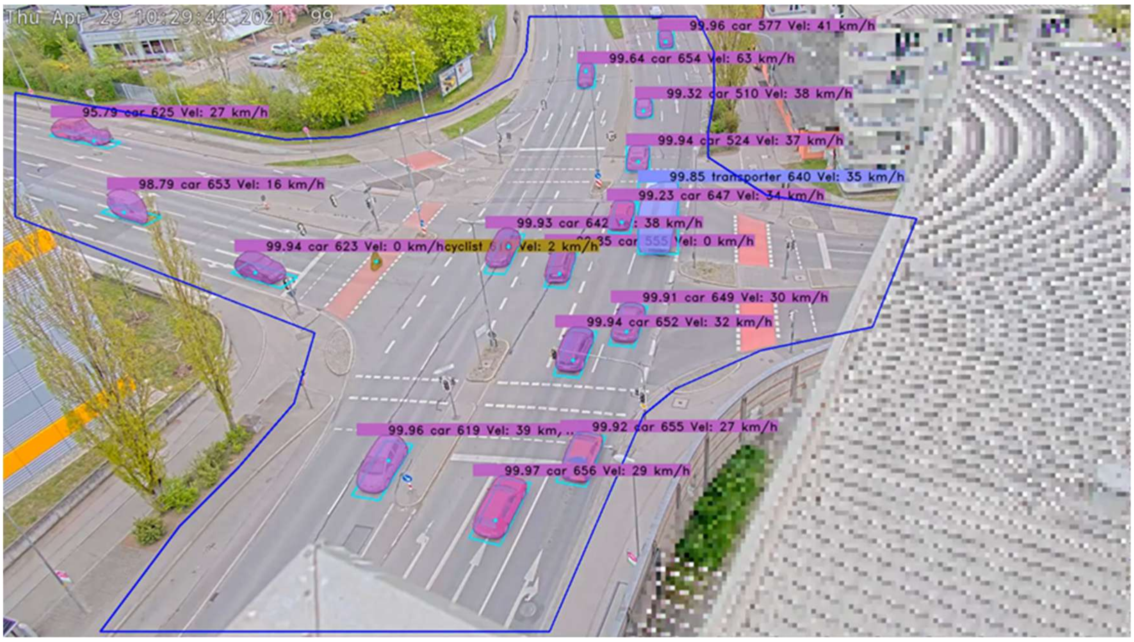

Figure 3.7: Camera image of the observed intersection with semantic segmentations of the traffic participants.

Figure 3.8: Camera image of the observed intersection without semantic segmentations of the traffic participants.

Figure 3.8 but with segmented vehicles. The blue frame around the intersection defines the area in which vehicles are segmented.

In order to obtain information on the relevant signal groups of the traffic light as well as on the logical relations of the vehicles to each other, the geo-referenced trajectory points of the objects have to be matched onto a map. However, in contrast to xFCD, the intersection trajectory data is not matched to the SUMO simulation network. Instead, a measured high-resolution OpenDRIVE [ASSOCIATION FOR STANDARDIZATION OF AUTOMATION AND MEASURING SYSTEMS, 2022] map of the intersection with an accuracy of 5 mm is available. The matching of the trajectory data is analogous to the matching of the xFCD to the SUMO network and also incorporates the logical connections within the intersection. The follower-leader relationships can be determined from the assignments of the objects to the individual lanes of the map using the logical connections of the lanes. Finally, the available traffic light information for the intersection is assigned to the objects in an analogous process as already described for the xFCD.

The entire data set, consisting of 610 xFCD approaches and 20860 trajectories from the camera recording enriched with additional information, is downsampled to a step size of 0.1 s and filtered for the training of LSTM. On the one hand, trajectories that do not have enough data points in the area before the stop line are sorted out. On the other hand, trajectories are removed from the final data set that must give priority to temporally and spatially competing vehicle flows within an intersection. The reason for this is that the developed driver model does not have any information about the right-of-way on the intersection ahead. The enrichment of the model with this data is conceivable in the future. The final data set consists of 334 xFCD trajectories as well as 8065 trajectories from the camera data. This data set is split 80 % for training, 10 % for validation during training and 10% for testing. In addition, in order to pay special attention during the training to situations where a considered vehicle must stop, data augmentation is applied to such trajectories from the training data set. These trajectories are added to the training data set with slight variation by multiplying the parameters describing the driving state of the considered vehicle as well as the potential preceding vehicle ahead by the multipliers 0.94, 0.96, 0.98, 1.02 and 1.04.

A categorical comparison of both the xFCD and the trajectory data set is done in Section 2.4.4. For the training of LSTM, both data types are used in order to compensate for the deficiencies of each type.

3.3.2 Data Analysis and Modeling of the Yellow-Light Dilemma

A special situation when approaching a signalized intersection is the state change from green to yellow to red. Depending on the time headway to the intersection, a dilemma situation may arise for the driver as to whether he/she can still pass the traffic light in time before the change to red or whether he/she can come to a stop in front of the traffic light. When different drivers are confronted with this dilemma, different behaviors are shown in identical situations [GATES et al., 2007]. The bifurcation of whether drivers pass through a yellow light or come to a stop can be described mathematically by a logistic regression model. Using the model, a probability of passing or stopping can be estimated as a function of the Time-To-Junction TTJ (TTJ = $d_{tl,n}$ / v_n) at the time of change from green to yellow. The entire data set on xFCD and trajectory data was examined with respect to the dilemma situations and a logistic regression model was determined. Figure 3.9 shows the probability distribution of stopping in front of the traffic light as a function of TTJ.

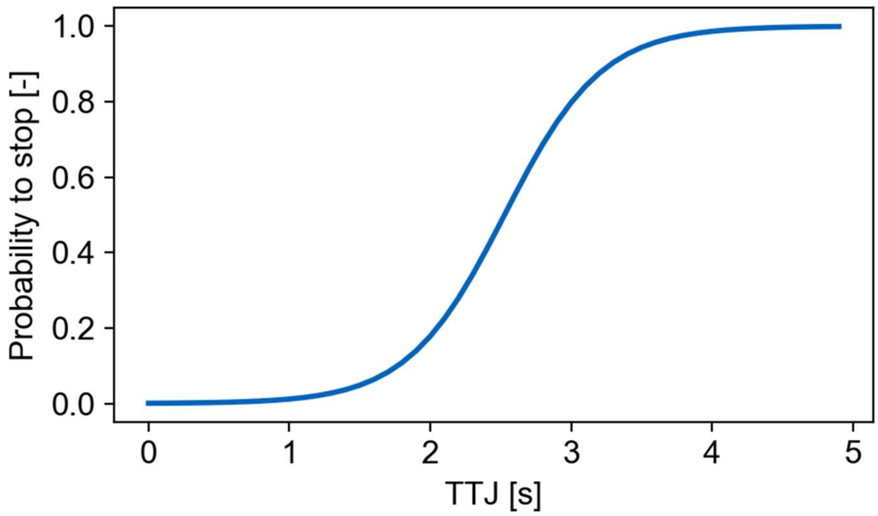

Figure 3.9: Logistic regression model for the yellow-light dilemma of stopping at the traffic light; probability 0 means crossing, probability 1 means stopping.

Figure 3.9 shows that the probability of stopping in front of the traffic light increases with increasing TTJ. Below 1 s TTJ, almost all drivers pass the traffic light. At a TTJ of 2.53 s, half of the drivers decide to stop and the other half to go through. Above 3.5 s, according to the model, almost all drivers stop. The duration of the yellow time in Germany is 3 s in inner cities. In the data, hardly any red crossings can be detected. However, passing the traffic light after 3 s would mean a red crossing. It can therefore be concluded that the model slightly distorts the probability in the range between 2.5 s and 4 s. Nevertheless, it provides information about the general distribution of stopping and passing drivers in the dilemma situation. The information from the regression model is used for decision making of the developed driver behavior model in the dilemma situation. The implementation is explained in Section 3.6.

3.4 Modeling of Individual Driver Behavior

As described in the state of the art in Section 2.3.4, there are two categories of influences on the driver and thus the observable driver behavior. On the one hand, individual permanent differences, such as the driver's age and gender, as well as vehicle attributes, such as performance capability and vehicle size, influence individual driving behavior. On the other hand, situational factors, which are divided into driver-individual factors, such as fatigue, and factors regarding the vehicle environment, such as weather influences, affect the driver. As a result, each driver reacts to all prevailing influences with an individual driver behavior. From the existing data set of xFCD and trajectory data from the observed intersection, no conclusions can be drawn about the majority of individual and situational individual factors acting on the driver. Instead, the salience of the influencing factors that manifests itself in a specific driver behavior can be quantified. In the following, groups of different driver types are formed between the individual drivers from the existing data set on the basis of these observed characteristics. From the data of these groups, separate driver behavior models are trained that describe the behavior of the respective group. Due to multiple driver changes during the trips, xFCD cannot be assigned to a single driver. Therefore, all approaches to traffic lights from xFCD are considered as separate trajectories which are assigned separately to the driver type groups. For dividing the set of drivers into different driver types, the following parameters are analyzed:

- Maximum speed
- Maximum acceleration
- Absolute value of the maximum deceleration
- Critical Time-To-Collision (TTC) (preferably positive, otherwise negative)
- Minimum time headway (THW)

While driving, a driver may show different behavioral characteristics depending on the situational influences affecting him/her, which allow him/her to be temporarily assigned to one or more groups of driver types. In addition, the short observation interval available for classifying drivers poses a particular challenge in the assignment process. Not all observed drivers have a preceding vehicle during the recording. The values for TTC and THW cannot be determined for these drivers. Therefore, for a consistent assignment to a driver type, the missing variables are determined by a multivariate imputation, the so-called MICE algorithm [VAN BUUREN AND GROOTHUIS-OUDSHOORN, 2011]. MICE allows an estimation of these missing variables based on the existing data of a considered driver as well as the characteristics in the rest of the data set. After completing the data set, a k-means clustering, also used for the same purpose by CHEN et al. [2020], is applied to the data set in combination with a Principal Component Analysis (PCA) to assign all drivers from the data set to the driver types. In order to make the five parameters, which have different scales, comparable with each other, they are standardized in a first step using the Z-score (cf. Equation (13)). Subsequently, PCA is applied to the normalized data set. PCA reduces the dimensionality of the data set from five to three dimensions without losing the meaningfulness. Instead of the original five parameters three so-called principal components are defined that are orthogonal to each other in the threedimensional space. The number of principal components is determined by a trade-off so that the original number of parameters is reduced and yet the new principal components explain as much variance in the data as possible. A k-means clustering with different numbers of clusters k is performed on the three resulting principal components. In order to determine the optimal number of clusters, the silhouette coefficient [ROUSSEEUW, 1987] is used. The silhouette coefficient gives information about the quality of the clustering. For the existing data set, the optimal cluster number is four with a silhouette coefficient of 0.27. As a result, the 8399 drivers are divided into four clusters with the following sizes: Cluster 0 contains 2759 drivers, Cluster 1 3094 drivers, Cluster 2 874 drivers and Cluster 3 1672 drivers. Figure 3.10 shows the kernel density estimates (KDE) of the considered parameters maximum speed, maximum acceleration, absolute value of maximum deceleration, minimum THW and critical TTC for all clusters. Figure 3.10 illustrates the distributions of the five selected parameters for each cluster. Cluster 0, shown in red, has a speed distribution that is mostly in the range of the prevailing speed limit of 50 km/h (≈13.9 m/s). The distributions of maximum acceleration and maximum deceleration are in comparatively small value ranges. The distributions of minimum THW and critical TTC mostly lie in uncritical value ranges. The drivers of the driver population covered by cluster 0 are referred to as *Moderate drivers* in the following. Compared to the other groups, the drivers from cluster 1 show an excessive maximum speed. Therefore, they are titled *Speeders* in the following. Moreover, the values of TTC and THW are in non-critical ranges. Since the data of the group of *Speeders* contain only a small number of situations that require a stronger deceleration and especially a stop in front of the traffic light, these situations in the training data are additionally augmented. The augmentation is done analogous to the one described in Section 3.3.1. The augmented data is added to the data set of the *Speeders*. Cluster 2 describes *Timid drivers*, which are mainly characterized by large time headways and negative TTC values. Finally, cluster 3, the *Aggressive drivers*, show high values of maximum acceleration and deceleration. Furthermore, the distributions of minimum THW and critical TTC are in comparatively low value ranges.

It is possible that, due to the short observation time, drivers could be assigned to other clusters depending on the situation, which is significantly influenced by the signal state of the traffic light. The quality of the clustering results could certainly have implications for the robustness and accuracy of all the models investigated in this work.

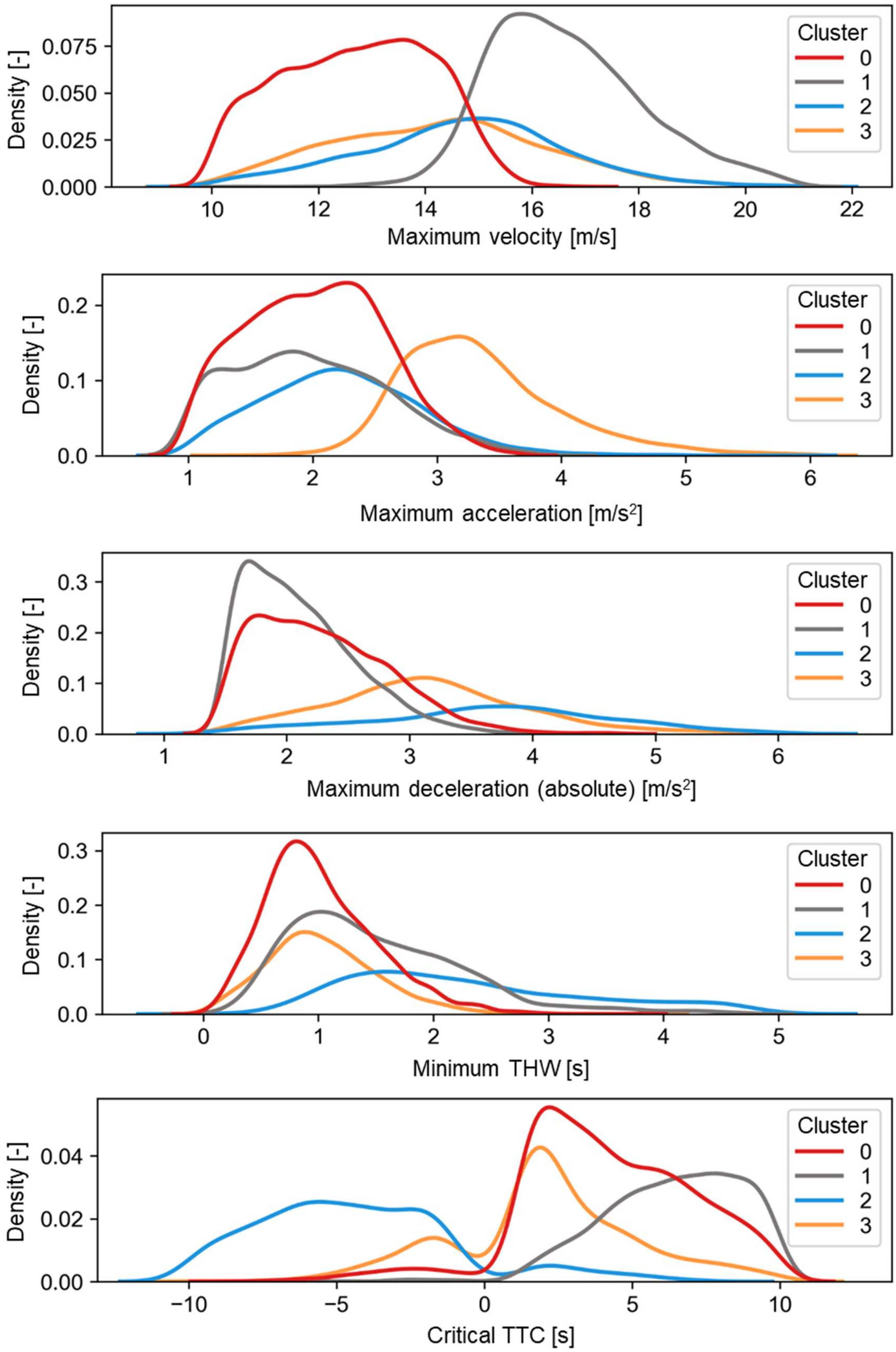

Figure 3.10: Kernel density estimates of the parameters maximum velocity, maximum acceleration, maximum absolute deceleration, minimum THW and critical TTC for all four clusters.

3.5 Training of the Developed Driver Behavior Model

The training of the developed driver model is carried out separately on the data of the four clusters presented in the previous Section 3.4. As already described in Section 3.3, 80 % of the data serve as training data, 10 % for validation and 10 % for testing. The separation of the data into the three data sets is done after the classification into the different clusters. For each cluster, different models are trained for the driving states of free driving and car-following. The main objective in the design of the training process and the loss function is to reproduce human characteristics in the execution of the driving task. The design of the training process and the modular loss function are described in detail in the following.

3.5.1 Design of the Training Process

For training neural networks, the existing training data set is usually divided into independent data snippets. Each of these data snippets contains a sequence of input parameters as well as the corresponding target parameter, which should be determined as accurately as possible by the network. These independent snippets are aggregated into batches, which the network uses to learn and update its weights. For the use-case in this thesis, this means that each traffic light approach is divided into time intervals with a duration of 1 s using a moving window. Additionally, an acceleration value delayed by the reaction time of also 1 s serves as the target variable in the learning process. The interval of input data of 1 s covered by 10 time steps corresponds to the capacity of the short-term memory of the trained LSTM. In addition, the driving state covered by the time snippet is projected into the future for another 2 s or 20 time steps, respectively, as described in Section 3.2.2. The network predicts an output acceleration value for each input data snippet, consisting of real historical and anticipated data. Figure 3.11 illustrates this type of training process.

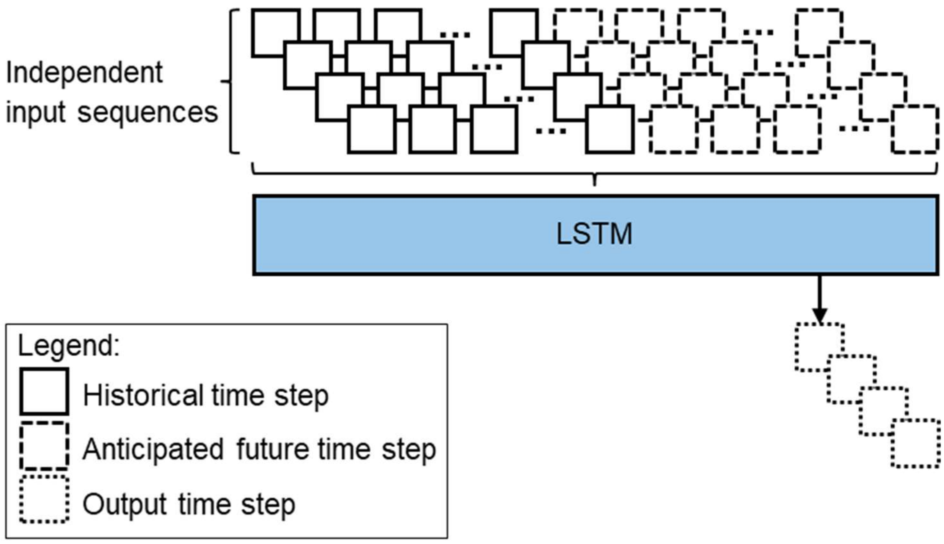

Figure 3.11: Open-loop training process of the LSTM model.

In the figure, the conventional training process is shown with a batch size of 4 for illustration. The black boxes with a solid frame line describe the input data of each time step, the boxes with a dashed line the anticipated data and the boxes with a dotted line below the LSTM the predicted output values.

This type of training process causes that the temporal relationship of the individual data snippets and their association with a particular approach to a traffic light are lost. Instead, the driving state of the vehicle is always reset to ground truth values in each training interval. Therefore, the network cannot learn from the accumulation of errors resulting from its own prediction. In particular, for the considered highly dynamic scenario of approaching a signalized intersection, the continuous correction of arising deviations from a realistic and plausible driving condition is essential. For example, the driver model must counteract when it is on a collision course with another vehicle due to prior predictions or when it might run a red light. For these reasons, the training is performed in a closed-loop approach. Figure 3.12 shows the training process schematically for an exemplary batch size of 4, which was selected in the figure for better representability. The batch size actually selected for the training is 20 input sequences.

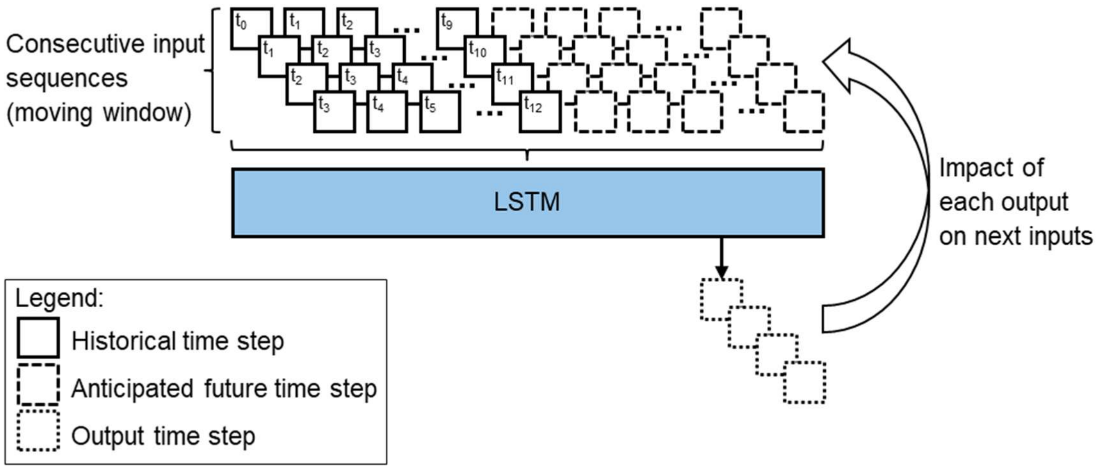

Figure 3.12: Closed-loop training process of the LSTM model.

To preserve temporal and logical relationships, temporally related segments of each approach to a traffic light that can be assigned to the same driving state (car-following / free driving) are used in their entirety during the training process. The first 20 time steps serve as the initial state for training with the respective trace. From the first ten time steps t_0 to t_9 of the trace segment, which are shown in Figure 3.12 as boxes with a solid line, the input parameters of the initial input sequence are extracted. These are projected into the future for 20 time steps, shown with a dashed line, using the methodology described in Section 3.2.2. The ten remaining historical time steps of the initial state are bridged by PRT. From the input sequence, an acceleration value is predicted, which is represented by a box with dotted lines in Figure 3.12. Using this acceleration value, the parameters that depend on the own driving state, such as velocity, distance to the traffic light and distance to the leading vehicle, are separately calculated for the time of prediction t (cf. Figure 3.5). For all training steps, all variables not related to the own vehicle, such as position and velocity of the leading vehicle, are taken from the real data as ground truth values. In the further training process, the ten time step interval of the input data is moved step by step as a moving window, so that the model has data delayed by PRT in each case. From the acceleration value predicted by LSTM, the own vehicle movements are determined and looped back via the closed-loop approach as input data for the next training step. Initially, the input data consists solely of real observed data. As soon as the moving window reaches the 21st time step, the already simulated data successively form the input data. The update of the weights of the neural network is done in 2-second intervals, i.e. a batch size of 20 single training steps, or when the end of the trajectory is reached. The evaluation of the quality of the predictions of LSTM is done with the help of the loss function, which is described in detail in the following Section 3.5.2.

3.5.2 Design of the Loss Function

As described in Section 3.5.1, the training is performed with consecutive sequences of the traces of the training data set of the respective cluster. The predicted movements of the driver have an influence on the input data of subsequent training steps. Therefore, the training process must be aligned with the closed-loop approach. Furthermore, accumulated errors compared to the real trajectories must be weighted much stronger than in the approach described in Section 3.5.1, which uses independent data snippets.

The complexity of the considered scenario of approaching signalized urban intersections additionally places special requirements on the definition of the loss function used during the training of LSTM. Depending on the traffic volume and the signal state as well as the state changes of the traffic light, the scenario comprises a multitude of possible dynamic traffic situations the developed driver model is confronted with. In addition to almost unaffected trips passing a green traffic light, there are also highly dynamic situations with potentially multiple state changes and decelerations to a standstill. In order to cope with all potentially occurring situations, the loss function is modularly structured into three parts. The individual parts of the loss function are only active in certain traffic situations which occur during the training process. In addition, domain knowledge is used for a part of the situations to support the learning process of LSTM and to ensure a safer resulting driving behavior.

Equation (14) represents the loss function used for training LSTM in this work:

$$\text{loss} = \alpha \cdot \text{loss}_{eco} + \beta \cdot \text{loss}_{cog} + \gamma \cdot \text{loss}_{res}. \tag{14}$$

The factors $\alpha = 1$, $\beta = 3$, $\gamma = 3$ preceding the individual terms of the loss function are used to weight the individual aspects which are addressed by the terms. The values of the weights were determined iteratively. The individual components of the loss function are presented in detail below.

The first so-called ecological term $loss_{eco}$ describes the Mean Squared Error (MSE) of the predicted acceleration compared to the observed acceleration. Thus, $loss_{eco}$ results in

$$loss_{eco}= \frac{1}{n} \cdot \sum_{i=1}^{n} (a_i - \hat{a}_i)^2, \tag{15}$$

with the observed acceleration a_i at time step i, the predicted acceleration $\hat{a}_i$ and the batch size n. The ecological term is intended to prevent strong acceleration variations compared to the observed behavior. The goal is to set a realistic acceleration resistance that results in a similar theoretical fuel consumption. $loss_{eco}$ is used in an identical way for training the driver models for car-following and free driving.

The second term of the loss function is active depending on the driving state. The cognitive term $loss_{cog}$ represents the reaction of the following vehicle to movements of the leading vehicle with respect to keeping an average realistic time headway. For free driving, the term is 0 by definition. At slower speeds, where the minimum value of the speed falls below 2 m/s during car-following, $loss_{cog}$ is determined for the entire batch size as a function of the distance to the leading vehicle as follows:

$$loss_{cog,v\_low} = \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{|\Delta x_i - \Delta \hat{x}_i|}{\Delta x_i}. \tag{16}$$

Δx_i represents the real observed distance and $\Delta \hat{x}_i$ represents the predicted distance at time step i.

During car-following in higher speed ranges of the following vehicle of consistently more than 2 m/s, another term becomes active. For this case, $loss_{cog}$ is calculated using the following equation (17):

$$loss_{cog,v\_high} = \frac{1}{N} \cdot \sum_{j=0}^{N-1} \left(\int_{j \cdot t_{osc}}^{(j+1) \cdot t_{osc}} THW_j - \widehat{THW}_j \, dt \right), \tag{17}$$

with the observed time headway THW_j at time interval j, the time headway resulting from the prediction of the driver model $\widehat{THW}_j$, the oscillation duration t_{osc} and the number of time intervals N. Individual time steps Δt in equation (17) are aggregated to time intervals with oscillation period t_{osc}. This is done in order to compare an average deviation over a time interval j rather than individual THW values at each time step. The time aggregation takes into account that even the real human driver cannot adjust THW at intervals of Δt = 0.1 s. In addition, even the same driver in an identical situation will behave slightly differently depending on many influences. Consequently, LSTM should not learn to adapt the time gap ideally to a given real human trajectory at each time step. Instead, the time gap should be reproduced as accurately as possible on average over an oscillation period t_{osc}. The considered oscillation period t_{osc} is defined by:

$$t_{osc} = \Delta t \cdot n_{osc}. \tag{18}$$

The number of aggregated time steps n_{osc} is set to 10 in this thesis. Using a time step Δt = 0.1 s, the aggregated time interval has a duration of t_{osc} = 1 s.

Figure 3.13 shows an exemplary comparison between the real THW curve in black and the curve predicted by LSTM, $\widehat{THW}$, which is shown with a dashed blue line. The figure shows three oscillation periods t_{osc}.

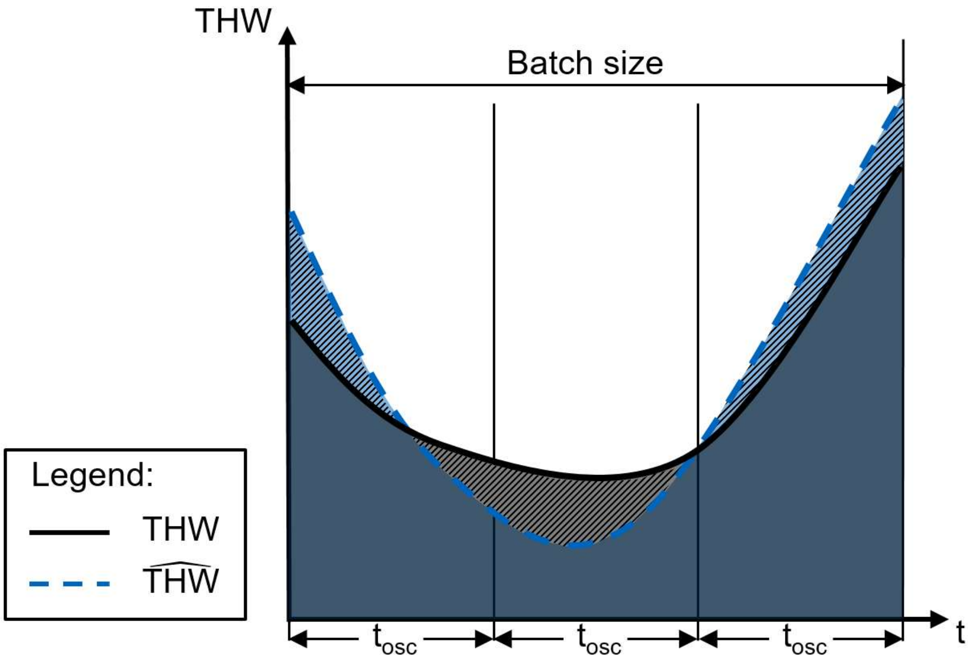

Figure 3.13: Exemplary graphical representation of the term $loss_{cog,v\_high}$ of the loss function. The mean area deviation of the hatched areas of the three intervals represents the term.

In Figure 3.13, the respective area below the graph of THW is shown in gray and the area below $\widehat{THW}$ is drawn in blue. The overlapping areas are shown in dark blue. All area deviations between both graphs are hatched. For the depicted example, the mean area deviation of both curves over the three intervals gives the cognitive term $loss_{cog,v\_high}$ of the loss function.

The total $loss_{cog}$ is a function of the minimum of all velocities v_i which occur in the respective batch.

$$\text{loss}_{\text{cog}} = \begin{cases} \text{loss}_{\text{cog,v\_low}}\text{ , if}\min(v_i) \leq 2\text{ m/s.} \\ \text{loss}_{\text{cog,v\_high}}\text{ , if}\min(v_i) > 2\text{ m/s.} \end{cases} \tag{19}$$

The third component of the loss function loss_{res} addresses situations in which the driver approaches an obstacle in a very close distance. For the case of free driving, this obstacle represents a stop line at a red light. During car-following, a close approach to a leading vehicle is considered. The respective term describes the deceleration necessary to enable timely braking before the stop line or to prevent an imminent collision.

The term $\text{loss}_{\text{res,ff}}$ for free flowing becomes active as soon as the driver is 65 m (≈4.7 s when traveling at 13.9 m/s) away from a traffic light. The distance is selected so that the driver can come to a stop at the stop line with a comfortable deceleration of approx. 1.5 m/s^2. $\text{loss}_{\text{res,ff}}$ is calculated by:

$$\text{loss}_{\text{res,ff}} = \frac{1}{n} \cdot \sum_{i=1}^{n} (a_{\text{des,ff,i}} - \hat{a}_i)^2 , \tag{20}$$

with

$$a_{\text{des,ff,i}} = -\frac{v_i^2}{c_{\text{ff}} \cdot d_{\text{tl,i}}} . \tag{21}$$

$a_{\text{des,ff,i}}$ represents the desired deceleration, which is determined from the speed v_i , the distance to the stop line $d_{\text{tl,i}}$ and a coefficient c_{ff}. The coefficient c_{ff} is set to 1, which ensures a stronger deceleration than necessary to directly stop at the stop line as soon as $a_{\text{des,ff,i}}$ is reached. In equation (20), $\hat{a}_i$ represents the respective predicted acceleration value at time step i and n the batch size. Especially at the beginning of the training, the driver model may get into a situation where it crosses the stop line despite a red light due to accumulated errors in the prediction. However, the training data does not include situations, in which the driver approaches a red light at high speed so that braking with moderate deceleration is no longer possible. Therefore, a model without the term $\text{loss}_{\text{res,ff}}$ cannot retrieve a learned behavior. In addition, the terms loss_{eco} and loss_{cog} punish such misbehavior too little, since these terms are aimed at other effects of driving behavior. This circumstance is intercepted via $\text{loss}_{\text{res,ff}}$ by setting $a_{\text{des,ff,i}}$ to -5 m/s^2 at each time step in which the stop line is crossed while the traffic light is red. This use of domain knowledge creates a high penalty in the loss function for crossing the stop line at a red light analogous to a reinforcement learning approach. LSTM learns from this that it must avoid such a situation in order to minimize the loss function.

Similarly, the component $\text{loss}_{\text{res,cf}}$ for the car-following mode is calculated according to the following equation (22), as soon as TTC falls below 5 s:

$$\text{loss}_{\text{res,cf}} = \frac{1}{n} \cdot \sum_{i=1}^{n} (a_{\text{des,cf,i}} - \hat{a}_i)^2 , \tag{22}$$

with

$$a_{\text{des,cf,i}} = -\frac{\Delta v_i^{\,2}}{c_{cf} \cdot \Delta x_i} , \tag{23}$$

where Δv_i represents the relative speed and Δx_i the distance between the leading and the following vehicle. The parameter c_{cf} is set to a value of 1. In equation (23), values of c_{cf} less than 2 result in greater decelerations than would be required to avoid a collision. Consequently, $a_{\text{des,cf,i}}$ represents the deceleration that is greater than necessary to compensate for the relative speed to the leading vehicle and to avoid a potential collision. Analogous to the approach described for free driving, an additional penalty is introduced for the case of a collision occurring during training. For this case $a_{\text{des,cf,i}}$ is set to -5 m/s^2. LSTM learns from this to avoid collisions in order to minimize the loss function.

Finally, loss_{res} results from the components for free driving or car-following associated with the respective driving state.

$$\text{loss}_{\text{res}} = \begin{cases} \text{loss}_{\text{res,ff}} \text{ , in free flowing mode.} \\ \text{loss}_{\text{res,cf}} \text{ , in car-following mode.} \end{cases} \tag{24}$$

Table 3.1 gives an overview of the calculation of the individual terms of the loss function.

Term	Condition	Car-following	Free flowing
Ecological	Always active	$loss_{eco} = \frac{1}{n} \cdot \sum_{i=1}^{n} (a_i - \hat{a}_i)^2$	
Cognitive	$\min(v_i) \leq 2$ m/s	$loss_{cog,v\_low} = \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{\lvert\Delta x_i - \Delta\hat{x}_i\rvert}{\Delta x_i}$	-
	$\min(v_i) > 2$ m/s	$loss_{cog,v\_high} = \frac{1}{N} \cdot \sum_{j=0}^{N-1} \left(\int_{j \cdot t_{osc}}^{(j+1) \cdot t_{osc}} THW_j - \widehat{THW}_j \, dt \right)$	-
Restrictive	$d_{tl,l} \leq 65$ m $img_{r,l}$ = "red"	-	$loss_{res,ff} = \frac{1}{n} \cdot \sum_{i=1}^{n} (a_{des,ff,i} - \hat{a}_i)^2$
	$TTC_i \leq 5$ s $img_{r,l}$ = "red"	$loss_{res,cf} = \frac{1}{n} \cdot \sum_{i=1}^{n} (a_{des,cf,i} - \hat{a}_i)^2$	-
Total	Always active	$loss = \alpha \cdot loss_{eco} + \beta \cdot loss_{cog} + \gamma \cdot loss_{res}$	

Table 3.1: Overview of the calculation of the terms of the loss function; "-" means that there is no such term.

3.6 Performance Evaluation

Virtual testing of automated vehicles places high demands on the quality of the driver behavior models used in the vehicle environment. The human driver behavior must be reproduced realistically by a used model. In addition to the underlying machine-learning method used, the developed driver behavior model differs from conventional engineering models by the selected input parameters as well as by the integration of HFs. When selecting the input parameters, particular attention is paid to ensuring that the human driver can estimate them as well as possible. However, HFs, such as reaction time, pose challenges for LSTM that must be interpreted, resolved, and compensated for by LSTM to produce realistic behavior. In contrast to the developed model, current engineering models used in traffic simulation strongly idealize the human driving process and neglect important HFs. Additionally, these models use parameters such as acceleration or relative speed, which can be easily determined by vehicle sensors but are difficult for humans to assess. As a result, such models do not adequately represent important characteristics of human driving (cf. Section 2.4.2). In this section, the performance of the developed model is evaluated in comparison to two benchmark engineering models described in Section 2.4.1: the Krauss model and the modified FVD model. In the following, it is investigated whether the developed driver model, which takes HFs into account, can reproduce the behavior of the human driver better than typical models used in traffic simulation. Figure 3.14 illustrates the comparison process of both driver model types with respect to real observed data.

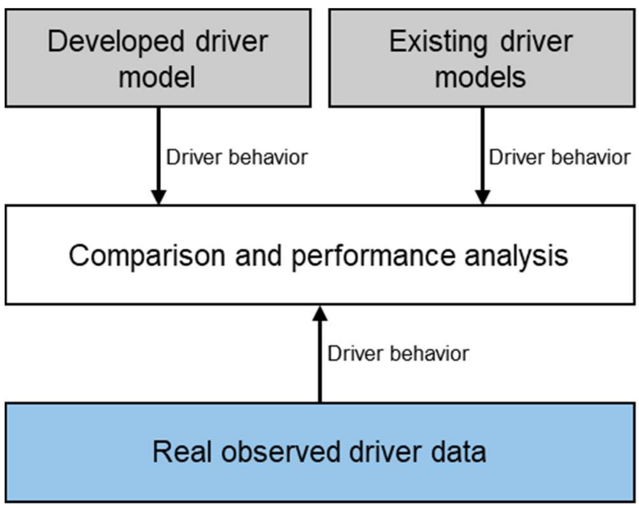

Figure 3.14: Comparison of the developed driver model with existing driver models and validation against real data.

The parameters of both engineering models are calibrated separately for each cluster on the respective training data set, analogous to the training of LSTM. In order to ensure a realistic behavior of the Krauss model, presented in Section 2.4.1, the minimum standstill distance s_0 is added as an additional calibratable parameter, analogous to the implementation in SUMO. The parameter s_0 reduces the actual distance s_n. As a result, a driver controlled by the model does not drive up to the bumper of the leading vehicle or to the stop line when decelerating to a standstill, but maintains a certain minimum distance. Due to the situation of the approach to a signalized intersection, the handling of the phase transition of the traffic light from green to yellow should be emphasized. In accordance with the logistic regression model presented in Section 3.3.2, a TTJ of 2.53 s at the transition to yellow is the threshold at which 50 % of the drivers still pass the traffic light, while the other half stops in front of the traffic light. Therefore, for modeling both engineering models, in the case of the yellow-light dilemma, for a TTJ of less than 2.53 s, the traffic light is assumed to be green, and for more than 2.53 s, the traffic light is assumed to be red. In the case of a red light, a virtual stationary vehicle is placed on the stop line to serve as an obstacle. For the calibration of all parameters, a genetic algorithm is used to find the optimal set of parameters for both models for each cluster. Equation (25) describes the fitness function to be minimized.

$$f = 0.5 \cdot NRMSE_{vel} + 0.5 \cdot NRMSE_{gap}, \tag{25}$$

with

$$NRMSE_{vel} = \frac{RMSE_{vel}}{\bar{v}} = \frac{\sqrt[2]{\frac{\sum_i^n (\hat{v}_i - v_i)^2}{n}}}{\bar{v}} \tag{26}$$

and

$$NRMSE_{gap} = \frac{RMSE_{gap}}{\Delta\bar{x}} = \frac{\sqrt[2]{\frac{\sum_i^n (\Delta\hat{x}_i - \Delta x_i)^2}{n}}}{\Delta\bar{x}}. \tag{27}$$

f describes the fitness value, $NRMSE_{vel}$ the Normalized Root Mean Square Error of all velocity values $\hat{v}_i$ predicted by the respective model with respect to the real observed values v_i. $\bar{v}$ represents the mean value of all real velocity values. Similarly, $NRMSE_{gap}$ is calculated from the deviations of the simulated distances to the leading vehicle $\Delta\hat{x}_i$ and the real distances Δx_i in relation to the mean value $\Delta\bar{x}$.

Table 3.2 summarizes the calibrated parameters of the Krauss model for all four driver type clusters (cf. the equations in Section 2.4.1). For simplicity, the index n of the current vehicle is omitted from the table.

Cluster	Parameter				
	s_0 [m]	τ_k [s]	a_{max} [m/s^2]	b_{des} [m/s^2]	v_{des} [m/s]
0 (*Moderate*)	2.20	1.20	1.43	1.99	13.26
1 (*Speeder*)	1.01	1.02	0.74	3.50	15.41
2 (*Timid*)	3.83	1.79	0.94	5.00	12.75
3 (*Aggressive*)	1.89	1.49	1.71	3.72	12.38

Table 3.2: Calibrated parameters of Krauss model; units in parenthesis.

The parameters of the Krauss model allow a simple interpretation of the individual values. The numerical values of the calibrated parameters of the *Moderate* driver group lie between the values of the parameters of the other groups without strong characteristics. Only the maximum deceleration b_{des} shows the lowest value of all clusters. The *Speeder* group is characterized by the highest value of the desired speed. While the *Timid drivers* show the largest minimum distance s_0 as well as the largest reaction time τ_k, the value of the maximum deceleration b_{des} represents the highest of all values. Finally, the group of *Aggressive drivers* is characterized by comparatively high values in the maximum acceleration a_{max} and maximum deceleration b_{des}.

Table 3.3 shows the calibrated parameters of FVD.

Cluster	Parameter											
	κ_r [1/s]	λ_r [1/s]	V_{1r} [m/s]	V_{2r} [m/s]	C_{1r} [1/m]	C_{2r} [-]	κ_g [1/s]	λ_g [1/s]	V_{1g} [m/s]	V_{2g} [m/s]	C_{1g} [1/m]	C_{2g} [-]
0 (*Moderate*)	0.38	0.44	4.16	11.73	0.14	0.90	0.31	0.49	4.75	10.68	0.09	0.69
1 (*Speeder*)	0.50	0.50	6.83	12.69	0.20	0.88	0.31	0.30	4.19	10.34	0.20	0.97
2 (*Timid*)	0.48	0.55	6.66	14.37	0.07	0.70	0.38	0.56	4.19	10.12	0.09	0.58
3 (*Aggressive*)	0.38	0.37	5.8	11.47	0.07	0.67	0.38	0.55	4.26	10.43	0.11	0.55

Table 3.3: Calibrated parameters of FVD; units in parenthesis; [-] means no unit.

No characteristics in driving behavior can be derived from the calibrated parameters of FVD.

For the evaluation of the performance of the developed driver model, here called LSTM, compared to Krauss and FVD, the models are simulated using the trajectories of the test data set. During each simulation, the initial state is given by the test data. Then, the movements of a potential leading vehicle as well as map and traffic light information are taken from the real data. The response and interaction of the modeled vehicle are performed by the respective model. Analogous to the procedure for the engineering models, the yellow-light dilemma is explicitly modeled for the evaluation with a TTJ limit of 2.53 s (see above). In dilemma situations, not enough data are available so that LSTM can sufficiently learn the bifurcation caused by the dilemma. The evaluation criterion of the performance of the models is the ability to reproduce the real trajectories as accurately as possible. This ability is evaluated by the Root Mean Square Error (RMSE) distributions of acceleration, velocity and distance to the leading vehicle. Figure 3.15 presents the results of the evaluation of the performance of the three models for all four clusters on the respective test data set as boxplots. Outliers are neglected in Figure 3.15.

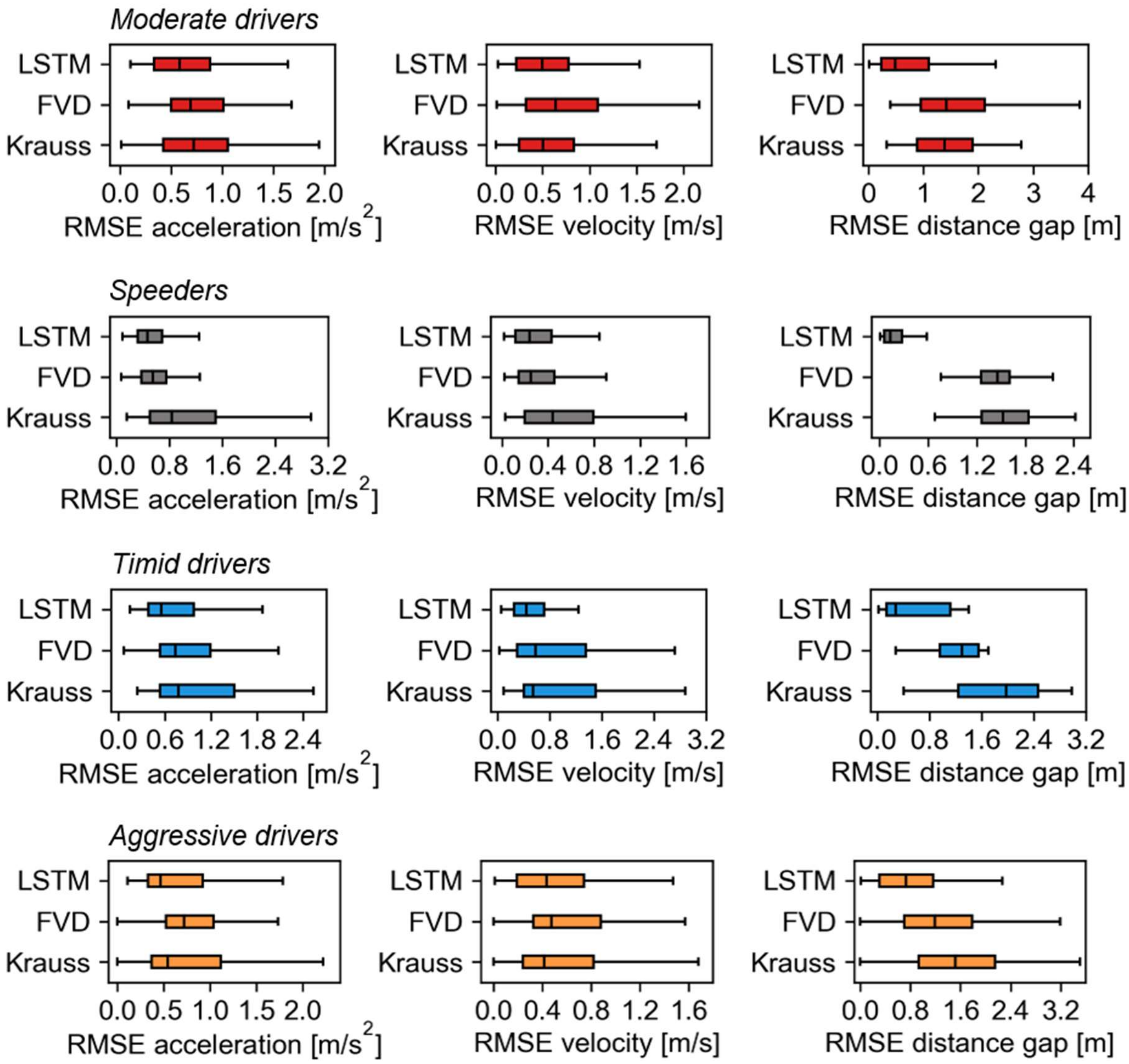

Figure 3.15: RMSE distributions for acceleration, velocity and distance errors for LSTM, FVD model and Krauss model for all four clustered driver types.

The RMSE distributions in Figure 3.15 indicate that LSTM performs at least comparably and mostly even better than the engineering models despite the input data being delayed by PRT. The evaluations of the errors of the *Moderate drivers*, the *Timid drivers* and *Aggressive drivers* reveal in all distributions both lower median errors and a lower variance of the errors. For the *Speeders*, the median values of FVD in the distributions of the RMSE values of acceleration and velocity are equal to those of LSTM. In contrast, the RMSE values of the distance behavior of LSTM stand out. The error distribution shows significantly lower errors with respect to the distributions of the compared engineering models. There is a general trend across all clusters that LSTM sets distances more realistically than FVD and Krauss. This suggests that LSTM outperforms both engineering models in terms of reproducing real trajectories, especially in car-following mode, for which FVD and Krauss were originally intended. Performance differences are evident between FVD and Krauss depending on the cluster. While FVD particularly stands out in the *Speeders* group, Krauss shows comparable performance to LSTM in reproducing acceleration and speed for the *Aggressive drivers*. Nevertheless, Krauss shows the widest distribution in ten of the twelve evaluations. The differences in the error distributions of the engineering models can be attributed in particular to the fact that the considered scenario of approaching a signalized intersection involves a large number of possible situations, depending in particular on the presence of a leading vehicle as well as on the traffic light state. While Krauss uses one parameter set in all possible situations, the extended FVD can access two adjusted parameter sets depending on the traffic light state. LSTM shows the greatest adaptability to the challenges posed across all simulated trajectories. The performance of FVD and Krauss could be improved by using additional parameter sets for different situations.

In particular, to investigate situations that are strongly influenced by the traffic light, the performance of the three compared models is additionally evaluated on a subset of trajectories from the test data set in which the traffic light performs at least one signal state change. Due to the phase transition and the associated immediate change in conditions, this subset poses a potential challenge to the models. In particular, LSTM must be able to recognize the changing requirements associated with the phase transition and must have learned an appropriate response to handle such a situation. Figure 3.16 shows the error distributions of the three models for the subset of trajectories in the same categories as Figure 3.15. The distance gap distributions of the *Speeders* and the *Timid drivers* contain only a statistically non-significant set of trajectories and are included here only for completeness. These two evaluations are therefore not considered in the following.

Figure 3.16 shows in comparison to Figure 3.15 that the absolute errors of velocity and acceleration for LSTM and Krauss decrease for the subset, while FVD shows similar error distributions like in the complete data set. It is noteworthy that despite switching between different sets of parameters, FVD shows a worse performance than Krauss in situations characterized by the influence of the traffic light. The error distributions of Krauss are significantly better compared to the complete data set and even outperform LSTM, e.g., for the

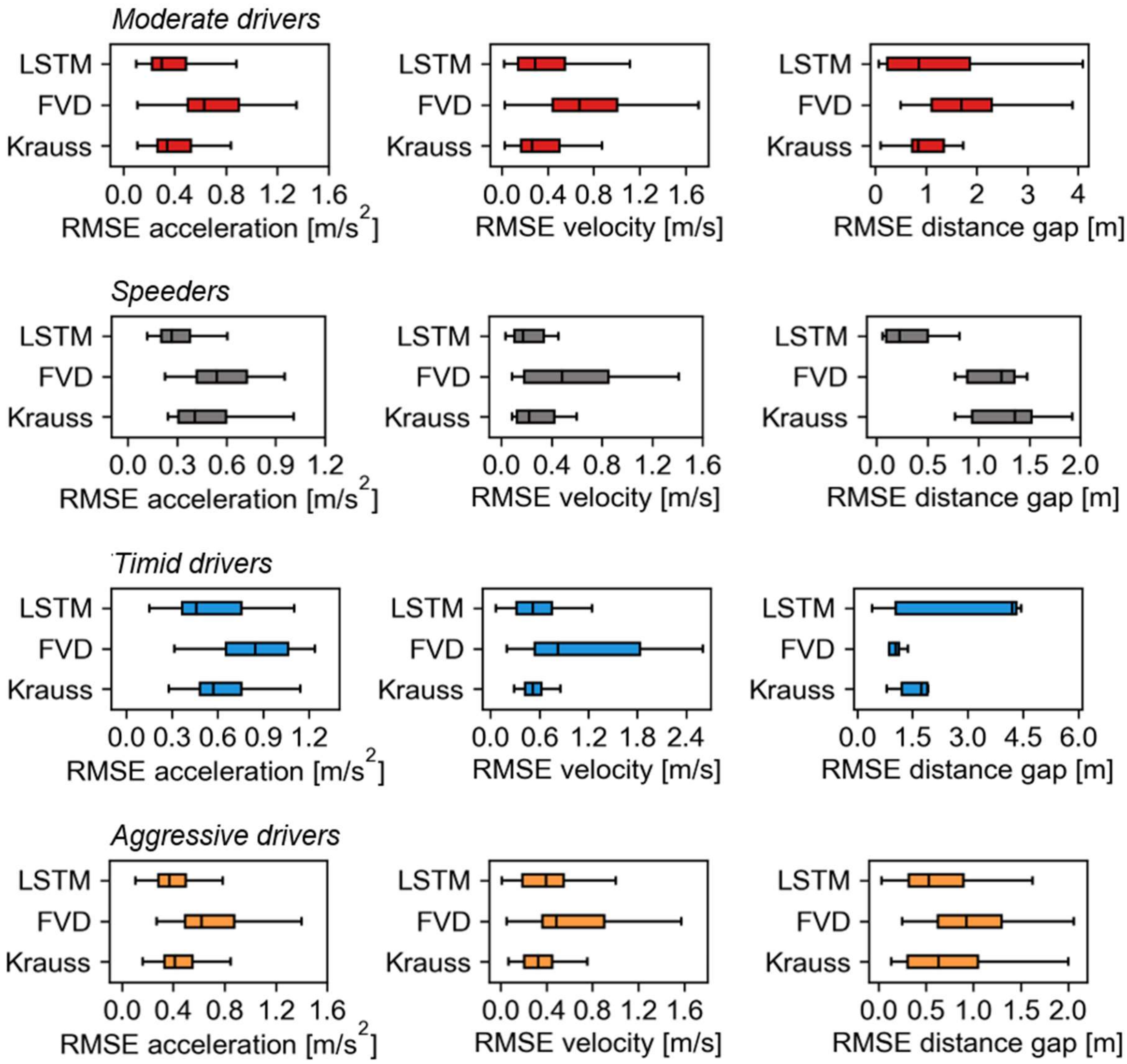

Figure 3.16: RMSE distributions for acceleration, velocity and distance errors for the LSTM, FVD model and Krauss model for all four clustered driver types. Only situations with a signal change of the traffic light are considered. *Speeders* and *Timid drivers* do not provide sufficient data for a valid evaluation.

velocity of *Aggressive drivers*. However, LSTM shows better median values than the engineering models for the majority of the shown distributions. The distance error distribution of the *Moderate drivers* additionally includes values at the right margin, which are not shown as outliers in Figure 3.15. For this reason, the distance distribution in Figure 3.16 is wider. Nevertheless, the median value is lower than that of the two compared engineering models.

The shown evaluations in Figure 3.15 and Figure 3.16 stress the superiority of the data-driven modeling of the driver behavior compared to the engineering models. It is remarkable that despite the imperfect perception as well as the delay by PRT, LSTM performs mostly better than the engineering models, which have ideal perception and do not need to compensate for reaction time. This also implicitly allows the model to cover surprising situations, such as a phase change of the traffic light which results in a delayed reaction. This effect has a significant influence in the study of traffic safety. Both engineering models are unable to replicate these characteristics due to their direct response behavior to a sudden stimulus. The generalizability of the developed driver model can be further increased by incorporating additional data covering other traffic situations besides the situation considered here. Furthermore, the ability to learn a plausible decision making in the yellow-light dilemma is possible with additional data. In contrast, the data-driven model requires more computing power. This results from the pre-processing of the delayed input data as a sequence which includes the projection of the current driving state into the near future. However, an increasing simulation time can be counteracted by more powerful simulation computers. Furthermore, in a virtual test of automated vehicles, only relevant vehicles in the immediate proximity of a VuT can be replaced by the LSTM-based model, while peripheral vehicles are controlled by an engineering model that applies simpler computational algorithms.

4. Traffic Model of Ingolstadt

Valid virtual tests of automated vehicles place high demands on the simulation environment used for the respective test. For the simulation of real-world traffic conditions, it is necessary to set up a traffic simulation in which spatially and temporally realistic traffic can be provided for the respective test. During the research project SAVe: [BUNDESMINISTERIUM FÜR DIGITALES UND VERKEHR, 2018] and its successor project SAVeNoW [BUNDESMINISTERIUM FÜR DIGITALES UND VERKEHR, 2021], a traffic model of the German city of Ingolstadt has been developed based on real data using the traffic flow simulation tool SUMO [LOPEZ et al., 2018]. The main purpose of the traffic model is to serve as a virtual test field for the validation of automated vehicles in realistic urban traffic. The replication of real-world traffic implies the essential advantage of the comparability of the results of simulation and reality. Chapter 4 describes the calibration process of the traffic model. The objective is to hourly reproduce real past traffic states in the simulation. In order to demonstrate the usability of the developed driver behavior model within a whole traffic system in an exemplary virtual test, the model is embedded into the traffic model, see Chapter 5. Subsequently, mixed traffic conditions of automated and manually controlled vehicles are simulated and evaluated with respect to traffic safety and efficiency.

In the following, the setup, calibration and validation of the Ingolstadt traffic model is described. Basically, the process can be divided into the steps map association, traffic light emulation and traffic calibration. These steps are explained in more detail in the following sections. Figure 4.1 shows the calibration process schematically.

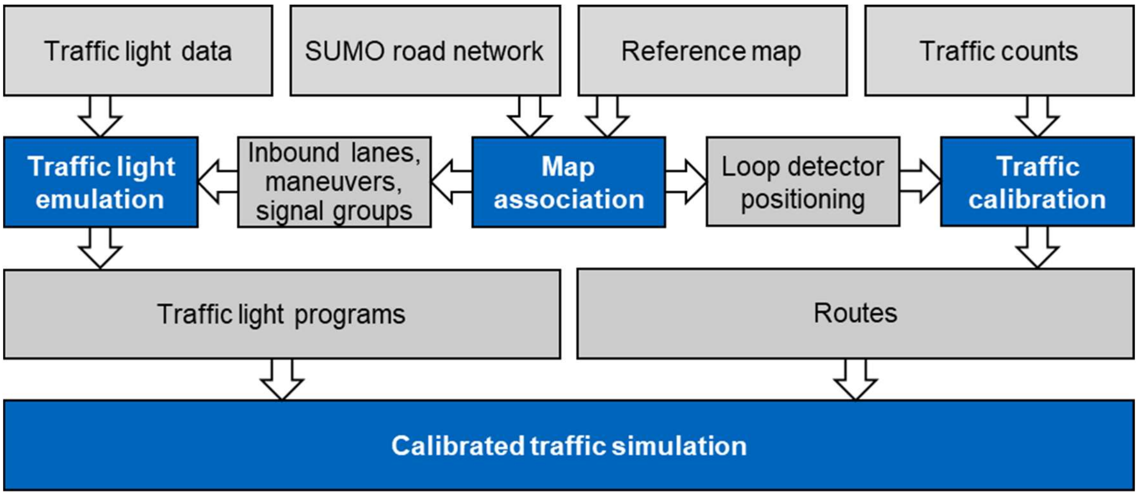

Figure 4.1: Calibration procedure of the traffic simulation according to HARTH, LANGER, et al. [2022].

The creation of the model has already been published and presented in [HARTH, LANGER, et al., 2021] and [HARTH, LANGER, et al., 2022].

4.1 Map Association

The traffic simulation in SUMO is based on the road network of the entire city of Ingolstadt. The dimensions of the considered map section are about 10 km by 10 km. The network imported from OpenStreetMap contains more than 12000 edges covering urban and rural roads as well as highways and 5600 intersections with 120 signal controlled intersections. Figure 4.2 shows a partial section of the network used.

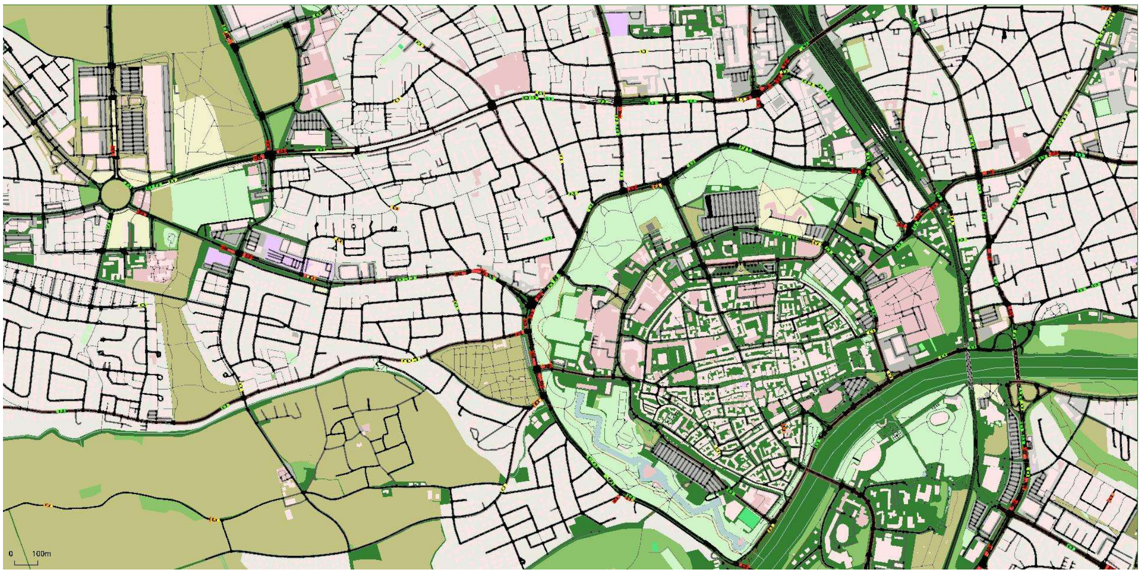

Figure 4.2: Network section in the traffic flow simulation of Ingolstadt in SUMO.

The essential process for the calibration of the traffic simulation is the association of real observations, like traffic counts as well as switching states of traffic lights, and their representations in the simulation. Only by precisely positioning of the real observations in the virtual world, it is possible to calibrate realistic traffic conditions including realistic traffic flows and traffic light programs. In order to locate the real observations in the simulation network, the network is first transformed into a GeoDataFrame based on the Python library GeoPandas [NUMFOCUS, 2024]. This format can be used to perform map matching tasks by not only applying a geometrical matching. Instead, logical connections and properties of the map are taken into account, such as subsequent or previous lanes and lane types.

In order to transfer the real signal states of the traffic lights into the simulation, the real signal groups must be assigned to the incoming lanes and to the subsequent internal connecting lanes of the signalized intersections in the simulation. Signal groups bundle the signal status of one or more associated incoming or internal connecting lanes. For the locally and temporally accurate mapping of traffic flows, traffic counts from detector loops in the network must also be located on the respective lanes. For the association of signals groups and detector counts, a reference topology, also called “reference map” in the following, is provided by the traffic light service company in Ingolstadt. The reference topology provides information about all incoming lanes, the associated signal groups and possible maneuvers across the intersection. Additionally, all geometrical positions of the loop detectors are available. The information from the reference topology is used to transfer the real-world information into the simulation.

The assignment of signal groups into the simulated network is done in several steps. In a first step, signalized intersections from SUMO have to be identified in the reference topology. This is done by comparing the geometric centers of the respective intersection taking into account small deviations between the two maps. While SUMO stores the assignment of traffic light signals as an attribute of the internal lanes within intersections, the reference topology assigns the signal groups to the incoming lanes. Therefore, in order to transfer the real signal groups to SUMO, the incoming lanes of SUMO must first be matched with those of the reference topology. The requirement for a valid matching is that the number of incoming roads and lanes from each signalized intersection must be identical for both maps. However, due to geometric deviations between the reference topology and the simulation network, the respective lanes of both maps do not match exactly. Therefore, in two further steps, the incoming roads of both maps must first be paired and then the lanes within each pair of roads must be matched.

In the SUMO network, the unique identifiers of each incoming road and lane imply the respective numbers of roads and lanes. The numbers can be compared to numbers in the reference topology. Discrepancies in the number of incoming roads and lanes from both maps are logged and require manual rework of the network. To determine the corresponding incoming roads of both maps, the aggregated center points of the stop lines of all lanes of each incoming road are compared. The corresponding incoming roads are determined by minimizing the Euclidean distance between the geometric centers of each pair of incoming roads of both maps. Finally, for pairing the incoming lanes, the total distance between all lanes of each pair of roads of both maps is minimized. The total distance serves as a quality parameter of the matching procedure to achieve an optimal matching result for all lanes. Figure 4.3 shows the results of matching roads and lanes for an example intersection of the SUMO network and the reference topology.

The symbolic markers in Figure 4.3 represent the center points of the stop lines of the incoming lanes of both maps. The markers on the outside are derived from the reference topology and the markers on the inside are taken from the SUMO network. Lanes of one road are shown in the same color. In addition, identical symbols with the same color imply related lanes between both maps.

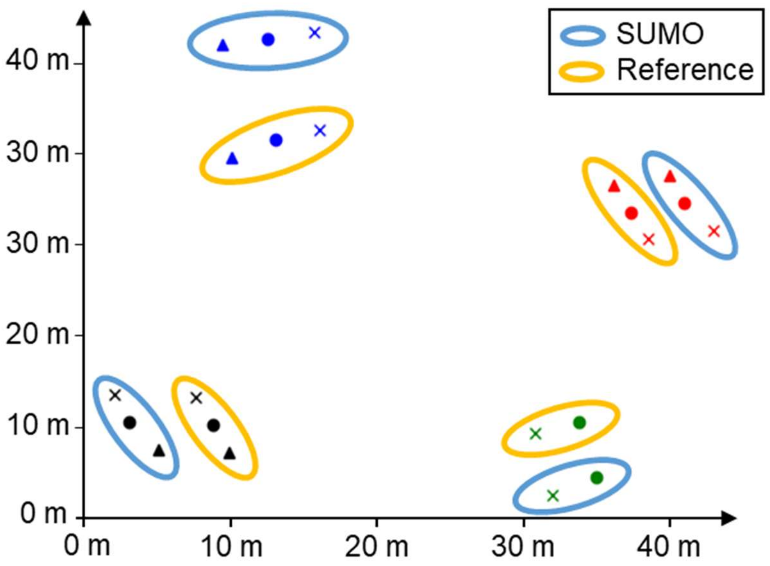

Figure 4.3: Exemplary result of the matching of roads and lanes [HARTH, LANGER, et al., 2022].

The method for the geographic positioning of the loop detectors also applies grouping and pairing. First, the detectors are grouped based on their incoming road in the SUMO network. Then, the entire group of detectors is associated while minimizing the total distance to the nearest incoming lanes. While the matching of inbound roads and lanes between the reference topology and the SUMO network requires identical numbers of roads and lanes, detectors are matched even if the number of detectors is different from the number of incoming lanes. If there are more detectors than lanes, it is also possible to place several detectors on the same lane depending on their geometric position.

4.2 Traffic Light Emulation

Especially in urban traffic, there is a large influence of traffic control on the traffic flow within a network. In order to take this influence into account in the simulation, it is essential to emulate real traffic light programs. In Ingolstadt, traffic lights adapt the green times to the traffic flow and use public transport prioritization. Instead of emulating the exact algorithms, for the Ingolstadt simulation, a fixed-time logic for the traffic signal programs is implemented for each hour that is based on the mean real green times of the real traffic lights. As the simulated traffic does not necessarily have to match the real traffic exactly, a simple replay of the real traffic light programs from the historical data is omitted. The mechanisms of adaptive green time extension and public transport prioritization would otherwise possibly take effect in inappropriate situations. Instead, the fixed time logic for each hour is determined from real historical traffic light status data and implemented in the simulation. As a result, the influences of green time extension and public transportation prioritization are neglected, but their manifestation is expressed in the mean green time. The red and green phases of each signal group of 75 signalized intersections in Ingolstadt can be queried via an interface to a traffic light backend of the city of Ingolstadt with a resolution of 1 second. For the other traffic signals in the network for which no data is available, a fixed time program is manually created based on empirical values that can handle the prevailing traffic demand.

A traffic light program is a continuously repeating circulation consisting of the temporal sequence of green and red times of the different signal groups, which may partially or completely overlap. Both the sequence of the signal groups and their average green time have to be simulated for the determination of the realistic traffic light programs. In a first step, the data of the considered one-hour time interval are examined with respect to the first signal group changing to green, which is defined as reference signal group in the following. This represents the starting point of a traffic light program cycle. On this basis, the most common switching sequence of the signal groups occurring in the data is determined. The most common sequence on the one hand starts with the reference signal group and on the other hand requires that all signal groups have switched to green at least once. Based on the average green and red time and the total green time within the considered time interval, the parameters of the traffic light program in the simulation are determined. First, a cycle time $t_{cycle,i}$ is determined separately for each signal group i depending on the number of occurrences n_i in the most common switching sequence, the mean green time $\bar{t}_{g,i}$ and the ratio of green time to the duration of the considered one-hour interval $p_{total,i}$.

$$t_{cycle,i} = \frac{n_i \cdot \bar{t}_{g,i}}{p_{total,i}} \tag{28}$$

The total cycle time t_{cycle} of the generated traffic light program is calculated from the individual cycle times $t_{cycle,i}$ weighted by $p_{total,i}$.

$$t_{cycle} = \frac{\sum (t_{cycle,i} \cdot p_{total,i})}{\sum p_{total,i}} \tag{29}$$

The weights stress the relevance of each signal group in the whole traffic light cycle depending on the ratio of green time in the considered one-hour interval. In order to match the determined cycle time t_{cycle}, the emulated green times $\tilde{t}_{g,i}$ of each signal group are adjusted based on their average green times $\bar{t}_{g,i}$ and red times $\bar{t}_{r,i}$.

$$\tilde{t}_{g,i} = \frac{\bar{t}_{g,i}}{(\bar{t}_{g,i} + \bar{t}_{r,i})} \cdot t_{cycle} \tag{30}$$

Finally, the green times are arranged in the cycle based on the mostly common switching order as well as the mean overlap of the observed green times. Between two consecutive signal groups that have no overlap with each other, an additional clearing time of 6 s is considered. Yellow phases are not part of the recorded traffic light data, but are considered as part of the red phases. The duration of the yellow phases is determined according to the German legislation depending on the speed limit.

The integration of the emulated traffic light programs in SUMO uses the methods described in Section 4.1 to assign the signal groups to the connecting lanes within the intersection. In addition, the possible maneuvers associated with the signal groups are compared with those of the connecting lanes in SUMO to achieve the assignment of a realistic signal control of an intersection. The final traffic signal programs resulting from the determined temporal sequences as well as the spatial assignment are stored in an XML format readable by SUMO.

4.3 Traffic Assignment

The calibration of traffic flows in the simulation is performed using a two-step methodology. First, a pool of routes is initialized from origin-destination (OD) dependencies based on statistical data. Then, this pool is refined using real traffic counts from Ingolstadt. In the first step, OD matrices are determined based on various statistical information. These matrices reflect the traffic condition in Ingolstadt at hourly intervals. The underlying statistics include distributions of inhabitants and jobs in different city districts as well as data on the attractiveness of city districts in terms of leisure activities, shopping facilities and educational institutions. In addition, aggregated data is available from AUDI AG, which provides information on the residence locations and working time models of all 44,000 Audi employees in Ingolstadt. Due to the numerical relevance of the employer AUDI AG in the Ingolstadt area, these data provide valuable information on the spatial and temporal distribution of trips with the purpose of work. For the temporal and spatial distribution of trips which are not related to the trip purpose work, as well as the modal distribution, no exact data collection exists for Ingolstadt. Therefore, Germany-wide statistics are used to estimate the temporal and spatial distributions of these trips as well as the modal distributions. The considered trip purposes are work, business, leisure, shopping and education. In order to be able to capture traffic flows beyond the boundaries of the network, commuter statistics regarding the surrounding counties are also taken into account.

For the generation of the OD matrices, a procedure that takes all of this information into account is applied. Each individual trip is created based on the source city or county district, the intended trip purpose, the temporal distribution depending on the trip purpose, and the length depending on the modal distribution. Then, the created trips are assigned to a destination city or county district according to their length as well as the attractiveness of the districts for each trip purpose. The modal distribution is divided into vehicles, public transport, bicycles, and pedestrians. From the temporally distributed OD relationships, hourly resolved OD matrices are generated for an average workday and for each modality type. Using the SUMO built-in tools od2trips and duarouter, routes for the network are generated from the OD matrices. Due to statistical imperfections as well as daily variations in traffic characteristics, the pool of routes generated from the OD matrices only approximately represents real traffic patterns, but does not reflect daily fluctuations. Therefore, this pool has to be further refined in a second step based on real observed traffic counts. Since only traffic counts of road vehicles are available, the calibration described in the following is performed only for the routes of the vehicles.

For the refinement, the traffic counts of each road, on which induction loops are located, are aggregated into one value for each of the lanes. In addition, the routes determined from the statistical data are evaluated with respect to the number of crossings over these roads. For the further calibration process, all routes in the pool are first copied twice and then successively removed until a state is reached in which the number of crossings and the real traffic count values match as closely as possible. The evaluation metric for the calibration is the difference between the number of crossings and the real traffic counts for each road. The calibration is performed in two steps. In the first step, only routes are removed from the pool that exclusively pass roads with an overestimation of crossings compared to the real traffic counts. This removal is repeated until no route remaining in the pool meets the criterion. In the subsequent second step, additional routes are removed whose removal has a positive effect on the overall error. This step is performed iteratively until any removal of a route would increase the total error. The procedure leads to a balancing of over- and underestimation of the error across all roads in the network. A final pool of routes emerges from the calibration, containing between one and three copies of each route originally derived from the statistical data. Finally, the routes are temporally shuffled for each hour to prevent copied routes from starting at the same time.

4.4 Evaluation of the Calibrated Traffic Simulation

For the evaluation of map association, traffic light emulation and traffic assignment, simulated observations are compared to the corresponding real-world measures. The evaluation presented in the following Sections 4.4.1 – 4.4.3 quantifies the quality of each part of the proposed calibration procedure.

4.4.1 Evaluation of Map Association

To verify the assignment of signal groups to the correct internal lanes from SUMO, both the correct assignment of incoming roads and lanes from the reference map and SUMO network as well as the matching of feasible maneuvers are verified at all considered 75 signalized intersections. After minor rework at 18 intersections in the SUMO network due to geographic deviations from the reference map as well as differences in the intersection logic, all considered intersections are successfully matched.

In contrast to the assignment of signal groups, detectors are also matched if the number of incoming lanes is different from the number of detectors. This leads to more errors in the matching process due to the geometrical deviations between the reference map and the SUMO network. All available 529 detectors can be assigned to a lane in the network. Of these detectors, 319 are correctly placed, while the positioning of 210 detectors needs to be readjusted manually. As long as the number of detectors matches the number of incoming lanes and a separate lane is provided for each detector, a corresponding group of detectors is always correctly assigned. In case of different numbers of detectors and lanes, a corresponding group is assigned to the next set of incoming lanes according to the minimization of the total distance. Thus, whole groups can be placed shifted by one or more lanes.

4.4.2 Evaluation of Traffic Light Emulation

While the emulated traffic light programs have a fixed time logic, the real traffic lights in Ingolstadt are controlled traffic-adaptively. The influences of the traffic-adaptive control as well as the public transport prioritization are implicitly considered by the influence of these effects on the average green times. In the validation of the reconstruction of the traffic signal programs in the simulation, a direct second-by-second comparison between the recreated and simulated traffic signal program is not useful. Instead, the green times of the individual signal groups are compared within one-hour intervals. For this comparison, the share of the green time in the total cycle time of the respective traffic light program is determined for each signal group for reality and simulation. Since deviations of the cycle times as well as varying switching behavior occur in reality, the real green time ratio is calculated over the entire one-hour interval. The presented evaluation of the traffic light programs follows the assumption that similar green time ratios have a comparable influence on traffic.

The quality of the emulation of the traffic light programs is evaluated below for a selected week from Monday to Friday. Figure 4.4 shows the distributions of the deviations of the green time ratios from simulated in comparison to real signal groups. The evaluation is aggregated for all weekdays in the considered one-hour intervals from 7 to 8 a.m., 9 to 10 a.m., 4 to 5 p.m. and 6 to 7 p.m.

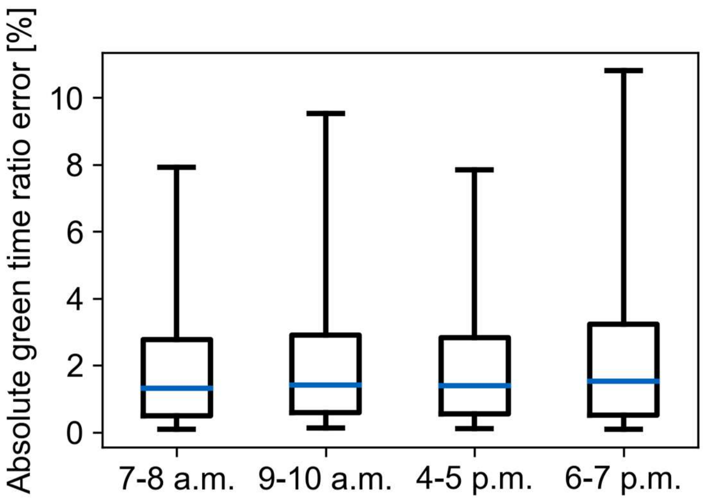

Figure 4.4: Absolute errors in green time ratio in percent [HARTH, LANGER, et al., 2022].

While the error distributions in the peak traffic hours of 7 to 8 a.m. and 4 to 5 p.m. vary between 0 and 10 %, the error distributions for the intervals with lower traffic volumes between 9 and 10 a.m. and 6 and 7 p.m. show a slightly wider range. Due to the lower traffic volume, the proportion of public transport is larger in these intervals, resulting in a larger influence on the green time duration as well as the switching order. This results in a greater diversity of real traffic signal programs, which explains the wider distribution of errors. The median values of the distributions range from 1.6 % to 1.9 %. Both the width of the distributions and the median values suggest that the emulated traffic light programs show realistic behavior for the considered time intervals.

Instead of simply adjusting the traffic light logics to manage the calibrated traffic flows, the described realistic emulation of the traffic lights based on real data poses a particular challenge. An incorrect calibration of traffic volume or traffic light programs would immediately show up in the simulation, since an underestimation of the capacity at the traffic light controlled intersections or an overestimation of the traffic volume would lead to traffic jams that do not occur in reality. However, neither pronounced congestion nor deadlock situations show up in the calibrated simulation not even in peak traffic hours. The working interaction between traffic flows and emulated traffic lights is an indicator for a realistic representation of the traffic lights in Ingolstadt, in addition to the quantitative analysis of the green times.

4.4.3 Evaluation of Traffic Assignment

To evaluate the quality of the calibration of the traffic assignment, the real observed and the simulated traffic counts are compared. Since in reality, traffic volume is measured only at specific positions where an induction loop is located, the quality of the calibration can be evaluated only at these measurement points. The metric of evaluation is the absolute value of the relative error e_{rel} of the simulated traffic counts according to:

$$e_{rel} = \frac{|count_{real} - count_{sim}|}{count_{real}}. \tag{31}$$

$count_{real}$ represents the aggregated real traffic count target value for each road and $count_{sim}$ the simulated count. The individual detector values are aggregated for each road. The resulting error value is calculated for each of the 254 roads with detectors.

The traffic in the simulation is calibrated with traffic counts from a one-hour interval to match the simulated traffic counts within the same interval. Accordingly, the calibration results in routes that start within the interval but cannot be completed. This problem is especially apparent during periods of high traffic volume, which potentially leads to more congestion and thus incomplete trips.

In order to evaluate the underlying calibration methodology, the traffic counts obtained from the simulation are evaluated at an extended time interval so that all vehicles in the simulation can complete their trips. In Figure 4.5, the distributions of the relative errors are shown for the same week as for the evaluation of the traffic light programs from Monday to Friday in time intervals from 7 to 8 a.m., 9 to 10 a.m., 4 to 5 p.m. and 6 to 7 p.m.

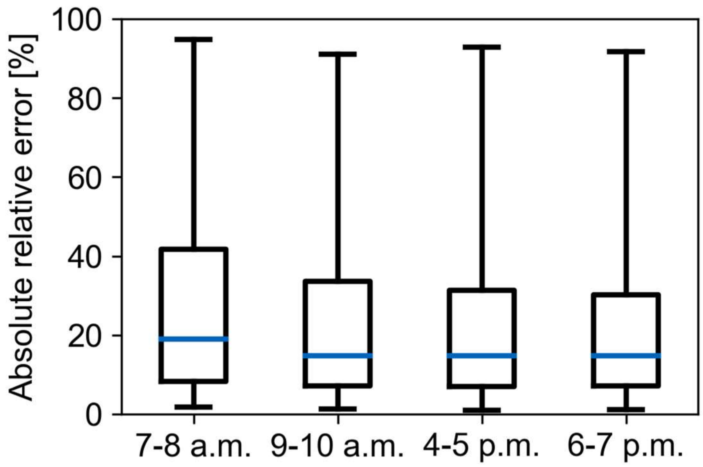

Figure 4.5: Distributions of the absolute value of the relative errors of the traffic counts after a clearing period [HARTH, LANGER, et al., 2022].

The time intervals with lower traffic volumes between 9 and 10 a.m. and 6 and 7 p.m. as well as the period from 4 to 5 p.m. with denser traffic show similar relative error distributions. In contrast, the distribution of errors between 7 and 8 a.m. shows a wider spread of errors. A similar range emerges for the median values, which vary between 13 % and 19 %. For the remaining deviations of the simulated count values, two main causes are identified: corrupt input data and rerouting. In the real traffic counts, a small number of unrealistically low target values can be observed, which strongly affect the calibration result. This is attributable to the goal of the calibration to balance overestimation and underestimation of routes passing the induction loops in order to achieve an overall error of zero. For example, a significantly underestimated target value at one road causes the calibration algorithm to overestimate the calibrated count value at that location. Simultaneously, counts at surrounding roads are underestimated to maintain the balance. However, higher absolute errors consequently arise at the individual measurement points. Furthermore, especially during hours with high traffic volumes, overestimation of traffic locally could lead to exceeding the capacity, causing congestion and deadlock situations. To counteract this problem caused by corrupt input data, the rerouting probability of SUMO vehicles is set to 20 %. This allows 20 % of the vehicles during the simulation to select a different route than originally intended to reach their destination. In comparison, in the work of LOBO et al. [2020] who also model Ingolstadt in SUMO, a rerouting probability four times larger is chosen to compensate for inaccuracies in the calibration as well as the input data.

The GEH analysis takes into account that absolute errors of low traffic counts lead to a higher percentage deviation than absolute errors of the same magnitude of higher traffic counts. The GEH analysis results in a non-linear weighting of the errors, so that at high traffic volumes the relative deviation determines the GEH value and at low traffic volumes the absolute deviation is more relevant. The GEH statistic results from

$$\mathrm{GEH} = \sqrt{\frac{2 \cdot (\mathrm{count}_{\mathrm{real}} - \mathrm{count}_{\mathrm{sim}})^2}{\mathrm{count}_{\mathrm{real}} + \mathrm{count}_{\mathrm{sim}}}} \,. \tag{32}$$

Figure 4.6 shows the distributions of the GEH statistics considering an extended time interval to let all vehicles finish their routes.

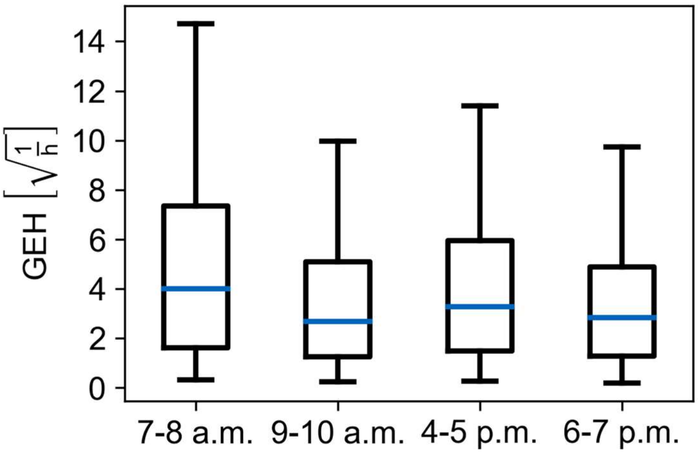

Figure 4.6: Distribution of GEH statistics of the traffic counts after a clearing period [HARTH, LANGER, et al., 2022].

Figure 4.6 shows broader distributions of the GEH value in intervals with higher traffic volumes from 7 to 8 a.m. and 4 to 5 p.m. than in the hours with less traffic. The median values resulting from the evaluation are between 2.4 and 4, indicating a good quality of the calibration.

To compensate for the incomplete routes in the evaluation, a simulation that includes a preceding one-hour warm-up phase instead of a subsequent clearing period is also evaluated. On the one hand, the warm-up phase provides initial conditions in which the simulated network is already filled with traffic at the beginning of the evaluation. On the other hand, routes that cannot be completed in the warm-up phase are originally calibrated for the previous hour and might not fully represent the current traffic conditions. Figure 4.7 shows the distributions of the absolute relative errors of the simulated traffic counts for the previously introduced time intervals considering a warm-up phase.

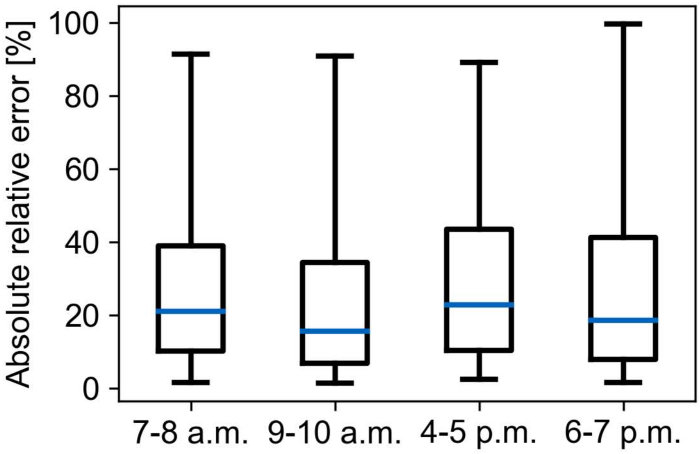

Figure 4.7: Distribution of the absolute value of the relative errors of the traffic counts considering a warm-up phase [HARTH, LANGER, et al., 2022].

Figure 4.7 shows similar deviations in intervals with high traffic volumes compared to hours with lower traffic volumes. Compared to Figure 4.5, in which a clearing period is taken into account, slightly higher errors are found in all distributions. The higher errors result from a quantitatively different traffic volume in the respective previous hour and thus contribute to a higher error. In particular, traffic between 9 and 10 a.m. shows a high similarity to the previous hour from 8 to 9 a.m., resulting in the lowest absolute observable errors. A partial overlap of routes beyond the respective calibrated time interval could contribute to an improvement of the presented methodology.

5. Empirical Assessment of the Driver Behavior Model and Application to Virtual Assessment of Automated Vehicles

In order to show the applicability of the developed driver modeling methodology described in Chapter 3 for virtual testing, the driver model was embedded in the traffic model of Ingolstadt presented in Chapter 4. Chapter 5 shows a potential use case for the developed driver model in the surrounding traffic of an endurance run for testing automated vehicles. For the integration of the driver model, a Python framework was implemented which enables the driver model to interact with the simulation. The structure of the framework is described in Section 5.1. Furthermore, various adaptations and extensions to the driver model were required for use in the simulation. These are explained in Section 5.2 in more detail. Since critical situations play an essential role in virtual tests of automated vehicles, the behavior of LSTM in these situations in the traffic simulation is highlighted in Section 5.3. The behavior is validated based on real observations.

The central content of this chapter is the simulation of mixed traffic of automated vehicles and human drivers as an example of a virtual endurance run of automated vehicles. In particular, it is investigated to what extent a driver type recognition procedure available to the automated vehicles influences traffic safety and efficiency. The procedure recognizes the respective driver type driving ahead and enables an adaptation of the driving behavior. The description of the implementation of the automated vehicles is given in Section 5.4, while different expansion stages of the recognition procedure of the driver types are described in Section 5.4.2. In Section 5.5, the quality of the recognition procedures is analyzed by simulation. Finally, Section 5.6 evaluates the benefits of the detection methods in terms of traffic safety and efficiency.

5.1 Software Implementation

The integration of the driver model as well as a model of automated vehicles into the simulation requires the implementation of a framework in which both models are embedded and can communicate bidirectionally with the simulation via an interface. For this coordination of the simulation, a Python component *SimulationManager* is introduced to perform the control of the simulation. Attached to the *SimulationManager* are other modules that obtain data from the simulation, control the instances of dynamic objects in the simulation, and finally record data to be evaluated. The whole framework uses the Python interface TraCI, which allows bidirectional data exchange between the traffic simulation in SUMO and the modules that the *SimulationManager* combines. Figure 5.1 shows an overview of the framework including the connection to the simulation.

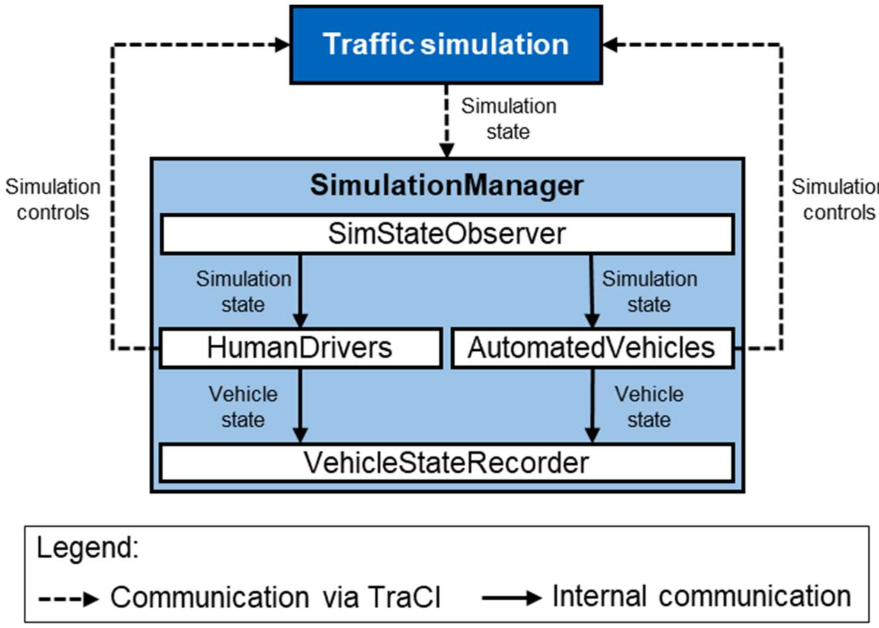

Figure 5.1: Overview of the simulation framework, its components and the communication. Dashed arrows indicate communication via the SUMO Python interface TraCI. Solid arrows mean internal communication.

Figure 5.1 shows the data exchange in the simulation framework via the TraCI interface. At each time step, the *SimulationManager* queries information about the simulation state from the running simulation. The information is processed by the four modules that the *SimulationManager* coordinates to control the next time step of the simulation. The first module is represented by the *SimStateObserver*. In this module, the static and dynamic state of all objects in the simulation, such as vehicles and traffic lights, is retrieved from the simulation state and made available to the other modules. The *HumanDrivers* module creates and manages the instances of human drivers in the simulation. Each human driver in the simulation is controlled by a separate instance of the developed LSTM-based model. *AutomatedVehicles* controls the instances of automated vehicles. These two modules and their integration into the framework are discussed in Sections 5.2 and 5.4 in more detail. Finally, the driving states of the human drivers and the automated vehicles are recorded in the *VehicleStateRecorder* and made usable for further evaluations of the simulation. In particular, data on the dynamic driving state of each vehicle, surrogate measures of safety regarding the relationships between the vehicles, and driver type recognition, which is described in Section 5.4.2, are captured.

5.2 Integration of the Developed Driver Behavior Model into the Traffic Model of Ingolstadt

In the framework described in Section 5.1, all instances of the developed driver behavior model are created and controlled by the *HumanDrivers* module. On the one hand, this integration into the framework requires the technical prerequisites for the data exchange with the simulation, which is described in the following Section 5.2.1 in more detail. On the other hand, adaptations of the driver behavior and of the incoming data are necessary for situations in the simulation, which the model could not or only insufficiently learn from the given training data. These are described in Section 5.2.2.

5.2.1 Technical Integration of the Developed Driver Behavior Model to the Simulation Framework

For the integration of the developed driver model, the data from the traffic simulation must first be retrieved at runtime and subsequently made available to the model in the required format. The *SimStateObserver* module is responsible for providing the data, which includes all relevant parameters for the application of the driver model in each time step for each vehicle in the simulation. The observed data include the own driving state, the current route, information about a potential leading vehicle, the geometry and the speed limit of the currently driven road as well as information about a traffic light ahead. This information is stored by the *SimStateObserver* over several time steps in a short-term memory. The stored data is the basis for generating the input sequence of LSTM, which includes the composition of historical and anticipated data (cf. Sections 3.2.1 and 3.2.2). The simulation runs with a time step of 0.1 s, which was also used for training LSTM. Since LSTM draws conclusions for the desired driver behavior especially from temporal trends in the input variables, an identical time step size of training and application is essential.

At each time step, LSTM predicts a desired acceleration from the processed input data. This is converted into a desired speed and passed on to SUMO via TraCI. In SUMO, the vehicle controlled by LSTM executes the desired speed in the next time step.

5.2.2 Adaptations to the Application of the Developed Driver Behavior Model

The traffic occurring in the Ingolstadt traffic model generates situations that are outside the application area of the model or are not covered by the training, e.g., because they are not included in the training data. In order to be able to cope with such influences, adjustments have to be made to the execution of the model, which apply in situations that the model cannot handle sufficiently.

The integration of the developed driver model into the traffic model in SUMO requires, in addition to the longitudinal control on the guidance and stabilization level (cf. model according to DONGES [1982] in Section 2.4), also the task of navigation as well as lateral control. The execution of both tasks, which are explicitly not covered by the driver model, is left to SUMO. The navigation task is largely determined by the routes obtained from the calibration of the simulation (cf. Section 4.3). It is already completed before the simulation starts. In addition, as described, a potential rerouting of 20 % of the vehicles is performed, so that these vehicles can change their route during the simulation to reduce travel time. SUMO performs the routing completely autonomously. Accordingly, no modifications are required for the routing when integrating the driver model. Furthermore, the default behavior of SUMO is maintained so that all vehicles obey the right-of-way rules at all right-of-way intersections. This means that a necessary braking response to a right-of-way to be granted is imposed on a vehicle by SUMO. In this case, all potential interventions by LSTM are overruled. In addition, a default SUMO setting is switched off, which prevents vehicles from passing through red traffic lights, sometimes with unphysically high decelerations.

Since the driver model is not able to represent the lateral behavior of the driver, this task is left to SUMO and adapted for the application of the driver model. At the strategic level, which concerns navigation, strategic lane changes are always allowed. Lane changes are performed using SUMO's default lane change model, which performs an immediate change from one lane to the other. To prevent situations where slow leading vehicles perform unnecessary lane changes and force other vehicles behind them to break heavily, lane changes are prohibited at speeds below 5 m/s via the lane change behavior settings in SUMO. Driving in the lane at the guidance and stabilization level is generally performed automatically in SUMO, since vehicles always drive in the middle of the respective lane.

The simulation potentially involves a variety of situations that cannot be handled by the driver model. These deficiencies result from the scope of the available training data as well as the selection of input parameters of the model. The collected data mostly consist of trajectories at the intersection shown in Figure 3.8 in Section 3.3.1. For the majority of the other intersections in Ingolstadt, there is only sparse coverage due to additional xFCD. Therefore, the model behavior cannot be generalized to all possible intersection types in Ingolstadt due to the scope of the training data. For this reason, situations are defined in the following that are not part of the scope of the developed driver model. In these situations, the driver behavior is controlled temporarily by the Krauss model, which uses the calibrated parameter set of the respective driver type (cf. Table 3.2) and an additional emergency deceleration of -6 m/s^2. The emergency deceleration is only used in situations that cannot be resolved by the actual desired deceleration b_{des} without a collision. Furthermore, the parameter σ, which reflects the imperfection, is set to 0.2 for all driver types. This choice imposes noise on the driver behavior to counteract monotonic driving characteristics. The Krauss model is optimized for use in SUMO and is therefore able to handle all situations which have not been learned by LSTM.

One application of the Krauss model is a right-of-way situation at a signalized intersection. Since LSTM does not learn the right-of-way rules at an intersection and the associated driver response to a prioritized oncoming vehicle, the model can only be used in situations where the controlled vehicle has the right-of-way. Therefore, the right-of-way conditions are analyzed at all signalized intersections in the traffic model. It is investigated which lanes have the right-of-way across an intersection. In addition, for all non-priority lanes, it is investigated whether the respective traffic signal program causes a conflict with another lane or whether it is a protected green phase. On all lanes that guarantee right-of-way or are protected by the traffic signal, LSTM is used in the approach to the respective intersection. If there is no priority when crossing the intersection, the Krauss model is used. SUMO uses an intersection model that is tuned with the Krauss model to handle a safe crossing of an intersection.

During the approach to a traffic light, situations may occur in which a stronger deceleration is required than LSTM is able to execute. Such situations can occur, for example, due to a strategic lane change in front of a traffic light, which forces a following vehicle to decelerate sharply. Therefore, in highly critical situations at TTC values below 1 s and in low dynamic situations, as soon as the distance falls below the minimum distance s_0 of the Krauss model of the respective driver type, the system switches to the Krauss model. In addition, in free-flow situations when approaching a red light, the system also switches to the Krauss model if the TTJ is less than 1 s, in order to ensure a timely stop in front of the stop line.

As already explained in detail in Section 3.6, for the application of LSTM, the dilemma situation at the phase transition from green to yellow is explicitly resolved. For TTJ values below 2.53 s at the time of switching to yellow, the traffic light state is considered green for the model. In contrast, a traffic light is interpreted as red for a TTJ greater than 2.53 s when switching to yellow. This approach addresses the fact that not enough dilemma situations are available in the training data for LSTM to adequately learn the bifurcation of the decision process.

In the simulation, situations occur in which the model comes to a stop in a long queue in front of a signalized intersection. Sometimes, LSTM does not succeed in starting in such a situation, because standstill situations at a long distance from the intersection rarely occur in the training data. Therefore, in such situations where the model does not start without necessity at a large distance from the leading vehicle, the Krauss model is used until a speed of 10 km/h is reached.

In this thesis, situations are considered in which a driver approaches a traffic light. Therefore, the training data only contain trajectories in the approach to a traffic light controlled intersection up to the stop line. In the traffic model, however, vehicles cross the stop line and subsequently accelerate to the respective desired speed. In urban areas, the area beyond the stop line can often be interpreted as the approach to the next traffic light controlled intersection. Nevertheless, LSTM has not been able to learn this acceleration process from a low speed range to the desired speed. Therefore, the Krauss model is used in the traffic model after crossing the stop line until a speed of 40 km/h is reached. In order to select a realistic starting acceleration, the averaged maximum starting accelerations during the first 3 s of all observed vehicles are evaluated in the real data in free driving conditions. When comparing the calibrated maximum acceleration values a_{max} of the Krauss model of the four driver types (cf. Table 3.2) with the determined real acceleration values, significant deviations are found in some cases. The mean maximum starting acceleration for the *Moderate drivers* is 1.82 m/s^2 (+0.39 m/s^2 compared to the previously calibrated value), for the *Speeders* 1.70 m/s^2 (+0.96 m/s^2), for the *Timid drivers* 1.94 m/s^2 (+1.0 m/s^2) and for the *Aggressive drivers* 2.17 m/s^2 (+0.46 m/s^2). The deviations of the calibrated maximum accelerations a_{max} to the measured values arise in particular, because a_{max} in addition influences not only the starting behavior but also the behavior after the starting process. In the training data, in addition to starting processes, other situations occur in which the driver selects a lower acceleration in higher speed ranges. Therefore, a_{max} corresponds to a compromise of all observed acceleration situations. Up to a speed of 40 km/h, the parameter a_{max} is thus replaced by the mean maximum starting acceleration determined in each case. Otherwise, especially for the *Speeder* and *Timid driver* groups, unrealistically low acceleration rates would result without the correction. After reaching 40 km/h, LSTM takes over control of the vehicle again. Under these conditions, it is possible for LSTM to accelerate to the desired speed.

Since two different models have been trained for free driving and car-following, criteria are established for defining the driving state and the application of one of the models. Car-following defines a state characterized by the fulfillment of **one** of the following conditions:

- THW $\leq$ 3 s
- TTC $\leq$ 10 s
- $\Delta x \leq$ 20 m

A condition in which a potential leading vehicle is outside the space defined by the conditions is considered free driving.

Since two different models were trained for free driving and car-following with a different number of input parameters of LSTM, the transition between the driving states must be modeled explicitly. For the transition from car-following to free driving, the parameters concerning the leading vehicle are retroactively removed from short-term memory from the point in time at which the driver detects free driving after the PRT has elapsed. Consequently, only all remaining variables are used as input parameters of the free-driving model. In the opposite case, when changing from free driving to car-following, virtual data concerning the movements of the emerging leading vehicle are emulated. For the emulation, all movements of the leading vehicle are projected into the past for the last second in such a way that the relative velocity is kept constant. From this retrospective, the input parameters required for the car-following model are determined. In the following time steps, the emulated data are successively replaced by the observed data. Figure 5.2 schematically shows the emulation of the input data of LSTM during the transition process for both described transitions.

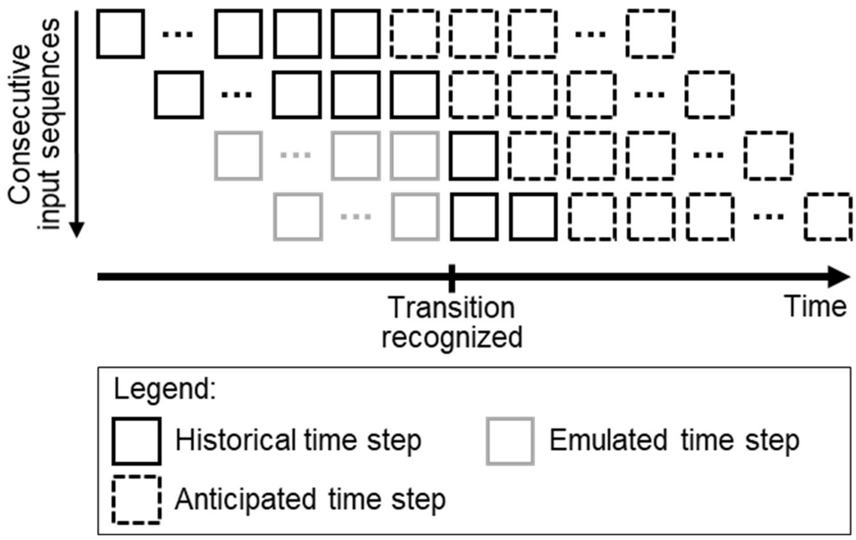

Figure 5.2: Emulation of LSTM input data for the transition from car-following to free driving and vice versa. The figure shows only one transition but applies to both transition directions.

In Figure 5.2, the aggregation of the input data in the individual time steps is described by chronologically ordered boxes. The observed historical input data are represented by black boxes with a solid line and the anticipated time steps by a dashed line. As soon as the driver detects a transition from car-following to free driving or vice versa, the already observed input data are changed retroactively according to the described procedure. These emulated time steps are shown in Figure 5.2 with gray boxes.

5.3 Simulative Study of the Interaction Among Human Drivers in Safety-Critical Situations

For the use of LSTM as a driver model for the surrounding traffic in virtual tests of automated driving systems, it is essential that the model behaves realistically in uncritical as well as critical situations. Critical situations are of special importance for the validation of an automated vehicle. Due to the large number of simulated vehicles, a representative number of critical situations occurs in the traffic simulation. This makes traffic simulation a valuable tool for recording and analyzing critical situations.

The aim of the following investigation is to validate the usability of LSTM in safety-critical situations in comparison to reality. Section 5.3.1 presents the simulation setup used for the investigation. In Section 5.3.2, the vehicle interactions are evaluated and validated with real data.

5.3.1 Simulation Setup

In order to analyze the driver model as well as its interaction with human drivers, three simulations were performed. All three simulations covered the calibrated real traffic volume of an average weekday for the period between 7 and 8 a.m. according to the calibration methodology presented in Chapter 4. In the simulation, a half-hour warm-up period is considered, during which the network already fills up with vehicles traveling on calibrated routes before the start of the evaluation period. The proportions of the four driver types are calculated according to the results of the driver type clustering (cf. Section 3.4). This results in 33 % *Moderate drivers*, 37 % *Speeders*, 10 % *Timid drivers* and 20 % *Aggressive drivers*. In the performed simulations, the developed driver model is used to control all simulated vehicles, considering the adaptations to the model explained in Section 5.2.2.

5.3.2 Evaluation of the Developed Driver Behavior Model in Safety-Critical Traffic Situations

In order to investigate how the developed model behaves in critical situations, the three simulations described in Section 5.3.1 are evaluated with respect to real data. For this evaluation, the reference is the trajectory data observed with the camera in reality, as described in detail in Section 3.3.1. The simulated data are evaluated in the same area where the real data were collected. Nevertheless, in order to generate realistic traffic flows at the area of interest, the whole city of Ingolstadt is simulated. Since the developed driver model addresses only the longitudinal behavior, while the lateral behavior is controlled by SUMO, only the car-following behavior is considered for the evaluation.

In the following, a traffic situation is classified as a “conflict” if the TTC falls below the value of 2 s at least once. To be sure, not all conflicts are critical: For example, in a traffic light scenario, a vehicle performs a controlled brake maneuver which can temporarily result in low TTC values at lower speeds. Nonetheless, the conflict number is a very useful statistical indicator as seen below.

Two parameters for the comparison between simulation and reality are examined in more detail: The first parameter is the average number of critical conflicts per hour and per vehicle. This parameter can be used to evaluate the frequency of conflicts in the simulation compared to reality. The second parameter is the average dwell time in critical situations for a single vehicle during one hour of driving. The mean dwell time complements the investigation of the number of conflicts, since longer conflicts below the TTC threshold of 2 s are potentially more critical than shorter critical situations. Figure 5.3 depicts the average number of conflicts per hour and per vehicle in reality and in the simulation.

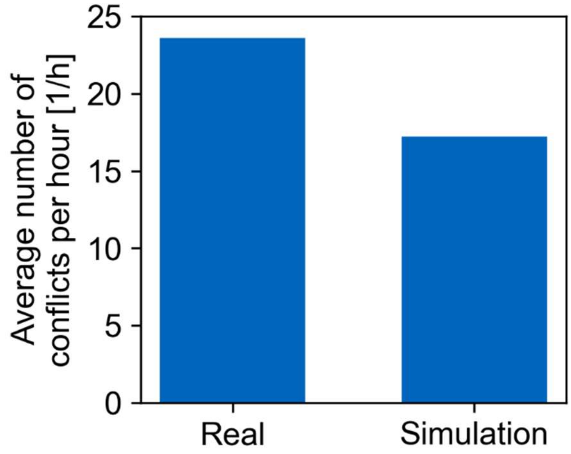

Figure 5.3: Average number of conflicts per hour and per vehicle in reality and in the simulation. Conflicts are defined by a TTC of less than 2 s.

Figure 5.3 shows that each vehicle encounters about 23.6 conflicts per hour in reality and about 17.2 conflicts per hour in simulation. At first glance, the total number of conflicts with a TTC below 2 s seems very high. However, as explained above, in the considered scenario (approach to a traffic light), uncritical low TTC values at lower speeds can occur, keeping in mind that the TTC statistic assumes constant vehicle dynamics.

Comparing both bars in Figure 5.3, a vehicle in reality is faced with about 6.4 conflicts per hour more than a vehicle in simulation. It can be concluded that LSTM generates a more defensive driving style that leads to a lower number of critical situations. In order to evaluate the overall criticality, Figure 5.4 shows the average dwell time in critical situations per vehicle and per hour which is weighted by each vehicle's travel time.

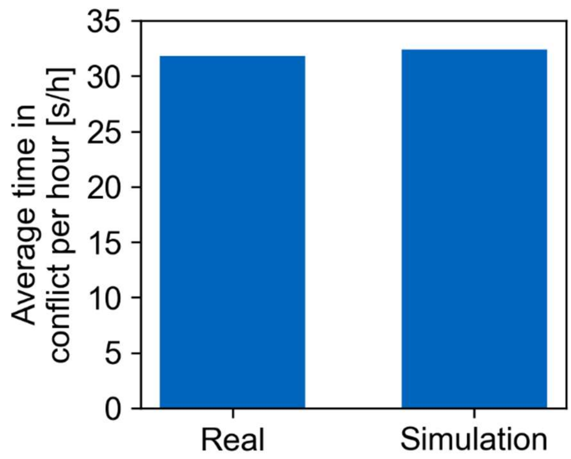

Figure 5.4: Average dwell time in a conflict per vehicle during an hour weighted by each vehicle's individual travel time. Conflicts are defined by a TTC of less than 2 s.

Figure 5.4 shows that per hour each vehicle in simulation spends about 32.4 s in critical situations while real vehicles spend 31.8 s in conflicts on average. From the combination of Figure 5.3 and Figure 5.4, it can be concluded that the average duration of a simulated conflict is 1.9 s and 1.3 s for a real conflict. Overall, LSTM on the one hand behaves slightly more defensively than a real driver. On the other hand, in a critical situation, LSTM takes longer to mitigate criticality.

The results show that LSTM can reproduce human driver behavior in critical situations. However, as the large majority of situations in the data pool is uncritical, more conflict data would be useful in order to further improve the performance of LSTM in these situations. The results of Figure 5.3 and Figure 5.4 indicate that simulated conflicts stay critical for a longer period. This behavior could even be used to specifically stress an automated vehicle in a virtual test.

5.4 Implementation of Automated Vehicles in the Traffic Model of Ingolstadt

For the integration of automated vehicles into the traffic simulation for a virtual endurance run, the *AutomatedVehicles* module of the framework presented in Section 5.1 takes control of each simulated automated vehicle. Section 5.4.1 describes the implementation of automated vehicles based on the state of the art described in Section 2.5. In Section 5.4.2, a procedure is introduced that recognizes the driver type of the vehicle ahead by observing the leading vehicle. Based on the driver type recognition, this section sets the foundation for investigating how an adaptation of the driving behavior of automated vehicles affects traffic safety and efficiency.

5.4.1 Model of Automated Vehicles

In order to reach the maximum validity of the following investigations of the interactions between human drivers and automated vehicles, the model of a real driving function that can handle all situations arising in the simulation would be required in the simulation. However, such a model for urban areas is not available for this thesis. Instead, in line with the current state of the art, for modeling automated vehicles (cf. Section 2.5), the Krauss model is used as a typical engineering model. Regarding the parameter set of Krauss, the parameter τ_{k} is set to 1 s. Consequently, according to the algorithm of the Krauss model, a desired THW of 1 s is maintained while driving at a similar speed as the leading vehicle. Thus, a human-like distance-keeping behavior is generated. This behavior is intended to prevent a potentially oppressive effect on a human-controlled leading vehicle, which in reality would potentially strongly influence driver behavior. It is also assumed that tailgating would reduce the acceptance of automated driving in the population. The human drivers controlled by LSTM could not recreate such effects as LSTM does not react to the movements of the vehicle behind. The remaining parameters of the Krauss model controlling the automated vehicles are chosen as follows: The parameter for maximum acceleration a_{max} is set to 2 m/s^2, the maximum deceleration b_{des} is set to 4 m/s^2 and the minimum standstill distance s_0 is set to 1.5 m. The desired speed v_{des} is selected depending on the speed limit. In addition, the parameter σ, which determines the imperfection of driving in SUMO, is set to 0. As a result, the simulated automated vehicles exhibit a smooth driving behavior. The control of the vehicles is fully handled by SUMO in the longitudinal as well as in the lateral direction. Analogous to the models of human vehicles, the standard lane change model of SUMO is applied for the lane change behavior.

5.4.2 Driver Type Recognition Procedures

An important motivation for the introduction of automated vehicles into road traffic is to sustainably improve traffic safety. In the following, it is investigated whether traffic safety can be improved by a procedure that enables automated vehicles to classify the driver in front into one of the four presented driver types. When specific driver types are detected, a correspondingly more cautious behavior is selected in order to mitigate or completely avoid potential critical situations caused by strong sudden decelerations. In addition, it is investigated how a change in the behavior of the automated vehicle affects a human-controlled vehicle driving behind it.

In the simulation, each automated vehicle permanently observes the leading vehicle. On the one hand, it stores the parameters that exclusively characterize the driving state of the leading vehicle. These include the maximum speed, the maximum acceleration and the maximum absolute deceleration. On the other hand, the minimum THW is recorded and serves as a quantity that describes the interaction with the vehicle in front of the leading vehicle. In comparison to the clustering described in Section 3.4, the critical TTC is not considered for the redetection of the drivers. To determine the critical TTC for clustering, negative TTC values are only considered if no positive value is observed. As soon as a positive value is registered, a negative value is overwritten. In the traffic simulation, positive TTC values result for each driver, especially in rear-end situations for vehicles stopped in front of a traffic light. Therefore, the observation of the TTC values leads to a significantly different TTC characteristic especially for the *Timid* drivers, for which a majority of negative TTC values were observed from the data analysis of the real data. This results in a significantly worse detection rate in the detection procedure explained below. For this reason, for driver type recognition, the five-dimensional hyperspace spanned by the parameters chosen for clustering is reduced by the hyperplane of the critical TTC. The information from the other four parameters remains for the detection. If one or more of these parameters have not yet been observed while following a manually controlled vehicle, the missing parameters are added in analogy to Section 3.4 with a multivariate imputation [VAN BUUREN AND GROOTHUIS-OUDSHOORN, 2011] using the complete available data set resulting from the clustering.

After a minimum observation time of 5 s and if the leading vehicle has driven faster than 10 m/s at least once, a classification of the driver type is made based on the knowledge gained about the four observed parameters. For this classification, the center points of the respective clusters, which are found in the k-means clustering procedure in Section 3.4 from the analysis of the entire available data set, are compared with the observed parameters considering PCA. The cluster whose center has the smallest Euclidean distance from the observed parameter set is selected. In the following, the assumed cluster of the leading vehicle is updated every 10 s based on the total observation.

The adaptation of the driving behavior of automated vehicles is applied with respect to the classified driver types for detected *Aggressive* and *Timid drivers*. The *Aggressive* driver group drives with low THW and TTC values compared to the other driver types and therefore exhibits a high potential for abrupt braking maneuvers. *Timid drivers* exhibit pronounced braking decelerations that are interpreted as anxious behavior. This unpredictability creates the potential for abrupt braking maneuvers that can lead to critical situations. Once an *Aggressive* or *Timid* driver type has been detected using the described methodology, the time headway of an automated vehicle is increased from τ_k = 1 s to τ_k = 1.8 s. This time headway corresponds to the rule of thumb commonly used in Germany to maintain half the value of the speedometer in m to the leading vehicle in order to maintain a safe distance. All other parameters of the Krauss model remain unaffected by the adjustment of the driving behavior. In order to increase traffic safety, it would also be theoretically conceivable that a higher time headway is maintained with respect to all driver types in order to induce an even safer driving behavior. However, a consistently more cautious behavior of automated vehicles may also have a negative impact on traffic efficiency. Under certain circumstances, this can lead to a reduced acceptance towards automated vehicles among the population. Accordingly, a trade-off must be made between an overly cautious and a risk-affine approach to adjusting driving behavior. In order to be able to investigate the effects of risk-averse behavior on traffic safety and efficiency, the effects of a permanent time headway of τ_k = 1.8 s towards all driver types is analyzed in the following.

In order to evaluate the potential of the described methodology with respect to traffic safety, five layers of cluster detection are investigated independently. Figure 5.5 shows an overview of the five layers in increasing detail from bottom to top.

At the zeroth layer, the automated vehicles do not perform any cluster detection nor adjust driving behavior. Instead, an increased desired time headway of τ_k = 1.8 s is set towards all driver types. The zeroth layer serves as a reference for the layers 2-4 reflecting a very cautious driving behavior.

The first layer also describes no observation of the leading vehicle and therefore no potential adjustment of the driving behavior. A desired time gap of τ_k = 1 s is set. This level serves as a reference state for the comparison with all other levels on which a detection of the leading vehicle as well as an adaptation of the driving behavior take place.

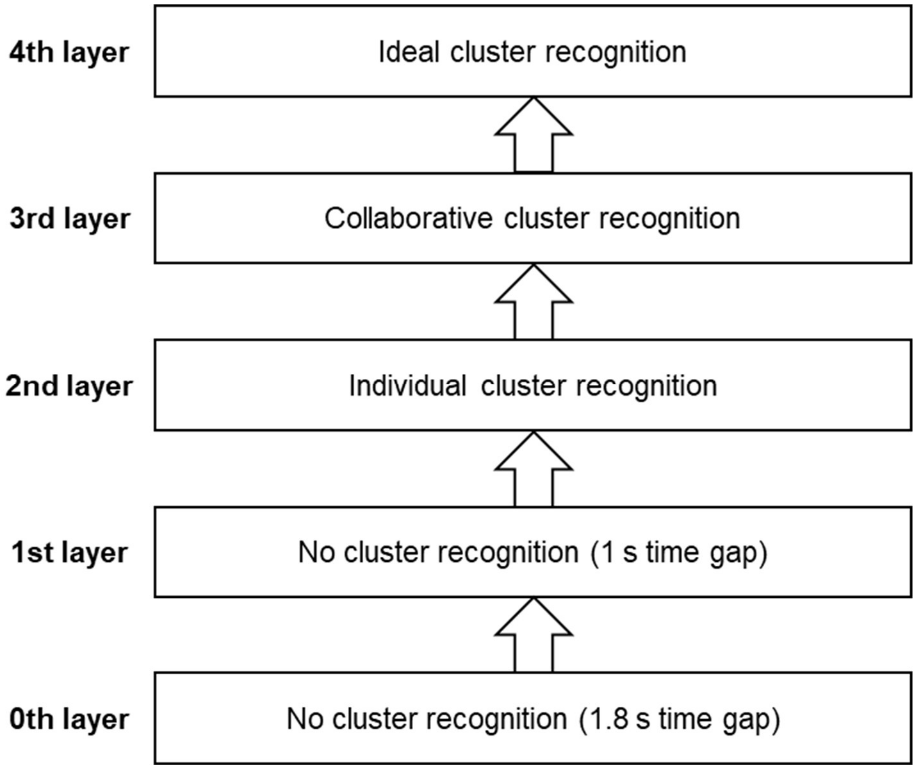

Figure 5.5: Investigated five layers of cluster recognition.

On the second layer, each automated vehicle individually collects information regarding the leading vehicle in order to determine the respective driver type and adapt the driving behavior. There is no exchange of information between the automated vehicles using this procedure. The collected information is deleted after the car-following process is completed.

On the third layer, all automated vehicles collaborate and bundle their collected findings on the human-controlled vehicles in a common data pool. This type of data management has the advantage of increasing the total time vehicles are observed. The information obtained is not deleted after the car-following process is completed, but is made available to the other vehicles for further use. However, the major disadvantage of the described methodology is that the data collected are highly sensitive personal data. Therefore, the actual implementation of the described methodology is questionable. Nevertheless, this thesis hypothetically examines the case that all data protection obstacles could be overcome.

At the top layer, an ideal recognition algorithm is available to the automated vehicles to determine the driver type ahead. In the performed simulation, the automated vehicles access the ground truth cluster data of the respective leading vehicle for a potential adaptation of the driving behavior. Provided that the leading vehicle is an *Aggressive* or *Timid* driver, an adaptation of the driving behavior is done without a minimum observation time. The described procedure represents the best possible detection, which promises the best possible effects on traffic safety. This is an idealization, since a detection algorithm will never act perfectly and drivers do not always show the behavior that is most characteristic for their specific type.

5.5 Simulative Study of Driver Type Recognition Quality

In the coming years, traffic will be composed of a mixture of manually controlled and automated vehicles for a longer period of time [WAGNER, 2015]. Therefore, the study of this traffic condition is of great importance. In order to evaluate the interactions of human-controlled and automated vehicles, a mixed traffic was simulated in the traffic simulation of Ingolstadt. The human-guided vehicles were controlled by the developed driver model considering the adaptations presented in Section 5.2 and the automated vehicles were modeled according to Section 5.4 equipped with one of the driver type recognition procedures. Section 5.5.1 presents the used simulation setup. In Section 5.5.2, the driver type recognition methods of layers 2 and 3 (cf. Figure 5.5) are compared with each other in terms of recognition quality dependent on the observation time.

5.5.1 Simulation Setup

In the simulation, the proportion of the four driver types was chosen according to the values determined from the clustering (cf. Section 3.4). The proportions result in 33 % *Moderate drivers*, 37 % *Speeders*, 10 % *Timid drivers* and 20 % *Aggressive drivers*. The use of the developed driver model follows the assumption that a human driver does not significantly change his/her own behavior in the presence of an identified automated vehicle. The share of automated vehicles in the total traffic is set to 10 %. It is assumed that the distribution of driver types among human-controlled vehicles remains unchanged by the introduction of automated vehicles into traffic. It is additionally assumed that the total amount of traffic is approximately constant with low penetration rates of automated vehicles and that the distributions of routes through the city do not change significantly. Therefore, the simulation used routes calibrated under traffic conditions in which only human drivers were encountered. It is currently the subject of research to determine under what conditions a higher occupancy rate of automated vehicles can lead to a reduction in overall traffic volumes [LITMAN, 2022].

For the conducted studies, the traffic volume of a weekday was considered for two periods in peak traffic hours between 7 and 8 a.m. and 4 and 5 p.m. and in moderate traffic between 9 and 10 a.m. according to the calibration methodology presented in Chapter 4. A half-hour warm-up period was considered in each case. Three simulations were performed for each cluster detection method for each time period. The performed simulations correspond to the "virtual endurance run" presented in Section 2.2, in which a large number of automated driving systems are stressed by virtual traffic without explicitly testing specific scenarios.

The use of LSTM as a driver model requires significantly higher computing power than the use of an engineering model. In particular, data preparation and prediction by LSTM represent computationally expensive processes. Therefore, a strategy must be considered to avoid these costly calculations. The focus of the endurance run is on the analysis of interactions between automated vehicles equipped with different driver type recognition techniques and human-controlled vehicles. In all areas in the simulation without interaction between automated and manually controlled vehicles, realistic traffic volumes must prevail, but the requirements for the quality of individual driver behavior are lower. For these reasons, LSTM was only used for human-controlled vehicles that directly or indirectly interact with an automated vehicle. These relevant areas of direct or indirect interaction were defined in the field of view of automated vehicles with a range of 100 m in the direction of travel and to the rear. Not only the direct leading or following vehicle, but all vehicles in the field of view of an automated vehicle were controlled by LSTM. In turn, all human-controlled vehicles that were not within the visual range of an automated vehicle were controlled by the Krauss model. For each driver type, the corresponding calibrated parameters shown in Table 3.2 are selected. Figure 5.6 shows the described usage of LSTM and Krauss.

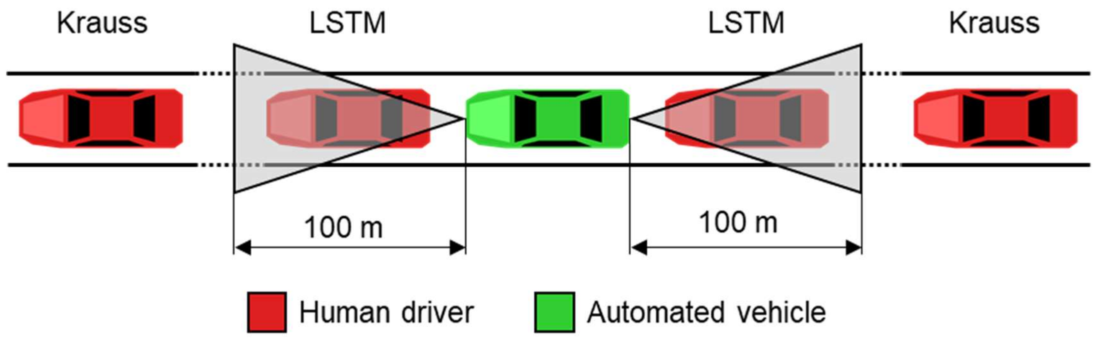

Figure 5.6: Usage of LSTM and Krauss model to reduce the computational effort during the simulation. LSTM is used within the sensor field of view of any automated vehicle in the direction of travel and to the rear. The Krauss model is used outside the sensor field of view of any automated vehicle.

Although the usage of LSTM was limited, data preparation at each time step and for each human driver during the simulation of the entire traffic model still prevented the simulation from being executed in real time. Instead, the simulation of up to 9000 vehicles ran with a factor of about 1/10 of real time. However, in general, the performance of the simulation of each individual human driver or automated vehicle falls far below the real time.

It should be emphasized that the Ingolstadt traffic model, despite the integration of the developed driver model and the automated vehicles, which inevitably leads to a different driver and traffic behavior, did never show a breakdown of the traffic at any time. In the calibration presented in Chapter 4, instabilities in the traffic state occurred even with small adjustments in driver behavior. The fact that the simulation maintains a stable state indicates that the driver behavior model can be used for further analysis with the adjustments made as described in Section 5.2.2.

Figure 5.7 shows an exemplary image from the Ingolstadt traffic model. Automated vehicles are depicted in green, *Moderate drivers* in red, *Speeders* in grey, *Timid drivers* in blue and *Aggressive drivers* in orange.

Figure 5.7: Section of the traffic model of Ingolstadt in SUMO. Automated vehicles are depicted in green, *Moderate drivers* in red, *Speeders* in grey, *Timid drivers* in blue and *Aggressive drivers* in orange.

5.5.2 Evaluation of Driver Type Recognition Quality

As described in Section 5.5.1, simulations were performed for the periods between 7 and 8 a.m., 9 and 10 a.m., and 4 and 5 p.m., in which the different driver type detection methods were used (cf. Figure 5.5). From the evaluation of the simulations in which individual (2nd layer) and collaborative cluster recognition (3rd layer) were used, conclusions can be drawn about the quality of the respective recognition method. Layers 0 and 1 serve as references which do not apply any recognition procedure. Therefore, these two layers are not considered in the following detection quality analysis. The ideal detection (4th layer) is a completely error-free reference procedure and is therefore also not shown in the following analysis.

For the four driver types *Moderate drivers*, *Speeders*, *Timid drivers* and *Aggressive drivers*, the sensitivity, the specificity, the positive predictive value (PPV) and the negative predictive value (NPV) are determined for each simulation for both the individual and the collaborative driver type recognition methods. The calculation is performed after 5 s, 15 s, 30 s, 60 s, 120 s and finally after 300 s observation time. For both considered cluster detection methods, the observation period is recorded separately for each vehicle-vehicle constellation. Thus, also for the collaborative detection method, the observation time of a preceding vehicle starts at 0 s for each initial vehicle-vehicle constellation, even if data for the preceding vehicle are already available in the data pool. The investigated statistical parameters are averaged over all three simulations of each period for each observation time point. There are only minor differences between the different time periods. In the following, the results for the period from 4 to 5 p.m. are presented. The other results for the periods between 7 and 8 a.m. and 9 and 10 p.m. can be found in the Appendix A.

The sensitivity, also known as the true positive rate, of a method describes the probability with which a certain driver type is correctly assigned to the correct cluster. Figure 5.8 shows the sensitivities for all four driver types as a function of the observation time. The solid lines represent the individual detection method and the dashed lines the collaborative method.

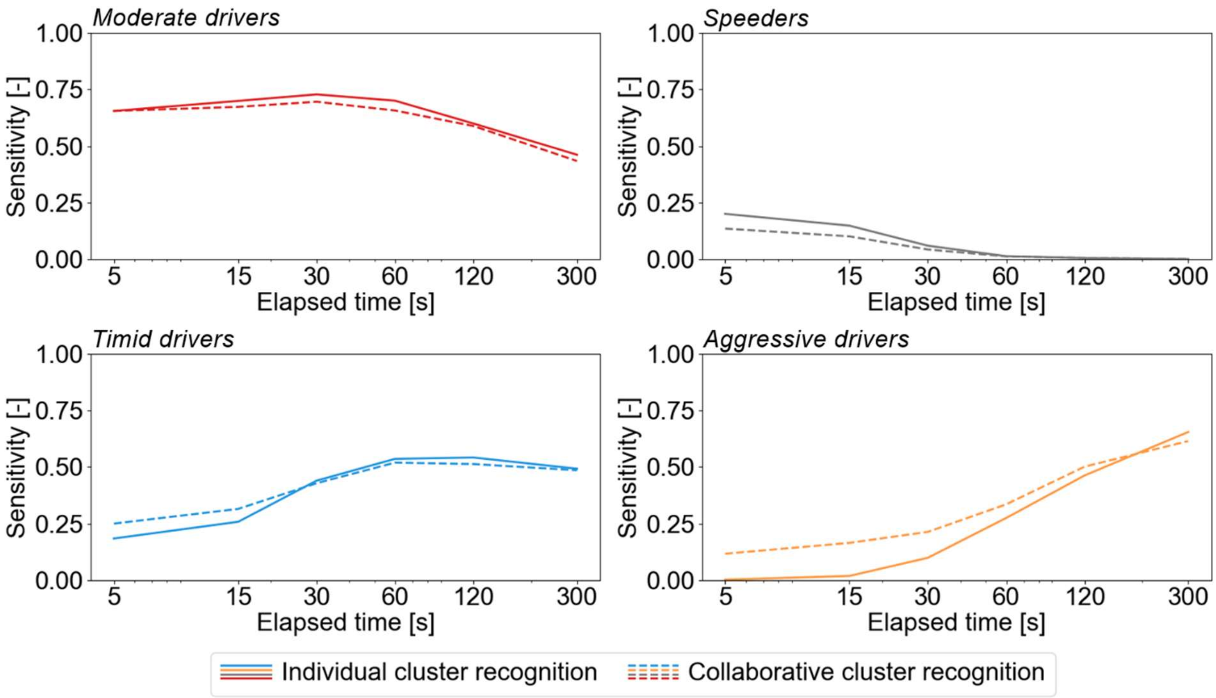

Figure 5.8: Sensitivity of the investigated driver type recognition procedures for all four driver types as a function of observation time.

The sensitivity of the detection of the *Moderate drivers* is about 0.65 for both methods at the beginning of the observation. Up to an observation time of 30 s, the sensitivity of both methods increases. Subsequently, both graphs drop until an observation time of 300 s. It can be concluded that *Moderate drivers* also show characteristics of other driver types after a certain time and are therefore assigned to one of the other types. Furthermore, it is noticeable that the graph of the collaborative cluster detection method is always below the graph of the individual method except for the first data point. In the collaborative method, vehicles have information available that potentially already comes from longer observation. However, since sensitivity decreases after longer observation, this is reflected in a lower sensitivity already at the start of observation of the collaborative recognition procedure.

Both methods show a low sensitivity of 0.2 and 0.14 already at the beginning of the observation of the *Speeders*. The sensitivity shows a decreasing course for both methods over the observation time. A probable reason is the characteristic of the *Speeders* to choose an excessive speed. The higher speed makes the *Speeders* escape from the sensor field of view of the automated vehicles. It can be observed that for longer observation times very few *Speeders* are observable. Furthermore, in congested traffic, *Speeders* are sometimes unable to realize their desired speed and can thus be falsely assigned to other driver types.

The sensitivity of the collaborative cluster recognition of the *Timid drivers* is already 0.07 higher than the individual recognition at the beginning of the observation. It can be concluded that the already existing information about the human drivers contributes to a better recognition quality of the *Timid drivers*. For both methods, an increase in sensitivity over time can be seen. After about 30 s, the individual driver type recognition shows a slightly higher sensitivity than the collaborative method. Both sensitivity values stabilize at about 0.5 after a longer observation time.

The *Aggressive drivers* are almost not detected at all at the beginning of the observation with the individual method. In contrast, the use of the data pool leads to a significantly higher detection rate of 0.12 already at the beginning. In the following, the sensitivities of both methods increase monotonously up to a value of 0.62 at 300 s observation time. Apart from the last observation point, the sensitivity of the collaborative method always shows higher values than that of the individual method. This effect is also shown in the evaluations of the simulations between 7 and 9 a.m. as well as between 9 and 10 a.m. (cf. Figure A.1 and Figure A.5).

The specificity, also called true negative rate, indicates the probability of correctly not assigning a certain driver type to a cluster to which it does not belong. An exemplary method that always outputs the same fixed driver type without any intelligence would have a sensitivity of 1 for that driver type, but a specificity of 0. Figure 5.9 shows the specificities of the two detection methods for all driver types over time.

Especially at the beginning of the observation, a large part of the driver population is initially wrongly categorized as *Moderate drivers*. The specificity of 0.33 of the individual method corresponds to the proportion of *Moderate drivers* in the simulation. Only with increasing observation time, driver types show their specific characteristics and can be assigned to other groups. Accordingly, the specificity of *Moderate drivers* is initially low and increases over time. The additional observations of other vehicles cause a higher specificity of the collaborative procedure compared to the individual procedure at each observation time point. The same effect is evident in the evaluation of the time period between 7 and 8 a.m. and for most parts also for 9 to 10 a.m. (cf. Figure A.2 and Figure A.6 in Appendix A).

The specificity of the *Speeders* is already high at the beginning of the observation with values of 0.8 for the individual and 0.86 for the collaborative procedure. Over time, both curves converge towards a specificity of 1.

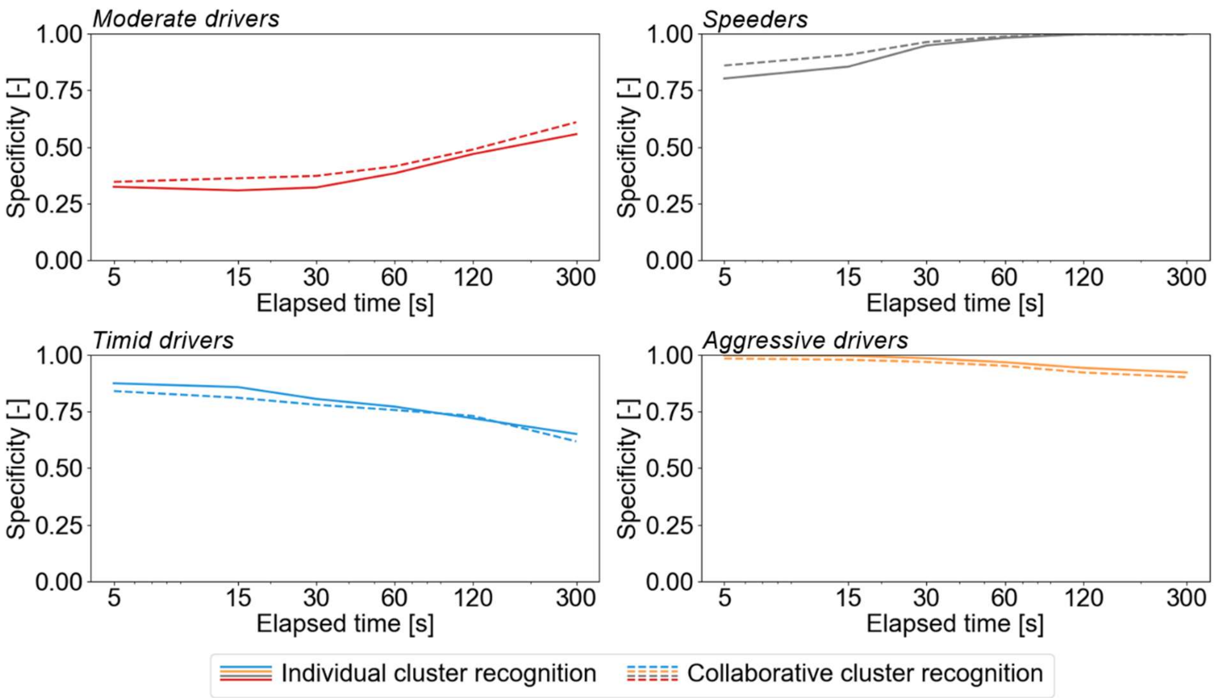

Figure 5.9: Specificity of the investigated driver type recognition procedures for all four driver types as a function of observation time.

For the *Timid drivers*, a decreasing trend in specificity for both methods over time is evident. With increasing time, *Timid drivers* show characteristics of other driver groups and are erroneously assigned to the wrong cluster. The use of the data pool slightly amplifies this effect, so that the specificity of the collaborative procedure is slightly lower than that of the individual procedure. The detection methods in the other two investigated time intervals show similar results (cf. Figure A.2 and Figure A.6).

Finally, the specificity of the *Aggressive drivers* already shows a value of nearly 1 for both methods at the beginning of the observation. In the further course, both curves are very close to each other and drop slightly. It can be concluded that the characteristics of the *Aggressive drivers* are very distinct and differ significantly from those of the other driver types. Therefore, hardly any other driver types are wrongly assigned to the *Aggressive drivers*.

The third considered classifier is the positive predictive value (PPV). PPV describes the probability of a driver being correctly assigned to a certain cluster among all drivers who are also assigned to this cluster. PPV thus indicates the proportion of a driver type that is correctly classified. The worst possible PPV would occur when simply guessing the driver type with a uniformly distributed probability of 25 %. In this case, PPV would be the proportion of the driver group in the total driver population. In the following diagrams of Figure 5.10, which show the PPV values of the detection methods for all four driver types, the proportion of the driver type is therefore marked with a dotted black line serving as a reference.

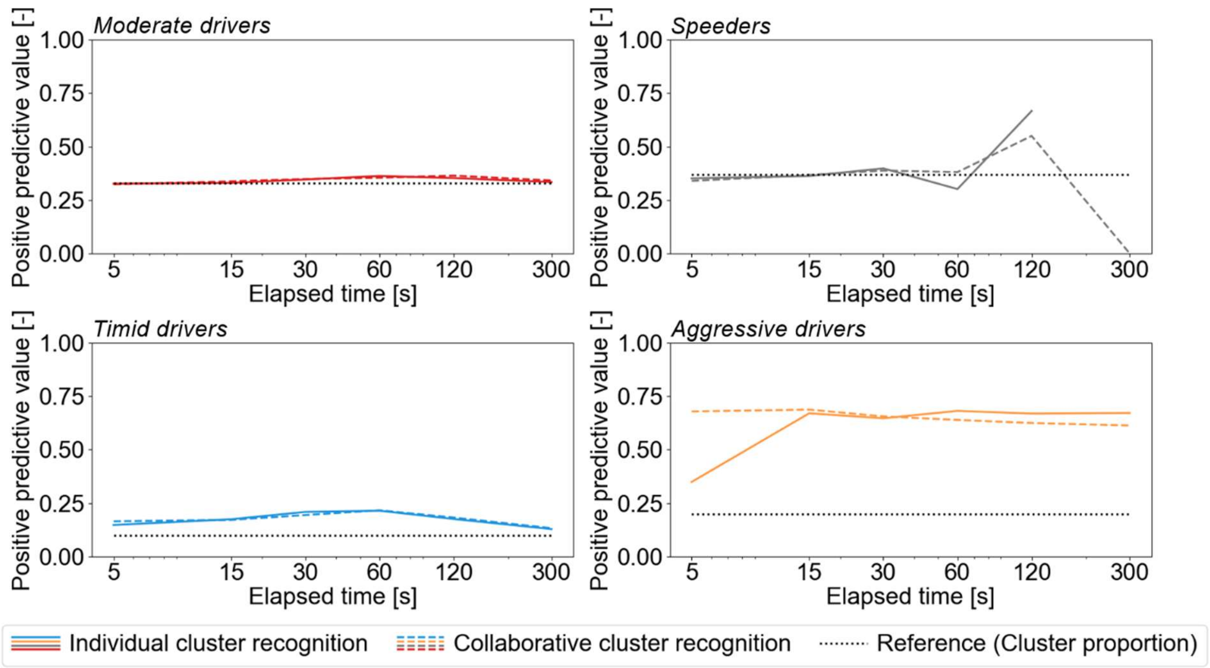

Figure 5.10: PPV of the investigated driver type recognition procedures for all four driver types as a function of observation time.

For both *Moderate drivers* and *Speeders*, the PPV value over time is not significantly better than the reference. For long observation durations, the number of observed *Speeders* is so small that a valid calculation of the PPV can no longer be made (cf. Figure A.3 and Figure A.7 in Appendix A).

The PPV values of both methods for the *Timid drivers* are in the range between 0.12 and 0.22, while the cluster proportion is 0.1. Accordingly, both methods offer a higher quality than the reference. There is no significant difference between the individual and the collaborative detection method over time. However, the collaborative detection method is slightly more precise at the beginning of the observation.

The detection methods show the highest prediction accuracy in the detection of the *Aggressive driver* type. A large PPV difference can be seen between the two procedures, especially at the beginning of the observation. The PPV value of the individual detection procedure already starts at a value above the cluster proportion. With increasing time, PPV increases monotonically and reaches a value of 0.68. In contrast, the PPV of the collaborative method is already at 0.68 at the beginning of the observation and remains static over time. At the other simulated observation times, both graphs also converge towards about 0.7 (cf. Figure A.3 and Figure A.7 in Appendix A). The prediction accuracy is always above the cluster proportion in all cases. It can be seen in the comparison of all driver types that *Aggressive drivers* are best detected by both detection methodologies caused by their distinct characteristics.

The imperfect detection of *Moderate drivers*, *Speeders*, but also *Timid drivers* results from the comparably short observation interval of the drivers in the original camera data used to form the clusters, compared to the observation time by the vehicles in the simulation. The nature of the clustering procedure, which only considers maximum and minimum values, also contributes to drivers being assigned to an incorrect group in the case of outliers. In the simulation, outliers can appear due to traffic conditions, which are not covered by the actual cluster.

The last considered classifier is the negative predictive value (NPV). NPV describes the probability of a driver being correctly not assigned to a particular cluster among all drivers who are also not assigned to this cluster. NPV thus indicates the proportion that is correctly not classified to a driver type. For reference, the respective aggregated proportion of the remaining driver types is plotted as a dashed line in each diagram. This value corresponds to the NPV with a theoretical procedure that randomly draws a driver type. Figure 5.11 shows the NPV for all four driver types over time.

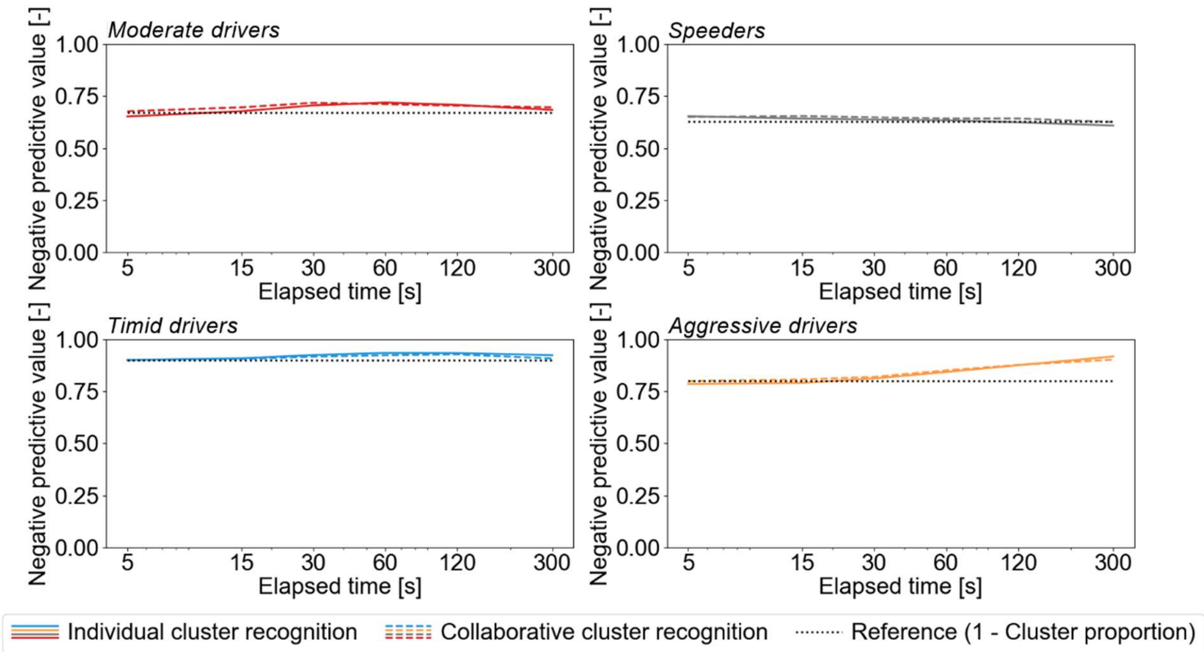

Figure 5.11: NPV of the investigated driver type recognition procedures for all four driver types as a function of observation time.

Figure 5.11 indicates that the NPV for all driver types is very similar for both detection methods. Furthermore, all curves, except for the *Aggressive drivers*, have an almost constant course. The NPV of the correct non-classification of the *Timid drivers* is constantly the highest with about 0.9. The NPV of the *Aggressive drivers* increases from 0.79 to 0.9. It can be concluded that both methods show a better separation ability for both the *Timid drivers* and the *Aggressive drivers* than for the other two driver types.

In summary, the collaborative detection method provides significant added value over the individual method in detecting *Timid* and *Aggressive drivers*, especially at the beginning of the observation. Early detection of these driver types allows the automated vehicles to early adjust driving behavior and potentially avoid critical situations. For *Moderate drivers* and *Speeders*, there is no significant difference between the individual and the collaborative detection procedure. However, these two groups are also not the focus of further analyses due to their less pronounced characteristics. Basically, the fact that the defined driver types can be recognized by the two procedures shows that LSTM can reproduce their main characteristics. Nevertheless, the selected methods show weaknesses in the recognition quality. These are based on the nature of the procedure, which uses maximum and minimum values to classify drivers. Slight outliers thus quickly lead to incorrect grouping. It is also conceivable that better recognition can be achieved with input data for longer observation periods.

5.6 Evaluation of the Interaction Between Human Drivers and Automated Vehicles

To investigate the effectiveness of the different driver type detection methods, as described in Section 5.5.1, multiple simulations were performed at three different time intervals as a virtual endurance run. In the previous sections, the prerequisites for this exemplary virtual test were outlined. Section 5.3 shows the capability of the developed human driver model to reproduce safety-critical driving situations in similar frequency and dwell time compared to reality. Section 5.5 proves the usability of the different driver recognition methods (cf. Figure 5.5) for the automated vehicles. Based on these findings, in the following, it is investigated how the introduction of the different detection methods affects traffic safety and traffic efficiency. Since the methods address the interactions between human drivers and automated vehicles, an extensive statistical evaluation of these interactions is required to analyze the influence of the methods on traffic. The performed simulations without cluster detection and with a desired time headway of 1 s (layer 1) serve as a reference group for evaluating the other procedures.

In Section 5.6.1, the concept for the evaluation of traffic safety is introduced including covered conflict scenarios and analyzed quantities. Section 5.6.2 evaluates the impact of the driver detection methods on minimum TTC values per conflict. Section 5.6.3 analyzes the impact of the methods on the number of critical conflicts per vehicle and hour. In Section 5.6.4, traffic safety and traffic efficiency differences are compared in order to investigate potentials of the different detection methods. From the comparison, the accident avoidance potential of the respective method is estimated. In Section 5.6.5, these potentials are determined and expressed in monetary terms based on the economic costs saved by accident avoidance.

5.6.1 Evaluation Concept

The differences in driver recognition methods are expected to have a measurable impact on traffic safety and efficiency. As described in Section 5.3, a longitudinal dynamic situation is assumed to be a critical conflict as soon as the TTC falls below 2 s. Three quantities are chosen as a measure for the evaluation of the influence of the driver type detection methods on traffic safety: The distribution of the minimal TTC value per conflict shows the shift in criticality for the different cluster recognition procedures. The average number of conflicts per vehicle and hour provides information on the quantitative frequency of conflicts, but not on the duration of the individual conflicts. Therefore, the average dwell time each vehicle spends in a critical conflict within one hour is analyzed. Both variables complement each other in order to draw conclusions about traffic safety. In addition, average velocity differences serve as the measure for traffic efficiency changes. All variables are analyzed for the following two traffic situations (a) and (b):

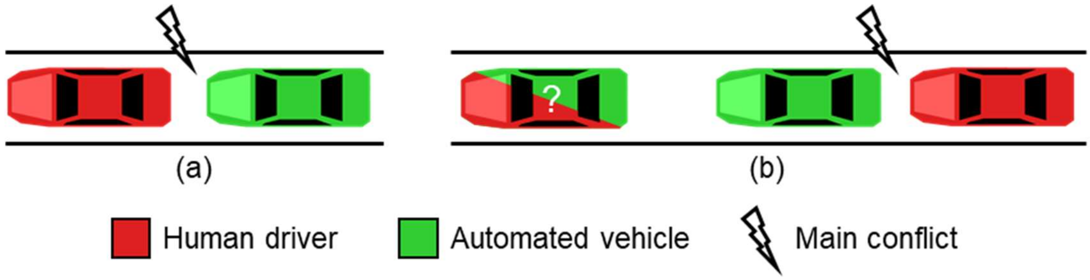

Figure 5.12: Investigated conflict situations: In (a), an automated vehicle follows a human-driven vehicle and eventually adapts its behavior according to a cluster recognition method. In (b), a human driver follows an automated vehicle which eventually adapts its behavior towards a potential leader (human driver / automated vehicle / none).

The first considered traffic situation (a) covers the conflict between an automated vehicle following a human-driven vehicle. The automated vehicle might adapt its driving behavior towards its leader according to a driver detection method. In the second analyzed conflict (b), a human driver follows an automated vehicle. The automated vehicle might adapt its behavior towards a leading human-driven vehicle. However, there could also be another automated vehicle or no leader which in both cases does not trigger an adaptation of the behavior.

5.6.2 Evaluation of Minimal TTC Values per Conflict

Figure 5.13 shows the probability density distributions of the minimal TTC value per conflict in a range between 0 s and 5 s for situation (a) for all investigated driver detection procedures.

Figure 5.13 shows two distinct peaks for layers 1 to 4. Apart from layer 0, the distributions of all driver recognition methods show a first peak at a TTC value of about 1.75 s. This peak seems to be characteristic for the Krauss model used for the automated vehicles in critical situations with a desired time headway of 1 s. The distribution of the individual cluster recognition shows a similar number of critical situations in the low TTC range as the distribution

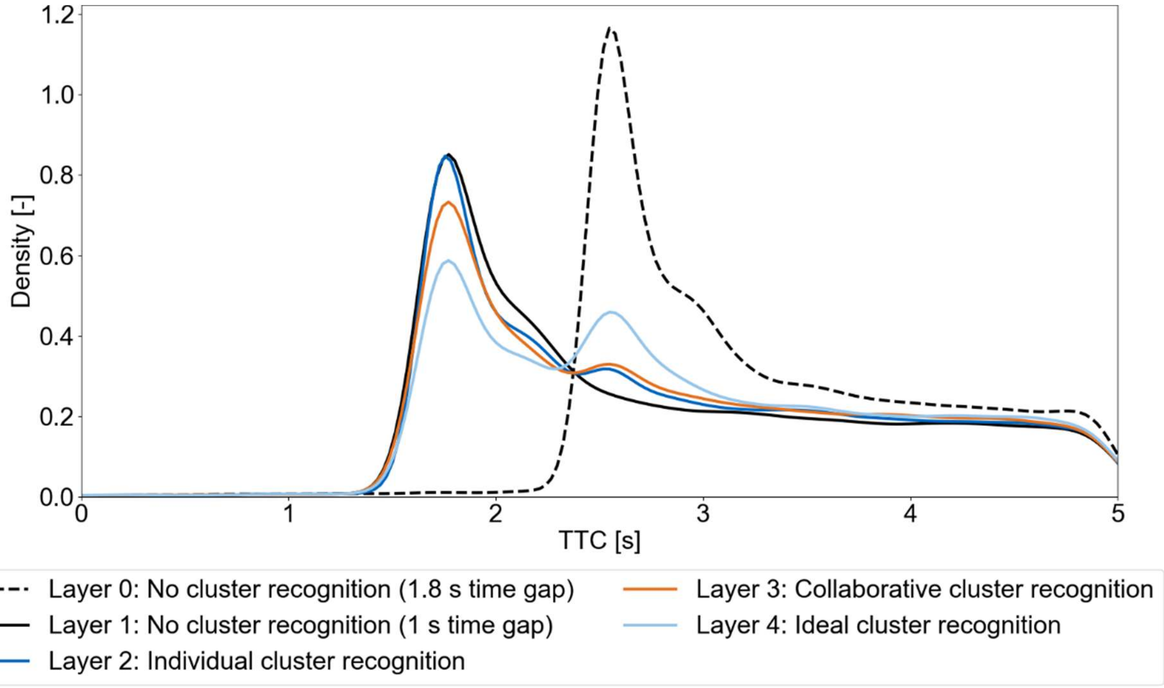

Figure 5.13: Distributions of the minimal TTC value per conflict in a range between 0 s and 5 s of automated vehicles in interaction with leading human-driven vehicles for all investigated driver type recognition procedures.

of layer 1. Automated vehicles using the collaborative method are rarely exposed to critical situations of less than 2 s compared to layers 1 and 2. Most critical situations are avoided by the ideal cluster detection. The second significant peak is at a TTC value of approximately 2.5 s. This peak results from the adjustment of the desired time gap to 1.8 s. It is the only local maximum in the distribution of layer 0 in the low TTC area. For layers 2 – 4, it is noticeable that there is a shift from the peak at 1.75 s to the non-critical range with TTC values around 2.5 s: The lower the proportion at 1.75 s, the higher it is at 2.5 s. The large number of values in small TTC ranges of less than 2 s result from the type of the considered situation which is the approach to a traffic light. On uninfluenced road sections without a traffic light, the distributions would shift towards higher TTC values.

In addition to the interactions of the automated vehicles with the leading vehicle, in situation (b), the impacts of the recognition methods in conflicts with the human-driven vehicle behind are analyzed. It is investigated whether a more cautious driving behavior towards the leading vehicle has an influence on critical situations with the vehicle behind. Figure 5.14 shows the probability density distributions of minimal TTC between a human-driven vehicle and a preceding automated vehicle.

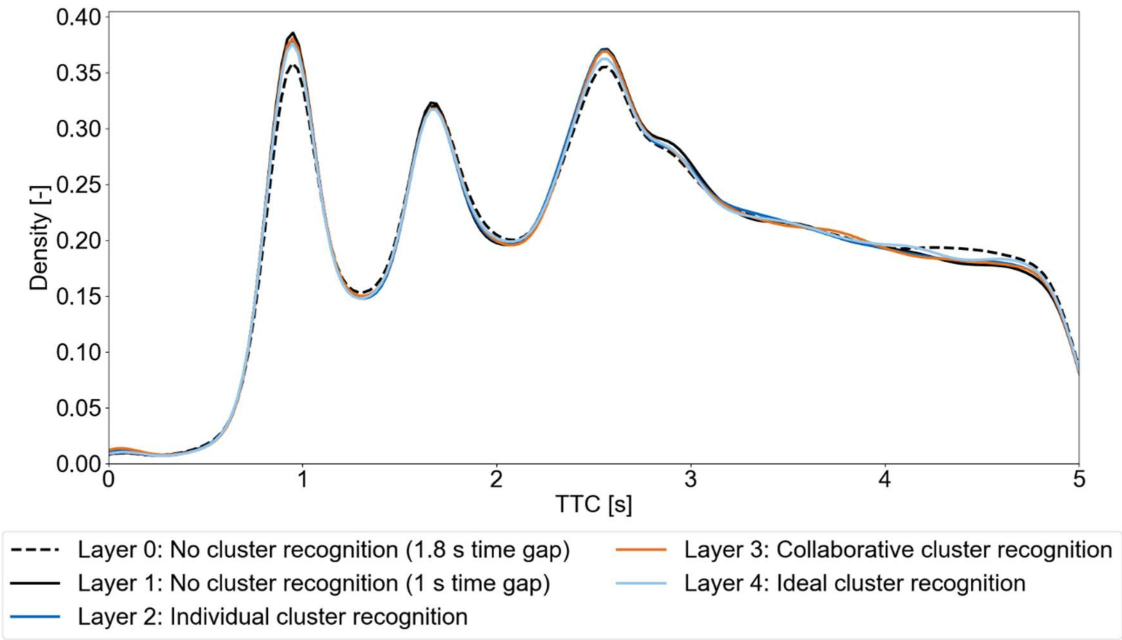

Figure 5.14: Distributions of the minimal TTC value per conflict in a range between 0 s and 5 s of human-driven vehicles with preceding automated vehicles for all investigated driver type recognition procedures. The recognition procedures of the automated vehicles only address the leading vehicle.

Figure 5.14 shows three distinct peaks with a TTC of around 0.95 s, 1.7 s and 2.55 s for all five considered layers. As in situation (a), layer 1 indicates the largest proportion of low TTCs. Furthermore, the application of any detection method reduces the proportion of low TTCs. Compared to the other methods, the ideal driver type detection method can avoid most situations with low TTCs. However, the overall effect of driver detection on traffic safety improvement is significantly smaller than in Figure 5.13. As in situation (a), the use of layer 0 has the greatest impact on avoiding critical situations. By keeping a permanent time gap of 1.8 s to the leading vehicle, critical situations with the following vehicle can be significantly reduced due to the more anticipatory driving behavior.

Comparing Figure 5.13 and Figure 5.14, the question might arise why there are three peaks in all TTC distributions of the human drivers. The explanation for the occurrence and position of the three distinct peaks can be found if the individual clusters are considered separately. Figure 5.15 shows the distributions of minimal TTC for situation (b) exemplarily for the use of the ideal cluster detection method separated for the four driver types. As the distributions for the other detection methods are almost identical, they are not shown here.

The distributions of the individual driver types each show two local maxima which partly cover different TTC ranges. The first peak in Figure 5.14 at around 0.95 s results from the superimposed peaks of the *Aggressive drivers*, the *Moderate drivers*, the *Speeders* and, slightly offset, the *Timid drivers*. The peak can be attributed to the approach of the drivers

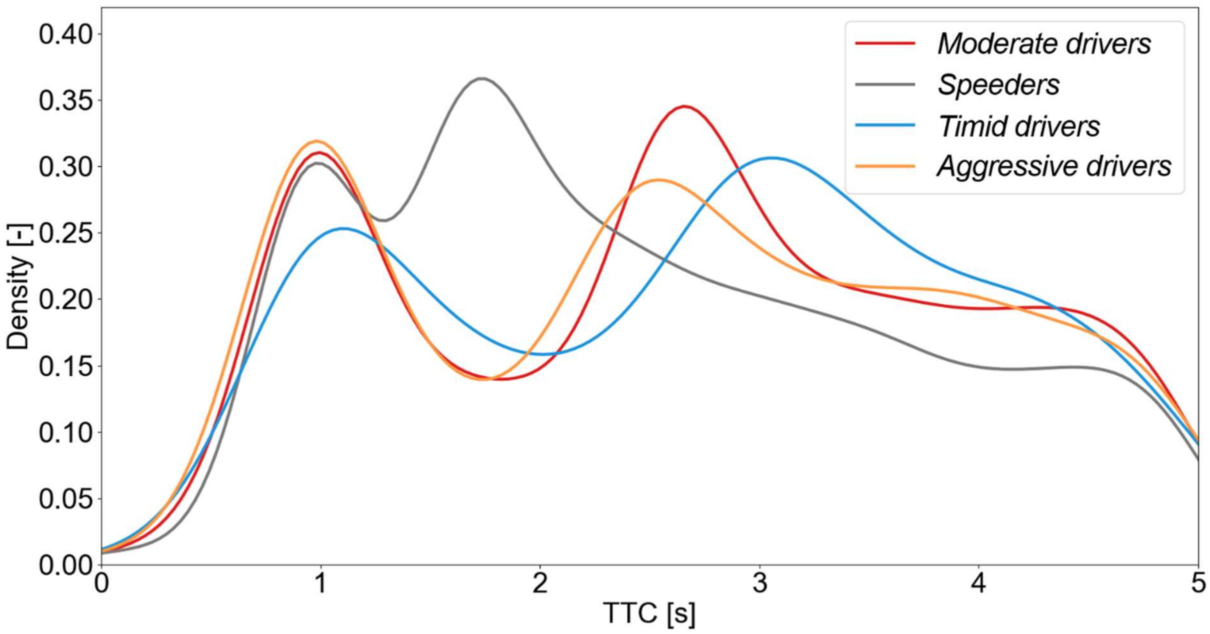

Figure 5.15: Distributions of the minimal TTC value per conflict in a range between 0 s and 5 s of human-driven vehicles with preceding automated vehicles broken down to the four driver types exemplarily using the ideal driver type recognition procedure. The recognition procedures of the automated vehicles only address the leading vehicle.

towards a queue in front of a traffic light. Apart from the *Timid drivers*, the other groups do not show pronounced driver type differences for the first peak. The *Aggressive drivers* show the highest proportion of low TTC values. However, due to their cautious driving behavior, the *Timid drivers* show a significantly lower proportion of low TTCs compared to the other driver types. The second peak in Figure 5.14 is determined exclusively by the *Speeders*. All other driver types show a minimum in the TTC distribution at this point. The third peak results from the superposition of the shifted maxima of the *Aggressive drivers*, the *Moderate drivers* and the *Timid drivers*. As with the first peak, the *Timid drivers* are characterized by the most non-critical TTC values. The second and third peaks in Figure 5.14 are referred to driver type characteristics in dynamic driving situations with the leading automated vehicle. All drivers react differently to stimuli of the automated vehicles. Therefore, the second peak of each driver group in Figure 5.15 is located at different TTC ranges.

From Figure 5.13 and Figure 5.14, it can be concluded that the use of a driver detection method reduces low-TTC situations in both considered situations (a) and (b). The collaborative method prevents more critical conflicts than the individual method. Both are outperformed by the ideal method. However, permanently keeping a larger time headway shows the best impact on traffic safety. Not only “expected” critical situations due to the presence of an *Aggressive* or *Timid* driver type can be reduced, but also “unexpected” conflicts that might occur with all driver types. However, it can be assumed that other phenomena that occur due to the larger time headways and are not covered by the traffic simulation, like sudden cut-ins, could impair traffic safety.

5.6.3 Evaluation of Average Number of Critical Conflicts per Vehicle and Hour

The second analyzed parameter is the number of critical conflicts per vehicle and per hour for each of the layers. Figure 5.16 depicts the average conflict numbers of automated vehicles and leading human-driven vehicles per hour.

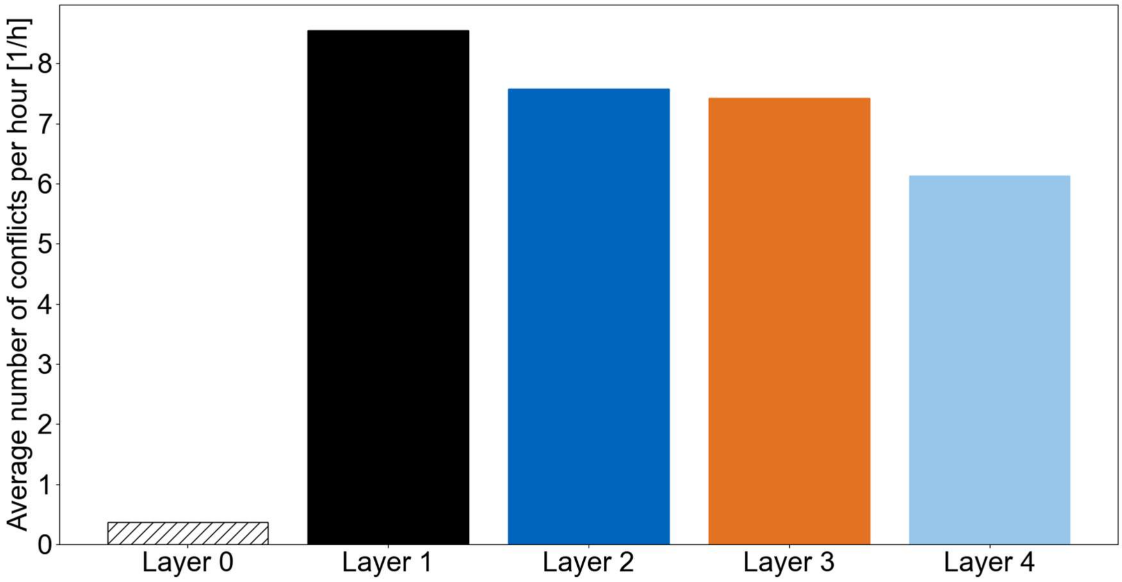

Figure 5.16: Average number of critical conflicts (TTC < 2 s) per vehicle and hour of automated vehicles in interaction with leading human-driven vehicles for all investigated driver type recognition procedures.

In Figure 5.16, layer 0 provides significantly lower conflict numbers than the other layers. Automated vehicles constantly keeping 1.8 s time gap are only faced with 0.4 critical conflicts per simulated hour. Layer 1 exhibits the most occurrences of critical situations with a value of 8.5 critical conflicts per hour, followed by layers 2 - 4 in this order. The difference between the individual and the collaborative layer is small, but still significant ($p = 0.02$). There is a large improvement in reduction of critical situations for the ideal cluster recognition method. Vehicles equipped with the ideal recognition method encounter 6.1 critical conflicts per hour, while those using the individual method face 7.6 conflicts per hour and those using the collaborative method 7.4 conflicts per hour.

Figure 5.17 shows the average critical conflict numbers for situation (b).

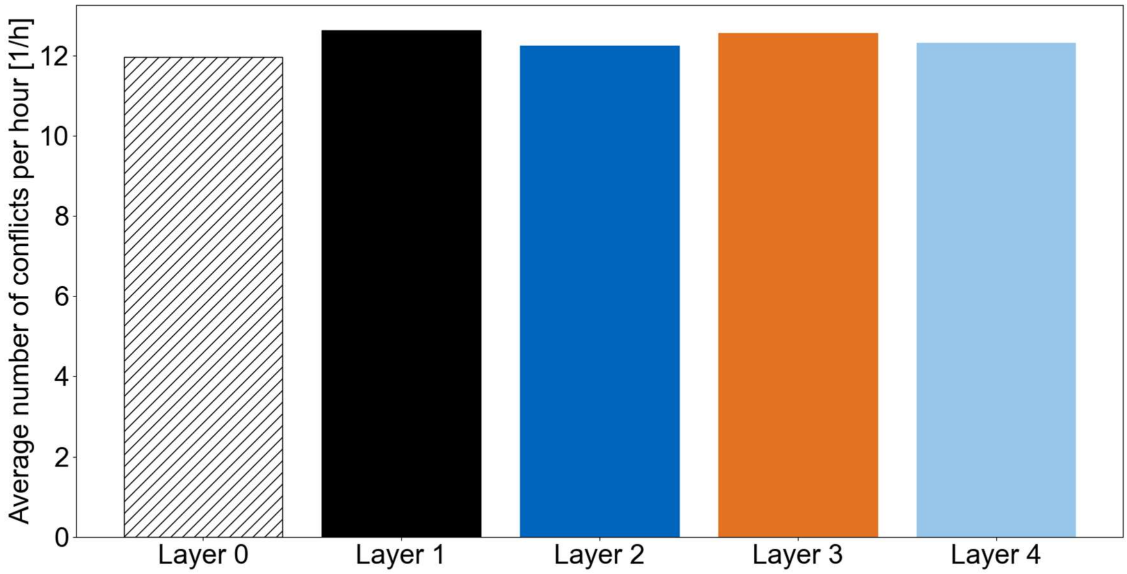

Figure 5.17: Average number of critical conflicts (TTC < 2 s) per vehicle and hour of human-driven vehicles with preceding automated vehicles for all investigated driver type recognition procedures. The recognition procedures of the automated vehicles only address the leading vehicle.

In the figure, all layers show very similar average critical conflict numbers at about 12 conflicts per vehicle and hour. The comparison does not reveal any significant differences between the conflict numbers of the layers. However, layer 0 shows a small difference in comparison with the other layers. Furthermore, it is evident that the total numbers in Figure 5.17 exceed the conflict numbers in Figure 5.16. Two effects mainly lead to the higher conflict numbers. On the one hand, data of human drivers are only evaluated when following an automated vehicle. In the data, the drivers are therefore increasingly exposed to conflict situations. On the other hand, human drivers controlled by LSTM must overcome PRT before they can react to surprise events like sudden breaking maneuvers of the leading vehicle. In comparison, automated vehicles are able to instantly react and consequently avoid lower TTCs. Therefore, critical conflict numbers of human drivers exceed the numbers of automated vehicles. Regarding the results of Figure 5.17, it is plausible that there could be a stronger effect on traffic safety for the human-driven vehicle in situation (b). However, considering the limited scope of a traffic simulation compared to detailed vehicle simulation and the chosen surrogate measure TTC for traffic safety, a more detailed simulation of the situation could reveal other effects that cannot be covered by traffic simulation.

For the automated vehicles, the results shown in Figure 5.16 emphasize the results from Section 5.6.2. The more sophisticated the driver type recognition method the more successfully critical conflicts can be avoided. However, for the human drivers, Figure 5.17 does not show considerable differences for the recognition methods. Layer 0 stands out from the other layers as it shows a significant reduction of critical conflicts in all cases. Average numbers of conflicts appear higher for the human drivers, as humans have to compensate for HFs such as PRT, whereas automated vehicles have the ability to react instantly.

5.6.4 Evaluation of Time in Conflict and Difference in Average Velocity

Minimal TTC value distributions and conflict numbers covered in Sections 5.6.2 and 5.6.3 provide information about the maximum intensity and quantitative number of conflicts. However, the temporal course of the conflict is not taken into account by these measures. Therefore, the average dwell time in conflict per vehicle and hour, while TTC is always below 2 s, is analyzed. For the analysis of traffic safety, the average dwell times are weighted by the travel times of the vehicles so that vehicles that have just departed in simulation are considered less. In order to investigate the impact on traffic efficiency, average velocity is selected as a key indicator for traffic flow.

Figure 5.18 depicts the difference in average velocity of automated vehicles with respect to layer 1 on the y-axis. The average velocity is measured during the entire simulation time and does not only refer to conflicts. On the x-axis, Figure 5.18 shows the average dwell time in conflicts per hour and vehicle weighted by each vehicle’s travel time for situation (a).

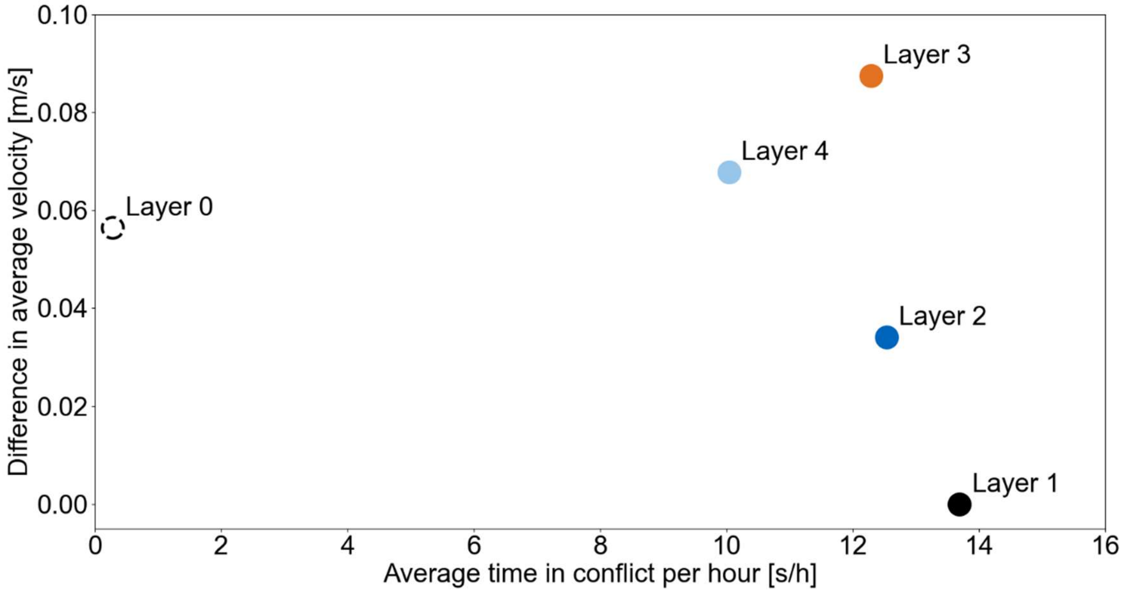

Figure 5.18: The difference in average velocity compared to layer 1 is shown over the average dwell time in critical conflicts (TTC < 2 s) per vehicle and hour which is weighted by each vehicle’s travel time. Both quantities are related to automated vehicles in interaction with leading human-driven vehicles for all investigated driver type recognition procedures.

Regarding traffic safety, the average dwell time in conflict decreases with a more elaborated driver detection method. Without any adjustments in layer 1, automated vehicles remain in critical conflicts for 13.2 s per hour. Layer 2 and 3 do not show large differences in reduction of the time in conflict. Both layers share a time in conflict at about 12 s per hour. The ideal cluster recognition method is able to reduce the time period to 9.7 s per hour which is a significant decrease in comparison with the other methods. However, the automated vehicles of layer 0 can almost totally avoid conflicts below a TTC of 2 s. These vehicles are only confronted with 0.3 s per hour in critical conflicts.

The shown differences in average velocity are only marginal in a range of 0.1 m/s. It appears that automated vehicles that keep larger time gaps show slightly higher average velocities than vehicles of layer 0. These automated vehicles might act as a buffer, making the traffic flow smoother, offering gaps for tactical lane changes and avoiding unnecessary stops. In contrast to the general relationship in traffic theory in which traffic safety and efficiency are mutually dependent, the driver recognition methods seem to address WARDROP's [1952] second principle which states a social and cooperative equilibrium. The larger time headways between vehicles appear to shift the traffic system slightly towards a more optimal state.

Figure 5.19 shows the average velocity difference of human-driven vehicles with respect to human-driven vehicles which follow automated vehicles of layer 1. The average velocity differences are plotted over the average dwell time in conflicts per hour and vehicle for each recognition method. Figure 5.19 addresses situation (b).

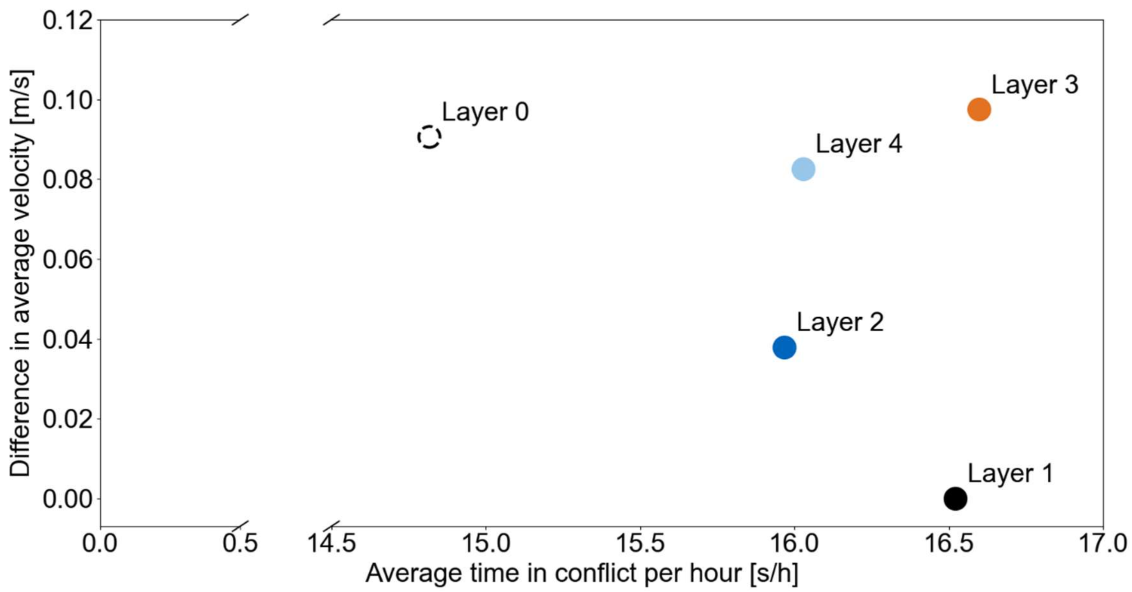

Figure 5.19: The difference in average velocity compared to layer 1 is shown over the average dwell time in critical conflicts (TTC < 2 s) per vehicle and hour which is weighted by each vehicle's travel time. Both quantities are related to human-driven vehicles with preceding automated vehicles for all investigated driver type recognition procedures. The recognition procedures of the automated vehicles only address the leading vehicle.

In situation (b), the reductions in dwell time are lower for each driver type detection method compared to the results of situation (a) (cf. Figure 5.18). The application of the collaborative method does not show any dwell time reduction for the human-driven vehicle behind. The usage of the individual and the ideal methods only slightly decreases the dwell time. It can be assumed that the shift of layer 3 towards higher dwell time is due to statistical dispersion. Layer 0 shows the largest reduction in dwell time compared to the other layers.

The differences in average velocity are below 0.1 m/s. As already noticed for situation (a), automated vehicles keeping a larger time gap might indirectly encourage a more cooperative behavior by buffering the traffic and enhancing the traffic flow. This could also affect the vehicle driving behind.

5.6.5 Economic Benefit Calculation from the Accident Avoidance Potential

The reduction of dwell time in critical situations determined in Section 5.6.4 can be used to estimate the accident avoidance potential of urban rear-end collisions for each driver detection method. In order to illustrate this potential, the economic benefit within one year is calculated for the simulated area of Ingolstadt. Assuming that Ingolstadt serves as a model city in Germany, the numbers can be extrapolated to a nationwide estimate of the cost reduction in inner-city rear-end collisions. This traffic economic assessment serves as an exemplary use case for the further application of the developed driver behavior model beyond the virtual validation of automated vehicles.

The following three assumptions are made for the relationship between critical situations and the accident density in longitudinal traffic. The accident density $d_{collision}$ in a defined period of time (e.g., one year) is defined by the total number of accidents $n_{accidents}$ in this period related to the average dwell time in critical situations per vehicle per hour t_{krit}.

1. All longitudinal traffic accidents are attributed to dangerous conditions. This results in a high accident density $d_{collision}$ in these critical situations.
2. The accident density in dangerous conditions $d_{collision,critical}$ is constant. In other words: A linear relationship between the number of accidents $n_{accidents}$ and the average dwell time in critical situations per vehicle per hour t_{krit} is assumed.
3. The accident density in non-hazardous conditions $d_{collision,uncritical}$ is low, so that it is neglected in the following.

Equation (33) shows the assumptions 1. - 3. in mathematical form:

$$d_{collision} = d_{collision,critical} + d_{collision,uncritical} \approx d_{collision,critical} = \frac{n_{accidents}}{t_{krit}} = \text{const.} \tag{33}$$

Regarding traffic safety, the traffic state that results from the introduction of any driver type recognition procedure x, compared to the reference state without cluster recognition, is given by:

$$\frac{n_{accidents,x}}{t_{krit,x}} = \frac{n_{accidents,reference}}{t_{krit,reference}}. \tag{34}$$

Accordingly, the number of accidents for a traffic state which results by the introduction of the driver type recognition procedure x can be calculated from the number of accidents in the reference state and the mean dwell times in critical situations of both states. The number of accidents in the reference state is determined from the German accident atlas ("Unfallatlas") [STATISTISCHE ÄMTER DES BUNDES UND DER LÄNDER, 2023]. In order to compensate for annual fluctuations, the accident numbers for accidents in longitudinal traffic in Ingolstadt from 2016 to 2021 are considered and averaged over the years. Since the accident atlas only includes accidents involving personal injury, the total number of accidents in Ingolstadt is additionally extrapolated from data from the German Federal Statistical Office using the proportions of personal accidents to all other accident types. [STATISTISCHES BUNDESAMT DESTATIS, 2023a]. The calculated accident numbers for all accident severities with and without personal injury as well as incurred costs can be found in Appendix B.

For the calculation of the cost reduction from the accident avoidance potential, an additional assumption is made:

4. The ratio of the different accident types remains constant when the dwell time t_{krit} in critical conflicts is reduced. According to assumption 2, this means that the costs are linear to the number of accidents $n_{accidents}$.

The German Federal Highway Research Institute (BASt) yearly publishes the economic per-accident costs broken down to the different accident types [BUNDESANSTALT FÜR STRAßENWESEN, 2022]. With the accident type specific costs $c_{per\_accident,i}$ from 2022 and the accident numbers $n_{accidents,i}$ for each accident type i, the current total accident costs $c_{total,current}$ can be calculated using equation (35):

$$c_{total,current} = \sum n_{accidents,i} \cdot c_{per\_accident,i}\,. \tag{35}$$

However, these current costs do not consider the presence of automated vehicles in traffic. RÖSENER [2020] states that automated vehicles are able to reduce the number of collisions with a leading standing or moving vehicle in urban traffic by the factor of about 50 %. Taking into account assumption 4, the accident costs of the reference group, layer 1, can be calculated from the total current costs split to automated vehicles (AV) and human drivers (HD):

$$c_{AV,reference} = f_{red,AV} \cdot p_{AV} \cdot c_{total,current}, \tag{36}$$

$$c_{HD,reference} = (1 - p_{AV}) \cdot c_{total,current}, \tag{37}$$

where $f_{red,AV}$ is the mentioned reduction factor of 0.5 and p_{AV} is the proportion of automated vehicles in traffic which is set to 0.1 in all simulations. For the ratio of accident costs and dwell time, taking into account equation (34) and assumption 4 yields:

$$\frac{c_{AV,x}}{t_{AV,krit,x}} = \frac{c_{AV,reference}}{t_{AV,krit,reference}}, \tag{38}$$

$$\frac{c_{HD,x}}{t_{HD,krit,x}} = \frac{c_{HD,reference}}{t_{HD,krit,reference}}. \tag{39}$$

The cost reduction with respect to layer 1 can be expressed by:

$$c_{AV,\Delta} = c_{AV,reference} - c_{AV,x}, \tag{40}$$

$$c_{HD,\Delta} = c_{HD,reference} - c_{HD,x}. \tag{41}$$

The combination of equations (38) and (40) as well as (39) and (41) results in the following expressions for the cost reduction:

$$c_{AV,\Delta} = c_{AV,reference} \cdot \frac{t_{AV,krit,reference} - t_{AV,krit,x}}{t_{AV,krit,reference}}, \tag{42}$$

$$c_{HD,\Delta} = c_{HD,reference} \cdot \frac{t_{HD,krit,reference} - t_{HD,krit,x}}{t_{HD,krit,reference}}. \tag{43}$$

The following Figure 5.20 shows the estimated cost reduction in Ingolstadt per year of automated vehicles and human drivers for the different driver recognition procedures as a stacked bar plot. As mentioned above, all evaluations are conducted for a penetration rate of automated vehicles of 10 %. The lower non-dotted bars represent the cost savings of automated vehicles that apply the respective procedure within situation (a), cf. Figure 5.12. The dotted areas above show the cost reductions for the human drivers driving behind the automated vehicles addressing situation (b), cf. Figure 5.12. All cost reductions are derived from the dwell time differences shown in Figure 5.18 for the automated vehicles and Figure 5.19 for the human drivers according to equations (42) and (43). Note that for the cost reduction of human drivers of layer 3, the savings of layer 2 are considered. Due to statistical dispersion, layer 3 does not show any dwell time deviation compared to the reference layer 1. However, a similar or even higher effect is assumed for layer 3 in comparison to layer 2 for plausibility reasons. All accident costs are obtained from the tables given in the Appendix B.

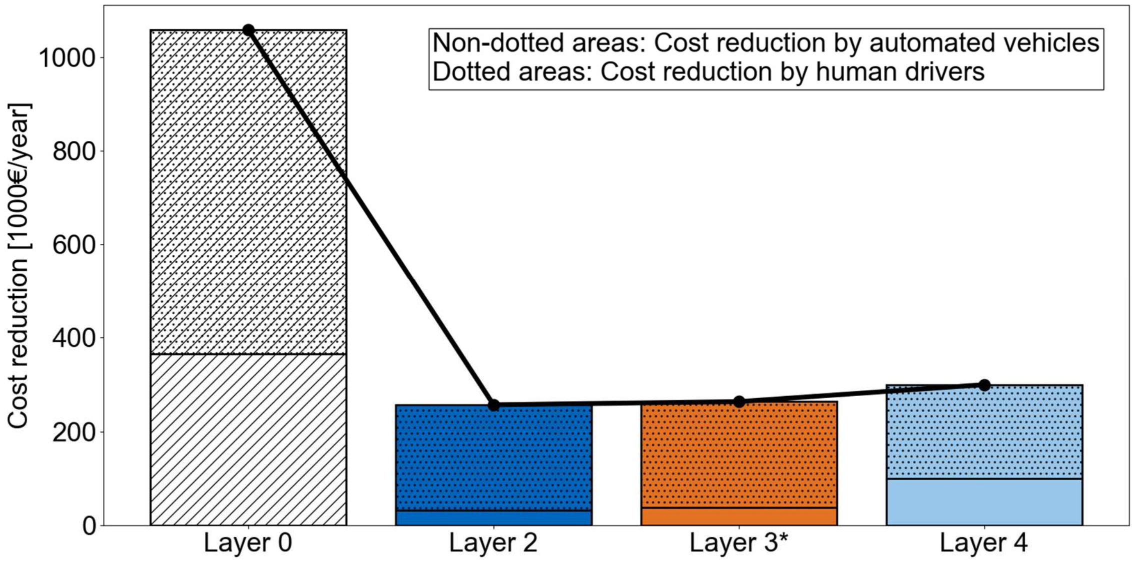

Figure 5.20: Estimated cost reduction in Ingolstadt per year by avoidance of inner-city rear-end collisions due to the usage of the different driver type recognition procedures in comparison to layer 1 (reference). The dotted areas represent the cost reduction by human drivers and the non-dotted areas the cost reduction by automated vehicles with a penetration rate of 10 %. *Layer 3 assumes the same cost reduction of human drivers as layer 2.

Table 5.1 provides the values shown in Figure 5.20. Additionally, the total costs are extrapolated using the numbers of registered vehicles in Ingolstadt [Stadt Ingolstadt, 2018] and Germany [Statista Research Department, 2023a] in 2018:

Layer	Cost reduction AV [1000 €]	Cost reduction HD [1000 €]	Total cost reduction [1000 €]	Extrapolated total cost reduction for Germany [Mio €]
0	366	693	1 059	515
2	32	225	257	125
3	38	225*	263*	128*
4	100	200	300	146

Table 5.1: Values of estimated cost reduction for Ingolstadt (in 1000 €) that are visualized in Figure 5.20. The last column estimates the total cost reduction for Germany (in Mio €). *Layer 3 assumes the same cost reduction of human drivers as layer 2.

In Figure 5.20, the largest reduction is achieved by layer 0 which reduces the total costs by about one million euros. The other layers, which apply the driver type detection, show total values between 250 and 300 thousand euros. The proportions of saved costs by automated vehicles range from 12 % to 34 % of the total cost reductions. This percentage range with respect to the overall costs is notable, because the penetration rate of automated vehicles is only 10 % and the reduction factor $f_{\text{red},AV}$ even lowers the costs by 50 %, cf. equation (36). The results emphasize the effectiveness of the driver detection methods on traffic safety. For the human drivers in situation (b), the effects are significantly lower under consideration of the share of human drivers in traffic.

Regarding the quantitative values in Table 5.1, the cost reduction of each procedure seems very low compared to published overall accident costs in traffic, e.g., 32.5 billion euros in Germany in 2021 [STATISTA RESEARCH DEPARTMENT, 2023b]. However, the cost reductions shown in Figure 5.20 and Table 5.1 only address rear-end collisions in the urban area of Ingolstadt. Rear-end collisions especially in urban areas do usually not lead to severe accidents with personal injuries but rather car body damages which cause considerably less costs for the national economy. Furthermore, the procedures probably affect other correlated accident types that have not been evaluated within this thesis, e.g., accidents in cut-in scenarios. The consideration of other accident types could increase the cost reductions significantly. In this thesis however, the focus is on the general methodology of realistic driver behavior modeling, especially HFs, and the recognition of the driver types with respect to an exemplary impact on traffic safety and efficiency without claiming full coverage of all possible effects.

In order to estimate the nationwide cost savings, all values are extrapolated using the numbers of registered vehicles in Ingolstadt (approx. 96000) and Germany (approx. 46.5 million). The calculation is based on the assumption that Ingolstadt serves as a model city for other German urban areas. Regarding the last column of Table 5.1, the cost reduction potential ranges from 125 million euros when considering layer 2 up to 515 million euros under the conditions of layer 0. The ideal recognition procedure can reach cost savings of 146 million euros in Germany. Although all calculated values represent only rough estimates of the cost reductions, the potential of the driver detection methods which are assessed in a virtual test, in which the human drivers in the surrounding traffic are controlled by realistically behaving driver models, becomes obvious.

The shown cost reductions are contrasted by the incurred costs that are associated with the driver type detection methods. These include costs for series development of the detection methods, integration into the vehicle, testing and equipment with the needed sensors. However, scaling the software results in relatively low costs for the individual vehicle. Furthermore, no additional sensors are required that are not already installed in modern vehicles. If most tests are done virtually, the costs for testing can be kept low. Regarding the collaborative approach, further costs for a data protection compliant solution of a data base arise. All in all, the costs are relatively small compared to the benefits which result from a traffic safety improvement. Car manufacturers can further promote these safety gains in advertising to attract more customers.

6. Conclusion and Future Research

This chapter concludes the dissertation by first summarizing the key findings and discussing the results in Section 6.1, followed by suggestions for potential future research areas in Section 6.2.

6.1 Conclusion

In this thesis, a data-driven driver model for longitudinal driver behavior has been developed, in which HFs and especially individuality were integrated. The motivation for the development of the model was based on the high requirements for the validity of simulations posed by virtual testing of automated vehicles. Therefore, in the simulated dynamic environment of a VuT, driver models need to realistically reproduce human driver behavior. Engineering models used in traffic simulation idealize the behavior by assuming accurate perception of ground truth and immediate response to stimuli [SAIFUZZAMAN AND ZHENG, 2014]. Proper modeling of these features including their relationship to HFs is essential for a realistic virtual assessment of automated vehicle performance, particularly regarding safety. These challenges motivated previous efforts to integrate HFs, including latency, into driver behavior models [TREIBER et al., 2006]. In recent years, especially data-driven approaches have shown that they have the potential to outperform classical engineering driver models in their ability to reproduce human behavior [MORTON et al., 2017]. Therefore, in this thesis, a data-driven driver behavior model based on LSTM has been developed that takes HFs into account.

This thesis comprised three methodological parts. In the first part of the thesis, the methodology of the developed driver model incorporating HFs was described in detail. The focus was on modeling driver behavior when approaching a signalized intersection. According to the psychological concept, the developed model was divided into three components: perception, information processing and action. The perception module focuses on input parameters that are perceivable by humans with a certain degree of inaccuracy. The information processing module includes a short-term memory of 1 s as well as the ability of anticipation of 2 s into the future. The main component of the action module is an LSTM that uses the preprocessed information to determine a desired acceleration. However, the input data are not provided to LSTM immediately, but delayed by the PRT.

The training of LSTM was performed on a trajectory data set obtained from camera observations of a signalized intersection and from xFCD. From the data set, four driver types were identified using k-means clustering. These types were designated as *Moderate drivers*, *Speeders*, *Timid drivers*, and *Aggressive drivers*. In order to learn the characteristics of each driver type, LSTM was trained separately for each type. The training of LSTM was performed in a closed-loop approach, allowing the model to learn from its own (mis-)behavior. The chosen loss function addresses realistic car-following behavior, but also takes into account known aspects of human driving, such as a smooth acceleration profile and safe driving features.

Regarding the modeling of the driver behavior model, from the first two research questions

RQ 1.1: What methods can be used to model human driver behavior considering HFs?

RQ 1.2: How can the individuality of individual and different drivers be captured and replicated in driver behavior modeling?

it can be concluded that

- a data-driven method based on LSTM is suitable because it is able to process temporary sequences of input parameters a human driver is able to perceive, which go beyond the parameters commonly used in engineering models,
- the limitations of LSTM related to implicitly learning the human psychological processes must explicitly be modeled and considered in the training process, and
- individual driver behavior can be achieved by clustering the driver population and training the model on different driver type specific data sets.

Assessment of the developed model was performed separately for each driver type. It was evaluated how well the model can reproduce real trajectories compared to FVD and Krauss models, which were also calibrated for each driver type. LSTM was found to perform better in reproducing acceleration, speed, and distance keeping compared to the engineering models. In addition, the behavior of the models was analyzed on a subset of the data set where drivers were challenged by traffic light phase transitions. Even in these situations, LSTM showed the smallest deviations from real human driver behavior in almost all comparisons.

Regarding the first research question related to the application of the developed driver behavior model

RQ 2.1: Can the developed driver behavior model, which takes into account HFs in its architecture, reproduce human driver behavior more realistically than existing driver behavior models?

it can be answered that

- LSTM outperforms the investigated engineering models in reproducing real-world trajectories regarding the whole test data set, and
- the performance of the LSTM is still superior to the other models when considering only a subset of the test data set in which HFs play an important role.

The methodological second part of the thesis described the setup and calibration of the traffic simulation of the German city of Ingolstadt using SUMO. The simulation served as the virtual test environment in the third part of the thesis. The three-step calibration methodology consisting of map association, traffic light emulation and traffic assignment led to an hourly calibration of the simulation. The map association of signal groups of traffic lights and detector loops enabled the use of real observed data to recreate real traffic light programs and traffic flows in simulation. The calibration of traffic flows was based on OD matrices obtained from statistical information. The traffic flows were further iteratively refined using real traffic counts. For trips that could not be finished within the calibrated one-hour interval, a clearing period was introduced. As seen in Section 4.4, overall median errors of traffic flow with respect to real observations ranged from 13 % to 18 %, part of which reflect natural statistical variations. Traffic lights showed median deviations of 1.6 % to 1.9 % in green time relative to the cycle time. This very high accuracy of green times suggests that the key dynamical properties of traffic flow at intersections come close to real traffic.

In the last part of the thesis, the developed driver model was integrated into the traffic simulation of Ingolstadt. The comparison of real observed and simulated conflict situations between human drivers demonstrated the suitability of the developed driver behavior model for analysis of safety critical conflicts. In the next step, automated vehicles were implemented into the traffic model for virtual testing. The automated vehicles were equipped with three different variants of driver type recognition methods that monitored the leading vehicle. The detection of the *Aggressive* or *Timid* driver type induced the time headway of the automated vehicles to be increased from 1 s to 1.8 s. While the *Aggressive drivers* were reliably detected, there were uncertainties with the other driver types, which could be resolved by improving the detection procedure in future. The recognition of driver types and adaptation of the time headway led to a significant reduction of critical rear-end conflicts for the automated vehicles. Furthermore, a slight reduction in criticality was also observed for human drivers following automated vehicles. While traffic safety is improved by the recognition methods, traffic efficiency is not impaired. Moreover, adapted distance keeping of the automated vehicles could provide a key mechanism for smoothing traffic flow. In order to show the economic benefit of the driver behavior modeling and recognition, the cost reductions implied by the traffic safety gains were estimated in a range from 257 thousand € to 1 million € for Ingolstadt and 125 million € to 515 million € for Germany per year assuming 10 % penetration of automated vehicles with the driver type recognition algorithm.

Regarding the last two research questions related to the application of the developed model

RQ 2.2: Can the developed driver model reproduce the behavior of specific driver types so that the type can be recognized by an automated vehicle?

RQ 2.3: Does the adaptation of driving behavior of automated vehicles to specific recognized driver types impact traffic safety and efficiency?

it can be concluded that

- the developed model is able to reproduce the different behavior patterns of different driver types so that automated vehicles are able to recognize the type,

- the adaptation of the driving strategy to the driver types with the most potential for pronounced decelerations leads to an improved safety for the automated vehicles,
- the safety for human drivers is not significantly impacted by the adaptations, and
- no significant influence of the adaptations on traffic efficiency could be determined.

To conclude, the findings of this thesis show that a data-driven approach which incorporates HFs is able to reproduce human driver behavior. The model is even able to learn the behavior of different driver types. The developed driver model can be included into a virtual traffic system for the tests of automated vehicles. From these tests, potentials and deficiencies of the automated vehicles can be derived.

6.2 Future Research

In this thesis, the developed driver behavior model has addressed urban approaches towards signalized intersections. The scope of the model could be further extended in the future. On the one hand, lateral motion could be added; however, the strong interdependence of longitudinal and lateral movements could present a substantial challenge. On the other hand, other aspects of urban traffic like turning or intersection behavior could be incorporated. Furthermore, the model could be extended to rural or highway scenarios in order to cover other types of roads which show different traffic characteristics. For all mentioned aspects, the collection of further data would be needed so that the model could be extended with all parameters which are relevant for the addressed situation. However, regarding the data collection, it could be a challenge to cover the whole spectrum of situations (from uncritical to critical) which could appear in each traffic scenario. It is further suggested to split the model into different models which apply to the different fields of application. This means, for example, that there could be different models for the approach to a traffic light and the subsequent behavior within the intersection. As a result, the different models can better learn the specific characteristics of the situation they are trained for, instead of reflecting one generalized behavior. However, switching models can also pose a challenge, as it is demonstrated in this thesis regarding the transition between car-following and free driving (cf. Section 5.2.2). Future work generalizing the phenomena considered in this thesis to additional maneuvers could reveal additional safety performance potential in automated driving.

The performance of the developed driver behavior model is currently limited considering the data set used for training. More data and longer observation periods of more traffic situations could extend the scope of the model, e.g., an implicit learning of anticipation. A larger quantity of data could also improve the quality of driver type clustering in order to reduce the number of drivers who are assigned to the wrong cluster. Additional data could also improve the driver type recognition quality of the automated vehicles within the conducted endurance run. The performance of the developed driver behavior model could also be increased by adding more information like the vehicle in front of the leader or vehicles in adjacent lanes. Additionally, there are other HFs such as perceptual thresholds, variable PRT, cognitive load, or distraction which offer the potential to further approximate the behavior of the model to real human driver behavior. When simulating a whole traffic system, the data preparation for the developed driver model requires high computing power. To counter the computational effort, either parallel data preparation on multiple cores, the usage of stronger machines or both could be applied.

Like most deep learning models, LSTMs are often viewed as "black boxes." Understanding why the model made a particular prediction about a driver's behavior can be difficult, which is a significant drawback for safety-critical applications like autonomous driving where explainability is crucial. Validation of data-driven methods used for behavior models in virtual testing of automated vehicles is therefore an extensive future research field.

The requirements of a virtual testbed towards the degree of detail exceed the scope of a traffic simulation software like SUMO, e.g., in simulation of vehicle physics. In the future, virtual tests of automated vehicles could be performed in a co-simulation with a high-precision vehicle simulation where SUMO generates a realistic traffic flow in the network. Such a simulation could also incorporate a high-precision road environment including buildings and nature which could influence the driver, e.g., through shadow and light changes or sun reflections on glass facades of buildings. For a realistic endurance run, other traffic participants including VRUs must also be considered. For a valid simulation, these traffic participants must also be reflected by realistic behavior models.

Further development of realistic driver behavior models along these lines could contribute enormously to the validity of virtual testing in the field of automated driving concepts.

List of Figures

List of Tables

List of Terms and Abbreviations

ABS	Anti-Lock Braking System
ACC	Adaptive Cruise Control
AI	Artificial Intelligence
AR	Anticipation reliance
Car2Car	Car to car communication
Car2X	Car to infrastructure communication
ESC	Electronic Stability Control
EU	European Union
FCD	Floating Car Data
FVD	Full Velocity Difference model
GEH	Traffic volume measure according to Geoffrey E. Havers
GHR	Gazis-Herman-Rothery model
HDM	Human Driver Model
HF	Human Factor
IDM	Intelligent Driver Model
JND	Just Noticeable Difference
KDE	Kernel Density Estimate
LSTM	Long Short-Term Memory
MICE	Multiple Imputation by Chained Equations
MSE	Mean Squared Error
NGSIM	Next Generation Simulation
NRMSE	Normalized Root Mean Square Error
NPV	Negative Predictive Value
OD	Origin-Destination
OVM	Optimal Velocity Model
PCA	Principal Component Analysis
PPV	Positive Predictive Value
PRT	Perception-Reaction-Time
R-CNN	Region-based Convolutional Neural Network
RNN	Recurrent Neural Network
RMSE	Root Mean Square Error
TD	Task Demand
THW	Time Headway
TraCI	Traffic Control Interface
TTC	Time-To-Collision
TTJ	Time-To-Junction
SA	Situational Awareness
SAE	Society of Automotive Engineers
SEEV	Salience, Effort, Expectancy, Value

SUMO	Simulation of Urban MObility
VRU	Vulnerable Road User
VuT	Vehicle Under Test
xFCD	Extended Floating Car Data

Statement about the Use of AI

In this thesis, AI was used for the following tasks:

- Modeling of human driver behavior using LSTM networks with the TensorFlow library in Python
- Object detection within camera images using Mask R-CNN with the Detectron2 library in Python
- Support with programming using the tool ChatGPT
- Grammar and spell check using the tool Deepl

Publications

LANGER, MARCEL; SCHIEN, THOMAS; HARTH, MICHAEL; KATES, RONALD; BOGENBERGER, KLAUS (2020). "An Improved Moving Observer Method for Traffic Flow Estimation at Signalized Intersections". In: *IEEE Intelligent Vehicles Symposium (IV), Proceedings*

REMLINGER, WOLFRAM; HARTH, MICHAEL; LANGER, MARCEL (2020). "Verfahren zum Betrieb eines Fahrzeugs und Fahrzeug". DE 10 2019 126 109 A1. Deutsches Patent- und Markenamt. https://register.dpma.de/DPMAregister/pat/register?AKZ=1020191261090&CURSOR=48

REMLINGER, WOLFRAM; HARTH, MICHAEL; LANGER, MARCEL (2020). "Verfahren zur Verarbeitung von Verkehrs-Informationen, Ego-Kraftfahrzeug und Computerprogrammprodukt". DE 10 2019 133 467 A1. Deutsches Patent- und Markenamt. https://register.dpma.de/DPMAregister/pat/register?AKZ=1020191334675&CURSOR=49

HARTH, MICHAEL; LANGER, MARCEL; BOGENBERGER, KLAUS (2021). "Automated Calibration of a Microscopic Traffic Flow Simulation Using Real-World Observations". In: Transportation Research Board.

SCHREINER, KARL; HANDEL, CAROLINE; OBSTBAUM, MARTIN; BEUTENMÜLLER, FRANK; MORICH, LUTZ; LANGER, MARCEL; HARTH, MICHAEL (2021). „Virtual Validation of Autonomous Driving Functions in Urban Areas with Mixed Traffic". In: *21. Internationales Stuttgarter Symposium: Automobil- und Motorentechnik*. pp. 261-270. Wiesbaden. Springer Fachmedien Wiesbaden.

HARTH, MICHAEL; SAJID ALI, MUHAMMAD; KATES, RONALD; BOGENBERGER, KLAUS (2021). "Data-Driven Modelling of Car-Following Behavior in the Approach of Signalized Urban Intersections". In: *IEEE Conference on Intelligent Transportation Systems, Proceedings*.

HARTH, MICHAEL; BIN AMJAD, UZAIR; KATES, RONALD; BOGENBERGER, KLAUS (2022). "Incorporation of Human Factors to a Data-Driven Car-Following Model". In: *Transportation Research Record* 2676.10. pp. 291-302.

HARTH, MICHAEL; LANGER, MARCEL; BOGENBERGER, KLAUS (2022). "Automated Calibration of Traffic Demand and Traffic Lights in SUMO Using Real-World Observations". In: *SUMO User Conference 2021* 2. Pp. 133-148.

Bibliography

ASSOCIATION FOR STANDARDIZATION OF AUTOMATION AND MEASURING SYSTEMS (2022). *"ASAM OpenDRIVE"*. https://www.asam.net/standards/detail/opendrive/

BAGSCHIK, GERRIT; MENZEL, TILL; MAURER, MARKUS (2018). "Ontology based Scene Creation for the Development of Automated Vehicles". In: *IEEE Intelligent Vehicles Symposium, Proceedings*, pp. 1813–1820. DOI: 10.1109/IVS.2018.8500632

BAIRD, JOHN C.; NOMA, ELLIOT (1978). *Fundamentals of Scaling and Psychophysics*. John Wiley & Sons.

BANDO, MASAKO; HASEBE, KATSUYA; NAKANISHI, KEN; NAKAYAMA, AKIHIRO (1998). "Analysis of optimal velocity model with explicit delay". In: *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* 58.5, pp. 5429–5435. DOI: 10.1103/PhysRevE.58.5429

BEN-AKIVA, M.; CHOUDHURY, C.; TOLEDO, T. (2006). "Lane Changing Models". In: *Proceedings of the International Symposium of Transport Simulation*.

BOCK, JULIAN; KRAJEWSKI, ROBERT; MOERS, TOBIAS; RUNDE, STEFFEN; VATER, LENNART; ECKSTEIN, LUTZ (2019). *"The inD Dataset: A Drone Dataset of Naturalistic Road User Trajectories at German Intersections"*. http://arxiv.org/abs/1911.07602

BROOKHUIS, KAREL A.; DE WAARD, DICK (2010). "Monitoring drivers' mental workload in driving simulators using physiological measures". In: *Accident Analysis and Prevention* 42.3, pp. 898–903. DOI: 10.1016/j.aap.2009.06.001

BRUNNER, PASCAL; DENK, FLORIAN; HUBER, WERNER; KATES, RONALD (2019). "Virtual safety performance assessment for automated driving in complex urban traffic scenarios". In: *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 679–685. https://api.semanticscholar.org/CorpusID:208634175

BUBB, HEINER; BENGLER, KLAUS; GRÜNEN, RAINER E.; VOLLRATH, MARK (2015). *Automobilergonomie*. Springer Vieweg. DOI: 10.1007/BF02027395

BUNDESANSTALT FÜR STRAßENWESEN (2022). *"Volkswirtschaftliche Kosten von Straßenverkehrsunfällen in Deutschland"*. DOI: 10.1055/s-0034-1381987

BUNDESMINISTERIUM FÜR DIGITALES UND VERKEHR (2018). *"Research Projekt SAVe:"*. https://www.bmvi.de/SharedDocs/DE/Artikel/DG/AVF-projekte/save.html

BUNDESMINISTERIUM FÜR DIGITALES UND VERKEHR (2021). *"Research Project SAVeNoW"*. https://www.bmvi.de/SharedDocs/DE/Artikel/DG/AVF-projekte/savenow.html

CALDERONE, JACK B.; KAISER, MARY K. (1989). Visual acceleration detection: Effect of sign and motion orientation. In: *Perception & Psychophysics* (Vol. 45, Issue 5, pp. 391–394). Psychonomic Society. DOI: 10.3758/BF03210711

CALVERT, SIMEON C.; SCHAKEL, WOUTER J.; VAN LINT, J. W. C. (2020). "A generic multi-scale framework for microscopic traffic simulation part II – Anticipation Reliance as compensation mechanism for potential task overload". In: *Transportation Research Part B: Methodological* 140, pp. 42–63. DOI: 10.1016/j.trb.2020.07.011

CAMPBELL, MARK; EGERSTEDT, MAGNUS; HOW, JONATHAN P.; MURRAY, RICHARD M. (2010). "Autonomous driving in urban environments: Approaches, lessons and challenges". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 368.1928, pp. 4649–4672. DOI: 10.1098/rsta.2010.0110

CHEN, XIAOYUN; SUN, JIAN; MA, ZIAN; SUN, JIE; ZHENG, ZUDUO (2020). "Investigating the long- and short-term driving characteristics and incorporating them into car-following models". In: *Transportation Research Part C: Emerging Technologies* 117.

COLYAR, JAMES; HALKIAS, JOHN (2007). *"NGSIM - US Highway 101 Dataset"*. https://www.fhwa.dot.gov/publications/research/operations/07030/07030.pdf

CONSTANTINESCU, Z.; MARINOIU, C.; VLADOIU, M. (2010). "Driving style analysis using data mining techniques". In: *International Journal of Computers, Communications and Control*

5.5, pp. 654–663. DOI: 10.15837/ijccc.2010.5.2221

DEUTSCHES ZENTRUM FÜR LUFT- UND RAUMFAHRT E.V. (DLR); DEUTSCHES INSTITUT FÜR WIRTSCHAFTSFORSCHUNG (DIW) (2022). *Verkehr in Zahlen 2022/2023* (Kraftfahrt-Bundesamt (ed.); 51st ed.). Bundesministerium für Digitales und Verkehr. DOI: 10.1007/978-3-642-59063-4_10

DONGES, EDMUND (1982). "Aspekte der aktiven Sicherheit bei der Führung von Personenkraftwagen". In: *Automob-Industrie* 27.2.

DÖRR, DOMINIK; GRABENGIESSER, DAVID; GAUTERIN, FRANK (2014). "Online driving style recognition using fuzzy logic". In: *2014 17th IEEE International Conference on Intelligent Transportation Systems, ITSC 2014*, pp. 1021–1026. DOI: 10.1109/ITSC.2014.6957822

ECKSTEIN AND ZLOCKI (2013). "Safety Potential of ADAS: Combined Methods for an Effective Evaluation". In: *The 23rd International Technical Conference on the Enhanced Safety of Vehicles (ESV) Seoul, Republic of Korea, May 27-30, 2013*, pp. 13- 0391- W. http://publications.rwth-aachen.de/record/225495

ENDSLEY, MICA R. (1995). "Toward a Theory of Situation Awareness in Dynamic Systems". In: *Human* 37.1, pp. 32–64.

EUROPEAN COMMISSION (2017). *"Road Safety Contry Overview - Germany"*. pp. 11.

EUROPEAN COMMISSION (2022). *"Road safety statistics 2022 in more detail"*. https://transport.ec.europa.eu/background/road-safety-statistics-2022-more-detail_en

FAN, PENGCHENG; GUO, JINGQIU; ZHAO, HAIFENG; WIJNANDS, JASPER S.; WANG, YIBING (2019). *"Car-Following Modeling Incorporating Driving Memory Based on Autoencoder and Long Short-Term Memory Neural Networks". Idm*.

FRITZSCHE, HANS-THOMAS (1994). "A model for traffic simulation". In: *Traffic Engineering & Control* 35.5.

GASSER, T.; ARZT, C.; AYOUBI, M.; BARTELS, A.; BÜRKLE, L.; EIER, J.; FLEMISCH, FRANK; HÄCKER, D.; HESSE, T.; HUBER, W.; LOTZ, C.; MAURER, M.; RUTH-SCHUMACHER, S.; SCHWARZ, J.; VOGT, W. (2012). Rechtsfolgen zunehmender Fahrzeugautomatisierung. In: *Berichte der Bundesanstalt für Straßenwesen, Unterreihe "Fahrzeugsicherheit" Heft F83, Januar 2012*.

GATES, TIM J.; NOYCE, DAVID A.; LARACUENTE, LUIS; NORDHEIM, ERIK V (2007). "Analysis of Driver Behavior in Dilemma Zones at Signalized Intersections". In: *Transportation Research Record* 2030.1, pp. 29–39. DOI: 10.3141/2030-05

GAZIS, DENOS C.; HERMAN, ROBERT; ROTHERY, RICHARD W. (1961). "Nonlinear follow-the-leader models of traffic flow". In: *Operations Research* 9.4, pp. 545–567.

GIPPS, P. G. (1981). "A behavioural car-following model for computer simulation". In: *Transportation Research Part B: Methodological* 15.2, pp. 105–111. https://www.sciencedirect.com/science/article/pii/0191261581900370

GIPPS, P. G. (1986). "A model for the structure of lane-changing decisions". In: *Transportation Research Part B* 20.5, pp. 403–414. DOI: 10.1016/0191-2615(86)90012-3

GREEN, MARC (2000). ""How Long Does It Take to Stop?" Methodological Analysis of Driver Perception-Brake Times". In: *Transportation Human Factors* 2.3, pp. 195–216. DOI: 10.1207/sthf0203_1

HALKIAS, JOHN; COLYAR, JAMES (2006). *"NGSIM - Interstate 80 Freeway Dataset"*. https://www.fhwa.dot.gov/publications/research/operations/06137/06137.pdf

HAMDAR, SAMER (2012). "Driver Behavior Modeling". In: *Handbook of Intelligent Vehicles* 1–2, pp. 537–558. DOI: 10.1007/978-0-85729-085-4

HARTH, MICHAEL; ALI, MUHAMMAD SAJID; KATES, RONALD; BOGENBERGER, KLAUS (2021). "Data-Driven Modelling of Car-Following Behavior in the Approach of Signalized Urban Intersections". In: *IEEE Conference on Intelligent Transportation Systems, Proceedings*.

HARTH, MICHAEL; AMJAD, UZAIR BIN; KATES, RONALD; BOGENBERGER, KLAUS (2022). "Incorporation of Human Factors to a Data-Driven Car-Following Model". In: *Transportation Research Record* 2676.10, pp. 291–302. DOI: 10.1177/03611981221089316

HARTH, MICHAEL; LANGER, MARCEL; BOGENBERGER, KLAUS (2021). "Automated Calibration of a Microscopic Traffic Flow Simulation Using Real-World Observations". In: *Transportation Research Board*.

HARTH, MICHAEL; LANGER, MARCEL; BOGENBERGER, KLAUS (2022). "Automated Calibration of Traffic Demand and Traffic Lights in SUMO Using Real-World Observations". In: *SUMO Conference Proceedings* 2, pp. 133–148. DOI: 10.52825/scp.v2i.120

HE, KAIMING; GKIOXARI, GEORGIA; DOLLÁR, PIOTR; GIRSHICK, ROSS (2020). "Mask R-CNN". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42.2, pp. 386–397.

HINRICHS, JAMES V.; KRAINZ, PATRICIA L. (1970). "Expectancy in choice reaction time: Anticipation of stimulus or response?". In: *Journal of Experimental Psychology* 85.3, pp. 330–334. DOI: 10.1037/h0029713

HUANG, XIULING; SUN, JIE; SUN, JIAN (2018). "A car-following model considering asymmetric driving behavior based on long short-term memory neural networks". In: *Transportation Research Part C: Emerging Technologies* 95.February, pp. 346–362. https://doi.org/10.1016/j.trc.2018.07.022

JOHNSON, DERICK A.; TRIVEDI, MOHAN M. (2011). "Driving style recognition using a smartphone as a sensor platform". In: *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pp. 1609–1615. DOI: 10.1109/ITSC.2011.6083078

KAHNEMAN, DANIEL (1973). *Attention and Effort*. Prentice-Hall.

KESTING, ARNE; TREIBER, MARTIN (2009). "Calibration of Car-Following Models Using Floating Car Data". In: *Traffic and Granular Flow'07*.

KESTING, ARNE; TREIBER, MARTIN; HELBING, DIRK (2007). "General lane-changing model MOBIL for car-following models". In: *Transportation Research Record 1999*, pp. 86–94. DOI: 10.3141/1999-10

KIM, BUMSIK; HEASLIP, KEVIN P. (2023). "Identifying suitable car-following models to simulate automated vehicles on highways". In: *International Journal of Transportation Science and Technology* 12.2, pp. 652–664. DOI: 10.1016/j.ijtst.2023.02.003

KRAJZEWICZ, DANIEL; RICHTER, ANDREAS; BEHRISCH, MICHAEL; ERDMANN, JAKOB (2013). *"Abbildung des Umgebungsverkehrs in einem Fahrsimulator"*. pp. 1–10.

KRAUß, STEFAN (1998). *Microscopic modeling of traffic flow: Investigation of collision free vehicle dynamics*. University of Cologne.

LEE, D. N. (1976). "A theory of visual control of braking based on information about time to collision". In: *Perception* 5.4, pp. 437–459. DOI: 10.1068/p050437

LINDORFER, MANUEL; MECKLENBRAUKER, CHRISTOPH F.; OSTERMAYER, GERALD (2018). "Modeling the Imperfect Driver: Incorporating Human Factors in a Microscopic Traffic Model". In: *IEEE Transactions on Intelligent Transportation Systems* 19.9, pp. 2856–2870.

LITMAN, TODD (2022). *Autonomous Vehicle Implementation Predictions Implications for Transport Planning*.

LOBO, SILAS C.; NEUMEIER, STEFAN; FERNANDEZ, EVELIO M. G.; FACCHI, CHRISTIAN (2020). "InTAS -- The Ingolstadt Traffic Scenario for SUMO". In: *SUMO User Conference 2020*, pp. 1–20. http://arxiv.org/abs/2011.11995

LOOMIS, JACK M.; DA SILVA, JOSÉ A.; FUJITA, NAOFUMI; FUKUSIMA, SERGIO S. (1992). "Visual Space Perception and Visually Directed Action". In: *Journal of Experimental Psychology: Human Perception and Performance* 18.4, pp. 906–921. DOI: 10.1037/0096-1523.18.4.906

LOPEZ, PABLO ALVAREZ; BEHRISCH, MICHAEL; BIEKER-WALZ, LAURA; ERDMANN, JAKOB; FLÖTTERÖD, YUN PANG; HILBRICH, ROBERT; LUCKEN, LEONHARD; RUMMEL, JOHANNES; WAGNER, PETER; WIEBNER, EVAMARIE (2018). "Microscopic Traffic Simulation using SUMO". In: *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pp. 2575–2582. DOI: 10.1109/ITSC.2018.8569938

LU, QIONG; TETTAMANTI, TAMÁS; HÖRCHER, DÁNIEL; VARGA, ISTVÁN (2019). "The impact of autonomous vehicles on urban traffic network capacity: an experimental analysis by microscopic traffic simulation". In: *Transportation Letters* 00.00, pp. 1–10. DOI: 10.1080/19427867.2019.1662561

MA, LIJING; QU, SHIRU (2020). "A sequence to sequence learning based car-following model for multi-step predictions considering reaction delay". In: *Transportation Research Part C: Emerging Technologies* 120. https://doi.org/10.1016/j.trc.2020.102785

MAI, MARCUS; TÜSCHEN, THOMAS; PROKOP, GÜNTHER (2015). "A physiological based Driver Model for longitudinal Vehicle Guidance and its Challenges in Validation". In: *Proceedings of the 3rd International Symposium on Future Active Safety Technology Towards Zero Traffic Accidents*, pp. 175–182.

MOONEY, JOHN G.; JOHNSON, ERIC N. (2014). "A Comparison of Automatic Nap-of-the-earth Guidance Strategies for Helicopters". In: *Journal of Field Robotics* 25.February, pp. 1–17. DOI: 10.1002/rob

MORTON, JEREMY; WHEELER, TIM A.; KOCHENDERFER, MYKEL J. (2017). "Analysis of Recurrent Neural Networks for Probabilistic Modeling of Driver Behavior". In: *IEEE Transactions on Intelligent Transportation Systems* 18.5, pp. 1289–1298. DOI: 10.1109/TITS.2016.2603007

NASH, CHRISTOPHER J.; COLE, DAVID J.; BIGLER, ROBERT S. (2016). "A review of human sensory dynamics for application to models of driver steering and speed control". In: *Biological Cybernetics* 110.2–3, pp. 91–116. DOI: 10.1007/s00422-016-0682-x

NUMFOCUS (2024). *"Geopandas"*. https://geopandas.org/

OLSTAM, JOHAN; ELYASI-POUR, ROYA (2013). "Combining traffic and vehicle simulation for enhanced evaluations of powertrain related ADAS for trucks". In: *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pp. 851–856. DOI: 10.1109/ITSC.2013.6728338

OLSTAM, JOHAN JANSON; TAPANI, ANDREAS (2004). Comparison of car-following models for simulation. In: *Transportation Research Record* (Vol. 960). Swedish National Road and Transport Research Institute. DOI: 10.3141/1678-15

OPENAI (2023). *"ChatGPT"*. https://openai.com/blog/chatgpt

OSSEN, SASKIA; HOOGENDOORN, SERGE P.; GORTE, BEN G. H. (2006). "Interdriver differences in car-following a vehicle trajectory-based study". In: *Transportation Research Record 1965*, pp. 121–129. DOI: 10.3141/1965-13

PALINKO, OSKAR; KUN, ANDREW L.; SHYROKOV, ALEXANDER; HEEMAN, PETER (2010). "Estimating cognitive load using remote eye tracking in a driving simulator". In: *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, pp. 141–144. DOI: 10.1145/1743666.1743701

PIPES, LOUIS A. (1953). "An operational analysis of traffic dynamics". In: *Journal of Applied Physics* 24.3, pp. 274–281.

PORFYRI, KALLIRROI N.; MINTSIS, EVANGELOS; MITSAKIS, EVANGELOS (2018). "Assessment of ACC and CACC systems using SUMO". In: *EPiC Series in Engineering* 2, pp. 82–93. DOI: 10.29007/r343

PRECHT, LISA; KEINATH, ANDREAS; KREMS, JOSEF F. (2017). "Identifying the main factors contributing to driving errors and traffic violations – Results from naturalistic driving data". In: *Transportation Research Part F: Traffic Psychology and Behaviour* 49, pp. 49–92. DOI: 10.1016/j.trf.2017.06.002

RASMUSSEN, JENS (1983). "Skills, Rules, and Knowledge; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models". In: *IEEE Transactions on Systems, Man, and Cybernetics* 13.3, pp. 257–266.

REUSCHEL, AV (1950). "Fahrzeugbewegungen in der Kolonne". In: *Österreichisches Ingenieur Archiv* 4, pp. 193–215.

ROBERT KRAJEWSKI, JULIAN BOCK, LAURENT KLOEKER AND LUTZ ECKSTEIN (2018). *"The highD Dataset". November*, pp. 2–9. http://www.highd-dataset.com

ROCKWELL, T. H. (1972). "Eye Movement Analysis of Visual Information Acquisition in Driving: an Overview". In: *Australian Road Research Board The Ohio State University, College of*

Engineering, Columbus 6, pp. 316–331. papers3://publication/uuid/EB737A11-A97D-427A-B7FC-64BFBF0B46ED

RÖSENER, CHRISTIAN (2020). *A Traffic-based Method for Safety Impact Assessment of Road Vehicle Automation*. RWTH Aachen.

ROUSSEEUW, PETER J. (1987). "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis". In: *Journal of Computational and Applied Mathematics* 20, pp. 53–65. DOI: https://doi.org/10.1016/0377-0427(87)90125-7

RUNESON, SVERKER (1974). "Constant velocity — Not perceived as such". In: *Psychological Research* 37.1, pp. 3–23. DOI: 10.1007/BF00309076

SAIFUZZAMAN, MOHAMMAD; ZHENG, ZUDUO (2014). "Incorporating human-factors in car-following models: A review of recent developments and research needs". In: *Transportation Research Part C: Emerging Technologies* 48, pp. 379–403.

SCHAKEL, WOUTER J.; KNOOP, VICTOR L.; VAN AREM, BART (2012). "Integrated Lane Change Model with Relaxation and Synchronization". In: *Transportation Research Record* 2316.1, pp. 47–57. DOI: 10.3141/2316-06

SCHLAG, BERNHARD; HEGER, RALF (2002). *Empfehlungen zur Berücksichtigung physiologischer und psychologischer Fähigkeiten und Grenzen der Kraftfahrer bei der Straßenplanung in Brandenburg*.

SCHMIDT, ROBERT F.; LANG, FLORIAN (2013). Physiologie des Menschen. In: *Springer-Verlag*. DOI: 10.1038/078601b0

SCHULDT, FABIAN (2017). *"Ein Beitrag für den methodischen Test von automatisierten Fahrfunktionen mit Hilfe von virtuellen Umgebungen"*. DOI: 10.1017/CBO9781107415324.004

SCHWAB, BENEDIKT; BEIL, CHRISTOF; KOLBE, THOMAS H. (2020). "Spatio-Semantic Road Space Modeling for Vehicle–Pedestrian Simulation to Test Automated Driving Systems". In: *Sustainability* 5.9. DOI: 10.3390/su12093799

SEMRAU, MARC (2018). *Untersuchung zur Modellierung von chinesischem Fahrverhalten auf Autobahnen für den Test pilotierter Fahrfunktionen* [TU Braunschweig]. DOI: 10.1007/978-3-658-23761-5

SEMRAU, MARC; ERDMANN, JAKOB; FRIEDRICH, BERNHARD; WALDMANN, RENE (2016). "Simulation framework for testing ADAS in Chinese traffic situations". In: *SUMO 2016 Conference Proceedings*, pp. 103–113.

SIDDIQUI, FAIZ; MERRILL, JEREMY B. (2023). *"17 fatalities, 736 crashes: The shocking toll of Tesla's Autopilot"*. The Washington Post. https://www.washingtonpost.com/technology/2023/06/10/tesla-autopilot-crashes-elon-musk/

STADT INGOLSTADT (2018). *"Hohe PKW-Dichte in Ingolstadt"*. https://www.ingolstadt.de/Home/Hohe-Pkw-Dichte-in-Ingolstadt.php?object=tx,2789.5&ModID=7&FID=3052.11361.1&NavID=2789.411

STATISTA RESEARCH DEPARTMENT (2023a). *"Anzahl zugelassener Pkw in Deutschland von 1960 bis 2023"*. https://de.statista.com/statistik/daten/studie/12131/umfrage/pkw-bestand-in-deutschland/

STATISTA RESEARCH DEPARTMENT (2023b). *"Volkswirtschaftliche Kosten der Straßenverkehrsunfälle in Deutschland nach Schadensart von 2010 bis 2021"*. https://de.statista.com/statistik/daten/studie/1203615/umfrage/kosten-der-strassenverkehrsunfaelle-in-deutschland-nach-schaden/

STATISTISCHE ÄMTER DES BUNDES UND DER LÄNDER (2023). *"Unfallatlas"*. https://unfallatlas.statistikportal.de/_opendata2020.html

STATISTISCHES BUNDESAMT DESTATIS (2023a). *"Verkehrsunfälle nach Unfallkategorie und Ortslage im Zeitvergleich"*. https://www-genesis.destatis.de/genesis/online?sequenz=tabelleErgebnis&selectionname=46241-0001

STATISTISCHES BUNDESAMT DESTATIS (2023b). *"Verkehrsunfälle und Verunglückte im*

Zeitvergleich (ab 1950)". https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Verkehrsunfaelle/Tabellen/liste-strassenverkehrsunfaelle.html

TOLEDO, TOMER; KOUTSOPOULOS, HARIS N.; BEN-AKIVA, MOSHE E. (2003). "Modeling Integrated Lane-Changing Behavior". In: *Transportation Research Record 1857*, pp. 30–38. DOI: 10.3141/1857-04

TREIBER, MARTIN; HENNECKE, ANSGAR; HELBING, DIRK (2000). "Congested traffic states in empirical observations and microscopic simulations". In: *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* 62.2, pp. 1805–1824.

TREIBER, MARTIN; KESTING, ARNE; HELBING, DIRK (2006). "Delays, inaccuracies and anticipation in microscopic traffic models". In: *Physica A: Statistical Mechanics and Its Applications* 360.1, pp. 71–88. DOI: 10.1016/j.physa.2005.05.001

VAN BUUREN, STEF; GROOTHUIS-OUDSHOORN, KARIN (2011). "mice: Multivariate imputation by chained equations in R". In: *Journal of Statistical Software* 45.3, pp. 1–67.

VAN LINT, J. W. C.; CALVERT, S. C. (2018). "A generic multi-level framework for microscopic traffic simulation—Theory and an example case in modelling driver distraction". In: *Transportation Research Part B: Methodological* 117, pp. 63–86. DOI: 10.1016/j.trb.2018.08.009

WACHENFELD, WALTHER; WINNER, HERMANN (2015). Die Freigabe des autonomen Fahrens. Technische, rechtliche und gesellschaftliche Aspekte. In: M. Maurer, J. C. Gerdes, B. Lenz, & H. Winner (Eds.), *Autonomes Fahren* (pp. 439–464). Springer Vieweg. DOI: 10.1007/978-3-662-45854-9_21

WAGNER, PETER (2015). Steuerung und Management in einem Verkehrssystem mit autonomen Fahrzeugen. In: M. Maurer, J. C. Gerdes, B. Lenz, & H. Winner (Eds.), *Autonomes Fahren* (pp. 313–330). Springer Vieweg.

WAKABAYASHI, DAISUKE (2018). *"Self-Driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam"*. The New York Times. https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html

WANG, XIAO; JIANG, RUI; LI, LI; LIN, YILUN; ZHENG, XINHU; WANG, FEI YUE (2018). "Capturing Car-Following Behaviors by Deep Learning". In: *IEEE Transactions on Intelligent Transportation Systems* 19.3, pp. 910–920. DOI: 10.1109/TITS.2017.2706963

WARDROP, JOHN GLEN (1952). "Some theoretical Aspects of road traffic research". In: *Proceedings of the Institute of Civil Engineers, Part II* 1, pp. 325–378. Wardrop.1952.pdf

WATAMANIUK, SCOTT N. J.; HEINEN, STEPHEN J. (2003). Perceptual and oculomotor evidence of limitations on processing accelerating motion. In: *Journal of Vision* (Vol. 3, Issue 11, pp. 698–709). Assn for Research in Vision & Ophthalmology (ARVO). DOI: 10.1167/3.11.5

WEBER, E. H. (1834). "De pulsu, resortione, auditu et tactu". In: *Annotationes Anatomicae et Physiologicae* 1.

WHEELER, TIM A.; ROBBEL, PHILIPP; KOCHENDERFER, MYKEL J. (2016). "Analysis of Microscopic Behavior Models for Probabilistic Modeling of Driver Behavior". In: *Intelligent Transportation Systems Conference*, pp. 1604–1609. DOI: 10.1109/TITS.2016.2603007

WICKENS, CHRISTOPHER D.; HOLLANDS, JUSTIN G.; BANBURY, SIMON; PARASURAMAN, RAJA (2015). *Engineering psychology and human performance*. Psychology Press.

WIEDEMANN, R. (1974). "Simulation des Straßenverkehrsflusses". In: *Schriftenreihe Des Instituts Für Verkehrswesen Der Universität Karlsruhe* 8.

YILMAZ, EMRE H.; WARREN, WILLIAM H. (1995). "Visual Control of Braking: A Test of the τ Hypothesis". In: *Journal of Experimental Psychology: Human Perception and Performance* 21.5, pp. 996–1014. DOI: 10.1037/0096-1523.21.5.996

YOUNG, KRISTIE; REGAN, MICHAEL (2007). "Driver distraction : A review of the literature". In: *Distracted Driving*, pp. 379–405.

ZHANG, JIAN; TANG, TIE QIAO; WANG, TAO (2019). "Some features of car-following behaviour in the vicinity of signalised intersection and how to model them". In: *IET Intelligent Transport Systems* 13.11, pp. 1686–1693. DOI: 10.1049/iet-its.2018.5510

ZHOU, MOFAN; QU, XIAOBO; LI, XIAOPENG (2017). "A recurrent neural network based microscopic car following model to predict traffic oscillation". In: *Transportation Research Part C: Emerging Technologies* 84, pp. 245–264. DOI: 10.1016/j.trc.2017.08.027

Appendix

A Driver Type Detection Quality

Figure A.1 to Figure A.4 show the driver type detection quality parameters for a time interval between 7 and 8 a.m. Figure A.5 to Figure A.8 show the same parameters for an interval between 9 and 10 a.m.

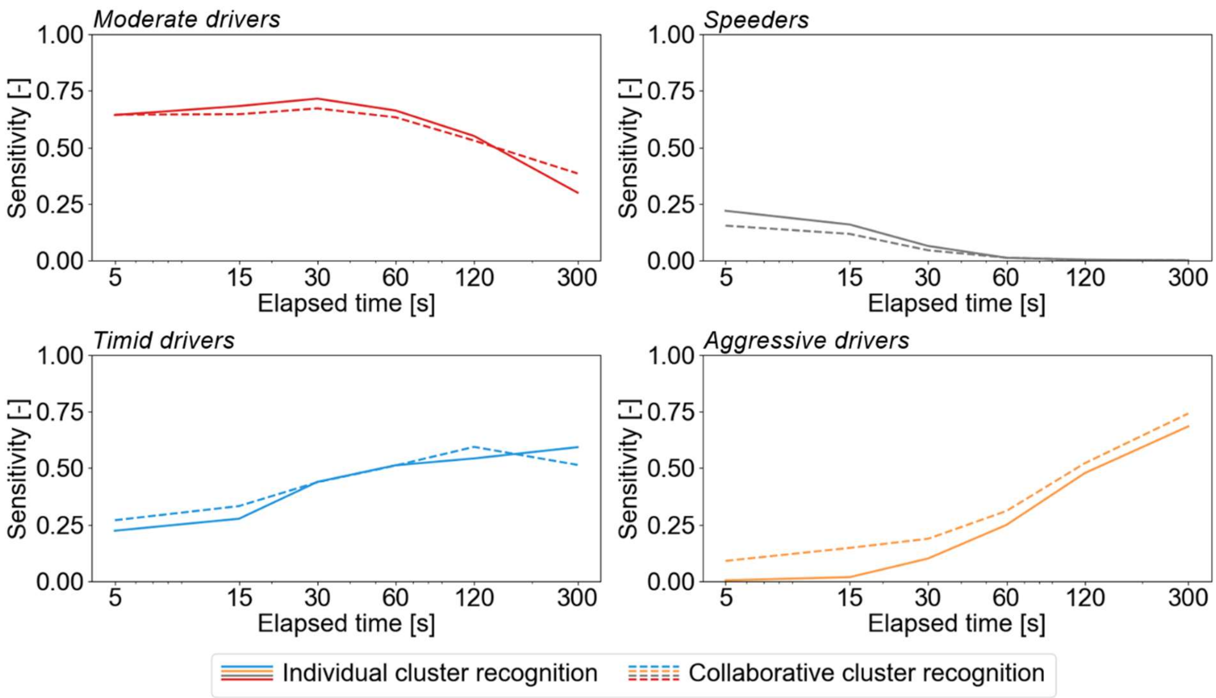

Figure A.1: Sensitivity of the investigated driver type recognition procedures between 7 and 8 a.m. for all four driver types as a function of observation time.

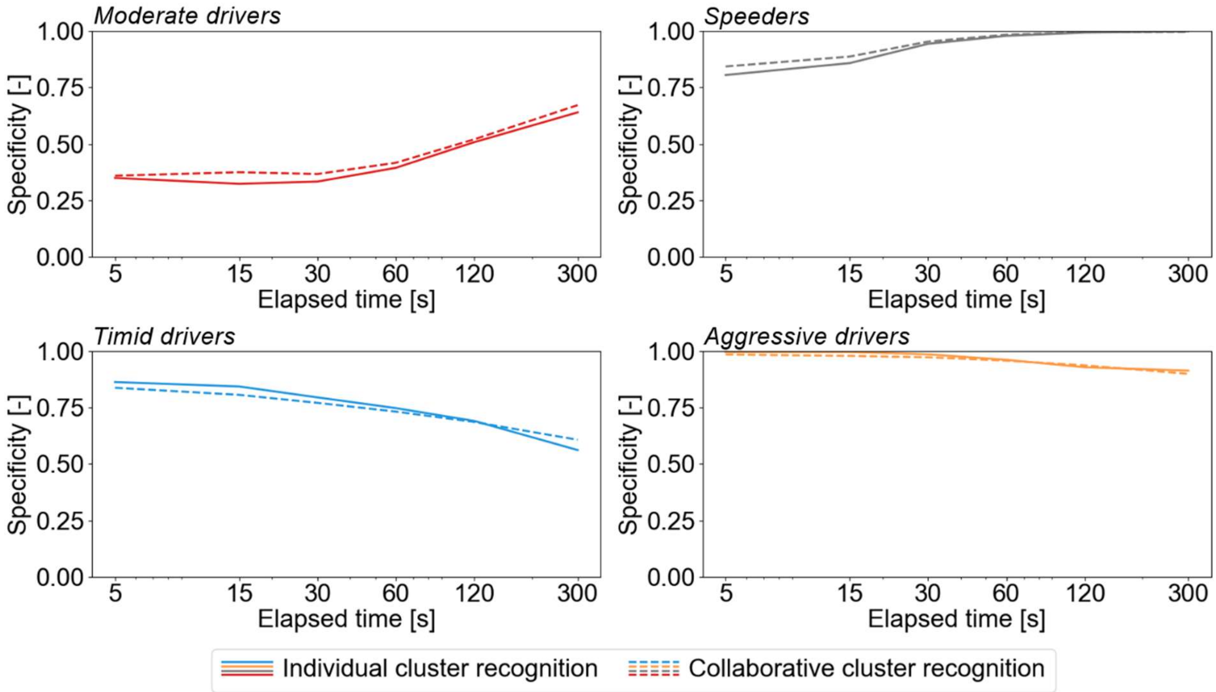

Figure A.2: Specificity of the investigated driver type recognition procedures between 7 and 8 a.m. for all four driver types as a function of observation time.

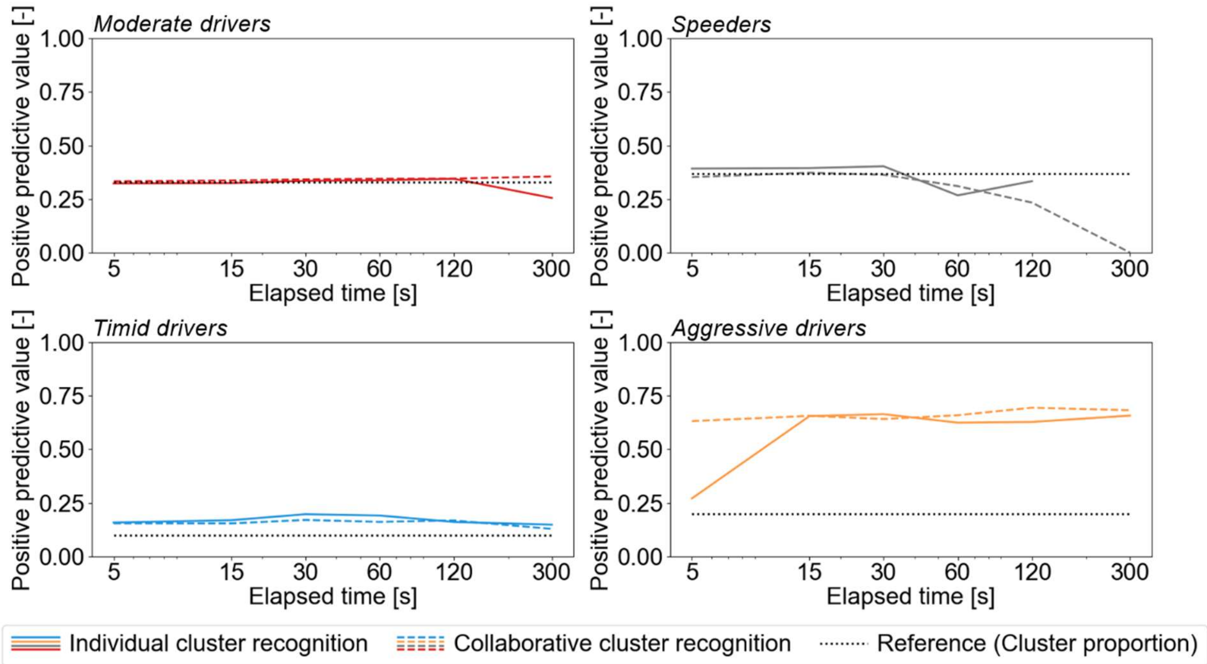

Figure A.3: PPV of the investigated driver type recognition procedures between 7 and 8 a.m. for all four driver types as a function of observation time.

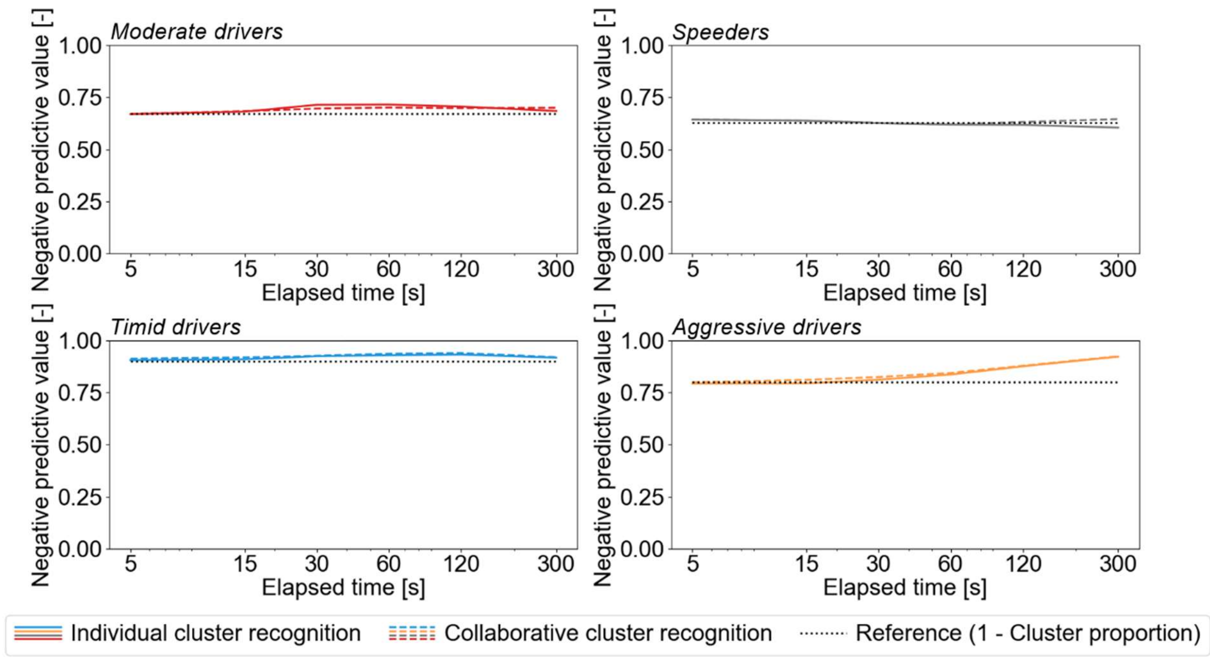

Figure A.4: NPV of the investigated driver type recognition procedures between 7 and 8 a.m. for all four driver types as a function of observation time.

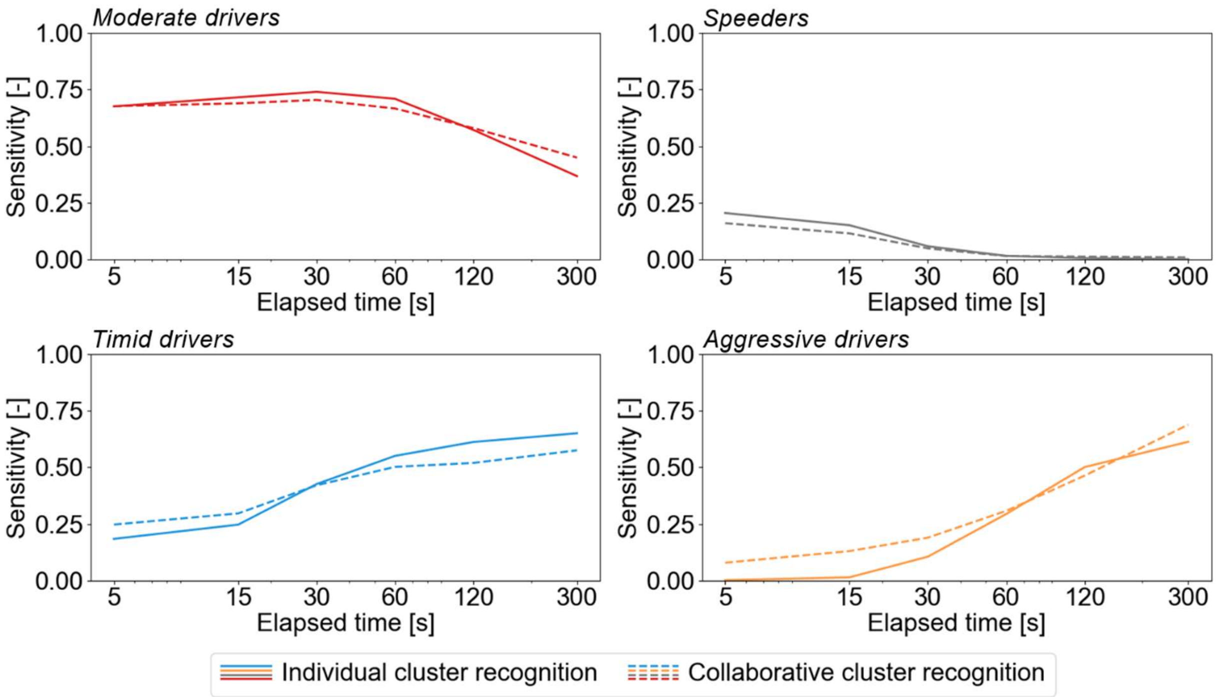

Figure A.5: Sensitivity of the investigated driver type recognition procedures between 9 and 10 a.m. for all four driver types as a function of observation time.

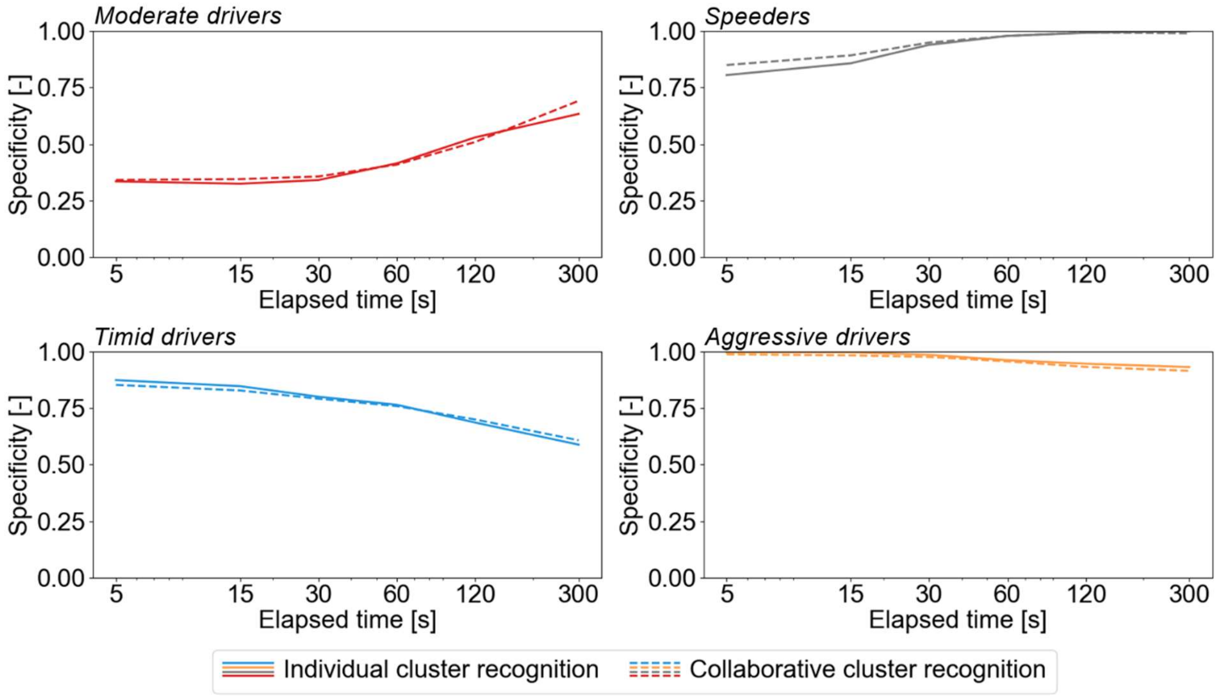

Figure A.6: Specificity of the investigated driver type recognition procedures between 9 and 10 a.m. for all four driver types as a function of observation time.

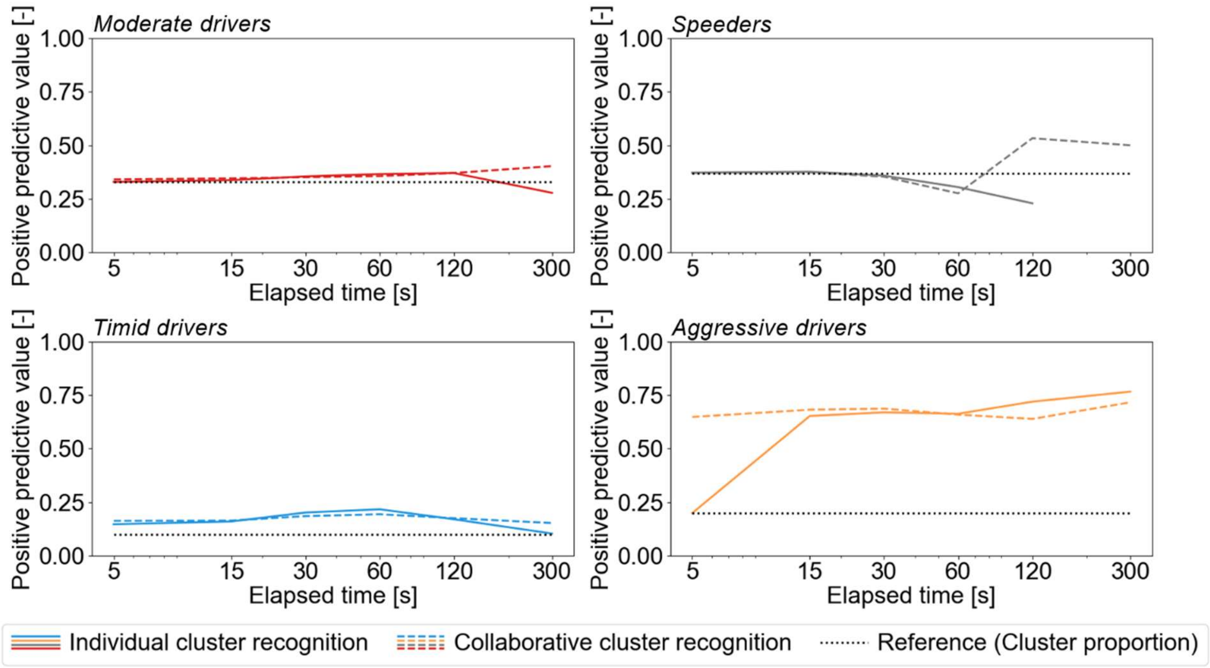

Figure A.7: PPV of the investigated driver type recognition procedures between 9 and 10 a.m. for all four driver types as a function of observation time.

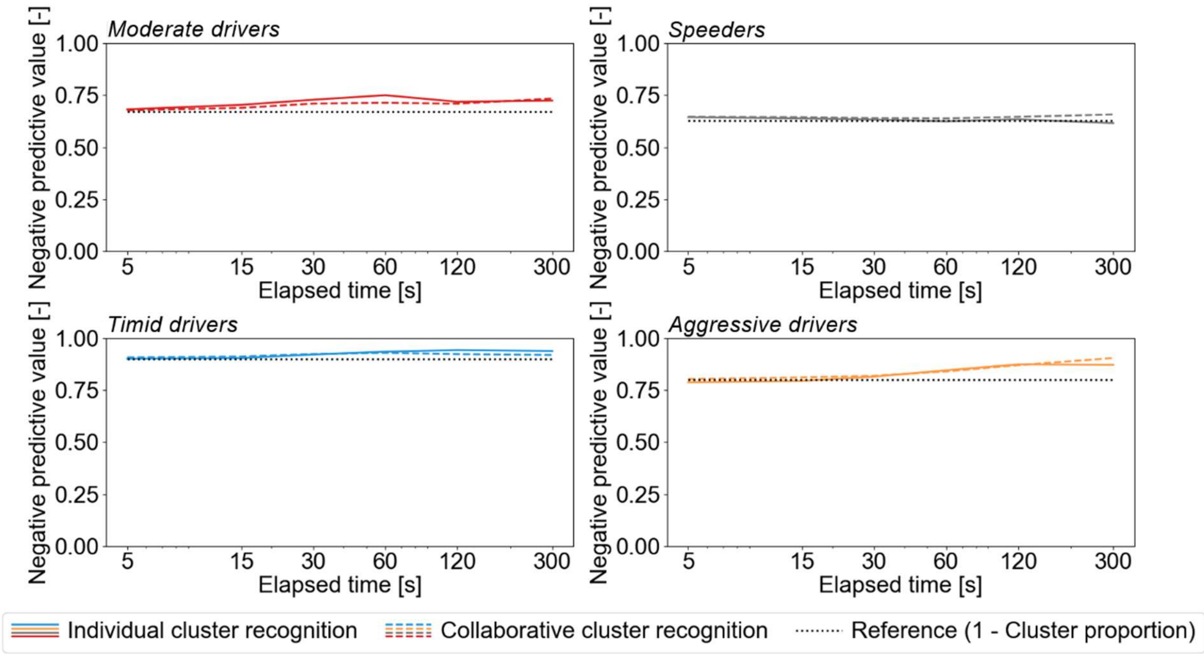

Figure A.8: NPV of the investigated driver type recognition procedures between 9 and 10 a.m. for all four driver types as a function of observation time.

B Average Accident Numbers and Costs in Ingolstadt

Tables B.1 and B.2 show the average personal injury and property damage costs for accidents in Ingolstadt per year.

Accident category	Average number of accidents in longitudinal traffic in Ingolstadt per year	Average costs per accident in Germany (2021) [€]	Total average costs per year [€]
Accidents with personal injury with fatalities	0	1 284 160	0
Accidents with personal injury with severe injuries	2.2	128 878	283 532
Accidents with personal injury with minor injuries	87.5	5 652	494 550

Table B.1: Average personal injury costs for accidents in Ingolstadt per year.

Accident category	Average number of accidents in longitudinal traffic in Ingolstadt per year	Average costs per accident in Germany in 2021 [€]	Total average costs per year [€]
Accidents with property damage with fatalities	0	46 546	0
Accidents with property damage with severe injuries	2.2	22 649	49 828
Accidents with property damage with minor injuries	87.5	14 204	1 242 850
Serious accidents with only property damage	16.2	21 799	353 143
Accidents with only property damage	675.5	7 474	5 048 687

Table B.2: Average property damage costs for accidents in Ingolstadt per year.