Quantified markers for degraded automation performance

Deliverable D1.3 – WP1 – Public

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 814735.

Quantified markers for degraded automation performance

Work package 1, Deliverable D1.3

Please refer to this report as follows:
Mano, D., Berger, N., Larsson, A., Brännström, M., Knauss, A., Toffetti, A., … Christoph, M. (2021). Quantified Markers for Degraded Automation Performance, Deliverable D1.3 of the H2020 project MEDIATOR.

Project details:

Project start date: **Duration:** **Project name:**	01/05/2019 48 months MEDIATOR – MEdiating between Driver and Intelligent Automated Transport systems on Our Roads
Coordinator:	Prof. Dr. Nicole van Nes | SWOV – Institute for Road Safety Research Bezuidenhoutseweg 62, 2594 AW, The Hague, The Netherlands
	This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 814735.

Deliverable details:

Version: **Dissemination level:** **Due date:** **Submission date:**	Final Public (PU) 30/04/2021 30/04/2021

Lead contractor for this deliverable:
Annika Larsson – Veoneer, Sweden

Report Author(s):	Mano, D., Larsson, A. – Veoneer, Sweden Berger, N., Brännström, M., Knauss, A. – Zenseact, Sweden Toffetti, A., Khorramian, K. – Fiat Chrysler Automobiles, Italy Bakker, B. – Cygnify, Netherlands Christoph, M. – SWOV Institute for Road Safety Research, Netherlands

Revision history

Date	Version	Reviewer	Description
16/04/2021	Draft for QA	Prof. Dr. Ann Williamson, University of New South Wales, Sydney, Australia	External review
16/04/2021	Draft for QA	Johan Karlsson, Autoliv, Sweden	Internal review
29/04/2021	Final draft	Ingrid van Schagen (QA Officer) & Ernst Verschragen (Project management officer) – SWOV Institute for Road Safety Research	Final editing and checks
30/04/2021	Final deliverable	Michiel Christoph (Deputy project coordinator) – SWOV Institute for Road Safety Research	Approved

Table of contents

List of Abbreviations

ADS	Automated Driving System (SAE levels 3+)
AF	Automation Fitness
CM	Continuous Mediation (automation level)
DAS	Driving Automation System (SAE levels 1-5)
DDT	Dynamic Driving Task
HD	High Definition
HMI	Human Machine Interface
KPI	Key Performance Indicator
LIDAR	LIght Detection And Ranging
ODD	Operational Design Domain
OEM	Original Equipment Manufacturer
RADAR	RAdio Detection And Ranging
SAE	Society of Automobile Engineers
SB	Driver Standby (automation level)
TTAF	Time To Automation Fitness
TTAU	Time To Automation Unfitness
TtS	Time To Sleep (automation level)
V2X	Vehicle to cloud

About MEDIATOR

MEDIATOR is a 4-year project led by SWOV. It started in May 2019. MEDIATOR will develop a mediating system for drivers in semi-automated and highly automated vehicles, resulting in safe, real-time switching between the human driver and automated system based on who is most fit to drive. MEDIATOR pursues a paradigm shift away from a view that prioritises either the driver or the automation, instead integrating the best of both.

Visions

Automated transport technology is developing rapidly for all transport modes, with huge safety potential. The transition to full automation, however, brings new risks, such as mode confusion, overreliance, reduced situational awareness and misuse. The driving task changes to a more supervisory role, reducing the task load and potentially leading to degraded human performance. Similarly, the automated system may not (yet) function in all situations. The objective of the Mediator system is to intelligently assess the strengths and weaknesses of both the driver and the automation and mediate between them, while also taking into account the driving context.

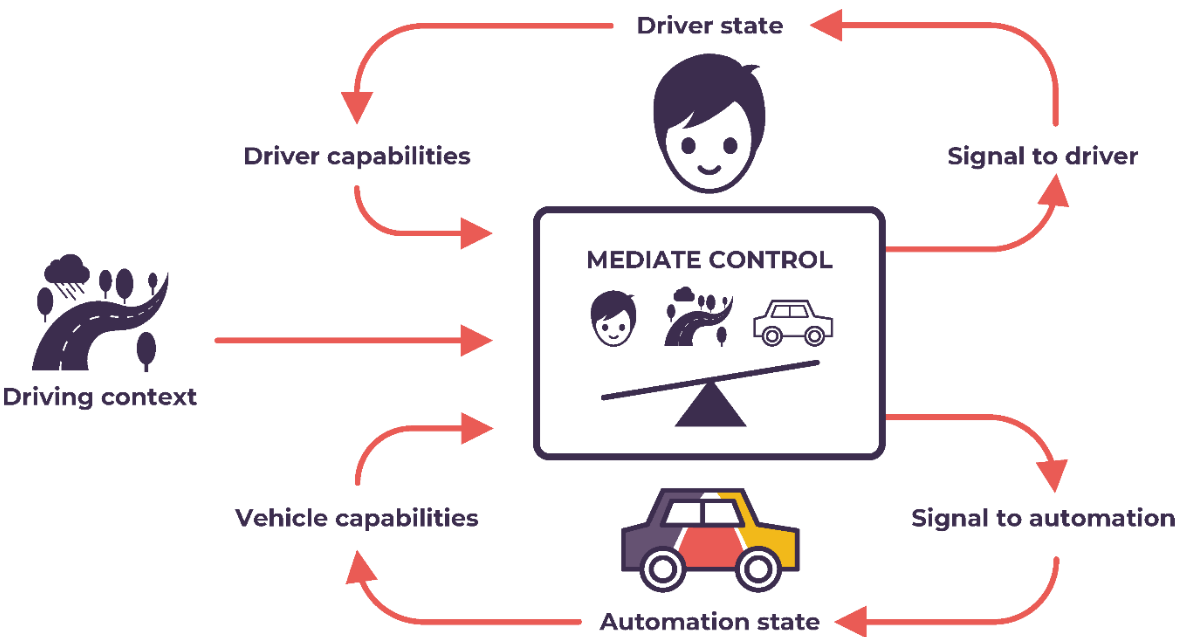

Figure 1. The Mediator system will constantly weigh driving context, driver state and vehicle automation status, while personalising its technology to the drivers' general competence, characteristics, and preferences.

MEDIATOR will optimise the safety potential of vehicle automation during the transition to full (level 5) automation. It will reduce risks, such as those caused by driver fatigue or inattention, or on the automation side imperfect automated driving technology. MEDIATOR will facilitate market exploitation by actively involving the automotive industry during the development process.

To accomplish the development of this support system MEDIATOR will integrate and enhance existing knowledge of human factors and human machine interface (HMI), taking advantage of the of expertise in other transport modes (aviation, rail and maritime). It will develop and adapt available technologies for real-time data collection, storage and analysis and incorporate the latest artificial intelligence techniques, such as deep learning.

Partners

MEDIATOR is being carried out by a consortium of highly qualified research and industry experts, representing a balanced mix of top universities and research organisations as well as several OEMs and suppliers. The consortium, supported by an international Industrial Advisory Board and a Scientific Advisory Board, also represents all transport modes, maximising input from, and transferring results to, aviation, maritime and rail (with mode-specific adaptations).

Executive summary

Vehicle automation has the potential to improve driving safety and driver comfort. The Mediator system aims to aid the realization of this potential by mediating between the driver and the automation on who is fittest to drive. Making this trade-off in a timely and safe manner requires both driver and automation fitness to be detected and predicted for the near future. Within this document, a method to quantify automation fitness, including time to automation fitness and time to automation unfitness are defined. In contrast to the driver state, which has known degraded performance markers such as fatigue or distraction as well as methods to quantify them, there are no established methods for assessing automation fitness. This deliverable provides the foundations and concepts that will allow the automation state module to estimate the automation fitness. The ideas and concepts developed in this deliverable will be implemented and validated at a later stage of the MEDIATOR project.

Firstly, “degraded automation performance” is defined by situations in which the driver disengages or overrides the driving automation system (DAS) due to perceived ill-fitting actions, or a situation in which the DAS shuts itself or goes into some fallback due to within-system quality triggers, or a situation where the automation causes a crash with another road user or the infrastructure.

Degraded performance as it manifests in markers (visible system behaviour) is the consequence of internal or external conditions to the DAS, which we refer to as factors. These factors expose functional limitations of the DAS and thereby have an impact on its fitness to drive. The resulting degradation of information quality circulating throughout the interconnected components will introduce uncertainty, at some point affecting the automation fitness.

Understanding the factors leading to degraded automation performance and the resulting effects throughout the various components of the driving automation system (such as perception or decision making) is a key element of the process for estimating how long the automation may remain fit to drive. We have identified two categories of factors relevant for the automation state module as they can be measured and predicted and therefore used as indications for an upcoming degraded performance (illustrated in Figure 2):

1. Factors related to system input such as adverse weather, dense traffic or roadworks, which can be measured and be predicted using driving context information such as weather forecast from an online service,
2. Factors related to internal states of the system, which can be measured using information from the driving automation system to compute performance self-assessment indicators.

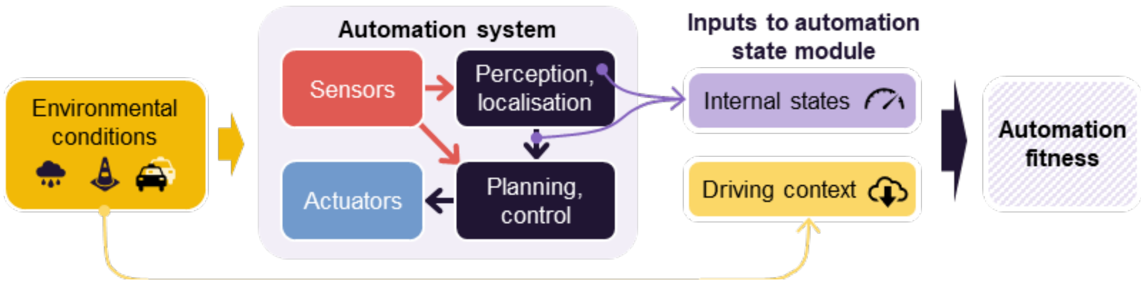

Figure 2. Simplified system architecture and inputs of the automation state module

Taking the (simplified) system architecture of typical automation systems into account as illustrated in Figure 2, we estimate the current and predicted automation fitness using the correlation between the performance self-assessment indicators, the annotated vehicle behaviour, and the driving context.

To quantify the automation fitness, an automation fitness scale is introduced, which corresponds to the rate of automation system deactivations or overrides (following the definition of degraded performance stated above). The higher on the scale, the less frequent are the system deactivations and so the more fit the automation. Using collected and annotated data for the driving automation system to be assessed, the goal is to correlate both the automation indicators and the driving context with the number of occurrences of system deactivations/overrides/fallback initiations per time unit normalized on the automation fitness scale

The final outcome of the method is an estimation of an automation fitness score using both online observations of the automation indicators as well as online observations and predictions of the driving context. The estimation of the automation fitness score is then used to predict the time to automation (un)fitness using cut-off thresholds; for instance the time to automation unfitness (TTAU) would be the shortest time when the estimated automation fitness score becomes lower than a cut-off threshold.

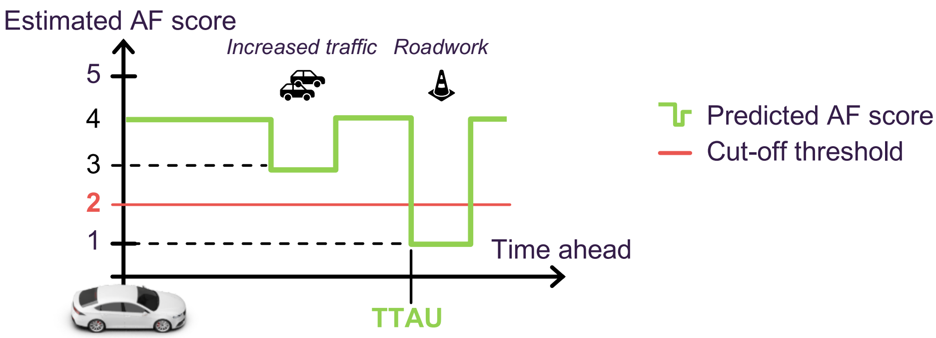

Figure 3. TTAU prediction using automation fitness score

As an example in Figure 3 above, the AF score (in green) is estimated in a near future using the predicted driving context which forecasts a limited increased traffic density followed later on by a section of roadwork (black icons). The AF score is estimated to decrease during these two sections based on the correlation built with the prior data collection. TTAU can then be estimated as the shortest time when the predicted AF score becomes lower than a predefined cut-off threshold (in red), which will happen when reaching the roadwork.

This research led to the main outcome of this deliverable, the functional requirements for the automation state module of the Mediator system. These are summarised in Table 1. Further refinements to the work presented here will be done as part of the actual development of the automation fitness module.

Table 1. Functional requirements for the Mediator automation state module

Functional requirements (software related)
AUTOMATION STATE MODULE
The system shall estimate *worst*, *likely* and *best*-case *time to automation (un)fitness* based on automation fitness estimates for the current driving context ▪ The system shall estimate the *current* automation fitness 　○ The system shall estimate the current automation fitness *score* 　○ The system shall access relevant information from the automation system to estimate the current automation fitness score ▪ The system shall estimate the *predicted* automation fitness 　○ The system shall estimate the predicted automation fitness *score* 　○ The system shall access relevant external driving context information ▪ The system shall estimate when the driving automation system is *unfit to drive* 　○ The driving automation system is deemed unfit to drive if it can no longer execute its defined dynamic driving task due to degraded automation performance (*low* automation fitness score) ▪ The system shall estimate the *time to automation unfitness* as the shortest time when the estimated automation fitness score becomes lower than a cut-off threshold ▪ The system shall estimate the *time to automation fitness* as the shortest time when the estimated automation fitness score becomes greater than a cut-off threshold ▪ The system shall estimate *worst*, *likely* and *best*-case scenarios of *time to automation (un)fitness* using the reliability of its inputs, both internal and external
The system shall determine the *active automation level* as either none, supervised (CM), or unsupervised (SB, TtS)
The system shall determine the *automation state class*, i.e. the reason for an upcoming change in automation availability
The system shall determine the *appropriate intervention type*, i.e. a possible way to improve the automation fitness
The system shall extract and collect *context relevant information* from the driving automation system to the context module
RECOMMENDATIONS
Drivers shall comply with the speed limit or lower
The automation state module shall know the planned route

1. Introduction

Vehicle automation has the potential to improve driving safety and driver comfort. To this aim, the Mediator system mediates between the driver and the automation on who is fittest to drive and supports the driver during his or her driving task. In order to determine who is fittest to drive, the Mediator system needs to estimate current and predicted driver fitness and automation fitness in a comparable manner. To this end, MEDIATOR needs to define numeral representations of driver and automation fitness; *time to automation fitness (TTAF)* and *time to automation unfitness (TTAU)* and the equivalents for driver (un)fitness.

The work described in this deliverable is focused on providing the foundations and concepts for the automation state module. The automation state module is able to judge the current status of the driving automation system (DAS) considering the current and near future context information. More precisely, to allow a judgement on the current and near future status of the driving automation system, a quantification of automation fitness is substantial. *Time to automation unfitness* is defined for the currently active automation level as the estimated time until the automation is no longer able to perform its driving task. Conversely, *time to automation fitness* is defined as estimated time before an available but yet inactive automation level is able to perform the automated driving task.

Part of the quantification for the time to automation unfitness is the definition and measurement of factors resulting in degraded automation performance. A factor is defined as a characteristic related to the driving automation system, which correlates with degraded performance.

Chapter 2 clarifies the details on what degraded performance of a driving automation system means. The chapter also provides two categories of factors relevant for the automation state module as they can be measured and predicted and therefore used as indications for an upcoming degraded performance:

1. Factors related to system input such as adverse weather, dense traffic or roadworks,
2. Factors related to internal states of the automation.

Chapter 3 gives an overview of the role and responsibilities of the driving automation system as well as driver's participation in the driving task. As there are no driving automation systems with SAE level 3 and 4 available in the European market to date (April 2021), some assumptions will be taken all along the deliverable for these levels of automation. In MEDIATOR, the distinction for use cases has been made between Continuous Mediation "CM" (SAE Level 2), Driver Standby "SB" (SAE Level 3) and Time to Sleep "TtS" (SAE Level 4). In this chapter, we will explain why we will consider SB to be a subcase of TtS, and SAE Level 3 and 4 to have the same technological requirements for the driving automation system (DAS). We will also introduce the concept of "supervised" and "unsupervised" DAS to provide an explanation of the reasons behind the technical choices made.

As illustrated in Figure 4 below, the automation state module relies on various information sources to feed its algorithms. This information is composed of:

- The driving context used to predict the occurrence of factors related to system input and detailed further in Chapter 4,

- Information from the DAS to compute performance self-assessment indicators used to detect factors related to internal states of the automation and discussed in Chapter 5. Some context relevant information from the driving automation system could also be provided as output to the context module (such as the current ego-vehicle position and speed).

Chapter 6 focusses on developing a methodology which combines both types of inputs to estimate the automation fitness and derive the various outputs of the automation state module. Such outputs will be based on *worst*, *likely* and *best*-case scenarios (similarly to the driver state module of the Mediator system) and built on the reliability of the inputs and outputs of the automation state module.

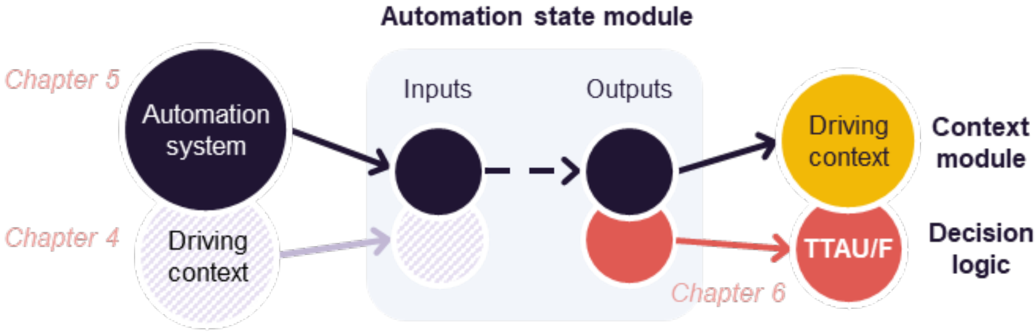

Figure 4. Inputs and outputs of the automation state module

Finally, Chapter **Fout! Verwijzingsbron niet gevonden.** summarises the main conclusions of this deliverable alongside open points, recommendations and assumptions to be taken into account during the development and evaluation of the automation state module. The chapter is concluded by the resulting functional requirements for the automation state module of the Mediator system.

2. Automation performance

To rate the automation fitness to drive, we first need to define what we mean by automation performance. Performance can be classified in a multitude of different ways, depending on the perspective taken. There is the performance of the automation as defined by the system designer, focusing on whether the system limitations are kept within the designed limitations of the system. Then, there is the performance of the automation as perceived by the user, which of course is highly subjective, but is not the same as the designer's perspective. Then, there can be things which is within the design, within what is ok for the user, but perhaps not optimal from a safety perspective.

2.1. MEDIATOR definition of automation performance

It is difficult to define good performance, more so than bad performance. Bad performance can be linked to risky situations and is easily defined with hindsight. Good performance may, on the other hand, be defined as good only until the context suddenly turns risky, which makes definitions difficult and problematic to put on a scale.

For the automation state module in MEDIATOR, we will define BAD automation performance as:

- A situation in which the driver disengages automation due to perceived ill-fitting responses OR
- A situation in which the automation shuts itself off OR goes into some fallback function due to within-system quality triggers OR
- A situation where the automation causes a crash with another road user or the infrastructure.

For GOOD automation performance, we will define this as:

- The driver is comfortable with the system's actions AND
- System behaving according to design in the targeted Operational Design Domain AND
- The system keeps a speed and distance that allows it to avoid or at least mitigate crashes (depending on the suddenness of the event).

Automation availability as well as assistance competence form a baseline of possible automation capability, which draws up the limits for automation performance. Automation performance is always judged within this, as a system cannot perform better than its specification.

2.2. Assessing degraded automation performance

In the context of this deliverable, a marker is defined as a manifesting characteristic of the driving automation system which correlates with BAD performance (see Section 2.1 above). To be able to use these markers to estimate degraded automation performance, we need to quantify and predict them. Markers for *degraded* automation performance are subjective and based on driver experience of "good" and "bad" automation performance. This can be contrasted with BAD and GOOD performance which we have explicit definitions of in 2.1. One example of a marker for degraded (but not necessarily BAD) vehicle performance is the vehicle accelerating without a known (to the driver) reason, or the vehicle driving over a lane marker if it is designed with the intention to always stay inside lane markers.

Degraded performance of the driving automation system (DAS) as it manifests in markers (system behaviour) is the consequence of internal or external conditions, which we will refer to as *factors*. Factors expose the functional limitations of the driving automation system and can have an impact on its fitness to drive (BAD or GOOD performance). Such functional limitations can result in degradation of the quality of the information that circulates throughout the interconnected components of the DAS and will introduce uncertainty, eventually affecting the automation fitness.

Taking the simplified system architecture of typical automation systems into account as illustrated in Figure 5, an external (or environmental) condition such as rain is a factor for degraded performance as it causes sensor interference, which is a functional limitation of sensors. This makes automation unable to accurately detect and predict the behaviour of road participants. If sensing (perception) is affected by adverse weather, the decision capabilities (planning and control) are also affected – and, thereby, the automation fitness to drive.

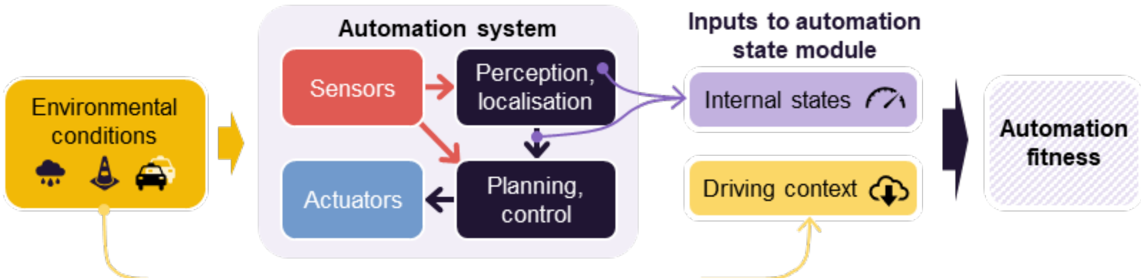

Figure 5. Simplified system architecture and inputs of the automation state module

Two categories of factors will be used by the automation state module and described further in this document (in particular the elaboration of their quantification methods):

1. Factors related to the DAS input as detailed in Chapter 4:
    a. Internal conditions, like hardware issues or sensor unavailability, which could be detected by the automation system (but possibly difficult to predict),
    b. External conditions, like rain, which could be quantified using a weather forecast online service,
2. Factors related to internal states of the automation, which could be quantified by performance self-assessment indicators e.g. a measure of how well a component is performing. These factors are described in Chapter 5.

The interest in using both categories of factors is that they have complementary strengths and weaknesses about fitness assessment. The factors relating to system input can be measured online and be predicted to some extent depending on the sources they rely on. The internal state related factors on the other hand are only measured online, but give a more accurate estimate of the current automation fitness since measured at the core of the system, but are dependent on the self-awareness of the system. Combining these factors and their quantifications to derive the time to automation (un)fitness (BAD or GOOD performance) predictions will be developed in Chapter 6.

Please note that there will be no attempt at detecting unknown unknowns, that is, situations where the driver may believe the system behaves in a non-optimal way but which cannot be detected by the Mediator system or automation system sensors. This instead forms a limitation of the automation state assessment.

3. Automation systems

Vehicle automation refers to systems that use steering, braking and acceleration to assist drivers in the task of operating a vehicle. SAE, the Society for Automotive Engineers, defines six levels of automation, SAE levels 0-5. In MEDIATOR, the distinction for use cases has been made between CM (SAE Level 2), SB (SAE Level 3) and TtS (SAE Level 4). In this chapter, we will explain why we will consider SB to be a subcase of TtS, and SAE Level 3 and 4 to have the same technological requirements for the driving automation system (DAS). We will also introduce the concept of "supervised" and "unsupervised" DAS to provide an explanation of the reasons behind the technical choices made.

3.1. Supervised and unsupervised DAS

In this section we will introduce the concept of "supervised" and "unsupervised" DAS, which are the two categories of DAS that will be used in the remainder of the report. First, we will make the connection between the SAE levels and the technical requirements on the system. Then, we will cover the human factors reasons behind that conclusion. Finally, we will discuss the availability of supervised and unsupervised systems on the European market at the time of writing (April 2021).

3.1.1. Connections between SAE levels and supervised / unsupervised systems

In this document, we will hereafter use two broad categories of DAS to clarify the system design and where the responsibility lies for driving safety. When the driver gets steering / brake / acceleration support from the DAS, we will refer to it as a *supervised* type system. When the automation operates the vehicle and the driver allowed to look away from the road and perform other tasks, we will refer to it as an *unsupervised* type system. In supervised systems, we include SAE Levels 1 and 2, and in the unsupervised systems we have SAE Levels 3 to 5. This distinction is made due to a need of determining the responsibility of the ADS and the driver in each of the cases, as we consider SAE Level 3 and 4 to have the same basic technological requirements for the driving automation system (DAS).

In SAE Level 3 a handover request can be given by the ADS to the driver at any time, e.g. due to unexpected changes in the driving conditions. Since the handover request can be given at any time, the driver of the vehicle need to be prepared to take over control within e.g. 10 - 60s and can hence theoretically relax while in this mode but not sleep. Possibly, this level might not have limited and explicitly mapped ODD (Operational Design Domain). That is, the vehicle can at any time request the driver to switch from unsupervised driving. The vehicle remains responsible for the driving task until the driver has completed the switch from to driving again. This would include "chauffeur" type functions for driving in traffic jams on motorways.

In SAE Level 4, a handover request is only given at planned locations/times or not at all, thus enabling the driver to relax or sleep. Examples of SAE Level 4 systems are "robot taxis" with a limited ODD, e.g. a certain part of a city, a parking garage, or others. An SAE Level 4 system can also constitute a system that parks a vehicle using full automation without any need for handover requests or a driver present in the vehicle.

In SAE Level 5 there is no need for handover requests ever; the vehicle is always able remain in control. The main difference between SAE Level 4 and 5 is that in SAE Level 5 systems can operate in an unlimited ODD. It is questionable if we formally will ever experience an SAE Level 5 system since no manually driven vehicle can operate in an unlimited ODD today, e.g. in theory ranging from highways to tight parking spaces, off-road driving on icy paths, hill climbing in the desert and driving on Mars.

3.1.2. Human factors aspects of supervised and unsupervised DAS

Research on the human factors as well as automation side of DAS has advanced since the initiation of the MEDIATOR project. Human factors research has over the years pointed out issues with hand-over requests in SAE Level 3, where the driver is expected to take back control in e.g. 10 seconds (starting with e.g. Blanco, Atwood, Vasquez et al., 2015). In a SAE level 3 DAS, the system is responsible for both executing and monitoring the driving task, and for asking the driver to take over when it assesses it cannot handle a situation (the driver is responsible for fallback). There is however no requirement on the minimum time between the request of such a handover and the time the system needs to detect a system failure. If such a time is not specified, a 0 second take-over-time must always be assumed for the worst case. Additionally, the driver may not necessarily take control as instructed.

These two cases lead us to the conclusion that drivers cannot be allowed to perform other tasks if they simultaneously may need to handle a safety-critical situation at any moment. As a 0 second take-over-time or a lack of driver action at the end of a transition period may cause a risky situation, we do not accept this behaviour from the system. Instead, we will need to design a Level 3 system with a fallback.

For the automation state module, we have therefore decided not to consider the MEDIATOR 'driver standby' (SB) level as a level separate from 'time to sleep' (TtS) for unsupervised systems. From a technological standpoint, the system capabilities for SB and TtS are the same. If the DAS needs to handle the fallback task, the DAS needs to know its ODD and be fully capable in all ways to handle the fallback task safely within that ODD. In neither SB nor TtS we can fully rely on the driver being ready and available to resume control – the DAS always needs to be able to cope alone should something happen to the driver. For an unsupervised system in an evolving situation (context) where the system would need to enter a fallback, a driver might also prefer to resume control and continue driving instead.

There are high demands placed on the HMI in the DAS to ensure that the driver is capable of resuming control when choosing to do so, to be fully aware of the need to monitor or not, how a transition to/from unsupervised driving is made, and so on. The MEDIATOR project is providing essential research on some difficulties with the SB subcase of unsupervised driving, like driver out-of-the-loop or driver on-the-loop and how it can be mitigated by HMI design (see previous work in MEDIATOR). These same issues arise in a TtS situation, and in neither of the two cases we know exactly how far ahead in time we might have a risky situation the driver needs to handle. Either way, for a driver to be allowed to resume control, something like the Mediator system will be essential to identify driver readiness, to support the driver in transitions, and to design automation to facilitate the transitions between system and driver.

3.1.3. Availability of supervised systems

All DAS available in Europe at the time of writing (April 2021) are supervised systems. In other words, they require that a driver is monitoring traffic as well as system performance and intervenes if necessary. This maps to the Mediator term "CM", "Continuous Mediation", where the project focus lies on finding and maintaining the optimal task load for the human driver. A CM or supervised system in Mediator terms may fail at any given time, so the driver always needs to be there and aware of traffic as well as the system's capabilities or lack thereof.

Table 2. Available supervised ADS

Available supervised DAS	Driver supervision	Fallback functions	Included in MEDIATOR?
Active distance + steering assist *Speed range: 0 – 130 kph*	Hands on wheel required	If hands off wheel detected, speed reduction (but steering continues with possible stop in lane functionality)	Yes
	Hands off / Driver eyes on-road required Mapped to highways only (GM Supercruise)	While still active, inform driver to look back at the road if eyes off-road glance duration > x seconds	No, not currently on the market in Europe

3.1.4. Availability of unsupervised systems

To date (April 2021) there are no unsupervised systems on the market in Europe for personal vehicles. Unsupervised systems comprise both MEDIATOR terms “SB” (driver standby) and “TtS” (time to sleep, driver long time out of the loop), and both allow the driver to completely disengage from the driving task for a shorter or longer period of time. In terms of SAE levels, this corresponds to SAE L3 and upwards. An unsupervised system requires that the automation system is available and competent for every case where it allows itself to be activated, as it cannot rely on the driver to resume control in a timely fashion. An unsupervised system always needs to have a fallback strategy if anything unexpected happens, as it cannot rely on the driver will resume control in a timely fashion.

3.2. Implications for automation performance estimation

For automation performance and state estimations, the automation will need a much better self-estimation for unsupervised driving than for supervised driving.

For supervised driving, the driver has the final say for vehicle behaviour. Thus, speed recommendations might not need to be followed, and if not followed may result in worsening automation performance and a need for the driver to be more alert. For unsupervised driving, the automation system is responsible for the fallback task. Specifically, the unsupervised system will need to cope with changes in environmental conditions, speed changes and other actions, and automation performance will always need to be good enough within the bounds of what automation decides. This will not be possible for the driver to challenge without resuming control from automation. For an unsupervised system, the driver might not appreciate the fallback task initiated by the system, and choose to resume control instead.

For automation state estimation, we will as previously explained consider SB and TtS in the same “unsupervised driving” category. Any system which allows the driver to be out of the driving loop by performing a different task will require the same high level of performance within its operational design domain as one which allows the driver to sleep. It will also need to be able to perform a fallback task if the driver does not resume control when asked.

In addition to the automation state estimation, i.e. when and where supervised or unsupervised system functionality should be made available to the driver, the automation needs to be able to estimate its automation performance while in supervised driving mode. The automation performance in supervised systems is, simply put, how capable it is at executing a Dynamic Driving

Task (DDT). The better the performance the vehicle can provide, the higher the requirements on a HMI solution to keep the driver in the loop and doesn't make them "check out" (Victor et al., 2018). Moreover, the automation performance level might be needed in order to change the HMI or even alter system behaviour to provide other types of information for anticipated or detected changes in the automation performance, e.g. due to changes in external conditions or the types on road the vehicle is currently driving on.

For the unsupervised mode, automation state estimation will need to assess the time available until the initiation of a fallback task, which the driver may or may not allow to continue. If the driver does not wish the car to enter the fallback, they will need to resume control instead. Thus, automation state estimation for the unsupervised systems will focus on determining the end of the ODD. The methodology will by and large otherwise remain the same.

4. Factors for degraded performance

A functional limitation is a technology constraint that can originate either in the system design (sensors chosen for the system, elements defined and taken into account in the decision making process, maximum torque or acceleration), or in the physical property or resolution limits of the components (e.g. cameras becoming blurred by rain).

A factor for degraded system performance is an internal or external condition, which exposes the functional limitations of the driving automation system (DAS) and has an impact on its fitness to drive. Functional limitations originate in the component(s) directly affected by the factor, resulting in effects which propagate throughout the interconnected components of the DAS architecture.

As an example, an external (or environmental) condition such as rain is a factor for degraded performance as it causes sensor interference, which is a functional limitation of sensors. This leads to the fact that with bad enough sensor performance, automation cannot accurately detect and predict the behaviour of road participants, affecting the decision capabilities (planning and control) – and, at some point, the automation fitness to drive.

Factors and their impact on performance are described further in sections 4.1 and 4.2. The first section discusses the functional limitations of sensors, which are an important contribution in degraded performance. The second section summarises these limitations alongside with detailed examples of factors affecting performance.

Assessing the occurrence of such factors ahead of time is important for the Mediator system to estimate the time to automation (un)fitness as they are indications for potential degraded system performance. Section 4.3 focusses on establishing a quantification of the external environment, the *driving context*, which can be used to predict the occurrence of factors.

From the automation perspective, performance self-assessment is also a valuable input for the automation state module. Self-assessment can continuously highlight degraded performance, not only when the system is presumed to undergo the effects of factors estimated through the driving context. Chapter 5 will address such self-assessment capabilities and introduce self-assessment indicators that could be used to quantify the automation performance.

Finally, Chapter 6 will outline a methodology to estimate the time to automation (un)fitness using both the factor assessment and self-assessment indicators, and which will be included in the automation state module.

4.1. Sensor limitations

Driving automation systems use a multitude of sensors to interpret the surrounding environment near the vehicle. Perception uses sensors to continuously scan and monitor the environment, similar to human vision and other senses. Localization and mapping algorithms calculate the global and local location of the ego-vehicle and map the environment from sensor data and other perception outputs. In general, robust and reliable perception, localization and mapping are required in order to make accurate and reliable decisions for vehicle control.

All sensors, including the most intricate sensor technologies, such as radar (RAdio Detection And Ranging), lidar (LIght Detection And Ranging) and camera, have points of strength and weakness in relation with perceiving environmental situations. All mentioned sensors are based on electromagnetic waves and they thus all share similar properties. The referred weakness are due to the selected implementation (design), which is based on automotive sensor requirements (including cost, size, weight and power consumption).

In Table 3, main advantages and disadvantages of different types of perception sensors for automotive applications are listed.

Table 3. Advantages and disadvantages of perception sensors (from Mohammed et al., 2020)

Sensor	Advantages	Disadvantages
Radar	• The sensor makes it possible to see for long distances ahead of the car in poor visibility conditions, which can help avoid collisions. • The sensor is small, lightweight, and affordable. • The sensor requires less power than a lidar sensor since it has no moving parts. • The sensor is more robust to failure compared to lidar. • Radar is less expensive than lidar.	• The obtained images have low accuracy and resolution. Information on detected objects is limited (such as neither precise shape nor color information). • Increasing power may solve radar attenuation in a precipitate environment, but increasing power is not a viable economic solution. • The mutual interference of radar sensors is a growing issue. • The azimuthal and elevation resolution of automotive radars is poor, and this makes the detailed mapping of scenes and object classification difficult and error prone. • The sensor cannot give a 360^0 measurement of the surroundings.
Lidar	• The sensor can see long distances ahead of the car in good visibility conditions (neither rain nor fog). • The sensor can provide a full 360^0 and 3D point clouds. • The images have good accuracy and resolution. • There are no significant interferences in multiple lidar sensors.	• Lidar is more expensive than radar and camera. • Transmission is sparse (not dense), due to which small objects (like wires and bars) remain undetected. • Due to oscillating components, mechanical maintenance is high. • When detecting wet surfaces, lidar shows poor discrimination of contrast compared to dry surfaces. • The sensor requires more power than a radar sensor. • The sensor is influenced by varying climatic conditions.
Ultrasonic	• Ultrasonic sensors are useful for the detection of transparent objects and non-metal objects. • Not influenced by varying climatic conditions. • Low in cost and small in dimensions. • At short ranges, higher resolution can be expected. • Unlike cameras, ultrasonic sensors overcome pedestrian occlusion problems.	• They are available for short-range distances only. • Sensitive to temperatures and windy environments. • Interference and reverberation are problematic when two ultrasonic sensors operate in two cars or are placed close together. • Noise from environments may interfere with measurements.
Camera	• Cameras maintain high resolution and color scales across the complete field of view. They offer a colorful perspective of the environment that helps to analyze the surroundings. • Stereo cameras can provide a 3D geometry of objects. • Cameras can robustly monitor and maintain information from surroundings over time. • They are small in dimensions. • Compared to lidar, they are cost-effective and easy to deploy on a vehicle.	• The camera data require a powerful computation system to extract useful data. • The sensor is sensitive to heavy rain, fog, and snowfall, which reduces the capability of the computer system to reliably interpret the surrounding scene. • The distance to obstacle accuracy is limited.
Far-Infrared	• Far-infrared (FIR) camera images depend on the target temperature and radiated heat. Therefore, light conditions and object surface features do not influence them. • Compared to lidar, FIR sensors are cheaper and smaller. • They have improved situational awareness at night. • FIR sensing range can cover up to 200 m or more horizontally and detect possible hazards ahead. • They have a better vision through dust, fog, and snow compared to cameras.	• FIR camera data require demanding computation sources and robust algorithms to extract useful data. • This sensor is expensive, compared to Charging Coupling Device (CCD) or the Complementary Metal Oxide Semiconductor (CMOS) cameras. • The resolution of the FIR camera is low in comparison to the visible camera and provides images in grayscale. Due to this, fast-changing moments of objects are quite challenging to detect and classify in real time. • Since FIR systems calculate based on temperature differences, it is often difficult to distinguish between specific targets of interest in cold climate scenarios. • Partial occlusion of the target causes classifiers to ignore the target (like a pedestrian standing behind a car or a group of pedestrians overlapping each other). Solutions to overcome this problem have been studied and, besides, they cannot provide information about the distance to obstacles.

4.1.1. Radar

Radar uses radio waves to detect objects and determine relative positions and relative speeds, and most current automotive radars are capable of measuring target angle (in azimuth). Radar performance in adverse weather is degraded compared to in good weather conditions. Rain, wet snowfall and spray from other vehicles have the most impact, causing a reduced range of radar detections and interference possibly creating of false targets. Although the radar performance in wet surroundings degrades, it outperforms lidar and cameras under the influence of rain. The radar is not impacted by the presence of airborne particles (e.g., dust and smoke) because of its wavelengths, which are much larger than the characteristic dimensions of dust.

Radars may be a great option for all weather conditions, but active signal interference is still a matter of concern. For instance, in high-density traffic conditions, radar systems may pick up other vehicles' radar signals, causing false detections, interference and uncertainty. Beyond weather conditions, radar handles darkness conditions well, but it has poor resolution, making it difficult to distinguish pedestrians, especially children. Radar also cannot reliably detect stationary objects (e.g., pedestrians waiting to enter a roadway)

4.1.2. Lidar

Lidar uses scanning lasers to measure distances to surfaces, producing a three-dimensional map of detailed shapes which can be constructed after moving through the scene and taking multiple snapshots (also true for radars and cameras). Lidar is capable of object detection in low/no-light conditions, but like cameras, is unreliable in adverse weather and when road surfaces are wet or reflective. Lidar performance in extreme weather conditions is not as strong as expected, because adverse weather conditions increase the transmission loss and decrease the reflectivity of the target. In particular, fog has the greatest impact on the ability of lidar, more than snowfall and rain. Performance can also be reduced with background illumination such as the one coming from the sun (especially at lower wavelengths).

The challenge in some fog conditions (which varies depending on a combination of fog molecule size and density, and the electromagnetic wavelength) is that many transmitted signals are lost, resulting in reduced reflected power from the target. Reduced power alters the signal-to-noise ratio of the lidar sensor and reduces the likelihood of target detection, which leads to degraded perception performance. The raindrops' intensity, size and shape drastically influence the attenuation rates of lidar. As rainfall intensity increases, the likelihood of lidar false-positive errors increase (although there are techniques of mitigating the effects of rain).

4.1.3. Camera

The human driving task is based on the visual analysis of the surrounding vehicles, obstacles and road signs and cameras can provide some of this information useful for automated operation. The camera can acquire the image and record contour, texture, colour distribution and other information of the object from a certain angle. Therefore, cameras are used to complete target recognition and target tracking tasks, including lane detection, pedestrian and vehicle identification and local path planning, but the camera is very affected by the adverse climatic conditions. In an aerosol environment, the camera decreases its visibility and contrast, and it is unreliable for object recognition. Moreover, a camera is not recommended for environmental detection and vehicle control tasks in foggy weather.

Snow also affects the mechanical operation of the camera when positioned outside the vehicle. It is also difficult to use data generated by the camera for lane detection, due to frost or droplets of moisture on the glass in front of cameras if mounted behind the front windshield. Moreover, the camera functions poorly in low-light conditions (from dusk to dawn), on slick surfaces where there are problems of glare and in low standing sun scenarios. The camera also requires a reasonable contrast between the target and the background.

4.1.4. Other limitations

Finally, yet importantly, the system could experience some technical problems that could at least partially compromise the proper functioning. Driving automation systems need to be able to detect and identify their sensor and perception failures and/or degradations to safely manoeuvre the vehicle accordingly to the level of the failure.

Sensor fusion permits the use of multi-sensor information to calculate, recreate the environment, and generate dynamic responses, resulting in a consistent and accurate representation of the vehicle's surroundings and position for a safer navigation: some sensors may be redundant under some environmental conditions and other may be complementary, assisting in positive cooperation to reach an accurate obstacles detection. This means that just because one sensor type or one individual sensor is degraded, the system as a whole might be able to balance and still have good environmental perception.

When automated systems have to rely on the visual road markings (e.g., road lane lines), under heavy rainfall and snowfall, reliability is an issue not easy improved. Ultrasonic sensors combined with short-range radar can enable the vision system to have an increased performance in lane marking detection.

To overcome the single technology gaps and to increase the DAS perception robustness a fusion perspective is advised, to make the DAS be able to cope with different environmental driving scenarios (e.g., weather, light, road geometry, road changes…). Perfect perception will always be dependent on what the different sensors have been trained on, and gaps may always remain.

4.2. Factors and functional limitations

A factor for degraded performance generally affects one or several components of the automation by exposing the functional limitations of these components. Such functional limitations result in degrading either the quality of the information that circulates throughout the interconnected components or introduces uncertainty, eventually affecting the automation fitness.

Taking the (simplified) system architecture of typical automation systems into account as illustrated in Figure 5, environmental factors do not point to a specific component since any of them can be directly affected depending on the factor in question. However, once a component is affected, the effect will always propagate following the same interconnected components. Each automation system relies on a specific design chosen by the DAS provider (e.g., sensors, system architecture, component logic and limitations) to fulfil its driving task. Functional limitations, and therefore also factors for degraded performance, are specific to each automation design.

As a first example, if an external factor reveals functional limitations in a sensor, then the effect might propagate through the sensor fusion to the perception (and localisation) component and to the decision making that might stop controlling actuators. Another example of a functional limitation is the maximum steering torque assistance; the steering torque is measured via a vehicle sensor and could exceed a threshold in some conditions e.g. high ego speed in a sharp curve. If so, the decision making may deactivate the steering assistance. Such factors do not affect sensors or perception, but still affects the performance of the automation.

The factors and functional limitations identified in Table 4 are based on Table 4.1 in D1.1 (Christoph et al., 2019) which summarises general knowledge on actual supervised systems and on sensors limitations discussed in Section 4.1. The content will vary depending on DAS provider and, on the type and ODD of every DAS.

For readability, the factors in Table 4 are grouped based on their origin (in bold). It must be noted that factors with the same origin might not lead to the same functional limitation (examples of functional limitations are often listed).

Table 4. Factors and functional limitations

Factor	Functional limitations
EXTERNAL CONDITIONS	
Weather conditions e.g. rain, snow, fog	Interferes with or reduces maximum range and signal quality of common perception sensors
Light conditions e.g. glare, darkness	Reduces maximum range and signal quality of camera systems
Obstructions e.g. dirt or ice, surrounding buildings, tunnel	Interferes with or reduces maximum range and signal quality of common perception sensors
Road design e.g. sharp curves, intersections, degraded lane markings, roadworks, holes and bumps	Affects lane detection of camera systems, impacts system decision capabilities
Traffic & Road participants e.g. vehicle cutting-in, dense traffic, vehicle standing still in lane	Interferes or reduces maximum range and signal quality of perception sensors, impacts system decision capabilities
INTERNAL CONDITIONS	
System decision and control e.g. speed out of range, excessive steering torque	Features might only function under certain ranges or within other system limitations (e.g. acceleration, steering torque)
Hardware & Connectivity issues e.g. sensor unavailable, hardware issues	Impacts sensors operability or system decision capabilities

A description of some use cases is provided below in which factors from Table 4 and their functional limitations are illustrated:

A. Roadwork (road design): a *roadwork* is a possible location for an unstructured road network with suboptimal lane markings (in the form of remaining of old lane markings or absence of lane markings). Such condition will affect the lane detection of the camera system, directly influencing the perception / localisation of the ego vehicle and the decision making capability of the DAS. However, there are roadworks with a structured road network, so all roadworks do not affect the automation performance equally. Moreover, not all roadwork locations might be available from the external sources of context information to be used by the automation state module, affecting time to discovery.
B. Vehicle standing still in lane (traffic and road participants): a *vehicle standing still in lane* can directly affect the decision making component if there is no possibility to overtake the stationary vehicle (sensors, actuators and perception are generally not affected unless the stationary vehicle is revealed suddenly at high ego vehicle speed). Such situations will require the driver to take over or the system to eventually activate collision avoidance features.

C. Excessive steering torque (system decision and control): the automation could be limited in the amount of *steering torque* it supports the driver with (for the lane centering assistance) which could be exceeded while in a sharp curve or a roundabout, for instance. Such a situation can lead the decision making component to deactivate the steering assistance. Predicting the steering torque (e.g. by identifying curves on the route in relation to the predicted vehicle speed), could help determine locations where the steering torque threshold could be exceeded.

There are as previously mentioned, two types of factors, internal and external:

- Internal factors that originate for instance from sensors' hardware issues or sensor unavailability. Internal factors concerning components such as the perception or planning/control, can also be responsible for degraded performance and are addressed in Chapter 5 as part of the automation performance self-assessment.
- External factors, which can cause the automated driving system to reveal its functional limitations by reaching ODD limitations through e.g. heavy precipitation, sand storms or sharper curves than expected. External factors may also be things such as lack of lead vehicles, bad lane markings, or other aspects related to the driving context.

Considering each factor independently does not indicate how the contribution of other factors, which may occur simultaneously, is affecting the automation. It is, for instance, unknown how a section of "roadwork" affects the automation in combination with differing weather conditions, light conditions or traffic densities. It is therefore necessary to assess the factors as a combined entity forming the driving context, as defined in introduction of this chapter, which will be used when estimating the time to automation (un)fitness in chapter 6. Building such driving context is developed in the following section.

4.3. Driving context

Assessment of the driving context involves the collection of pieces of external information needed to assess each factor's relevance and weight for determining the capability of the DAS.

Building the driving context is achieved in three steps:

a. Identify the relevant factors believed to affect the automation system,
b. Identify the necessary pieces of context information needed to assess the current state for each factor, the whole forming the driving context. Additionally, assessing the confidence in how these pieces of context information are obtained is also valuable for discussing the effectiveness of the predictions,
c. Finally, to make predictions of the driving context ahead, the automation state module will be helped by knowing the route.

a. Identify the relevant factors:

Let's assume that the relevant factors are the three use cases described in the previous section: roadwork (A), vehicle standing still in lane (B) and excessive steering torque (C). This simplification is made only to demonstrate the principles for building a driving context but still could be used to generalise to larger driving context.

b. Identify the context information:

The pieces of context information are mainly extracted from external sources (cloud-to-vehicle or vehicle-to-vehicle) to the automation system such as an online service providing weather forecasts. The pieces for the three use cases are described in Table 5.

Table 5. Factors and associated context information

Factor	Piece of context information	Source
A. Roadwork	Roadwork location	Known locations e.g. using a HD map(1)
B. Vehicle standing still in lane	Vehicle location	Known locations by cloud-to-vehicle or vehicle-to-vehicle communication
C. Excessive steering torque	Steering torque measurement	Prediction of steering torque could be inferred from: + Predicted speed (see next row) + Road curvature using HD map However, such prediction might not be reliable enough since steering torque also depends on other characteristics than speed and road curvature. Alternatively, known locations of possible exceeding could also be used: + Sharp and S-curves using HD map + Roundabouts using HD map
	Ego speed	Predicted speed at location using 1. Traffic speed forecast API 2. Speed limit from HD map
	Steering torque threshold	Available from the DAS documentation

(1) The term “HD map” is a cloud-to-vehicle service which provides enriched map information about the road network.

Some of the sources of context information shown in Table 5 may be difficult to obtain with a good enough level of confidence for the following reasons:

- The source might be inaccurate. Roadwork locations could be available reliably for lasting roadworks whereas mobile or temporary roadworks might not be available consistently.
- Weather forecasts can provide information such as “chance of rain: 40 %”, so how to assess if it is going to rain or not?
- Some pieces of information can be estimated from other pieces, thereby degrading the overall confidence of the estimation because of the method or assumptions used to derive the estimated pieces, like for the steering torque

c. Additional inputs:

Additionally, the automation state module might need to know two other inputs to predict the driving context ahead:

- The planned route,
- The (estimated) time to reach any location on the route.

Without any planned route, the automation state module is limited to the immediate options for driving, for instance by the next intersection ahead, since multiple paths can be taken afterwards. A lack of a known, planned, route would make the development of this module much more difficult and is best left for a later stage. For the purpose of this deliverable, therefore, we will hereafter always assume that the system has access to a known destination with a defined route, as it is essential in order to predict the driving context with a larger time horizon.

Another major piece of information needed is the predicted time to reach any location on route. Knowing the distance to a certain location on route, e.g. the next occurrence of a roadwork, is useful information to predict the distance to a location that may lead to degraded performance.

Determining the time when the ego vehicle will reach for example a roadwork is complex as it depends on the ego speed along the route, itself dependent of the driving style as well as the traffic conditions. One possible solution is to subscribe to an online service which provides an estimated time to arrival at a specific location on route, usually using traffic speed collected from other users reporting to the service, and which represents an average time over several users. This might therefore differ from what will be observed by the ego vehicle. Another method, using HD map only, could be to extract the speed limit on the road segments to compute a lower bound of the estimated time to reach the roadwork (assuming the driver does not exceed speed limits). With this method, the estimation would likely be less accurate than using the online service, but it would give the shortest time to reach the location. Therefore, this could be employed as a conservative approach. Both methods are relevant depending on the strategy considered by the automation state module.

4.3.1. Predicting the driving context

The prediction of the driving context is an essential part of the automation state module in estimating the time to automation (un)fitness which will be described in chapter 6.

Developing such driving context prediction does come with challenges in finding the relevant pieces of context information, possibly combining them to derive other pieces, and assessing the confidence in these predictions. Associated methods will be developed and evaluated further during the implementation of the automation state module.

The prediction of driving context is used in the *worst*, *likely* and *best*-case scenarios of the outputs of the automation state module. Confidence or probability measures that certain conditions will occur, will be used in order to make the distinction between the different scenarios. For instance, if the weather forecast along the route is “chance of rain: 40%”, the driving context for the *worst* case could consider heavy rain, the *likely* case could assume light rain whereas the *best* case will presume it will not rain. Again, the exact logic for this cannot be determined at the present stage. Those decisions will be made when defining methods, testing and validating their implementation, as part of the development of the automation state module.

Internal factors such as hardware failures, which are not part of the driving context, cannot be predicted although they might be detected by the automation system. In the case that such internal factors would appear, it may or may not be known by the automation state module.

Lastly, a proposal for the relevant context information that could be used to predict the driving context is described in Table 6. It contains the pieces of information provided above for the three use cases as well as information sources required for other factors in Table 4 that have not been detailed previously.

Table 6. Driving context information

Piece of context information	Source
Road attributes: Road type Number of lanes Lane width Speed limit	Known attributes by HD map
Location of road items: Intersections (on-ramps, off-ramps, weaving sections, roundabouts) Sharp and S curves Tunnels, tolls Roadworks	Known locations by HD map
Weather conditions	Weather forecast API
Ambient light conditions	Ambient light conditions API
Traffic density and speed	Traffic forecast API
Vehicle standing still in lane location	Known locations by cloud-to-vehicle or vehicle-to-vehicle communication
Ego speed	Predicted speed at location using 1. Traffic speed forecast API 2. Speed limit from HD map
Steering torque	Prediction of steering torque could be inferred from: + Predicted speed + Road curvature using HD map
Heading	Could be derived from GPS trace in planned route
Steering torque threshold	Available from the automation system
Other system limitations (road type, speed range)	Available from the automation system

5. Assessing automation performance

This section discusses automation performance assessment with the focus on self-assessment. The objective is to define performance self-assessment measures, which could be used by the automation state module to estimate the automation fitness and understand the capabilities and limitations of that assessment.

5.1. Automation availability and assessment basics

To determine whether a driving automation system (DAS) is available, the system will assess its own ability to drive in every given second. Activation is only permitted when sufficient information is available. Whenever the system is then switched on, it is continuously evaluated on whether it has enough sensor inputs, and will disengage when sensors or sub-systems are failing to provide required information.

The most basic version of system self-assessment will be based on current sensor and sub-system information. To know when supervised driving assistance fails or unsupervised driving will come to an end, the system will also need to have a continuous Operational Design Domain (ODD) assessment. Thus, determining what the system knows about itself and its limitations is an important aspect when designing an availability measure.

The availability is not necessarily the same as the system performance. Availability will be dependent on the sanity check of the system itself, and its assessment of whether it is getting enough reliable information to start. A system may be available, but at the same time according to the user not perform very well.

A supervised system may decide it is available based on the here-and-now information, but an unsupervised system will need to be much more certain of its future performance. What is classified as ‘reliable’ performance as well as automation fitness will always need to be related to the actual system in question. Systems have different performance depending on their design – not always possible to relate on the same scale. Thereby, what is ‘reliable’ performance will always need to be defined and qualified within an ODD before trying to assess system ‘degradation’. In addition, reliable system performance will be dependent on the intended capabilities of the system, where performance will be different between different set-ups.

One precondition for reliable performance is that the conditions that matter for the system in question are possible to detect. Any system will therefore need to detail what it can be expected to detect and what it cannot. In a very basic sense, this means that to assess automation performance we will need to detect as many unknown unknowns (to the DAS) as possible. As an automation system is designed today, less than perfect system status means the system will try to perform anyway until it can’t, and when engineers see the issue they try to fix it for the next version of the system. Detection of bad performance is generally done through annotating actual drives in the real world.

Thus, the first and most important factor for reliable performance (“GOOD” performance as detailed in Section 2.1) is the current conditions and stability in road conditions including lane markers, light conditions, traffic, and weather. There will always be some uncertainty about system performance when conditions change, and whether the system is able to detect that change or not.

5.2. System information needed for assessment

Some items are necessary to know in order to monitor the DAS capabilities:

- What are the operational limits of the system (torque, braking, accelerating)?
- Does the system have a geofenced ODD? In that case, what is it?
- Where are the sensing limitations of the system? (Distance, angles outside the vehicle, weather, light conditions etc.)
- What infrastructure can the system or sensors identify? (Roadworks, temporary markings, etc.)

System limitations can be annotated during test drives over time to feed into algorithm development. A continuous monitoring of automation capabilities requires some kind of input into the system that it is performing well or less than well according to a driver, and here personal preference will play a large role.

This could be done by a driver, annotating somehow that they are disengaging the system due to bad performance during a drive. This would help the system distinguish between deactivations and overrides because the driver wishes to do so, and deactivations or overrides due to worsening system performance in the traffic conditions.

5.3. Estimating current automation performance

As discussed in Section 2.1, the measure for automation performance is the number of driver overrides (or system deactivations or system fallbacks if applicable) per hour. As these measures need to be recorded over extended periods of time, they are not an optimal measure to calculate current and upcoming performance.

When the driver overrides the system, it can be due to an uncomfortable or (perceived) unsafe behaviour. These system behaviours in turn are the result of either too little or too much acceleration or steering for a certain context. It could be said that if the automation fitness is at its highest level, there is by definition no unwanted behaviour, and no driver interventions are needed. However, there may of course be interventions that are not related to automation behaviour as well, such as driver preferences. By studying driver interventions, lower levels of automation fitness could be estimated by looking at the factors that cause automation behaviour that leads drivers to intervene (that is, BAD performance). The behaviour of a DAS is the result of two main elements, perception (how well can the system perceive the world around it) and decision logic (given a view of the world, what action should be taken).

Given perfect perception, the implemented decision logic determines the ODD of that particular DAS. For example, regardless of whether a red traffic light is detected or not, if the decision logic does not consider that input, it will not stop for it. In this case, traffic lights are not part of the ODD of the DAS. If the system was designed to stop for red traffic lights, but the perception system failed to detect that there was a traffic light, or that it was red, this would cause a ‘failure’: the system would not brake the car to a stop and thus it would not behave as designed.

As the fitness to drive of the automation is judged by the rate of failures, or the rate of needed driver interventions or overrides, estimating the current automation fitness can be achieved at least partly by looking at the current perception performance, and the factors that influence it. This definition also entails that some automation systems may never get a high automation fitness rating in any ODD, as drivers consider their performance to be poor most of the time. There is always a higher chance of the driver having to intervene in scenarios where systems have not been developed to identify and handle specifically. This lack of performance cannot be estimated any other way than by iteratively test-driving and collecting examples for every specific system.

5.3.1. The influence of perception

In the example above the perception system completely missed detecting the traffic light and its colour. Similarly, the DAS could fail to detect a nearby pedestrian or a stationary vehicle. The opposite could also be true, the DAS could detect a pedestrian or vehicle where there is none. An important factor therefore is the “Perception Confidence”: how sure is the DAS that the perception is true, that an object, lane marker, or other detection exist where detected and how sure is it that none exist where none are detected.

Another scenario to this example would be a late detection, meaning the traffic light was detected but at a point where hard braking would be needed in order to stop for it. Therefore, the current “Perception Range” (the distance at which the world can be perceived) versus the “ideal” range (that the system was designed for) is an important factor for automation fitness.

The proposal in this report is therefore to estimate the current automation fitness by creating a measure for perception range and perception confidence. If deemed necessary, other internal states could be added during development.

5.3.2. Perception Range

The perception range can be explained as “how far can the DAS perceive the world”. In an ideal situation with clear skies, perfectly functioning sensors and low traffic density, the DAS should be able to detect the road and other traffic users around it at a long enough distance to make accurate decisions on what to do. For example, detecting an upcoming curve with the road bending to the right gives the DAS ample opportunity to plan a smooth path following the road. Conversely, a very short range of road estimate means a curve could suddenly pop up requiring an uncomfortable amount of steering or braking to adjust for the curvature. In other words, a perception range close to the ideal range that can be expected from the sensors should be one indicator for a high automation fitness.

The perception range itself can be estimated by the furthest distance detection of any sensor that can be fused into usable and reliable road or object data. An example could be a lane marker detected at 50 meters in front of the vehicle and used by the perception fusion system to get a road model. Another example would be a cluster of radar detections that are fused into an object detection at 100 meters. Given a detection in the ego-vehicle reference frame, at longitudinal distance x and lateral distance y, the sensor range R can be computed as:

$$R_{sensor\ i} = \max\left(\sqrt{{x_1}^2 + {y_1}^2}, \dots, \sqrt{{x_k}^2 + {y_k}^2}\right)$$

As different sensors are impacted differently by context and environmental factors mentioned earlier in Chapter 4, a conservative estimate of perception range could be made by instead using the shortest of the fused perception ranges (road, objects or free space, for example).

$$R_{perception} = \min\left(R_{sensor\ 1}, \dots, R_{sensor\ k}\right)$$

5.3.3. Perception Confidence

The perception confidence is the measure of how sure the DAS perception system is of the presence or absence of a detection. Both are important and cause two different types of false behaviour: false positive and false negative behaviour. In the case that an object does not exist, but the perception system reports one, it can cause unnecessary braking and/or steering, which would be called false positive behaviour. The connected confidence measure is then the True Positive Confidence. In case an object is not detected the DAS would not steer to avoid and/or brake for it if necessary, which would be false negative behaviour. The connected confidence measure for that is the True Negative confidence.

For the first measure of perception confidence, the True Positive (TP) confidence, the sources of detection will be looked at. Given that the DAS has access to different sensors, the more sensor detections confirm an object, the higher the confidence in the detection of that object is. If all reported objects are confirmed by all available (and applicable) sensors, then the TP confidence will have maximum value. The more objects reported and confirmed by less than all available sensors, the lower the measure will be, as the chance of false positive behaviour rises. For object i, confidence $C_{TP,i}$ will be:

$$C_{TP,i} = \frac{k_{confirming\ sensors,i}}{n_{applicable\ sensors,i}}$$

And overall confidence C_{TP} for total number of objects m:

$$C_{TP} = \frac{\sum_{i=0}^{m} C_{TP,i}}{m}$$

The second measure, the True Negative (TN) confidence, is the inverse of the number of dissociated detections; those detections made by one or more sensors but deemed false and therefore discarded by the ADS perception fusion system.

$$C_{TN} = \frac{1}{1 + c_1 \cdot n_{discarded}}$$

Where c_1 is a tuning parameter to adjust for the noisiness of a certain sensor, and should be set based on ground-truth data collection.

These two measures are inversely related to each other, as the fusion system can be tuned to discard a lot of low confidence detections. This would result in a high true positive confidence and a low false positive behaviour rate, but increase the chance that an actual true object is missed. Therefore, it makes sense to take the convolution of these two measures as the overall measure for perception confidence C:

$$C_{perception} = C_{TP} \cdot C_{TN}$$

5.4. Estimating future automation performance

Decision capability can only be calculated if using intrinsic knowledge of the system design, and if those situations or events that the system is not designed to handle can be derived directly from sensor information or indirectly through vehicle to cloud (V2X) such as mapped data of traffic congestion or infrastructure.

Given a system which is not designed to handle close cut-ins from on-ramps, the proximity of an on-ramp including the statistics on how many close cut-ins happen there (based on for example traffic density) can be used to estimate the chance of a driver override at such a location. Here personal preference will play a role as to the degree at which drivers feel the system needs to be overridden or not.

The example of cut-ins from on-ramps also clarifies the need for a statistical approach to estimate the future automation performance. The automation performance estimation, in turn, relies on the prediction of the traffic situation (traffic context) as it will be perceived by the DAS in a near future, and such prediction will have some limitations. In the MEDIATOR project, the driving context prediction is based on the driving context as detailed in Chapter 4, which will provide a limited and probabilistic representation of that traffic situation. The driving context data may provide the next on-ramp location and the traffic density/speed to the automation state module. This information is however not sufficient to understand also the lane markings at the junction with the driven road, the curvature of the on-ramp, or the location and speed of the other road participants.

A very accurate descriptive representation of the driving context might however cause other issues. High complexity will make it difficult to predict if the DAS will be adversely affected. The estimation of driving context therefore needs to be extended carefully, not taking too much into account at the beginning.

Consequently, in Chapter 6 a generic statistical method is introduced which involves data collection to assess the automation fitness for various states of the driving context. This will allow the automation state module to estimate a probability of the DAS to be fit or unfit to drive in the near future.

6. Automation fitness

The objective of this chapter is to describe a methodology to measure and predict an estimation of the automation fitness used to derive the outputs of the automation state module such as the various *time to automation (un)fitness* for each of the *worst*, *likely* and *best* cases and for the various automation levels used in MEDIATOR (CM, SB and TtS).

The estimation of the automation fitness is similarly based on estimating an automation fitness score on an automation fitness scale:

- The AF scale is defined as a number of occurrences of system deactivations, fallbacks or overrides per time unit,
- The (continuous) measurement of the AF score is based on the perception confidence C and range R, which act like the amplitude of the waves on the seismographs (Wikipedia contributors, 2021). However, no mathematical formula is available at this time to use the perception indicators as input to obtain a fitness score. Instead, the automation fitness scale will be developed through later data collection and the correlation of the perception indicators with the occurrences of system deactivations/overrides/fallbacks,
- Much like the Richter scale, the AF scale will need to be unbounded, at least in the beginning, as we do not yet know what a “maximum” on the scale would be.
- The estimation of the automation fitness score is then used to predict the time to automation (un) fitness using cut-off thresholds to be determined later.

The methodology relies on the definition of BAD and degraded performance provided in chapter 2, the prediction of the driving context as detailed in chapter 4 and the monitoring of the automation perception confidence C and range R defined in chapter 5.

The methodology comprises two phases:

- Development phase: using collected data, the goal is to correlate both the perception indicators and the driving context with the number of occurrences of system deactivations/overrides per time unit normalized on an <u>automation fitness scale</u>,
- Implementation phase: using knowledge of these correlations, the goal is to estimate an <u>automation fitness score</u> on the scale using both online observations of the perception indicators and online observations / predictions of the driving context. The estimation of the automation fitness score is then used to predict the *time to automation (un)fitness* using cut-off thresholds.

The methodology can be applied to both supervised and unsupervised systems with some exceptions described in Section 6.3, as it relies on common foundations (driving context, perception indicators and system deactivations/overrides). Development of the module will only be done for supervised systems, as no unsupervised systems are available in Europe to date (April 2021). Instead, assumptions will need to be made for unsupervised systems.

6.1. Automation Fitness Scale

The automation fitness scale is a rating of how well the automation is doing in relation to a traffic situation. The rating can be represented e.g. from 1 to 5 with each level corresponding to a varying number of occurrences of system deactivations/overrides (or BAD performance as defined in chapter 2.1) per a certain time unit, as illustrated in Figure 6. The higher the value, the better the automation performance. It would also be possible to include the degraded performance or the factors correlating with BAD or degraded performance. The exact scale as detailed below is one example, and the actual development will need to be iterative and subject to change. The “maximum” performance as illustrated below will also need to be adapted to each automation system, so that any updates that improve performance can be accommodated by extending the scale.

5 Maximum performance
Rare occurrences of system deactivations/overrides/fallbacks per X hours
…
1 No performance
Frequent occurrences of system deactivations/overrides/fallbacks per X hours

Figure 6. Automation fitness scale representation

Defining the scale (rating values and the number of occurrences per category, as well as the factors associated with them) is done after the data collection during which the occurrences of system deactivations, fallbacks, and overrides will be annotated.

6.2. Estimating automation fitness score

After performing the data collection, a correlation analysis needs to be performed in order to estimate the automation fitness (AF) score on the scale. The correlation analysis will be performed independently for both the observed perception indicators and observed driving context with the occurrences of system deactivations/overrides/fallbacks for both types of inputs. Alternatively, the observed driving context could also be correlated with the perception indicators as an intermediate step of degraded automation or not to set a correlation with an AF score.

At this stage, the current AF score will be estimated by online observations of both perception indicators and driving context. However, the estimation using perception indicators, taken at the core of the system itself, is possibly more representative of the actual present automation performance than the estimation based on driving context. Nevertheless, predicting the future AF score can only be achieved using the driving context.

The estimation of the AF score in the automation state module therefore consists of two parts:

- Estimating the current AF score with the online observation of the perception indicators,
- Predicting the variation of AF score (relative to the current AF score estimation) with the online observation and predictions of the driving context.

6.3. Estimating time to automation (un)fitness

The *time to automation (un)fitness* can be predicted using the current AF score, its predicted variation due to the extent of degraded automation fitness and a cut-off threshold that sets the boundary between fitness and unfitness.

As an example in Figure 7 below, the current AF score is estimated at 4 (in yellow) using the real time observation of the perception indicators. The driving context ahead is identical to its real time observation except for two sections where the traffic density is expected to vary from no to moderate traffic density followed later on by a section of roadworks (black icons). Using the predicted driving context and its relationship to variations of AF score, the AF score can be estimated along the route (in green) with a decrease of AF score of -1 and -3 in the sections of moderate traffic and roadworks respectively. TTAU can then be estimated as the shortest time when the predicted AF score becomes lower than a predefined cut-off threshold (in red).

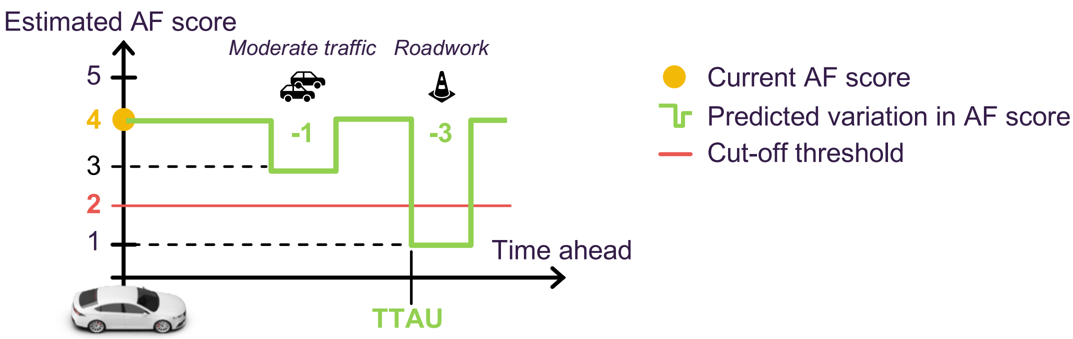

Figure 7. TTAU prediction using automation fitness score

There are a few exceptions to this method:

- **Worst case TTAU in supervised driving (CM)**: Given that any DAS up to and including SAE level 2 gives no guarantee on positive system performance (it will always rely on the driver to supervise the system) the worst case TTAU that can be expected in CM mode will always be 0 seconds,
- **All cases in unsupervised driving (SB and TtS)**: In a SAE level 4 ADS the system is by definition always fit to drive within its ODD, both for handling the DDT, the monitoring and the fallback. This means that the TTAU would always be infinite. Practically however, a driver will prefer to take over from the system before a fallback happens, for example at the end of the ODD. As previously explained, we will also consider SB as a subcase within TtS, where the driver is not asleep but where the ODD is coming to an end and fallback will be initiated. Therefore we will adapt the definition of TTAU in SB and TtS to reflect the time until fallback rather than time to automation unfitness.

In Figure 8 below, we illustrate the estimation of TTAU for the various scenarios (worst, likely and best cases) for the CM level. In this example, it is assumed that the automation state module receives the information that a dense traffic is predicted ahead for a short duration (represented with the three black vehicles) followed by a roadwork later on. The TTAU predictions are built as follows:

- **TTAU worst case** is always equal to zero second (independently of the current or predicted AF score estimation),
- **TTAU likely case** could consider the traffic density prediction “as is”, i.e. taking into account the dense traffic in the middle part,
- **TTAU best case** could consider a more favourable (or optimistic) traffic prediction, such as light or moderate traffic density, which would lead to a lower variation of the AF score.

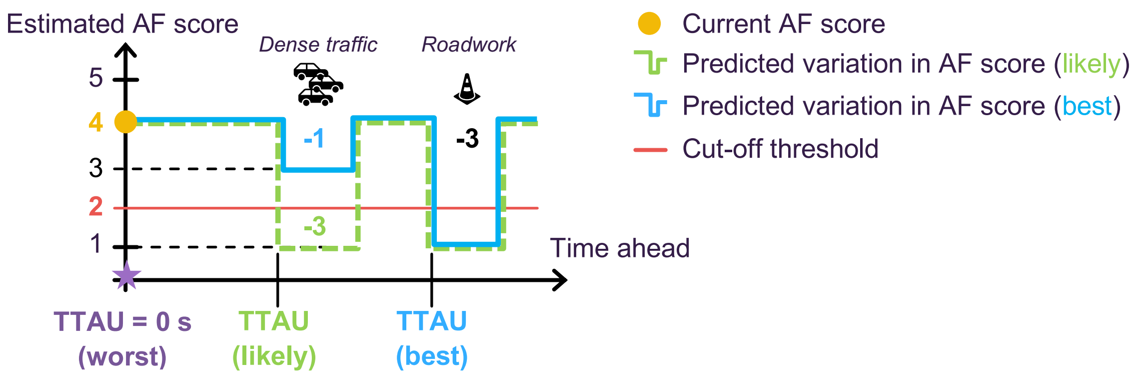

Figure 8. TTAU prediction for continuous mediation level (CM)

Table 7 summarises the TTAU estimations for the worst, likely and best case scenarios and for all three MEDIATOR automation levels.

Table 7. TTAU estimations for the worst, likely and best case scenarios.

TTAU estimation	Certainty level	Underlying assumptions	Supervised (CM)	Unsupervised (SB, TtS)
Worst case	High	Conservative	0 s [1]	Function of AF score
Likely case	Medium	Pragmatic	Function of AF score	
Best case	Low	Optimistic	Function of AF score	

[1] As stated above in the same section, TTAU estimations in the worst case for CM is always considered to be zero second as the driver is always in the loop.

6.4. Other MEDIATOR relevant outputs

Two other outputs can also be derived while doing the automation fitness score estimation:

- The *automation state class* as the reason for upcoming change in automation availability, like a change from fitness to unfitness and conversely. See MEDIATOR deliverables Christoph et al., 2019 and Cleij et al., 2020.
- The *appropriate intervention type* defined as a possible way to improve the automation fitness by increasing TTAU. See (Christoph et al., 2019).

Both outputs can be used by the HMI module to improve system transparency to the driver and possibly to require an action from the driver, in case the appropriate intervention type requires the driver rather than automation to act.

6.4.1. Automation state class

To compute the automation state class, it might be possible to identify which piece of the driving context information weighs the heaviest in a variation of the automation fitness score (using knowledge on the correlation built with the data collection). In the simple example in Figure 7, we assume that the driving context is composed of weather conditions and roadwork locations. As there is a roadwork location ahead and the weather is unchanged, the decrease in automation fitness score estimation from 4 to 1 is therefore linked to the roadwork in the system logic.

Cases might not always be so clear-cut though. It is likely that some roadworks will not constitute a problem for the DAS in good visibility, but at night, in poor visibility, or without a lead vehicle they might. The automation state class in this case might therefore need to be something more abstract such as the probability of “bad perception”, “bad lane markings” and “roadworks”.

In any more complex case where there has been no previous decision of what external context constitutes the issue for the DAS, the automation state class will probably become “end of ODD”. The automation state class outputs and timings must be developed further for improved messaging to the user, and in collaboration with the HMI component and decision making component of the Mediator system.

6.4.2. Appropriate intervention type

As the automation state module knows when (TTAU/F) and why (automation state class) there is an upcoming change in automation fitness, it can propose possible actions it thinks are suitable to take to improve the automation fitness.

As stated in Section 4.3.3 in (Christoph et al., 2019), the Mediator system can propose to activate or deactivate the DAS depending on whether the system is becoming fit or unfit to drive.

In the case of the system becoming unfit, other alternative actions could possibly increase the time during which the automation remains fit (i.e. TTAU). In the example above about the roadwork, these actions an others could be appropriate:

- “Reduce speed” to increase the time to reach the roadwork location,
- “Change lane” if the roadwork is known to only affect the current lane,
- “Take a different route” to avoid the roadwork, having checked beforehand that the alternate route leads to a prolonged fitness, comparatively.

6.4.3. Key Safety Indicators

The automation state module outputs time to automation fitness and unfitness, automation state class, current automation level and relevant context information. To assess the safety impact of the current automation fitness score, key safety indicators (KSI) will be used.

The KSIs basically constitute the same as the markers for degraded performance (like driving over a lane marking one should not), as there is a need to have an objective measure not occluded by hindsight. The markers for degraded performance as annotated by drivers and observers in the data collection can therefore be used as KSIs in conjunction with annotations of the actual driving context. For example, swerving slightly in lane is not unsafe if there is no traffic, but may be unsafe if there is heavy traffic and very narrow lanes.

6.5. Summary of the automation fitness score assessment

The estimation of the automation fitness score relies on various elements:

- The prediction of the driving context,
- The estimation of the current automation performance,
- The automation fitness scale,
- The correlation of the first two with the automation fitness score.

The capabilities and limitations of both the prediction of the driving context and the estimation of the current automation fitness score are detailed in their respective chapters (4 and 5). By transitivity, the correctness of the automation fitness score estimation is also affected by the accuracy of its inputs.

Additionally, the definition of the automation fitness scale as well as the correlation relationships are highly dependent on the significance of the data collection and so is also the estimation of the automation fitness score.

To build a very accurate estimation of the automation fitness score would require an extensive data collection to obtain a lot of exposure in diverse driving contexts when the driving automation system is affected by various factors. This is assuming a highly accurate prediction of the driving context and estimation of the current automation performance. For a reliable correlation for a large number of traffic situations and contexts, there is not only the need to be able to identify such traffic situations through sensing and sensor specifications. There is also a need to collect extensive data with multiple combinations of weather conditions, traffic conditions, road type, light conditions, roadworks, and so on. Such data collection would also contribute a number of occurrences of degraded and BAD automation performance to determine the automation fitness scale.

However, a data collection of such significance is beyond the possibilities of the MEDIATOR project. Nevertheless, to have a working automation state module for the MEDIATOR project, in particular for the vehicle prototype evaluation happening later in the project, we can focus the data collection on the traffic conditions in which the vehicle prototype will be evaluated.

6.5.1. Evaluation of the correctness of the outputs

The automation state module outputs time to automation fitness and unfitness, automation state class, current automation level and relevant context information. To assess the correctness of these outputs, key performance indicators (KPI) will be used.

For instance, time to automation (un)fitness could be described as a classification problem on whether the automation state module has correctly predicted a change in automation fitness versus the observation therefore involving standard metrics based on true/false positive/negative. Moreover, assuming that a change in automation fitness was correctly predicted by the automation state module at a certain location in the planned route, it could also be of interest to evaluate how good was the time prediction before reaching the location compared to the time it actually took. As described in Chapter 4.3, the estimated time to reach a certain location is dependent on multiple factors such as the traffic speed provided by an online source. Making such estimated time assessment therefore involves to also assess the accuracy of the various pieces of context information used to build the driving context.

Since the methodology to estimate the automation fitness detailed in this chapter is new, the KPIs will be developed in more details during/after the data collection part of the development of the automation state module. The analysis of the collected data to build the correlations between the automation internal states, driving context and the automation fitness (from which the outputs of the automation state module are derived) will provide insights that will drive the definition KPIs.

7. Concluding remarks

7.1. Main conclusions

This deliverable elaborates on the concepts and foundations for assessing the automation fitness of a driving automation system to be developed and implemented in the automation state module of the Mediator system.

The proposed framework relies on two main aspects. The first is the assessment of the current (real time) automation fitness which involves the creation of measures (indicators) based on information regarding the internal state of the driving automation system. The main focus lies on the sensors and decision-making software inside the automation module since automation fitness will be adversely affected if these fail or begin to degrade. The second is the assessment of the future (predicted) automation fitness, using information on the driving context in a near future, especially environmental conditions, road and route characteristics. These factors can affect the operation of various components of the driving automation system, such as sensors, which in turn will adversely affect the automation fitness. The deliverable also provides a review of the types of factors that can compromise the sensor function as well as known functional limitations of a range of automotive perception sensors.

The notion of automation fitness score is introduced, which allows the possibility to quantify the automation fitness on an automation fitness scale built on the rate of system deactivations/fallbacks/overrides. The deliverable proposes a methodology to create estimates of the current and predicted automation fitness score by examining the relationships between the automation self-assessment measures, the driving context and the automation system behaviour using collected and annotated data. These estimates then provide input to the automation state module to enable it to create the fitness score, to assess the future automation performance. This is achieved through the creation of a predictive metric derived from the fitness score, Time To Automation (Un)fitness, estimate the worst, likely and best case scenarios for various automation levels.

7.2. Open points to address during development

The deliverable focuses on the elaboration of the foundations for the automation state module. These concepts must therefore be addressed further during the actual automation state module development, in particular for the establishment of the correlation relationships between the inputs of the automation state module and the estimation of the automation fitness score.

The driving context as defined in Chapter 4, is a collection of various information sources mainly used in the prediction of the estimated automation fitness score for the worst, likely and best case scenarios. Understanding the accuracy and reliability of the sources that will be considered by the automation state module is critical to determine the confidence in this prediction and its derived outputs such as time to automation (un)fitness.

The estimation of the current (real time) automation fitness score is based on performance self-assessment indicators such as the perception range and confidence defined in Chapter 5. Further work will need to be done to assess the effectiveness of such indicators, make adjustments in their definition if needed as well as define other indicators if relevant to improve the assessment of the automation performance.

The definition of the automation fitness scale and the threshold used for characterising the automation as fit or unfit, as introduced in Chapter 6, will evolve during the development phase of the automation state module to adapt for instance to the size of collected data and to the statistical methods used to determine the correlation relationships. As stated in Section 6.5.1, key performance indicators will also be developed in more detail during the development phase of the automation state module.

7.3. Recommendations for the driver

The development of an effective mediator-type system needs to establish some recommendations to drivers whether they are the human supervisor or constitute the unsupervised automation system, so that the Mediator system can be as supportive as possible. These recommendations include but are not exclusive to:

- Drivers or the driving automation system should comply with the speed limit or lower,
- Drivers must input into the Mediator system where they are going and the route they plan to use.
- More recommendations may follow during the development phase.

7.4. Functional requirements for the automation state module

The functional requirements define the function of the system and its components**Fout! Verwijzingsbron niet gevonden.** summarize the functional requirements of the automation state about fitness to drive, and are based on the work described in this deliverable. These functional requirements (see Table 8) provide input to guide the further design and development of the automation state module.

Table 8. Functional requirements for the Mediator automation state module

Functional requirements (software related)
AUTOMATION STATE MODULE

The system shall estimate *worst*, *likely* and *best*-case *time to automation (un)fitness* based on automation fitness estimates for the current driving context
▪ The system shall estimate the *current* automation fitness
    ○ The system shall estimate the current automation fitness *score*
    ○ The system shall access relevant information from the automation system to estimate the current automation fitness score
▪ The system shall estimate the *predicted* automation fitness
    ○ The system shall estimate the predicted automation fitness *score*
    ○ The system shall access relevant external driving context information
▪ The system shall estimate when the driving automation system is *unfit to drive*
    ○ The driving automation system is deemed unfit to drive if it can no longer execute its defined dynamic driving task due to degraded automation performance (*low* automation fitness score)
▪ The system shall estimate the *time to automation unfitness* as the shortest time when the estimated automation fitness score becomes lower than a cut-off threshold
▪ The system shall estimate the *time to automation fitness* as the shortest time when the estimated automation fitness score becomes greater than a cut-off threshold

▪ The system shall estimate *worst*, *likely* and *best*-case scenarios of *time to automation (un)fitness* using the reliability of its inputs, both internal and external
The system shall determine the *active automation level* as either none, supervised (CM), or unsupervised (SB, TtS)
The system shall determine the *automation state class*, i.e. the reason for an upcoming change in automation availability
The system shall determine the *appropriate intervention type*, i.e. a possible way to improve the automation fitness
The system shall extract and collect *context relevant information* from the driving automation system to the context module
RECOMMENDATIONS
Drivers shall comply with the speed limit or lower
The automation state module shall know the planned route

References

Blanco, M., Atwood, J., Vasquez, H. M., Trimble, T. E., Fitchett, V. L., Radlbeck, J., & Morgan, J. F. (2015). *Human factors evaluation of level 2 and level 3 automated driving concepts.* (Report No. DOT HS 812 182). Washington, DC: National Highway Traffic Safety Administration.

Brummelen, J.V., O'Brien, M., Gruyer, D., & Najjaran, H. (2018). Autonomous vehicle perception: The technology of today and tomorrow. *Transportation Research Part C: Emerging Technologies*, 89, 384-406.

Christoph, M., Cleij, D., Ahlström, C., Bakker, B., Beggiato, M., Borowsky, A., … Van Nes, C.N (2019). Mediating between human driver and automation: state-of-the art and knowledge gaps, Deliverable D1.1 of the H2020 project MEDIATOR

Cleij, D., Bakker, B., Borowsky, A., Christoph, M., Fiorentino, A., van Grondelle, E., van Nes, N. (2020). Mediating System and Functional Requirements, Deliverable D1.4 of the H2020 project MEDIATOR

Combs, T. S., Sandt, L. S., Clamann, M. P., & McDonald, N. C. (2019). Automated vehicles and pedestrian safety: Exploring the promise and limits of pedestrian detection. *American Journal of Preventive Medicine*, 56 (1), 1-7. https://doi.org/10.3390/s20226532

ISO 20473:2007(E), *Optics and Photonics - Spectral Bands*, (International Organization for Standardization, 2007)

Mohammed AS, Amamou A, Ayevide FK, Kelouwani S, Agbossou K, Zioui N. The Perception System of Intelligent Ground Vehicles in All Weather Conditions: A Systematic Literature Review. *Sensors*. 2020; 20(22):6532. https://doi.org/10.3390/s20226532

Tchoya Florence Kone, Eric Bonjour, Eric Levrat, Frédérique Mayer, Stéphane Géronimi. Safety demonstration of autonomous vehicles: a review and future research questions. *Proceedings of 10th Complex Systems Design & Management, CSD&M 2019*, Dec 2019, Paris, France

Victor, T. W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., & Ljung Aust, M. (2018). Automation Expectation Mismatch: Incorrect Prediction Despite Eyes on Threat and Hands on Wheel. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, (August), https://doi.org/10.1177/0018720818788164

Wang, Z., Wu, Y, and Niu, Q. Multi-Sensor Fusion in Automated Driving: A Survey. *IEEE Access*, vol. 8, pp. 2847-2868, 2020. https://doi: 10.1109/ACCESS.2019.2962554

Wikipedia contributors. (2021, April 8). Richter magnitude scale. *Wikipedia, The Free Encyclopedia*. Retrieved from: https://en.wikipedia.org/w/index.php?title=Richter_magnitude_scale&oldid=1016634571 2021-04-16,09:25.