139

Karlsruher Schriftenreihe
Fahrzeugsystemtechnik

Steffen Metzger

Development of an assistant system for regulating the mental strain state of agricultural machinery operators

KIT Scientific Publishing

Steffen Metzger

Development of an assistant system for regulating the mental strain state of agricultural machinery operators

Karlsruher Schriftenreihe Fahrzeugsystemtechnik
Band 139

Herausgeber
FAST Institut für Fahrzeugsystemtechnik
Prof. Dr.-Ing. Martin Cichon
Prof. Dr. rer. nat. Frank Gauterin
Prof. Dr.-Ing. Marcus Geimer
Prof. Dr.-Ing. Frank Henning
Prof. Dr.-Ing. Luise Kärger

Das Institut für Fahrzeugsystemtechnik besteht aus den Institutsteilen Bahnsystemtechnik, Fahrzeugtechnik, Leichtbau und Mobile Arbeitsmaschinen.

Eine Übersicht aller bisher in dieser Schriftenreihe erschienenen Bände finden Sie am Ende des Buchs.

Development of an assistant system for regulating the mental strain state of agricultural machinery operators

by
Steffen Metzger

Scientific Publishing

Karlsruher Institut für Technologie
Institut für Fahrzeugsystemtechnik

Development of an assistant system for regulating the mental strain state of agricultural machinery operators

Zur Erlangung des akademischen Grades eines Doktors der Ingenieurwissenschaften (Dr.-Ing.) von der KIT-Fakultät für Maschinenbau des Karlsruher Instituts für Technologie (KIT) genehmigte Dissertation

von Steffen Metzger, M.Sc.

Tag der mündlichen Prüfung: 17. März 2026
Erster Gutachter: Prof. Dr.-Ing. Marcus Geimer
Zweite Gutachterin: Prof. Dr.-Ing. Barbara Deml

Impressum

KIT Scientific Publishing

Karlsruher Institut für Technologie (KIT)
KIT Scientific Publishing
Straße am Forum 2
D-76131 Karlsruhe

KIT Scientific Publishing is a registered trademark of Karlsruhe Institute of Technology.
Reprint using the book cover is not allowed.

www.bibliothek.kit.edu/ksp.php | E-Mail: info@ksp.kit.edu | Shop: www.ksp.kit.edu

Print on Demand 2026 – Gedruckt auf FSC-zertifiziertem Papier

ISSN 1869-6058
ISBN 978-3-7315-1482-4
DOI 10.5445/KSP/1000192087

Vorwort des Herausgebers

Die Megatrends Bevölkerungswachstum und Digitalisierung beeinflussen unmittelbar die Landwirtschaft. Immer mehr Menschen leben auf der Erde, die mit Nahrungsmitteln versorgt werden müssen. Gleichzeitig nimmt die bewirtschaftete Fläche ab. Diesem sich verstärkenden Trend kann nur mit einer zunehmenden Produktivitätssteigerung entgegengewirkt werden. Dies hat zum Ersten eine zunehmende Mechanisierung zur Folge und zum Zweiten müssen Prozesse automatisiert werden, damit der Landwirt nicht überfordert wird. Die Digitalisierung fördert hierbei die Automatisierung der Landmaschinen, da Sensorik, Schnittstellen und Steuerungssysteme kostengünstig zur Verfügung stehen.

Die Karlsruher Schriftenreihe Fahrzeugsystemtechnik widmet sich Themen der Entwicklung, Automatisierung und Effizienzsteigerung von Fahrzeugen. Für die Fahrzeuggattungen Pkw, Nfz, Mobile Arbeitsmaschinen und Bahnfahrzeuge werden in der Schriftenreihe Forschungsarbeiten vorgestellt, die Fahrzeugtechnik auf vier Ebenen beleuchten: das Fahrzeug als komplexes mechatronisches System, die Fahrer-Fahrzeug-Interaktion, das Fahrzeug im Verkehr und Infrastruktur sowie das Fahrzeug in Gesellschaft und Umwelt.

Ein moderner Mähdrescher kann heute 60 bis 100 Tonnen Getreide pro Stunde ernten. Eine manuelle Steuerung dieses Prozesses ist mit einer hohen mentalen Belastung verbunden, so dass moderne Mähdrescher hoch automatisiert sind. Der Bediener wird dadurch einerseits entlastet, er erhält aber zahlreiche Informationen, die er für eine effiziente Ernte verarbeiten muss. Ist die Maschine einmal auf die Erntebedingungen eingestellt, so obliegt dem Bediener „nur“ noch die Überwachung der Maschine, was zu einer Monotonie und Langeweile führen kann. Die Dissertation von Herrn Metzger hat das Ziel, die mentale Beanspruchung des Bedieners zu messen und dann auf diese Beanspruchung so zu reagieren, dass der Bediener in einem angenehmen Beanspruchungszustand gehalten werden kann. Hierzu schlägt er eine Methode auf Basis eines Machine-Learning-Modells vor, die auf Basis des aktuellen

Beanspruchungszustands, der aktuellen und erwarteten Umweltbedingungen und der Vorlieben des Benutzers Handlungsvorschläge macht.

Karlsruhe, im März 2026 *Prof. Dr.-Ing. Marcus Geimer*

Abstract

The increasing automation of agricultural machinery leads to fluctuating phases of mental underload and overload for operators, which can impair performance, safety, and well-being. This dissertation describes the development and evaluation of an adaptive assistance system that monitors, evaluates, and regulates the mental strain of machine operators in real time.

As an application example, a highly automated combine harvester was used, simulated in an immersive demonstrator cabin for experimental studies. Multimodal sensor data, including physiological metrics such as eye-tracking parameters and cardiovascular signals, were utilized to train a machine learning model capable of classifying the mental state. This model was integrated into a modular virtual assistant, which provides context-dependent task recommendations via voice and touch interfaces in order to prevent critical mental strain states such as underload or overload.

Experimental studies in the demonstrator cabin demonstrated that the system can detect mental strain and regulate it through targeted task recommendations. A subsequent field study on a real combine harvester confirmed practical applicability and identified optimization potential regarding system integration and user acceptance.

The results indicate that adaptive mental strain regulation can specifically reduce critical stress conditions, thereby enhancing the safety, efficiency, and satisfaction of operators. In the future, the developed concept can be transferred to other agricultural machinery as well as other highly automated work environments, thereby creating new approaches for adaptive, human-centered assistance systems.

Keywords: Adaptive assistance systems, Mental strain, Virtual assistant, Ergonomics, Human-machine interface, Combine harvester

Kurzfassung

Die zunehmende Automatisierung landwirtschaftlicher Maschinen führt bei Bedienenden zu wechselnden Phasen mentaler Unter- und Überforderung, die Leistung, Sicherheit und Wohlbefinden beeinträchtigen können. Die vorliegende Dissertation beschreibt die Entwicklung und Evaluation eines adaptiven Assistenzsystems, das den mentalen Beanspruchungszustand von Maschinenführern in Echtzeit erfasst, bewertet und reguliert.

Als Anwendungsbeispiel diente ein hochautomatisierter Mähdrescher, der in einer immersiven Demonstrator-Kabine für experimentelle Studien simuliert wurde. Multimodale Sensordaten, darunter physiologische Messgrößen wie Eye-Tracking-Parameter und kardiovaskuläre Signale, wurden genutzt, um ein Machine-Learning-Modell zu trainieren, das den mentalen Zustand klassifizieren kann. Dieses Modell wurde in einen modularen, virtuellen Assistenten integriert, der situationsabhängige Handlungsempfehlungen über Sprach- und Touchschnittstellen bereitstellt, um Unter- oder Überforderung gezielt zu vermeiden.

Experimentelle Studien in der Demonstrator-Kabine zeigten, dass das System mentale Beanspruchung erkennen und durch gezielte Handlungsempfehlungen beeinflussen kann. Eine anschließende Feldstudie an einem realen Mähdrescher bestätigte die praktische Anwendbarkeit und identifizierte Optimierungspotenziale hinsichtlich Systemintegration und Nutzerakzeptanz.

Die Ergebnisse verdeutlichen, dass die adaptive Beanspruchungsregulation kritische Beanspruchungszustände gezielt reduzieren und damit Sicherheit, Effizienz und Zufriedenheit der Bedienenden steigern kann. Perspektivisch kann das entwickelte Konzept auf andere landwirtschaftliche Maschinen sowie auf andere hochautomatisierte Arbeitsumgebungen übertragen werden, wodurch neue Ansätze für adaptive, menschenzentrierte Assistenzsysteme entstehen.

Stichworte: Adaptive Assistenzsysteme, Kognitive Beanspruchung, Virtueller Assistent, Ergonomie, Mensch-Maschine-Schnittstelle, Mähdrescher

Danksagung

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter und Promotionsstudent am Institutsteil Mobile Arbeitsmaschinen (Mobima) des Karlsruher Instituts für Technologie (KIT).

Mein besonderer Dank gilt meinem Doktorvater Prof. Dr.-Ing. Marcus Geimer für die kompetente wissenschaftliche Betreuung sowie für die Übernahme des Hauptreferats. Der mir im Rahmen dieser Arbeit gewährte Freiraum am Institut sowie das mir entgegengebrachte Vertrauen bei der Projektbearbeitung sind keineswegs selbstverständlich. Sowohl fachlich als auch persönlich konnte ich in diesen sechs Jahren außerordentlich viel lernen.

Prof. Dr.-Ing. Barbara Deml, Leiterin des Instituts für Arbeitswissenschaft und Betriebsorganisation (ifab) am KIT, danke ich herzlich für ihr Interesse an der Arbeit sowie für die Übernahme des Korreferats. Ebenso danke ich Prof. Dr.-Ing. Zäzilia Seibold, Professorin des Instituts für Fördertechnik und Logistiksysteme (IFL) am KIT, für die Übernahme des Prüfungsvorsitzes.

Meinen Kolleginnen und Kollegen am Institut spreche ich ebenfalls meinen aufrichtigen Dank aus. Vielen Dank für die fachliche wie auch persönliche Unterstützung in zahlreichen Situationen, in denen ihr mir stets freundschaftlich mit Rat und Tat zur Seite gestanden habt. Namentlich möchte ich mich ganz besonders bei meinem Bürokollegen Patrick Lehr für die gemeinsame Zeit im Büro 028 sowie für die ausgezeichnete Zusammenarbeit während der gesamten Projektlaufzeit bedanken. Darüber hinaus gilt mein Dank allen Projektpartnerinnen und Projektpartnern des Projekts „Fahrerkabine 4.0“, die maßgeblich zu einem außerordentlich erfolgreichen Projektabschluss beigetragen haben.

Meinen studentischen Hilfskräften sowie den Abschlussarbeiterinnen und Abschlussarbeitern, die an meinem Thema mitgeforscht haben, danke ich herzlich für ihre engagierte Unterstützung. Ohne euch wäre die Durchführung vieler Arbeiten nicht möglich gewesen.

Mein ganz besonderer Dank gilt meiner Familie, die mich während meiner gesamten Promotionszeit uneingeschränkt unterstützt hat. Für den damit verbundenen bedingungslosen Rückhalt bin ich sehr dankbar. Hervorheben möchte ich hierbei insbesondere meine Partnerin Nadine. Du hast mich nicht nur während der Schreibphase mit deiner fürsorglichen, liebevollen und geduldigen Art unterstützt, sondern standest mir auch fachlich jederzeit mit Rat und Tat zur Seite, selbst dann, wenn deine eigene Dissertation dabei zurückstehen musste. Diese Arbeit wäre ohne dich nicht möglich gewesen.

Ein weiterer großer Dank gilt meinen Eltern Birgit und Oliver, die meinen Lebensweg stets unterstützt, mir mein Studium ermöglicht haben und mir während der Promotion wie auch in allen anderen Lebenssituationen jederzeit hilfreich zur Seite standen.

Zu guter Letzt danke ich all jenen, die bei allen wissenschaftlichen und technologischen Innovationen den Menschen stets im Blick behalten.

Prima di essere ingegneri voi siete uomini.
Bevor ihr Ingenieure seid, seid ihr vor allem Menschen.

Francesco de Sanctis
Eidgenössisches Polytechnikum in Zürich, 1856

Karlsruhe, im März 2026 *Steffen Metzger*

Contents

List of Figures

List of Tables

List of Abbreviations

1. Introduction

Demographic change and the increasing demands for reconciling professional and private life present agricultural businesses with significant structural and personnel challenges. In recent years, there has been a decline in both agricultural specialists and young farm owners. Between 2013 and 2023, the proportion of farm managers over the age of 55 increased by more than 15%, while the total number of agricultural workers declined by approximately 14% during the same period [25]. Concurrently, the number of farms has decreased considerably, while the remaining farms have increased in size. This has resulted in a concentration of work and an increase in mental stress [26], [27]. These developments illustrate that maintaining agricultural performance increasingly depends on the attractiveness of the working environment.

Simultaneously, advancing automation is fundamentally changing working conditions in agriculture. Modern harvesting machines, such as combine harvesters, already feature highly automated control and assistance systems that make it possible to maintain the work process largely without operator intervention over extended periods [28]. As a result, the demands placed on machine operators are shifting [29]. Phases of intense mental strain alternate with long periods of low mental strain. This dynamic brings new ergonomic and psychological challenges. Overload can lead to impaired performance and reduced situational awareness. Conversely, underload, monotony and boredom can result in fatigue and reduced attention. Both extremes negatively affect well-being, safety, and performance [30]–[32].

The design of work systems that promote a balanced level of mental strain is therefore gaining importance. Studies show that the optimal performance state of a person lies within a moderate range of mental strain, where cognitive demands and individual resources are in balance [30]. A central goal of ergonomic work design is to avoid both overload and underload, thereby stabilizing the level of mental strain throughout the workday.

With the increasing integration of automated and digital functions in agricultural work processes, new possibilities are emerging to detect, assess, and specifically influence mental strain. Sensor-based methods, such as eye tracking or other physiological measurements, allow for real-time detection of an operator's Mental strain state (MSS) [33]. On this basis, adaptive assistance systems can be developed that respond to overload or underload situations. In phases of low mental strain, targeted task recommendations or supplementary activities can counteract monotony, while in times of high mental strain, supportive measures can help relieve the operator.

Adaptive work system design can not only increase safety and work efficiency but also enhance employee satisfaction, thereby making agricultural work more attractive overall.

The objective of this dissertation is to develop and evaluate methods and models for assessing and regulating human mental strain in highly automated agricultural mobile machines. It investigates how adaptive assistance functions can be designed to optimally support the operator in different strain situations and to create a sustainable and attractive work environment. The combination of ergonomic, psychophysiological, and technical approaches opens up an interdisciplinary research field that redefines the collaboration between human, machine, and work environment. A highly automated combine harvester serves as the application example for the described methods, models, and investigations. This machine is selected because of the various assistance systems available to automate the work process, thereby creating alternating mental strain states in the operator. This harvester is represented through a detailed simulation in the form of a demonstrator cabin. Within this demonstrator cabin, three studies were conducted to develop and validate both the MSS monitoring model and the complete assistance system. Finally, the assistance system was evaluated for its practical applicability, performance, and transferability to less automated machines through a field study using a real combine harvester. The structure of this dissertation is as follows.

To begin with, chapter 2 presents the current state of research and technology based on a literature review. Research work and guidelines on general design principles of assistance systems, remote sensing methods, and selected assistance systems in user-centered applications are presented. Additionally, the theoretical and ergonomic foundations for investigations of mental strain are established, followed by current findings and methods for strain prediction and detection.

From this state of research, chapter 3 identifies a research demand that leads to a research hypothesis. To address this hypothesis, four research questions are formulated.

Since the presented methods for regulating mental strain require a highly automated machine, chapter 4 describes suitable procedures for environmental monitoring. These procedures are based on low-cost and freely available data sources to ensure the broadest possible applicability.

A central component of the assistance system developed in this dissertation is the regulation of the operator's mental strain. For this purpose, the MSS must first be detected. In chapter 5, a study is used to develop a model capable of classifying mental strain. This model is then integrated as a module into the assistance system.

In chapter 6, the technical implementation of the Virtual Assistant (VA), as a comprehensive and adaptive component of the Human-Machine Interface (HMI), is explained in detail. The software architecture, interaction possibilities, and various interfaces are described.

The control intelligence for the regulation of mental strain in the assistance system is discussed in chapter 7. First, the fundamental properties of Task Recommendations (TRs) and the quantitative characteristics determined through a subject study are presented. Subsequently, the control algorithm is described.

The validation of the overall system is carried out through a comprehensive system study within the demonstrator cabin and in a real machine as part of a field study. The methods, procedures, and results are described and discussed in chapter 8.

The scientific contribution of the presented methods, studies, and results are summarized and discussed in a concluding chapter. Chapter 9 presents this discussion, including a review of the research hypothesis based on the results of the research questions.

The dissertation concludes with a comprehensive summary and an outlook on future research and development opportunities in chapter 10.

Partial results of this work have already been published in various publications [1]–[7].

2. State of research and theoretical background

This chapter presents the state of the art in the scientific and technical fields relevant to this work. In section 2.1, the different types, research objectives, and technical implementations of assistance systems in agriculture and other technical domains are examined. Particular emphasis is placed on general design principles, acceptance requirements, and safety considerations in the development of user-centered assistance systems. The domain of ergonomics and human factors, especially the theoretical foundations of subjective mental strain experienced by users, is described in section 2.2.2. Historical and current research on mental stress and mental strain states, as well as on situation awareness in selected application cases, is analyzed. Building on this, section 2.2.3 provides a detailed discussion of the technical possibilities for detecting these subjective mental strain states of the user. Finally, a technical analysis of environmental monitoring is conducted in section 2.3, due to the essential requirement for environmental monitoring in the context of this work.

2.1. User-centered Assistance Systems

The following subsections present the fundamental concepts and research required for the successful development of a user-centered assistance system. The focus is on general design principles and application examples from research and industry.

2.1.1. Development principles for user-centered assistance systems

Agricultural machinery is already highly automated and relies on a wide range of assistance systems to simplify or automate work and driving processes [28]. In particular, combine harvesters have long been equipped with, for example, steering automation, curve automation, and extensive harvesting process automation, so that operators are increasingly being transformed into supervisors [34]. However, this transformation of the operator has a major impact on situational awareness and must be taken into account in the development of assistance systems. Devitt examined cognitive factors influencing the acceptance of Robotic and Autonomous Agricultural Technologies (RAAT), including trust, loss of expertise, and social cognition. The results show that the level of automation plays a central role. With increasing autonomy, mental stress decreases, yet situational awareness may also decline, potentially impairing strategic decision-making. Therefore, semi-autonomous systems require intuitive user interfaces and clear, comprehensible information displays to promote trust and acceptance. [35]

Bashiri et al. analyzed the relationship between automation level and situational awareness in greater detail than Devitt. Their study involved 30 young, experienced drivers during the seeding process, using automation systems such as automatic steering. The work process was divided into four automation levels, from Level 1: provision of sensor data to Level 4: fully automated task execution. This approach covered the entire range from manual driving to fully automated operation with the driver in a pure supervisory role. The results show that automation systems can increase the driver's situational awareness compared to manual driving. However, when the system made changes autonomously without driver intervention, situational awareness dropped back to the level of manual driving. The highest situational awareness was achieved when the system suggested settings that had to be actively confirmed by the driver. [36]

Ahmann et al. explored how assistance systems in livestock farming should be designed to effectively support farmers in decision-making. The starting point was the observation that existing models often focus solely on technical aspects such as sensor technology and automation while neglecting human decision-making and the interaction between humans, animals, and machines. Therefore, the authors extended an existing model by adding a decision-making layer that places the human more at the center. The results show that assistance systems must not only provide data but also actively involve the user in decision-making processes. [37]

This research shows that the operator must be brought more into focus in order to develop efficient and safe assistance systems. At the beginning of the development process of a user-centered assistance system, it is essential to clarify and define the technical, legal, and ethical framework conditions. Various scientific studies and publications address this topic to provide fundamental principles and requirements for the design process of assistance systems.

Apt et al. examined the technical, legal, and organizational framework conditions for the development and implementation of digital, user-centered assistance systems. Their study emphasizes that the acceptance and effectiveness of such systems largely depend on the early involvement of users, the consideration of data protection and ethical aspects, and the adaptation to organizational work structures. They show that the challenges lie less in technological feasibility and more in the integration into existing workflows and the assurance of adequate employee qualifications. [38]

Sabattini et al. also reach similar conclusions with their approach to the design of user-centered collaborative assistance systems. They emphasize that assistance systems must not only be technically efficient but should also include adaptive interaction mechanisms, context-dependent user support, and an intelligent, adaptable system behavior. Sabattini et al. argue that systems should monitor the state, abilities, and learning curve of users and adjust their support accordingly, for example by reducing task load or by modifying control interfaces when overload becomes apparent. In this context, user-centered design means that the system is not only operable but also actively adapts to the users in order to maximize acceptance, usability, and effectiveness. Speech control and speech output represent suitable means to make systems more user-adaptive [39]. Both Apt et al. and Sabattini et al. emphasize in their findings the importance of a user-centered and interdisciplinary development process that integrates technical, legal, and social factors, such as user dynamics as well as learning and mental stress states, at an early stage [38], [39].

Forster et al. investigated the influence of different auditory warnings on drivers' responses to takeover requests in conditionally automated vehicles.

They compared a generic warning melody with a combined speech-and-melody alert. The speech-supported variant elicited faster reactions, with drivers placing their hands on the steering wheel earlier and terminating secondary tasks more quickly. Gaze reaction time did not differ significantly between the warning types. Furthermore, drivers rated the speech warning as more understandable and pleasant. The findings indicate that speech-supported auditory warnings enhance both takeover performance and user acceptance in semi-automated driving [40].

Voice User Interfaces (VUIs) have gained increasing importance as a medium for human-machine interaction in recent years. Despite their widespread adoption, particularly on smartphones, usage statistics indicate relatively low acceptance rates. This is primarily attributable to misunderstandings in speech recognition as well as concerns regarding privacy and data security. Nevertheless, a subset of intensive VUI users exists who exhibit distinct requirements and expectations for such systems. Klein et al. emphasize that designing a positive User Experience (UX) for VUIs is still in its early stages. To increase acceptance, it is crucial to consider the usage context and individual user needs. This includes developing context-sensitive interaction models that account for users' capabilities and requirements. Technological advances, such as the introduction of AI models, offer new opportunities for VUI development. For the future, the authors propose developing evaluation tools for VUIs that enable an in-depth analysis of the user experience. This could be achieved through the introduction of a UX toolkit tailored specifically to the characteristics of VUIs. Such an approach would not only improve VUI design but also contribute to the further development of assistance systems and personal assistants. [41]

The fourth Industrial Revolution (Industry 4.0) has significantly transformed industrial production and workforce support in recent years through the use of Virtual Assistants (VAs) as part of an assistance systems. VAs provide contextualized real-time information, supporting employees on the production floor as well as administrative staff, supervisors, and production managers. They contribute significantly to increasing productivity and efficiency in modern smart factories. A literature review shows that VAs provide both virtual and physical assistance. Virtual assistance delivers real-time or static information, e.g., for training purposes, while physical assistance directly intervenes in task execution and supports decision-making. Virtual assistance is more common, as it only requires software development, whereas physical assistance requires additional hardware. In practice, VAs mainly offer two types of services: integration into existing information processing systems and the processing of static information for various applications. Key limitations include security concerns in handling sensitive production data and difficulties in adapting to noisy or unstable production environments, which are often insufficiently addressed in current solutions. Future research should focus on integrating virtual and physical assistance, utilizing modern AI models such as large language models, and incorporating the Industry 5.0 paradigm. Despite the limited number of studies available, the analysis provides valuable insights into the design, functionality, and challenges of assistance systems in industrial production. [42]

In summary, it becomes evident that technological performance alone is insufficient to ensure acceptance and effectiveness. The early involvement of users, consideration of legal and ethical requirements, and adaptation to organizational conditions are crucial. Studies show that adaptive and context-sensitive interaction capabilities can improve users' trust, user experience and situational awareness. At the same time, an appropriate level of automation is necessary to avoid overload and effectively support decision-making processes.

2.1.2. User-centered Assistance Systems in practical applications

This subsection presents existing solutions and approaches to user-centered assistance systems and relates them to the assistance system described in this work. Many of the existing approaches focus on specific application areas, such as aviation or commercial vehicles, or are still largely in a conceptual phase in terms of their development status.

Li et al. examined the development of a virtual co-pilot for single-pilot operations in the aircraft cockpit, which is based on a multimodal Large Language Model (LLM). A case study was implemented in which the system processes cockpit instrument images as well as pilot speech and text instructions, and retrieves the appropriate procedures from a specialized database of manuals and checklists. [43]

The increasing automation of truck driving functions (SAE Level 2–3) creates opportunities to relieve the driver by taking over driving-relevant tasks, but at the same time presents new challenges regarding safety, vigilance, and user acceptance. Research shows that partially automated systems carry the risk that drivers neglect monitoring and engage in unauthorized non-driving activities. At the same time, different automation levels require adaptive assistance systems that neither under- nor overload the driver. The TANGO project developed an Attention and Activity Assistant (AAA) as a central assistance system. The AAA continuously evaluates the driver state, the current driving situation, and the automation level in order to offer secondary tasks in a targeted manner without compromising the system's safety margin. Central to this are the integration of driver modeling, Human-Machine Interface (HMI), and the automated driving system, which allows dynamic adaptation of the assistance. [44], [45]

The implementation followed a user-centered approach:

- Driver monitoring through video-based and 3D sensor technology enabled the detection of gaze direction, facial expressions, gestures, and interactions with devices.
- HMI concepts provided information context-sensitively and controlled the selection of secondary tasks.
- Iterative evaluation in simulators and test vehicles with professional truck drivers ensured validation of usability, acceptance, and safety aspects.

Research shows that personalized assistance systems, which consider driver state, driving situation, and automation level, make a significant contribution to the safety and user acceptance of automated driving functions. [44], [45]

Büchner et al. developed a personalized, adaptive driver assistance system for truck reverse docking. The goal is to support drivers through assistance adapted to real-time driving performance data. A Virtual Reality (VR) simulation platform was used, which captures vehicle and driver data and controls an adaptive user interface. Initial results show that the system can distinguish between inexperienced and experienced drivers using a machine learning-based expertise estimation model, allowing the assistance to be adapted to the situation. [46]

Lee et al. investigated how driver assistance systems for partially autonomous vehicles can be improved by capturing driver state and emotions. They developed a system that analyzes biosignals (Photoplethysmography (PPG) and galvanic skin response (GSR)) and uses a convolutional neural network to recognize driving scenarios and emotions, allowing real-time adaptation of vehicle control. The results show that the system responds reliably and recognizes four driving scenarios and eight emotions. [47]

Kumar et al. investigated personal assistance systems that support farmers through voice-controlled interaction. These systems use artificial intelligence, natural language processing, and machine learning to provide farmers with real-time information, tailored recommendations, and task guidance. This includes weather forecasts, crop management advice, market prices, and best agricultural practices. Through the intuitive voice-based interface, farmers can access relevant information regardless of educational background or technical experience. This accelerates decision-making processes, optimizes resource usage, and reduces operational risks. The adaptability of the assistance to different agricultural contexts and languages further contributes to broad applicability. The implementation of such systems promotes increased productivity, more efficient operations, and supports long-term sustainable agriculture and food security. Overall, the research demonstrates that AI-based smart voice assistants as personal assistance systems can provide significant value by offering farmers practical, data-driven support and advancing the digital transformation in agriculture. [48]

In summary, user-centered assistance systems have already been successfully implemented in various research studies, contributing to increased safety, enhanced user acceptance, and effective support for operators. The mental strain condition of the operators is only taken into account to a minor extent, although it has a significant influence on performance and comfort.

2.2. Mental strain in user-centered applications

This chapter examines and describes the foundations of ergonomics and human factors in the context of mental strain, as well as key established methods for predicting and detecting mental strain. It takes into account all the research and theoretical foundations necessary for a comprehensive understanding of the subsequent investigations and development steps in the following chapters of this work.

2.2.1. Theoretical background of mental strain

This subsection provides the relevant literature necessary on mental strain in the context of Human Factors Engineering, also known as Ergonomics. This is necessary for understanding the present work and the assistance system that has been introduced. It then outlines the current state of art on methods for detecting mental strain in individuals across various use cases and domains.

In Ergonomics, the distinction between mental stress and mental strain represents a central theoretical foundation for the evaluation and design of work systems. This differentiation originates from Rohmert's classical "stress-strain concept", which defines mental stress as the totality of external conditions and demands within the work system that act upon the working individual. Mental strain, on the other hand, describes the individual reaction to these external influences. While mental stress represents an objectively measurable quantity, for instance in the form of work tasks, environmental conditions, or organizational factors, mental strain refers to the resulting physiological, psychological, or behavioral reactions, which can vary subjectively between individuals. [31], [49], [50]

According to DIN EN ISO 10075, mental strain is defined as the immediate effect of mental stress on a person. It is influenced by individual factors such as abilities, skills, motivation, and current condition. This means that the same level of mental stress can lead to very different levels of mental strain across individuals. This insight constitutes a fundamental ergonomic principle, according to which working conditions should be adapted to the individual characteristics and resources of employees in order to maintain health, motivation, and performance, and to enhance the efficiency of the work system. [31], [32]

In research on mental strain, the relationship between task demands and individual response is further differentiated. De Waard describes mental strain as the ratio between task demands and available cognitive resources. Accordingly, mental strain does not arise solely from the absolute level of stress but from the balance between the intensity of the demands and the available resources of the individual. High levels of mental stress only lead to high mental strain when the available resources are insufficient to cope with the demands efficiently. This perspective highlights the dynamic interaction between task, individual, and situation, which is crucial for maintaining a balanced stress–strain relationship. [30]

Research on work performance and efficiency shows that mental strain also largely depends on how individuals manage limited resources. Zijlstra emphasizes that efficiency in work behavior is not achieved merely by reducing mental stress, but through the optimal use of available cognitive and physical resources, taking into account the current cognitive state. This underlines the necessity of designing work activities in such a way that they enable a balance between demands and resources, preventing both overload and underload. [51]

An additional theoretical framework is provided by Wickens' Multiple Resource Theory [52], which conceptualizes mental strain as a function of multiple simultaneously utilized resources. Relevant is not only the magnitude of the strain but also the types of resources such as visual, auditory, cognitive, or motor that are activated by the task demands. Two tasks with an identical objective workload can lead to different levels of mental strain if they draw on distinct resource pools or compete for the same ones. This perspective explains why multitasking, task switching, or sensory interference situations often cause increased strain, even when the objective workload remains constant [52].

De Waard also describes how mental strain influences performance and indicates that productivity increases with rising demand or strain until a moderate level is reached. At low strain levels, individuals may feel understimulated and perform poorly because their focus is insufficient. With increasing strain, attention, motivation, and cognitive resources reach their peak, which enables optimal performance. If mental strain continues to rise beyond this point, strain begins to overwhelm cognitive control. This leads to errors, reduced working-memory capacity and impaired decision-making. The resulting relationship between strain and performance takes the form of an inverted U and emphasizes the importance of balanced strain for achieving favourable results [30], [53]. This is consistent with the arousal theory proposed by Yerkes and Dodson in 1908 [54].

Another empirical finding regarding the effects of work-related influences on mental strain and performance emerges from research on electronic monitoring in the workplace. Backhaus, in a meta-study analyzing various qualitative and quantitative studies, found that electronic monitoring can increase performance by 14%, while simultaneously increasing mental strain by 34% and reducing employee trust and commitment by more than 37%. These findings emphasize that attention should not only be paid to objective stress factors but also to the psychological consequences of control and surveillance. [55]

In summary, mental stress encompasses the external, objectively describable influences of work, while mental strain represents the resulting individual response, depending on personal disposition, experience, and resource availability. This distinction and the consideration of psychological effects have far-reaching implications for the design of human-centered work systems. The goal is not the general minimization of stress, but the optimal alignment between stress and the individual's resources, so that mental strain remains within a beneficial range. Only under these conditions can health, motivation, and performance be maintained in the long term, thereby increasing the overall efficiency of the work system.

2.2.2. Methods of mental strain prediction

As early as 2003, Neerincx developed a model that enables a qualitative prediction of mental strain [56]. This Cognitive Task Load (CTL) is a function of three central stress factors. The factor "Time occupied" describes the extent to which the available time is consumed by tasks. Building on the findings of Beevis, it is assumed that an occupancy of more than 70–80% of the available time leads to overload and significantly reduces performance [57]. The second factor, "Level of information processing", is based on Rasmussen's "Skill-Rule-Knowledge" framework and is divided into three levels [58]. At the "skill-based" level, actions are largely automated, resulting in low strain. The "rule-based" level is characterized by the application of learned rules and "if-then relationships". The "knowledge-based" level imposes the highest cognitive demands, as it requires analytical problem-solving and creative thinking. Since more complex problems often consist of many smaller subproblems, a third factor, "Task set switches", is included, describing how frequently and abruptly one must switch between different tasks, goals, or contexts. High switching rates lead to increased mental strain, as each switch requires additional cognitive resources [58]. From this model, four different critical strain regions result, each characterized by a specific set of stress factors. Figure 2.1 shows these regions in the Cognitive Task Load (CTL) space.

Building on this work, Jeschke developed an extended analytical model for predicting mental strain in process control [59]. The objective of this further

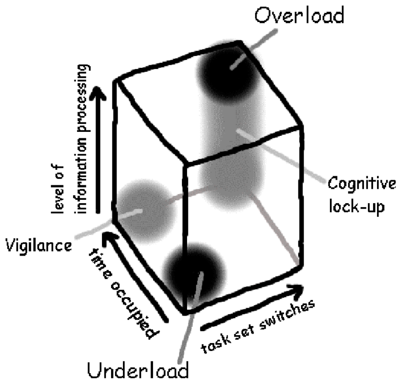


Figure 2.1.: Critical regions inside the cognitive task load space by Neerincx [56]

development was to determine the optimal level of strain for operators and to enable a prediction of the expected mental strain already during the initial task planning phase. Similar to the original CTL model, this model is based on the three stress factors “Time occupied”, “Level of information processing”, and “Task set switches”. The factors are combined into a quantitative estimation, in contrast to the qualitative estimation of Neerincx of the strain state. Moreover, the model extends the critical strain regions to include the area of mental satiation defined in DIN EN ISO 10075 [32], [59].

However, both Neerincx and Jeschke do not sufficiently account for the psychological and physical characteristics of individuals. Their approaches aim to predict strain by considering only objective stress factors. While this enables an analytical computation of strain prediction, it contradicts the definitions by Rohmert, Schlick, and DIN EN ISO 10075 presented in section 2.1.1.

To integrate the individual characteristics of a person into the assistance system presented in this work, a strain monitoring system is required that can capture the subjective mental strain of each user. Therefore, the following chapter presents key research findings and methods for detecting and classifying Mental strain states (MSSs).

2.2.3. Methods of mental strain detection

The concept of detecting the MSS has gained increasing relevance in recent years, particularly in the field of driver assistance systems. This can be attributed to the fact that various physiological and psychological factors can have a significant impact on driving behavior and accident risk. Therefore, reliable detection of the driver's condition is of central importance in order to initiate safety-related measures if necessary [60]. Comparable approaches can also be found in other safety-critical domains, such as air traffic monitoring, where multidimensional analysis systems are being tested for real-time assessment of the operator condition [33]. In addition to safety-related applications, adaptive methods for user state analysis also offer potential for improving efficiency and user comfort when interacting with technical systems. For example, a research project in the field of truck transportation investigates how integrating mental strain detection and providing additional activities can enhance driving comfort and optimize the overall user experience [44], [45].

In general, mental strain detection methods can be classified into three categories [61], [62].

Performance measurements:
By requiring participants to perform a secondary task concurrently and assessing their performance, researchers can indirectly estimate subjective mental strain [63]. Continuous task-performance measures provide a direct and robust indicator of human performance, even during extended activities. However, when used in isolation, such measures exhibit limited diagnostic value for identifying specific sources of workload and thus have restricted utility. When integrated with complementary measures, they become valuable dependent variables, as their variations can be meaningfully interpreted in relation to other indicators. This combined approach enhances researchers' ability to formulate and test causal explanations of mental strain. [53]

Self-report measurements:
Self-report measures are widely employed because they are easy to administer across diverse contexts and computationally inexpensive to evaluate, even when they include multiple dimensions. Their scores can typically be derived using simple mathematical operations, making them accessible to nonexperts.

However, these measures are most often applied after task completion, which avoids interference with the primary activity but limits temporal granularity and reduces reliability for longer tasks. Moreover, because the resulting scores reflect subjective perceptions, comparisons across participants on an absolute scale are challenging [64]. To capture subjective mental strain at defined intervals, participants commonly complete questionnaires such as Zijlstra's Rating Scale Mental Effort (RSME), a one-dimensional instrument on which individuals rate their perceived mental effort from 0 ("no effort") to 150 ("extreme effort"). Despite its simplicity, the Rating Scale Mental Effort (RSME) has demonstrated strong validity and allows researchers to reliably differentiate cognitive strain across tasks, conditions, and individuals, making it well suited for psychological and ergonomic research [51]. Another widely used tool is the multidimensional NASA Task Load Index (NASA-TLX), which provides a more comprehensive assessment of perceived workload [65].

Psychophysiological and neurophysiological measurements:
Particularly the psychophysiological measurements show high application potential, as the collected data are difficult to manipulate, can be continuously recorded and enable real time analysis [33]. Various studies have demonstrated a strong relationship between psychophysiological parameters and mental strain [66]–[70].

Psychophysiological measurements cover a wide range of indicators. These include cardiovascular activity reflected in heart rate and its variability, respiratory activity measured through breathing rate and oxygen uptake and ocular responses captured through pupil size, blinking and gaze direction. They also encompass skin related measures such as electrodermal activity and skin temperature, neuroendocrine markers including salivary cortisol and amylase as well as brain activity assessed with electroencephalography and functional near infrared spectroscopy. Each signal type possesses distinct characteristics and offers specific advantages and limitations for the assessment of mental strain [53].

Cardiovascular signals are commonly employed due to their minimal intrusiveness and relative ease of acquisition. However, they are highly sensitive to physical exertion and additional confounding influences such as emotional arousal or respiration. Physical factors must therefore be controlled carefully in experimental settings [71], [72]. Blood pressure is not used as often as other cardiac measures because of its intrusive nature [73].

Respiratory measures form another category. The respiration rate is a measurement of the number of breaths taken per unit time. It generally increases in response to an increase in mental strain [73], [74]. Respiratory rate, like heart rate and its variability, is a vital sign that can be easily measured and recorded with minimal invasiveness. It is essential that the level of physical exertion remains consistent throughout the experiment, since variations in physical effort have been shown to alter the respiratory rate. Researchers have determined that the amount of oxygen a person uses can effectively measure mental strain. This appears to be directly associated with the level of exertion experienced by the individual [75].

The category of ocular measures is well established and encompasses a range of eye-related activities, including the blink rate, blink closure duration, gaze angle, pupil size, pupil diameter and pupillary responses [76]. The blink rate indicates the frequency at which the eyes close in a given time period. The blink closure rate, on the other hand, denotes the time spent blinking [73]. Research has shown that ocular signals respond sensitively to variations in mental strain. In particularly, pupil diameter increases in line with mental strain, and is sensitive to several demands and emotional states [53], [61], [62]. They are however strongly affected by environmental conditions such as illumination and can exhibit saturation effects under very high workload levels, which reduces their discriminatory power [77].

Neurophysiological measures are widely employed in the assessment of mental strain, largely due to the fact that Electroencephalography (EEG) provides a direct measure of brain activity, rather than relying on peripheral physiological responses mediated by the brain. Despite this advantage, the use of neurophysiological measures presents practical challenges, including substantial equipment demands, the necessity for extensive signal processing, and limited feasibility for field applications [78]. Recent developments are beginning to reduce these limitations [53].

Skin-based measures constitute another important category in the assessment of mental strain. Monitoring temperature across different body regions is a well-established approach. For instance, Hancock suggested that auditory canal temperature may reflect overall changes in mental strain, although its utility is constrained by the signal's inherent inertia [53], [79], [80]. Skin temperature more generally reflects peripheral sympathetic nervous system activity, which is activated under mental strain and can therefore serve as an indicator of workload. However, its interpretation is complicated by environmental temperature effects, and sensitivity varies across body regions, for example, forehead skin is less responsive to different types of strain compared with nasal skin [81]. In addition to temperature, other physiological indices have been developed as potential markers of mental strain, including electrodermal activity and galvanic skin response [53], [82]–[84].

Driven by significant advances in sensor technology and signal processing techniques, physiological measures are becoming increasingly important in experimental research. These devices allow for the continuous monitoring of physiological reactions and neurophysiological activity without disrupting the performance of the primary task. However, implementation require often more administration than other measurement approaches, as extensive pre-processing is usually required to minimize artifacts originating from the body or the surrounding environment. Furthermore, such pre-processing can require significant computing power and is typically performed offline, limiting its use in real-time mental strain evaluation [53]. Physiological parameters and eye tracking data are additionally subject to high interindividual variability, which means they do not provide universally valid indicators for determining mental strain. Calibration tasks are therefore generally required when using eye tracking systems to allow valid conclusions about individual mental strain [85]. Similarly, defining an individual reference value can improve the detection of mental strain when collecting physiological data [86].

Three research studies by Funk et al. form a consecutive series focusing on detecting and evaluating psychological strain states within the project "Fahrerkabine 4.0", based on psychophysiological measurements [7]. The aim of these studies was to develop a system for monitoring mental strain that can reliably assess the current mental strain state of operators. Initially, an experimental environment was developed that generated a controllable spectrum of mental load. Stress was induced through a screen-based monitoring task with varying difficulty and, in a second stage, by adding a visual secondary task in the form of a mental rotation task. Using the RSME, a quasi-linear increase in perceived strain from low to high effort was observed, confirming the validity of the experimental setup [69]. Subsequently, the method was extended. In addition to the visual secondary task, an auditory one was introduced, enabling a broader spectrum of mental strain. Three laboratory studies ($N_1 = 17; N_2 = 8; N_3 = 21$) revealed significant differences between task conditions in variance analyses. Subjective strain increased with task complexity and was associated with longer reaction times and higher error rates. The method thus served as a valid foundation for developing physiological measurement systems [70]. Finally, the concept was translated into a model for strain detection. For 21 participants, 20 physiological indicators from ocular and cardiovascular measurements were collected and linked to classified RSME values using Naive Bayes algorithm. The individually trained models achieved an average prediction accuracy of 57.4% [68]. These studies resulted in a validated, algorithmically applicable system for inducing, measuring, and predicting psychological strain. Thus, Funk et al. established a scientifically grounded basis for the studies conducted in this work, for the development of the MSS monitoring model in chapter 5 and for future work systems.

The presented fundamentals and studies to detecting mental strain demonstrate that reliable measurement of mental strain is possible. However, due to individual differences and external influencing factors, careful selection of psychophysiological parameters and the experimental environment is required.

2.3. Environmental monitoring in agriculture

The chapter on the state of the art concludes with a comprehensive review of the diverse environmental sensing systems that have been developed for the purpose of monitoring the machine surroundings and acquiring supplementary field and ground information for the work process. This review is limited to systems and methods that do not require any further machine-mounted sensor systems, as this approach is also pursued in the work presented here.

To further advance the necessary automation of agricultural machinery, comprehensive environmental sensing and monitoring are required. Without fully automated operational phases of agricultural machinery, the provision of task recommendations and thus an improvement in work-life balance and work efficiency can only be realized with difficulty. For this reason, this work presents methods for environmental monitoring that have been integrated into the described assistance system. The following sections outline the theoretical background and key research in the fields of remote sensing and precision farming.

Liaghat et al. investigated the use of remote sensing technologies as central tools in precision agriculture. They demonstrated that, when combined with a Global Positioning System (GPS) and a Geographic Information System (GIS), as well as multispectral and hyperspectral sensors, these technologies can help to capture field and crop variability, enabling more precise management of fertilization, irrigation and crop protection. Liaghat et al. emphasize that remote sensing does not provide all agronomic information but serves as a valuable complement to soil and operational data, contributing to increased efficiency, environmental compatibility, and sustainability. [87]

Wang et al. developed a method for precise mapping of rice cultivation areas by combining a U-Net model for automatic field boundary detection from high-resolution satellite images with a Random Forest classification of Sentinel-1 radar data. This combination of deep learning and time series analysis enabled highly accurate detection of field structures and rice fields, with a Jaccard Index (JI) of 0.801 for field boundaries and 0.953 for rice field mapping. [88]

Watkins et al. compared various Object-Based Image Analysis (OBIA) methods for the automatic delineation of agricultural fields from Sentinel-2 satellite images. Segmentation algorithms were used to group image pixels into homogeneous objects based on spectral and spatial features such as NDVI, texture, and color differences. The study showed that well-tuned segmentation parameters are crucial for precise boundary detection and that OBIA approaches provide a reliable foundation for automated field mapping. [89]

The presented research demonstrates that remote sensing data, combined with intelligent algorithms, has significant potential to complement soil data and support automatic field mapping and field boundary detection.

3. Research Demand

Despite numerous advances in the development of user-centered and adaptive assistance systems, there remains a significant research demand regarding their integration into the agricultural context. The current state of art in section 2.1 shows that existing approaches often focus on the technical performance and functional scope of the systems, while the well-being and individual mental strain of users are only given secondary consideration. Although systems that measure the user state and react contextually already exist in other application domains, such as the automotive sector, these primarily target safety and attention aspects rather than actively promoting human well-being. So far, there is no transferable concept that addresses both the users' mental strain and their well-being in agricultural working environments.

As shown in section 2.3, there is a demand for additional research in the field of environmental sensing. High-resolution sensor systems for monitoring the agricultural environment, such as for vehicle navigation, are currently associated with high investment costs and are only limitedly available in series-production machines [90], [91]. Remote sensing, in combination with machine learning methods and multispectral image analysis, offers great potential for reducing costs and expanding the data base. Especially the linkage of remote sensing data with internal operational data and machine information opens new perspectives for the development of adaptive assistance systems. In addition, comprehensive and functional environmental sensing is essential for actively distracting operators with task recommendations and thus forms a necessary basis for this work.

Research on the detection of mental strain has made significant progress in recent years. Various studies from section 2.2 demonstrate the validity of psychophysiological measurement methods as well as the possibility of classifying and predicting Mental strain states (MSSs) using machine learning methods. Nevertheless, no integrated assistance systems currently exist that make use of these findings to detect the operator's strain state in real time and, based on this, generate adaptive task recommendations for the active regulation of mental strain. Such a system would not only react to state changes but would proactively aim to maintain mental strain within an optimal range, for instance through situationally adjusted information presentation, adaptive task recommendations, and intuitive user interfaces. The transition from passive mental strain detection to active, user-oriented mental strain regulation therefore represents a central research focus.

The present work introduces an assistance system that not only provides the desired technical performance but also takes the individual mental strain of the operator into account. This system is designed to adapt to the operator's specific needs, detect the current strain state, and intervene in a regulating manner through targeted task recommendations. Interaction with the operator is ensured through a Virtual Assistant (VA). The aim is to maintain the mental strain state within an optimal range to maximize both well-being and performance of the operator. This requires an interdisciplinary combination of engineering insights from current research on assistance systems and ergonomic methods of strain detection and psychophysiological state analysis under the specific conditions of agricultural working environments.

Based on this, a research hypothesis can be formulated, which will be examined and either confirmed or refuted in the following chapters:

"It is possible to deliberately influence the cognitive strain of an agricultural machine operator by means of an assistance system."

To verify the hypothesis, four overarching research questions are formulated. Their answers enable a comprehensive examination of the hypothesis:

1. *"How can critical situations in harvesting operations be reliably detected by the assistance system?"*
   In chapter 4, two different approaches and methods are presented that allow this research question to be answered.

2. *"How can the mental strain state of an agricultural machine operator be monitored or classified?"*
   This second research question is addressed in chapter 5 as part of a participant study for the development and evaluation of a model for strain detection.

3. *“How can the influence of Task Recommendations (TRs) on mental strain be reliably quantified and generalized?”*
   To answer this research question, a study analyzing selected TRs is presented in chapter 7.

4. *“How can an assistance system be designed to reliably regulate the mental strain state of agricultural machine operators?”*
   Finally, the last research question is addressed in chapter 8 by means of a comprehensive system study.

The final verification of the research hypothesis, a summary of the scientific contribution and a detailed discussion are presented in chapter 9.

4. Environmental monitoring

Comprehensive environmental monitoring is essential for the assistance system presented in this work. The following sections describe suitable approaches and methods for implementing such monitoring.

4.1. Motivation

The increasing automation in the agricultural sector is fundamentally changing the role of the machine operator. As already discussed in chapter 1, this development leads to operators of agricultural machinery increasingly experiencing phases of low mental strain. In such phases, the targeted integration of assistance systems can help provide active distraction, for instance through information offerings or interactions that improve operators' well-being and work-life balance. However, this requires that the operator can temporarily turn away from the driving task without compromising safety. This requires reliable and safe machine environment monitoring to prevent possible damage to the machine or hazards in the surrounding area.

A key challenge in implementing such systems lies in the detection of field boundaries and relevant obstacles. These environmental factors are particularly significant around field edges, where obstacles such as vegetation and traffic routes are more frequent. Within the field area, irregular structures such as lodged crops, power poles, or wind turbines also occur, which demand increased attention and precise detection.

In addition, there are high demands on the transferability of the systems to different types of machines. While combine harvesters tend to perform specialized and repetitive tasks during the harvest season, tractors are much more versatile in their use due to their wide range of attachments. This increases the relevance of robust, adaptable environment and obstacle detection for different types of machines. Despite all the progress that has been made, humans remain indispensable in many situations, especially when it comes to non-automated or safety-critical functions.

The aim of this chapter is therefore to investigate approaches for field boundary and obstacle detection using the example of a combine harvester. As high-quality environmental sensors are currently only implemented in a few autonomous prototype machines and are associated with high costs, two cost-efficient alternatives have been developed and are discussed below [90], [91]:

- Mapping based on historical user and machine data (Retrospective approach), and
- remote sensing using satellite data (Remote sensing-based approach).

After detecting both obstacles and field boundaries, the assistance system calculates the remaining time before the vehicle reaches these elements. This interval represents the time available to the operator for executing the relevant Task Recommendations (TRs) and is incorporated into the system's process for selecting appropriate TRs. The calculation of the available time is described as an example in the first approach.

4.2. Mapping methods

4.2.1. Retrospective approach

A promising approach for determining the field contour and detecting obstacles is based on the retrospective analysis of historical user inputs in combination with Global Navigation Satellite System (GNSS) data. In this method, past operator actions are linked with the corresponding Global Navigation Satellite System (GNSS) position data to derive conclusions about future events. This approach enables offline and largely autonomous analysis as well as mapping of the field to be processed. This approach is based on a Master's thesis from Tammo Kaper, supervised by Metzger [92].

4.2.1.1. Material and methods

A fundamental requirement is a functional and accessible GNSS module. However, since the first GPS-based parallel guidance systems were introduced in the late 1990s and automatic steering systems for tractors have been available since 2003, this requirement can be considered fulfilled in almost all cases. Moreover, positioning accuracy has improved significantly over the past 20 years and, depending on the boundary conditions and Differential Global Navigation Satellite System (DGNSS) correction data, ranges between 50 cm and 2.5 cm [93].

The method is designed to record and process all relevant user inputs and machine events and link them to the current GNSS coordinates. Since the developed method accounts for potential measurement inaccuracies and maintains a safety margin from detected obstacles as well as from the field boundary, absolute GNSS accuracy is of minor importance. [92]

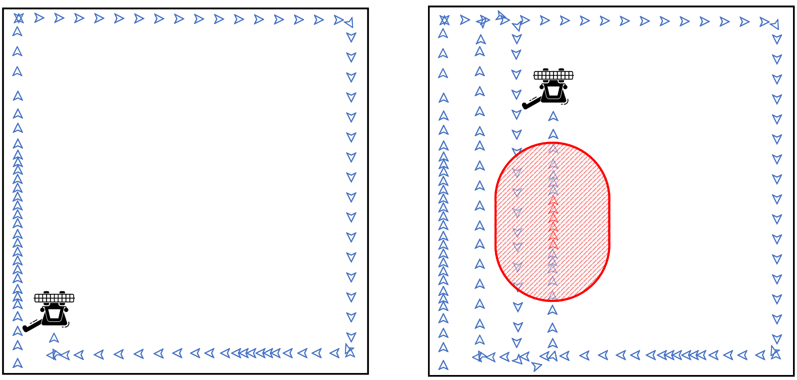
To detect user inputs, Controller Area Network (CAN) data from the carrier vehicle, in this case a combine harvester, are utilized. Consequently, access to the CAN bus must be ensured.

Mental stress detection:
The idea of this approach is to link situations in which the user experienced an exceptionally high amount of mental stress with the GNSS coordinates and thus make them usable for future events and calculations. To determine mental stress, the weighted amount of different operator interactions with the machine within a sliding time window is evaluated. All control elements of the multifunction armrest are included. These include settings for the threshing drum, fan, sieve, and rotor, as well as buttons for functions such as all-wheel drive, differential lock, and opening or closing the grain tank cover. The operation of the comfort panel (e.g., multimedia volume) and the touch display (e.g., Claas CEBIS) is also considered. At the control lever (e.g., Claas C-MOTION), all buttons except for the lever position are monitored. These include the controls for the cutting unit (including reel), unloading auger, and grain tank discharge. The actual driving speed is used instead of lever position. If specific acceleration or deceleration thresholds are exceeded, this is, analogous to a button press, counted as operator interaction. The threshold values were empirically determined based on real driving data. All registered interactions within the defined period are summed and converted into a

stress value. If this value exceeds a predefined threshold, the corresponding position is marked as a Point of Interest (POI). [92] The method presented can be combined with the Mental strain state (MSS) monitoring system of the demonstrator cabin described in chapter 5, thus providing an additional layer for assessing the operator's state. In contrast to the MSS monitoring system introduced in chapter 5 the method described here captures objective mental stress indicators. It is therefore particularly suitable for obstacle detection, as such situations are often associated with increased operator workload.

Points of Interest (POIs):

For the identification of Points of Interest (POIs), a prior harvesting of the headland is required. Since headland harvesting already requires the operator's full attention, this prerequisite does not significantly limit the usefulness of the method. The GNSS data recorded during this process define the field boundary and reduce the area to be considered, as shown in fig. 4.1a. If a stress threshold is exceeded or an increased mental strain is detected by the mental strain monitoring system, the corresponding point in the grid is marked, as shown in fig. 4.1b.



(a) Headland harvesting

(b) Registration of a POI

Figure 4.1.: Process of POI registration based on Kaper [92]

The goal is to detect obstacles (e.g., trees) or abrupt changes in the crop (e.g., lodged grain or varying crop densities) and mark them as POIs. Due to spatial correlations, such events are likely to occur in neighboring driving lanes as well. The area around a detected point is modeled as circular, since no precise information about the shape of the underlying object is available. The radius of this circle is determined by the working width of the cutting unit to ensure that adjacent driving lanes are also covered. When multiple POIs are located close to each other, their zones overlap, resulting in a deviation from the idealized circular shape. To predict the available time, the future route of the combine harvester is forecasted. [92]

Distance determination:
To predict the remaining time until the next critical point, the future route of the combine harvester is extrapolated based on compass data from the GNSS module and the distance already traveled. The calculated driving line is extended until it intersects a grid cell with a high mental stress value. Figure 4.2 illustrates this method graphically.

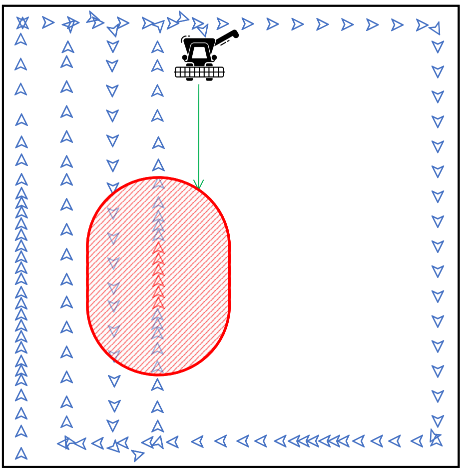


Figure 4.2.: Driving time prediction based on Kaper [92]

Using the position data of the POI grid cell and the current position data of the combine harvester (CH), the available distance is determined using the equirectangular projection [94]. Figure 4.3 illustrates, by way of example, point “P” and the angular relationships including the associated latitude circle.

First, the arc distances for geographic longitude ($d_{\Delta\lambda}$) and latitude ($d_{\Delta\phi}$) are calculated separately. Here, λ and ϕ are given in radian. [92]

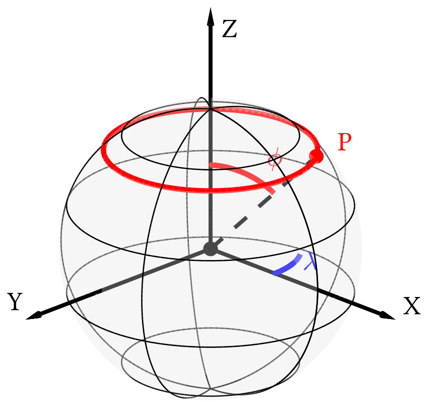


Figure 4.3.: Angular relationships in a spherical coordinate system

$$\begin{aligned} \Delta\phi &= \phi_{cell} - \phi_{CH} \\ d_{\Delta\phi} &= r_{Earth}\Delta\phi \end{aligned} \tag{4.1}$$

The calculation of the arc distance along the latitude is more complex due to the dependence of the radius of the latitude circles on the latitude ($r_{lat}\left(\phi\right)$). As a simplification, the latitude of the POI grid cell and the current position data of the combine harvester (CH) is averaged (ϕ_{mean}).

$$\begin{aligned} \Delta\lambda &= \left(\lambda_{cell} - \lambda_{CH}\right) \\ r_{lat}\left(\phi\right) &= r_{Earth}\cos\left(\phi\right) \\ d_{\Delta\lambda,mean} = r_{lat}\left(\phi_{mean}\right)\Delta\lambda &= r_{Earth}\cos\left(\tfrac{\phi_{cell}+\phi_{CH}}{2}\right)\Delta\lambda \end{aligned} \tag{4.2}$$

Subsequently, a simplified resulting distance d is calculated from the two coordinate distances. r_{Earth} denotes the average radius of the Earth and is approximated as $r_{Earth} = 6\,371.009\,\mathrm{km}$.

$$d = \sqrt{\left(d_{\Delta\phi}\right)^2 + \left(d_{\Delta\lambda,mean}\right)^2} = r_{Earth}\sqrt{\left(\Delta\phi\right)^2 + \left(\Delta\lambda\cos\left(\frac{\phi_{cell}+\phi_{CH}}{2}\right)\right)^2} \tag{4.3}$$

Since the unloading process requires increased attention from a combine harvester operator, even with possible assistance systems, the remaining range of the grain tank is also taken into account. The shortest resulting distance (d_{min}) is converted into the remaining time (t_{rem}) using the current speed (v_{CH}).

$$t_{rem} = \frac{d_{min}}{v_{CH}} \tag{4.4}$$

Since obstacles and changing field conditions, as well as entry and exit into the crop, demand particular attention from the operator, these factors are considered in the time estimation. During the extension of the driving line, it is checked whether it intersects a previously driven section. The headland represents a natural boundary. If the projected driving line intersects an already driven track, the intersection angle is calculated. The point is only used for distance determination if this angle exceeds a defined threshold. This prevents parallel tracks, e.g., caused by fluctuations in the GNSS compass, from being falsely interpreted as a POI. The method reduces potential misclassifications, particularly on long driving paths. If no relevant intersection is found, the analysis ends at the field boundary. At the end of each calculation step, the prediction model outputs the calculated remaining time as well as additional relevant information, such as the distance. [92]

4.2.1.2. Results and discussion

This section presents the results of the proposed method based on a test drive in Mecklenburg-Vorpommern. The goal is to verify the functionality of the algorithm under real operational conditions and to critically assess its performance as well as potential weaknesses.

POIs and prediction of the remaining time:
To validate the method, previously recorded machine data from a harvesting operation with a total duration of 4.75 hours were used. During this time, two fields were worked on, with the combine harvester spending the majority of the time in the second field. Figure 4.4 shows the zones of increased driver mental stress detected by the system during the operation of this second field. It is visible that the thresholds for detecting POIs were exceeded mainly near the field boundaries and the access point (top right in the figure). This clustering at the field edge is plausible, as turns and increased attention are typically required here. Additionally, POIs can also be observed in the

interior of the field. These are attributed to increased user interactions, such as adjustments of machine settings. Since only CAN data were used and no subjective driver feedback (e.g., regarding crop quality or storage grain) is available, the exact cause of individual POIs cannot be determined with certainty. [92]

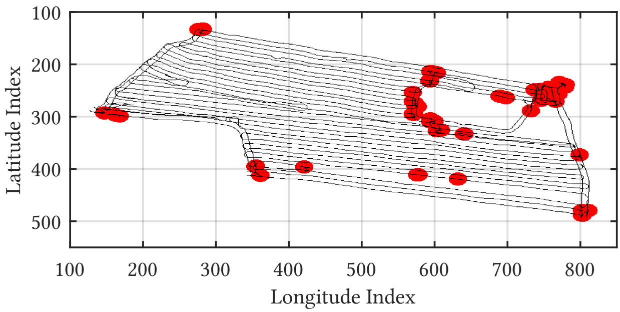


Figure 4.4.: POIs during the test drive based on Kaper [92]

The remaining time available for the next TR is the central output of the algorithm. This output is based on the predicted driving line and the distance to detected mental stress zones or field boundaries. Figure 4.5 shows the predicted remaining time for an exemplary driving lane. Only from the first relevant intersection with an already traversed lane does the system provide usable values. Before this point, the field boundary is unknown, leading to a lack of meaningful information. [92]

The displayed time series shows a typical triangular pattern: ideally, the remaining time decreases linearly with each passing second. This pattern is clearly visible between seconds 2450 and 2650. Fluctuations before and after this interval can mainly be attributed to variations in speed and steering corrections. The closer the machine approaches the end of the driving lane, the more accurate the calculated time becomes. [92]

The variability of the prediction is discussed based on several representative points from Figure 4.6. Point “1” marks the transition from a turning maneuver to a new driving lane. Due to the reduced entry speed, a comparatively high remaining time is initially predicted, which rapidly decreases with increasing speed. Point “4” illustrates that even within a lane, changes in speed can lead to noticeable fluctuations in the predicted time.

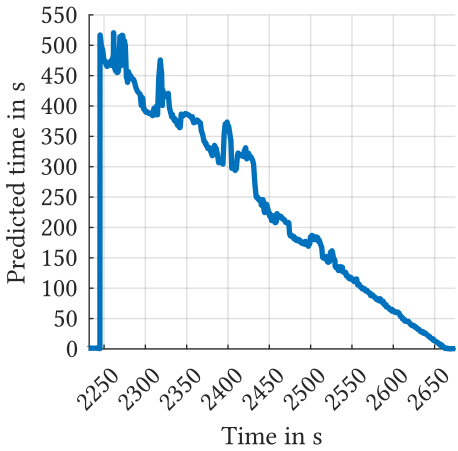


Figure 4.5.: Predicted remaining time based on Kaper [92]

Points "2" and "3" on the other hand, show incorrect detections of the end of the predicted lane. These occur, among other reasons, when a neighboring lane or a POI is falsely recognized as the end of the current lane due to compass errors. The field geometry and the harvesting strategy have a significant impact on this. For example, a curved harvest edge increases the likelihood of incorrect intersection calculations with parallel lanes.

Further deviations in the time prediction arise from lane cancellation or the temporary reversing of the vehicle, for instance, to process sub-areas of the field separately.

Limits and errors:
Due to its operation with geographic coordinates, the system exhibits a critical line along the 180th meridian. When using decimal degrees, a transition occurs along this line from positive to negative counting. Crossing this line leads to errors based on the calculation method used. Because the area is mostly over the sea, sparsely populated, and subject to distortions in common satellite images, this critical line can be considered irrelevant. [92]

Further inherent sources of error arise from the use of a grid system for area modeling and from measurement errors of the GNSS module. The target positions correspond to the center of each grid cell. This results in a maximum

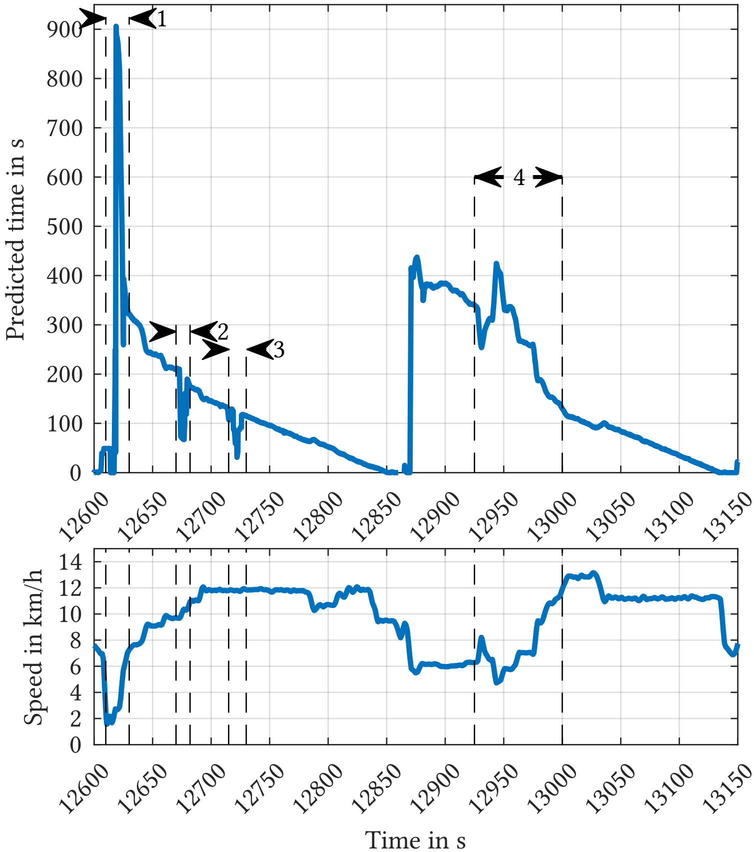


Figure 4.6.: Influence of speed on prediction based on Kaper [92]

positional error equivalent to half the length of a grid cell's diagonal. Figure 4.7 illustrates these relationships. [92]

The maximum theoretical deviation x for an edge length a is given by:

$$x = \sqrt{\left(\frac{a}{2}\right)^2 + \left(\frac{a}{2}\right)^2} = \frac{1}{\sqrt{2}}a \tag{4.5}$$

Additionally, the GNSS accuracy (s_{GNSS}) affects both the start and end points of distance calculations. Deviations for current systems used in agriculture

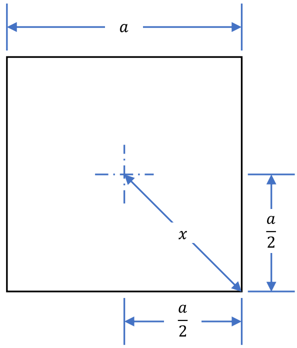


Figure 4.7.: Deviation within the grid cell based on Kaper [92]

are in the range of 2.5 cm to 30 cm [93]. The update interval (Δt_{Update}) of the GNSS module also influences accuracy, as the combine harvester moves between two position recordings. Considering these factors, the theoretically maximum error is:

$$F_{max,res} = \frac{1}{\sqrt{2}} a + 2 \cdot s_{GNSS} + v_{CH} \cdot \Delta t_{Update} \tag{4.6}$$

Under conservative assumptions of $s_{GNSS} = 5\,\mathrm{cm}$, $a = 1.5\,\mathrm{m}$, $\Delta t_{Update} = 1\,\mathrm{s}$, and a speed of $v_{CH} = 10\,\mathrm{km\,h^{-1}}$, the maximum error is 3.9 m, corresponding to a temporal error of approximately 1.4 s. This error is considered minor in the overall context and has only a limited impact on the decision-making system, particularly compared to uncertainties in route projection or driving dynamics changes. [92]

Conclusions:

The presented method for detecting obstacles or mentally demanding situations during harvesting, as well as for predictive calculation of available time for TRs, demonstrates high potential for improving the assistance system presented in this work. Analysis of real machine data over an extended operation period shows that central areas of mental stress, particularly during turning maneuvers and field access, can be identified. The localization of POIs follows a plausible pattern and correlates with known labor-intensive phases, such as lodged crops and turning maneuvers, during harvesting. This is reflected by clusters within the field and accumulations along the field boundary.

The additional time prediction provides a usable forecast for the remaining time for the next possible TR, which shows an almost linear decrease under optimal driving conditions. At the same time, limitations of the method become apparent, especially regarding variations in driving speed, field geometry, and the detection of exact lane endpoints. These effects lead to locally increased prediction deviations but do not impair the fundamental functionality of the system.

The conducted error analysis confirms that the main system-inherent uncertainties, such as those caused by GNSS inaccuracy, grid structure, or position update intervals, are quantifiable and manageable within the context of the application. The methodological weaknesses lie less in the technical foundation and more in representing real field-specific uncertainties.

Overall, this chapter demonstrates that the combination of mental stress analysis and time prediction provides a solid basis for developing adaptive assistance systems. To further enhance practical utility, future work should focus on more robust handling of complex field geometries, integration of machine-specific condition data, and validation through subjective driver feedback.

4.2.2. Remote sensing-based approach

This second approach focuses on the automated detection of field contours and relevant obstacles (POIs), relying solely on satellite imagery and GNSS data. The segmentation is achieved through a seed-based region growing algorithm combined with morphological operations and multitemporal image analysis. The following summary outlines the method, its evaluation across various fields, and the key findings regarding detection performance and optimization potential. The detailed methods and all findings already have been published at agricultural-engineering.eu [4] and is based on a Bachelor's thesis from Benedikt Jochum, supervised by Metzger [95].

4.2.2.1. Materials and methods

Initial satellite data:
Of the various Earth observation missions, Sentinel-2 was selected due to its high spatial and temporal resolution, its 13 available spectral bands, and its free usage (see table 4.1). Two metrics were tested to quantify the health and density of vegetation: the widely used Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Water Index (NDWI). The NDWI provided superior segmentation performance and was therefore used in this approach.

Table 4.1.: Overview of selected earth observation missions with the respective resolution of the satellites [4]

Mission	Spatial resolution in m	Spectral resolution in channels	Time resolution in days
Sentinel-2	10 20 60	13	5
Landsat 8	15 30	11	16
Spot 7	1.5 6	5	1-3
World View 3	0.31 1.24 3.7 30	29	< 1
Rapideye	6.5	5	< 5.5

The detection algorithm, shown in fig. 4.8, is divided into three major process steps:

Pre-processing:
Satellite data is retrieved dynamically using the Sentinel Hub API, centered around the GNSS location of the machine [8]. A 3.5 km square area ensures full coverage of typical field sizes. Images from April to September of the last two vegetation periods, with up to 10% cloud cover, are used to improve

robustness. This has the particular advantage that many cloud-free days are available and the decisive growth phases of the plants are covered [9], [96]. Finally, the main seed point and, optionally, additional seed points are calculated and defined for the main process.

Main Processing:
Four segmentation approaches were analyzed: pixel-based, model-based, edge-based, and region-based. Region-based segmentation, specifically, seed-based region growing, was chosen as the most suitable.

The method begins with a seed-based region growing algorithm starting from the current GNSS position, shown as "Field contour detection" in fig. 4.8. A dynamically adjusted gray-level threshold is computed using the standard deviation of pixel values. To refine the results, morphological operations (erosion and dilation) are applied to counteract leakage and noise [97].

Obstacle detection follows and is confined to the previously segmented field. A Wiener filter is applied to the image, followed by Otsu binarization and morphological closing to improve POI shape detection [98], [99].

Post-processing:
In the final step of the process, all generated binary masks undergo a quality-based selection and are then merged into a single, robust result. Initially, masks in which the seed point lies outside the detected field are discarded, as are those with implausibly small or large areas. This ensures that only spatially consistent and representative detections are considered.

Further filtering is based on the Otsu threshold difference, which serves as an indicator of segmentation reliability. Masks derived from low-contrast images are excluded, as they tend to produce inaccurate obstacle detections.

The selected masks are then fused by computing a pixel-wise mean across all inputs, with those from the current year receiving higher weight to reflect possible changes in field layout or crop type.

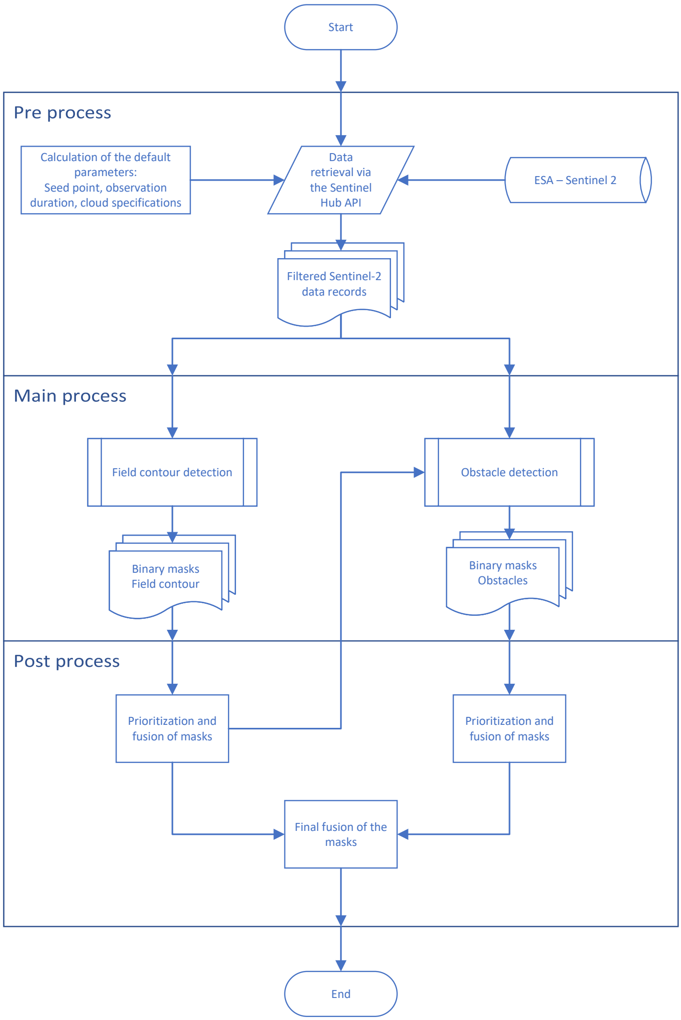


Figure 4.8.: Flow of the overall algorithm [4]

A fixed threshold is applied to this grayscale composite to generate the final binary masks for field contours and POIs. This selective fusion across multiple temporal instances enhances robustness and reduces the influence of outliers and noise in individual detections.

4.2.2.2. Results and discussion

The method was tested on 10 fields in northern Germany. For each field, 2 475 parameter combinations were evaluated. In order to minimize the relevance of the seed point choice, five seed points were selected for the evaluation. Care was taken to ensure that these were not positioned at the immediate edge of the field, since the detection rate is significantly reduced in this region due to an inhomogeneous pixel environment. The Jaccard Index (JI) values of the respective seed points were subsequently averaged for each field. Using the JI as the performance metric, the best median JI achieved across all fields was 0.913, with optimal parameters $\sigma_{RG} = 0.24$ and $PT_t = 0.4$ and five selected seed points. An example of an analyzed field is shown in Figure 4.9 [4].

Detection quality varied across fields. For example, field 2 was under-segmented, while fields 7 and 10 experienced leakage. A detailed table in the original paper presents metrics such as precision, recall, and JI per field [4].

The seed point location was found to be the most influential factor for detection quality. To improve robustness, six additional seed points were introduced on two concentric circular paths around the GNSS location, spaced at 120° intervals, as shown in fig. 4.10. To investigate the effect of this approach, each pixel of the source image was selected as a seed point. This calculation involved considering between 788 and 6 060 seed points per field, corresponding to the number of pixels within the ground area of the analyzed fields. Subsequently, the JI values of all seed points were averaged for each field. This "multiple seed point" strategy improved the median JI from 0.710 (single SP) to 0.837 (multiple SP). The comparison between the "multiple seed point" and "single seed point" strategies is illustrated in fig. 4.11. Varying the weight of seed points had minimal effect, but optimizing the radii of circular paths ($r_{R1} = 2$ px, $r_{R2} = 7$ px) significantly impacted performance.

Compared to machine learning-based approaches like U-Net ($JI = 0.801$) [88] and edge detection algorithms presented from Watkins and Niekerk ($OA = 0.929$) [89], the proposed method performs similarly or better ($JI = 0.913$). It

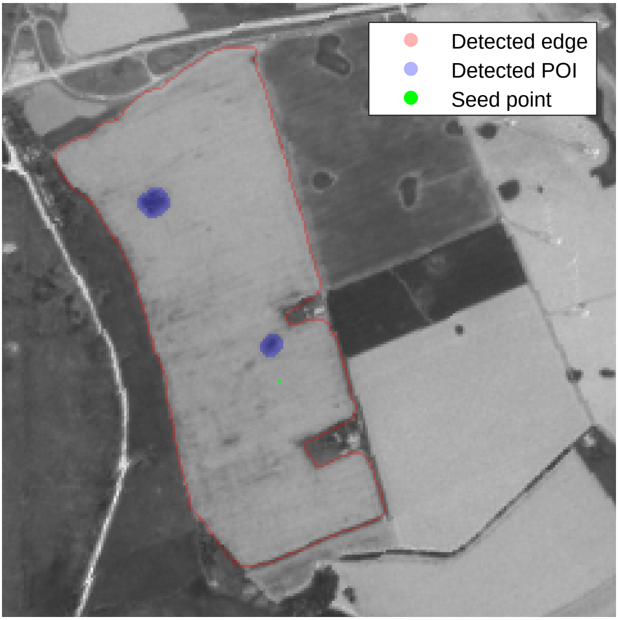


Figure 4.9.: Detected field (Modified Copernicus Sentinel data 2022/Sentinel Hub) [4]

eliminates the need for labeled training data and can be embedded directly into machine-based systems using GNSS data in real time.

Conclusion:

The presented method enables reliable detection of field boundaries and obstacles using only freely available Sentinel-2 satellite data in combination with GNSS positioning. Without relying on onboard sensor systems or training data, it offers a low-cost and scalable approach suitable for a range of agricultural machines, particularly combine harvesters.

Detection quality is primarily influenced by the satellite data used, the choice of seed points, and parameter settings. Especially the integration of multiple seed points significantly improves robustness. While some challenging field scenarios revealed limitations, such as leakage or under-segmentation, the overall results are competitive with established, data-driven methods.

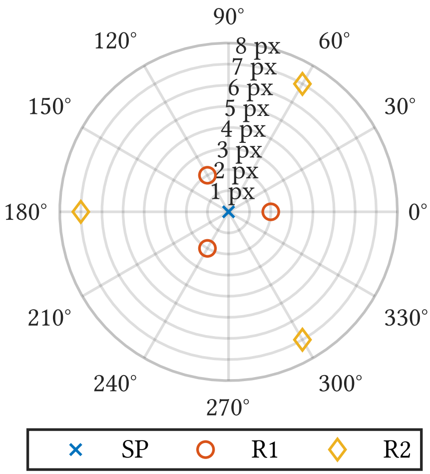


Figure 4.10.: Seed point arrangement with multiple seed points [4]

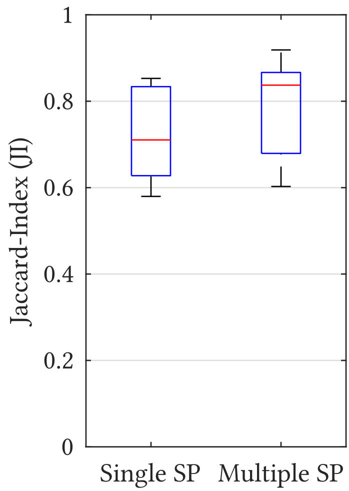


Figure 4.11.: Comparison seed point method per field [4]

Rather than replacing real-time safety systems, this approach complements them by providing spatial context and early awareness of critical zones such as field edges or obstacle locations.

5. Monitoring of mental strain

This chapter provides a detailed examination of the detection of the current mental strain state of a combine harvester operator. Since the overarching goal of the developed assistance system, as introduced in chapter 1, is the regulation on the mental strain state, this state must first be correctly monitored. For this purpose, the methods described in section 2.2 are further developed and adapted to the specific application of a combine harvester. Ocular and physiological measurement systems in combination with machine learning methods have proven to be reliable and promising approaches for mental strain detection. This approach is also pursued in this chapter.

5.1. A Demonstrator cabin as validation platform

At the beginning of developing a suitable system for detecting the mental strain state of combine harvester operators, an environment must first be created in which mental strain states can be induced in a reproducible manner. Funk et al. (2021) presented an experimental setup that enables the validated and targeted induction of mental strain states in test subjects [69], [70]. Building on this experimental environment and incorporating the experiences and results of these studies, a new comprehensive system was developed in the form of a modified combine harvester cabin, hereinafter referred to as the demonstrator cabin. This system allows the testing and validation of various sensor systems for measuring mental strain. In addition, it enables the creation of an immersive simulation environment for combine harvester operators, which serves as a test and validation platform for all subsystems of the assistance system presented in this work. Figure 5.1 illustrates the system in use during the study described in section 5.2.

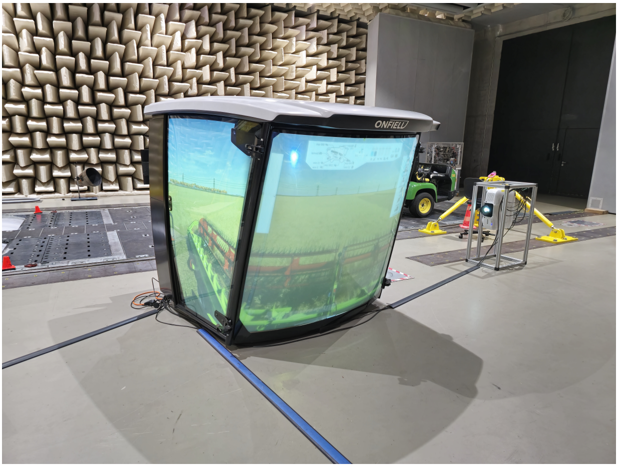


Figure 5.1.: Front view of the demonstrator cabin

Cabin design and seat concept:
The central element of the cabin is the centrally positioned driver's seat, which serves as the workstation for the participants. It is rotatable and automatically adjusts its orientation depending on the selected mode to meet ergonomic requirements and ensure optimal visibility conditions [6], [7].

The three available seating modes in the demonstrator cabin are:

- **Harvesting mode:** The seat faces forward, allowing a direct view of the harvesting process as well as the machine's working environment.
- **Office mode:** The seat is rotated to the right to provide access to the integrated office workspace in the cabin, for example for documentation and administrative tasks.
- **Relaxation mode:** The seat is rotated to the left, providing more legroom and a more relaxed sitting posture, particularly during highly automated operating phases.

In the research project "Fahrerkabine 4.0", these modes are used to investigate how different work contexts can be ergonomically and functionally integrated into a common cabin environment. In the study described here, these modes cannot be used because the eye-tracking system requires the operator to look at the front projection. This is due to the setup with four cameras. With more cameras, a larger field of view would be possible. To avoid calibration issues, the driver's seat remains fixed in the harvesting position during all experimental phases. [7]

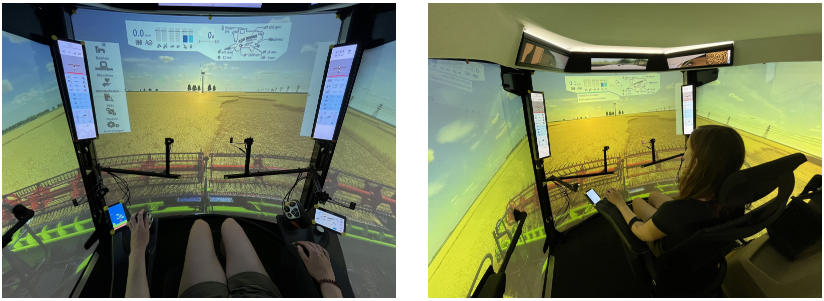
Control Elements and Display Units:
The machine is operated via two ergonomically shaped Multifunction Armrests (MFAs), which serve as the primary control units. The left MFA contains a joystick for steering as well as switches and buttons relevant to driving. The right MFA features a multifunction joystick for controlling machine functions and a touch display for adjusting machine parameters [6]. Using these control elements, steering, speed, and the settings of the reel and cutting unit can be managed. Three displays are integrated into the cabin ceiling, projecting rear-view images to provide participants with visual feedback on the area behind the machine. Additionally, two more displays are installed in the cabin's A-pillars. These show content in a tile format, including a navigation display specifically developed for the driving task. Further information, such as machine status or the participants' heart rate, is also displayed. [6]

For interaction with system settings and the Virtual Assistant (VA), participants have access to two additional touch displays integrated into the armrests. Figure 5.2a shows the arrangement and design of the MFAs. Detailed information on interaction with the VA is provided in section 6.3. [6]

During operation, the cabin is embedded within a virtual environment. The simulation is projected onto the front and side windows, which are equipped with a polymer-dispersed liquid crystal (PDLC) layer that can switch between transparent and opaque states as needed [100].

The projections include a virtual field environment, head-up displays with context-related information, text outputs from the VA, and task recommendations in the form of videos, websites, or images. These projections, illustrated in fig. 5.2b, create an immersive working environment where real control elements are combined with virtual information displays [6].

Technical interfaces and data processing:
The virtual cabin environment is based on a modified version of the "Farming Simulator 22" (GIANTS Software GmbH). The predefined test field includes a



(a) View of the MFAs

(b) Interior setup of the demonstrator cabin with a study participant [7]

Figure 5.2.: Interior view of the demonstrator cabin

driving lane that participants are instructed to follow as precisely as possible during manual operation.

To supply the cabin with the necessary hardware infrastructure, a control cabinet is installed at the rear of the cabin as a central hardware backbone. This ensures proper power supply and connectivity of all technical components. [7]

Control and operational commands are transmitted via a CAN bus interface as well as analog and digital signals to an I/O board with a microcontroller. This board converts incoming signals into virtual joystick and button commands for the simulation software [6]. Through targeted modifications, machine parameters, environmental data, as well as position and speed information can be read and stored. The collected data is further processed in a Matlab/Simulink model, which, among other functions, controls the adjustment of the threshing systems in real time. For increased realism, a GNSS control system and a dedicated machine simulation are additionally employed. This includes feedback on system adjustments and a simulation of harvest losses, allowing users to immediately observe and understand the effects of their settings on the threshing process. [7]

An overview of the installed technical components required for data processing in the demonstrator cabin is shown in fig. 5.3.

The modular structure of the validation environment allows both individual subsystems and the overall system to be tested flexibly. External data sources,

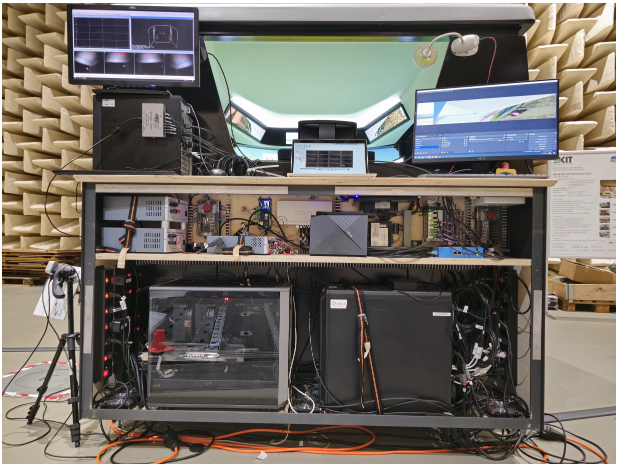


Figure 5.3.: Rear view of the demonstrator cabin

such as weather or farm management information, can be simulated as well as scenarios for critical operational situations. In this way, the interaction with the VA and the effect of the adaptive human-machine interface can be examined under conditions that closely resemble real-world operation [6].

5.2. Applied methods for mental strain monitoring

The following chapters describe the methods used for developing a model to monitor mental strain. The focus is initially on data acquisition and processing within the framework of a participant study, followed by the training of a machine learning model (section 5.2.2). The training results are subsequently validated and analyzed in section 5.2.3 and section 5.2.4.

5.2.1. Introduction

The assistance system presented in this work aims to improve the cognitive performance as well as the mental and physical health of combine harvester operators. This improvement is intended to be achieved through targeted influence or regulation of the operator's mental strain via task recommendations. In order to propose task recommendations that are precisely tailored to the user's cognitive state, it is first necessary to monitor the user's mental strain. More detailed explanations can be found in chapter 1 and chapter 3. Further information on the ergonomics and technical foundations is provided in section 2.2. This mental strain monitoring system should be minimally invasive and comfortable for the user, while also being precise. Additionally, it should be possible to monitor the current Mental strain state (MSS) in real time. Since Funk et al. have already successfully demonstrated that mental strain of operators can be measured using a combination of ocular and cardiovascular signals, the following work builds upon this [68]. The technical requirements at the interfaces for continuous monitoring of mental strain are examined in more detail in section 6.4.

From this starting point, the second research question of chapter 3 is addressed: "How can the mental strain state of an agricultural machine operator be monitored or classified?"

The following sub-questions are considered for a comprehensive answer to the research question:

1. **Sensor systems:** *"Is a combination of ocular and cardiovascular signals suitable for mental strain monitoring?"*
2. **Classification algorithms:** *"Which classification algorithms are suitable?"*
3. **Detection range:** *"Which mental strain states can be detected most and least effectively?"*

By answering these questions, the accuracy and applicability of the model to monitor mental strain for the assistance system presented in this work can be demonstrated. Potential optimization opportunities for future applications can also be identified and discussed.

5.2.2. Method

The participant study for mental strain monitoring, hereafter referred to as the mental strain study, was conducted as a quantitative investigation in the demonstrator cabin presented in section 5.1. The demonstrator cabin served as an immersive test environment to provide each study participant with an environment that was as identical as possible.

5.2.2.1. Participants

A total of $N = 40$ participants took part in the mental strain study. Four participants could not be considered due to technical issues. The age of the remaining $N = 36$ participants ranged from 20 to 37 years ($M = 25.89$; $SD = 3.70$). The sample consisted of 31 male and 9 female participants. Recruitment was conducted among students and staff of the Karlsruhe Institute of Technology (KIT). All participants had no prior experience with operating combine harvesters. Before the study began, participants were thoroughly informed about the procedure, potential risks, and their rights. Additionally, they were assured of the pseudonymity of their data. Following this briefing, all participants gave their informed consent to participate. Furthermore, it was ensured that all participants either had unrestricted vision or used appropriate visual aids during the study.

5.2.2.2. Materials

The demonstrator cabin, described in more detail in section 5.1, served as the main experimental environment. Additional, several modifications were made to meet the requirements of the mental strain study and to address the associated research questions. The modifications primarily concern the sensor systems for monitoring the mental strain. The sensor system includes an eye-tracking system and a heart rate sensor in the form of a fitness tracker.

The Smarteye system, consisting of four cameras and three infrared LEDs, was used as the eye-tracking system [10]. To ensure correct positioning of the cameras and LEDs, several aluminum profiles were magnetically attached to the A-pillars of the cabin. Subsequently, a 3D world with a world coordinate system (WCS) was measured and created using the Laser Chessboard Tool (LCT). This 3D world served as the system's virtual experimental environment. The camera positions and gaze parameters of each participant were calibrated and verified before each study run. Figure 5.2b in section 5.1 shows the exact camera positions as well as the general interior layout of the demonstrator cabin.

Another key system for monitoring mental strain is the fitness tracker used. A “TICKR FIT” fitness tracker from Wahoo Fitness, LLC was employed [11]. The advantage of this fitness tracker is that it does not rely on a proprietary protocol, but instead publishes its raw data in real time via Bluetooth Low Energy (BLE). Further details are provided in section 6.4. Participants wore the fitness tracker on their forearm for the entire duration of the experiment.

5.2.2.3. Experimental design

In the study, different scenario drives were defined, which varied in terms of driving style, pause duration following a task recommendation of the type machine settings (TR_{MS}), and additional mental stress because of arithmetic tasks. These scenario drives were designed in a preliminary study to induce three different mental strain states (“0–37”, “38–71”, and “72–150”).

Table 5.1 describes these different scenario drives.

Table 5.1.: Overview of scenario drives

Scenario Drive	Pause after TR_{MS} in s	Arithmetic Tasks
Autonomous Drive	-	no
Manual Drive	-	no
TR_{MS10}	10	no
TR_{MS5}	5	no
TR_{MS2C}	2	yes
TR_{MS0}	0	no

In four of the six scenario drives, participants had to perform additional machine settings tasks alongside the harvester steering task. These tasks were to be completed as quickly as possible. An example of a task recommendation of the type machine settings (TR_{MS}) would be: “Change threshing drum speed to 950 $\frac{1}{min}$”. This request was presented both visually on the windshield and audibly via the VA of the demonstrator cabin. The task was executed using the display in the right armrest of the demonstrator cabin (fig. 5.4). In scenarios with TR_{MS}, the pause durations varied systematically. In the TR_{MS10} scenario, the pause after a TR_{MS} was 10 s, while in the TR_{MS5} scenario it was shortened to 5 s. In both cases, no additional arithmetic tasks were included. In the TR_{MS2C} scenario, the pause after a TR_{MS} was reduced to

2 s, and additional arithmetic tasks were presented, creating an additional mental strain for the participant. These arithmetic tasks were completed on the touch display in the left armrest. Figure 5.5 shows this interface. Finally, in the TR_{MS0} scenario, the pause after a TR_{MS} was eliminated entirely (0 s), with no additional arithmetic tasks.

Figure 5.4.: Machine silhouette in the right armrest

In autonomous driving, the vehicle controlled itself, and the driver did not perform any manual interventions. In manual driving, the vehicle was controlled by the driver alone. In both scenario drives, no machine settings tasks were required.

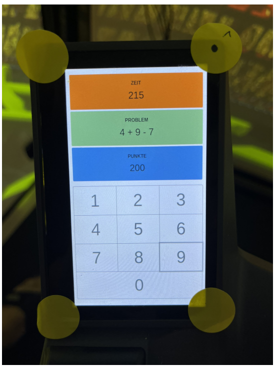


Figure 5.5.: Arithmetic task on the left touch display

5.2.2.4. Experimental procedure

A single study session consisted of three different phases. The first phase was an introduction phase, during which the participant was familiarized with the operation and procedure.

This was followed by a practice phase. This practice phase consisted of two levels, with the first being scenario drive “Manual Drive” and the second being TR_{MS5}. After each level, the current mental strain was assessed using the Rating Scale Mental Effort (RSME) [51]. Since the RSME captures only a single value, it is significantly simpler and faster to apply than, for example, the NASA-TLX. For this reason, the RSME is used as a self-report measurement. The interface shown in fig. 5.6 was used for this, and participants had 20 seconds to adjust the slider and rate their current MSS. The practice phase lasted 6 minutes, allowing participants to get used to the experimental environment.

The third phase consisted of the study drive. The study drive comprised a total of 17 levels and lasted 41 minutes and 40 seconds. Each level consisted of one of six different scenario drives and lasted 120 seconds. After each level, mental strain was assessed using the RSME, as in the practice phase.

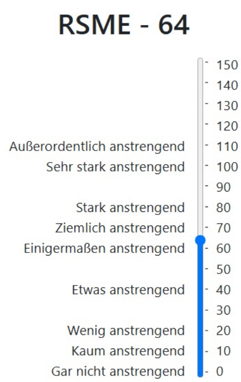


Figure 5.6.: RSME request interface

5.2.2.5. Data preprocessing

The raw data from the Smarteye eye-tracking system are transmitted directly to the server of the assistance system via the User Datagram Protocol (UDP) network protocol. These packets are converted into individual raw data IDs according to the Smarteye specification [101]. The data from the “TICKR FIT” fitness tracker are transmitted via BLE and also converted into heart rate data according to the respective specification. Matlab® R2023b was used for this purpose.

The following describes the data preprocessing based on the used predictors. All predictors refer to the past 120 seconds of the respective level. This means that each level has one value for each predictor.

Predictors:
From the two subsystems of mental strain monitoring, a total of 22 potential predictors arise that are suitable for training a machine-learning model. Table 5.2 presents this selection along with the corresponding unit.

A review of related scientific studies guided the selection of predictors used in this study. For example, Yousefi et al. focus primarily on pupil-diameter, fixation-based, and electrodermal-activity predictors. However, electrodermal-activity predictors were excluded from the presented system because they are too intrusive for users [102]. A study conducted of Kaczorowska et al. examines a broader range of predictors, including fixation-, saccade-, blink-, and pupil-based features [66]. Based on these insights and the results of Funk et al. [68], the predictors listed in Table 5.2 reflect all features considered relevant and well-supported by previous research. Heart rate was also included as a predictor, as it has been shown to be strongly associated with mental strain [67].

The predictors NNI, LFHF, ICA, and IPA require extensive preprocessing of the raw data and are explained in more detail below. All other predictors are calculated using arithmetic means, standard deviations and summations.

- **NNI:** The Nearest Neighbor Index (NNI) is a measure for quantifying the spatial distribution of points and is used in this work to analyze the arrangement of fixation points in a given 3D environment. Various studies show the relationship between NNI and mental strain, for example [103]. The calculation is based on the fixation points of the

Table 5.2.: Overview of the considered predictors and their units

Predictor	Unit
Gaze Direction Quality mean	-
Pupil Diameter Quality mean	-
Saccade Duration mean	s
Saccade Duration sd	s
Saccade Duration acc	s
Saccade Rate	s
Saccade Velocity mean	s
Saccade Velocity sd	s
Fixation Duration mean	s
Fixation Duration sd	s
Fixation Duration acc	s
Fixation Rate	-
NNI	-
Blink Duration mean	s
Blink Duration sd	s
Blink Duration acc PERCLOS	s
Blink Rate	-
LFHF	-
ICA	Hz
IPA	Hz
Heart Rate mean	Hz
Heart Rate sd	Hz

past 120 seconds. For each fixation, the x-, y-, and z-coordinates averaged over the fixation duration are determined. The resulting fixation points are modeled as a point cloud to calculate the distances to their respective nearest neighbors. The average Distance to the Nearest Neighbor (d(NN)) is then divided by Average Random Distance (d(ran)) to determine the NNI. Equation (5.1) shows the corresponding calculation in three-dimensional space. An $NNI < 1$ indicates a clustered distribution, $NNI \approx 1$ a random distribution, and $NNI > 1$ a regular arrangement of points [103].

$$NNI = \frac{d(NN)}{d(ran)} = \frac{d(NN)}{0.55396\sqrt[3]{\frac{V}{N_{NN}}}} \tag{5.1}$$

- **LFHF:** The ratio $\frac{LF}{HF}$ of pupil diameter is a promising method for assessing mental strain that is independent of lighting conditions. While pupil diameter is widely used as a non-intrusive measure to monitor cognitive state in complex environments, its application is significantly limited by the pupillary light reflex. However, spectral analysis of the pupil diameter using the Welch method allows the signal to be decomposed into frequency-specific components, minimizing the influence of lighting conditions. The power spectral density of the pupil signal is calculated and then divided into the low-frequency range (LF: 0 - 1.6 Hz) and the high-frequency range (HF: 1.6 − 4 Hz). The mean power of these frequency bands is determined, and their ratio $\frac{LF}{HF}$ is computed. Previous studies have shown that this ratio is sensitive to mental strain but is not affected by varying lighting conditions. Therefore, the $\frac{LF}{HF}$ ratio provides a robust and potentially real-time metric for assessing mental strain. [104]
- **ICA:** The Index of Cognitive Activity (ICA) is based on changes in pupil dilation during interaction with visual stimuli. To calculate the ICA, the pupil signal is decomposed into its detail coefficients using a wavelet transform to separate rapid, cognitively induced pupil fluctuations from slower physiological or light-induced changes. A denoising method (Minmax-Threshold) is then applied to reduce noise and highlight relevant pupil changes. The number of detected peaks in the filtered signals is normalized over the level duration (120 s) to determine the ICA value. Since the ICA is less affected by external lighting conditions than absolute pupil sizes, it provides a robust method for real-time monitoring of mental strain in dynamic environments [105], [106].
- **IPA:** The Index of Pupillary Activity (IPA) is a novel metric for assessing mental strain based on the frequency of pupil diameter oscillations. Unlike existing proprietary methods such as the ICA, the IPA offers a transparent and reproducible calculation method based on wavelet

analysis and modulus-maxima detection. To calculate the IPA, the pupil signal is first decomposed into approximation and detail coefficients using a Symlets16 wavelet transform. The modulus maxima of the detail coefficients are then identified to separate rapid, cognitively induced pupil fluctuations from slower physiological changes. These maxima are filtered using a universal threshold method ($\lambda_{univ} = \hat{\sigma} \cdot \sqrt{2 \log n}$, where $\hat{\sigma}$ is the standard deviation of the noise) to minimize unwanted signal noise. The final IPA value is determined from the number of remaining significant oscillations, normalized to the measurement duration. Initial studies indicate that the IPA can detect differences in mental strain depending on task complexity, making it an open alternative to existing methods [107].

The selection of the predictors examined in this work is primarily based on their prevalence and relevance in recent scientific publications. In total, 22 predictors were considered that are frequently used in research related to the measurement of mental strain using eye-tracking systems [31], [68], [85], [104], [107]–[113].

To further reduce the number of predictors and improve model performance, a feature selection method was applied. In this context, the χ^2 test is particularly suitable because it quantifies the statistical significance of the relationship between a predictor (feature) and the target variable. This makes it possible to distinguish between true correlation and random noise [114].

The Matlab® function `fscchi2` was used to perform the χ^2 test. This function examines each individual feature for independence from the target variable and calculates individual χ^2 tests. To assess predictor relevance, the value $-\log p$ is used, where p represents the p-value of the test. A small p-value indicates that the corresponding feature has a significant dependence on the target variable and can therefore be considered particularly important for model construction. The predictors are subsequently ranked according to their $-\log p$ values. Based on the determined predictor ranking, the number of features used is empirically varied during the subsequent training process. The objective is to identify the predictor configuration that provides the best generalization performance of the model.

5.2.2.6. Classification methods

This section describes and compares the individual classification algorithms used in this study. Subsequently, the training and validation methods are explained in more detail.

- **Support Vector Machine (SVM):** SVMs are among the most powerful and widely used methods in the field of supervised machine learning, particularly for classification tasks. Their fundamental principle is to find an optimal decision boundary, the so-called hyperplane, that separates the training data points of different classes with a maximum margin. SVMs are particularly suitable for high-dimensional problems and are characterized by their ability to model both linear and nonlinear decision boundaries through the use of kernel functions [115], [116].

  In the context of multi-class classification problems, when more than two classes must be distinguished, the originally binary SVM is extended by strategies such as One-vs-One or One-vs-All. MATLAB provides a powerful implementation for such cases through the function `fitcecoc` (Error-Correcting Output Codes). The `fitcecoc` procedure combines several binary SVM classifiers to create a multi-class classification model. For hyperparameter optimization, the parameters "BoxConstraint", "KernelScale", and "Standardize" are considered.

- **Decision Tree:** Decision trees are among the most intuitive and widely used methods in supervised machine learning. They are based on a recursive partitioning principle, where the input features of the training data are successively selected to split the data into increasingly homogeneous subsets. Each internal node of the tree represents a decision rule based on a feature, while the leaves correspond to class membership. Decision trees are particularly advantageous for nonlinear decision problems and offer high interpretability of the resulting model [117]–[119].

  For classification tasks involving more than two classes, decision trees can also be generalized using the Error-Correcting Output Codes (ECOC) approach. Similar to the SVM implementation in MATLAB, this is achieved through the `fitcecoc` function, which combines multiple binary classifiers to produce a robust multi-class model. This combination benefits from the inherent robustness and flexibility of individual trees. To optimize the decision tree, the hyperparameter "MinLeafSize" can be adjusted to achieve a balanced trade-off between model complexity and generalization ability.

- **Naive Bayes:** Naive Bayes is a simple yet often powerful classification method based on Bayes' theorem, which assumes strong independence between input features. The model calculates, for each class, the posterior probability based on the likelihood of the features and the prior probabilities of the classes. The Naive Bayes classifier often achieves remarkable results, particularly for high-dimensional or text-based data [120].

  In MATLAB, Naive Bayes is implemented using the function `fitcnb`. This function also allows targeted optimization of the hyperparameters "DistributionNames", "Standardize", and "Width".

A uniform training procedure was applied to all classification algorithms. The dataset, consisting of recordings from 36 participants, was divided into five equally sized subsets using K-fold cross-validation with $K = 5$. The chosen value $K = 5$ is a well-established standard in the literature and provides a balanced compromise between training efficiency and generalizability [121].

As part of the cross-validation, hyperparameter optimization was conducted. In each optimization iteration, the model was trained five times on four of the five subsets and subsequently evaluated on the remaining subset. The roles of the training and test data were systematically rotated across the five folds.

The actual optimization of the hyperparameters was carried out using Bayesian optimization, a probabilistic approach for optimizing expensive objective functions. This procedure was performed for a total of 60 iterations. The goal was to identify the combination of hyperparameters that yielded the lowest value of the Bayesian objective function. Bayesian optimization has been established in the literature as an effective and robust method for automated hyperparameter tuning [122].

5.2.3. Results

The models were trained with each of the considered classification algorithms for both a three-class MSS (3C) and a ten-class MSS (10C) target variable structure. The basis for this was the RSME values reported by the participants, which were divided into the corresponding classes.

To ensure comparability among all trained models, the posterior probabilities of the ten MSS classes (10C) were aggregated according to the three main MSS classes. The division of the ten classes follows the literature ([30], [51]) and is defined as follows: “0–2”, “3–13”, “14–25”, “26–37”, “38–57”, “58–71”, “72–85”, “86–102”, “103–112”, and “113–150”. These MSS classes were then combined into three higher-level categories: “0–37”, “38–71”, and “72–150”.

Furthermore, the number of predictors used was systematically varied based on the previously determined χ^2 ranking (see section 5.2.2.5), ranging from 1 to 22 predictors. The notation “19P” refers, for example, to a model trained with the 19 highest-ranked predictors according to the χ^2 test.

Table 5.3 shows a comparison of the best-performing models for each of the investigated classification methods. It should be noted that this represents a multiclass classification, in which each instance is uniquely assigned to one class. The reported means and corresponding standard deviations are based on the $K = 5$ cross-validation described in section 5.2.2. As evaluation metrics, the most commonly used measures in the literature were applied: “Recall” (also known as True Positive Rate (TPR)), “Precision” (also referred to as Positive Predictive Value (PPV)), and the well established F_1-score. Since the dataset is assumed to be imbalanced, macro-averages were calculated to aggregate the metrics across all classes. This ensures that all classes contribute equally to the overall performance, regardless of their frequency. The results of the ten-class models (10C) were determined by aggregating the classes into the same classes like the three-class models (3C).

With respect to Recall, the Naive Bayes model with three classes and 20 highest-ranked predictors achieved the highest mean value (0.5962), closely followed by the Decision Tree model with ten classes and 19 highest-ranked predictors (0.5956). The latter also exhibited the highest Precision (0.5979) and the best F_1-score (0.5949). It is noteworthy that the SVM model with three classes and the 21 highest-ranked predictors showed the lowest variance across all metrics (0.0146 for Recall, 0.0295 for Precision, and 0.0262 for the F_1-score), although its mean F_1-score was comparatively lower (0.5504). The Decision Tree models overall showed higher variability across cross-validation folds, which is reflected in the relatively large standard deviations.

Table 5.3.: Comparison of classification algorithms

Classifier	Recall		Precision		F_1-score	
	Mean	SD	Mean	SD	Mean	SD
SVM (3C[1], 21P[2])	0.5752	**0.0146**	0.5820	**0.0295**	0.5504	**0.0262**
SVM (10C, 13P)	0.5622	0.0359	0.5596	0.0450	0.5375	0.0535
Decision Tree (3C, 11P)	0.5618	0.0466	0.5696	0.0524	0.5638	0.0479
Decision Tree (10C, 19P)	0.5956	0.0647	**0.5979**	0.0670	**0.5949**	0.0661
Naive Bayes (3C, 20P)	**0.5962**	0.0341	0.5854	0.0309	0.5800	0.0288
Naive Bayes (10C, 14P)	0.5700	0.0296	0.5627	0.0366	0.5640	0.0335

[1] Classes
[2] Predictors

Since the Decision Tree model (10C, 19P) achieved the highest F_1-score (0.5949) among all tested configurations, a detailed analysis of its characteristics and results is presented below.

Figure 5.7 shows the confusion matrix of the model for the ten-class (10C) classification. It becomes evident that the recognition rate strongly depends on the specific class. In particular, the edge classes “0–2” and “113–150” show comparatively poor recognition rates. In contrast, middle classes such as “3–13” ($Recall$ = 0.573), “14–25” ($Recall$ = 0.388), and “86–102” ($Recall$ = 0.321) are recognized more accurately. Nevertheless, the Precision for all predicted classes remains below 0.4, indicating a relatively high number of false-positive predictions.

A clear improvement in prediction quality can be observed when the ten original classes are aggregated into three main MSS classes (3C). Figure 5.8 shows the corresponding confusion matrix. Through this reduction, the Recall for the class “0–37” increases to 0.811. The middle class “38–71”, however, exhibits the lowest Recall value (0.422), which may be due to overlaps with neighboring classes. The Precision is distributed more evenly across the classes and ranges between 0.494 and 0.724, indicating an overall improvement in the reliability of class predictions.

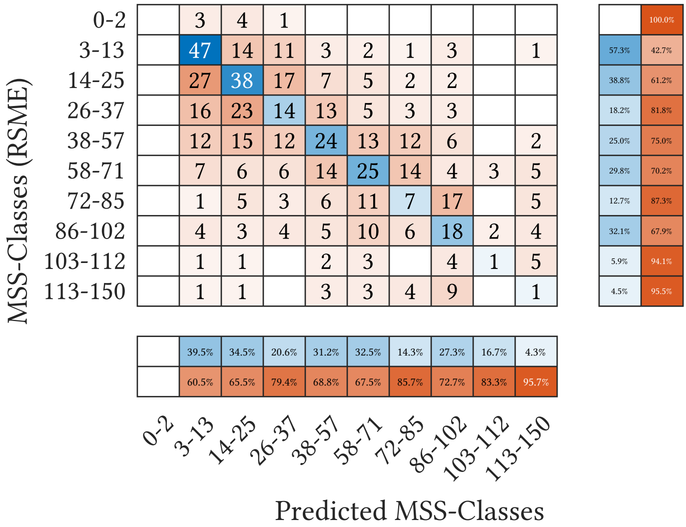


Figure 5.7.: Confusion chart of Model Decision Tree (10C, 19P)

5.2.4. Discussion

To address the central research question of this study, *"How can the mental strain state of an agricultural machine operator be monitored or classified?"* (see section 5.2.1), the corresponding sub-research questions are discussed sequentially below.

Sensor systems:
The first sub-research question, *"Is a combination of ocular and cardiovascular signals suitable for mental strain monitoring?"*, deals with the selection of appropriate sensor systems for measuring mental strain. Section 5.2.2 describes the sensors used in this study. A comparison with current literature ([66], [102]) particularly highlights the strong potential of eye-tracking and heart rate monitoring systems, thus supporting the sensor choice in this work. The achieved classification performance with an F_1-score of 0.5949 is also within the range reported by other studies in different contexts using similar measurement devices. Hillege et al. used remote photoplethysmography (rPPG) to predict three levels of mental strain based on heart rate, achieving

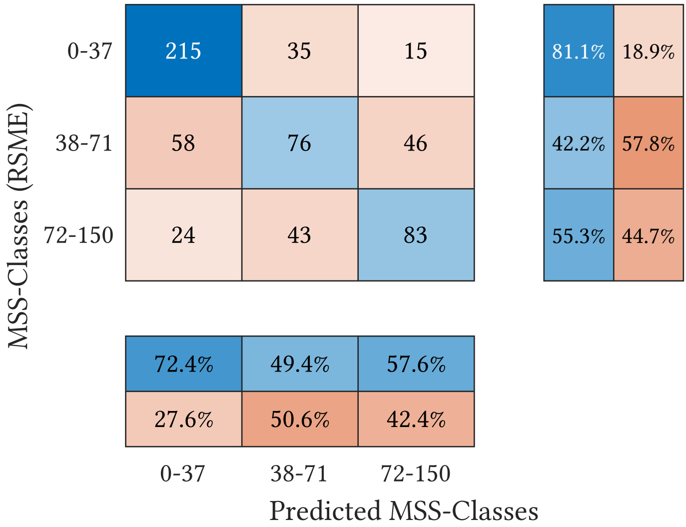


Figure 5.8.: Confusion chart of Model Decision Tree (10C, 19P) reduced to three Classes

an accuracy of 60% [123]. Wu et al. were able to achieve a prediction accuracy of 84% using an eye tracker. However, the study only distinguished between two levels of mental strain [124]. By combining eye-tracking parameters, electrodermal activity, and heart rate, Luong et al. achieved a prediction accuracy of approximately 65% [125]. A strong correlation among various cardiovascular signals was also demonstrated by Ettema, Zielhuis, and Solhjoo et al. [67], [126]. Therefore, the first sub-research question can be answered positively, as the presented results demonstrate that a combination of ocular and cardiovascular signals is suitable for monitoring mental strain.

Classification algorithms:
To answer the second sub-research question, *"Which classification algorithms are suitable?"*, a comparison with relevant research literature is considered. Laine et al. applied an artificial neural network to predict the mental strain of aviation personnel using an extensive sensor setup [127]. Chen et al. employed kernel-based classifiers such as SVM and Extreme Learning Machine (ELM) for stress detection during driving tasks [128]. Atasoy and Yildirim used K-Nearest-Neighbor (kNN) and Decision Trees to classify cognitive

demands in various tasks [129]. Naive Bayes has also been successfully used in the context of human-computer interaction [130]. Based on these findings, three widely used classification methods (SVM, Decision Trees, and Naive Bayes) were investigated in this study. Table 5.3 shows that the performance of all tested algorithms is comparable. The Decision Tree model (10C, 19P) achieved the best performance with an F_1-score of 0.5949.

Detection range:
The third sub-research question, *"Which mental strain states can be detected most and least effectively?"*, examines the MSS classification accuracy as a function of the target classes. The results in fig. 5.7 and fig. 5.8 show that edge classes in the ten-class variant (10C) and the middle class in the three-class aggregation exhibit lower recognition rates. This is likely due to an imbalanced distribution of training data, where MSS classes with more samples are predicted more often and with higher accuracy than those with fewer samples. Additionally, it must be noted that the RSME values are subjective and therefore prone to errors and bias. These errors can have a stronger impact, especially in underrepresented classes. In future research, these errors can be mitigated by collecting more samples.

In summary, this study demonstrates that the mental strain state of agricultural machine operators can be monitored and classified, directly addressing the central research question *"How can the mental strain state of an agricultural machine operator be monitored or classified?"*. By combining non-invasive ocular and cardiovascular sensors with suitable machine learning classifiers, in particular the Decision Tree model, mental strain could be predicted reliably with an F_1-score of 0.5949. Although detection accuracy differed between classes, the results indicate that integrating physiological measurements with appropriate algorithms constitutes a feasible and practical approach for real-time recognition of mental strain in agricultural operators.

6. Virtual Assistance

This chapter provides a comprehensive description and explanation of the Virtual Assistant (VA) as well as all associated subsystems and modules of the presented assistance system.

6.1. Motivation

The increasing automation and digitalization of agricultural work processes leads to a continuous extension of functions in modern agricultural machinery cabins [29]. With the integration of numerous assistance systems, sensors, and control elements, the operator's cabin becomes a complex, highly networked workplace. However, this development carries the risk of growing cognitive strain for the machine operator. Parallel processes, changing work situations, and a multitude of simultaneously displayed information require continuous attention and rapid, situation-appropriate decisions. Under unfavorable conditions, such as extended working hours, the risk of incorrect decisions, delayed reactions, and safety-relevant omissions increases [2].

At the same time, the advancing automation fundamentally changes the role of the operator. Many control and regulation tasks are taken over by automated functions, so that the human increasingly shifts to an observing and monitoring position. This shift can lead to underload in certain work phases if no interaction is required for extended periods. Such monotonous monitoring work carries the risk that attention and readiness to respond decrease. In critical situations, this can result in delayed interventions or errors [30], [59].

The introduction of a virtual assistance system in this context brings both challenges and opportunities. Central challenges include the reliable interpretation of relevant data, the clear definition of areas in which the assistance system is allowed to make independent decisions, and the protection of sensitive data while ensuring safety and data privacy [131].

Simultaneously, a well-designed assistance system offers the opportunity to increase safety, reduce mental strain, and serve as an interface to networked systems to support the operator in a targeted manner [131]. An assistance system in the form of a VA in the operator cabin can support the driver in processing information and help maintain situational awareness even during passive phases, thereby ensuring performance throughout the entire work shift. By filtering, prioritizing, and presenting relevant data in an adapted form, the assistant is intended to prevent both overload and underload and to regulate attention according to demand.

In addition to adapting to cognitive resources, supporting consistently high work quality is also important. In agricultural work processes, a faulty or delayed reaction can have immediate economic and qualitative consequences, such as harvest losses, increased machine wear, or inefficient resource use. A networked and adaptive assistance system helps to recognize critical states in time and provide the operator with clear task recommendations.

Furthermore, the operator cabin is evolving from a purely controlling workplace to a supervision and management center, where the human acts as a decision-maker and coordinator of complex human-machine systems [132]. In this context, the VA serves as a link between technical system logic and human decision-making, thereby creating the foundation to fully exploit the potential of modern agricultural machines without overloading or underloading the operator.

6.2. Architecture

The advantages, challenges, and opportunities of a VA presented in Section 6.1 are now integrated as effectively as possible into the assistance system developed in this work. The VA is primarily used for communication and interaction with the operator. In the following, the technical implementations, interfaces, and modules required to realize such a system are described.

The assistance system presented here is based on a modular algorithm divided into four functional main components: Input, Selection, Run, and Output. This modular architecture allows a clear separation of responsibilities and enables both scalable further development of the system and efficient adaptation to different application environments. The modules are connected in a sequential process chain, with data always flowing from Input to Output and being influenced in the intermediate stages by decision-making and control logic.

Module “Input”:
The input module handles the initialization phase and the establishment of all connections to external systems. In this phase, all relevant data sources are connected, data streams are captured, and unified in a temporary storage. This primarily includes:

- User interactions with the VA and status messages from the Human-Machine Interface (HMI),
- User preferences from the user database,
- Machine data of the combine harvester via a CAN bus interface,
- Results from the Mental strain state (MSS) monitoring system via WebSockets according to RFC 6455 [133],
- Information from external service providers (e.g., weather and farm management systems) via standardized web APIs.

These data are prepared in temporal synchronization and then passed on to the selection module. The preprocessing serves both to reduce redundant information and to ensure a uniform data structure for the subsequent processing steps. Further details on the technical aspects and implementation of the interfaces can be found in section 6.4.

Module “Selection”:
The selection module implements the decision-making core component of the algorithm. Its goal is to determine the task recommendation that keeps the user in an optimal cognitive state. The decision-making process proceeds in several stages:

1. Checking trigger conditions:
   It is checked whether defined trigger conditions are met. These include direct inputs or commands from the user, time-based events, critical MSS or threshold exceedances of external data streams (e.g., weather changes, machine signals, incoming emails).

2. Utilization of the Cognitive Task Load (CTL) model and MSS:
   The preselected task recommendations are further filtered based on the CTL model described in section 7.3.2 and the MSS monitoring system of chapter 5. The CTL model combines collected mental stress and mental strain data into quantified and processable information for the selection process. One or more task recommendations are selected to help the operator achieve or maintain an optimal cognitive state.

3. Consideration of user preferences:
   Finally, preferences stored in the user profile are taken into account, including the history of previous executions and the task category. This ensures that the operator can carry out task recommendations that are both pleasant and tailored to the individual.

More detailed information about this selection process can be found in section 7.3.3.

Module “Run”:
The run module is responsible for initializing and executing the selected task recommendation. To avoid conflicts between ongoing and new activities, each activity has authorization parameters that define whether ongoing processes may be interrupted. Critical alerts (e.g., warnings) are always prioritized to ensure they reach the user.

The execution structure follows a polyhierarchy in the form of a Directed Acyclic Graph (DAG), a directed graph that contains no cycles and therefore allows for a clear hierarchical or linear ordering of its nodes [134]. The specific execution path of a task recommendation can be influenced by user decisions and external data streams. This adaptive process design increases user acceptance and enables situationally optimized interaction.

Module “Output”:
The output module ensures user interaction and result presentation. Here, the user interfaces and interaction possibilities are rendered. Additionally, this module executes and updates the speech synthesis and visual displays of the virtual assistant.

The output modalities are designed so that they neither overload nor underload the user, but rather maintain the optimal cognitive state determined by the selection module. Detailed information on the input and output modules of the demonstration cabin can be found in section 5.1.

6.3. Options for interaction

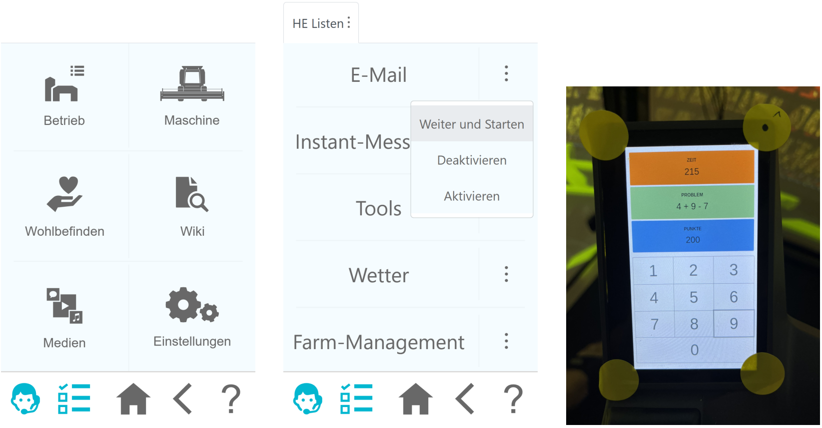
Interaction with the virtual assistance can occur through two different methods. It can be operated either via the two touch displays integrated into the Multifunction Armrest (MFA) or through voice input. The following section first describes the haptic control options in more detail.

6.3.1. Touch control

The functions of the two touch displays integrated into the demonstrator cabin are clearly separated. The display integrated into the left armrest is used exclusively for controlling comfort functions, such as multimedia, lighting, climate control, and interaction with the VA. The touch display of the right MFA, in contrast, provides adjustment options for machine and process parameters. This separation supports both functional clarity and user-friendliness.

Example representations of the user interface of the left touch display are shown in fig. 6.1. If no active interaction with the VA is taking place, a tile view of the different categories of possible Task Recommendations (TRs) is displayed, as shown in fig. 6.1a. Through these tiles, users can directly access and execute activities and actions with the VA. Furthermore, fig. 6.1b illustrates a context menu that can be displayed during an ongoing interaction. Figure 6.1c shows the display in use during a user study, in which a task recommendation in the form of mental arithmetic tasks is displayed directly.

The right touch display of the MFA allows only indirect interaction with the VA by enabling adjustments to process or machine parameters, such as the threshing drum speed. Since the virtual assistant is closely linked to the machine settings, these changes are incorporated into the input and selection modules explained in section 6.2. Figure 6.2a shows the display during a study, while fig. 6.2b exemplarily shows the adjustment window for the threshing drum speed.

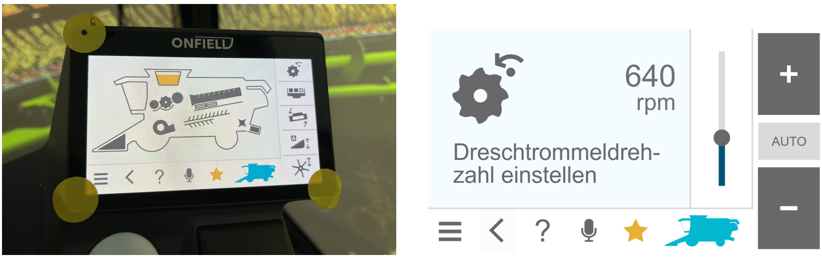


(a) Overview of the categories

(b) Selection of a TR with context menu

(c) During a study

Figure 6.1.: Left touch display



(a) During a study

(b) Adjusting the threshing drum speed

Figure 6.2.: Right touch display

6.3.2. Voice control

A central interaction method for modern assistance systems is voice control. In particular, in the context of virtual assistants, it offers the advantage that users can keep their hands free for other tasks while still issuing commands or retrieving information in natural language. This is especially important in environments with high demands on attention and motor skills, such as operating a combine harvester in the field [135].

However, implementing a robust voice interface involves several challenges. On the one hand, it must be ensured that recognition performance remains reliable under realistic operational conditions, such as high noise levels. On the other hand, aspects such as data privacy, offline availability, and the quality of speech synthesis play an important role in the acceptance and everyday usability of the system [41].

Against this background, two approaches to implementing voice control are considered below: the use of the Web Speech API [136] as a cloud-based solution and the use of local offline models. Both approaches differ in their architecture, strengths, and limitations and are analyzed with regard to their application in the agricultural assistance system.

Web Speech API:
The Web Speech API, developed within the Speech API Community Group of the W3C, provides a JavaScript interface that enables web applications to perform both speech synthesis and speech recognition [136]. The API includes two independently usable modules:

- **SpeechRecognition:**
  The SpeechRecognition interface of the Web Speech API enables the conversion of spoken language into text by using the device's microphone and sending the recorded audio data to a speech recognition engine, in this case a Google Cloud engine [136].
- **SpeechSynthesis:**
  The SpeechSynthesis interface of the Web Speech API is used to convert written text into spoken language and thus represents a central component for the speech output of the virtual assistant. It allows customization of the speech output through parameters such as language, voice, and volume, making it adaptable to different scenarios. Furthermore, it supports the use of Speech Synthesis Markup Language (SSML), which enables fine control of pauses, emphasis, or intonation. This makes it possible to conduct all communication with the virtual assistant in a bilingual manner [136], [137].

Offline models:
In the context of a combine harvester in field operations, a continuous network connection cannot always be assumed, which motivates the consideration

of an offline approach. Additionally, this approach alleviates data privacy concerns, as all recordings can be processed locally. Martin Waldschmidt, in a Bachelor's thesis supervised by Metzger [138], investigated whether it is possible to develop an offline speech recognition model as an alternative to the Web Speech API approach described in the previous section.

Waldschmidt's thesis followed a clearly structured methodological approach, consisting of a general analysis and comparison phase, as well as two targeted optimization approaches.

Initially, two open-source frameworks, Mozilla DeepSpeech [12], based on findings by Hannun et al. [139], and NVIDIA QuartzNet [140], were selected and examined in their basic configuration. The choice of these systems was based on their offline capability, open architecture, and the advantages of neural networks over classical Hidden Markov models reported in the literature. In the baseline tests, the pretrained models were evaluated using a clear dataset (Clear) recorded in a quiet room, a noisy dataset (Real) recorded on an operating tractor, and a clear dataset with added background noise (Clear-Mix). These tests served to assess the baseline performance of the systems and to identify key weaknesses, particularly the significant drop in recognition performance under realistic background noise. Table 6.1 presents these initial test results using the Word Error Rate (WER) metric [138].

Table 6.1.: Initial Comparison DeepSpeech and QuartzNet [138]

Voice-Dataset	$WER_{QuartzNet}$	$WER_{DeepSpeech}$
Real	**0.60**	0.65
Clear	**0.16**	0.28
Clear-Mix	**0.34**	0.46

Since the initial tests with up to 60% WER did not yield satisfactory results, two possible optimizations were subsequently conducted and compared.

In the first optimization approach, the language model was adapted to the operational context. Two variants were developed using KenLM [141]. The first is an application-specific language model based solely on predefined commands, while the second is an extended language model that incorporates colloquial language data to increase robustness against variable expression. The language models were integrated as external scorers in the recognition process and systematically evaluated for their impact on WER [138].

In the second optimization approach, the acoustic model was improved through targeted training. A web application was developed to collect speech data from around 30 speakers, resulting in several thousand recordings of predefined commands. To realistically replicate the acoustic conditions of a combine harvester, these recordings were supplemented with tractor and machinery noise. Subsequently, a hyperparameter search was conducted to fine-tune the DeepSpeech model on this extended dataset [138].

The results showed that background noise has a significant impact on recognition performance. While pretrained models still delivered acceptable WER in quiet environments, the error rate increased by 15–45% under realistic noise conditions. Adaptation of the language model reduced WER in quiet environments to below 10%. The decisive improvement, however, came from retraining the neural network with noisy speech data. In one of the most challenging test datasets (Real 95), WER dropped from over 50% to around 15%, an improvement of more than 35%. Significant improvements were also achieved in the other datasets, demonstrating that combining a domain-specific language model with targeted retraining significantly enhances recognition performance in noisy environments [138].

Despite these advances, the optimized solution did not yet achieve the frequently cited goal of < 10% WER in particularly difficult scenarios [13]. Waldschmidt attributes this primarily to the limited amount of real training data and the high variability of background noise in agricultural operations. A logical next step for future research is therefore the expansion of datasets through extensive field studys and the testing of more modern model architectures [138].

For the assistance system presented here, the Web Speech API is used as the speech recognition engine due to the insufficient recognition performance of the offline solution and the wired network connection of the demonstrator cabin. Undocumented initial tests and gathered experiences of the Web Speech API during the development process showed a subjectively clearly better recognition performance in comparison to the offline models.

6.4. Technical interfaces

This chapter provides a detailed description of the technical aspects and the concrete implementation of the interfaces of the VA. The focus is on the three central interfaces for MSS monitoring, the machine and external services. The VA functions as a central coordination unit in which all data streams are consolidated and processed.

To ensure uniform management of the data streams, the WebSocket protocol is used [133]. A dedicated WebSocket server is implemented for each of the essential data streams. This architecture enables parallel processing and ensures that each subsystem can simultaneously access the data relevant to its function via a secure connection. At the same time, computing resources are conserved, as subsystems only interact with the servers required for their tasks. The following section explains the technical implementation of the five most important data streams.

- **User interactions with the VA and status messages from the HMI:** A custom-developed web server is used to provide all interface displays and interaction options. This server communicates via a custom WebSocket server with the machine and the VA. In this way, current machine data and feedback from the VA can be displayed visually and audibly with minimal delay. Furthermore, this approach allows the implementation of a control center application, through which all relevant information can be provided to the study management during a study.
- **User preferences from the user database:** User preferences and settings are stored in an AES-256-encrypted, structured JSON file [142]. After login via the user interface, the respective data are read from the database and updated if changed. Communication occurs securely (TLS) via a WebSocket server. Each user has access only to their own profile.
- **Machine data from the combine harvester:** In real operation, machine data are transmitted robustly to the VA via a CAN interface. In the demonstrator cabin, however, the data are provided by the simulation environment "Farming Simulator 22" and the user interface. Details can be found in section 5.1. Table A.1 provides an overview of the machine data processed by the VA. These form the basis for all

machine-related TR. Communication also occurs via WebSockets in this case.

- **Results of MSS monitoring system:** The MSS monitoring system described in chapter 5 is based on two sensor systems: a Bluetooth Low Energy (BLE) fitness tracker ("TICKR FIT" by Wahoo Fitness) and an eye-tracking system by Smarteye. The sensors are initialized at the beginning, and their predictors are subsequently fused. The eye-tracking system provides data at a sampling rate of 60 Hz. To ensure reliable and low-latency transmission, UDP is used instead of Transmission Control Protocol (TCP) [101], [143], [144]. The fitness tracker uses the open BLE interface with a sampling rate of approximately 1 Hz. All recorded data points are stored in a dynamic list with a length of 120 s (section 5.2.2). This process is illustrated in fig. 6.3.

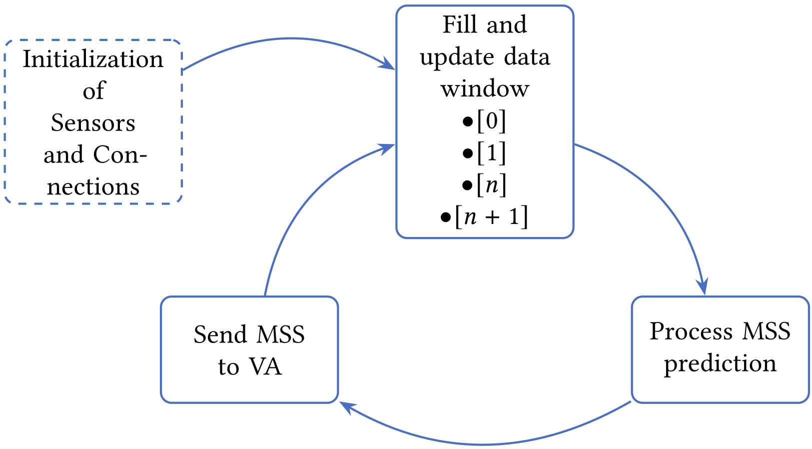


Figure 6.3.: Flowchart of the MSS monitoring system

Due to the different sampling rates, heart rate values are held until the next measurement and averaged over the time window. Once at least 115 s of the list is filled, the mental strain monitoring process starts. This time window provides tolerance for short-term failures in eye detection. After calculating the most probable MSS, the result is transmitted to the VA via a WebSocket connection [133]. The process is continuously restarted with a maximum wait time of 2 s, ensuring that the VA is regularly supplied with current condition status information. This wait time was determined empirically and represents an appropriate compromise between data volume and update rate. This approach meets real-time requirements, and the VA is informed in real time about the operator's mental strain [145].

- **Information from external service providers:** The VA is connected to various external services via public REST APIs. This extends the functionality and the potential number of different TR without significantly increasing the complexity of the assistance system [146]. Additionally, external services can enhance the user experience and increase user acceptance. The processing of these interfaces takes place in outsourced subsystems that communicate with the VA via WebSockets. Examples of implemented services include:
  - "OpenWeather" for adverse weather alerts ([14])
  - "365FarmNet" as a farm management tool ([15])
  - RSS feeds from user-selected providers ([16])
  - "SentinelHub" for the environmental monitoring described in chapter 4 ([8])
  - SMTP/IMAP for email communication ([147], [148])
  - Web radio ([17])
  - "Spotify" as a media service provider ([18])
  - "Matrix" for instant messaging ([19])
  - "DeepL" and "Pons" as translators ([20], [21])
  - "OpenThesaurus" as a dictionary ([22])
  - "Wikipedia" as an encyclopedia ([23])

  These services are either accessed regularly (e.g., weather information) or on explicit request by the user within the scope of a TR (e.g., "Spotify", web radio).

7. Regulating the operator's mental strain state

Building on the technical foundations and implementations presented in chapters 4 to 6, this chapter provides a detailed description of the logic of the Virtual Assistant (VA). The focus is particularly on decision-making processes and the selection of suitable Task Recommendations (TRs).

7.1. Pool of possible task recommendations

In order to influence and optimize the mental strain of an operator, it is necessary to interact with the operator. One possible form of interaction is with TRs. This form of interaction has the advantage over, for example, acoustic signals or manipulation of the driver's seat, that the operator gains a direct benefit from completing such a task, especially when it is business-relevant, regardless of whether the optimization of mental strain was successful or not. Additionally, the operator can immediately perceive a practical relevance, which can lead to intrinsic motivation in completing such TRs and increases user acceptance of the presented assistance system. This is also supported by various surveys and analyses conducted by Metzger et al. [1]. In this context, it is necessary to create a pool of possible activities or TRs in order to offer meaningful and, above all, strain-relevant suggestions in as many different situations as possible. Metzger et al. conducted a series of surveys and interviews with farmers and farm owners in 2019 and 2020 [1]. These identified various areas for potential activities. In particular, activities such as "checking the weather forecast" or "reviewing machine information" were frequently mentioned. Activities in the areas of communication, management, and office work were also repeatedly highlighted [1]. Based on these results, a total of seven categories were developed to classify the activities and TRs in the TR pool:

1. **Business:** This category collects all business-relevant TRs, including organization, communication, and farm information provided by an Farm Management Software (FMS).
2. **Machine:** This category includes various TRs related to the currently operated machine, primarily involving adjustments to threshing parameters and machine settings.
3. **Wellbeing:** The Wellbeing category gathers exercises aimed at improving physical wellbeing.
4. **Wiki:** This category groups TRs that can provide the operator with guidance or assistance.
5. **Media:** All TRs related to media, such as music, audiobooks, and videos, are collected in this category.
6. **Settings:** TRs and activities that can adjust the behavior of the VA are included here.
7. **Other:** All remaining TRs that do not directly relate to the other categories are grouped in this category.

To describe the characteristics of individual TRs in more detail, table 7.1 presents a small selection of TRs implemented in the assistance system. In addition to names, descriptions and categories, the authorized operator groups and the amount of linked actions are also indicated. The complete table can be found in appendix A.2. To enable the creation and maintenance of individual TRs as efficiently as possible, an object-oriented approach was chosen for programming the database. This has the advantage that all created TR objects are generated based on a blueprint, specifically a class, ensuring that all objects contain the necessary attributes and methods defined in this class. Additionally, this approach allows TRs to be created dynamically during operation by the user or the assistance system (VA) and added to the database. The following sections describe the general structure, attributes, and methods of the TR objects.

The structure of a TR object can be divided into three parts. First, the general attributes should be mentioned. These are listed below:

- ID: A unique alphanumeric string.
- Name: The name of the TR. Examples can be found in table 7.1.

Table 7.1.: Excerpt from Full TRs Table (see appendix A.2)

Name	Description	Category	Sub-category	Operator Group	Action Count
E-Mail	Communication via email	01: Business	Communication	Manager & Contractor	6
Machine Settings	Reminds the operator to check harvest parameters	02: Machine	Parameters	All drivers	2
Exercise	Short exercises for relaxation or stretching	03: Wellbeing	Wellbeing	Manager & Employees	10
User Manual	Shows the manual or tutorial video of the machine	04: Wiki	Help	Employees	1
News	Access to various news sites via the browser	05: Media	News	Manager	3
Seat Position	Change the current seat position	06: Settings	Seat	All drivers	11

- Description: A concise, expressive description of the activity.
- Disabled: A boolean value indicating whether the TR has been disabled or not. With the necessary permissions, the operator is able to disable TRs.
- Category: One of the seven categories mentioned above.
- SubCategory: The subcategory of the TR.
- LIP: The Level of Information Processing (LIP) is part of the Cognitive Task Load (CTL) model and is described in more detail in section 7.3.
- OperatorGroup: One or more assignments to predefined operator groups. Only operators in these groups have access to this TR.

- CallHistory: All previous calls with timestamps. This information is used in the "Selection" module of the VA. More details can be found in section 6.2.

In addition, each TR object contains a workflow structure consisting of three different stage types. The entire workflow logic of a TR is implemented as a method within the TR object, ensuring that the attributes and actual execution of the TR are closely linked. Initially, a TR passes through the "Trigger-Stages", which can be distinguished as "passive" and "active". Subsequently, any number of "Main-Stages" can be executed. These stage types are also defined as objects to leverage the advantages of object-oriented programming. The overarching stage object consists of the following attributes:

- TextOutput: Displayed text of the current stage in the VA User Interface (UI).
- SSMLOutput: Speech output of the VA in the form of Speech Synthesis Markup Language (SSML).
- PowerLevelInterruptible: Authorization value between 0 and 9 indicating whether this TR can be interrupted.
- View: Virtual display where visual outputs of the current TR stage are shown.
- Action: Definition of the programmatic action to be executed by the stage.
- Background: Indicates whether the stage should run in the background.
- Disabled: Boolean value indicating whether this stage is disabled or not.
- Saveable: Indicates whether the TR can be minimized in the current stage.
- Default: A default option that can be defined to be triggered automatically after a specified time.
- Options: All possible interactions in this stage, including the textual representation of the option and the subsequent stage that follows this option.

The “Trigger-Stages” inherit all these attributes but additionally include properties specific to these stage types. For “Active Trigger-Stages”, this includes:

- TriggerWords: Contains all regular expressions to start this stage via voice control [149]. For the TR “Weather”, example TriggerWords are “`what.*temperature.*(right now|at the moment|currently|now)`” and “`how.*(hot|warm|cold).*(right now|at the moment |currently|now)`”.
- PowerLevelInterruptive: Authorization value between 0 and 9 to interrupt other TRs.
- CallHistory: All previous calls of this stage with timestamps.

The “Passive Trigger-Stages” inherit from the “Active Trigger-Stages” and add further properties:

- Timer: Specified time at which the TR in this stage is triggered.
- Interval: Specified time interval at which the TR in this stage is repeatedly triggered.

The final part of a TR object is a list of all integrated sub-TRs. These sub-TRs contain the actions executed during a designated stage. Examples include an SMTP function for sending an email or API calls. Sub-TRs are completely separate from the TR object and stored in a separate database. They are only linked within a TR object. This approach allows other TRs to access generic actions without creating duplicates in a different context.

7.2. Influence of task recommendations

7.2.1. Introduction

In this participant study, the influence of individual TRs on the mental strain of participants is systematically examined. In the later real-world application of the assistance system, the investigated TRs form the basis for specifically influencing the user’s mental strain state. However, this requires quantifying the effects of individual TRs on the mental strain state. The investigation serves as a preliminary study for the overall system study described in section 8.1, as the findings directly contribute to system integration. Based on the study's objectives and previous research findings, the central research question of this study and the third research question of chapter 3 is formulated as follows:

"How can the influence of TRs on mental strain be reliably quantified and generalized?"

To address this main question, three specific sub-questions were formulated:

1. **Impact of the historical mental strain state:** *"Does the preceding mental strain state affect the impact of a TR?"* This analysis is essential for the future selection of context-specific suitable TRs.
2. **Categories as a generalization approach:** *"Is it possible to define categories of TRs so that results can be transferred to task recommendations not yet investigated?"* Since only a limited number of TRs are studied, an analysis of generalizability is indispensable.
3. **Diversification of selected TRs:** *"Can TRs be distinguished in their effects on mental strain?"* The objective is to analyze the investigated TRs individually and compare their effects.

Answering these questions contributes to demonstrating the applicability of individual TRs as well as potentially universal categories for the overall system. Moreover, the findings can be used to validate the method for quantifying TRs within the algorithm presented in section 7.3.

7.2.2. Method

The study was conducted as a quantitative experiment. The methodology of this study is explained in detail below.

7.2.2.1. Participants

A total of 54 individuals participated in the study, of whom six had to be excluded from analysis due to technical issues. The remaining data of $N = 48$ participants (16 female, 32 male) were fully or partially included in the analysis.

To increase the number of TRs examined, participants were randomly divided into two groups. Group 1 consisted of $N = 25$ participants (6 female, 19 male) aged between 20 and 55 years ($M = 26.92; SD = 6.98$). Group 2 consisted of $N = 23$ participants (10 female, 13 male) aged between 18 and 59 years ($M = 29.65; SD = 12.34$).

Participants were primarily recruited from the group of students and employees at KIT as well as from their personal networks (friends and relatives) in the region of Karlsruhe. All participants had little to no prior experience with combine harvesters. Before the study began, they were thoroughly informed about the procedure, potential risks, their rights, and data pseudonymization. Subsequently, all participants provided their informed consent to participate.

7.2.2.2. Materials

The study was conducted in the demonstrator cabin of the research project "Fahrerkabine 4.0". Detailed information on the general setup of the demonstrator cabin is provided in section 5.1. This test environment enables reproducible scenarios for all participants, allowing a detailed assessment of the effects of the TRs under controlled conditions. To prevent accidental operation or access to irrelevant functions, the functionality of the VA and the user interface was limited to the minimum necessary for the study. This ensured a standardized and valid procedure for all participants.

7.2.2.3. Experimental procedure

The procedure was identical for all participants and comprised eleven Scenarios (S). Each Scenario lasted five minutes and consisted of two consecutive Levels: one Scenario Drive (SD) (Level 0) and one TR (Level 1). The detailed procedure is illustrated in fig. 7.1.

Level 0 consists of 180 s of one of two possible Scenario Drives (SD) and concludes with a 30 s Rating Scale Mental Effort (RSME) and Self-Assessment Manikin (SAM) assessment [51], [150]. This RSME assessment is used to determine the mental strain before the TR.
The two possible Scenario Drive (SD) types differed in the following aspects:

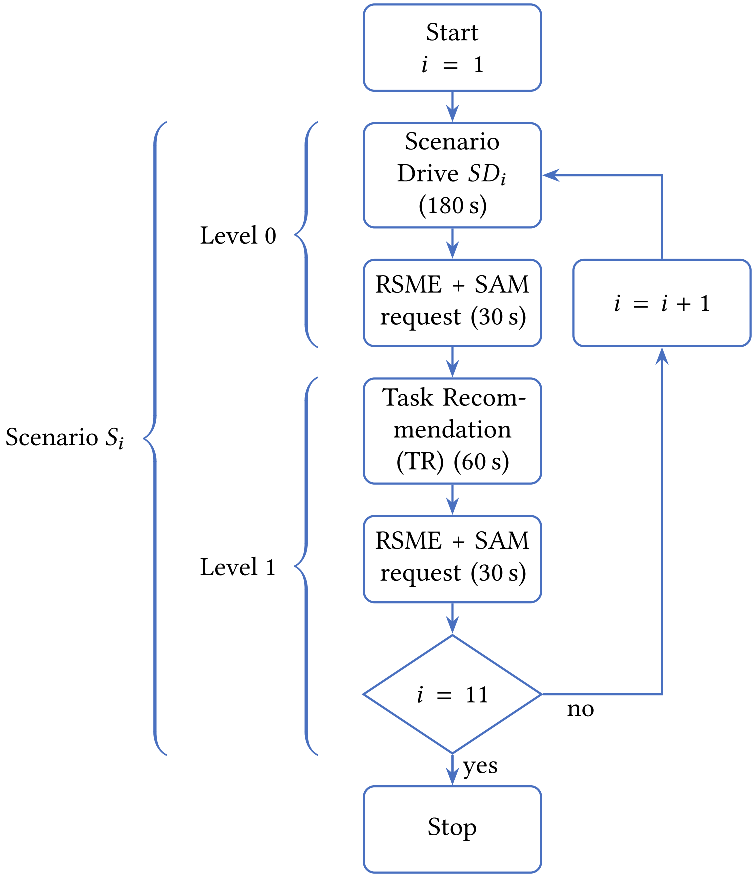


Figure 7.1.: Procedure plan of the task recommendation study

- Relaxed SD:
  In the "Relaxed SD", the participant assumed a purely observational role. This Scenario Drive (SD) type was identical to the "Autonomous Drive" described in section 5.2. The participant was completely relieved of active input and did not interact with the VA. The driving task was executed automatically.

- Demanding SD:
  In the “Demanding SD”, the participant had to steer manually and followed a predefined track in the test field at a constant speed. It had to be ensured that the center of the cutting unit aligned with the field boundary. If the participant deviated from the track by more than one meter, the virtual assistant issued regular reminders that the track must not be left and that the steering performance was insufficient. Additionally, the participant was required to modify machine settings at two-second intervals. The prompts for these adjustments were provided both auditorily and visually on the front windshield by the virtual assistant. An example of such a prompt is: “Dreschtrommeldrehzahl auf ‘950’ $\frac{1}{\text{min}}$ ändern.” The possible adjustment options are illustrated in simplified form in fig. 5.4 in section 5.2. In parallel, the participant also had to solve as many arithmetic problems as possible on the left touch display, as illustrated in section 5.2 (fig. 5.5). It was specified that the driving task always had the highest priority, followed by the machine settings. Solving arithmetic problems had the lowest priority but still had to be performed simultaneously with the other tasks. Each task should be performed as well as possible. This Scenario Drive (SD) type was identical to the Scenario Drive (SD) “TR_{MS2C}” described in section 5.2.

The subsequent TR is conducted in Level 1 for 60 s. Five different categories of TRs in Level 1 could be analyzed. Each category of TRs was examined under both scenario types in Level 0, resulting in ten scenarios. This includes five “Relaxed SD” scenarios and five “Demanding SD” scenarios. This allows the impact of the different preceding scenario types to be investigated. Additionally, an eleventh scenario was conducted as a control scenario, which contained a “Demanding SD” in both Level 0 and Level 1. Together with the previously existing ten scenarios with TRs, this results in a total of eleven scenarios investigated. The order of the eleven scenarios was randomized for each participant, with each scenario presented exactly once. After each TR, the participant’s cognitive state was again determined using the RSME assessment. The scenario then ended, and the next scenario began. The study concluded for the participant after they had completed all eleven scenarios.

During the selection of the TRs, care was taken to ensure that they could be carried out in a potential future practical application. In the “Office” category, for example, participants were required to read a weather report for different locations and recite details to the experiment management or to compose a social media post about the previous TR. The "Relaxation" category included a stretching and breathing exercise. In the "Infotainment" category, participants read various Wikipedia articles or listened to different informative audio articles. In the final category, "Entertainment", humorous videos were projected onto the windshield or a simple game was presented on the left touch display. A detailed breakdown of the individual TRs and their assignment to the corresponding groups is provided in table 7.2. Since the TR "Autonomous Drive" appeared in both groups, nine distinct TRs were examined overall.

Table 7.2.: Overview of the Task Recommendations (TRs) and their corresponding categories

Category	Group 1	Group 2
Relaxed Drive	Autonomous Drive	Autonomous Drive
Office	Read Weather Report	Write Social Media Post
Relaxation	Stretching Exercise	Breathing Exercise
Infotainment	Read Wikipedia Article	Listen to Audio Article
Entertainment	Mini Game	Watch Videos

7.2.2.4. Data processing and analysis

All data generated during the study were centrally recorded and stored by the assistance system. Since the implementation of the assistance system was primarily carried out in Matlab® R2023b, the same software environment was used for subsequent data processing and analysis. The collected raw data included pseudonymized participant IDs, the TRs performed, and the corresponding RSME and SAM values before and after each TR.

7.2.3. Results

The following section presents the results of the study. The analysis proceeds step by step, starting with a manipulation check, followed by an examination of the influence of the scenario type on mental strain, an evaluation of the individual TR categories, and finally post-hoc analyses between the categories.

Manipulation check:
At the beginning, it was examined whether the two different Scenario Drive (SD) types successfully induced different Mental strain states (MSSs) in the study. The RSME values reported by the participants of both groups after the Scenario Drive in Level 0 are shown in fig. 7.2. The “Relaxed SD” led to an underload state ($Mdn = 0; IQR = 2.75$), while the scenario “Demanding SD” was at the threshold of overload ($Mdn = 70; IQR = 17$) and exhibited greater variability. The Wilcoxon signed-rank test indicated a significant difference between the two conditions ($W = 0; p < 0.001$).

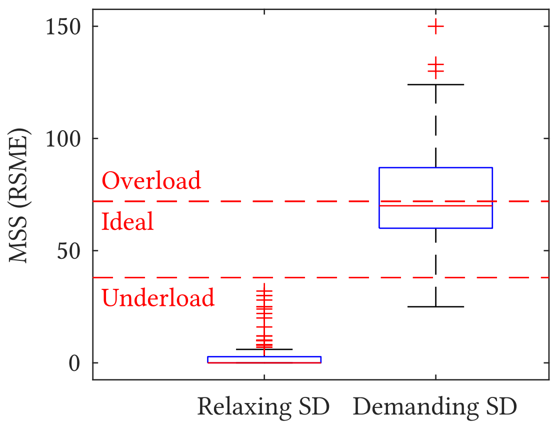


Figure 7.2.: RSME in Level 0 of both scenario types

Influence of Scenario Drive (SD) type on RSME values:
Since it has now been established that the different scenario types have distinct effects on the participants’ MSSs, it was further examined whether these effects persisted after the TRs in Level 1. Figure 7.3 shows the RSME values of both groups in Level 1 (after the TR) for each of the possible preceding scenario types in Level 0. When the preceding level involved the “Relaxed SD” scenario, the mental strain after the TR was rated lower ($Mdn = 5; IQR = 14$) than after the “Demanding SD” scenario ($Mdn = 20; IQR = 30$). According to the Wilcoxon signed-rank test, the differences between the two scenario types are significant ($W = 3130; p < 0.001$).

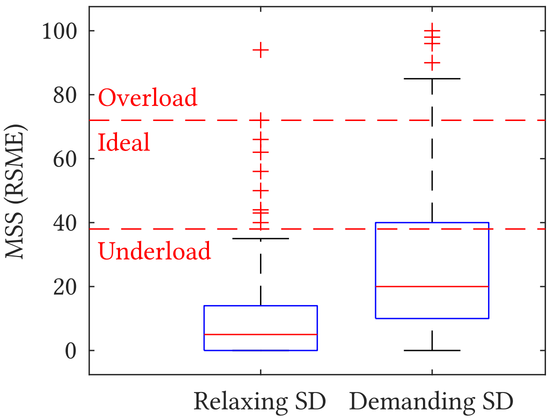


Figure 7.3.: RSME in Level 1 of both scenario types

Analysis of the categories:
To evaluate the effect of each TR, the following section examines the participants' MSS in the form of RSME values after each respective TR. These TRs are divided according to their categories and averaged for both scenario types.

As an additional test of the study, the TR "Autonomous drive" can be considered. Figure 7.4 shows the RSME values for both groups. Since both groups performed the same TR, these values should not differ. The differences between the groups are, as expected, not significant ($W = 2248; p = 0.152$).

Figure 7.5 presents the evaluation of all remaining categories for each group, allowing for a comparison of each TR within each category. In the following analysis, the individual TRs are denoted as "$Category_{GroupNumber}$". A detailed breakdown of the individual TRs is provided in table 7.2.

The first category to be examined in fig. 7.5a, "Office", shows different effects on the median RSME values for both groups ($Mdn_1 = 26.5; IQR_1 = 32$ and $Mdn_2 = 20; IQR_2 = 23$). In both groups, individual outliers with RSME values above 80 can be observed. A Wilcoxon signed-rank test indicates that these differences are not statistically significant ($W = 2556; p = 0.170$).

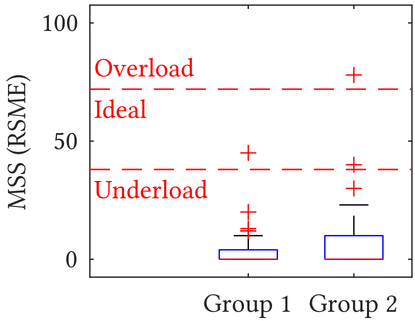


Figure 7.4.: RSME values for category "Autonomous drive"

The category "Relaxation" in fig. 7.5b shows clearly smaller deviations with identical medians and similar Interquartile Range (IQR) values ($Mdn_1 = 10; IQR_1 = 18$ and $Mdn_2 = 10; IQR_2 = 23.75$). These differences are also not significant ($W = 2431; p = 0.818$). The maximum RSME values are $\text{Office}_{1,max} = 66$ and $\text{Office}_{2,max} = 70$.

The third category also shows only minor differences in fig. 7.5c ($Mdn_1 = 12.5; IQR_1 = 14$ and $Mdn_2 = 11; IQR_2 = 19$). The maximum values are $\text{Infotainment}_{1,max} = 90$ for Group 1 and $\text{Infotainment}_{2,max} = 80$ for Group 2. The Wilcoxon signed-rank test reveals no significant difference ($W = 2432; p = 0.665$).

The category "Entertainment" shows clear differences in the recorded RSME values. As seen in fig. 7.5d, Group 1 shows higher RSME values than Group 2 ($Mdn_1 = 26.5; IQR_1 = 27$ and $Mdn_2 = 4; IQR_2 = 16$). In this category, the maximum RSME values are $\text{Entertainment}_{1,max} = 85$ and $\text{Entertainment}_{2,max} = 78$. These differences are statistically significant ($W = 3146; p < 0.001$).

In all listed categories, the minimum RSME value is $\text{allCat}_{min} = 0$. The category "Entertainment" is the only one that shows statistically significant differences within the category.

Since the categories differ in some cases significantly, the categories cannot be considered together in the following. In the subsequent data analysis, both groups are now considered separately.

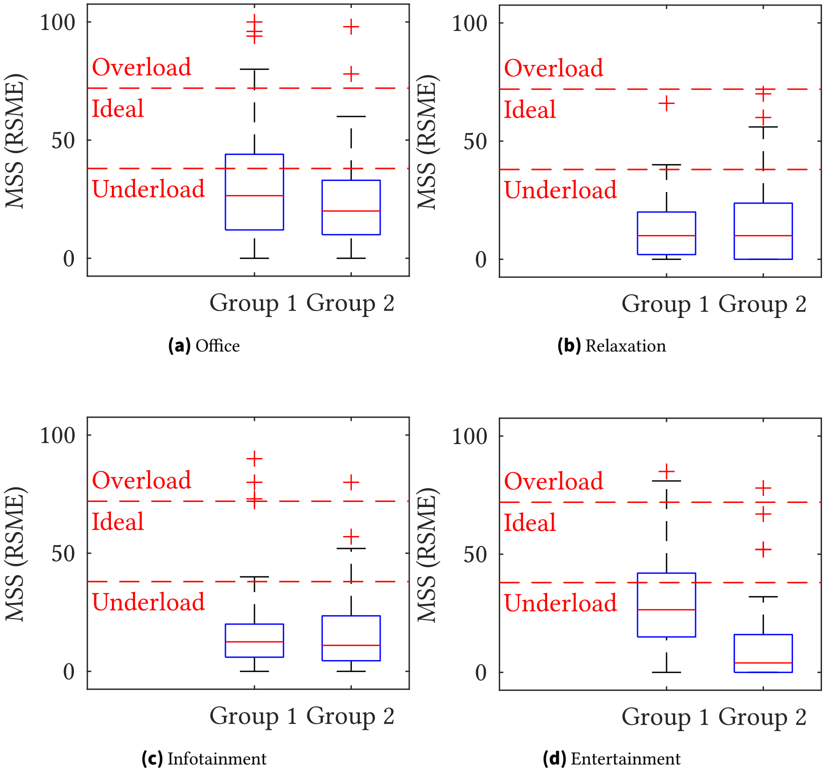


Figure 7.5.: RSME values in Level 1 for all categories

Comparison of all TR:
A Friedman test was conducted for each group to analyse differences between the categories within each group [151]. These tests show that there are statistically significant differences in the RSME values between the categories for both Group 1 ($\chi^2 = 52.56; p < 0.001$) and Group 2 ($\chi^2 = 29.82; p < 0.001$).

To obtain detailed comparison results of the individual categories within each group, post-hoc comparisons were carried out. For this purpose, the MATLAB `multcompare` function was used, which performs pairwise comparisons based on the Friedman test results using the "Tukey-Kramer honestly significant difference procedure" [24]. The results of these tests are listed in table 7.3 and table 7.4. Bold values indicate statistical significance.

The post-hoc comparison results for Group 1 (table 7.3) show that two comparisons are not statistically significant (Office$_1$ <-> Entertainment$_1$ and Relaxation$_1$ <-> Infotainment$_1$). All other comparisons show significant differences. Office$_1$ and Entertainment$_1$ cause significantly higher RSME values than Relaxation$_1$ and Infotainment$_1$.

Table 7.3.: Post-hoc matrix for Group 1

	Office$_1$	Relaxation$_1$	Infotainment$_1$	Entertainment$_1$
Office$_1$	–	< **0.001**	< **0.001**	0.992
Relaxation$_1$		–	0.317	< **0.001**
Infotainment$_1$			–	< **0.001**
Entertainment$_1$				–

In Group 2, the post-hoc comparisons show no significant differences for three comparisons (Office$_2$ <-> Infotainment$_2$, Relaxation$_2$ <-> Infotainment$_2$, and Relaxation$_2$ <-> Entertainment$_2$). Office$_2$ yields significantly higher RSME values than Relaxation$_2$ and Entertainment$_2$. Conversely, Infotainment$_2$ results in significantly higher RSME values compared to Entertainment$_2$.

Table 7.4.: Post-hoc matrix for Group 2

	Office$_2$	Relaxation$_2$	Infotainment$_2$	Entertainment$_2$
Office$_2$	–	< **0.001**	0.131	< **0.001**
Relaxation$_2$		–	0.400	0.452
Infotainment$_2$			–	**0.013**
Entertainment$_2$				–

When analyzing the individual TRs while considering intergroup and intercategory significant differences, it is possible to group them using the Compact Letter Display (CLD) method (fig. 7.6). Three distinct groups (blue, green, and red) were formed, which showed significant differences between each other, but not within groups. These groups are used in section 7.3 for the quantification of the TRs in the regulation logic of the assistance system.

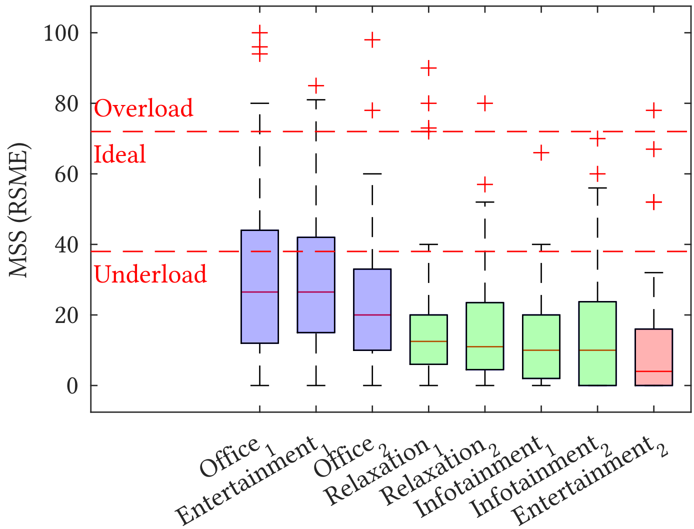


Figure 7.6.: Ranking of all TRs using the CLD methodology

7.2.4. Discussion

This chapter relates the study results to the initial sub-research questions in order to answer the main research question. The objective of the present study was to systematically investigate the influence of different TRs on the mental strain of participants. The intention was to establish a methodological basis for quantifying the effect of individual recommendations and to make the results applicable for later system integration. The following section summarizes the key findings, critically discusses them, and places them in the context of the initial research objectives.

Impact of the historical mental strain state:
First, it should be emphasized that the study confirmed that the two scenario types, relaxed and demanding drive, reliably induced different levels of mental strain. This is clearly reflected in the reported RSME values, as shown in fig. 7.2. While the Relaxed SD scenario resulted in very low mental strain values, the Demanding SD reached the threshold of overload. This

manipulation check underlines that the experimental setup was suitable for creating a differentiated baseline for investigating the TRs. It should be noted that these differences were not limited to Level 0 but also persisted after performing the respective TRs (Level 1), as illustrated in fig. 7.3. The mental strain after a TR was significantly lower when participants had previously completed a relaxed drive than after a demanding one. This result highlights that the effectiveness of TRs cannot be considered independently of the preceding mental strain state. Therefore, the first sub-research question, *"Does the preceding mental strain state affect the impact of a TR?"*, can be considered answered, as a clear influence of the preceding MSS was identified. Since the dynamic effect of the preceding MSS would have too great an impact on the complexity of the assistance system at its current stage, the results of the two scenario types are not considered separately. Consequently, in the following discussion the results of Relaxed and Demanding SD scenarios are combined. Since the aim of this study is only a coarse quantification of the TRs, this approach is not problematic. However, this effect should be taken into account in future research. In practical terms, this means that TRs must be adaptively adjusted to the situational context, for example by integrating a mental strain monitoring system (chapter 5).

Categories as a generalization approach:
Another important finding concerns the possibility of classifying TRs into categories to generalize their effects. The analyses show that this is feasible in principle, however there are limits for some TRs. In particular, the categories "Office", "Relaxation" and "Infotainment" proved to be consistent. The categories "Relaxation" and "Infotainment" resulted in comparatively low and stable mental strain values with no significant differences across both groups. In contrast, the category "Entertainment" showed distinct differences between the groups, and thus between the two TRs it contained. One possible explanation is that the TR "Mini Game" imposes a substantially different type of cognitive demand than the TR "Watch Videos" (active task vs. passive task). This discrepancy may have led to a high mental strain state in one group, while producing a low state in the other. This finding is important as it reveals the limits of a purely category-based generalization. The effect of TRs within a category evidently depends strongly on their specific design and the associated cognitive demands. Therefore, only a partial transfer of the results to contents within the same category is possible. However, generalization through categories could be improved if more specific categories are defined

that take the specific design of the TRs and the associated cognitive demands into account. Consequently, the second sub-research question, *"Is it possible to define categories of TRs so that results can be transferred to task recommendations not yet investigated?"*, can also be considered answered, as the study demonstrates both the feasibility and the limitations of such categorization, in addition to presenting an improved approach.

Diversification of selected TRs:
The detailed analysis of individual TRs further showed that their effects can differ considerably. To analyze these effects, an assessment using the RSME was conducted. Some TRs, such as specific relaxation exercises, tended to induce low mental strain, whereas others, particularly in the Office category, could have the opposite effect. The post-hoc analyses confirmed these differences and demonstrated that the TRs can even be ordered according to their effects. Based on the CLD grouping and sorting by the median of all TRs, the TRs can be ranked according to their effects, thus answering the third and final sub-research question, *"Can TRs be distinguished in their effects on mental strain?"*.

In summary, the results demonstrate that it is possible to quantify the influence of TRs on mental strain and to group them accordingly. It also became evident that partial generalization across categories can be feasible, though subject to clear limitations. The preceding mental strain state plays a crucial role, as it affects the impact of the TRs. Furthermore, the investigated TRs can be differentiated and comparatively classified, providing valuable insights for later system integration. The IQR of the individual TRs additionally shows that the induced mental strain is highly individual and that the quantification should therefore only serve as a coarse initial reference value. The individual effects of the respective TRs on the operator should be monitored live during operation and should be corrected or updated accordingly in the user profile. Furthermore, none of the TRs was able to reliably achieve an ideal state directly. This means that for the actual regulation of the MSS, a combination of multiple TRs should be considered. The main research question, *"How can the influence of TRs on mental strain be reliably quantified and generalized?"*, can thus be considered answered with reference to the sub-research questions answered above. These findings make an essential contribution to validating the quantitative models developed in the project and offer concrete starting points for the user-centered design of future assistance systems. For practical applications, this means that successful implementation requires an adaptive, context-sensitive selection of TRs supported by a MSS monitoring system. However, some limitations of the study should be considered. The number of investigated TRs was limited, representing only a subset of possible measures. Although categories such as “Office”, “Relaxation” and “Infotainment” proved to be relatively consistent, it remains unclear whether similar effects would be observed with a larger and more diverse selection. In particular, care should be taken to ensure that the TRs have similar cognitive demand types and similar resources, since according to Wickens these can cause different mental strains [52]. Therefore, future work should extend the investigation to a broader range of TRs and different categories. Additionally, it should be examined whether individual preferences, prior experiences, or situational factors should be given greater consideration than was possible within the scope of this study.

7.3. Control loop

7.3.1. Overview

The following section explains the method for regulating the cognitive user state. This method is illustrated using a control loop defined in accordance with DIN IEC 60050-351 [152]. A generic control loop of this type is shown in fig. 7.7.

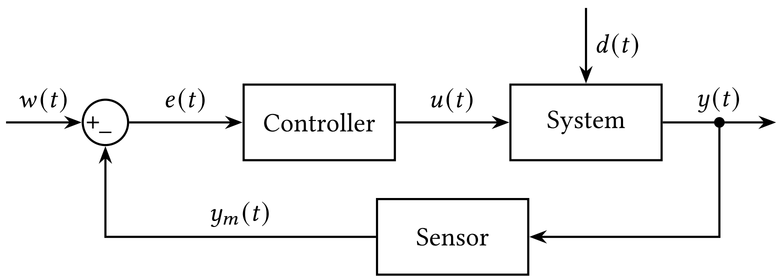


Figure 7.7.: Control loop according to [152]

First, the reference variable $w(t)$ must be defined. To do so, quantification of the desired user state is required. The objective of the assistance system, and thus of the depicted control loop, is to maintain or move the operator's cognitive strain state within an optimal range. The ideal range in which optimal performance and well-being of the operator are achieved has been extensively investigated and described in section 2.2. With the RSME and the findings of Funk et al. [68], the ideal mental strain state can be defined within the range $RSME_I \in \mathbb{N} \mid RSME_I \in [38, 71]$. However, the mental stress is not taken into account in this context. A combination of mental strain and mental stress in the reference variable $w(t)$ allows the underlying causes of a critical mental strain to be assessed and thus enables targeted manipulation of the mental stress via TRs.

To represent the operator's mental stress and mental strain state, the CTL model by Neerincx and Jeschke is used [56], [59]. The "Load Space", a three-dimensional space, defined by the CTL model is shown in fig. 7.8. The critical stress regions Underload (UL), Vigilance (V), Mental Satiation (MS), Cognitive lock-up (CL), and Overload (OL) are marked with black hatching. More information on how mental stress and mental strain are represented in the CTL model can be found in section 7.3.2.

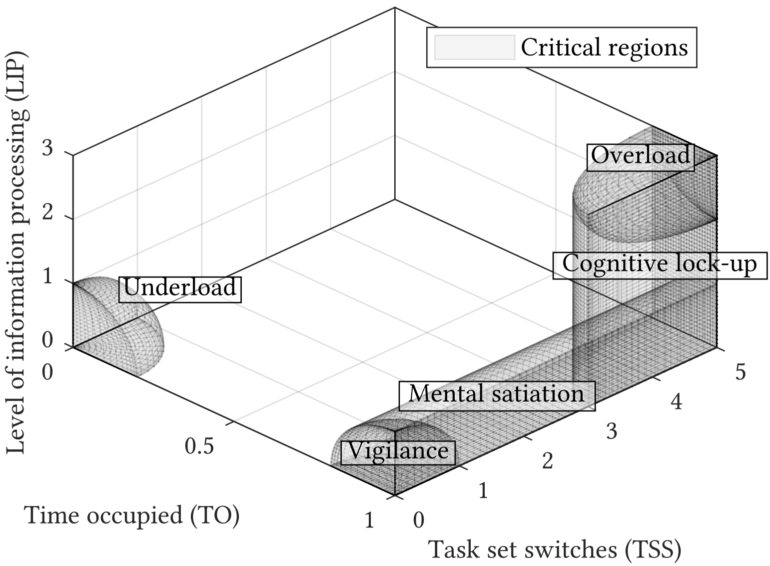


Figure 7.8.: CTL space according to Jeschke [59]

Consequently, the reference variable $w(t)$ is defined such that the operator's mental stress and mental strain state does not lie within these critical regions ($w(t) \notin \{UL, V, MS, CL, OL\}$). Together with the output $y_m(t)$ resulting from the "Sensor" block, the control deviation $e(t)$ is determined. Both $y_m(t)$ and $e(t)$ can be separated into mental strain ($y_{MSS}(t)$, $e_{MSS}(t)$) and mental stress components ($y_{MStS}(t)$, $e_{MStS}(t)$). For better clarity, these variables are presented in a consolidated manner in fig. 7.7. This control deviation serves as the input variable for the "Controller". Subsequently, a control variable $u(t)$ is supplied to the "System", which in turn results in the sensor input variable $y(t)$ while taking potential disturbances $d(t)$ into account.

7.3.2. Sensor

The "Sensor" block comprises an entire system of different sensors. This block captures and processes both all user interactions and the predictors required for monitoring the current MSS, as described in chapter 5.

To quantify and compare the user's mental strain and mental stress, the results of the MSS monitoring system, based on eye tracking data and fitness tracker data, as well as the stress data derived from user interactions, historical data, and current TR data, are transferred into the three dimensional CTL space. The output variable $y_m(t)$ is calculated for both the Mental stress state (MStS) and the MSS. Figure 7.9 illustrates the general structure and the processing of the sensor data of the "Sensor" block. The transfer process of the two state types is explained in more detail below.

Mental stress state (MStS):
To calculate the current MStS, all three parameters of the CTL model must be determined. These parameters are considered over an empirically derived time window of $\Delta t = 120$ s, which was defined in the undocumented preliminary studies of the study presented in section 7.2.

First, the "Task Set Switches (TSS)" are examined. According to Neerincx [56], these are defined as amount of tasks and activities running in parallel. In the present case, all active as well as historical TRs performed within the observed time frame are considered. In addition, background or minimized TRs and the current driving or steering mode of the machine are included in the calculation. To incorporate adjustments to machine settings into the TSS

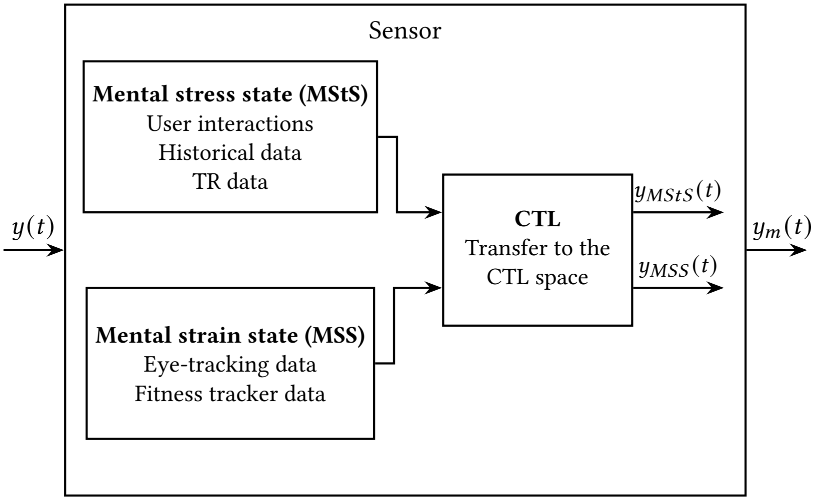


Figure 7.9.: Structure of the "Sensor" block

parameter, all user actions within the observed time window are integrated into the assistance system. For this purpose, all inputs are multiplied by a weighting factor $w_{UI} = 0.125$, which was empirically determined during the undocumented preliminary studies of the study described in section 7.2. Finally, all values are summed so that $TSS(t) \in \mathbb{R} \mid TSS(t) \in [0, 5]$ can be determined. This range is based on the quantification of the CTL model proposed by Jeschke [59].

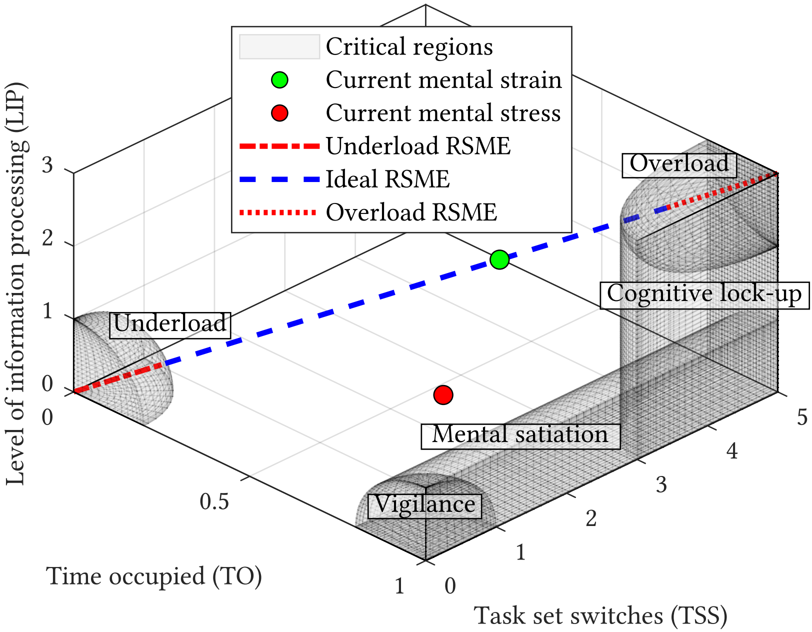
The second parameter to be considered is the "Time Occupied (TO)". This describes the percentage of time during which the operator is occupied. A value of $TO(t) = 1$, where $TO(t) \in \mathbb{R} \mid TO(t) \in [0, 1]$, means that the operator was fully occupied during the observed time window [59]. Similar to the TSS parameter, all active and historical TRs, as well as concurrent background and driving tasks, are considered. User actions in the UI related to machine settings are incorporated into the calculation via a time constant. If the calculated value exceeds the maximum of 100% due to simultaneous activities, it is limited accordingly.

Finally, the parameter "Level of Information Processing (LIP)" is determined. For this purpose, the Skill-Rule-Knowledge framework according to Rasmussen [58] is applied. This framework divides tasks and activities into three levels:

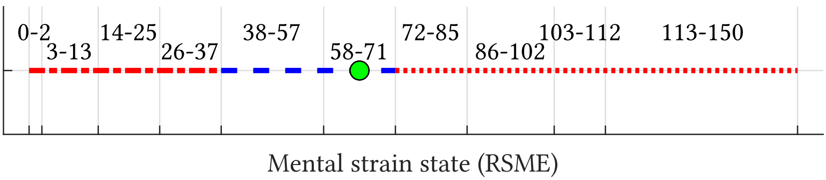
- **Skill-based level:** Activities at this level require minimal cognitive effort, as information is processed largely automatically. The implemented LIP value interval is $LIP_S(t) \in \mathbb{R} \mid LIP_S(t) \in [0, 1[$.
- **Rule-based level:** At this level, incoming information is processed through habitual or predefined procedures, allowing efficient problem-solving with moderate cognitive effort. With $LIP_R(t) \in \mathbb{R} \mid LIP_R(t) \in [1, 2[$.
- **Knowledge-based level:** At this level, incoming information serves as a basis for analyzing problems and generating possible solutions, particularly in new or unfamiliar situations. This level is associated with significant cognitive effort and with the assigned interval $LIP_K(t) \in \mathbb{R} \mid LIP_K(t) \in [2, 3]$.

The driving task can be assigned to the “skill-based level” ($LIP_D = 0.66$) due to its characteristic features. Adjustments to machine settings generally follow predefined rules and therefore belong to the “rule-based level” ($LIP_{MS} = 1.33$). The LIP values of selected TRs were assigned based on the results of the study presented in section 7.2. The ordering and grouping from fig. 7.6 and the characteristic features of the TR within the Skill-Rule-Knowledge framework of Rasmussen were taken into account [58]. The blue group was assigned $LIP_{blue} = 1.66$, the green group $LIP_{green} = 1.00$, and the red group $LIP_{red} = 0.33$. For the resulting LIP value, the highest occurring value within the observed time window is used. Active and historical TRs in the time window are considered equally.

Once all parameters are determined, a mental stress point can be calculated within the three-dimensional load space of the CTL model. Figure 7.10 shows an example of such a point in red: $(TSS, TO, LIP) = (1.5, 0.75, 1)$. In this example, the point lies outside a critical area, resulting in a control deviation of the Mental stress state (MStS) of $e_{MStS}(t) = (0, 0, 0)$. However, if the current stress point is located within a critical area, the minimum distance vector, that is, the smallest deviation required to leave the critical region, is calculated as the control deviation. This value ($e_{MStS}(t)$) is then passed to the “Controller”.



(a) Mapping of mental strain to the CTL space



(b) Number line of mental strain mapping

Figure 7.10.: Concept of mental strain mapping

Mental strain state (MSS):
To transfer the operator's current mental strain into the CTL space, a method must be established that extends the one-dimensional results of the mental strain measurement, obtained through the MSS monitoring system introduced in chapter 5, to three dimensions.

To do this, a diagonal is defined in the CTL space, extending from the underload region to the overload region (fig. 7.10). Along this diagonal, the RSME value range is mapped by a piecewise linear, continuous function. This mapping is illustrated in fig. 7.10a. Considering the assumptions of Funk et al. [68], a value range of $MSS_{RSME,UL} \in [0, 37]$ is defined within the underload region and a range of $MSS_{RSME,OL} \in [72, 150]$ within the overload region. All intermediate values ($MSS_{RSME,I} \in [38, 71]$) are classified as non-critical or ideal. Where all MSS values are natural numbers ($MSS_{RSME} \in \mathbb{N}$). Figure 7.10b shows the linear mapping function of the mental strain with the ten strain classes detectable by the mental strain monitoring system. As an example, a detected strain class of "58–71" is shown as a green point in fig. 7.10.

If the strain point transferred into the CTL space is located within a critical region (for example "overload"), the control deviation $e_{MSS}(t)$ is determined analogously to the calculation of the mental stress state. Both control deviations are then forwarded to the "Controller".

7.3.3. Controller

In the "Controller", the current system state is monitored and processed, and a control variable $u(t)$ is output in the form of a TR. For this purpose, all steps of the assistance system described in section 6.2 are executed. The key module of the assistance system is the selection module. In this module, the data streams of the current system state are examined for trigger conditions for TRs. These include machine states ($s_M(t)$), such as threshing parameters, environmental conditions ($s_E(t)$), such as weather changes, internal TR triggers initiated by timers or intervals ($s_I(t)$), and critical control deviations ($e_{MSS}(t)$). The status data streams are not included in the control loop described above (fig. 7.7), as they do not have a direct influence on the control process but only affect the user experience. However, they are listed in the detailed schematic of the controller in fig. 7.11.

Only critical control deviations caused by critical MSSs ($e_{MSS}(t)$) are considered, as different mental stress states may trigger varying levels of mental strain among different operators. This prevents a critical MStS from causing TRs to be suggested even though the measured MSS is ideal. Critical mental stress states ($e_{MStS}(t)$) are used solely to indicate the causes of a critical MSS and thus provide a solution to resolve this state in the ranking process. If valid trigger conditions for TRs are identified, the valid TRs are accumulated and marked for potential selection.

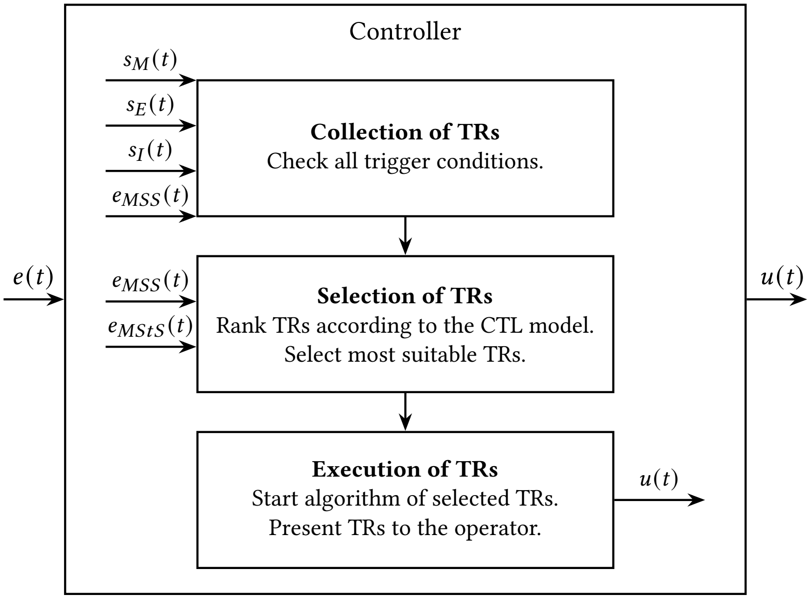


Figure 7.11.: Structure of the "Controller" block

In the selection process, the pool of accumulated TRs are then ranked as follows:

- Rank 1:
  The highest rank is assigned to those TRs that, through their respective properties, can independently bring both the operator's mental stress and mental strain state into a non-critical state. No additional TR is required.
- Rank 2:
  The second rank is reserved for TR combinations that, together with an additional TR, can bring both the mental stress and mental strain state into a non-critical condition. A combination of two TRs is selected and initiated.

- Rank 3:
  If it is not possible to optimize both the mental stress and mental strain states to non-critical levels using a single TR, only the mental strain state is considered. TRs that fulfill this condition receive the third rank.
- Rank 4:
  The lowest rank is assigned to TR combinations that can only optimize the mental strain state to "non-critical" with the help of an additional TR. A non-critical mental stress state is not achieved.

After all collected TRs and TR combinations have been assigned a rank, the TRs with the highest rank are further processed. All selected TRs are sorted by frequency of use and operator preferences within their respective category. If equally suitable options remain after this sorting, one TR is randomly selected. This approach allows optimal adaptation to the individual operator, thereby achieving the highest possible user experience.

Finally, the algorithm of the selected TR is started and presented $u(t)$ to the operator, who can then decide whether to accept or reject the proposed recommendation.

7.3.4. System

In the depicted use case, the operator assumes the role of the "System" block in the control loop shown in fig. 7.7. In a highly abstracted form, the structure of the "System" block can be represented as shown in fig. 7.12.

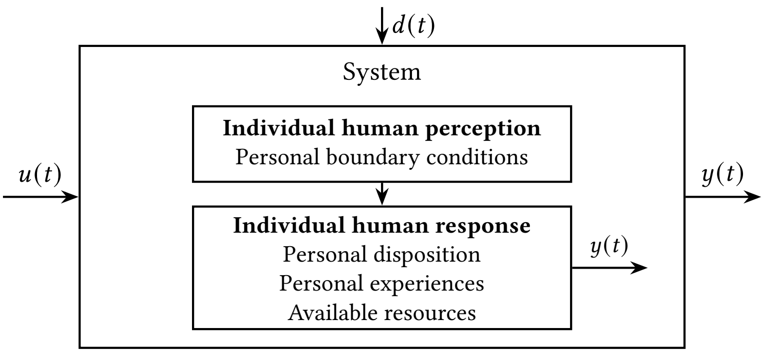


Figure 7.12.: Structure of the "System" block

Each TR ($u(t)$) has a direct influence on the Mental stress state (MStS) and Mental strain state (MSS) of the operator. This influence, the individual response and possible changes in the user state can be read out directly via the sensor input variable $y(t)$.

In addition, numerous disturbances ($d(t)$) act on this system, which can significantly affect and complicate the regulation process. Possible disturbances include both intrinsic and extrinsic uncontrollable distractions, such as phone calls or personal obligations. These influences directly impact the control performance of the entire system.
A central aspect of the disturbances is the operator's willingness to cooperate. If the operator is uncooperative, optimization of the MSS through the proposed assistance system is not possible.

8. Validation of the assistance system

For the final validation of the assistance system presented in this work, two studies are presented below that examine and analyze both the overall system and its transferability to practical application.

8.1. Overall system study

8.1.1. Introduction

In this comprehensive system investigation, the interaction of all system-integrated components is analyzed, with particular focus on the role of the virtual assistance and the monitoring of the Mental strain state (MSS). Furthermore, this study serves to verify and validate the findings presented in section 7.2. In section 7.2, an experiment with a demonstrator cabin was conducted, which showed that different task recommendations have varying effects on the participants' mental strain. The overall system investigation now makes it possible to evaluate this influence in the context of the control loop introduced in chapter 7, the Cognitive Task Load model, and the mental strain monitoring system. The model of mental strain monitoring was introduced in chapter 5. The present study now offers the opportunity to confirm these results or transfer them to other application contexts.

The study was divided into two sections to answer the study's research questions: a verification study that checks the correctness of the VA algorithm, and a validation study that examines the interaction of the individual subsystems and the functionality of the overall system.

Within the framework of this overall system study, a central research question was defined in chapter 3, derived from both previous research results and the objectives of the present investigation:

"How can an assistance system be designed to reliably regulate the mental strain state of agricultural machine operators?"

Based on this overarching research question, several sub-questions can be formulated. These address both the validation of existing findings and the evaluation of the virtual assistance in the new experimental setup:

1. **Monitoring of mental strain:** *"How can the results of the mental strain monitoring system from section 5.2 be reliably transferred to a new sample and new scenarios?"* This question analyzes whether the mental strain monitoring continues to operate with high accuracy under more realistic conditions.
2. **Reaction of the virtual assistance:** *"Does the virtual assistance act appropriately in response to the monitored MSSs?"* This question aims to examine the decision logic of the assistance system and identify potential misinterpretation in the interaction.
3. **Influence of the virtual assistance on the MSS:** *"To what extent can the virtual assistance deliberately influence the current MSS and thereby contribute to an ideal state?"* This part investigates whether and how the assistance system can help regulate the participants' mental strain.
4. **Comparison of successful and unsuccessful scenarios:** *"What differences can be observed between scenarios that ended in an ideal MSS and those that did not?"* The analysis of these differences aims to provide insight into which factors influence the success or failure of MSS regulation.

Answering these questions contributes to assessing the effectiveness and practical applicability of the mental strain monitoring and the virtual assistance and to identifying potential areas for optimization in future applications.

8.1.2. Method

The overall system study was conducted as a quantitative investigation. The methodology of this study is explained in detail below.

8.1.2.1. Participants

A total of 43 participants took part in the overall system study, of whom three had to be excluded from the analysis due to technical issues. The remaining data from $N = 40$ participants could be included at least partially in the analysis. The verification study involved $N = 10$ participants (8 male, 2 female) aged 18 to 25 years ($M = 21.1; SD = 2.34$). For the validation study, data from $N = 30$ participants (24 male, 6 female) aged 20 to 28 years ($M = 23.9; SD = 2.74$) were analyzed. Participants were recruited from the group of students and staff at KIT, with none having prior experience operating combine harvesters. Before the study began, participants were fully informed about the procedure, potential risks, their rights, and the pseudonymity of their data. All participants then provided informed consent. It was also ensured that all participants had unimpaired vision or used suitable visual aids during the study.

8.1.2.2. Materials

This study was conducted in the demonstrator cabin of the research project "Fahrerkabine 4.0". The demonstrator cabin served as the experimental environment, as this study aimed to analyze the interaction of all system components. For further information on the general setup of the demonstrator cabin, refer to section 5.1. All modifications to this system are explained below.

The implemented assistance system, i.e., the VA, was intended to make decisions autonomously in the respective study scenarios without study specific modifications. This means that the assistance system reacts to all conditions as if it were a real field operation. This allows testing the system's practical applicability to the greatest extent. While the participant was guided through the experiment, the VA suggested TRs solely based on sensor signals and the control deviation of the control loop from section 7.3. A detailed description of the technical implementation of the virtual assistance is provided in chapter 6.

In addition to the study in section 7.2, the mental strain monitoring system described in chapter 5 was implemented. Its setup is identical to the setup described in section 5.2. To enable a direct and continuous connection between the strain monitoring and the operation of the virtual assistant, the model developed in chapter 5 was extended with an algorithm that allows continuous computation. The raw data from the Smarteye system were transmitted to the monitoring algorithm via the UDP network protocol, and the raw data from the fitness tracker via Bluetooth Low Energy (BLE). The detailed methodology is described in section 6.4.

8.1.2.3. Experimental procedure

Each participant followed the same structured experimental procedure. The experiment consisted of nine Scenarios (S), which could be distinguished based on the Scenario Drive (SD) types: three "Underload Drive" Scenarios, three "Ideal Drive" Scenarios, and three "Overload Drive" Scenarios. The order of these Scenarios was randomized for each study participant. Figure 8.1 shows the schematic procedure.

The Scenario Drives (SD) of the three different scenario types differed as follows:

- Scenario "Underload Drive":
  The Scenario Drive (SD) of the "Underload Drive" scenario consisted of a fully autonomous drive. The participant only assumed an observing role, without making active inputs or interacting with the virtual assistant.

- Scenario "Ideal Drive":
  In the Scenario Drive (SD) of this scenario, the participant followed a pre-defined track in the test field at a constant given speed and had to steer independently. The center of the cutter bar had to align precisely with the field edge. If the participant deviated more than one meter from the track, the virtual assistant periodically instructed the participant to maintain the track and improve steering. Additionally, every five seconds the participant was prompted to adjust various machine settings, communicated both visually and auditorily on the front window by the virtual assistant. For further information and an example, see section 7.2.2. Participants were instructed to treat the driving task as the highest priority, with machine adjustments as a lower priority.

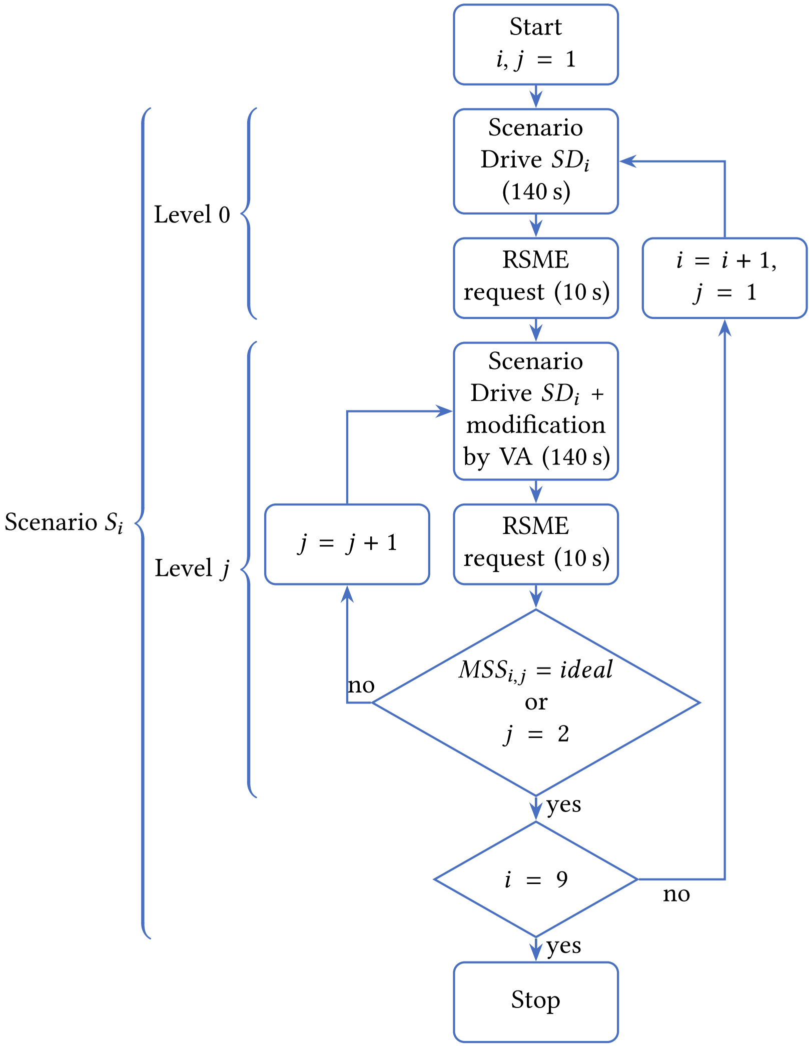


Figure 8.1.: Flowchart of the overall system study

- Scenario “Overload Drive”:
  In the Scenario Drive (SD) of the “Overload Drive” scenario, the requirements of the “Ideal Drive” scenario were increased and an additional

task was added. Instead of changing machine settings every five seconds, adjustments were now required every two seconds. In parallel, the participant was tasked with solving as many arithmetic problems as possible on the left touch display, as described in section 7.2.2. These arithmetic tasks were considered the lowest priority compared to the other tasks and were performed simultaneously with manual driving and machine settings.

Each Scenario (S) contained at least two and at most three consecutive levels. In the first level (Level 0), participants completed a 140-second Scenario Drive (SD). At the end of this level, the participants rated their current mental strain state on an interactive Rating Scale Mental Effort (RSME) interface displayed on the left touch display. Ten seconds were allotted for this assessment (see section 5.2 and fig. 5.6).

In the second level (Level 1), the VA attempted to regulate the participants' MSS through targeted TRs. For this purpose, the Scenario Drive (SD) of Level 0 was maintained, but the VA could additionally add TRs or remove activities, including those from the Scenario Drive (SD). If the participant was already in an ideal state at the beginning of this level, the VA did not intervene as long as this state persisted. This level also lasted 140 seconds and ended with an RSME query. If a participant reached the ideal cognitive state after Level 1, the Scenario (S) was terminated.

If the VA failed to shift the cognitive state into the ideal range, a further attempt was made in the third level (Level 2). This level was structurally identical to the previous one, but it built upon it. This means that all previously added or removed TRs or activities remained in place, while the VA could additionally add TRs or remove activities a second time. After completing Level 2, the Scenario (S_i) ended and the next Scenario (S_{i+1}) with randomized scenario type began immediately. Overall, the VA had the opportunity to optimize the participants' cognitive state after Level 0 between 9 (the participant always reached ideal state after Level 1) and a maximum of 18 times ($9 \cdot 2$) (the participant never reached ideal state after Level 1) throughout the experiment. A list of all possible activating TRs with the corresponding Level of Information Processing (LIP) value is shown in table 8.1.

This table includes only the TRs that would be selected in the case of underload, because in the case of overload activities or TRs would be reduced. The table is grouped according to the results of sections 7.2 and 7.3, where the blue group together with the "Arithmetic Tasks" forms the "High demand"

Table 8.1.: Overview of the TRs and their corresponding demands

Demand	Task recommendation	LIP
High demand	Read Weather Report	1.66
	Write Social Media Post	1.66
	Mini Game	1.66
	Arithmetic Tasks	1.66
Medium demand	Stretching Exercise	1.00
	Breathing Exercise	1.00
	Read Wikipedia Article	1.00
	Listen to Audio Article	1.00
	Machine Settings Adjustments	0.66
	Manual Driving	0.66
Low demand	Watch Videos	0.33

group, the green group together with the “Machine Settings Adjustments” and “Manual Driving” forms the “Medium demand” group, and finally the red group forms the “Low demand” group.

The difference between the verification and validation studies was the consideration of the RSME value by the VA. In the verification study, only the RSME value entered by the participant, was used instead of the MSS monitoring system for selecting task recommendations. This eliminated uncertainty from the mental strain monitoring system, allowing isolated verification of the virtual assistant system, including the decision logic for the task recommendations.

In contrast, the validation study considered the predicted values from the mental strain monitoring. Since all subsystems were interconnected in this study, it enabled a holistic assessment of the overall system. The RSME value served solely as a control measure for subsequent analysis.

8.1.2.4. Data preparation and analysis

All data collected during the study were centrally gathered and stored by the VA. Since the VA was primarily implemented in Matlab® R2023b, Matlab®

R2023b was also used for data preparation and analysis. The raw data included the results from the mental strain monitoring system, comprising model predictions as well as the associated model predictors, the data from the CTL model, RSME values from the RSME query interface and all interactions with the VA in the form of TRs. Data from three participants were entirely excluded due to technical issues during the experiment. For minor technical problems that did not cause a total system failure, only the affected scenario was discarded. The data from the other scenarios of these participants remained unaffected and could be used in the analysis. Overall, this affected 15 participants resulting in 19 scenarios. The loss of 19 scenarios was evenly distributed across the different scenario types, resulting in a total of 114 scenarios (verification study: 27, validation study: 87) of the “Underload Drive” type, 113 scenarios (verification study: 28, validation study: 85) of the “Ideal Drive” type, and 114 scenarios (verification study: 29, validation study: 85) of the “Overload Drive” type available for analysis.

The data analysis is conducted separately for the two sub-studies (verification study and validation study). This separation is necessary because the objectives and research questions of the two studies differ.

8.1.3. Results

8.1.3.1. Verification study

The following section presents the results of the verification study.

Manipulation check:
First, it is examined whether the different Scenario Drives (SD) in Level 0 lead to different degrees of mental strain in the participants. Figure 8.2 shows the confusion matrix for these Scenario Drives (SD).

The analysis reveals that the “Underload Drive” scenario achieved the highest success rate, with a match rate of 100%. This means that in 100% of the “Underload Drive” scenarios, the intended RSME value “0–37” was reported by the participants. The “Overload Drive” scenario resulted in an overload state for participants in 72.4% of the scenarios. The “Ideal Drive” scenario

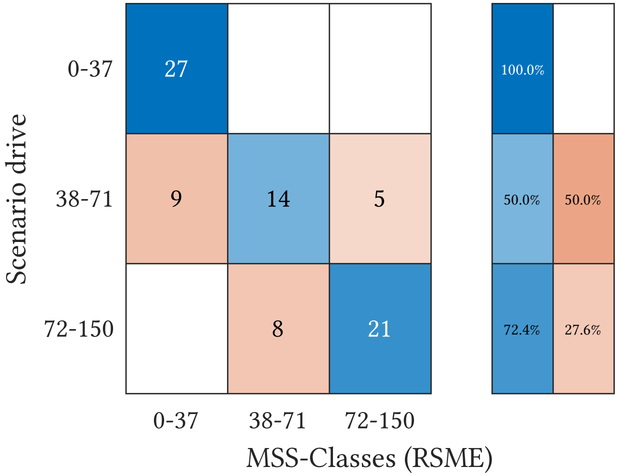


Figure 8.2.: Confusion chart for Scenario Drives (SD)

achieved a success rate of 50% in inducing the ideal mental strain. Of the remaining scenarios, 32.14% induced an underloaded state, and 17.86% induced an overloaded state.

Overall, the intended cognitive state was successfully induced in Level 0 in 73.81% of the scenarios.

To ensure a reliable basis for analyzing participants' mental strain in the subsequent evaluations, the actual reported or detected state is considered rather than the intended state. This allows inclusion of scenarios in which the intended cognitive state was not successfully induced. This strategy will also be applied to the validation study.

Henceforth, the scenarios "Underload Drive", "Ideal Drive", and "Overload Drive" refer to the actually measured mental strain of the participants. This adjustment changes the distribution to 155 scenarios (verification study: 36, validation study: 119) of the "Underload Drive" type, 107 scenarios (verification study: 22, validation study: 85) of the "Ideal Drive" type, and 79 scenarios (verification study: 26, validation study: 53) of the "Overload Drive" type.

Selection of TRs by the VA:
In the verification study, a total of 84 scenarios were conducted, resulting in 168 opportunities for the VA to issue a TR in either Level 1 or Level 2.

A closer look at the distribution shows that participants were underloaded in 67 of these situations and accordingly received an activating TR. In 39 cases, participants were overloaded, prompting the VA to reduce the amount of tasks. In the remaining 62 cases, no TR was issued because participants were already in an ideal state. Since the proposed or withheld task recommendation corresponded in all cases to the respective cognitive state class, the VA achieved a decision accuracy of 100%.

Effectiveness of the VA:
To assess the effectiveness of the virtual assistant, the RSME values reported by the participants after Levels 1 and 2 were analyzed. The goal of the VA is to shift participants toward an ideal cognitive state. Of the 84 scenarios conducted, participants were in an ideal state after Level 1 in 40 cases (47.61%). After Level 2, this number increased to 57 (67.85%).

By separately analyzing the different scenario types, the effectiveness of the virtual assistant can be evaluated as a function of the respective scenario.

The highest success rate was observed in the "Overload Drive", where 82.75% of participants were guided to the ideal user state. In comparison, the effectiveness was lower for the "Underload Drive" (62.96%) and the "Ideal Drive" (57.14%).

To understand why an ideal user state was not achieved at the end of Level 2 in 32.15% of the cases, the RSME values for each level are analyzed depending on the scenario type. This enables a detailed examination of the cognitive state progression during the scenarios.

Figure 8.3 shows the progression of RSME values for the underload scenario, considering only data from participants who did not reach an ideal user state by the end of the scenario. Across 13 scenarios in the initial drive (Level 0), the median RSME value was within the MSS class "0–2". As the levels progressed (Levels 1 and 2), the median increased to the class "14–25", with occasional deviations appearing in the "Overload" range. The values extend exclusively across the underload and overload ranges.

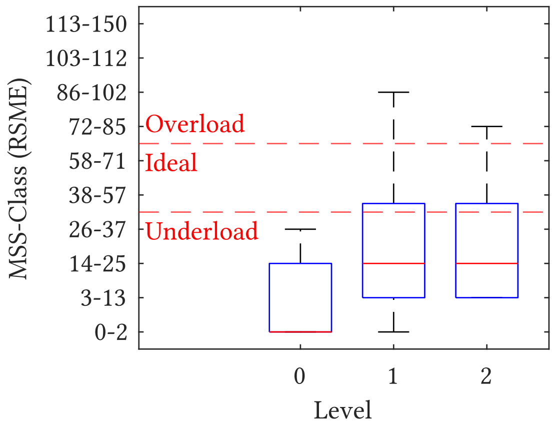


Figure 8.3.: RSME of underload scenarios that did not reach the ideal state

For the “Ideal Drive” scenario, the progression of RSME values differs notably from that of the underload scenario. Figure 8.4 illustrates this trend. In this scenario, eight scenarios did not reach the ideal state. In Level 0 (initial drive) the median RSME value lies within the expected MSS class “38–57”. With an increasing number of task recommendations, the distribution of values changes. In Levels 1 and 2 a significant increase in the standard deviation is observed, while the median increases only slightly. In Level 2 the RSME values show a wide dispersion across the entire observed range. It should be noted that the values extend exclusively across the underload and overload ranges.

In the “Overload Drive” scenario, the RSME progression is less consistent (see fig. 8.5). While in Level 0 the median, corresponding to the MSS class “86–102”, lies as expected in the overloaded range, the value changes as task recommendations are introduced. In Level 1 the RSME values vary considerably, and the median decreases to the classes “3–13” and “14–25”. In Level 2 an overshoot of the MSS is observed. The median increases sharply to the class “72–85”, with one outlier even staying in the underload range.

To better understand why certain scenarios did not end in an ideal state, the baseline RSME values (Level 0) are analyzed for each scenario type. The following figures compare RSME values between two groups: on the left,

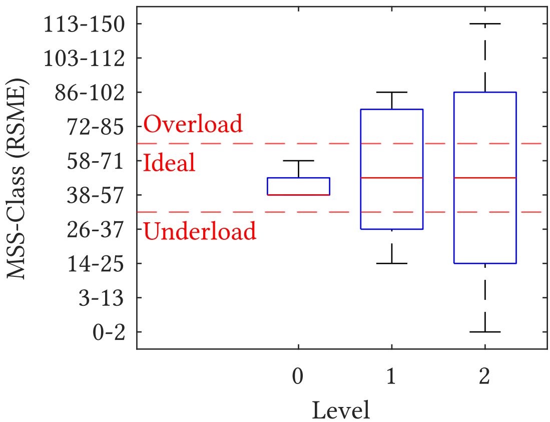


Figure 8.4.: RSME of ideal scenarios that did not reach the ideal state

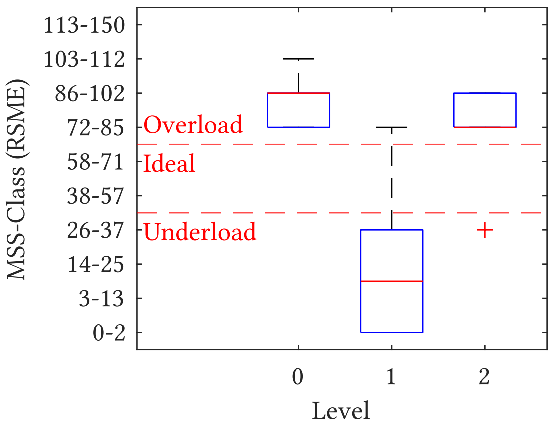


Figure 8.5.: RSME of overload scenarios that did not reach the ideal state

participants who reached an ideal state by the end of the scenario, and on the right, those who did not. Figure 8.6 shows this comparison for the underload scenarios. A direct comparison reveals that the median RSME value in the non-ideal scenarios is one class lower than the ideal scenarios ("3–13" vs.

"0–2"). However, standard deviation and overall dispersion show no notable differences between the two groups.

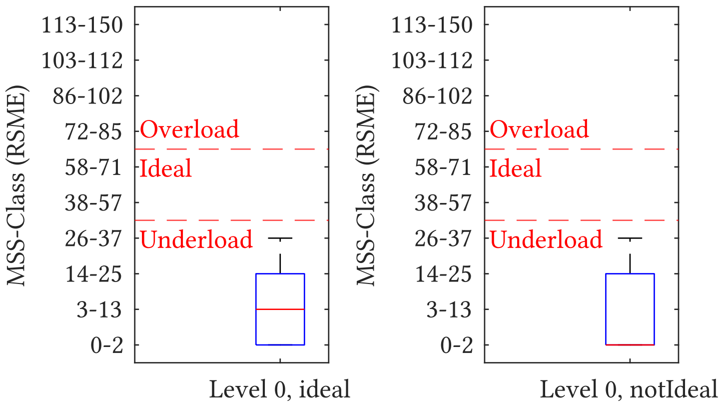


Figure 8.6.: RSME in Level 0 of underload scenarios

In Figure 8.7, no substantial difference between the two groups for the ideal scenarios is evident. The median is identical in both cases.

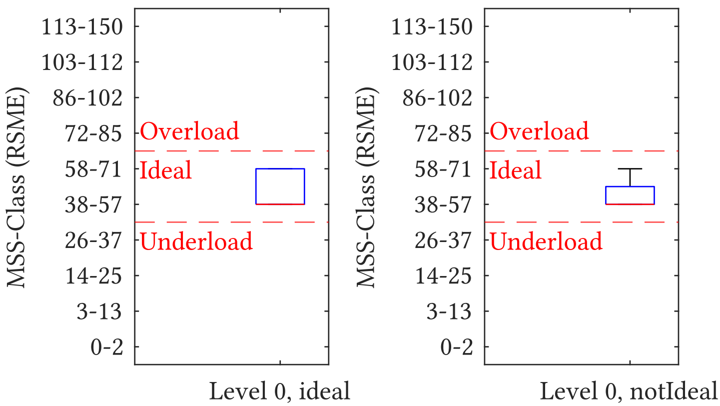


Figure 8.7.: RSME in Level 0 of ideal scenarios

In Figure 8.8, the median RSME values of the overload scenarios in Level 0 are identical, but both the standard deviation and total variance show slight differences between the two groups.

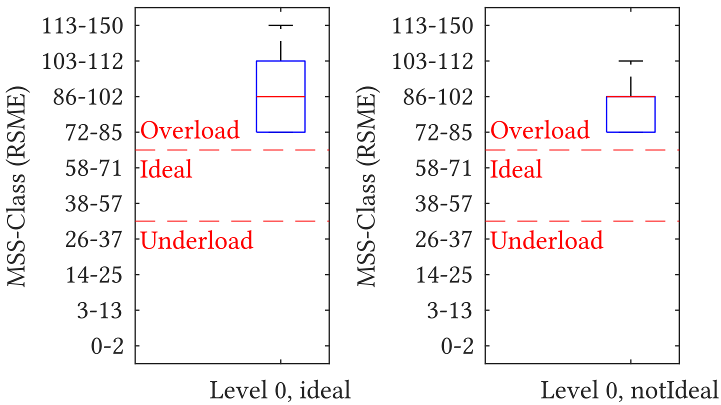


Figure 8.8.: RSME in Level 0 of overload scenarios

To analyze whether different task recommendations might have contributed to some scenarios not ending in an ideal state, Figure 8.9 illustrates the distribution of suggested task recommendations for ideal and non-ideal scenarios.

The focus here is on cases where participants were initially underloaded. This analysis is not easily possible for overload states because the Scenario Drive (SD) is directly manipulated in these cases, making the results highly dependent on the previous level. Since fewer participants took part in the verification study compared to the validation study, this comparison is more relevant and clearer in the validation study. The figure shows the ten most frequent TRs or combinations of TRs. The distribution of activating task recommendations shows no notable differences between the different scenario types.

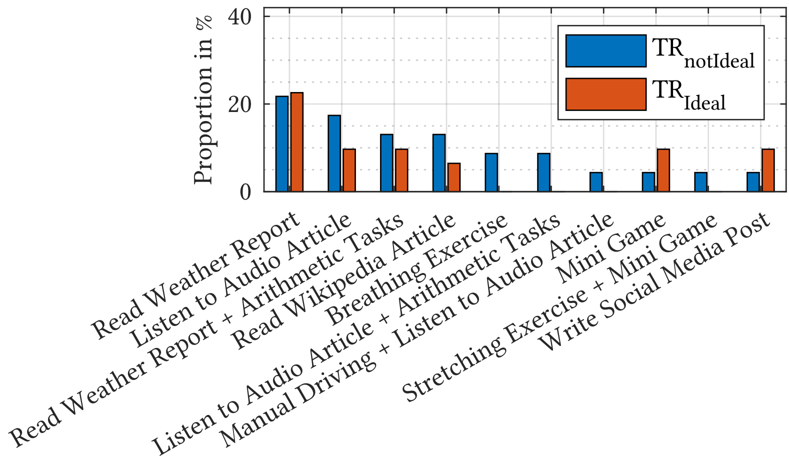


Figure 8.9.: TR Distribution

8.1.3.2. Validation study

Lastly, the validation study is examined, and the results are analyzed.

MSS monitoring:
To re-evaluate the results of the MSS monitoring studies from section 5.2 under new conditions and with an extended scope, the predicted values of the MSS monitoring were compared with the participants' reported RSME values. The reported and predicted values of all levels are taken into account.

Since the mental strain monitoring model from chapter 5 was trained with a total of ten classes, it seems reasonable to first analyze the results across all classes. Figure 8.10 illustrates these results in the form of a confusion matrix. The same metrics are used to achieve comparable results.

The detection rate, Recall or True Positive Rate (TPR) for the underload MSS classes "3–13" and "14–25" is 31.8% and 32.4%, respectively, while the Precision (Positive Predictive Value (PPV)) of these classes is 15.4% and 12.2%. The underload classes are frequently overpredicted, resulting in many false-positive outcomes. In particular, ideal RSME values are often incorrectly assigned to underload classes. This is also reflected in the TPR of the ideal

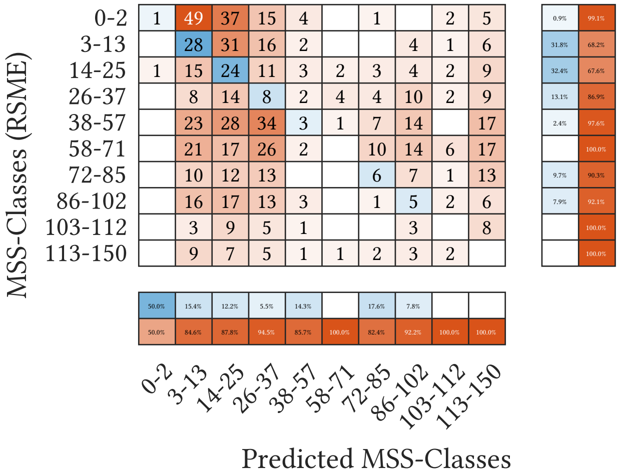


Figure 8.10.: Confusion chart MSS monitoring with ten classes

classes, which is very low with 2.4% for class "38–57" and 0% for class "58–71". Similarly, the two highest overload classes "103–112" and "113–150" are never correctly detected (0%). Overall, 75 out of 761 RSME values were correctly classified, corresponding to a total TPR of 9.9% and an overall macro-averaged F_1-score of 0.072.

By reducing the original ten classes to three broader classes "0–37", "38–71", and "72–150", the detection rate improves considerably. The overall TPR increases to 42.44% while the macro-averaged F_1-score increases to 0.315. Figure 8.11 presents the confusion matrix that visualizes this reduction in detectable classes.

A clear overprediction of class "0-37" can be observed. Furthermore, there are significant issues in correctly detecting the ideal condition class "38-71".

The detection metrics of the individual classes vary considerably. In Table 8.2, the most commonly used metrics are compared in an overview.

Selection of TRs by the VA:
In the course of the validation study, there were a total of 514 opportunities during 257 scenarios for the VA to issue a TR. A detailed analysis reveals the

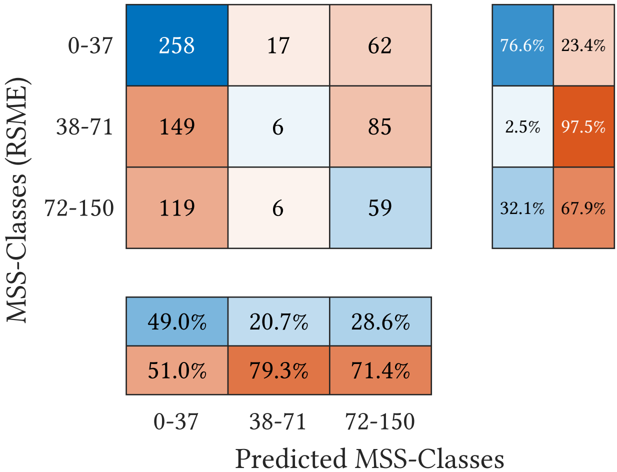


Figure 8.11.: Confusion chart MSS monitoring three classes

Table 8.2.: Detection metrics of the three classes

	TPR	PPV	F_1-score
Underload class “0-37”	76.6%	49.0%	0.598
Ideal class “38-71”	2.5%	20.7%	0.045
Overload class “72-150”	32.1%	28.6%	0.303

following: the VA issued task recommendations in 298 cases when participants were in an underload condition. In 189 situations, an overload was detected, prompting the VA to reduce the amount of issued tasks. The remaining 27 cases recorded no recommendations, as the participants were already in an ideal cognitive state. The number of situations where no modification by the VA was required is significantly lower compared to the previous study. This is due to the strong underprediction of the ideal class by the MSS monitoring. Since all task recommendations given or withheld by the VA corresponded correctly to the respective cognitive state class, the VA achieved a decision accuracy of 100%. This means that out of the 514 possible cases, the VA made the correct decision 514 times.

Effectiveness of the VA:
The analysis of the MSS monitoring system in this study shows that mental strain cannot be reliably detected. Therefore, the participants' reported RSME values are additionally considered. To avoid the influence of misclassified mental strain levels, all scenarios with mismatched classifications were excluded. Consequently, only scenarios in which the monitored mental strain prior to task recommendation selection aligned with the participants' reported RSME values were considered. This procedure reduces the number of scenarios available for analysis from 119 to 80 underload drive scenarios, from 85 to 21 ideal drive scenarios, and from 53 to 13 overload drive scenarios. Of these 114 considered scenarios, an ideal mental strain state was achieved in 29 scenarios (24.17%) at Level 1. This amount increased to 52 scenarios (45.61%) at Level 2. Broken down by scenario type, the following distribution is observed: in 37 scenarios (46.25%) of the underload scenarios, an ideal condition was reached after Level 2. Eleven scenarios (52.38%) of the ideal scenarios and four scenarios (30.77%) of the overload scenarios ended in an ideal condition.

Figure 8.12 shows the development of RSME values for participants during an underload scenario that did not reach an ideal mental strain state. Of the total 80 recorded underload scenarios, 43 were affected. Similar to fig. 8.3

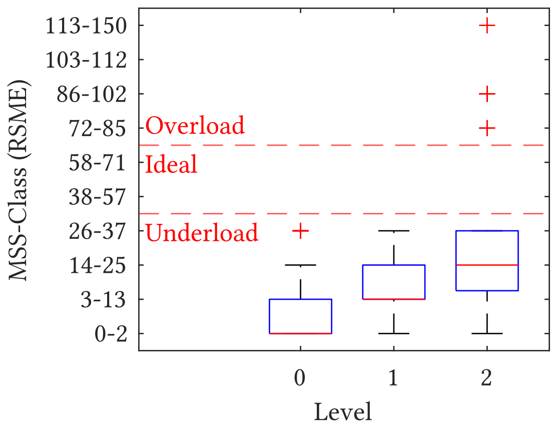


Figure 8.12.: RSME of underload scenarios that did not reach the ideal state

from the verification study, the RSME value increases with the number of levels. The median rises from class "0–2" (Level 0) through "3–13" (Level 1) to "14–25" (Level 2). In addition, three scenarios ending in overload can be observed at Level 2.

In fig. 8.13, the development of RSME values for participants during an ideal scenario that did not reach an ideal mental strain state is shown. A slight difference compared to the verification study is visible, as the medians of the reported RSME values vary more strongly across levels in this validation study. The figure includes the "ideal drives" scenarios in which no ideal mental strain state was achieved. In Level 1, the median decreases to class "14–25", while it slightly increases to "26–37" in Level 2. However, similar to fig. 8.4 from the verification study, the Interquartile Range (IQR) increases considerably after Level 0. The mental strain states separate into overload and underload states.

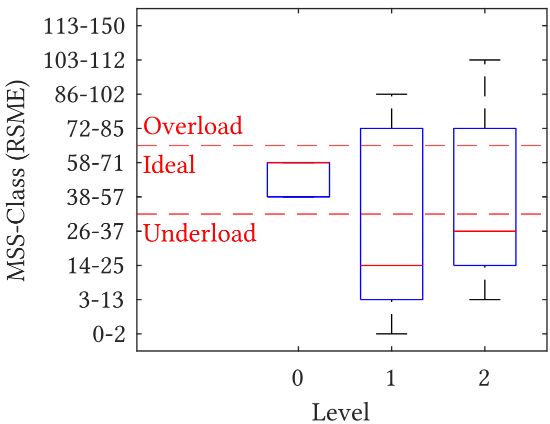


Figure 8.13.: RSME of ideal scenarios that did not reach the ideal state

Figure 8.14 shows "overload drive" scenarios that did not reach an ideal mental strain state. Compared to the verification study (fig. 8.5), the RSME trend decreases, and a larger IQR is observed, particularly in Level 2. The median drops to class "72–85" in Level 1 and thus remains in the overload region, and to "26–37" in Level 2.

The analysis of why a scenario failed to reach an ideal mental strain state follows the same procedure as in the verification study. The individual sce-

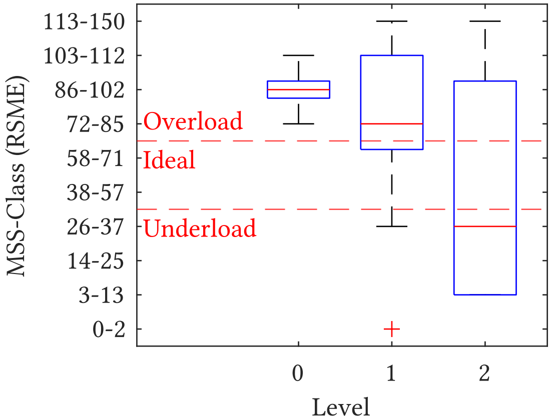


Figure 8.14.: RSME of overload scenarios that did not reach the ideal state

nario types are examined separately in the following sections. Again, only those scenarios are considered in which the monitored mental strain before the task recommendation selection matched the RSME values reported by the participants. To examine the influence of the initial drive (Level 0), the RSME values of the scenarios that ended in an ideal mental strain state are compared with those that did not.

The comparison of RSME values at Level 0 for the underload scenario (fig. 8.15) shows no major differences. The median of both groups lies in class “0–2”, with the range of values being slightly narrower for the scenarios that did not end in an ideal state.

A similar pattern is seen in fig. 8.16 for the “ideal drive” scenario. The IQR does not differ between the two groups, but the median of the scenarios that failed to reach an ideal mental strain state is slightly higher, at class “58–71”.

Similarly, for the overload scenario (fig. 8.17), no substantial differences are evident between scenarios that reached an ideal condition and those that did not. The median RSME value for scenarios that reached an ideal condition is slightly lower (“72–85” compared to “86–102”). These scenarios also show a wider spread and a larger IQR.

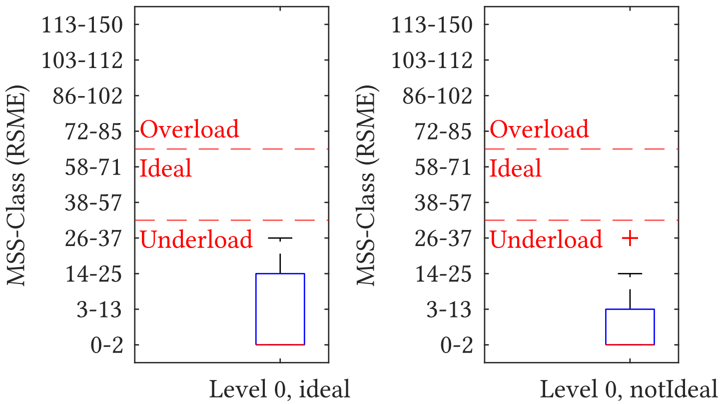


Figure 8.15.: RSME in Level 0 of underload scenarios

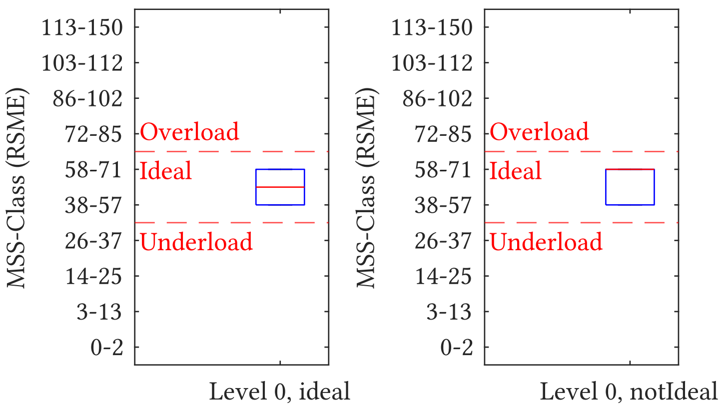


Figure 8.16.: RSME in Level 0 of ideal scenarios

Finally, it is examined whether differences in the selection of TRs could be a possible reason for the success rate of regulating the mental strain state. Figure 8.18 compares the distribution of TRs between both groups. The figure shows the ten most frequent TRs or combinations of TRs. There are no major differences in the selection of TRs between the two groups. In both,

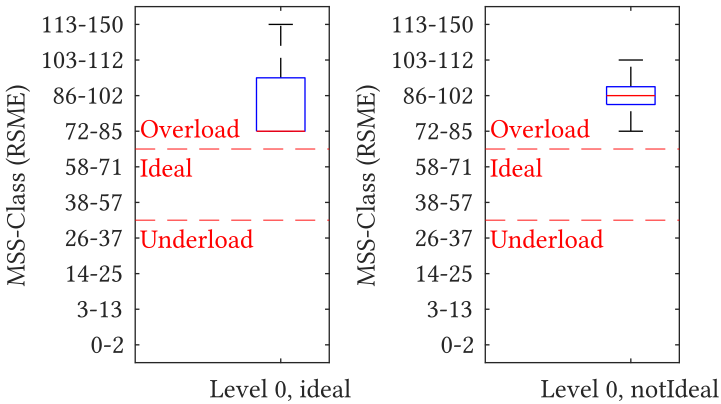


Figure 8.17.: RSME in Level 0 of overload scenarios

the task recommendations “Read Weather Report”, “Listen to Audio Article”, “Read Wikipedia Article”, and “Mini Game” are among the most frequently triggered. Furthermore, no outliers occur that appear exclusively in one of the two groups.

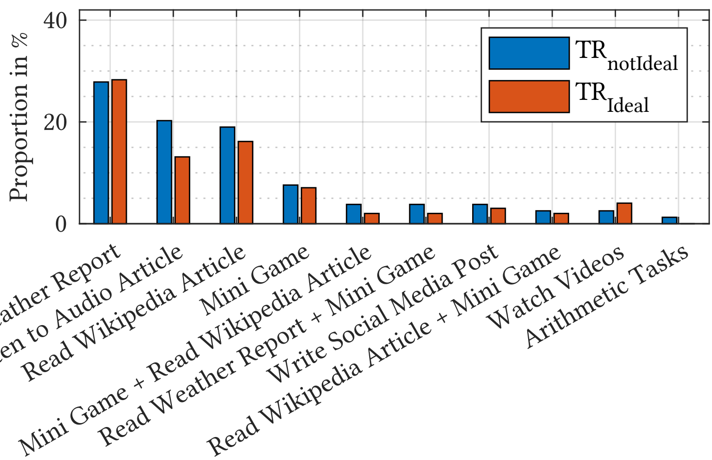


Figure 8.18.: TR Distribution

8.1.4. Discussion

In this section, the results of the verification and validation studies are related to each other in order to address and answer the initial research questions from section 8.1.1. The objective of the study was to investigate the interaction of the individual subsystems as well as the functionality of the overall system. For this purpose, one central research question and several sub-questions were defined. The aim was to examine to what extent the VA can influence the operator's mental strain and thus contribute to achieving an ideal state of mental strain. The first step was to analyze whether the implemented MSS monitoring system can operate with high accuracy under more realistic conditions so that it can serve as a reliable basis for the VA. Subsequently, the correct functioning of the decision logic of the VA and the underlying algorithm was to be verified. In addition, the effectiveness of the VA was to be quantified, and possible failures in regulating mental strain were to be analyzed.

The study was divided into two sub-studies, the verification study and the validation study, to better address different research questions and objectives. In the verification study, the subsystem for MSS monitoring could be explicitly excluded in order to verify the decision logic of the VA subsystem independently.

Monitoring of mental strain:
To determine whether the MSS monitoring system from chapter 5 also delivers satisfactory results in a more realistic context, the current mental strain of the participants was explicitly recorded after each level using the RSME. This allowed the MSS monitoring system to be independently verified.

Figure 8.10 clearly shows that reliable detection across all ten classes with which the model was trained is not possible. The detection rate (TPR) is only 9.9%. The MSS monitoring system shows particular difficulty in identifying ideal states, for which the detection rate is $\leq$ 2.4%. The extreme states "0–2", "103–112", and "113–150" are never correctly detected. This is mainly due to the small training dataset, as these classes occurred significantly less frequently than others. This is also consistent with the results from section 5.2.

By reducing the ten classes to the three main classes “Underload” (“0–37”), “Ideal” (“38–71”), and “Overload” (“72–150”), the overall detection rate increases significantly to 42.44%. The underload state is detected best with 76.6%, which is also reflected in the F_1-score of 0.598, compared to the overload state with a TPR of 32.1% and an F_1-score of 0.303. The ideal state lags far behind with a TPR of 2.5% and an F_1-score of 0.045.

The results of this study show that the current state of mental strain monitoring can only be transferred to new samples and scenarios to a limited extent. Although the underload state can be detected relatively well, its Precision of 49.0% is not sufficiently high and results in many false detections. Since the Recall and Precision of the other classes are even lower, it can be concluded that the mental strain monitoring model still needs to be improved and, in its current form, is not yet ready for operational use. This becomes even clearer when comparing the macro-averaged F_1-score of 0.315 with the F_1-score of 0.595 from the study in section 5.2. It should be noted, however, that considerably more TRs were used in this study than in the study from chapter 5. In particular, a combination of multiple TRs was not considered in chapter 5. This could also affect the performance of the MSS monitoring system.

To improve the condition detection system so that it can also be applied to new samples and scenarios, and thereby answer the first sub-question, *“How can the results of the mental strain monitoring system from section 5.2 be reliably transferred to a new sample and new scenarios?”*, the following changes are proposed:

- The training of the mental strain monitoring model should be conducted with a significantly larger training dataset comprising more participants. All classes to be classified should be adequately represented. This can be achieved by adapting the experimental design to include a broader range of task difficulty levels to induce greater variability in mental strain.
- The demographic composition of the training dataset should be adjusted to match the target user group of the system or generalized entirely to avoid a biased training dataset. This is primarily due to the fact that ocular parameters change considerably with age [153], [154].
- The eye tracking system and its associated software should be reviewed with respect to potential updates and improvements. An enhanced detection algorithm or an improved optical performance of the lens

system could potentially enhance the reliability of the MSS monitoring system [10].

- The position and number of eye tracking cameras should be analyzed and optimized. During the study, noticeable differences in eye tracking accuracy and detection quality were observed depending on the participant's head position. The system frequently lost the participants' gaze when they tilted their head too far downward or to the side.

Once a new model has been trained under these revised conditions, the resulting model should be scientifically evaluated in a new validation study.

Reaction of the virtual assistance:
Both the verification and validation studies showed that the underlying algorithm and the method for selecting TRs functioned correctly. Out of a total of 682 opportunities to propose a TR, the VA selected 603 TRs or combinations of TRs and correctly decided that no TR was required in the remaining 79 cases. The VA took the current mental strain state into account when selecting TRs and appropriately distinguished between activating TRs and task reducing actions. The Cognitive Task Load (CTL) model described in chapter 7 was consistently taken into account. The second sub-question, *"Does the virtual assistance act appropriately in response to the monitored MSSs?"*, can therefore be answered positively.

Influence of the virtual assistance on the MSS:
To answer the research question regarding the effectiveness of the VA and its ability to regulate users' mental strain in a targeted way, the results of the verification and validation studies are considered in combination.

The effectiveness of the VA, defined as the ratio of the number of scenarios that reached the ideal state in Level 2 to the total number of scenarios, differs markedly between the two studies. In the verification study, the success rate was 67.85%, which is more than 20% higher than in the validation study, where it was only 45.61%. The distribution of success rates across scenario types also differs markedly. Whereas the "Overload Drive" yielded the highest success rate in the verification study (82.75%), it exhibited the lowest rate in the validation study (30.77%). In contrast, the "Underload Drive" enabled a substantially larger proportion of participants to reach the ideal state (46.25% in the validation study; 62.96% in the verification study). The "Ideal Drive" scenario performed most consistently, achieving success rates of 52.38% in the validation study and 57.14% in the verification study.

One possible cause for this discrepancy could be the quality of the MSS monitoring. Although scenarios in which the detected user state did not match the self-reported RSME values from the previous level were excluded from the effectiveness analysis, potential effects on participants and their RSME ratings cannot be ruled out. It is plausible that frequent misclassifications of the MSS and the resulting incorrect selection of TRs reduced participants' trust in the system. Consequently, participants may have reported more extreme RSME ratings even in scenarios in which no misclassifications occurred. This assumption is supported by a comparison of RSME trends across levels: in the validation study, the values exhibit a wider spread across nearly all levels compared to the verification study. This is particularly evident in the underload (fig. 8.3 and fig. 8.12) and overload scenarios (fig. 8.5 and fig. 8.14).

In contrast, the selection of specific TRs differs little between the studies or between successful and unsuccessful scenarios. Therefore, the choice of TRs is not considered a likely cause of the failure to reach an ideal state.

The initial RSME values (Level 0) also show no substantial differences between the two studies. Likewise, there are no major differences in the initial RSME values between scenarios that achieved an ideal state at the end of the scenario and those that did not. It can thus be concluded that differences in initial conditions, such as participants with generally higher RSME ratings, cannot explain the differences between the studies or the failure to reach an ideal state.

Despite the identified differences between the two studies, the results demonstrate that it is fundamentally possible to influence and improve users' mental strain in a targeted manner. Since neither the selection of TRs nor the initial RSME values explain the observed differences, accurate MSS detection emerges as the key prerequisite for consistent success. Consequently, the research question *"To what extent can the virtual assistance deliberately influence the current MSS and thereby contribute to an ideal state?"* can be answered by stating that a targeted influence is possible in principle, while its reliable realization depends critically on successful MSS detection.

Moreover, the RSME progression of scenarios that did not reach an ideal state (fig. 8.3–fig. 8.5 and fig. 8.12–fig. 8.14) show that in underload scenarios, the VA tends to undercorrect, whereas in overload scenarios, it tends to overcorrect. This means that in TRs selected in the case of underload, the activating effect is too low, whereas the measures against overload have an excessively calming effect. This indicates that the currently available TRs do not provide sufficient diversity with respect to the LIP parameter, leading to recommendations that may not have been optimally aligned with participants' cognitive needs. To address this issue, the pool of TRs should be expanded to include options with more extreme LIP values, along with a finer granularity that more precisely accommodates the individual cognitive requirements of users. In addition, as already mentioned in the third research question, the initial RSME values in Level 0 and the selection of specific TRs do not differ between successful and unsuccessful scenarios. Consequently, the fourth and final research question *"What differences can be observed between scenarios that ended in an ideal MSS and those that did not?"* can be answered by concluding that unsuccessful scenarios are characterized by systematic over- or undercorrection, highlighting the need for a more diverse and finely tuned set of TRs.

In summary, the findings show that the presented assistance system can regulate operators' mental strain in a targeted manner if its core subsystems are reliably designed and well aligned. While the decision logic of the VA and the underlying control model function correctly, the overall system performance is currently limited by insufficiently reliable MSS monitoring and an overly coarse set of TRs. Accurate and transferable MSS detection, combined with a more diverse and finely graded pool of TRs, emerges as the decisive factor for consistent success under realistic conditions. Accordingly, the main research question *"How can an assistance system be designed to reliably regulate the mental strain state of agricultural machine operators?"* can be answered by concluding that reliable regulation requires dependable mental strain monitoring as a prerequisite, complemented by adaptive decision logic and a sufficiently granular pool of TRs that can be tailored to individual cognitive needs.

8.2. Field study

8.2.1. Introduction

The aim of the present field study was to evaluate the transferability of the developed assistance system to a real harvesting machine and to assess the user experience (UX) and the usability of the system under practice-oriented operating conditions.

For this purpose, the assistance system was installed on a “CLAAS LEXION 750 MONTANA”, which was equipped with standard machine assistance functions such as the header control “AUTO CONTOUR”, but did not feature advanced process automation (for example “CEMOS AUTOMATIC”). The virtual input and output interfaces of the assistance system were adapted to the existing series components of the machine to ensure seamless integration. Figure 8.19 shows the test vehicle used in the study.

Figure 8.19.: Test vehicle “CLAAS LEXION 750 MONTANA” [155]

Detailed information on the experimental setup, data collection, and complete results of this field study is documented in the final report of the project partner Institut für Agrartechnik (University Hohenheim) [155]. The following sections summarize the key findings and the basic structure of the study based on the published report [155].

8.2.2. Method

The field study was conducted as a qualitative experiment under real operating conditions in a combine harvester. All test runs were performed with the same driver to ensure consistent operation and comparability of the results. Insights that were not based on objectively collected data were derived from the subjective observations and experiences of the operator during the test runs.

8.2.2.1. Experimental design and material

A reduced assistance system adapted to the real machine environment was developed for the field study. It provided task recommendations while taking the operator's mental strain into account. To achieve this, the machine was connected to the VA, the required triggers for warnings and information messages were implemented, and interaction was made possible through several user interfaces. The essential adaptations to the assistance system and the machine can be summarized as follows:

- Integration of a joystick steering system on the left armrest to replicate the demonstrator cabin, including the option to switch between manual steering and Global Navigation Satellite System (GNSS)-based automated steering.
- Use of touch displays and a laptop for operating the VA, including interfaces for data acquisition and for displaying task recommendations.
- Integration of external displays for visualizing the machine silhouette, status information, and task recommendations.
- Setup of an dedicated network with a router and internet connection to connect all system components.
- Use of a MATLAB-based server for the VA and a CAN data gateway to capture and convert machine signals in real time.
- Implementation of a mental strain assessment system for continuous assessment of the user's mental stress, implemented in the 2024 harvest season using a haptic RSME slider.

- Integration of triggers for critical machine situations (for example grain tank fill level, grain losses, returns), supplemented by acoustic signals to reduce visual load.

The field tests were conducted during the 2023 and 2024 harvest seasons at two locations with an average field size of 4 ha:

- **Ihinger Hof, Renningen:** Total area 60.8 ha, field sizes above 10 ha
- **Meiereihof, Hohenheim:** Total area 18.3 ha, field sizes below 10 ha

Winter barley, rapeseed, spring barley, winter wheat, and triticale were harvested during the tests. Machine data, operator inputs, and the operator's subjective mental strain in the form of the RSME value were continuously recorded during operation. In the 2024 harvest season, mental strain was measured using a haptic slider, replacing the 2023 method of 120-second interval acoustic prompts, in order to minimize workflow interruptions.

8.2.2.2. Experimental procedure

During the test runs, the driver regularly interacted with the VA. It provided context-dependent task recommendations, for example for optimizing machine settings or querying current weather data. In critical operating situations, such as high grain tank fill levels or increased grain losses, the VA supplemented visual information with acoustic signals.

The operator's subjective MSS was continuously collected using the RSME slider on the left armrest. The entered values were transmitted to the VA in real time and served as an input for the adaptive selection of task recommendations. The evaluation of the RSME data was conducted only for time spans in which the working systems were active and the engine speed was high. Downtimes, such as machine transfer or setup operations, were not considered. The subjective mental strain level was entered by the operator using manual RSME input. Although this did not constitute a continuous measurement, the operator made sure to regularly assess their own mental strain.

In parallel, the standard machine functions were operated via the "CEBIS" display unit of the combine harvester. Header control, reel adjustment, cleaning, and driving speed were controlled through this interface. All relevant process and machine data, including CAN signals, interaction events, and mental strain data, were automatically recorded throughout the experiment. After each test run, data were backed up and the recorded measurements were prepared for the subsequent analysis.

This standardized procedure ensured that both the technical functionality of the assistance system and the interaction between the driver and the VA under real harvesting conditions could be fully captured and analyzed.

8.2.3. Results

Analysis of the datasets from the two fields showed that the reported mental strain values were predominantly within the medium, optimal range (RSME "38–71"). Lower values below 38 occurred only sporadically and were typically triggered intentionally to test the reaction logic of the virtual assistance. Higher values above "71" were observed only in isolated periods. The distribution of mental strain states for a single test run is shown in fig. 8.20. Detailed results for all test runs are documented in [155].

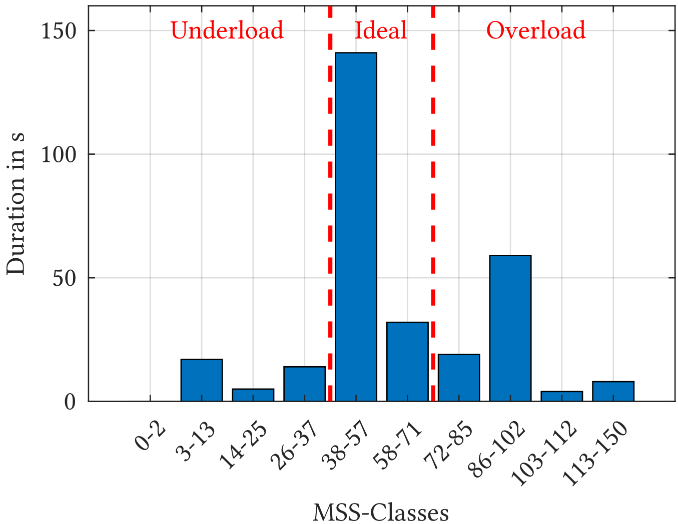


Figure 8.20.: Duration of different MSSs according to Böttinger [155]

The temporal analysis of the RSME values showed that the operator worked within an optimal mental strain range during most of the harvesting period.

Underload phases were documented only in a few measurements, while overload occurred only occasionally. The virtual assistance consistently reacted to both the entered RSME values and the recorded usage frequencies of machine functions. All system components operated reliably during the field tests.

The connection to the CAN bus and the transmission of machine data via the CAN-WebSocket gateway were stable, ensuring that all relevant machine parameters were available to the virtual assistance in real time. Differences between the signal definitions of the real machine and the demonstrator required occasional adjustments of the algorithm.

The integration of additional display and control elements into the cabin at series production level was successfully implemented. The touch display, which was mounted on the armrest, was positioned in such a way that the view of the header remained largely unobstructed. The additionally implemented joystick steering enabled complete control of the combine harvester without using the traditional steering wheel.

Acoustic and visual signals were correctly triggered in all critical situations. The local system architecture ensured stable data processing with minimal latency. Online access was carried out situationally via mobile or satellite-based internet connections.

8.2.4. Discussion

The field study shows that a reduced version of the assistance system implemented in the demonstrator cabin could be successfully developed and integrated into a cabin at series production level. This required both hardware and software adaptations, particularly with respect to network connectivity and the structural design of the overall system. Mental strain monitoring was performed manually through subjective user input. This approach is practical, but only partially reliable, as it requires driver attention and does not permit continuous data collection. Improvements can be achieved by integrating an automated, continuous real-time mental strain monitoring system (see chapter 5). [155]

The virtual assistance reacted appropriately to states of underload and overload, while no additional interventions or interactions occurred during phases of optimal mental strain. Overall, the system demonstrated stable and expected behavior. Observations during the field trials suggest that interacting with the system had a positive effect on the operator's subjective well-being, particularly due to the clear and comprehensible acoustic notifications and warnings. There is potential for further optimization through future integration of the assistance system into farm management systems, for example by linking processes of harvesting, feeding, or logistics control. [155]

9. Scientific contribution

To contextualize and evaluate the scientific contribution of the present work, the research questions formulated in chapter 3 are answered in the following. In doing so, the answers and insights from the corresponding chapters are summarized with a focus on the scientific contributions. Subsequently, the underlying research hypothesis is examined.

1. *"How can critical situations in harvesting operations be reliably detected by the assistance system?"*
   To answer this research question, two approaches for cost-effective environment monitoring are presented in chapter 4. The retrospective approach identifies mentally demanding phases such as turning maneuvers and obstacles and provides a usable prediction of the available time to perform Task Recommendations (TRs). The remote sensing based approach detects field boundaries and obstacles using Sentinel-2 data without requiring onboard sensors or training data. While its performance depends on satellite quality, seed point selection and parameter settings and may show leakage or undersegmentation in complex fields, it offers early awareness of points of interest. Overall, neither approach alone ensures sufficient reliability to serve as a standalone system for automated harvesting. However, they can provide useful supplementary information or, in combination with additional machine-based or drone-based sensor systems, form a practical, robust, and precise application.

2. *"How can the mental strain state of an agricultural machine operator be monitored or classified?"*
   In chapter 5, a study was conducted to answer this research question, forming the basis for the development and validation of a model for mental strain monitoring. As part of the study design, three sub-research questions were defined to support the answer to the main question. The data basis consisted of a combination of ocular and

cardiovascular signals, which were recorded in real time using an eye-tracking system and a fitness tracker. The developed model uses a Decision Tree machine learning algorithm to classify the individual mental strain states. To achieve the best results, the model was trained with ten classes and subsequently reduced to three classes for prediction. This approach demonstrated that mental strain states can be detected with an F_1-score of 0.5949, confirming that the selected sensor combination and classification method are fundamentally suitable. The applied methods and the results of the sub-research questions are presented in chapter 5.
However, the study in section 8.1 also revealed limitations in transferability. When the model was applied to a new context in the overall system study (section 8.1), the classification performance decreased substantially, achieving an F_1-score of only 0.315. This indicates that the current implementation may not generalize well across different operators or operational scenarios. Consequently, further improvements are necessary before the system can be considered ready for operational use, including larger and more balanced training datasets, demographic adaptation, and optimization of the eye-tracking setup. These refinements are expected to enhance both the reliability and practical applicability of the mental strain monitoring system.

3. *"How can the influence of Task Recommendations (TRs) on mental strain be reliably quantified and generalized?"*
   To answer this research question, a study on the Virtual Assistant (VA) was carried out in chapter 7, in which three sub-research questions were defined. The influence of the historical mental strain state as well as the possibility of categorizing and generalizing selected TRs were examined. The study demonstrates that the effects of individual TRs on mental strain can be quantified, differentiated, and comparatively ranked. For this purpose, the TRs are divided into different groups using the Compact Letter Display (CLD) method. These groups were subsequently assigned a Level of Information Processing (LIP) value for further use in the assistance system. Partial generalization of TRs across categories is feasible, with categories like "Office", "Relaxation" and "Infotainment" showing consistent effects, though TRs of other categories such as "Entertainment" reveal divergent effects due to differences in cognitive demand. Limitations include the restricted number of investigated TRs and the need to consider cognitive de-

mand, individual preferences, and situational factors in future work to enhance reliability and generalizability.

4. *"How can an assistance system be designed to reliably regulate the mental strain state of agricultural machine operators?"*
   In section 8.1, an overall system study was conducted to answer this research question, from which four sub-research questions were derived. For this purpose, a verification study and a validation study were performed. The sub-research questions included the monitoring of operator's mental strain, the responsiveness of the VA, the influence of the VA on the current mental strain state, and a comparison between successful and unsuccessful regulation attempts of the VA. The results show that the developed assistance system is fundamentally capable of classifying the mental strain state of an agricultural machine operator and regulating it through adaptive TRs. While the decision logic of the VA and the underlying control model function correctly, the overall system performance is currently limited by insufficiently reliable MSS monitoring and a coarse pool of TRs. Accurate and transferable MSS detection, combined with a more diverse and finely graded pool of TRs, represents the decisive factor for achieving consistent success under realistic conditions. Furthermore, section 8.1 provides suggestions for improvement and identifies potentials to make mental strain monitoring more reliable and precise in the future. In section 8.2, a potential practical transfer is additionally described based on a field study with a real combine harvester, in which optimization potentials for practical application are highlighted.

In summary, the findings of this work demonstrate that the formulated research questions can be answered in a manner that supports the research hypothesis. The first question showed that critical situations in harvesting operations can be detected using both retrospective and remote sensing approaches, providing useful information for the assistance system. The second question confirmed that mental strain can be monitored and classified using a combination of ocular and cardiovascular signals, while also identifying limitations and areas for improvement. The third question established that TRs can be quantified and partial generalized to influence mental strain. Finally, the fourth question demonstrated that an assistance system integrating these components is fundamentally capable of regulating the operator's cognitive strain. Finally, the formulated research hypothesis can be confirmed based on the detailed answers to all research questions:

"It is possible to deliberately influence the cognitive strain of an agricultural machine operator by means of an assistance system."

The presented assistance system shows, through the combination of ergonomics and engineering methods, that it is not only possible to monitor a machine operator's mental strain but also to regulate it using context-related task recommendations. While individual components have limitations in isolation, their integration within the system enables practical and effective regulation of cognitive strain. The results indicate that the approach works in principle, while also revealing several areas for improvement and providing concrete suggestions for further development. Overall, the system has the potential to enhance both the operator's well-being and performance, contributing not only to future scientific research but also offering economic and ergonomic value.

10. Summary and outlook

The present work addresses the development of an adaptive assistance system for agricultural machinery. The aim is to reliably monitor the Mental strain state (MSS) of machine operators and to regulate it through targeted task recommendations. An interdisciplinary research approach was applied for this purpose.

The decline in young farmers and the decreasing number of agricultural businesses lead to increasing work intensification and rising mental stress among employees. At the same time, the task profile of machine operators changes fundamentally due to the use of highly automated harvesting machines. Manual control tasks increasingly move into the background, while monitoring and supervisory activities dominate. This results in alternating phases of overload and underload, which can impair both the performance and the well-being of operators. Balanced mental strain becomes a central objective in modern work system design. Adaptive technical systems that monitor the cognitive state of an operator and provide situational support offer new opportunities to enhance safety, efficiency and satisfaction. Against this background, this work aims to develop a scientifically grounded, adaptive assistance system that monitors, classifies and actively regulates the mental strain of a machine operator. A highly automated combine harvester serves as the application example. The developed methods were experimentally examined and validated using an immersive demonstrator cabin with environmental simulation.

At the beginning, a detailed presentation of the theoretical background and an extensive literature review revealed a clear research demand. The literature shows that while numerous approaches exist for measuring mental strain, no holistic system is currently available that integrates monitoring, assessment and targeted regulation of mental strain in a real and complex work system such as an agricultural machine. Based on the resulting research hypothesis, further research questions were derived, which were then used to evaluate the hypothesis.

Subsequently, methods for cost-effective detection of field contours and obstacles based on GNSS and satellite data were presented. Although these methods establish a foundation for a safe machine environment, they must be combined with additional sensor systems to improve robustness and cannot operate reliably as standalone solutions. On their own, neither approach provides sufficient reliability for fully autonomous harvesting. Nevertheless, they can supply valuable supplementary information or, when integrated with other machine-based or drone-based sensor systems, support a practical, robust, and precise implementation.

In the further course, a system was developed for monitoring the mental strain. Using a specially designed demonstrator cabin, eye movements and heart rate were monitored to assess mental strain. These data were subsequently combined with subjective self-report mental strain data using the Rating Scale Mental Effort. Based on these data, a machine learning model was developed, which recognizes the cognitive state with three classes.

The developed assistance system consists of modular software components for data acquisition, decision logic and user interaction. It considers the current mental strain, mental stress, individual preferences and current environmental conditions to derive situation-specific task recommendations. The selection of these task recommendations is based on the Cognitive Task Load model. Interaction takes place multimodally via voice and touch interfaces in the form of a virtual assistant. In an experimental study, different task recommendations were tested to quantify their effects. The results show that task recommendations measurably influence the mental strain and can be generalized into groups.

Finally, the complete assistance system was tested in a comprehensive experimental study in the demonstrator cabin. The results show that the virtual assistant can recognize and regulate the cognitive state of operators. A subsequent field study on a real combine harvester highlighted the potential for practical application and identified opportunities for optimization regarding system integration and user acceptance. The results were then discussed with respect to the research questions and the underlying hypothesis. All four research questions could be answered positively, and the hypothesis is considered confirmed.

A central starting point for future work lies in expanding the data basis and increasing model reliability of the mental strain state monitoring system. The sensor methods employed in this dissertation, particularly cardiovascular and physiological measurements, proved suitable for reliably monitoring mental strain. Nevertheless, there is potential to further increase the prediction accuracy of the classification models by integrating additional sensor technology such as EEG and skin conductance, if in the future these are less invasive and uncomfortable for the operator, or by using machine-internal process data. The integration of such multimodal data sources could provide a more comprehensive picture of the user condition, enabling a more precise and stable monitoring of the mental strain state. This improvement could also make it possible to select more appropriate task recommendations and to better regulate the cognitive state of the operator. Additionally, the composition of the training dataset should generally be adjusted and demographically expanded. In the presented studies, the participants were predominantly young individuals, although age has a considerable influence on ocular parameters [153], [154].

Another key area of future research concerns the generalizability and transferability of the developed methods. Validation was carried out using a combine harvester, ensuring a direct reference to agricultural practice. Future studies should transfer the assistance system to other machine and work types such as tractors, field sprayers or forage harvesters to examine cross-machine applicability. Furthermore, long-term studies are required to analyse how adaptive mental strain regulation affects performance, fatigue and operator well-being over extended work shifts and harvesting seasons.

A third field of development concerns the further improvement of interaction possibilities of the virtual assistant. The combination of voice and touch interfaces implemented in this work forms a solid basis for user-centered communication. Future work should investigate how context-adaptive dialogue systems and adaptive interaction models based on modern language processing, for example through large language models, can be integrated. This could enable more personalized and context-sensitive task recommendations. The integration of emotion recognition or adaptive feedback mechanisms would also be conceivable, allowing more empathic, human-like and comfortable system interaction.

In addition, ergonomic and ethical questions arise that must be increasingly considered in the future. As technical systems respond ever more to psychophysiological human states, questions of data sovereignty, transparency and acceptance become central. The development of suitable data protection concepts and ethical guidelines is therefore essential to ensure acceptance in practice and to support a responsible handling of sensitive human-machine data.

Furthermore, integration into existing operational and agricultural management systems represents a promising goal. Closer linkage to farm management software can improve data exchange between machine, work planning and operational management. This can expand the pool of task recommendations and thus increase the regulatory capability of the assistance system.

In the long term, the concepts developed in this work can be applied beyond agriculture. The fundamental principles of mental strain adaptive assistance can be transferred to other highly automated work environments, for example in industrial production, vehicle cockpits, emergency control centers or aviation. In all these areas, the combination of condition monitoring, adaptive mental strain regulation and user-centered interaction opens up new perspectives for safe, efficient and human-oriented work systems.

Bibliography

Own References

[1] S. Metzger, P. Lehr, and M. Geimer, "Beanspruchungsadaptive Nutzerschnittstelle für die vernetzte Landwirtschaft", *ATZheavy duty*, vol. 15, no. 1, pp. 48–52, Jan. 1, 2022, ISSN: 2524-8782. DOI: 10.1007/s35746-021-0475-6.

[2] S. Metzger *et al.*, "Entwicklung einer adaptiven Benutzerschnittstelle zur Optimierung des kognitiven Benutzerzustands", in *Arbeit Unter Einem DA-CH: Der Landwirt Im 4.0-Modus*, L.-I. für Agrartechnik und Bioökonomie e.V., Ed., Mar. 8, 2022, pp. 79–90.

[3] P. Lehr *et al.*, "Validierungsumgebung für ein neues Kabinenkonzept am Beispiel eines Mähdreschers", *ATZheavy duty*, vol. 16, no. 1, pp. 34–37, May 1, 2023, ISSN: 2524-8790. DOI: 10.1007/s35746-023-1015-3.

[4] S. Metzger and M. Geimer, "Field contour and obstacle detection for automated agricultural machinery: A multi-temporal segmentation method based on Sentinel-2 satellite data", *agricultural engineering.eu*, vol. 80, no. 1, 1 Jan. 27, 2025, ISSN: 2943-5641. DOI: 10.15150/ae.2025.3329.

[5] S. Metzger, P. Lehr, and M. Geimer, "Stress-adaptive User Interface for the Networked Agriculture", *ATZheavy duty worldwide*, vol. 15, no. 1, pp. 48–51, Jan. 1, 2022. DOI: 10.1007/s41321-021-0468-5.

[6] P. Lehr *et al.*, "Validation Environment for a New Cab Concept Using the Example of a Combine Harvester", *ATZheavy duty worldwide*, vol. 16, no. 1, pp. 34–37, May 1, 2023, ISSN: 2524-8774. DOI: 10.1007/s41321-023-1013-5.

[7] M. Geimer *et al.*, “Agrarsysteme der Zukunft: Fahrerkabine 4.0 - Entwicklung einer beanspruchungsadaptiven Nutzerschnittstelle für Landmaschinenbetreiber - Teilprojekt A: Abschlussbericht”, Hannover : Technische Informationsbibliothek, Nov. 30, 2024. DOI: 10.34657/23443.

Internet References

[8] Sinergise Ltd. “Sentinel Hub Brochure”. (2019), [Online]. Available: https://www.sentinel-hub.com/docs/Sentinel_HUB_Brochure_2019_NEW.pdf (visited on 01/16/2024).

[9] Deutscher Wetterdienst. “RCC Node-CM Produktbeschreibung - Wolkenbedeckung”. (Jun. 1, 2023), [Online]. Available: https://www.dwd.de/DE/leistungen/rcccm/int/descriptions/cfc/pds_cfc_de.pdf?__blob=publicationFile&v=16 (visited on 12/06/2024).

[10] Smart Eye AB. “Smart Eye Pro – Remote Eye Tracking System”. (2025), [Online]. Available: https://www.smarteye.se/smart-eye-pro/ (visited on 01/15/2025).

[11] Wahoo Fitness LCC. “TICKR FIT Heart Rate Monitor Information and Setup”, Wahoo Fitness Support. (2025), [Online]. Available: https://support.wahoofitness.com/hc/en-us/articles/4406623020690-TICKR-FIT-Heart-Rate-Monitor-Information-and-Setup (visited on 03/31/2025).

[12] Mozilla Foundation. “DeepSpeech”. (Aug. 18, 2025), [Online]. Available: https://github.com/mozilla/DeepSpeech (visited on 08/18/2025).

[13] L. Lane. “Nvidia NeMo Curator Audio Quality Metrics”. (Sep. 25, 2025), [Online]. Available: https://docs.nvidia.com/nemo/curator/latest/about/concepts/audio/quality-metrics.html?utm_source=chatgpt.com#word-error-rate-wer (visited on 11/09/2025).

[14] OpenWeather. “Weather API - OpenWeatherMap”. (2025), [Online]. Available: https://openweathermap.org/api (visited on 09/01/2025).

[15] 365FarmNet. “Digitalisieren Sie Ihre Landwirtschaft mit 365FarmNet”, 365FarmNet. (2024), [Online]. Available: https://www.365farmnet.com/en/ (visited on 09/01/2025).

[16] Dave Winer. “RSS 2.0 Specification (RSS 2.0 at Harvard Law)”. (Jul. 15, 2003), [Online]. Available: https://cyber.harvard.edu/rss/rss.html (visited on 09/01/2025).

[17] Alex Segler. “API.radio-browser.info docs”. (Oct. 6, 2025), [Online]. Available: https://api.radio-browser.info/ (visited on 09/01/2025).

[18] Spotify AB. “Web API | Spotify for Developers”. (2025), [Online]. Available: https://developer.spotify.com/documentation/web-api (visited on 09/01/2025).

[19] Matrix.org. “Matrix Specification”, Matrix Specification. (Jun. 26, 2025), [Online]. Available: https://spec.matrix.org/v1.15/ (visited on 09/01/2025).

[20] DeepL SE. “DeepL API: skalierbare Übersetzung und Textoptimierung”. (2025), [Online]. Available: https://www.deepl.com/de/products/api (visited on 09/01/2025).

[21] PONS Langenscheidt GmbH. “Wörterbuch-API für 20 Sprachen | PONS”. (2022), [Online]. Available: https://de.pons.com/p/online-woerterbuch/fuer-entwickler/api (visited on 09/01/2025).

[22] Daniel Naber. “Webservice / API-Zugriff - OpenThesaurus”. (2025), [Online]. Available: https://www.openthesaurus.de/about/api (visited on 09/01/2025).

[23] Wikimedia Foundation Inc. “REST API Documentation”. (2025), [Online]. Available: https://de.wikipedia.org/api/rest_v1/ (visited on 09/01/2025).

[24] The MathWorks, Inc. “MATLAB - Multiple comparison test - multcompare”. (2024), [Online]. Available: https://de.mathworks.com/help/stats/multcompare.html (visited on 09/13/2025).

Other References

[25] *Situationsbericht 2025/26 Trends und Fakten zur Landwirtschaft*, 1. Auflage. Berlin: Deutscher Bauernverband e.V, 2025, ISBN: 978-3-9820166-7-2.

[26] U. Hemmerling and P. Pascher, *Situationsbericht 2017/18: Trends Und Fakten Zur Landwirtschaft*, 1. Auflage. Berlin: Deutscher Bauernverband e.V, Jan. 1, 2017, 232 pp., ISBN: 3-9812770-9-0.

[27] P. Pascher *et al.*, *Situationsbericht 2020/21: Trends Und Fakten Zur Landwirtschaft.* Berlin: Deutscher Bauernverband e.V, Jan. 1, 2020, 239 pp., ISBN: 978-3-9820166-2-7.

[28] M. Geimer and C. Pohlandt, *Grundlagen Mobiler Arbeitsmaschinen* (Karlsruher Schriftenreihe Fahrzeugsystemtechnik). Hannover; Karlsruhe; Karlsruhe: Technische Informationsbibliothek u. Universitätsbibliothek KIT Scientific Publishing, Jan. 1, 2014, vol. 22, 228 pp., ISBN: 978-3-7315-0188-6. DOI: 10.5445/KSP/1000039443.

[29] U. Hartl, *Branchenanalyse Landtechnik: Entwicklungstrends Und Herausforderungen* (Working Paper Forschungsförderung 052). Düsseldorf: Hans-Böckler-Stiftung, Jan. 1, 2017, vol. 052, 50 pp.

[30] D. Waard, *The Measurement of Drivers' Mental Workload.* Groningen: Traffic Research Centre Univ. of Groningen, Jan. 1, 1996, 125 pp., ISBN: 90-6807-308-7.

[31] C. Schlick, R. Bruder, and H. Luczak, *Arbeitswissenschaft.* Berlin, Heidelberg: Springer, 2018, ISBN: 978-3-662-56037-2. DOI: 10.1007/978-3-662-56037-2.

[32] M. Schütte, *Psychische Belastung und Beanspruchung am Arbeitsplatz: inklusive DIN EN ISO 10075-1 bis -3*, 2., erweiterte Auflage, Deutsches Institut für Normung, Ed. Berlin Wien Zürich: Beuth Verlag GmbH, 2021, 1 p., ISBN: 978-3-410-30286-5.

[33] J. Schwarz and S. Fuchs, "Multidimensional Real-Time Assessment of User State and Performance to Trigger Dynamic System Adaptation", in *Augmented Cognition. Neurocognition and Machine Learning*, D. D. Schmorrow and C. M. Fidopiastis, Eds., Cham: Springer International Publishing, 2017, pp. 383–398, ISBN: 978-3-319-58628-1. DOI: 10.1007/978-3-319-58628-1_30.

[34] S. A. Shearer, S. K. Pitla, and J. D. Luck, "Trends in the Automation of Agricultural Field Machinery", in *Proceedings of the 21st Annual Meeting of the Club of Bologna*, Bologna, Nov. 13, 2010.

[35] S. K. Devitt. "Cognitive factors that affect the adoption of autonomous agriculture". arXiv: 2111.14092 [cs]. (Nov. 28, 2021), [Online]. Available: http://arxiv.org/abs/2111.14092 (visited on 10/25/2025), prepublished.

[36] B. Bashiri and D. D. Mann, “Automation and the situation awareness of drivers in agricultural semi-autonomous vehicles”, *Biosystems Engineering*, vol. 124, pp. 8–15, Jan. 1, 2014, ISSN: 15375110. DOI: 10.1016/j.biosystemseng.2014.06.002.

[37] J. Ahmann *et al.*, “Landwirtschaftliche Assistenzsysteme zur Entscheidungsunterstützung in der Nutztierhaltung”, *agricultural engineering.eu*, vol. 79, no. 2, Apr. 15, 2024, ISSN: 2943-5641. DOI: 10.15150/ae.2024.3305.

[38] W. Apt, M. Schubert, and S. Wischmann, *Digitale Assistenzsysteme: Perspektiven Und Herausforderungen Für Den Einsatz in Industrie Und Dienstleistungen.* Berlin, Jan. 1, 2018, 82 pp., ISBN: 978-3-89750-181-2.

[39] L. Sabattini *et al.* “Methodological Approach for the Design of a Complex Inclusive Human-Machine System”. arXiv: 1706.08461 [cs]. (Jun. 26, 2017), [Online]. Available: http://arxiv.org/abs/1706.08461 (visited on 10/25/2025), pre-published.

[40] Y. Forster *et al.*, “Driver compliance to take-over requests with different auditory outputs in conditional automation”, *Accident; analysis and prevention*, vol. 109, pp. 18–28, Jan. 1, 2017. DOI: 10.1016/j.aap.2017.09.019. PMID: 28992451.

[41] A. M. Klein *et al.*, “Design and Evaluation of Voice User Interfaces: What Should One Consider?”, in *Design, Operation and Evaluation of Mobile Communications*, G. Salvendy and J. Wei, Eds., Cham: Springer Nature Switzerland, 2023, pp. 167–190, ISBN: 978-3-031-35921-7. DOI: 10.1007/978-3-031-35921-7_12.

[42] R. Pereira *et al.*, “Virtual Assistants in Industry 4.0: A Systematic Literature Review”, *Electronics*, vol. 12, no. 19, p. 4096, Jan. 2023, ISSN: 2079-9292. DOI: 10.3390/electronics12194096.

[43] F. Li *et al.* “Virtual Co-Pilot: Multimodal Large Language Model-enabled Quick-access Procedures for Single Pilot Operations”. arXiv: 2403.16645 [cs]. (Mar. 25, 2024), [Online]. Available: http://arxiv.org/abs/2403.16645 (visited on 10/27/2025), pre-published.

[44] M. Fischer *et al.*, “TANGO - Technologie für automatisiertes Fahren nutzergerecht optimiert : Schlussbericht : Vorhabenbezeichnung: "Neue Fahrzeug- und Systemtechnologien" des BMWI; Verbundprojekt "Technologie für automatisiertes Fahren nutzergerecht optimiert

- TANGO" : Teilvorhaben: Arbeitswissenschaftliche Aspekte und HMI-Styleguide : Laufzeit des Vorhabens: 3,83 Jahre, 01. Dezember 2016 bis 30. September 2020", [Stuttgart], 2020. DOI: 10.2314/KXP:1762351587.

[45] P. Alt *et al.*, "TANGO - Technologie für automatisiertes Fahren nutzergerecht optimiert : Schlussbericht : Vorhabenbezeichnung: "Neue Fahrzeug- und Systemtechnologien" des BMWI; Verbundprojekt "Technologie für automatisiertes Fahren nutzergerecht optimiert - TANGO" : Teilvorhaben: Aufmerksamkeits- und Aktivitätenassistent : Laufzeit des Vorhabens: 3,5 Jahre, 01. Dezember 2016 bis 31. Mai 2020", [Abstatt], 2020. DOI: 10.2314/KXP:1760418196.

[46] M.-A. Büchner *et al.*, "Toward the Design of Personalised Adaptive Driver Assistance for Truck Docking:" in *Proceedings of the 3rd International Symposium on Automation, Information and Computing*, Beijing, China, China: SCITEPRESS - Science and Technology Publications, 2022, pp. 131–137, ISBN: 978-989-758-622-4. DOI: 10.5220/0011907100003612.

[47] H.-G. Lee, D.-H. Kang, and D.-H. Kim, "Human–Machine Interaction in Driving Assistant Systems for Semi-Autonomous Driving Vehicles", *Electronics*, vol. 10, no. 19, p. 2405, Jan. 2021, ISSN: 2079-9292. DOI: 10.3390/electronics10192405.

[48] D. S. S. Kumar *et al.*, "Ai based smart voice assistant for farmers", *International Journal of Emerging Technologies and Innovative Research*, vol. 11, no. 5, g587–g594, May 2024, ISSN: 2349-5162.

[49] J.-H. Kirchner, "Belastungen und Beanspruchungen — Einige begriffliche Klärungen zum Belastungs- Beanspruchungs-Konzept", in *Arbeitsorganisation und Neue Technologien: Impulse für eine weitere Integration der traditionellen arbeitswissenschaftlichen Entwicklungsbereiche*, R. Hackstein, F.-J. Heeg, and F. von Below, Eds., Berlin, Heidelberg: Springer, 1986, pp. 553–569, ISBN: 978-3-642-93339-4. DOI: 10.1007/978-3-642-93339-4_23.

[50] W. Rohmert, "Das Belastungs-Beanspruchungs-Konzept", *Zeitschrift für Arbeitswissenschaft*, vol. 38, no. 4, pp. 193–200, 1984, ISSN: 0340-2444.

[51] F. R. H. Zijlstra, *Efficiency in Work Behaviour: A Design Approach for Modern Tools.* Delft: Delft University Press, Jan. 1, 1993, 186 pp., ISBN: 90-6275-918-1.

[52] C. D. Wickens, "Multiple resources and performance prediction", *Theoretical Issues in Ergonomics Science*, vol. 3, no. 2, pp. 159–177, Jan. 1, 2002, ISSN: 1463-922X. DOI: 10.1080/14639220210123806.

[53] L. Longo *et al.*, "Human Mental Workload: A Survey and a Novel Inclusive Definition", *Frontiers in Psychology*, vol. 13, Jun. 2, 2022, ISSN: 1664-1078. DOI: 10.3389/fpsyg.2022.883321.

[54] R. M. Yerkes and J. D. Dodson, "The relation of strength of stimulus to rapidity of habit-formation", *Journal of Comparative Neurology and Psychology*, vol. 18, no. 5, pp. 459–482, 1908, ISSN: 1550-7149. DOI: 10.1002/cne.920180503.

[55] N. Backhaus, "Context Sensitive Technologies and Electronic Employee Monitoring: A Meta-Analytic Review", in *2019 IEEE/SICE International Symposium on System*, IEEE, Ed., 2019, pp. 548–553. DOI: 10.1109/SII.2019.8700354.

[56] M. A. Neerincx, "Cognitive task load analysis: Allocating tasks and designing support", in *Handbook of Cognitive Task Design*, E. Hollnagel, Ed., Mahwah, NJ: Lawrence Erlbaum Associates, Jan. 1, 2003, pp. 283–305.

[57] D. Beevis, "Analysis Techniques For Man-Machine Systems Design", North Atlantic Treaty Organization - Defence Research Group, Brussels, Technical Report AC/243(Panel 8)TR/7, Jul. 31, 1992, 331 pp.

[58] J. Rasmussen, *Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering* (North-Holland Series in System Science and Engineering). New York, N.Y.: North-Holland, Jan. 1, 1986, vol. 12, 215 pp., ISBN: 0-444-00987-6.

[59] P. Jeschke, *Entwicklung Eines Analytischen Modells Zur Prognose Der Mentalen Beanspruchung in Der Prozessführung*. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Jan. 1, 2017, 260 pp. DOI: 10.21934/baua:bericht20171011.

[60] S. Mühlbacher-Karrer *et al.*, "A Driver State Detection System-Combining a Capacitive Hand Detection Sensor With Physiological Sensors", *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 4, pp. 624–636, Apr. 2017, ISSN: 1557-9662. DOI: 10.1109/TIM.2016.2640458.

[61] M. Schneider and B. Deml, "Analysis of a Multimodal Human-Robot-Interface in Terms of Mental Workload", in *Advances in Ergonomic Design of Systems, Products and Processes*, C. M. Schlick *et al.*, Eds., Berlin, Heidelberg: Springer, 2017, pp. 247–260, ISBN: 978-3-662-53305-5. DOI: 10.1007/978-3-662-53305-5_18.

[62] M. Schneider and B. Deml, "An Integrated Approach of Mental Workload Assessment", in *Advances in Ergonomic Design of Systems, Products and Processes*, B. Deml *et al.*, Eds., Berlin, Heidelberg: Springer, 2016, pp. 191–208, ISBN: 978-3-662-48661-0. DOI: 10.1007/978-3-662-48661-0_13.

[63] C. D. Wickens *et al.*, *Engineering Psychology and Human Performance*, 4th ed. New York: Psychology Press, Aug. 20, 2015, 544 pp., ISBN: 978-1-315-66517-7. DOI: 10.4324/9781315665177.

[64] L. Longo, "Formalising Human Mental Workload as a Defeasible Computational Concept", Dissertation, Trinity College Dublin. Dublin, Oct. 1, 2014.

[65] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research", in *Advances in Psychology*, ser. Human Mental Workload, P. A. Hancock and N. Meshkati, Eds., vol. 52, North-Holland, Jan. 1, 1988, pp. 139–183. DOI: 10.1016/S0166-4115(08)62386-9.

[66] M. Kaczorowska, M. Plechawska-Wójcik, and M. Tokovarov, "Interpretable Machine Learning Models for Three-Way Classification of Cognitive Workload Levels for Eye-Tracking Features", *Brain Sciences*, vol. 11, no. 2, p. 210, 2 Feb. 2021, ISSN: 2076-3425. DOI: 10.3390/brainsci11020210.

[67] S. Solhjoo *et al.*, "Heart Rate and Heart Rate Variability Correlate with Clinical Reasoning Performance and Self-Reported Measures of Cognitive Load", *Scientific Reports*, vol. 9, no. 1, p. 14 668, Oct. 11, 2019, ISSN: 2045-2322. DOI: 10.1038/s41598-019-50280-3.

[68] Y. Funk, H. Haase, and B. Deml, "Echtzeiterfassung Psychischer Beanspruchungszustände", in *Technologie Und Bildung in Hybriden Arbeitswelten*, ser. Kongress Der Gesellschaft Für Arbeitswissenschaft 68, G. für Arbeitswissenschaft, Ed., Magdeburg: GfA-Press, Mar. 2, 2022, ISBN: 978-3-936804-31-7.

[69] Y. Funk *et al.*, "Entwicklung und Validierung einer Experimentalumgebung zur Messung mentaler Beanspruchungszustände", in *Frühjahrskongress Arbeit HUMAINE Gestalten*, G. für Arbeitswissenschaft, Ed., Dortmund, Mar. 3, 2021, ISBN: 978-3-936804-29-4.

[70] Y. A. Funk *et al.*, "Entwicklung und Validierung einer computerbasierten Aufgabe zur Induktion eines psychischen Beanspruchungsspektrums", *Zeitschrift fur Arbeitswissenschaft*, pp. 1–17, Jan. 1, 2022, ISSN: 0340-2444. DOI: 10.1007/s41449-022-00304-y. PMID: 35287339.

[71] "Heart Rate Variability and Mental Workload Assessment", in *Advances in Psychology*, vol. 52, North-Holland, Jan. 1, 1988, pp. 101–115. DOI: 10.1016/S0166-4115(08)62384-5.

[72] A. Henelius *et al.*, "Mental workload classification using heart rate metrics", in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Sep. 2009, pp. 1836–1839. DOI: 10.1109/IEMBS.2009.5332602.

[73] S. Miller, *Workload Measures.* Iowa City: University of Iowa, National Advanced Driving Simulator, Aug. 2001.

[74] Y. Lean and F. Shan, "Brief review on physiological and biochemical evaluations of human mental workload", *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 22, no. 3, pp. 177–187, 2012, ISSN: 1520-6564. DOI: 10.1002/hfm.20269.

[75] D. Cárdenas-Vélez *et al.*, "El efecto de la carga de trabajo mental en la intensidad y la dinámica emocional del esfuerzo percibido", *Anales de Psicología / Annals of Psychology*, vol. 29, no. 3, pp. 662–673, 2013, ISSN: 1695-2294. DOI: 10.6018/analesps.29.3.175801.

[76] G. Marquart, C. Cabrall, and J. de Winter, "Review of Eye-related Measures of Drivers' Mental Workload", *Procedia Manufacturing*, 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015, vol. 3, pp. 2854–2861, Jan. 1, 2015, ISSN: 2351-9789. DOI: 10.1016/j.promfg.2015.07.783.

[77] B. Pfleging *et al.*, "A Model Relating Pupil Diameter to Mental Workload and Lighting Conditions", in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI '16, New York, NY, USA: Association for Computing Machinery, May 7, 2016, pp. 5776–5788, ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858117.

[78] B. Cain, “A Review of the Mental Workload Literature”, Defence Research and Development Canada - Toronto Human System Integration Section, Toronto, RTO-TR-HFM-121-Part-II, Jul. 1, 2007, p. 35.

[79] P. A. Hancock, N. Meshkati, and M. M. Robertson, “Physiological reflections of mental workload”, *Aviation, Space, and Environmental Medicine*, vol. 56, no. 11, pp. 1110–1114, Nov. 1985, ISSN: 0095-6562. PMID: 3907615.

[80] P. A. Hancock, “The Effect of Gender and Time of Day Upon the Subjective Estimate of Mental Workload During the Performance of a Simple Task”, in *Advances in Psychology*, vol. 52, North-Holland, Jan. 1, 1988, pp. 239–250. DOI: 10.1016/S0166-4115(08)62390-0.

[81] M. Ohsuga, F. Shimono, and H. Genno, “Assessment of phasic work stress using autonomic indices”, *International Journal of Psychophysiology*, Psychophysiology In, vol. 40, no. 3, pp. 211–220, Apr. 1, 2001, ISSN: 0167-8760. DOI: 10.1016/S0167-8760(00)00189-6.

[82] E. T. Pierce, *Mental Workload Measurement Using the Intersaccadic Interval*. Ottawa: Library and Archives Canada, 2011, ISBN: 978-0-494-71378-5.

[83] H. Zhang *et al.*, “Detection of variations in cognitive workload using multi-modality physiological sensors and a large margin unbiased regression machine”, in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2014, pp. 2985–2988. DOI: 10.1109/EMBC.2014.6944250.

[84] E. Wolf, “Psychophysiologische Profile in nutzerzentrierten Mensch-Maschine-Systemen: Extraktion kardialer und elektrodermaler Profile zur Bewertung der mentalen Beanspruchung”, 2021. DOI: 10.5445/IR/1000139423.

[85] M. Schneider, “Blickbasierte Beanspruchungsmessung : Entwicklung und Evaluation eines Kalibrierungssystems zur individuellen Bewertung der mentalen Beanspruchung in der Mensch-Technik-Interaktion”, 2019. DOI: 10.5445/KSP/1000083975.

[86] D. Giakoumis, D. Tzovaras, and G. Hassapis, “Subject-dependent biosignal features for increased accuracy in psychological stress detection”, *International Journal of Human-Computer Studies*, vol. 71, no. 4, pp. 425–439, Apr. 1, 2013, ISSN: 1071-5819. DOI: 10.1016/j.ijhcs.2012.10.016.

[87] S. Liaghat and S. K. Balasundram, “A Review: The Role of Remote Sensing in Precision Agriculture”, *American Journal of Agricultural and Biological Sciences*, vol. 5, no. 1, pp. 50–55, Mar. 31, 2010, ISSN: 1557-4997. DOI: 10.3844/ajabssp.2010.50.55.

[88] M. Wang *et al.*, “Agricultural Field Boundary Delineation with Satellite Image Segmentation for High-Resolution Crop Mapping: A Case Study of Rice Paddy”, *Agronomy*, vol. 12, no. 10, p. 2342, 10 Oct. 2022, ISSN: 2073-4395. DOI: 10.3390/agronomy12102342.

[89] B. Watkins and A. van Niekerk, “A comparison of object-based image analysis approaches for field boundary delineation using multi-temporal Sentinel-2 imagery”, *Computers and Electronics in Agriculture*, vol. 158, pp. 294–302, Mar. 1, 2019, ISSN: 0168-1699. DOI: 10.1016/j.compag.2019.02.009.

[90] CLAAS Gruppe, “Innovation Lab: CLAAS zeigt auf Agritechnica 2023 nachhaltige und autonome Technik für die Landwirtschaft von Morgen und Übermorgen”, CLAAS KGaA mbH, Presseinformation, Nov. 1, 2023.

[91] Laurel Caes, “John Deere Reveals Fully Autonomous Tractor at CES 2022”, John Deere GmbH & Co. KG, Jan. 4, 2022.

[92] T. Kaper, “Entwicklung Eines Prädiktionsmodells Zur Bestimmung Der Zeitlichen Verfügbarkeit Des Fahrers Am Beispiel Eines Modernen Mähdrescher”, Masterarbeit, Karlsruher Institut für Technologie. Karlsruhe, 2021.

[93] DLG-Ausschuss Arbeitswirtschaft und Prozesstechnik and P. Noack, *Satellitenortungssysteme (GNSS) in Der Landwirtschaft* (DLG-Merkblatt 388), 3rd ed. Frankfurt am Main: DLG e.V, May 2016, 28 pp.

[94] J. P. Snyder, *Flattening the Earth: Two Thousand Years of Map Projections*, TAschenbuchausgabe. Chicago London: The University of Chicago Press, 1997, 365 pp., ISBN: 978-0-226-76747-5.

[95] B. Jochum, “Entwicklung Eines Modells Zur Feldkontur- Und Hinderniserkennung, Um Eine Echtzeitfähige Prädiktion Des Nutzerverhalten Eines Mähdrescherbedieners Zu Ermöglichen”, Bachelorarbeit, Karlsruher Institut für Technologie. Karlsruhe, 2021.

[96] K.-G. Karlsson *et al.*, *CLARA-A3: CM SAF cLoud, Albedo and surface RAdiation dataset from AVHRR data - Edition 3*, NetCDF-4, version 3.0, Satellite Application Facility on Climate Monitoring (CM SAF), Apr. 27, 2023. DOI: 10.5676/EUM_SAF_CM/CLARA_AVHRR/V003.

[97] M. Werner, *Digitale Bildverarbeitung*. Wiesbaden: Springer Fachmedien Wiesbaden, Jan. 1, 2021, 480 pp., ISBN: 978-3-658-22184-3. DOI: 10.1007/978-3-658-22185-0.

[98] J. S. Lim, *Two-Dimensional Signal and Image Processing* (Prentice-Hall Signal Processing Series). Englewood Cliffs, NJ: Prentice Hall PTR, Jan. 1, 1990, 694 pp., ISBN: 0-13-935322-4.

[99] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, Jan. 1, 1979, ISSN: 0018-9472. DOI: 10.1109/TSMC.1979.4310076.

[100] H. Hakemi, O. Pinshow, and D. Gal-Fuss, "Evaluation of Polymer Dispersed Liquid Crystal (PDLC) for Passive Rear Projection Screen Application", *Recent Progress in Materials*, vol. 1, no. 3, pp. 1–1, Jan. 1, 2019. DOI: 10.21926/rpm.1903002.

[101] Smart Eye AB, *Programmer's Guide*, 9th ed. Gothenburg Sweden, Apr. 29, 2020, 79 pp.

[102] M. S. Yousefi *et al.*, "Stress Detection Using Eye Tracking Data: An Evaluation of Full Parameters", *IEEE Access*, vol. 10, pp. 118 941–118 952, 2022, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2022.3221179.

[103] P. Maggi and F. Di Nocera, "Sensitivity of the Spatial Distribution of Fixations to Variations in the Type of Task Demand and Its Relation to Visual Entropy", *Frontiers in Human Neuroscience*, vol. 15, Jun. 8, 2021, ISSN: 1662-5161. DOI: 10.3389/fnhum.2021.642535.

[104] V. Peysakhovich *et al.*, "Frequency analysis of a task-evoked pupillary response: Luminance-independent measure of mental effort", *International Journal of Psychophysiology*, vol. 97, no. 1, pp. 30–37, Jul. 1, 2015, ISSN: 0167-8760. DOI: 10.1016/j.ijpsycho.2015.04.019.

[105] S. Marshall, "The Index of Cognitive Activity: Measuring cognitive workload", in *Proceedings of the IEEE 7th Conference on Human Factors and Power Plants*, Sep. 2002, pp. 7–7. DOI: 10.1109/HFPP.2002.1042860.

[106] S. Marshall, “Method and Apparatus for Eye Tracking”, pat. WO/2000/054654, Sep. 21, 2000.

[107] A. T. Duchowski *et al.*, “The Index of Pupillary Activity: Measuring Cognitive Load *vis-à-vis* Task Difficulty with Pupil Oscillation”, in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, Montreal QC Canada: ACM, Apr. 21, 2018, pp. 1–13, ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3173856.

[108] R. W. Backs and K. A. Seljos, “Metabolic and cardiorespiratory measures of mental effort: The effects of level of difficulty in a working memory task”, *International Journal of Psychophysiology*, vol. 16, no. 1, pp. 57–68, Feb. 1, 1994, ISSN: 0167-8760. DOI: 10.1016/0167-8760(94)90042-6.

[109] F. Di Nocera, M. Camilli, and M. Terenzi, “Using the Distribution of Eye Fixations to Assess Pilots’ Mental Workload”, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 50, no. 1, pp. 63–65, Oct. 1, 2006, ISSN: 1071-1813. DOI: 10.1177/154193120605000114.

[110] M. Maltz and D. Shinar, “Eye Movements of Younger and Older Drivers”, *Human Factors*, vol. 41, no. 1, pp. 15–25, Mar. 1, 1999, ISSN: 0018-7208. DOI: 10.1518/001872099779577282.

[111] K. Pataki *et al.*, “Anwendung von Usability-Maßen zur Nutzeneinschätzung von Fahrerassistenzsystemen”, *Beiträge zur Mensch-Maschine-Systemtechnik aus Forschung und Praxis*, pp. 211–228, 2005.

[112] J. Abich IV, L. Reinerman-Jones, and G. S. Taylor, “Investigating Workload Measures for Adaptive Training Systems”, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 57, no. 1, pp. 2091–2095, Sep. 1, 2013, ISSN: 1071-1813. DOI: 10.1177/1541931213571466.

[113] L. L. D. Stasi *et al.*, “Saccadic Peak Velocity Sensitivity to Variations in Mental Workload”, *Aviation, Space, and Environmental Medicine*, vol. 81, no. 4, pp. 413–417, Apr. 1, 2010, ISSN: 0095-6562. DOI: 10.3357/ASEM.2579.2010.

[114] Y. Takefuji, “Chi-square and P-values versus machine learning feature selection”, *Annals of Oncology*, vol. 36, no. 2, pp. 227–228, Feb. 2025, ISSN: 09237534. DOI: 10.1016/j.annonc.2024.10.013.

[115] V. N. Vapnik, “The Support Vector method”, in *Artificial Neural Networks — ICANN’97*, W. Gerstner *et al.*, Eds., Berlin, Heidelberg: Springer, 1997, pp. 261–271, ISBN: 978-3-540-69620-9. DOI: 10.1007/BFb0020166.

[116] M. Awad and R. Khanna, “Support Vector Machines for Classification”, in *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*, M. Awad and R. Khanna, Eds., Berkeley, CA: Apress, 2015, pp. 39–66, ISBN: 978-1-4302-5990-9. DOI: 10.1007/978-1-4302-5990-9_3.

[117] S. B. Kotsiantis, “Decision trees: A recent overview”, *Artificial Intelligence Review*, vol. 39, no. 4, pp. 261–283, Apr. 1, 2013, ISSN: 1573-7462. DOI: 10.1007/s10462-011-9272-4.

[118] W.-Y. Loh, “Classification and regression trees”, *WIREs Data Mining and Knowledge Discovery*, vol. 1, no. 1, pp. 14–23, 2011, ISSN: 1942-4795. DOI: 10.1002/widm.8.

[119] J. R. Quinlan, “Induction of decision trees”, *Machine Learning*, vol. 1, no. 1, pp. 81–106, Mar. 1, 1986, ISSN: 1573-0565. DOI: 10.1007/BF00116251.

[120] F.-J. Yang, “An Implementation of Naive Bayes Classifier”, in *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, Dec. 2018, pp. 301–306. DOI: 10.1109/CSCI46756.2018.00065.

[121] T. Hastie, R. Tibshirani, and J. Friedman, “Model Assessment and Selection”, in *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, T. Hastie, R. Tibshirani, and J. Friedman, Eds., New York, NY: Springer, 2009, pp. 219–259, ISBN: 978-0-387-84858-7. DOI: 10.1007/978-0-387-84858-7_7.

[122] B. Shahriari *et al.*, “Taking the Human Out of the Loop: A Review of Bayesian Optimization”, *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148–175, Jan. 2016, ISSN: 1558-2256. DOI: 10.1109/JPROC.2015.2494218.

[123] R. H. L. Hillege *et al.*, “The Mental Machine: Classifying Mental Workload State from Unobtrusive Heart Rate-Measures Using Machine Learning”, in *Adaptive Instructional Systems*, R. A. Sottilare and J. Schwarz, Eds., Cham: Springer International Publishing, 2020, pp. 330–349, ISBN: 978-3-030-50788-6. DOI: 10.1007/978-3-030-50788-6_24.

[124] C. Wu *et al.*, “Eye-Tracking Metrics Predict Perceived Workload in Robotic Surgical Skills Training”, *Human Factors*, vol. 62, no. 8, pp. 1365–1386, Dec. 2020, ISSN: 1547-8181. DOI: 10.1177/0018720819874544. PMID: 31560573.

[125] T. Luong *et al.*, “Towards Real-Time Recognition of Users Mental Workload Using Integrated Physiological Sensors Into a VR HMD”, in *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Nov. 2020, pp. 425–437. DOI: 10.1109/ISMAR50242.2020.00068.

[126] J. H. Ettema and R. L. Zielhuis, “Physiological Parameters of Mental Load”, *Ergonomics*, vol. 14, no. 1, pp. 137–144, Jan. 1, 1971, ISSN: 0014-0139. DOI: 10.1080/00140137108931232. PMID: 4399026.

[127] T. Laine *et al.*, “Selection of input features across subjects for classifying crewmember workload using artificial neural networks”, *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 32, no. 6, pp. 691–704, Nov. 2002, ISSN: 1558-2426. DOI: 10.1109/TSMCA.2002.807036.

[128] L.-l. Chen *et al.*, “Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers”, *Expert Systems with Applications*, vol. 85, pp. 279–291, Nov. 1, 2017, ISSN: 0957-4174. DOI: 10.1016/j.eswa.2017.01.040.

[129] H. Atasoy and E. Yildirim, “Classification of Verbal and Quantitative Mental Tasks Using Phase Locking Values between EEG Signals”, *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 9, no. 7, pp. 383–390, Jul. 31, 2016, ISSN: 20054254, 20054254. DOI: 10.14257/ijsip.2016.9.7.34.

[130] A. Oschlies-Strobel *et al.*, “Preliminary classification of cognitive load states in a human machine interaction scenario”, in *2017 International Conference on Companion Technology (ICCT)*, Sep. 2017, pp. 1–5. DOI: 10.1109/COMPANION.2017.8287084.

[131] G. Lugano, “Virtual assistants and self-driving cars”, in *2017 15th International Conference on ITS Telecommunications (ITST)*, May 2017, pp. 1–5. DOI: 10.1109/ITST.2017.7972192.

[132] S. Hohagen and T. Steckel, “Entwicklung der landwirtschaftlichen Arbeitswissenschaft im Kontext der fortschreitenden Digitalisierung”, in *Arbeit Unter Einem DA-CH: Der Landwirt Im 4.0-Modus*, L.-I. für Agrartechnik und Bioökonomie e.V., Ed., Mar. 8, 2022, pp. 18–24.

[133] I. Fette, A. Melnikov, and R. Editor, “The WebSocket Protocol”, RFC Editor, 6455, Jan. 1, 2011. DOI: 10.17487/RFC6455.

[134] G. W. Furnas and J. Zacks, “Multitrees: Enriching and reusing hierarchical structure”, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI ’94, New York, NY, USA: Association for Computing Machinery, Apr. 24, 1994, pp. 330–336, ISBN: 978-0-89791-650-9. DOI: 10.1145/191666.191778.

[135] R. Jha *et al.*, “Analyzing the Effectiveness of Voice-Based User Interfaces in Enhancing Accessibility in Human-Computer Interaction”, in *2024 IEEE 13th International Conference on Communication Systems and Network Technologies (CSNT)*, Apr. 2024, pp. 777–781. DOI: 10.1109/CSNT60213.2024.10545835.

[136] Evan Liu *et al.*, “Web Speech API”, W3C, Draft Community Group Report, Jul. 7, 2025.

[137] Mark Walker, Daniel Burnett, and Andrew Hunt, “Speech Synthesis Markup Language (SSML) Version 1.0”, W3C, W3C Recommendation, Oct. 2004.

[138] Martin Waldschmidt, “Analyse Und Umsetzung Eines Offline-Fähigen Deep-Learning-Spracherkennungsmodells Für Den Einsatz in Einem Mähdrescher”, Bachelorarbeit, Karlsruher Institut für Technologie. Karlsruhe, Oct. 31, 2021.

[139] A. Hannun *et al.* “Deep Speech: Scaling up end-to-end speech recognition”. arXiv: 1412.5567 `[cs]`. (Dec. 19, 2014), [Online]. Available: http://arxiv.org/abs/1412.5567 (visited on 08/18/2025), pre-published.

[140] S. Kriman *et al.* “QuartzNet: Deep Automatic Speech Recognition with 1D Time-Channel Separable Convolutions”. arXiv: 1910.10261 `[eess]`. (Oct. 22, 2019), [Online]. Available: http://arxiv.org/abs/1910.10261 (visited on 08/18/2025), pre-published.

[141] K. Heafield, “KenLM: Faster and Smaller Language Model Queries”, in *Proceedings of the Sixth Workshop on Statistical Machine Translation*, C. Callison-Burch *et al.*, Eds., Edinburgh, Scotland: Association for Computational Linguistics, Jul. 2011, pp. 187–197.

[142] National Institute of Standards and Technology (US), “Advanced Encryption Standard (AES)”, National Institute of Standards and Technology (U.S.), Washington, D.C., NIST FIPS 197-upd1, May 9, 2023, NIST FIPS 197–upd1. DOI: 10.6028/NIST.FIPS.197-upd1.

[143] "User Datagram Protocol", Internet Engineering Task Force, Request for Comments RFC 768, Aug. 1980, 3 pp. DOI: 10.17487/RFC0768.

[144] W. Eddy, "Transmission Control Protocol (TCP)", Internet Engineering Task Force, Request for Comments RFC 9293, Aug. 2022, 98 pp. DOI: 10.17487/RFC9293.

[145] P. Scholz, *Softwareentwicklung eingebetteter Systeme: Grundlagen, Modellierung, Qualitätssicherung* (Xpert.press). Berlin, Heidelberg: Springer, 2005, ISBN: 978-3-540-23405-0. DOI: 10.1007/3-540-27522-3.

[146] S. Tilkov *et al.*, *REST und HTTP: Entwicklung und Integration nach dem Architekturstil des Web*, 3., aktualisierte und erweiterte Auflage. Heidelberg: dpunkt.verlag, 2015, 1 p., ISBN: 978-3-86490-120-1.

[147] J. C. Klensin, "Simple Mail Transfer Protocol", Internet Engineering Task Force, Request for Comments RFC 5321, Oct. 2008, 95 pp. DOI: 10.17487/RFC5321.

[148] M. Crispin, "Internet Message Access Protocol - Version 4rev1", Internet Engineering Task Force, Request for Comments RFC 3501, Mar. 2003, 108 pp. DOI: 10.17487/RFC3501.

[149] R. Mitkov, *The Oxford Handbook of Computational Linguistics*. OUP Oxford, 2004, 808 pp., ISBN: 978-0-19-927634-9. Google Books: yl6AnaKtVAkC.

[150] M. M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential", *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 25, no. 1, pp. 49–59, Mar. 1, 1994, ISSN: 0005-7916. DOI: 10.1016/0005-7916(94)90063-9.

[151] M. Friedman, "The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance", *Journal of the American Statistical Association*, vol. 32, no. 200, pp. 675–701, Dec. 1, 1937, ISSN: 0162-1459. DOI: 10.1080/01621459.1937.10503522.

[152] DKE Deutsche Kommission Elektrotechnik Elektronik Informationstechnik in DIN und VDE, "DIN IEC 60050-351 Internationales Elektrotechnisches Wörterbuch - Teil 351: Leittechnik (IEC 60050-351:2013)", *Deutsche Normen*, Sep. 2014. DOI: 10.31030/2159569.

[153] L. A. Abel and J. Douglas, "Effects of age on latency and error generation in internally mediated saccades", *Neurobiology of Aging*, vol. 28, no. 4, pp. 627–637, Apr. 1, 2007, ISSN: 0197-4580. DOI: 10.1016/j.neurobiolaging.2006.02.003.

[154] E. L. Irving *et al.*, “Horizontal Saccade Dynamics across the Human Life Span”, *Investigative Ophthalmology & Visual Science*, vol. 47, no. 6, pp. 2478–2484, Jun. 1, 2006, ISSN: 1552-5783. DOI: 10.1167/iovs.05-1311.

[155] S. Böttinger, “Agrarsysteme der Zukunft: Fahrerkabine 4.0 - Entwicklung einer beanspruchungsadaptiven Nutzerschnittstelle für Landmaschinenbetreiber - Teilprojekt F: Sachbericht zum Verwendungsnachweis”, Hannover : Technische Informationsbibliothek, 2025. DOI: 10.34657/22622.

A. Appendix

A.1. Used machine data parameter

Table A.1.: Input data machine

Category	Name	Description
SeatData		
	SeatPosition	Current seat position
EngineData		
	EngineSpeed	Engine speed
	EngineTorque	Engine torque
	Utilization	Engine utilization
MachineData		
	MainDriveStatus	Main drive status
	HeaderStatus	Header status
	VehicleSpeed	Vehicle speed
GrainTankStatus		
	GrainTankFillLevel	Grain tank fill level
	GrainTank100PStatus	Grain tank 100 % full
	GrainTank70PStatus	Grain tank 70 % full
	GrainTankFillRange	Grain tank fill range
GPSData		
	Latitude	GPS latitude
	Longitude	GPS longitude
	Altitude	GPS altitude
	CompassBearing	Compass bearing
	Pitch	Pitch angle

Category	Name	Description
MachineStatus		
	CruisePilot	Cruise pilot
	Korntankrohr	Unloading auger
	FuelLevel	Fuel level
	SeparatorRPM	Separator RPM
	ThreshdrumRPM	Threshing drum RPM
	ConcavePos	Concave position
	UpperSievePos	Upper sieve position
	LowerSievePos	Lower sieve position
	FanRPM	Fan RPM
	Perform_Cleaning	Cleaning and separation performance
	GrainMoisture	Grain moisture
	CutterHeight	Cutterbar height
	Reel	Reel position or setting

A.2. List of implemented Task Recommendations

Table A.2.: Overview of Task Recommendations (TRs)

Name	Short Description	Category	Sub-category	Operator Group	Action Count
E-Mail	Communication via email	01: Business; 05: Media	Communication	All drivers	6
Instant Messenger	Communication via instant messenger	01: Business; 05: Media	Communication	All drivers	8
Tools	Tools for planning and organization	01: Business; 05: Media	Organization	All drivers	7
Weather	Weather websites in the browser	01: Business; 05: Media	Organization	All drivers	8
Farm Management	365FarmNet	01: Business	Field information	All drivers	2
Information Messages	Displays various information messages	02: Machine	Help	All drivers	2
Machine Settings	Reminds the operator to check harvest parameters	02: Machine	Parameters	All drivers	2
Sunshade	Activates or deactivates the sunshade	02: Machine	Settings	All drivers	2
Exercise	Short exercises for relaxation or stretching	03: Wellbeing	Wellbeing	All drivers	10
Eating	Reminder to eat	03: Wellbeing	Wellbeing	All drivers	5
Drinking	Reminder to drink	03: Wellbeing	Wellbeing	All drivers	5

Name	Short Description	Category	Sub-category	Operator Group	Action Count
User Manual	Shows an informative manual or tutorial video about various functions, parameters or components of the combine harvester	04: Wiki	Help	All drivers	1
Translation	Translation using Pons	04: Wiki	Help	All drivers	4
Wolfram Alpha	Information using Wolfram Alpha	04: Wiki	Help	All drivers	3
Wikipedia	Information using Wikipedia	04: Wiki	Help	All drivers	4
Radio	Plays radio on request	05: Media	Multimedia	All drivers	7
Spotify	Plays Spotify on request	05: Media	Multimedia	All drivers	8
News	Access to various news sites via the browser	05: Media	News	All drivers	3
Social Media	Access to various social media platforms via web browser	05: Media	Social Media	All drivers	3
User Data	Shows the user data interface	06: Settings	User Data	All drivers	16
Virtual Assistant - Time	Announces the time and date	06: Settings	Recommendations	All drivers	1
Disable HEs	Enable and disable recommendations	06: Settings	Recommendations	All drivers	2
Virtual Assistant - Jokes	Jokes from the virtual assistant	06: Settings	Recommendations	All drivers	1

Name	Short Description	Category	Sub-category	Operator Group	Action Count
Virtual Assistant - Language	Language of the virtual assistant	06: Settings	Recommendations	All drivers	2
HE Lists	List recommendations	06: Settings	Recommendations	All drivers	8
Sleep Mode	The virtual assistant is deactivated	06: Settings	Recommendations	All drivers	3
Virtual Assistant - Misc	Sayings of the virtual assistant	06: Settings	Recommendations	All drivers	2
Seat Position	Change the current seat position	06: Settings	Seat	All drivers	11
Greeting Trade Show Scenario	User is greeted by Fabia	07: Misc	Trade Show Scenario	All drivers	1
End Trade Show Scenario	End of the trade show scenario is reached	07: Misc	Trade Show Scenario	All drivers	2
End Manual Driving	End manual driving	07: Misc	User Study	All drivers	1
Autonomous Driving	Autonomous driving	07: Misc	User Study	All drivers	1
Changes Needed	User study scenarios	07: Misc	User Study	All drivers	1
Changes Needed	User study scenarios without check	07: Misc	User Study	All drivers	1
User Study Breathing Exercise	User study breathing exercise	07: Misc	User Study	All drivers	1

Name	Short Description	Category	Sub-category	Operator Group	Action Count
Reminder Scenario	Driver takes too long to execute the recommendation	07: Misc	User Study	All drivers	2
User Study Scenario Funny Video	User study scenario funny video	07: Misc	User Study	All drivers	8
End User Study Minigame	End minigame in left armrest	07: Misc	User Study	All drivers	1
User Study Minigame	Minigame in left armrest	07: Misc	User Study	All drivers	4
Distance to Field Edge	Driver is too far from field edge	07: Misc	User Study	All drivers	2
Status Query	Current RSME and SAM status queried	07: Misc	User Study	All drivers	3
Changes Needed	End user study scenarios	07: Misc	User Study	All drivers	1
User Study Condition Assessment	User studies	07: Misc	User Study	All drivers	5
Changes Needed	User study scenarios	07: Misc	User Study	All drivers	3
User Study Weather	User study weather	07: Misc	User Study	All drivers	4
User Study Wikipedia	User study Wikipedia	07: Misc	User Study	All drivers	4

Name	Short Description	Category	Sub-category	Operator Group	Action Count
User Study Knowledge in 60 Seconds	User study knowledge in 60 seconds	07: Misc	User Study	All drivers	9

Karlsruher Schriftenreihe Fahrzeugsystemtechnik
FAST Institut für Fahrzeugsystemtechnik
(ISSN 1869-6058)

Eine vollständige Übersicht der Bände finden Sie im Verlagsshop

Band 76 Kai-Lukas Bauer
Echtzeit-Strategieplanung für vorausschauendes automatisiertes Fahren
ISBN 978-3-7315-0949-3

Band 77 Thomas Schirle
Systementwurf eines elektromechanischen Fahrwerks für Megacitymobilität
ISBN 978-3-7315-0995-0

Band 78 Dominik Dörr
Simulation of the thermoforming process of UD fiber-reinforced thermoplastic tape laminates
ISBN 978-3-7315-0998-1

Band 79 Dominik Robert Naake
Simulation of damage mechanisms in weave reinforced materials based on multiscale modeling
ISBN 978-3-7315-1005-5

Band 80 Martin Hohberg
Experimental investigation and process simulation of the compression molding process of Sheet Molding Compound (SMC) with local reinforcements
ISBN 978-3-7315-1007-9

Band 81 Benedikt Fengler
Manufacturing-constrained multi-objective optimization of local patch reinforcements for discontinuous fiber reinforced composite parts
ISBN 978-3-7315-1006-2

Band 82 Johannes Masino
Road Condition Estimation with Data Mining Methods using Vehicle Based Sensors
ISBN 978-3-7315-1004-8

Band 83 **11. Kolloquium Mobilhydraulik**
10. September 2020, Karlsruhe
ISBN 978-3-7315-1036-9

Band 84 Felix Weber
Beitrag zur Entwicklung von Konstantflusspumpen für Frischbeton unter genauerer Betrachtung der Dickstoffventile
ISBN 978-3-7315-1037-6

Band 85 8. Fachtagung
Hybride und energieeffiziente Antriebe für mobile Arbeitsmaschinen. 23. Februar 2021, Karlsruhe
ISBN 978-3-7315-1071-0

Band 86 Sebastian Fünfgeld
Vorausschauende Regelung von Fahrzeugsystemen durch stochastische Vorhersage der Fahrzeugdynamik
ISBN 978-3-7315-1060-4

Band 87 Isabelle Charlotte Ays
Development of a CO2e quantification method and of solutions for reducing the greenhouse gas emissions of construction machines = Entwicklung einer CO2e Quantifizierungsmethode und von Lösungen zur Reduzierung von Treibhausgasemissionen in Baumaschinen
ISBN 978-3-7315-1033-8

Band 88 Alexander Bernath
Numerical prediction of curing and process-induced distortion of composite structures
ISBN 978-3-7315-1063-5

Band 89 Nils Bulthaupt
Objektivierung des Schwingungskomforts schwerer Nutzfahrzeuge
ISBN 978-3-7315-1075-8

Band 90 Lars Brinkschulte
Assistenzsysteme zur Reduktion des Schädigungsverhaltens von Komponenten einer mobilen Arbeitsmaschine
ISBN 978-3-7315-1089-5

Band 91 Dominik Dörr
Adaptive Fahrhinweise für ein längsdynamisches Fahrerassistenzsystem zur Steigerung der Energieeffizienz
ISBN 978-3-7315-1090-1

Band 92 Jürgen Römer
Steuerung und Regelung des Lenkradmoments durch Nutzung radselektiver Frontantriebe
ISBN 978-3-7315-1104-5

Band 93 Christian Riese
Werkzeuge und Konzepte für die Untersuchung und Entwicklung zukünftiger Kfz-Bremssysteme
ISBN 978-3-7315-1125-0

Band 94 Yaoqun Zhou
Dynamisches Bremsverhalten des Reifen-Fahrwerk-Systems
ISBN 978-3-7315-1156-4

Band 95 Stefan Haug
Ganzheitliche Optimierung einer Axialkolbenpumpe durch bedarfsangepasste Entlastung tribologischer Kontakte
ISBN 978-3-7315-1150-2

Band 96 Stefan Scheubner
Stochastic Range Estimation Algorithms for Electric Vehicles using Data-Driven Learning Models
ISBN 978-3-7315-1166-3

Band 97 Yusheng Xiang
AI and IoT Meet Mobile Machines: Towards a Smart Working Site
ISBN 978-3-7315-1165-6

Band 98 Nils Meyer
Mesoscale simulation of the mold filling process of Sheet Molding Compound
ISBN 978-3-7315-1173-1

Band 99 Christian Timo Poppe
Process simulation of wet compression moulding for continuous fibre-reinforced polymers
ISBN 978-3-7315-1190-8

Band 100 Torben Fischer
Modellprädiktive Regelung eines innovativen Thermomanagement-Systems für batterieelektrische Fahrzeuge
ISBN 978-3-7315-1199-1

Band 101 Florian Wittemann
Fiber-dependent injection molding simulation of discontinuous reinforced polymers
ISBN 978-3-7315-1217-2

Band 102 Sebastian Watzl
Experimentelle und numerische Analyse des Körperschallübertragungsverhaltens von Aggregatelagerelementen im akustisch relevanten Frequenzbereich
ISBN 978-3-7315-1226-4

Band 103 Dominik Stretz
Vibroakustische Analyse eines elektrischen Radnabenmotors und Optimierung durch geeignete Steuerungsansätze
ISBN 978-3-7315-1245-5

Band 104 Mohamed Elgharbawy
Measurable Safety of Automated Driving Functions in Commercial Motor Vehicles - Technological and Methodical Approaches
ISBN 978-3-7315-1254-7

Band 105 Bernhard Schmiedel
Indirekte Schätzung des Fahrbahnnässegrads zur Detektion von gefährlichen Fahrzuständen
ISBN 978-3-7315-1258-5

Band 106 9. Fachtagung
Hybride und energieeffiziente Antriebe für mobile Arbeitsmaschinen. 28. Februar 2023, Karlsruhe
ISBN 978-3-7315-1260-8

Band 107 Patrick Riehm
Zur Wechselwirkung zwischen Fahrbahntextur und Laufstreifenmischung von Pkw-Reifen
ISBN 978-3-7315-1268-4

Band 108 Markus Tesar
Deep Reinforcement Learning zur Steigerung von Energieeffizienz und Pünktlichkeit von Straßenbahnen
ISBN 978-3-7315-1277-6

Band 109 Michael Mürken
Methode zur Bewertung der Zuverlässigkeit der elektrischen Energieversorgung in der automobilen Vorentwicklung
ISBN 978-3-7315-1298-1

Band 110 Julien Pinay
Experimental investigation of relevant road surface descriptors for tire-road noise measurements on low-absorbing road surfaces
ISBN 978-3-7315-1328-5

Band 111 Adrian Strigel
Methode zur Ermittlung optimaler Rad- und Reifendimensionen in der frühen Entwicklungsphase von Personenkraftwagen
ISBN 978-3-7315-1321-6

Band 112 Jens Jauch
Trajectory optimization based on recursive B-spline approximation for automated longitudinal control of a battery electric vehicle
ISBN 978-3-7315-1332-2

Band 113 Nicolas Fraikin
Methodik zur effizienten Applikation automatisierter Fahrfunktionen
ISBN 978-3-7315-1339-1

Band 114 Jan Siebert
Effizienzoptimierung mobilhydraulischer Load-Sensing-Systeme durch Reduzierung systembedingter Druckverluste am Beispiel eines Hydraulikbaggers
ISBN 978-3-7315-1343-8

Band 115 Tobias Sebastian Straub
Flottendatenbasierte physikalische Routenenergiebedarfsprognose
ISBN 978-3-7315-1348-3

Band 116 Michael Herrmann
Eine Methodik zur Definition von Zielkriterien am Beispiel des tieffrequenten Geräuschkomforts eines Fahrzeugs
ISBN 978-3-7315-1370-4

Band 117 Adam Thor Thorgeirsson
Probabilistic Prediction of Energy Demand and Driving Range for Electric Vehicles with Federated Learning
ISBN 978-3-7315-1371-1

Band 118 Alexander Jackstadt
Constrained-layer damping in hybrid fibre metal elastomer laminates and its tolerance to damage
ISBN 978-3-7315-1376-6

Band 119 **13. Kolloquium Mobilhydraulik 8./9. Oktober 2024, Karlsruhe**
ISBN 978-3-7315-1381-0

Band 120 Toni Wilhelm
Querdynamik von Velomobilen unter Berücksichtigung der Reifeneigenschaften
ISBN 978-3-7315-1382-7

Die Bände sind unter www.ksp.kit.edu als PDF frei verfügbar oder als Druckausgabe bestellbar.

Band 121 Alexander Brunker
Hochgenaue und robuste odometriebasierte Lokalisierung in einem Parkvorgang
ISBN 978-3-7315-1383-4

Band 122 Daniel Förster
Systemauslegung autarker Hybridantriebe unter Berücksichtigung kundenspezifischer Randbedingungen
ISBN 978-3-7315-1384-1

Band 123 Maurizio Mauro Festa
Objektive Fahrkomfortbewertung des Gesamtfahrzeugs unter Berücksichtigung menschlicher Wahrnehmungsaspekte
ISBN 978-3-7315-1392-6

Band 124 10. Fachtagung
Hybride und energieeffiziente Antriebe für mobile Arbeitsmaschinen. 19. Februar 2025, Karlsruhe
ISBN 978-3-7315-1403-9

Band 125 8. Fachtagung
MOBILE MACHINES – Sicherheit und Fahrerassistenz für Arbeitsmaschinen. 18. Februar 2025, Karlsruhe
ISBN 978-3-7315-1404-6

Band 126 Raphael Mieth
Kundenbedarfsgerechte Auslegung elektrischer Pkw-Antriebssysteme zur Steigerung der Energieeffizienz
ISBN 978-3-7315-1416-9

Band 127 Lukas Michiels
A simulation-based approach to the fluid-structure interaction inside fatigue cracks in hydraulic components
ISBN 978-3-7315-1424-4

Band 128 Felix Deufel
Optimales Energiemanagement mild elektrifizierter Antriebe unter realen Betriebsbedingungen mittels Prädiktionsalgorithmen aus dem Bereich des Maschinellen Lernens
ISBN 978-3-7315-1426-8

Band 129 Andreas M. Billert
Predictive Battery Thermal Management of Electric Vehicles using Deep Learning
ISBN 978-3-7315-1429-9

Band 130 Bastian Benjamin Jascha Schäfer
Macroscopic forming simulation of unidirectional non-crimp fabrics: Hyperelastic material modeling and 3D-solid-shell approach
ISBN 978-3-7315-1430-5

Band 131 Rene Armbruster
Untersuchung des Schwingungsverhaltens der Einheit Reifen, Rad, Radführung im komfortrelevanten Schwingungsbereich
ISBN 978-3-7315-1433-6

Band 132 Constantin Krauß
Direction-Dependent Result Data Transfer in Virtual Process Chains for Fiber-Reinforced Polymers and the Impact on Structural Simulation
ISBN 978-3-7315-1434-3

Band 133 Roland Herberth
Nutzerorientierte Optimierung des Over-the-Air-Updateprozesses vernetzter Kraftfahrzeuge
ISBN 978-3-7315-1438-1

Band 134 Jan Löwer
Untersuchungen zum Bremsverhalten von PKW-Reifen im Grenzbereich zwischen Nassgriff und Aquaplaning
ISBN 978-3-7315-1439-8

Band 135 Timo von Wysocki
Reduction of Selected Road Noise Phenomena by Modifying Suspension Kinematics Using Metamodels in the Digital Vehicle Development Phase
ISBN 978-3-7315-1447-3

Band 136 Stefan Schläfle
The Influence of Vehicle Operating Conditions on Tire–Road Particulate Matter Emissions
ISBN 978-3-7315-1462-6

Band 137 Fabian Weitz
Entwicklungsmethodik für Fahrwerksysteme neuartiger Fahrzeugkonzepte
ISBN 978-3-7315-1466-4

Band 138 Marco Wydra
Vorgehen zur Optimierung einer elektrohydraulischen Bedarfsstromsteuerung durch Erweiterung der Systemfreiheitsgrade
ISBN 978-3-7315-1476-3

Band 139 Steffen Metzger
Development of an assistant system for regulating the mental strain state of agricultural machinery operators
ISBN 978-3-7315-1482-4

Karlsruher Schriftenreihe
Fahrzeugsystemtechnik

Institut für Fahrzeugsystemtechnik

The increasing automation of agricultural machinery leads to fluctuating phases of mental underload and overload for operators, which can impair performance, safety, and well-being. This work describes the development and evaluation of an adaptive assistance system that monitors, evaluates, and regulates mental strain in real time. As an application example, a highly automated combine harvester was simulated in an immersive demonstrator cabin for experimental studies. Multimodal sensor data, including physiological metrics such as eye-tracking parameters and cardiovascular signals, were used to train a machine learning model capable of classifying mental states. This model was integrated into a modular virtual assistant providing context-dependent task recommendations via voice and touch interfaces to prevent critical strain states. Experimental studies demonstrated effective detection and regulation of mental strain. A subsequent field study confirmed practical applicability and identified optimization potential regarding system integration and user acceptance. The results indicate improved safety, efficiency, and operator satisfaction.

ISBN 978-3-7315-1482-4

9 783731 514824

ISSN 1869-6058
ISBN 978-3-7315-1482-4

Gedruckt auf FSC-zertifiziertem Papier