Grant Agreement N°: 952189

Topic: ICT-53-2020

5G BLUEPRINT

Next generation connectivity for enhanced, safe & efficient transport & logistics

D7.4: Evaluation of integrated technologies

Work package	WP7
Task	T7.4
Due date	31/12/2023
Submission date	22/12/2023
Deliverable lead	MIW
Version	2.0

Abstract

In the 5G-Blueprint project, we refer to teleoperation as a completely complementary technology to autonomous driving/sailing that can be used to provide human in the loop interventions and tackle edge scenarios that cannot be properly handled in the autonomous mode. However, teleoperation is a process that significantly relies on the quality of network connection, and thus, it reflects demanding connectivity requirements such as: i) uplink bandwidth for transferring video streams from cameras onboard to the teleoperation center, ii) low latency and ultra-reliable connection for relaying commands from the teleoperator to the remote vehicle/vessel, and iii) low interruption time when the teleoperated vehicle/vessel is crossing the border between two countries to ensure seamless connectivity and uninterrupted remote operation. Therefore, this deliverable provides i) a digest of extensive network performance evaluation presented in D5.4 deliverable, highlighting the main results related to network performance at different pilot locations, and ii) a comprehensive overview of final conclusions and analysis of pilot results for all use cases and enabling functions developed and deployed in the 5G-Blueprint project. The analysis of use cases and enabling functions consists of the service Key Performance Indicators (KPIs), which are described for each use case, and enabling function, along with expected target values, and discussion of the obtained measurements. Finally, the deliverable highlights the main lessons learned during the pilot activities.

Keywords: 5G SA, network slicing, teleoperation, seamless roaming, TOV, cross-border, safety, VRU, intelligent traffic lights, enhanced awareness, distributed perception

Document Revision History

Version	Date	Description of change	List of contributor(s)
V0.1	30/11/2023	All individual inputs from partners integrated and reviewed	Nina Slamnik-Kriještorac (IMEC), Rakshith Kusumakar (V-tron), Ghazaleh Kia (Seafar), Karel Kural and Bas Hetjes (HAN), Oliver Held (Roboauto), Steven Boerma (MAPTM), Linde Vande Velde (Be-Mobile), Ahmed Ahmed (IMEC), Sander Maas (Sentors), Tom van de Ven (Locatienet)
V0.2	8/12/2023	Comments included from Nina's review	Nina Slamnik-Kriještorac (IMEC)
V0.3	10/12/2023	Sections 1 and 2 written	Nina Slamnik-Kriještorac (IMEC)
V1.0	11/12/2023	Revised input from partners integrated, Missing input from Seafar	Nina Slamnik-Kriještorac (IMEC)
V1.0	12/12/2023	Lessons learned to be written	Nina Slamnik-Kriještorac (IMEC)
V1.1	19/12/2023	Comments from the reviewers included	Eric Oostendorp (KPN), Eva de Groote (Telenet), Wim Vandenberghe (MIW), Johann Marquez-Barja (IMEC)
V1.2	21/12/2023	All comments addressed and consolidated; Lessons learned written	Nina Slamnik-Kriještorac (IMEC), Steven Boerma (MAPTPM), Ghazaleh Kia (SF)
V1.3	21/12/2023	Section 3 added; Digest of	Nina Slamnik-Kriještorac (IMEC)

		network performance evaluation results made based on the results from D5.4	
V1.4	21/12/2023	New summary of results provided for Zelzate city area where EF3 is tested	Xhulio Limani (IMEC), Nina Slamnik-Kriještorac (IMEC), Steven Boerma (MAPTM)
V1.5	22/12/2023	The overall content harmonized and editorial changes made	Nina Slamnik-Kriještorac (IMEC)
V2.0	22/12/2023	Final version compiled	Nina Slamnik-Kriještorac (IMEC) Wim Vandenberghe (MIW)

Project co-funded by the European Commission under H2020-ICT-2018-20		
Nature of the deliverable:	R	
Dissemination Level		
PU	Public, fully open, e.g., web	√
CI	Classified, information as referred to in Commission Decision 2001/844/EC	
CO	Confidential to 5G-Blueprint project and Commission Services	

* R: Document, report (excluding the periodic and final reports)

DEM: Demonstrator, pilot, prototype, plan designs

DEC: Websites, patents filing, press & media actions, videos, etc.

OTHER: Software, technical diagram, etc

EXECUTIVE SUMMARY

The objective of this deliverable is to provide insights into final results obtained during comprehensive piloting activities performed during the final stage of the 5G-Blueprint project, including aspects related to 5G network performance, and their impact on performance of use cases and enabling functions.

Before deep-diving into performance evaluation results, we start with the overview of pilot sites and different testing locations within each of the three sites: Antwerp (BE), Vlissingen (NL), and Zelzate (BE-NL). This document provides a short description and the final status of these three sites and sub-sites that were incorporated to facilitate the piloting process.

To provide sufficient understand the 5G capabilities in the above mentioned pilot sites, we use a comprehensive network evaluation analysis provided in D5.4 as a reference, and summarize the main lessons learned when it comes to 5G performance in different pilot sites. From the results obtained in all three pilot sites, it is clear that the 5G SA network in the 3.5GHz range suffers from limited range, which offers good and stable signal quality but only up to 2km away from the gNB. This highlights the importance of proper placement of gNBs as good signal quality is essential for uplink throughput and end-to-end latency, required for latency-sensitive applications such as teleoperation. Also, the harsh environment in the busy port area is a significant impact factor for network performance. All presented results are promising as they show that both SA and NSA are able to support the teleoperation requirements (5Mbps uplink throughput per sensor/camera, below 30ms end-to-end latency for remote control commands, and below 150ms interruption time during handover process). Specific for the cross border site, service interruption time has been measured to evaluate how much time is needed for UE to continue using the previously established session in the home network when it attaches to the visiting one. The values obtained during testing show that optimized version of seamless handover brings significant improvements, and as both median and 95th percentile are significantly below 150ms, making service interruption time unnoticeable for cross-border teleoperation process.

After understanding how the network performs in each testing location, final conclusions and analysis of pilot results for all Use Cases (UCs) and Enabling Functions (EFs) developed and deployed in the 5G-Blueprint project are presented. In the case of 5G-Blueprint, we developed use cases such as Automated barge control (UC4.1), Autodocking of trucks and skid steer teleoperation (UC4.2), and Teleoperation-based platooning (UC4.3 & UC4.4), and several enabling functions (Enhanced awareness dashboard, Vulnerable Road User (VRU) Warning, intelligent Traffic Light Controllers (iTLCs), Distributed perception, Container ID recognition, and Estimated Time of Arrival Sharing) to test and validate 5G capabilities that could be leveraged large scale in future deployments. The relevant service KPIs are studied for all of these use cases and enabling functions, analyzing in particular the impact 5G network imposes on those service KPIs.

Finally, this WP7 deliverable summarizes valuable insights obtained during extensive piloting activities in real-life network settings. The summary covers all necessary technical elements in the 5G-enhanced teleoperation chain (network, teleoperation use cases, and enabling functions providing increased situational awareness), highlighting insights that will further pave the way towards achieving large-scale teleoperated transport based on uninterrupted in-country and cross-border 5G connectivity.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

5G NSA	5G Non-Standalone
5G SA	5G Standalone
AMF	Authentication Management Function
AuSF	Authentication Server Function
APF	Artificial Potential Function
CACC	Cooperative Adaptive Cruise Control
CAM	Cooperative Awareness Messages
CAN	Controller Area Network
CCU	Central Control Unit
C-ITS	Cooperative Intelligent Transportation System
Cloud App	Cloud Application
C-V2X	Cellular Vehicle-to-Everything
DBW	Drive-By-Wire
EAD	Enhanced Awareness Dashboard
ECU	Electric Control Unit
Edge App	Edge Application
EF	Enabling Function
eMBB	enhanced Mobile Broadband
ETA	Estimated Time of Arrival
GPS	Global Positioning System
HD	High Definition
HPLMN	Home PLMN
HR	Home-Routed
iTLC	intelligent Traffic Light Controller
IPKW	Industriepark Kleefse Waard
KPI	Key Performance Indicator
MAPem	MAP Extended Message
MEC	Multi-Access Edge Computing
MPC	Model Predictive Controller
MQTT	Message Queuing Telemetry Transport
MVP	Minimum Viable Platform
NR	New Radio
NRF	Network Repository Function
NSSF	Network Slice Selection Function
OBU	On-Board Unit
PC	Personal Computer

PCF	Policy Control Function
PLMN	Public Land Mobile Network
PPC	Pure Pursuit Controller
RAN	Radio Access Network
RSRP	Reference Signal Received Power
RTD	Round Trip Delay
RTK	Real-Time Kinematic
RTSP	Real-time Streaming Protocol
SMF	Session Management Function
SPaTM	Signal Phase and Timing Messages
SRM	Signal Request Messages
SRTI	Safety Related Traffic Information
SSM	Signal Status Messages
TCP	Transmission Control Protocol
TLEX	Traffic Light Exchange
ToV	Teleoperated Vehicle
UC	Use Case
UDAP	Urban Data Access Platform
UDM	Unified Data Management
UDP	User Datagram Protocol
UDR	Unified Data Registry
UE	User Equipment
uRLLC	ultra-Reliable Low-Latency Communication
V2V	Vehicle-to-Vehicle
VPLMN	Visited PLMN
VRU	Vulnerable Road User
WAN	Wide Area Network

1 INTRODUCTION

This deliverable is part of the documentation of the works on WP7, which is responsible for the integration of the effort developed in technical work packages (WP4-6), creating the end-to-end chain for 5G-enhanced teleoperation within and across country borders, and for managing the piloting activities. The goal of the document is to provide an overview of final results of all use cases and enabling functions, which are obtained during extensive testing in three pilot sites.

The final list of all Use Cases (UCs) and Enabling Functions (EFs) is presented in Table 1 and Table 2, respectively. These UCs and EFs are collocated within 5G-capable pilot sites, as presented in the latest overview of the overarching 5G-Blueprint architecture (D7.3) displayed in Figure 1, which combines the pieces of 5G Standalone network with seamless roaming mechanisms, and service/application components (use case and enabling functions). To clarify the terminology and 'collocation' of EFs and UCs, it is important to explain that EFs are not formally integrated with the UC chain, i.e., teleoperation. The collocation is therefore an intentional design choice that allows us to keep EFs and teleoperation chain as decoupled as possible, in order to increase the robustness of the system. This way, any possible propagation of a software fault or other similar issue from one domain to the other is prevented or at least minimized, which is essential for safety-critical applications such as teleoperation. As explained in D7.3, the EFs collocated with use cases provide valuable input for the remote driver in the form of enhanced awareness dashboard where detected obstacles, VRUs, signaling from the iTLCs, containers IDs to be loaded/unloaded, and relevant ETAs.

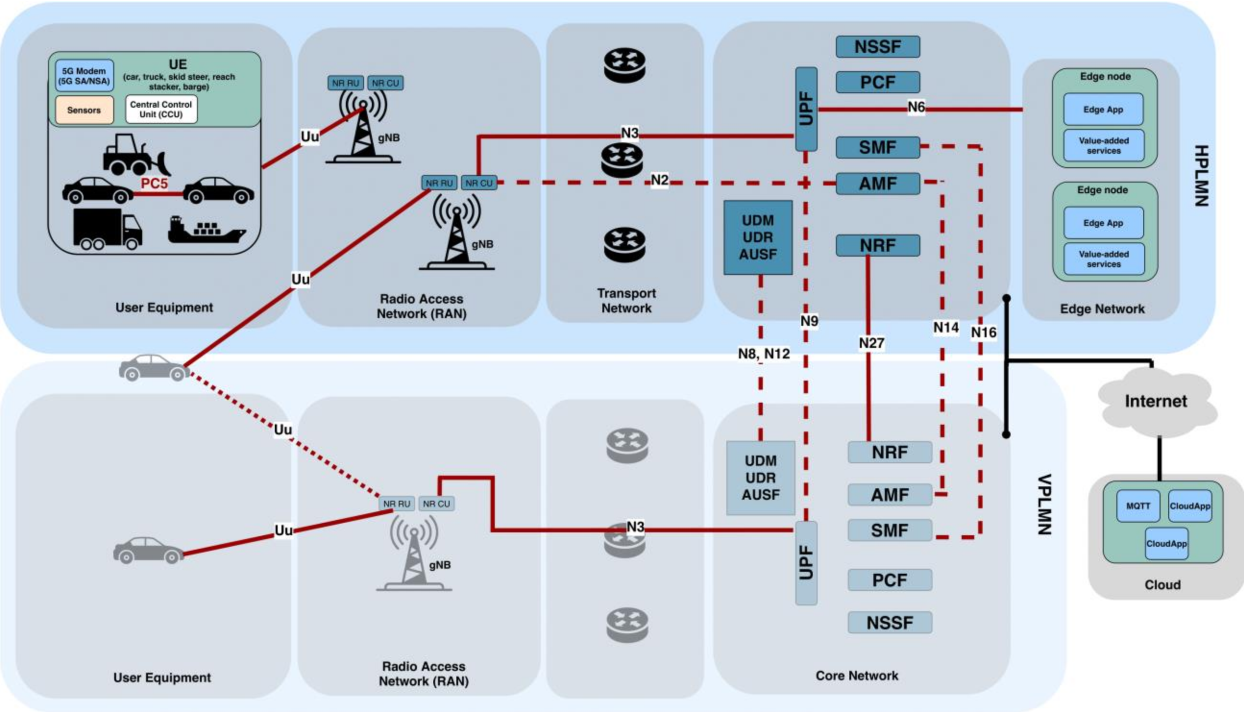


Figure 1 Final 5G-Blueprint architecture.

Table 1 List of Use Cases (UCs).

Use case ID	Full name
UC4.1	Automated barge control
UC4.2	Autodocking of full scale trucks and skid steers
UC4.3 & UC4.4	Teleoperation-based platooning

Table 2 List of Enabling Functions (EFs).

Enabling Function ID	Full name
EF1	Enhanced Awareness Dashboard
EF2	Vulnerable Road User (VRU) Warning
EF3	Intelligent Traffic Light Controller (iTLC)
EF4	Distributed Perception
EF5	Collision Avoidance System
EF6	Container ID Recognition
EF7	Estimated Time of Arrival (ETA) Sharing

In Section 2, this document provides a short description and the final status of three pilot sites, i.e., national sites: Antwerp (BE) and Vlissingen (NL), and the international or cross-border site: Zelzate (BE-NL). Based on the extensive 5G network evaluation analysis provided in D5.4, in Section 3 we create a 5G performance analysis digest, thereby summarizing the main lessons learned when it comes to 5G performance in different pilot sites. After understanding how the network performs in each testing location, Section 4 brings final conclusions and analysis of pilot results for all Use Cases (UCs) and Enabling Functions (EFs) developed and deployed in the 5G-Blueprint project. This section consists of the service Key Performance Indicators (KPIs), which are described for each UC, and EF, along with expected target values, and discussion of the obtained measurements. In Section 5, we conclude this deliverable and highlight the main lessons learned during the pilot activities. It is important to note that Annex of this document integrates the input on two EFs, i.e., EF1 and EF7, and network evaluation in Zelzate city center as part of the EF3 testing.

2 FINAL OVERVIEW OF THE PILOT SITES

This section provides a brief snapshot of the final pilot sites that have been extensively used in the piloting activities during the project, and especially in the third year. Offering both 5G Non Standalone (NSA) and Standalone (SA) capabilities at different sites, 5G-Blueprint enabled real-life testing and validation of teleoperation and autodocking aspects, along with a dynamic VRU and obstacle detection, intelligent traffic light controllers, and container ID recognition services, which altogether aim to optimize the transport & logistics operations in busy port environments.

As shown in Figure 2, and in detail reported in D7.2 [1], three pilot sites have been implemented to test and validate 5G impact on the teleoperation use cases and enabling functions, covering both national: Antwerp (Belgium) and Vlissingen (The Netherlands), and the international, i.e., cross-border site: Zelzate (Belgium-The Netherlands).

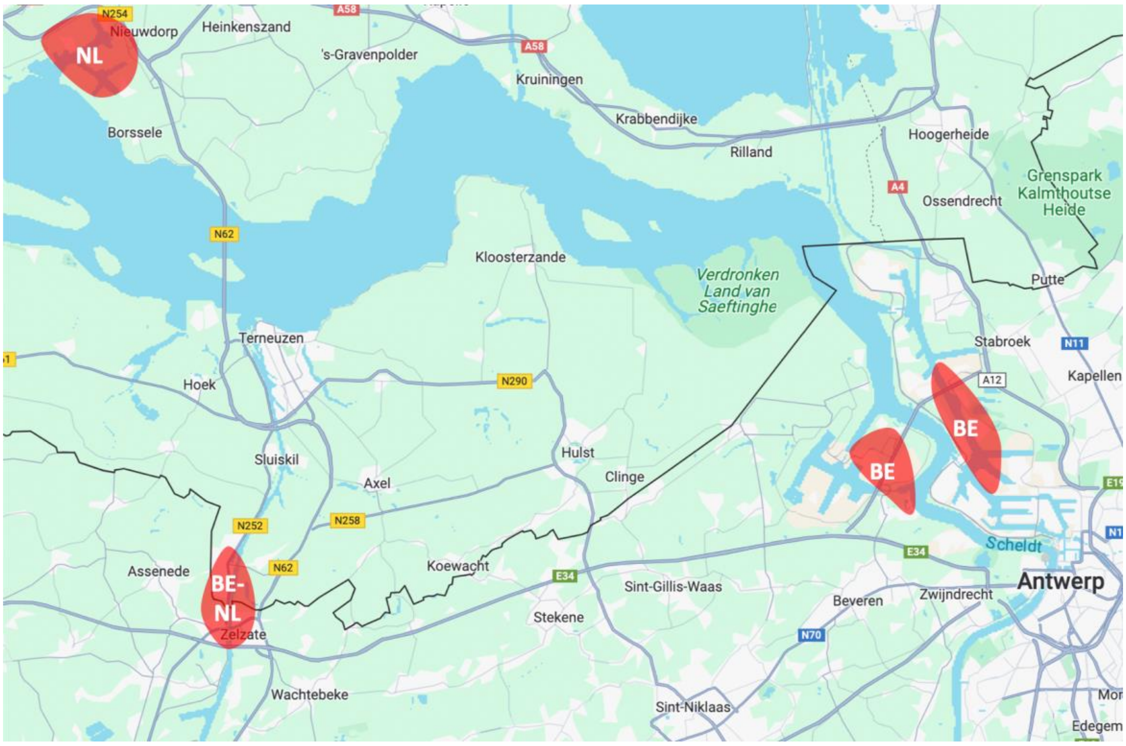


Figure 2 Geographical overview of three pilot sites.

Both Vlissingen and Antwerp pilot sites provide network coverage for both 5G Non-Standalone (NSA), and Standalone (SA) test network. Depending on the country and spectrum regulations, SA and NSA deployments exhibit different frequency ranges. While in Vlissingen 5G NSA is provided at 700 MHz (anchored 1800 MHz), Antwerp offers NSA connectivity 2.1 and 3.7GHz. In the case of SA, network is deployed at 3.7 GHz in both sites (center frequency with bandwidth of 100 and 50MHz, in Vlissingen and Antwerp, respectively), as well as in the cross-border site (2 gNodeBs deployed in close proximity from the geographical and administrative border between Belgium and the Netherlands).

To make the result and testing analysis in the following sections more comprehensible, we make a pilot overview in Table 3, mapping the pilot site and specific testing location covered by 5G SA and/or NSA within the site with the relevant use cases and enabling functions that leveraged test locations for their testing and validation. Concerning the **national site in the Netherlands**, **Vlissingen** offers three possible testing locations. For example, MSP Onions offers a unique 5G NSA-covered environment with a docking area (five docking stations) and parking spots, which was convenient for testing the combination of teleoperation and autodocking. To validate enhancements brought by 5G SA connection, teleoperation activities (of cars, trucks, skid steers) have been mostly performed in the Verbrugge Scaldia Terminal (second testing location in Vlissingen). Also, 5G SA connectivity has been used at three relevant locations for Container ID recognition (EF6) to validate the flexibility of the EF6 setup for dynamic container recognition while running the algorithm on the 5G edge. The driving rounds of teleoperation-based platooning over 5G NSA have been done on the public road, using the third testing location, i.e., public road from MSP Onions and the Kloosterboer terminal. On this road segment, intelligent traffic lights are deployed and thus the performance of the intelligent Traffic Light Controller (EF3) is tested at that location. Finally, Figure 3 includes also an additional testing site that is covered by KPN 5G NSA network, which has been deployed at a geographically suitable location for partners doing early testing in the Minimum Viable Platform (MVP) phase described in D7.2 [1].

Within the **national site in Belgium**, **Antwerp** offers the right bank site of the Scheldt river that has been used for periodic testing of *shadow-mode* teleoperated navigation of automated barge control (UC4.1) on both 5G SA and NSA. Here it is important to remind the reader about what shadow-mode testing means in the 5G-Blueprint project. Shadow mode testing is a term we have extensively used in the project when referring to a specific form of testing of direct control teleoperation. In this test form all subsystems of the teleoperation solution are active, meaning that the camera streams are sent to the remote operator station as normal, and the control signals created by the remote operator (steering wheel, pedals, joysticks) are sent to the teleoperated vehicle as normal. But the specific characteristic of shadow mode is that when these commands sent from the teleoperation center over 5G reach the UE in the vehicle, then their final translation to mechanical signals that induce steering, acceleration or braking, is disabled. As a result, the teleoperator who is giving commands as if he or she is in direct control of the vehicle, in reality is not in control. Instead, the safety driver in the vehicle is actually driving the vehicle. However, all created data streams over the network were identical to the situation where the teleoperator would have been in actual control. During this shadow mode testing, the KPIs can be measured, and the impact of network connectivity can be evaluated in real field operational conditions. From a networking perspective this gives much more realistic and valuable results then on a small closed test circuit. Shadow mode testing is usually used when there is no permit to perform direct remote control, such as in the case of public roads. When combined with insights captured on closed circuit testing regarding the relation between remote vehicle handling dynamics and network KPIs, it allows us to fully assess if teleoperation would have been possible on these public roads or not, in normal mixed traffic, without introducing any safety risks to the surrounding traffic or infrastructure.

The second Antwerp pilot location, i.e., the Transport Roosens Kallo site (hub for picking up and dropping off the containers) is on the left bank of the Scheldt river, and it offers a longer stretch of public road testing of teleoperation on both 5G NSA and SA. Apart from being used for use case testing, this location in Antwerp served as a suitable site for piloting and validation of Distributed perception (EF4) and VRU Warning (EF2) in an industrial setting. In addition, Antwerp pilot site spilled over to a new location with 5G SA coverage, i.e., Mechelen city center with one gNodeB deployed at the Telenet Headquarter (Figure 4). This location has been subsequently added for two purposes: i) to have an ad-hoc testing and debugging setup of 5G New Radio (NR), and ii) to create urban environment setting for VRU Warning testing (EF2).

Table 3 Mapping between pilot sites (test locations and network type) with the tested use cases and enabling functions.

Pilot site	Testing location	Network type	Use case/Enabling function	Network performance results	UC/EF results
Vlissingen	MSP Onions	5G NSA	Autodocking of trucks and skid steer teleoperation (UC4.2)	Section 3.2	Section 4.2
Vlissingen	Verbrugge Scaldia Terminal	5G SA and NSA	Teleoperation-based platooning (UC4.3 & UC4.4), Enhanced Awareness Dashboard (EF1), Container ID recognition (EF6), Estimated Time of Arrival (ETA) sharing (EF7)	Section 3.2	Section 4.1, Section 6.1, Section 4.7, Section 6.2
Vlissingen	Public road from	5G NSA	Teleoperation-based platooning	Section 3.2	Section 4.1,

	the MSP Onions to the Kloosterboer terminal		(UC4.3 & UC4.4), Enhanced Awareness Dashboard (EF1), Estimated Time of Arrival (ETA) sharing (EF7), Intelligent Traffic Light Controllers (EF3)		Section 6.1, Section 6.2, Section 4.5
Antwerp	Right bank of the Port of Antwerp-Bruges	5G SA and NSA	Automated barge control (UC4.1)	Section 3.1	Section 4.3
Antwerp	Transport Roosens Kallo site	5G SA and NSA	Teleoperation-based platooning (UC4.3 & UC4.4), Vulnerable Road User (VRU) warning (EF2): testing in industrial area, Distributed perception (EF4)	Section 3.1	Section 4.1, Section 4.4, Section 4.6
Antwerp	Extended Antwerp site in Mechelen city center	5G SA	Vulnerable Road User (VRU) warning (EF2): testing in urban area	Section 3.1	Section 4.4
Zelzate	Cross-border area around the canal Gent-Terneuzen (urban and industrial areas)	5G SA	Industrial area: Teleoperation-based platooning (UC4.3 & UC4.4), Automated barge control (UC4.1); Urban area: Intelligent Traffic Light Controllers (EF3)	Section 3.3 and Annex (Section 6.3)	Section 4.1, Section 4.3, Section 4.5

Figure 3 Pilot location in Vlissingen (red) and additional testing site covered by 5G NSA in Helmond (purple).

In terms of the network deployment, **Zelzate**, as the third pilot site in the 5G-Blueprint project, has gained most of the attention during the final year of the project, with the goal **to enable session and service continuity when crossing the border between Belgium and the Netherlands**. The final network setup is including one gNodeB installed at the Dutch side of the border (SA @ 3.5GHz, provided by KPN), and another one at the Belgian side (SA @ 3.5GHz, provided by Telenet). To enable advanced seamless roaming mechanisms, gNodeBs are connected to their respective 5G Core instance provided by TNO. The 5G-Blueprint roaming solution is described in more detail in D5.4 and D7.3 when explaining the final 5G-Blueprint network architecture. Briefly, by establishing interfaces between two TNO Core instances, in particular between Session Management Functions (SMFs) and User Plane Functions (UPFs), i.e., N16 and N9, respectively, the solution combines the Home-Routed (HR) roaming and the N2 handover over the N14 interface, i.e., between two Authentication Management Functions (AMFs). Significant reductions in downtime, i.e., interruption time induced by crossing the border and roaming from one operator's network to the other, were achieved due to a more efficient exchange of messages between peering 5G Core functions. The measurements obtained during network evaluation testing in Zelzate are reported in D5.4, and Section 3.3, and they display values lower than 150ms, which are sufficient for obtaining smooth teleoperation across country borders.

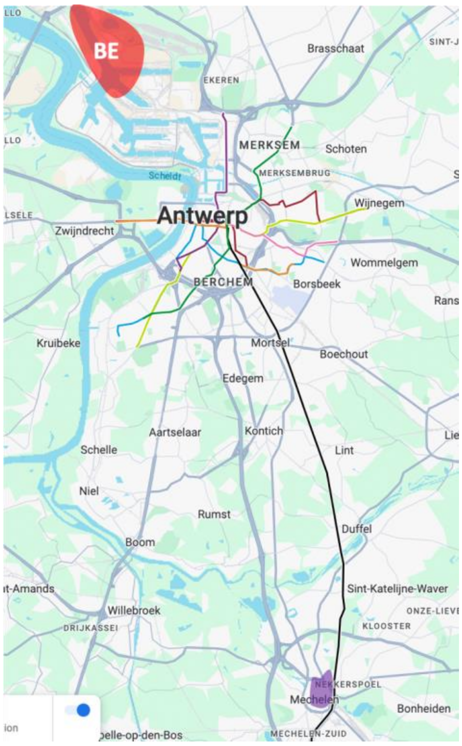


Figure 4 Pilot location in Antwerp (red) and additional testing site in Mechelen city center (purple).

At this cross-border pilot site, teleoperation-based platooning (UC4.3 & UC4.4) of vehicles has been extensively tested using the trajectory that stretches the country border and an industrial area. In parallel with the public road, there is a canal called Gent-Terneuzen, which is covered by the same network deployment. This canal contains a bridge on the border between two countries, which is an important obstacle in terms of connectivity due to which the piloting of automated barge control (UC4.1) needs to switch from an automated mode to teleoperation. Finally, in the urban part of this site, more tests have been performed in the Zelzate city center, thereby testing the capabilities of 5G SA for intelligent traffic lights or EF3 as well.

Table 3 provides dynamic references to sections that contain result analysis for all piloting activities mentioned in this section.

3 5G PERFORMANCE ANALYSIS DIGEST

To thoroughly evaluate the network performance, WP5 has conducted a series of testing sessions at different locations in Antwerp, Vlissingen, and Zelzate pilot sites, which are covered by 5G NSA and SA network. The performance results are in detail presented and discussed in D5.4, while in this section, we extract the main findings and highlight those results that refer to uplink throughput, end-to-end latency or Round Trip Time (RTT), and service interruption time, as those three metrics are considered the most critical for teleoperation processes. To remind the reader, the connectivity requirements for teleoperation use cases in 5G-Blueprint project (D5.1) are the following: i) uplink throughput of at least 30Mbps in case six cameras are simultaneously streaming High-Definition (HD) videos (or 5Mbps per camera/sensor), ii) ultra-low latency for remote control commands (RTT less than 35ms), and iii) service interruption time of less than 150ms. The results summarized in this section provide better understanding of 5G capabilities of different testing locations, which is a prerequisite for further understanding of results related to use cases and enabling functions (Section 4).

3.1 Antwerp pilot site

As reported in Section 2, the final 5G network outlook in Antwerp consists of two locations covered by 5G SA signal. Figure 5 shows the iterative deployment of 5G New Radio (NR) cells in the overall pilot site, where at the initial stages of the project one gNB was deployed for the network testing purposes. For the purpose of extending the coverage, six additional gNBs were installed. Due to the very dynamic environment in the port area, with massive metal constructions, container parks, trucks passing by on the roadside and large ships and vessels sailing on waterways, having proper understanding of 5G performance is essential. Such diversity in terms of obstacles creates challenging circumstances for propagation of 5G SA signal, which due to shadowing and fading phenomena can be blocked or reflected, thereby in turn impacting the quality of service experiences by teleoperated barges, cars, and trucks.

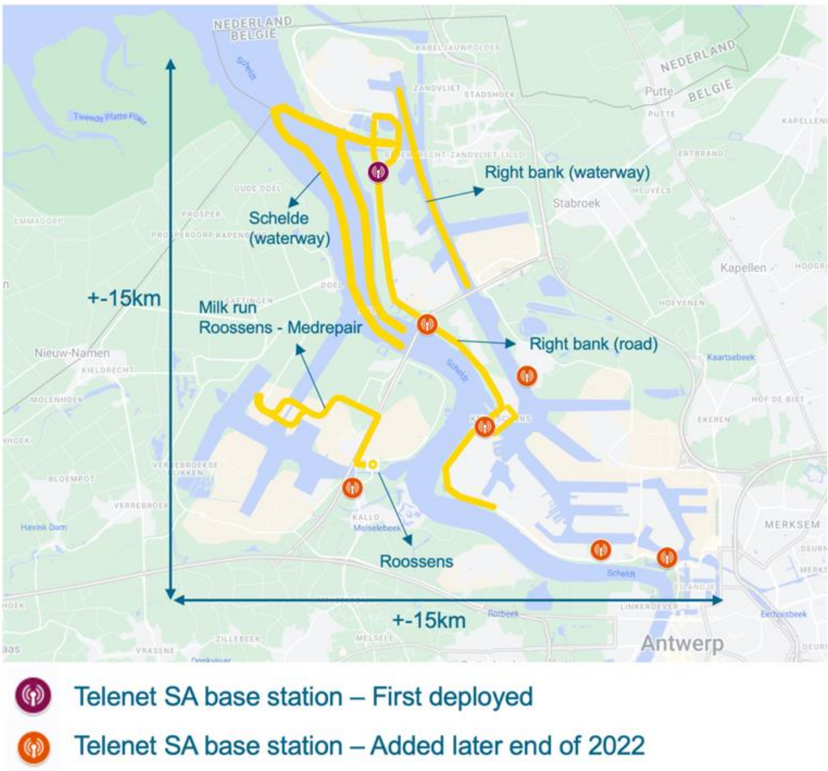


Figure 5 Iterative deployment of 5G NR in the Antwerp pilot site.

Although D5.4 provides a comprehensive analysis of metrics at various testing locations in

Antwerp, here we briefly summarize the main results related to the Right bank (test location for teleoperation of barges, obstacle detection of EF4, and VRU warning or EF2), and Transport Roosens site where teleoperation of cars happens (UC4.3 & UC4.4).

Table 4 Network performance at Antwerp pilot site.

KPI (Right bank/Roosens)	Min	Median	95th perc	Max	Average
RSRP (dBm)	-140/-140	-95/-97	-78/-76	-72/-60	-96.8/-95.3
TCP UL (Mbps)	0	4.94/4.76	29.9/28.8	36.5/45.9	9.14/10.2
UDP UL (Mbps)	0	11.5/5.33	29.5/30.9	32.1/36.7	12/11.6
RTT (ms) eMBB slice without background	13.4/17.4	27.4/19.3	36.7/35.7	99.3/11474	27.1/36.6
RTT (ms) eMBB slice with background	13.4/NA	29.4/NA	70.5/NA	666/NA	37.1/NA

Based on the results obtained for signal strength, i.e., RSRP, with the 95th percentile is of -78dBm at the Right bank, and -76dBm at Transport Roosens site, the majority of the testing trajectory is under good coverage. As expected, the minimum values of RSRP are obtained near the edges of the cells, where the signal naturally becomes significantly weaker.

Concerning round-trip time, which is reflecting how fast the control commands from the remote driver/skipper can reach the teleoperated vehicle/barge, or relevant messages from EFs to VRUs and teleoperated vehicles, the obtained value is 27.1ms on average (maximum one is 99.3 ms, measured at the edge of cell) in the case of Right bank, and 36.6ms at the Roosens site. The average value measured at Roosens is higher due to the big peak that occurs near the Medrepair site where the connection was lost, which is confirmed by relatively low median value of 19ms. Given the requirement of less than 35ms, it can be confirmed that remote commands and essential messages can be safely and reliably propagated over 5G SA network in the tested locations in case of no additional background load. When the background traffic is introduced, as done at the Right bank of Antwerp pilot site, the impact on the RTT values is noticed, resulting in peaks of up to 666ms. Nevertheless, even in these conditions, the average and median values remained within acceptable range.

In the case of uplink throughput measurements, similar observations are made at both Right bank and Roosens site. Although the average UDP uplink throughput is 12Mbps, the 95th percentile reaches 29.5Mbps, which is sufficient for . Similarly as in the case RTT, measurements obtained closer to the edges of the cell yield poorer network signal quality. However, as video streaming from cameras installed onboard is usually performed over Real-time Streaming Protocol (RTSP), which uses TCP as transport protocol, it is important to evaluate TCP uplink throughput as well. At the tested trajectory, the TCP uplink throughput significantly drops when the UE moves away from the gNB, resulting in mean value of only 9.14Mbps. Although the 95th percentile is close to 30Mbps, the trajectory for testing has been subsequently adjusted to avoid signal loss at cell edges.

From the extracted results presented and discussed above, it is clear that the 5G SA network in the 3.5GHz range suffers from limited range, which offers good signal quality up to 2km away from the gNB. Therefore, the placement of gNBs needs to be strategically planned as good signal quality is essential for uplink throughput and end-to-end latency, required for successful teleoperation of cars, trucks, and barges. Also, the harsh environment in the busy port area is a significant impact factor for network performance.

3.2 Vlissingen pilot site

Testing of network performance at the Vlissingen pilot site was performed at three different locations, i.e.,: Verbrugge terminal (relevant for trialing of teleoperation-based platooning and Container ID recognition), MSP Onions (relevant for testing autodocking capabilities), and public road between MSP Onions and Kloosterboer terminal in the Vlissingen pilot (relevant for trialing of teleoperation-based platooning in shadow mode, and intelligent traffic light controllers).

Figure 6 5G NSA coverage (left) and 5G SA testing area in Vlissingen (right).

Table 5 Network performance at Vlissingen pilot site.

KPI (Verbrugge/MSP Onions/public road)	Median	95th perc	Average
RSRP (dBm)	-89/-97/-96	-68/-96/-84	-86.6/-96.6/-95.3
TCP UL (Mbps)	(53.1, 37)/35.8/28.6	(61.8, 60.6)/38.5/57.8	(52.8, 38.3)/34.9/30.1
RTT (ms)	(14.4, 12.6)/23/24.7	(20.9, 17.8)/34.2/34.2	(15.0, 13.1)/24.1/29.6

Due to the very detailed performance analysis for the **Verbrugge terminal**, where SA coverage is present, here we extract only TCP uplink throughput, RTT, and RSRP, results. The values presented in Table 5 in pairs represent measurements for (eMBB, URLLC). Concerning throughput on the uplink, the mean value with no background traffic introduced is 52.8Mbps, which drops to 23.3Mbps when there is impact of the background traffic. This impact is significant, which can be seen due to 95th percentile being 62Mbps when there is no impact and 30.3Mbps when background traffic is introduced. As the impact of background traffic affects end-to-end latency as well (~80% increase in 95% cases), it is important to carefully dimension cells, and define gNB placement, with denser deployments at busy areas.

At **MSP Onions**, the network evaluation tests reported in detail in D5.4, were conducted at a stationary location near the truck docking gates (where the trialling of autodocking use case takes place). As this location is covered only with 5G NSA signal, there is no slicing involved. The TCP uplink throughput is 34.9Mbps on average, with 95th percentile of 38.5Mbps. The RTT measurements indicate the average value of 24.1ms, while the highest values stretch to 63.8ms which are considered outliers due to scheduling of resource blocks.

The third location, i.e., the stretch of the **public road** between the MSP Onions and the Kloosterboer terminal is presented in Table 5 as well. Similarly as in case of Antwerp pilot site, values obtained near the cell edges are lower in case of uplink throughput and end-to-end latency. The maximum value reached 69.5Mbps, signifying the important of good cell coverage for obtaining sufficient uplink throughput values, which are essential for safe teleoperation. The end-to-end latency is 29.6ms on average, with the 95th percentile of 34.2ms, which is slightly exceeding the expected values for the remote control of trucks at MSP Onions.

In overall, the 5G SA performance evaluation at Vlissingen pilot site shows that URLLC slice is more stable than eMBB when background traffic is introduced. This highlights the importance of proper network design and slice choice when it comes to latency-sensitive applications as teleoperation, where URLLC should be always used for transferring mission-critical messages and control commands. All presented results show that both SA and NSA networks at tested locations are able to support the teleoperation requirements. Similarly as in case of Antwerp pilot site, providing good coverage is crucial to meet the connectivity requirements, as moving towards cell edges drastically degrades signal quality, thus achieving lower throughput values.

3.3 Zelzate pilot site

The cross-border site consists of two gNBs that are covering both the waterways (Gent-Terneuzen canal) and the stretch of the public road in parallel with the canal. The locations of gNBs are shown in Figure 7. The driving tests with IMEC test vehicle and sailing tests with North Sea Port boat have been performed in the selected area between two base stations (Figure 7). It is important to note that more tests have been done on the roadways due to the availability of the North Sea Port boat only for one-day testing.

Figure 7 Cross-border area and locations of gNBs; Old Telenet gNB is located in the City Center area where EF3 has been tested (results for this location presented in Annex).

After resolving initial issues on the signal quality caused by interference in the cross-border area (illegal transmitter was identified and shut down) and installation of an active mobile antenna on the Telenet side, significant improvements on the uplink throughput are obtained. Given the presence of active antenna on the Belgian side of the border, the TCP uplink throughput values are higher on the Telenet than on the KPN network. Table 6 shows the statistical analysis of throughput values on the entire cross-border trajectory, with median value of 24.3Mbps and 95th percentile of 51.5Mbps. These values are high as the UE has been connected to Telenet network longer than on the KPN. The range of obtained values for both uplink throughput and end-to-end latency is acceptable for achieving safe teleoperation across country border, as per requirements mentioned in Section 1.

Table 6 Network performance at Zelzate pilot site.

KPI	5^{th} percentile	95^{th} percentile	Average
RSRP (dBm)	-103	-69	-86.3
TCP UL (Mbps) on roadways	1.42	51.5	24.1
RTT (ms) on roadways	12.7	40.1	24.4
TCP UL (Mbps) on waterways	NA	59.25	47.35
RTT (ms) on waterways	NA	36.227	24.38
Service interruption time (ms) – optimized	96.4	109.0	108.3
Service interruption time (ms) - unoptimized	124.2	150.95	119.615

Another important metric that is specific for the cross-border setting is the service interruption time, or network downtime. The metric is defined as the time between the last packet the UE could send while being connected to Home Public Land Mobile Network (HPLMN), and the first packet it sends over the Visited PLMN (VPLMN) when crossed the border. The values presented in Table 6 show that optimized version of seamless handover yields lower values of interruption time, and as both median and 95^{th} percentile are significantly below 150ms, service interruption time is unnoticeable for cross-border teleoperation process. This optimization is obtained by preparing the PDU session as much as possible beforehand to further reduce the downtime (fewer messages are exchanged between 5G Core functions in two networks). Finally, to further reduce the end-to-end latency (between UE and application servers), it would be important to study aspects of edge deployments and deploying teleoperation services and EFs at the network edge to avoid home-routing roaming, which is out of scope of the 5G-Blueprint project.

4 PILOT RESULTS: USE CASES AND ENABLING FUNCTIONS

This section is focused on the so-called service Key Performance Indicators (KPIs), which are measured to evaluate the performance of use cases and enabling functions. Therefore, each of the subsequent sections is providing overview of testing methodology, relevant KPIs, and analysis of the results at specific pilot locations. While Sections

4.1 Teleoperation-based platooning

The service KPIs measured for teleoperation-based platooning are grouped into two categories: i) Cooperative Adaptive Cruise Control (CACC)-based platooning (UC4.3) KPIs such as distance between the lead and following vehicle, distance error, and maximum speed, and ii) teleoperation (UC4.4) KPIs such as steering accuracy, pedals accuracy, and maximum speed. The extensive overview of the final blueprint , and performance results of these two use cases is presented in D4.3. In this deliverable, we provide a summary of the achieved results and the main lessons learned.

4.1.1 Cooperative Adaptive Cruise Control (CACC)-based platooning

Preconditions

The two Toyota test vehicles (Figure 8) equipped with Cellular Vehicle-to-Everything (C-V2X) boxes, facilitating communication between the lead, and following vehicles. Additionally, the following vehicle is equipped with the CACC system, with the necessary hardware and software. A safety driver is always present in the vehicle to take control during edge cases. The tests were conducted at a speed range of 40-60 kmph on isolated roads with no traffic and under acceptable weather conditions.

Figure 8 Pilot vehicle with installed cameras.

Procedure

During the testing, the lead vehicle is initially operated by a driver/teleoperator at low speeds. Upon reaching a stable speed, the CACC system is activated. The following vehicle, equipped with the CACC system on-board, follows the lead vehicle with a desired headway time. The CACC system is evaluated under three primary conditions,

- Gap closing: The acceleration of the lead vehicle will be increased gradually, and the behavior of the following vehicle will be monitored. The following vehicle is expected to close the gap created because of the acceleration.

- Following: The lead vehicle will be driven at a constant speed (zero acceleration), and the behavior of the following vehicle will be monitored. The following vehicle is expected to follow the current following distance (without large variations)
- Collision avoidance: The lead vehicle will be decelerated (to a complete stop), and the behavior of the following vehicle will be monitored. The following vehicle is expected to react and decelerate instantly and avoid a collision.

4.1.1.1 KPIs and Results

The overview of measured KPIs is shown in Table 7, while the measured values along with the statistical analysis are provided in Table 8 and Table 9, respectively. It is important to note that these measurements are obtained during the testing on the Vlissingen pilot site.

Table 7 CACC KPI measurement methodology.

KPI	Measurement Methodology
Following distance (Headway time)	Measured with the vehicle vision sensor in [s]
Distance error	Calculated based on the logged data. Distance is measured in [m]
Vehicle-to-Vehicle (V2V) communication latency (PC5 communication)	Calculated from the time stamp data measured in [ms]
Packet loss	Calculated based on the total number of packets sent and received
Maximum safe speed achieved	Measured from the vehicle Controller Area Network (CAN) bus

Table 8 CACC KPIs measured in Vlissingen site.

#	KPI	Definition	Target values	Measurement
1	Following distance (Headway time)	The minimum achievable headway to the lead vehicle	1 [s]	0.8 [s]
2	Distance error	Difference between actual and desired distance	Less than 5% (in steady state condition)	2 - 4 % (Mean error – 0.25 m)
3	Latency - V2V communication	Delay communicating the message from lead	20 [ms]	18 [ms]

		vehicle		(Average)
4	Packet loss	The number of packets lost in the V2V communication	Less than 5% (within 100 m distance)	2 %
5	Maximum test speed	Maximum achievable speed with CACC activated	80 [kmph]	60 [kmph] (Limited for testing purpose)

Table 9 CACC KPI Statistical values.

KPI	Measured values	Std. Deviation	95th Percentile
Following distance (Headway time)	0.8 [s]	0.3981	0.9009
Distance error	2-4% (Mean 0.25m)	0.5176	1.263
V2V communication latency	18 [ms] (average)	2.529	22
Packet loss	2 %	-	-
Maximum safe speed achieved	60 [Kmph]	-	-

Table 10 CACC KPI updates compared to MVP phase (D7.2 [1]).

KPI	Status	Reasoning
Packet loss	Added	The packet loss is a key factor for defining the robustness of the communication. This is crucial for such a system as large packet losses would lead to disengagement of CACC system.
Maximum acceleration / deceleration	Removed	From the previous results it was noted that the controller and the vehicle always stay within the maximum acceleration/deceleration rate, also due to the physical limitations of vehicle actuation. It is therefore decided as unnecessary for KPI measurement.
Number of human interventions	Removed	From the tests conducted, it was noted that there is no need for human interventions during the closed environment testing scenario. This is more applicable if the tests were performed in real dynamic traffic.

4.1.1.2 Discussion

CACC based platooning tests were only performed in the Vlissingen pilot site. Since this system is not dependent on the long-range network, the test location selection had no effects.

The tests were performed in daytime with clear weather and minimal traffic. Safety drivers were present in both the following and lead vehicles to take manual control whenever necessary. The point of interest was to monitor the following vehicle's behavior during acceleration, steady speed following and deceleration. The results obtained were consistent with previous test results, obtained during the MVP phase reported in D7.2 [1], which validates the robustness of the overall system.

Headway time:

During the maneuver, the following vehicle is set to follow the lead vehicle with a 0.8 second headway time. From Figure 9, it can be seen that the following vehicle's velocity matches closely with the lead vehicle's velocity. In at least 95% cases, the headway time is less than 0.9 second, and the controller is able to control the following vehicle with the set headway time with minimal distance error throughout the maneuver and was able to bring the vehicle to a complete stop at the end of the maneuver.

Distance error:

The distance error was close to zero during the steady state driving and an overall mean error was within 2 – 4 %. Analyzing the standard deviation and 95th percentile values, the results are close to the target values, thereby validating the overall performance of the CACC system. The distance error had a standard deviation of 0.5 which proves that the error deviation is very small and 95th percentile of 1.26. The CACC is a dynamic system, and the error values are expected to be higher during the transition phase when the speed is changing and stabilize during the steady state driving. This explains the comparatively bigger 95th percentile value.

Maximum speed achieved:

During the test, the lead vehicle was driven by the safety driver and the maximum speed was limited to 60 kmph (speed limit of the test route). The speed profile, as seen in Figure 9, was selected to validate the point of interest. The test starts at around 45 kmph and the lead vehicle accelerates until it reaches 60 kmph. After driving at a constant speed, the lead vehicle decelerates back to 45 kmph before it comes to a complete stop. Note that the system that has been tested for higher speeds until 100 kmph provides similar results.

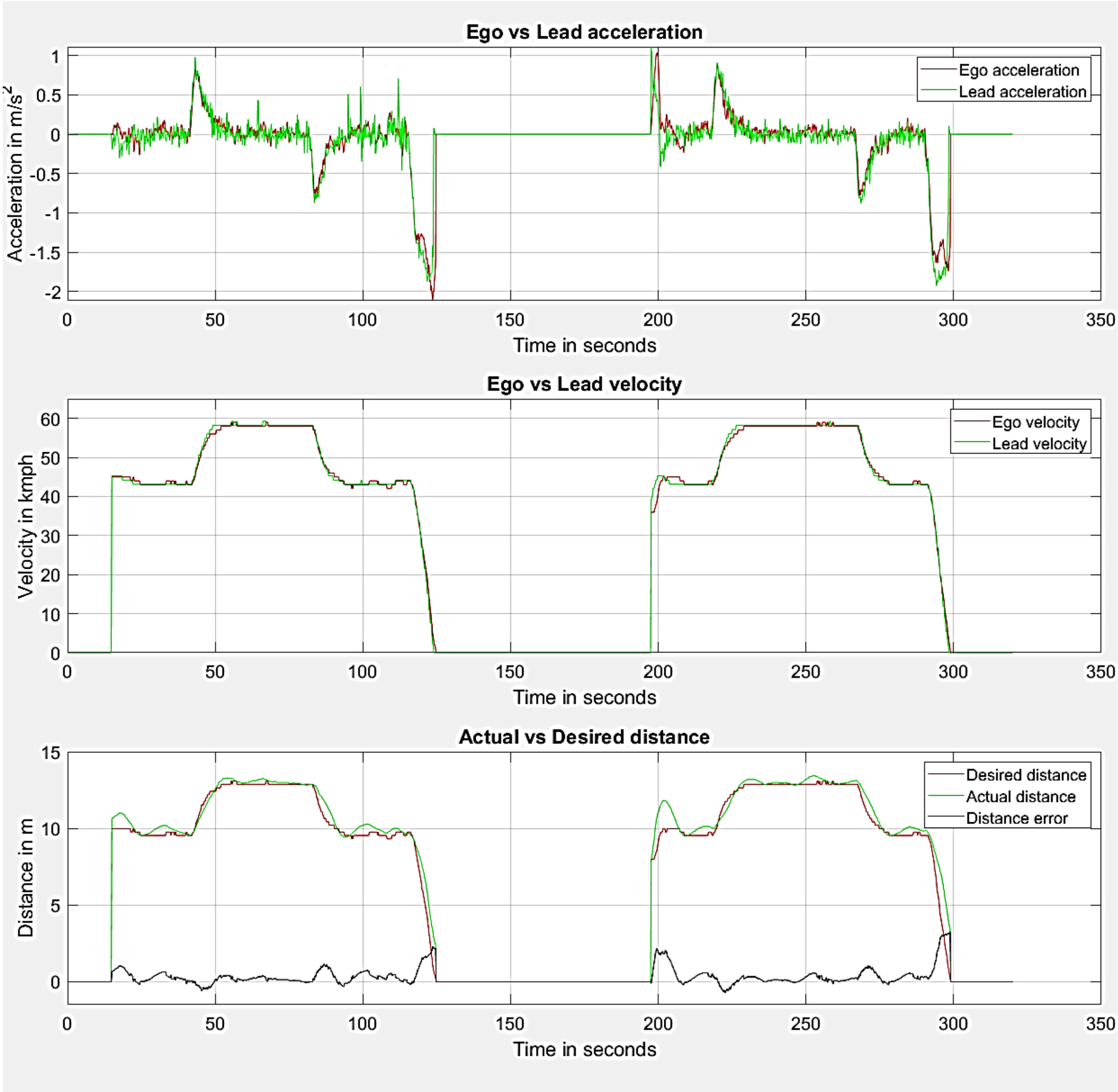


Figure 9 Test results of CCAC-based platooning in Vlissingen.

Communication latency and packet loss:

The Vehicle-to-Vehicle (V2V) performance is shown in Figure 10. It can be seen that the latency values are around 18 ms which is consistent with the defined KPI. The standard deviation and the 95th percentile values are analyzed and are close to the target values. The latency has a standard deviation of 2.53 and the 95th percentile was 22 which is close to the mean value, validating the consistent performance of the communication.

The system logs the total number of messages sent and received to calculate packet losses. The overall packet loss was close to 2%, which is acceptable for the system. Throughout the tests, the CACC system experienced no deactivation caused by communication delays or losses. This confirms that the measured values are within the threshold.

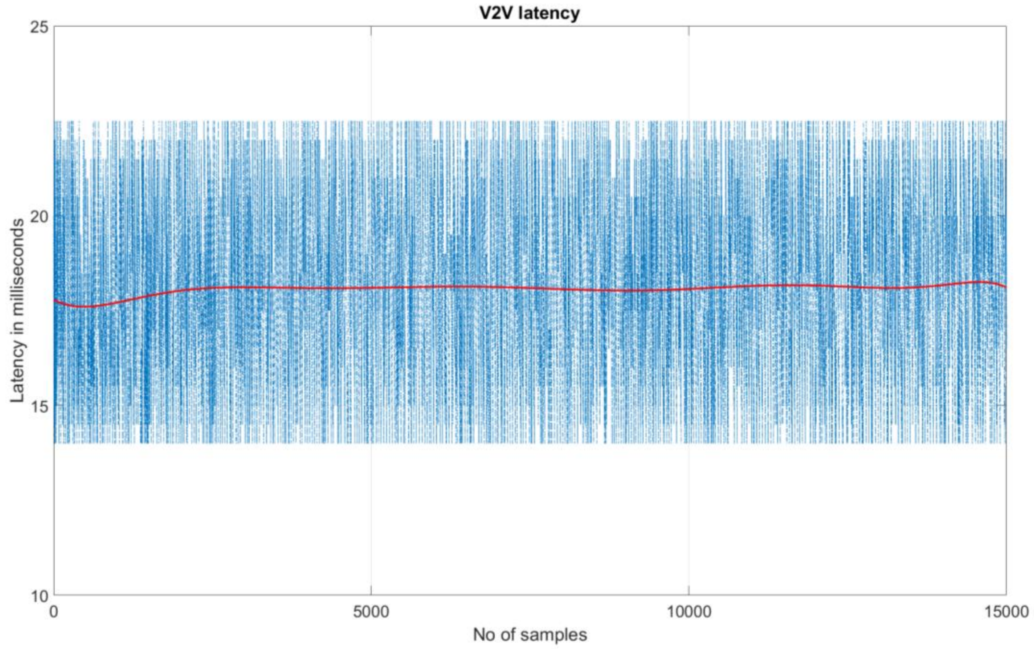


Figure 10 V2V latency plot for CACC-based platooning in Vlissingen.

4.1.2 Teleoperation

Preconditions

The test plan focuses on evaluating the performance of teleoperation and remote takeover functionalities, alongside their supporting functions. To ensure a comprehensive assessment, several preconditions must be met, including network setup on both the remote station and vehicle, Safety systems tests, including brake, throttle, and steering responsiveness tests and the overall teleoperation functionality tests. The testing environment involves a Toyota test vehicle equipped with teleoperation hardware/software, communicating with the remote station through the 5G network. Tests are performed on isolated roads with minimal traffic, and safety drivers are always present in the vehicle, to take manual control if needed.

Procedure

The comprehensive teleoperation and remote takeover test plan involves a series of critical assessments:

- Safety Systems Test: Evaluates brake, throttle, and steering responses in varied scenarios, both remotely and from within the vehicle, including manual steering override validation.
- Steering, Brake, and Throttle Responsiveness Tests: Assess the synchronization of vehicle responses to teleoperator inputs, focusing on minimal delay, stable transitions, and natural behavior.
- Driving Accuracy Test: Examines delays by comparing incoming messages and physical actuations, with a small delay indicating accurate actuator performance.
- Slow Speed and Regular Speed Maneuvering Tests: Simulate parking and everyday driving scenarios, respectively, validating actuators, tuning, and network stability.

4.1.2.1 KPIs and Results

The teleoperation KPIs listed in Table 11 are measured in both Vlissingen and Antwerp pilot sites (Table 12 and Table 13).

Table 11 Teleoperation KPIs measurement methodology.

KPI	Measurement Methodology
Steering accuracy	The steering wheel rotation is measured in degrees [º]
Pedals accuracy	The pedals mapped to a percentage [0-100%]
Maximum safe speed	Maximum possible speed for safe teleoperation in [kmph]

Table 12 Teleoperation KPI measured in Vlissingen & Antwerp site.

KPI	Definition	Target values	Measurement	
			Vlissingen	Antwerp
Steering accuracy	The input given through the driving station should be the same on the teleoperated vehicle.	Mean error < 0.1 [º] Mean Absolute Error (MAE) <3.0 [º] Root Mean Squared Error (RMSE) < 5.0 [º]	Mean error = 0.11 [º] MAE = 2.41 [º] RMSE = 3.85 [º]	Mean error = 0.077 [º] MAE = 4.56 [º] RMSE = 6.29 [º]
Brake / Throttle Pedals accuracy	The input given through the driving station should be the same on the teleoperated vehicle.	Mean error <1.0 [%] Mean Absolute Error (MAE) <4.0 [%] Root Mean Squared Error (RMSE) < 6.0 [%]	Mean error = 0.33 [%] / 0.88 [%] MAE = 0.51 [%] / 1.27 [%] RMSE = 1.08 [%], 2.09 [%]	Mean error = 0.32 [%] MAE = 0.702 [%] RMSE = 1.22 [%]
Maximum	Maximum possible speed for safe teleoperation.	25 [kmph]	Limited to 15kmph	>30kmph

safe speed				

Table 13 Teleoperation KPI statistical values.

KPI	Vlissingen		Antwerp	
	Std. Deviation	95th Percentile	Std. Deviation	90th Percentile
Steering accuracy	3.85 [º]	5.75	6.29 [º]	6.877
Brake/ Throttle pedal accuracy	1.04 / 1.91 [º]	2.41 / 1.17	1.17[º]	2.041
Maximum safe speed	-	-	-	-

Table 14 Teleoperation KPI updates.

KPI	Status	Reasoning
Network related KPIs for teleoperation	Removed	The network related KPIs are more significant for WP5, therefore included in the KPIs from WP5
Teleoperation Overridability	Removed	From different tests it was clear that this KPI was not characterizing any aspect of the system.

4.1.2.2 Discussion

The graphs in the Figure 11 show the comparison between the behavior of the vehicle as requested by the remote operator and the one accomplished by the vehicle. The results are obtained during testing in the Antwerp pilot site. This was done to show relevant data regarding the reactivity of the system with respect to steering angle and pedal position. The Vlissingen testing also provided comparable results. It can be seen how the graphs in Figure 12 are very closely related with minimal errors.

Steering accuracy

When analyzing the steering angle graphs, it is evident that there is a close correlation between them, with only minor errors. Specifically, the comparison between the requested steering angle and the vehicle's steering angle shows a Mean Absolute Error (MAE) of 4.56 degrees. While this value slightly exceeds the desired value, it is still within an acceptable range, given that the 90th percentile for steering is 6.877 degrees. Moreover, the sample size of 3001 further reinforces the validity of the results. Additionally, the standard deviation of 6.29 degrees highlights the system's ability to maintain a consistent level of accuracy, with most of the data points falling within an acceptable range of deviation from the requested steering angle. The weather conditions have very minimal impact on the results. It is important to note that there is an inherent error factor in the system, which cannot be eliminated due to physical actuation limitations. Nonetheless, the system's overall performance remains acceptable.

Pedals accuracy

Regarding the throttle pedal results, they are within expectation. This can be attributed to a standard deviation of 1.17% and a 90th percentile of 2.041. This is due to the fact that there is no physical actuation of the pedal, resulting in faster response. It is to be noted that, for Antwerp tests, only the throttle pedal has been shown, this is due to the fact that the brake pedal has a different unit output from the vehicle, which does not allow for direct comparison. This was rectified and presented in the Vlissingen tests, shown in Figure 12.

The MAE for brake pedal positions were 0.51 indicating that it was within the acceptable range. Although the brake pedal results were better, the throttle accuracy KPIs were still satisfactory. Comparing the results, the brake pedal showed greater similarity between teleoperation and the car.

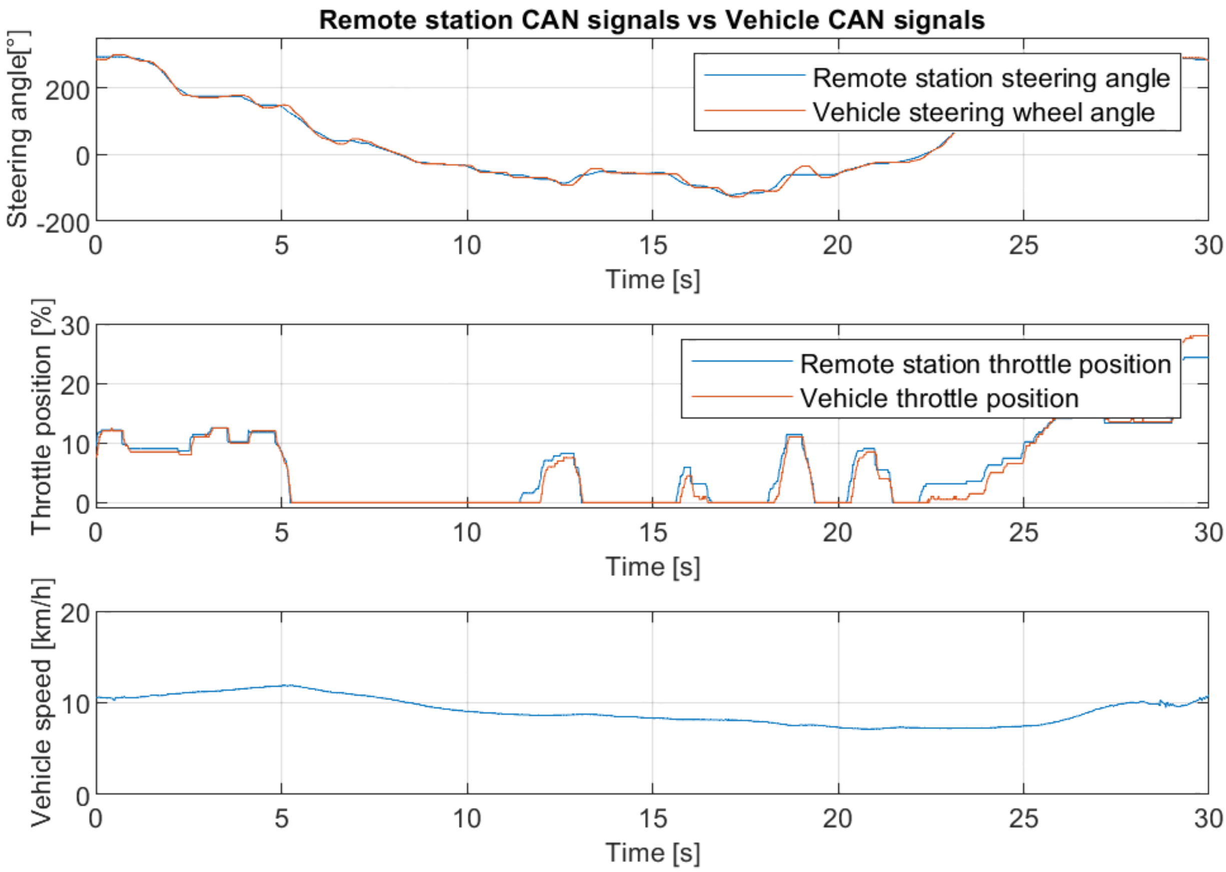


Figure 11 Teleoperation KPIs - Antwerp results.

Maximum safe speed

During the period of 30 seconds, the speed of the vehicle varies between 8 and 13 kmph. This limitation was in place due to the fact that the testing perimeter was not sufficiently large to drive at faster speeds during these maneuvers. The maximum safe speed indicator is based on the experience of the remote operator. It was shown that driving at 30 kmph still gave the remote operator a feeling of comfortability while driving.

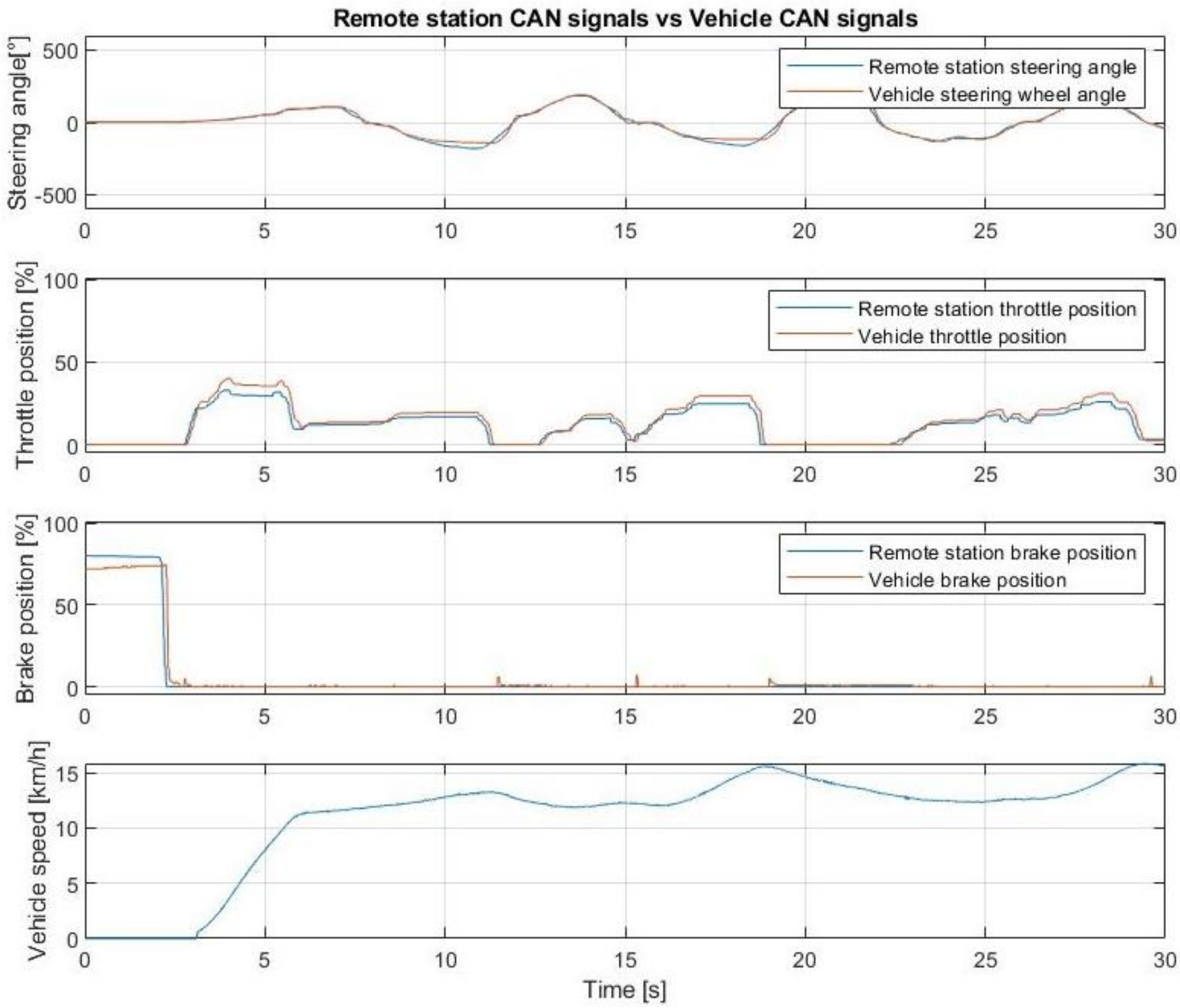


Figure 12 Teleoperation KPIs - Vlissingen results.

4.1.3 Cross-border testing

The Zelzate pilot is a crucial demonstration, where in addition to testing the technical capabilities, the focus is on the seamless handover of the network in the cross-border scenario. The test involved remotely driving the cars from the Netherlands to Belgium and back, with KPN and Telenet network towers strategically placed for the handover. The key factor assessed is the network downtime or service interruption time when crossing the border, as presented in Section 3.3.

Since it was tested in the public roads, teleoperation was performed in a shadow mode for safety purposes. Analysis of the results obtained during shadow mode teleoperation tests in Zelzate shows an interruption time of around 120 milliseconds, indicating a minimal downtime of service. Moreover, this low downtime does not hinder teleoperation, confirming the seamless ability for teleoperators to navigate cross-border situations.

4.2 Autodocking functionality testing

The autodocking functionality developed on a full-scale truck-trailer combination has been tested for two different controllers, Pure Pursuit Controller (PPC) & Model Predictive Controller (MPC), at two locations, MSP Onions in Vlissingen and Industriepark Kleefse Waard (IPKW) in Arnhem. The second location with commercial 4G/5G NSA connectivity is not part of any of the pilot sites, but due to the convenience of the location, partners from HAN used this small site for pre-testing of autodocking and truck teleoperation capabilities before moving the testing setup to Vlissingen (MSP Onions). The complete results of all the tests are described in detail in deliverable D4.2, and in this deliverable, we provide a summary of main results and lessons learned during piloting activities.

4.2.1 Testing Procedure

The test results of the PPC on the full-scale truck-trailer combination were gathered during the tests at MSP Onions in Vlissingen. The testing consisted over full automated teleoperator in-the-loop docking maneuvers executed multiple times to collect statistically significant data. Figure 13 shows a top view of how one autodocking test looked like.

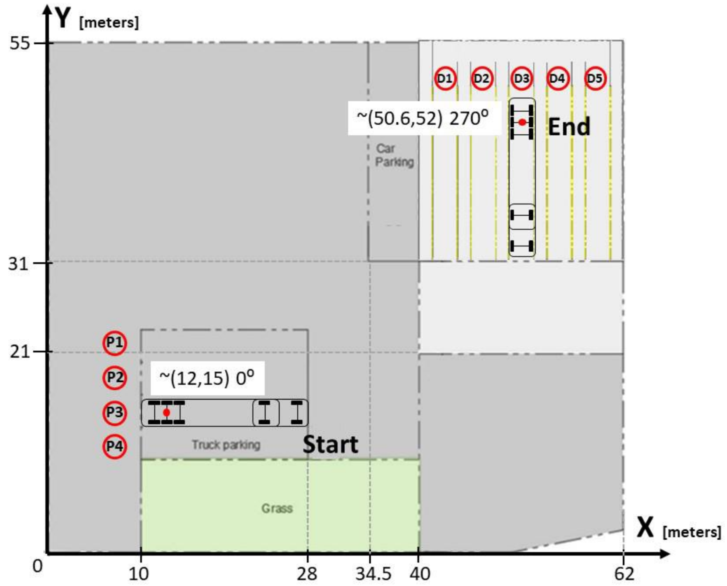


Figure 13 Top view of autodocking test at MSP Onions test site.

The vehicle combination was driven remotely on the premises of MSP to a starting position for the autodocking maneuver. Remote driving was done from the Teleoperation Centre located at the MSP Onions office. The docking maneuver was started at one of the Start points (either P1, P2, P3 or P4) and ended at the end point which is a dock (either D1, D2, D3, D4 or D5). The autodocking functionality was initiated from the teleoperation center and the truck started the docking maneuver. The overall maneuver consists of a forward path (curve to the right) and a rearward path (semi-straight line to the dock).

The test results of the MPC on the full-scale truck-trailer combination were gathered during the tests at IPKW in Arnhem. In essence, the tests were the same as at MSP Onions. The main difference were the total amount of docks (2 for IPKW) and the overall maneuver, consisting of a forward path (curve to the left) and a rearward path (curve to the right).

Throughout all the tests a safety operator was always in the vehicle to abort a test when required. The KPIs were measured with the use of the Real-Time Kinematic (RTK) Global Positioning System (GPS), the teleoperation Personal Computer (PC) and measurements by hand (i.e., tape measurements).

4.2.2 KPIs Definition

Before the measurements of the KPIs can be presented, it is important to discuss what the KPIs are and how they are measured. The KPIs for the autodocking functionality are listed in Table 15 along with the target values that are provided in Table 16.

Table 15 KPI measurement methodology.

#	KPI	Measurement Methodology
1	Path Planning Time	The path planning time was measured within the path planning software (MATLAB) by the use of a timer. The timer would start when the path planner started and would end when the path was planned. The measured time was logged in the datafile of that specific test.
2	Tracking Error Real Time	The tracking error was calculated within the path tracking controller software (MATLAB). The lateral tracking error was calculated by comparing the actual position of the truck-trailer combination (RTK-GPS) to the position where the truck-trailer combination should be according to the planned path.
3	Final Docking State error	The Final Docking State Error was measured by hand when the truck was finally stopped. The planned docking position was marked, and the difference was measured with a measuring tape. The final docking state error was also checked by comparing the final docking coordinates of the RTK-GPS with the end dock coordinates (pre-defined end point of path planner). This was done for both the lateral and the longitudinal errors. The Final Orientational Docking State error was calculated by comparing the orientation at the end point (logged from GPS) with the pre-defined path planner end orientation.
4	Elapsed Time	The total elapsed time was measured within the path planning and path tracking software (MATLAB) by the use of a timer. The timer would start when the path planner started and would end when the truck reached the end point and stopped moving. The measured time was logged in the datafile of that specific test.
5	GPS Position Accuracy	The GPS position accuracy was read out from the GPS system and logged in MATLAB. The accuracy was measured over the full test and averaged for a final value. The min. and max. values were also examined, but there were no outliers since the GPS position accuracy is very constant over time.
6	GPS Heading Accuracy	The GPS orientation accuracy was read out from the GPS system and logged in MATLAB. The accuracy was measured over the full test and averaged for a final value. The min. and max. values were also examined, but there were no outliers since the GPS orientation accuracy is very constant over

		time.

Table 16 Autodocking functionality KPIs definition.

#	KPI	Definition	Target values
1	Path Planning Time	The time it takes the path planner to plan the desired path for docking	< 60 [sec]
2	Tracking Error Real Time	The lateral (Y) deviation of the actual position of the axle of the trailing unit with respect to the generated path during maneuvering.	< 0.5 [m]
3	Final Docking State Error	The difference between the actual docking position and the planned docking position after the docking maneuver is performed. The Final Docking state error is divided into three parts: A) Lateral (Y) B) Longitudinal (X) C) Orientation angle (θ)	A = < 10 [cm] B = < 10 [cm] C = < 2 [deg]
4	Elapsed Time	The time between the initial movement and the final stop of movement at the end position.	< 150 [sec]
5	GPS Position Accuracy	The accuracy of the GPS positioning system in cm.	< 10 [cm]
6	GPS Heading Accuracy	The accuracy of the GPS Orientation in degrees.	< 1 [deg]

The target values are determined based on the time it takes for a regular driver to dock its truck-trailer combination and on the (limited) space available at a dock. For the available dock space, a dock at MSP Onions has been used as an example. The average clearances between tire and metal rails results in the tolerance of ±10cm absolute lateral error with respect to the center of the loading gate. Longitudinally, it is possible to overcome 10cm error with the tailgate of the trailer. And an orientation angle of more than 2 degrees will result in a possible collision with another truck-trailer combination located at the adjacent dock. The real time tracking error and GPS target values are based on the general available space in front of a dock and the final position / final state error target values.

4.2.3 Test results

Table 17 shows the average results of the KPIs together with the target values of both the Pure Pursuit Controller (PPC) testing at MSP Onions and the Model Predictive Controller (MPC) testing at IPKW. These are average values done over 48 successful autodocking tests at the MSP test site and 46 successful autodocking tests at the IPKW test site.

The 48 PPC tests were executed during three testing days at the end of February 2023. The weather was cloudy with no rain during the three days. No noticeable variations in weather were noted. The autodocking functionality was tested with the use of the 5G NSA network since the 5G SA network was not available at the test site. As described in the previous sections, the autodocking tests were performed using a forward movement (curve to the right) and a rearward movement (semi-straight line). The maneuver type was not changed during the 48 tests. However, the dock number (i.e., end position) was changed over time, and the starting position differed per test.

The 46 MPC tests were executed during three testing days at the beginning of September 2023. The weather was sunny with some clouds. No noticeable variations in weather were noted. The autodocking functionality was tested with the use of the 5G NSA network since the 5G SA network was not available at the test site. The autodocking tests were performed using a forward movement (curve to the left) and a rearward movement (curve to the right). The maneuver type was not changed during the 46 tests, yet the starting position was always kept random and the end position was changed over time.

Table 17 PPC and MPC KPI results (average).

#	KPI	Definition	Target values	Measurement PPC @MSP[1]	Measurement MPC @IPKW[Error! Bookmark not defined.]
1	Path Planning Time	The time it takes the path planner to plan the desired path for docking	< 60 [sec]	15.0 [sec]	11.0 [sec]
2	Tracking Error Real Time	The lateral (Y) deviation of the actual position of the axle of the trailing unit with respect to the generated path during maneuvering.	< 0.5 [m]	0.16 [m]	0.09 [m]
3	Final Docking State Error	The difference between the actual docking position and the planned docking position after the docking maneuver is performed. The Final Docking state error is divided into three parts: A) Lateral (Y) B) Longitudinal (X) C) Orientation angle (θ)	A = < 10 [cm] B = < 10 [cm] C = < 2 [deg]	A = 3.6 [cm] B = 8.4 [cm] C = 0.4 [deg]	A = 5.17 [cm] B = 8.33 [cm] C = 0.63 [deg]
4	Elapsed Time	The time between the initial movement and the final stop of movement at the end position.	< 150 [sec]	117.3 [sec]	153.59 [sec]
5	GPS Position Accuracy	The accuracy of the GPS positioning system in cm.	< 10 [cm]	3.7 [cm]	4.0 [cm]

[1] *Average of all the measured results*

6	GPS Heading Accuracy	The accuracy of the GPS Orientation in degrees.	< 1 [deg]	0.25 [deg]	0.07 [deg]

Table 18 shows the statistical relevance information per KPI of all the 48 tests performed with the Pure Pursuit Controller (PPC) at MSP Onions.

Table 18 KPI Statistical values (PPC controller).

KPI	Std. Deviation	95th Percentile	Unit
Path Planning Time	5.88	25.71	Seconds
Tracking Error Real Time	0.03	0.21	Meters
Final Lateral Docking State error	3.3	9.1	Centimeters
Final Longitudinal Docking State error	8.03	20.4	Centimeters
Final Orientational Docking State error	0.34	1.01	Degrees
Elapsed Time	13.07	135.88	Seconds
GPS Position Accuracy	0.17	4.00	Centimeters
GPS Heading Accuracy	0.29	1.00	Degrees

Table 19 shows the statistical relevance information per KPI for all the 46 tests performed with the Model Predictive Controller (MPC) at IPKW.

Table 19 KPI Statistical values (MPC controller).

KPI	Std. Deviation	95th Percentile	Unit
Path Planning Time	0.49	11.71	Seconds
Tracking Error Real Time	0.03	0.12	Meters
Final Lateral Docking State error	2.83	9.15	Centimeters
Final Longitudinal Docking State error	6.04	20.75	Centimeters
Final Orientational Docking State error	1.11	3.66	Degrees
Elapsed Time	4.91	160.40	Seconds
GPS Position Accuracy	0.01	0.07	Centimeters
GPS Heading Accuracy	0.02	0.08	Degrees

4.2.4 Discussion

It can be concluded that all KPIs are within the desired target values. Still, it is important to discuss the certain KPIs in a bit more detail. Table 20 gives an overview of all the KPIs of each individual version of the autodocking functionality that can be used to make the overall discussion more understandable. All the values shown in Table 20 are the absolute averages of all the tests performed.

Table 20 All KPIs for each version of the autodocking functionality.

KPI	MVP (1:3 scaled truck)	Full scale Truck (PPC)	Full scale Truck (MPC)
Path Planning Time	32 [sec]	15 [sec]	11 [sec]
Tracking Error Real Time	0.27 [m]	0.16 [m]	0.09 [m]
Final Lateral Docking State error	5.7 [cm]	3.6 [cm]	5.17 [cm]
Final Longitudinal Docking State error	10.2 [cm]	8.4 [cm]	8.33 [cm]
Final Orientational Docking State error	0.46 [deg]	0.4 [deg]	1.63 [deg]
Elapsed Time	153.4 [sec]	117.3 [sec]	153.59 [sec]
GPS Position Accuracy	3.8 [cm]	3.7 [cm]	4.0 [cm]
GPS Heading Accuracy	0.23 [deg]	0.25 [deg]	0.07 [deg]

Relying on good network quality

Throughout the tests performed at IPKW it became evident that the performance of the autodocking functionality is highly reliant on the network quality. The tests performed at IPKW used the 5G NSA network, which was a much poorer network than used at MSP Onions site. Figure 14 shows a screenshot of the network performance, where the download and upload latencies are in the order of hundreds of milliseconds (186 and 214 milliseconds respectively). This has a (negative) effect on the controller performance and overall KPIs.

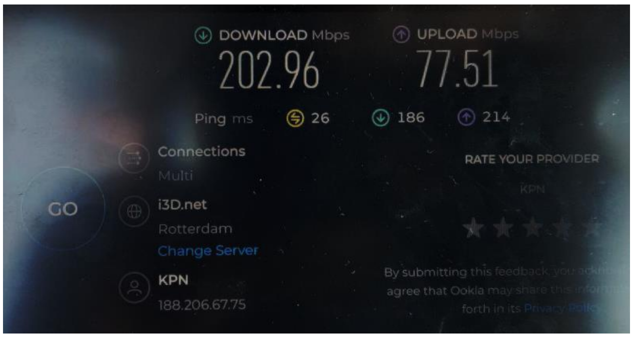


Figure 14 Network quality at IPKW.

During the tests at IPKW, the Round Trip Delay (RTD) values of the GPS related messages were logged. An analysis of the RTD was conducted to quantify the impact of network related delay on the MPC performance, see Figure 15. Although the average RTD (left graph) is around 70 [ms], it is not consistent. The spikes in the delay regularly are above 100 [ms], which is slower than the control loop. Furthermore, the variation in the delay is also very high. Variation causes undesirable behavior in systems, especially when it is greater than the control loop speed.

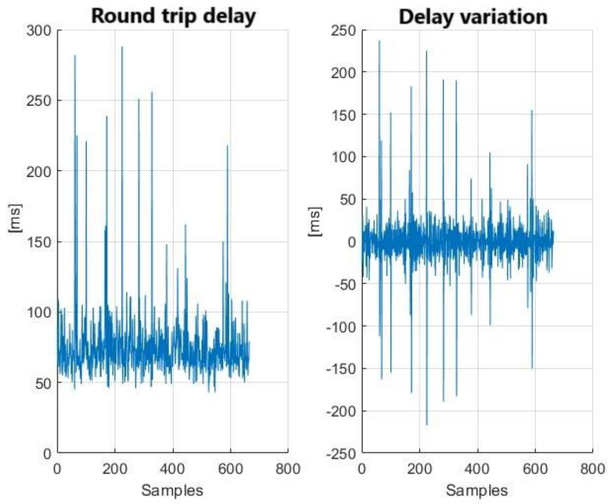


Figure 15 Round Trip Delay and Delay variation.

When the network speed fluctuates, messages from the remote station will also fluctuate in arrival timing at the destination. Figure 16 shows an example of this using logged data from March 2023 and September 2023. The graph shows the X position of the semitrailer for the same type of maneuver, at the same location, with the same network and settings in the system. The variation in samples received over time is clearly greater in September, than in March. This results in a higher standard deviation of delay variation.

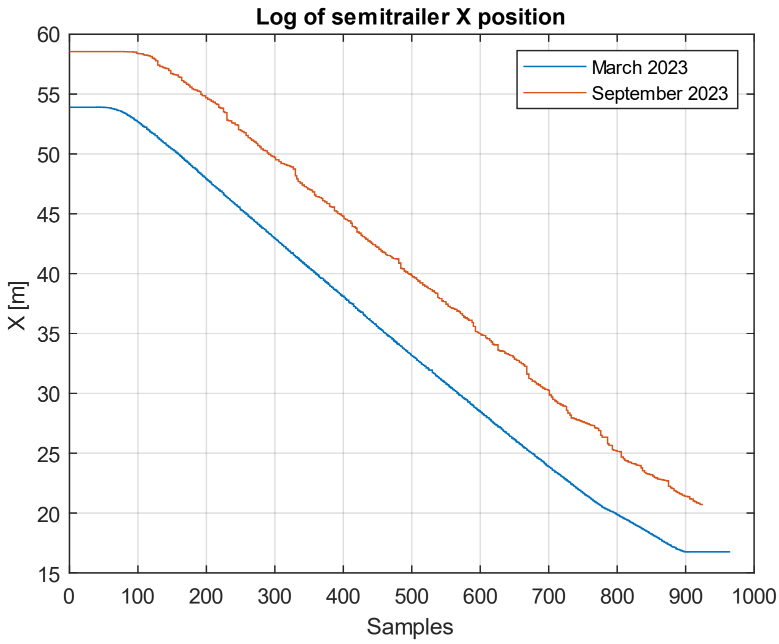


Figure 16 Variation in consistency of data samples received.

The delay variation not only applies to receiving input data to the controller, but also to the control outputs being sent to the truck. When steering command signals are sent from the autodocking controller PC to the truck in a varying frequency, the truck will become oscillatory. Oscillatory control outputs (in this case steering commands) are known to be caused by delays in communication. An example of this is shown in Figure 17, where the blue lines in the left graph show the oscillatory behavior. The result of this oscillatory behavior can directly be seen in the right graph, showing the tracking error.

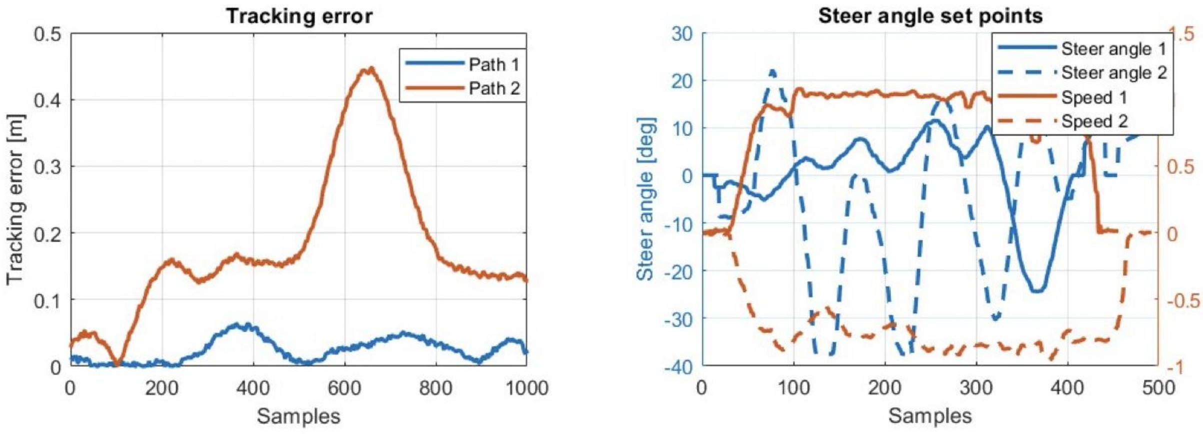


Figure 17 Effect of network quality on autodocking performance.

Statistically, looking at the standard deviation of the RTD gives more insights. Figure 18 shows a histogram comparison between 2 tests, one which was successful (left) and another where the KPIs were not satisfied (right). The variation of the unsuccessful test deviates much wider, with fewer counts close to 0 compared to a successful test.

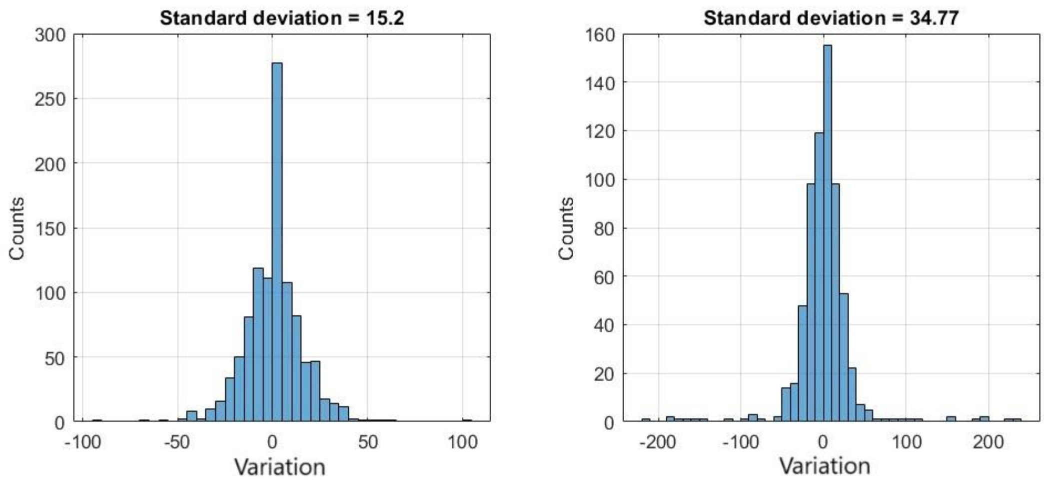


Figure 18 Histograms of the delay variation in [ms]. Left: successful test, Right: unsuccessful test.

It is evident that the performance of the autodocking functionality is highly reliant on the network quality. A stable network with a delay less than 100 [ms] (control loop speed and TC cut-off threshold) will of course give the best results. But this is not always the case. To still perform the autodocking, cost functions of the MPC can be changed. Without going into further details, the weightage given to the steering suppression was increased, resulting in a smoother, slower steering response. In other words, this slows down the behavior of the controller such that the delay variation will not cause rapid changes in control output.

It is important to note that this is only possible because of the advanced control strategy of the MPC. Furthermore, the tuning method only works due to the slow speed nature of the autodocking situation. If faster dynamics were to be controlled, such as a lane change maneuver at high speed, the network will definitely need to be consistent and fast.

Path Planning Time

The Path Planning KPI is well within the target value of under 60 seconds for each version of the autodocking functionality. The speed of the path planner is mainly determined by the strength of the computer it runs on. The better the processer of the computer, the faster the path planner. Additionally, looking at Table 20, there is also a big improvement noticeable between the MVP and the full-scale truck. The computer that runs the path planner did not change between MVP and full-scale truck, but the code generation for the path planner was improved which more than halved the path planning time. Since the path planning takes place locally on the remote PC, it is not affected by the network parameters. But it does of course affect the overall performance since it is part of the total end-to-end latency.

Tracking Error Real Time

The tracking error real time is the tracking error of the truck-trailer combination during the docking maneuver and indicates how well the truck-trailer combination can follow the planned path. This KPI is highly sensitive to the right path following controller parameters. The tracking error of both the PPC and MPC is very constant with a standard deviation of just 3 centimeters (Table 18 and Table 19). When the network quality is bad, a noticeable difference can be seen. When latency is high, or fluctuates, the truck deviates more from the planned path since the information for the path following controller (GPS coordinates) arrive later which in return delays the control commands that will also arrive later at the truck (as shown in Figure 16). So it is important to notice that network quality can be a bottleneck for autodocking, especially when the network quality is bad as explained in paragraph on Relying on good network quality.

Final lateral, longitudinal and orientational docking state error

The final docking state errors are very important KPIs when it comes to autodocking since these correspond to the end position of the trailer. Big errors mean that the truck trailer combination is not parked rightly at the dock. With only 10cm play at most docks laterally, these KPIs come very narrow. The errors are also highly dependent on network quality for similar reasons as described in the previous subsection. High latencies result in big offsets at the end position. This was especially noticed in narrow surroundings when the network quality was bad. When surroundings or places to dock are narrow, the paths that are planned are tight. If big offsets occur due to high latency, it will be quite challenging to park the truck-trailer combination straight, for example straight at a dock. With a good network quality (consistent delays and delays <100 [ms]), this is not the case. The autodocking functionality will still work properly. But, it can be concluded that network quality is a highly important factor of these KPIs, especially if test ground dimensions are tight. This was not the case during the testing at MSP onions since there was enough space and the network quality was good, but it was the case when testing the MPC controller at IPKW. Unfortunately, the network quality was very poor compared to the tests at MSP Onions. This resulted in worse KPI values as described in paragraph on Relying on good network quality.

Looking at Table 20, it can be noticed that the longitudinal docking error is higher compared to the lateral docking error for both the PPC and MPC controller. This can be explained by the fact that speed control is not yet implemented in the autodocking functionality, making stopping at exactly the right spot longitudinally quite challenging. As of now, the truck will stop with applying throttle some distance before the desired end point and apply the brakes at the end point. Tuning and timing this to be exact is challenging. If the truck is for example on a slope, or approaching with a bit higher speed, the longitudinal error changes. When implementing speed control (maintaining a fixed speed), determining the brake point is easier. During the testing, this was not yet implemented hence the higher error value for the longitudinal error. Additionally, network quality also plays an important role in the longitudinal error. If for example the latency is 0.1 seconds at a docking speed of 5 kmph, a delayed braking signal of 0.1 seconds already results in a 14 centimeter overshoot longitudinally. This will eventually result in a higher standard deviation which is a result of different docking slope, minor variations in approaching speed, GPS accuracy and Network quality which all effect the moment the truck will brake near the end point.

Elapsed time

The elapsed time KPI is not necessarily dependent on network quality, but there are other factors that affect it. For example, the path planning time is included in the total elapsed time so the factors that affect the path planning time also affect the elapsed time. Furthermore, the overall docking maneuver and especially the length of the paths affect it as well. Longer paths take longer to drive since the overall speed of the truck is ±5 kmph. Speed is therefore also an important factor. There is an improvement of 36.1 seconds to be seen when the MVP is compared to the full-scale truck (Table 20). This is mainly due to the reduction of path planning time as explained in the previous paragraph.

The elapsed time of the MPC testing is way higher compared to the PPC testing. The reason for this is the length of the planned path which was (much) larger at IPKW compared to the paths at MSP Onions. It can be acknowledged that this KPI is therefore not as meaningful. It is suggested that in future research, it may be more meaningful to use a KPI that is based on a prescribed minimal velocity. This change is proposed to create more objective metrics, as the current KPI is heavily influenced by the length of the maneuver in both directions, which was not consistently defined.

GPS position and orientation accuracy

The GPS accuracies are very important KPIs since they highly affect the overall errors and therefore the functionality of the autodocking system. These are listed as KPIs since the play at dock is on average 10 centimeters which calls for a highly precise GPS system. As can be seen in Table 18 and Table 19, the KPIs are well within the target values and are also very constant. The position accuracy KPI has a standard deviation of just 0.17 and 0.01 centimeters and the orientation KPI has a standard deviation of 0.29 and 0.02 degrees. Also, when looking at the 95th percentile, 95 percent of all the measured position and orientation accuracies are within the target values of 10 centimeters and 1 degree respectively.

It might be interesting to note that RTK GPS is a very suitable and robust method for localization for autodocking. Accuracies are high and very stable, which is needed for autodocking. However, the costs for a RTK system as used in the project are very high, and it is recommended to study alternatives. One of those alternatives is the 5G network itself, which can offer RTK GPS corrections over the cellular network. This feature was not yet fully developed to be implemented during the 5G-Blueprint project, but it is a very useful and promising feature of the 5G network for applications like the autodocking functionality that require high precision localization.

PPC vs MPC

Now that the KPIs are discussed and their reliance on network quality is addressed, it is important to discuss which form of controlling the autodocking functionality is preferred. Unfortunately, due to poor network quality at the IPKW test site, as explained in the section, the KPI values of the MPC controller are not as “set in stone” that they immediately prove that the MPC is preferred over the PPC controller. However, the lack of visible improvement in terms of KPI values is mainly a result of the poor network quality. The reasons why the MPC controller is still preferred over the PPC controller are listed below.

- **Less fluctuations in controller behavior:**
  MPC takes into account a dynamic model of the system and predicts future states over a specified prediction time horizon. It optimizes control inputs over this horizon to minimize a cost function. This predictive nature of MPC allows it to proactively account for potential disturbances and uncertainties. In contrast, PPC typically relies on fixed classical control laws and may react less effectively to changing conditions or disturbances. This can result in more fluctuations in the system's behavior. This can be seen in the tracking error real time KPI (Table 20), where the real time tracking error of the MPC is 9 centimeters compared to the 16 centimeters for the PPC even though the network quality was much poorer during MPC testing.

- **Smoother Path Following because of the ‘predictive’ component:**
  MPC generates control inputs by considering future states and the desired path, allowing it to plan and execute control actions more smoothly. It can anticipate upcoming changes in the path and adjust the control inputs accordingly. PPC, on the other hand, often relies on reactive control strategies that may lead to abrupt changes in control inputs when tracking a path. This can result in jerky or non-smooth behavior, which is undesirable. This was especially noticeable during the docking maneuvers where the steering output of the PPC controller was quite oscillatory. The MPC steering control is very smooth with almost fixed steering angles and less fluctuations.

- **Capability to work in tighter areas and make tighter curves:**
  MPC is particularly advantageous when navigating through tight, limited space spaces or when precise path tracking is required, as it can make finer adjustments to control inputs based on its predictive capabilities. PPC may struggle in such situations, as it may not have the ability to plan and execute control inputs as effectively in constrained environments or when navigating tight curves. Especially when reversing a truck, the tractor always needs to counter-steer to get the trailer to the right place. MPC accounts for this with its predictive capabilities and can therefore counter-steer ahead, something the PCC cannot do. The PPC will therefore be too late with counter-steering and therefore the trailer won’t end up in the desired spot when the area is narrow. This was very noticeable at the IPKW testing. The area there is tighter and the PPC controller had way more trouble to successfully dock the trailer where the MPC controller had no issues at all.

In summary, MPC's ability to predict and optimize control inputs over a horizon, its capacity to adapt to changing conditions, and its capability to provide smoother path tracking, makes it the preferable choice compared to conventional PPC, especially in applications where precision, adaptability, and smoothness of control are crucial, such as tight and narrow distribution centers for example. It might also be worth mentioning that even though the network quality was poor, the results of the MPC are still impressive and within KPIs. The MPC is therefore way more stable than the PPC when it comes to network quality. This is also a very big advantage of the MPC compared to the PPC.

4.3 Automated barge control

The goal of the tests presented in this section is to validate the importance of 5G Standalone (SA) connectivity for the teleoperation of barges and evaluate the capabilities of 5G SA for this purpose. This research and development activity has been done in the two main areas, where the waterway transport experience challenges in connectivity. The first one is the port area where there is traffic of other barges. The second area is the Zelzate cross border, where in the previously 4G-based setup, the connection could have been lost for a significant amount of time when crossing the border. Further details of the automated barge control use case, as well as the overall network performance with detailed testing steps, are presented in D4.1. As in the case of the previous two use cases, this deliverable summarizes the main achievements of UC4.1 as well.

4.3.1 Preconditions

The setup and getting the device in an active mode, requires several tests onsite, where there is enough coverage. To investigate the capabilities of 5G SA for teleoperation of the barges, we have done several tests with one commercial barge with a length of 110 meters (Figure 19) in the port area and we have done several tests with an urban barge with a length of 5 meters in the Zelzate cross border area (Figure 20).

Seafar has done 18 tests in total. Nine of these tests took place in the Port of Antwerp and nine other took place in the Zelzate cross border area. Some of these tests were done on car to test the network connectivity together with network providers to tackle the challenges such as NSA to SA switching. The details of these tests are explained in Deliverable 4.1.

Figure 19 Operational barge utilized by Seafar for the tests on 5G SA.

The illustrated barge is equipped with Seafar control system and all the required sensors and devices as explained in detail in Deliverable 4.1. The tests in Zelzate cross border area are done with an urban barge named AVATAR. This boat is also equipped with various sensors and devices to enable remote operation from Seafar control station in Antwerp. More photos and detailed information can be found in Deliverable 4.1.

Figure 20 AVATAR urban barge used for demo purposes in the cross-border area.

4.3.2 Test results

In the following tests, there is a part explaining the coverage map, showing circles of the signal quality over the trajectory. These circles are coded with colors. The code description is as follows.

The map shown in Figure 21 illustrates the strength and the quality of signals during a) test in Port of Antwerp and b) test in Zelzate. The Wide Area Network (WAN) quality on the map is colored in three circles depending on the values of latency, signal strength, and signal quality. The outer circle is indicating the latency (Green: Lower than threshold of 30ms, Red: higher than threshold). The middle circle indicates signal strength (Green: excellent, Dark Blue: very good, Light Blue: good, Orange: Fair, Red: poor, Dark Red: very poor). The inner circle is indicating the signal quality (Green: excellent, Dark Blue: very good, Light Blue: good, Orange: Fair, Red: poor, Dark Red: very poor). The most important KPIs for the barge teleoperation, which are presented in the following sections, include i) successful connection to the network, ii) good signal quality and coverage, and iii) end-to-end latency.

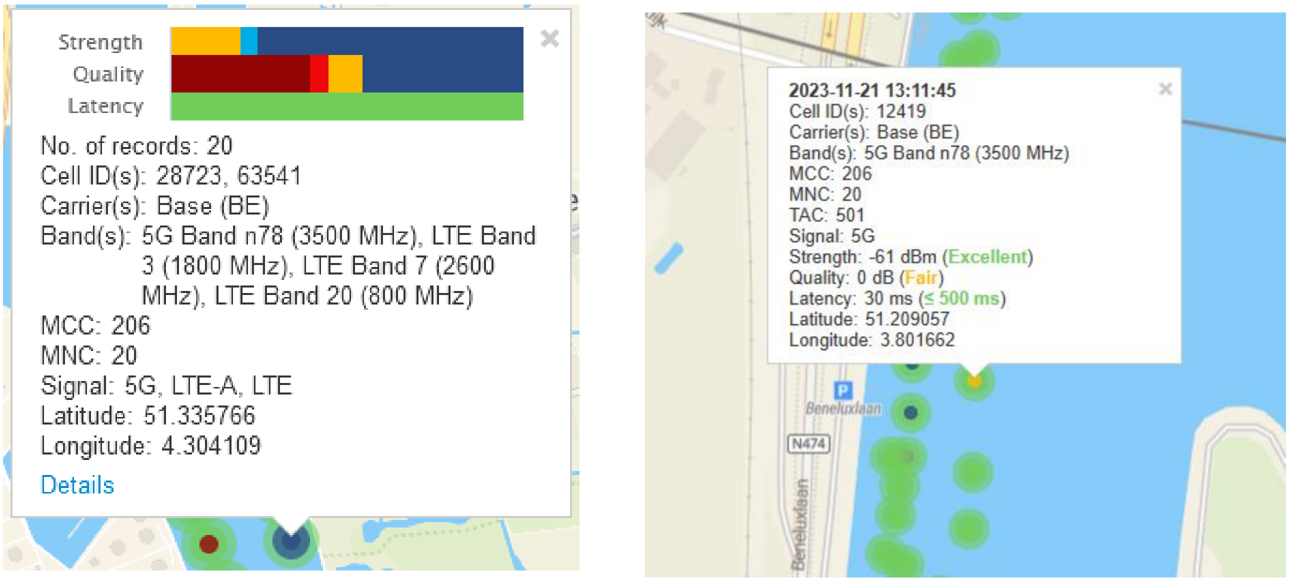


Figure 21 Example of color code for the measurements a) during a test in Port of Antwerp, and b) during a test in Zelzate.

4.3.2.1 Testing in Antwerp pilot site

Tests with IDs T1.6, T1.7, T1.8 were some of the successful tests for sailing of the vessels in Port of Antwerp area. The summary of these tests are as below.

Test T1.6

This test was to verify the results and report the signal quality of the 5G SA signal. The test with the barge shown in Figure 19 is done on September 19, 2023 from (14:50) to (15:30).

The WAN quality based on signal strength for the signals utilizing 5G SA dome is illustrated in Figure 22. Based on the figure, we can see that the signal strength fluctuates among Poor, Fair, Good, Very Good, and Excellent during the start of the test and stays on Very Good and Excellent during the rest of the test window.

Figure 22 5G SA signal quality during sailing in T1.6.

The end-to-end latency during the test is between 21 and 50ms as shown in Figure 23. The average latency is 25ms, which is sufficient for safe teleoperation.

Figure 23 5G SA latency during sailing in T1.6.

Table 21 UC4.1 test results in Antwerp pilot site (T1.6).

KPI	First Quartile	Mean	Third Quartile	Standard deviation
Latency (ms)	23	27.23	31	12.98

This test shows sailing over 5G SA. In Figure 25, we can see that at some point the signal is lost. During this failure, our network balance onboard switched to 5G NSA to keep the teleoperated sailing smooth and seamless. The early issues on switching can be read on Deliverable 4.1. Number of Data Points collected in this test is 1500 and the summary of the key results are presented in Table 21.Table 21 UC4.1 test results in Antwerp pilot site (T1.6).Table 21 UC4.1 test results in Antwerp pilot site (T1.6).

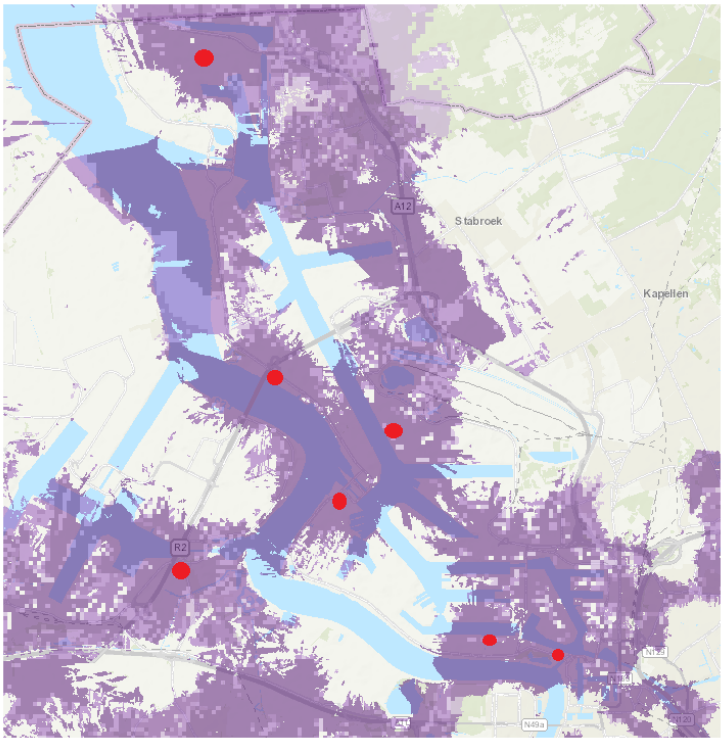


Figure 24 Coverage map provided by Telenet in Port of Antwerp area on April 26, 2023.

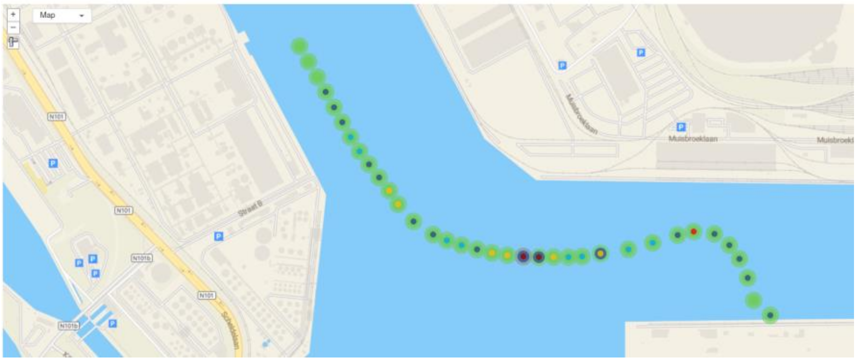


Figure 25 5G coverage map during sailing in T1.6.

Test T1.7

This test took place on September 19, 2023 similar to previous test using the same barge. Based on Figure 27, we can see that in the coverage area of Telenet there is good to fair coverage, with some fluctuations in signal quality but also good latency with an average of 30ms.

Figure 26 5G SA signal quality and latency in T1.7.

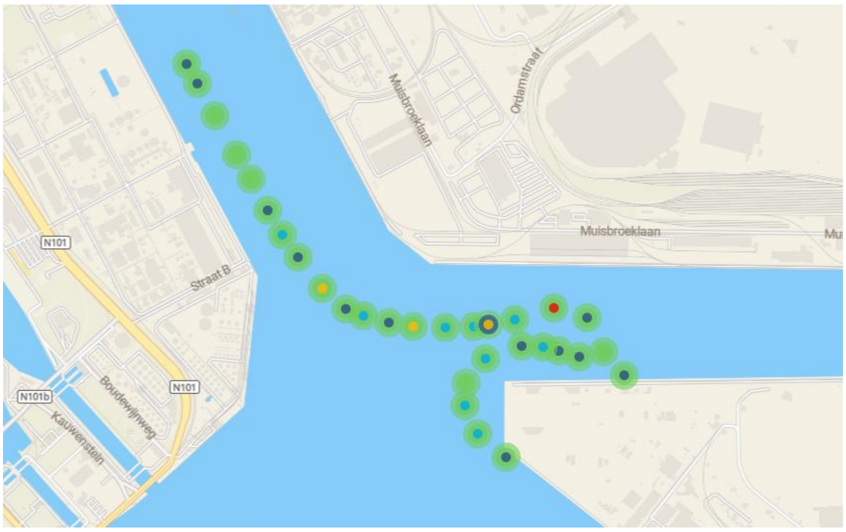


Figure 27 5G SA coverage and trajectory in T1.7.

Number of Data Points collected in this test is 950 and the summary of the key metrics is presented in Table 23.

Table 22 UC4.1 test results in Antwerp pilot site (T1.7).

KPI	First Quartile	Mean	Third Quartile	Standard deviation
Latency (ms)	24	33.35	40	59.36

Test T1.8

This test took place on October 30, 2023, and it is similar to previous test using the same barge. Considering Figure 28 and Figure 29, we can see that far from the SA base station there is low signal quality, high fluctuations in signal quality and latency happened. At some locations, the latency was even more than 300ms.

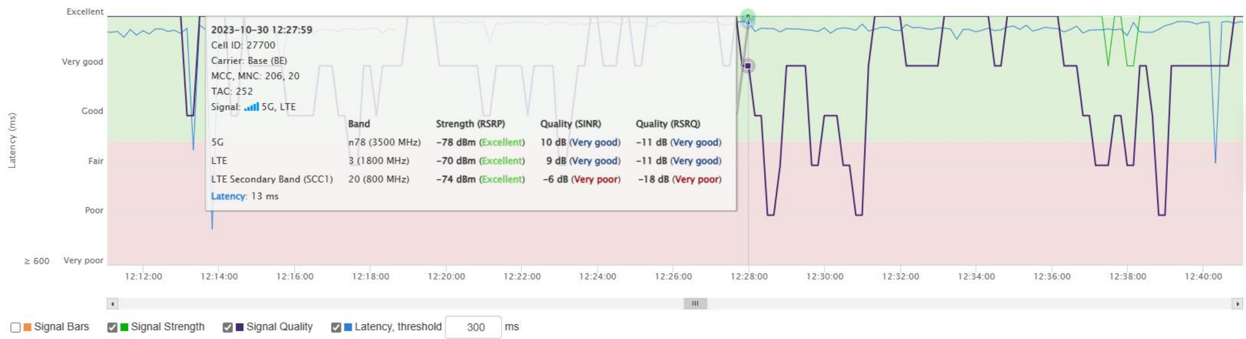


Figure 28 5G SA signal quality and latency in T1.8.

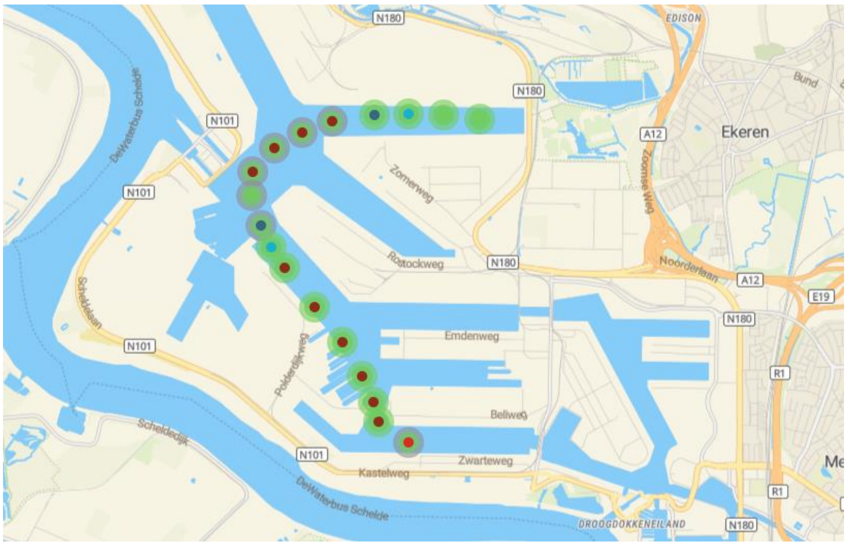


Figure 29 5G SA coverage and trajectory in T1.8.

The low quality of 5G SA signal was further checked with Telenet, and it was due to the 5G SA Core updates that happened on October 27. Number of Data Points collected in this test is 600 and the summary of the results is presented in Table 23.

Table 23 UC4.1 test results in Antwerp pilot site (T1.8).

KPI	First Quartile	Mean	Third Quartile	Standard deviation
Latency (ms)	15	20.02	21	26.97

4.3.2.2 Testing in Zelzate pilot site

Similar to the case of Port of Antwerp testing, nine tests were done in the Zelzate cross border area and some of them were done with a car together with network providers to test connectivity. The details of these tests are available in deliverable D4.1. Tests with IDs T2.5, T2.6, and T2.8, are some of the successful tests done with remote operation of the AVATAR urban barge. The summary of these tests are provided as below.

Test T2.5

This test was done on Nov 6th, 2023. The results in Figure 30 (signal quality) and Figure 31 (signal coverage) show an improved connectivity to KPN base station.

Figure 30 5G SA signal quality connected with KPN SIM card and sailing across the border in T2.5.

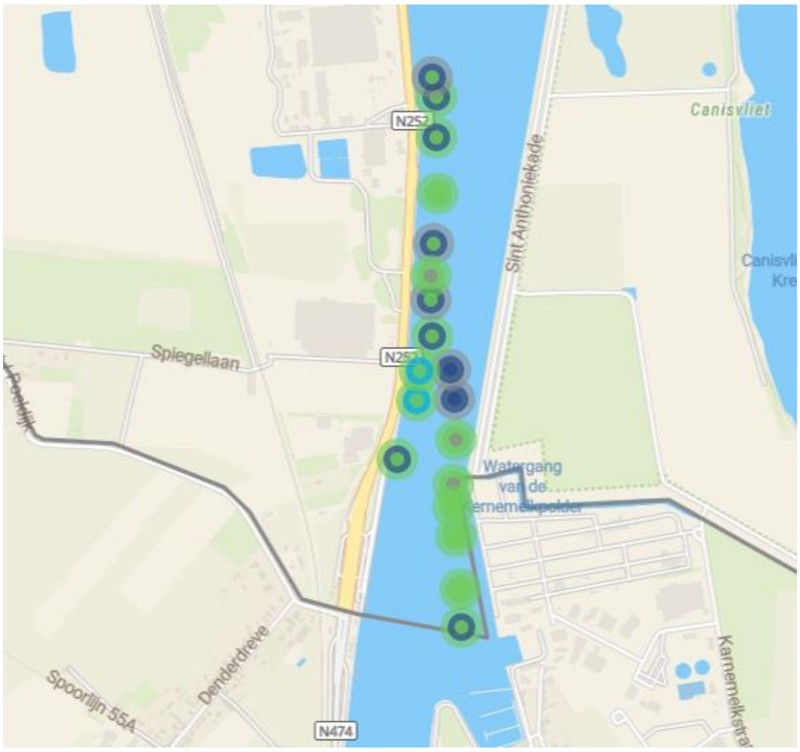


Figure 31 5G SA signal quality connected with KPN SIM card and sailing across the border in T2.5.

Number of Data Points collected in this test is1300 and the summary of the key results is presented in Table 24.

Table 24 Cross-border tests of UC4.1 (T2.5).

KPI	First Quartile	Mean	Third Quartile	Standard deviation
Latency (ms)	25	40.47	38	38.89

Test T2.6

Another test was conducted on Nov 7th with the urban barge while sailing in the cross border area. During this test, we switched the SIM cards and used Telenet SIM card instead of KPN and started the sailing from Belgium to the Netherlands. The results showed an improved connectivity. Videos were recorded from remote control station and the barge was sailing in the waterway, while the remote operator was controlling and steering the vessel from the Antwerp office. The latency is shown in Figure 32 and the 5G SA signal coverage is shown Figure 33.

Figure 32 5G SA signal quality connected with Telenet SIM card and sailing across the border in T2.6.

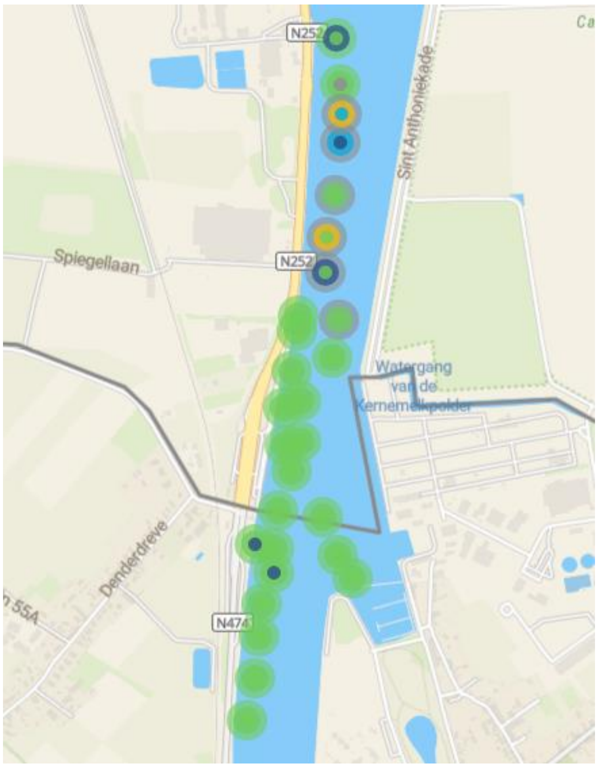


Figure 33 5G SA signal quality connected with Telenet SIM card and sailing across the border in T2.6.

One remark is that with Telenet connectivity we experienced higher signal quality, which is line with network performance evaluation reported in Section 3.3. During this test, some specific challenges such as antenna position on the boat and IP issues were detected, which are further explained in detail in deliverable D4.1.

Number of Data Points collected in this test is1400 and the summary of the key results is presented in Table 25.

Table 25 Cross-border tests of UC4.1 (T2.6).

KPI	First Quartile	Mean	Third Quartile	Standard deviation
Latency (ms)	35	39.07	43	6.05

Test T2.8

This test was done during the dry-run on Nov 20, 2023. The AVATAR barge was sailing in the cross border area from the NL to BE and the opposite direction. The remote captain was making continuous U-Turns so the connections can be tested sufficiently. In this test we realized that increasing the height of the antenna (the local on-vessel broadcasting antenna of the Peplink) and whitelisting all relevant IPs, solved the detected issues in previous tests. The connection was overall good and we were able to sail remotely from the office facing no issue around the border. The signal quality and latency are shown in Figure 34 and the signal coverage is shown in Figure 35.

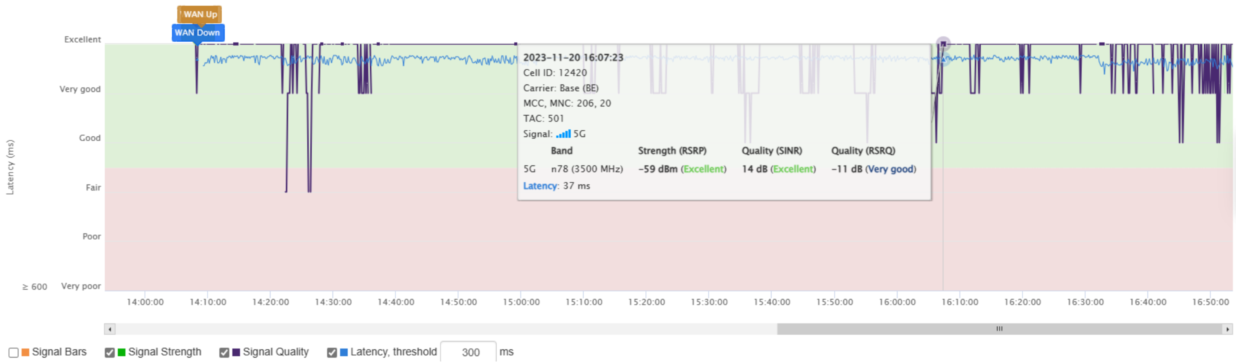


Figure 34 5G SA signal quality and latency during dry-run test in T2.8.

Figure 35 5G SA signal coverage and trajectory during dry-run test in T2.8.

Number of Data Points collected in this test is1200 and the summary of the key results is provided in Table 26.

Table 26 Cross-border tests of UC4.1 (T2.8).

KPI	First Quartile	Mean	Third Quartile	Standard deviation
Latency	29	35 (ms)	46	26.35

In cross-border areas, 5G SA exhibits low latency, crucial for seamless automated sailing, especially near structures like the Zelzate bridge. 5G SA demonstrates superior bandwidth, improved video quality, enhancing the remote navigation experience, particularly in busy port locations. Despite exceptional performance within the designated test area, the study emphasizes the need for a thorough examination of the broader application of 5G SA, considering its limited coverage and potential challenges in harsh port conditions.

4.4 Assessment of 5G capabilities for VRU Warning

Overall testing procedure

This trial validated the operation of the EF2 solution in different 4G and 5G network setups, operational settings and usage scenarios.

Application

The solution that was tested consists of the VectorDrive smartphone app and exchange service. The app automatically determines the most likely path of the Vulnerable Road Users (VRUs) and detects whether this path intersects with the likely path of TOVs at a specific moment in time. The TOV paths are generated based on live data of the TOVs (UC4.3) and itinerary data from EF7.

Information on the expected paths of VRUs and TOVs are exchanged through the exchange service. Each instance of the Vectordrive app detects potential collisions with all nearby TOVs. Potential collisions are posted to the exchange service, and subsequently presented to the TO in the EF1 dashboard.

A more elaborate description of VectorDrive is available in deliverables D6.1 [2], D6.2 [3] and D6.3 [4].

Goal

The goal of the trials was to establish whether the URLLC slice in the 3.5GHz band of a 5G stand-alone network[2], provides the performance that is required to deploy safety critical services such as VectorDrive, i.e., VRU warning, in operational environments.

Pilot sites

For each network setup a set of trial scenarios was completed in both an *industrial setting* (port area Antwerp) and *urban setting* (city centre of Mechelen) (Figure 36). The selection of the sites was based on the availability of operational 5G nodes.

Initially all field trials were to be carried out in Zelzate using the Telenet 5G SA node available in the city centre of Zelzate. Because availability of the 5G slice in Zelzate interfered with the setup for the cross-border UC trials, it was decided to move the EF2 trial sites to Antwerp and Mechelen respectively, as these sites provided continuous availability of the URLLC slice on a 5G SA network. An additional benefit of the selected sites was that they provided larger and more uniform industrial and urban testing environments than are available in Zelzate. Trials were carried out in June, July and September of 2023. Note that the Mechelen site is not an official pilot site of the 5G-Blueprint project, but, as described in Section 2, it is a result of extending the Antwerp pilot site to a new location with 5G SA coverage (one gNodeB deployed at the Telenet Headquarter, urban environment), for the purpose of: i) having an ad-hoc testing and debugging setup of 5G NR, and ii) creating urban environment setting for EF2 testing.

The Antwerp trial site is located in the port area and features multiple Telenet 5G base stations, interspersed with areas without 5G coverage. The area is dotted with large metal structures (warehouses, a container terminal, chemical plants) and crisscrossed by rail and road connections and metal fences (Figure 37). Between the structures there are large empty plots.

The Mechelen site is near the city centre and hosts one single 5G SA cell. Most neighbourhoods in the area feature low-rise buildings in narrow streets (Figure 38). In the centre of the area there

[2] Where "5G" is mentioned in this section, it should be read as the URLLC slice in the 3.5GHz band of a standalone 5G network.

are a few large buildings (Football stadium, Telenet HQ, prison). A canal and major urban road intersect the site on the south side.

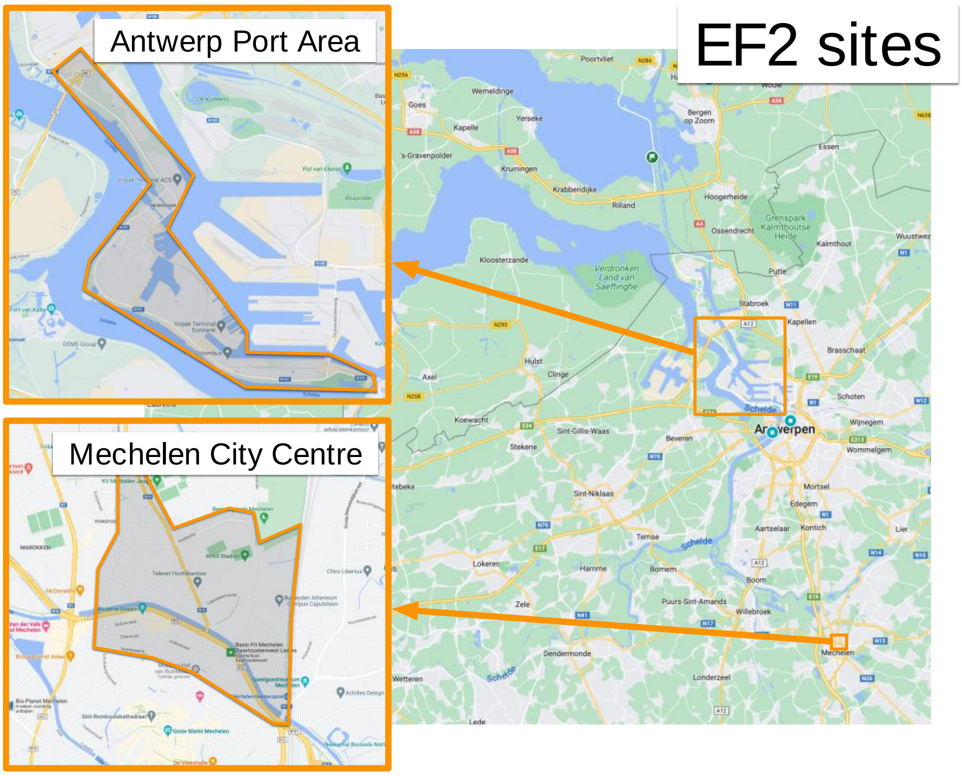


Figure 36 Location of the EF2 sites.

Figure 37 Impression of the ‘industrial’ site in the Port of Antwerp.

Figure 38 Impression of the trial site in the 'urban' site in Mechelen.

Trial setup

All trial probes (TOVs and VRUs) carried two identical handsets (Oppo X5 Pro 5G), one with a 5G SA-capable SIM that guides traffic to the exchange service via the Ultra-Reliable Low Latency Communications (URLLC) slice that provides high priority and low latency, and one with a commercial 4G SIM to provide benchmark data. The 5G handsets were configured to connect to the SA network (Figure 39).

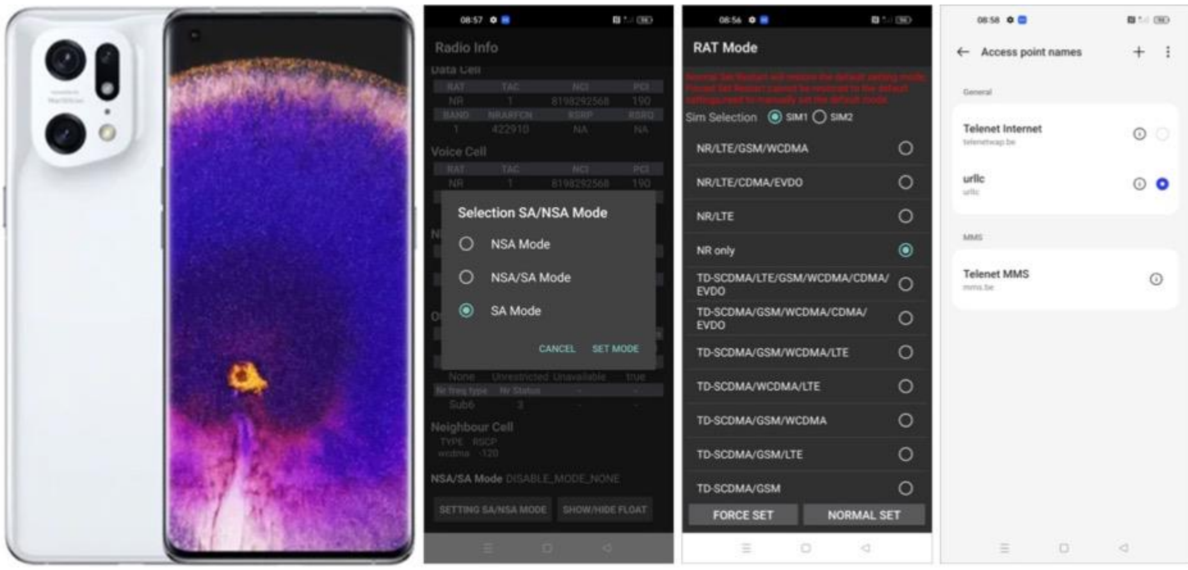


Figure 39 Oppo Find X5 Pro 5G configuration.

All VRU-probes carried out multiple runs per trial site, applying different usage scenarios (Figure 40):

1. Ideal circumstances
   a. Pedestrian: holding the handset in hand
   b. Cyclist: handset mounted on handlebar

2. Realistic usage conditions: carrying handset in a rucksack or in a pocket

Due to the limited availability of handsets that could connect to network, the trial combination of urban - ideal - pedestrian was trialed for 5G only, i.e., could not be benchmarked against 4G.

Figure 40 Trialing ideal and realistic usage scenarios.

Network traffic characteristics and messages exchanged on the application level were logged in a cloud data store (Figure 41). Trial runs were repeated until the exit conditions were met per trial site and usage scenario.

KPI data were collected as follows:

- By each individual app by logging user and datacom activity and characteristics on the central data store.
- All messages posted on the MQTT were stored for analysis.

A car driven by a person acted as live TOV. VRUs (on foot and bicycle) approached TOV paths perpendicularly and longitudinally, to trigger potential collision messages.

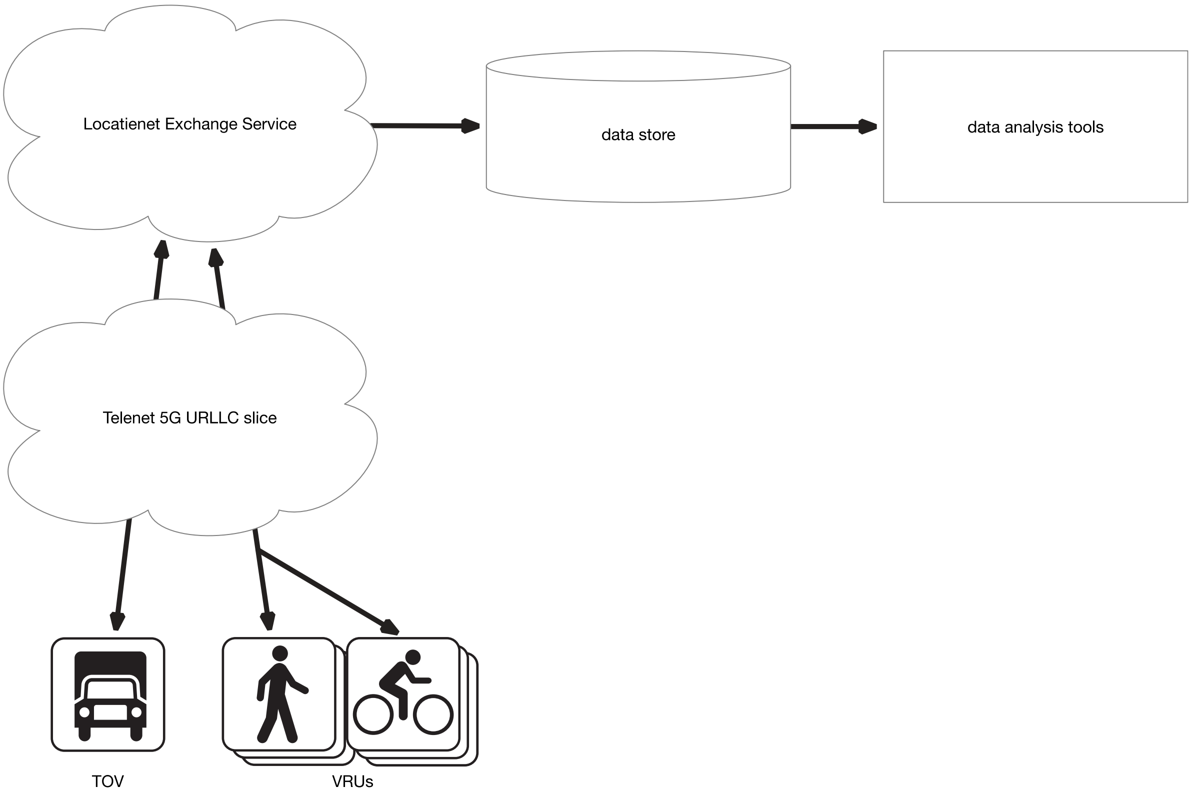


Figure 41 Setup for the logging of messages and network characteristics for analysis.

KPIs

Table 27 List of EF2 KPIs.

#	KPI	Definition	Target values	Measurement	Status
EF2-1	Service Continuity	Percentage of time during which smartphone apps were operational during each field trial.	98%	crash monitoring tool	Trialled in both trial sites in June, July, Sept 2023
EF2-2	Service Continuity	Percentage of time during which MQTT service was operational	98%	Server log file	Trialled in both trial sites in June, July, Sept 2023
EF2-3	Service Continuity	Network awareness: % of times the radio connection was reliable enough to timely warn VRUs	RSRP > -105 dBm in 98% of time	Network monitoring log	Trialled in both trial sites in June, July, Sept 2023
EF2-4	Reliability	Number of messages made available via MQTT Broker with position of VRU, and potentially warning, per hour	98% of 32,400 messages	Server log file	Trialled in both trial sites in June, July, Sept 2023
EF2-5	Latency	Roundtrip time for messages exchanged with MQTT server	<200ms	Server log file.	Trialled in both trial sites in June, July, Sept 2023

The definition and target values of these KPIs have been improved from those in D7.2 [1]:

- EF2-3:
  - Definition in D7.2: “Network awareness: % of times apps correctly warned VRU for network degradation”, target “98%”.
  - Reason for change: RSRP from the smartphone operating system is the only source available to determine the quality of the radio connection. The same source is used by the app to warn the end-user. Hence outcome of analysis would always be 100% and have no meaning.
- EF2-4:
  - Definition of target in D7.2: “3600 * 9 (# Quad tiles in detection zone)”
  - Reason for change: 3600 * 9 (= 32,400) is the actual number of messages that is exchanged per hour, hence the target definition in D7.2 corresponded to a 100% score. The service is designed that 2% of messages may be lost without inhibiting VRU safety, which corresponds to 2.4 seconds of the warning horizon of 2 minutes

(derived in D6.2, section 2.2.1).

- EF2-5:
  - Definition in D7.2: “Time between detection <500ms and warning to TO”, target: “<500ms”.
  - Reason for change: low latency is critical for all exchanged messages, i.e., not just warnings. The 500ms target value was not in line with the initial network requirement of 100ms latency one-way (defined in D6.3 and used in the design phase). Measuring the roundtrip time on the device proved to be the only way to reliably measure the latency in the application layer.

Pre-processing

5G coverage in the trial sites was limited to one 5G node (Mechelen) or dispersed 5G nodes (Port of Antwerp). This meant that trial probes sometimes went outside 5G coverage. All data records – 5G and 4G - captured outside the 5G coverage area of the nodes were deleted before starting the data analyses. The coverage area was assumed to be circular with the cell tower as center. The diameter of the circles was determined by determining the disconnect points on East-West and North-South transects that were driven before starting the trials. For the comparison tests equal data set sizes were created by removing incomplete sample records.

EF2-1 - Service Continuity Apps

Crashes of apps were registered by a crash monitoring tool (Crashlytics). Per trial site the average offline time for all devices was calculated.

Table 28 Details of KPI EF2-1.

#	KPI	Definition	Target values	Trial result
EF2-1	Service Continuity	Percentage of time during which smartphone apps were operational during each field trial.	98%	99.99%

During the trials the apps reported 1.75 crashes, and 11.8 non-fatal errors per user per month. Assuming an average restart time of 1 minute, the total availability of the apps was 99.99%.

EF2-2 - Service Continuity MQTT

Crashes of the exchange service were registered by a crash monitoring tool.

Table 29 Details of KPI EF2-2.

#	KPI	Definition	Target values	Trial result
EF2-2	Service Continuity	Percentage of time during which MQTT service was operational	98%	100%

During the trials no crashes were reported (100% uptime).

EF2-3 - Network Awareness

The Reference Signals Received Power (RSRP) value reported by the handset OS was registered and stored per exchanged message, together with information on the operational setting, modality, usage scenario, network used per trial run.

In cellular networks, a mobile device continuously measures the signal strength/quality of nearby cells to support cell selection/reselection and handover in the network. The RSRP is a key measure to express the signal level, hence it is indicative of how well the device is connected to the network. A radio connection with an RSRP of -105 dBm or higher (i.e., smaller negative number) is considered sufficient, an RSRP between -120 and -105 dBm is considered fair. RSRP values of the 5G SA network were benchmarked against 4G using Kruskal-Wallis, Wilcoxon and Dunn's Test for the different modalities, operational settings and usage scenarios. In total 299,670 messages were exchanged. These tests can be used for non-normal distributed data sets as is the case with the collected data.

The table below presents the overall KPI. The second table lists the results per usage scenario.

Table 30 Details of KPI EF2-3.

#	KPI	Definition	Target values	Trial result
EF2-3	Service Continuity	Network awareness: % of times the radio connection was reliable enough to timely warn VRUs	RSRP > -105 dBm in 98% of time	98.01%

Table 31 RSRP (dBm) statistics per operational setting, modality and usage scenario.

Operational setting	Modality	Usage scenario	5G SA URLLC RSRP				4G RSRP			
			1st Qrt	Median	Mean	3rd Qrt	1st Qrt	Median	Mean	3rd Qrt
Industrial	Pedestrian	Ideal	-79	-71	-72,6	-65	-87	-81	-81,09	-75
Industrial	Cyclist	Ideal	-80	-74	-74,2	-68	-82	-78	-77,35	-72
Industrial	Pedestrian	Realistic	-90	-90	-81,8	-72	-86	-86	-82,84	-79
Industrial	Cyclist	Realistic	-86	-76	-78,0	-69	-89	-80	-81,96	-76
Urban	Pedestrian	Ideal	-97	-92	-91,5	-86	-	-	-	-
Urban	Cyclist	Ideal	-86	-80	-80,4	-75	-90	-85	-84,47	-79
Urban	Pedestrian	Realistic	-102	-102	-102,0	-102	-96	-96	-95,77	-96
Urban	Cyclist	Realistic	-98	-84	-85,8	-76	-93	-91	-87,13	-82

Table 31 shows that:

- 5G in the trial areas in general provided a good to excellent radio connection, with median and average RSRP around or above -90.
- Overall, 4G signal offers better signal quality than 5G, as the difference in their performance is statistically significant (further elaborated below).
- The radio connection appears to be slightly better for both 5G and 4G in the more open industrial setting than in the denser built-up urban setting (further elaborated below).
- Pedestrians appear to experience a slightly less good radio connection on 5G than on 4G in the realistic usage scenarios, when they carry the handset in a rucksack or in their pocket, with RSRP values dropping to around -100 dBm, versus around 96 dBm for 4G.

Statistical tests were carried out on all records to determine if RSRP overall differs between 5G and 4G (n=299,670). Table 32 lists general statistics and Figure 42 shows the corresponding histograms. Though the overall statistics suggest comparable performance, the histogram shows 5G and 4G have deviating distributions. While the RSRP of 4G appears to be clustered around specific values, the 5G histogram shows a more gradual distribution. The Kruskal-Wallis (p-value of 2.2x10-16) and Wilcoxon (p-value of 2x10-16) tests also confirms a significant overall difference in RSRP between 5G and 4G, in favor of 4G.

Table 32 Comparison of RSRP (dBm) between 5G and 4G overall.

5G SA URLLC RSRP				4G RSRP			
1st Qrt	Median	Mean	3rd Qrt	1st Qrt	Median	Mean	3rd Qrt
-93	-85	-85,4	-77	-98	-88	-88,5	-81

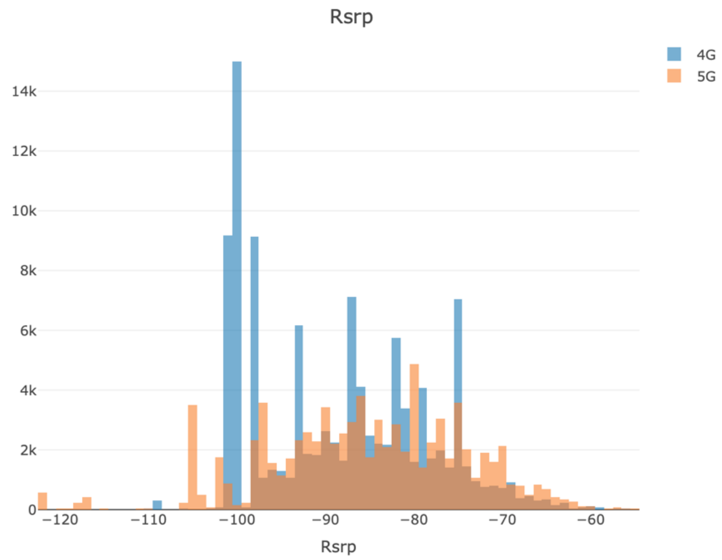


Figure 42 Histogram of RSRP (dBm) of 4G (blue) versus 5G (orange) for all operational settings and all usage scenarios.

Table 33 presents RSRP statistics per operational setting, regardless of modality and usage scenario (n=262,771). Figure 43 presents histograms of 5G and 4G RSRP for both operational settings. It shows that both networks exhibit lower RSRP values, hence poorer radio connection, in an urban setting. However, the degradation for 5G (median -9%, mean -8%) is lower than for 4G (median -12%, mean -11%) showing a robust performance by 5G.

Table 33 Comparison of RSRP (dBm) of 5G SA and 4G per operational setting.

Operational setting	5G SA URLLC RSRP				4G RSRP			
	1st Qrt	Median	Mean	3rd Qrt	1st Qrt	Median	Mean	3rd Qrt
Industrial	-90	-80	-81,1	-72	-88	-83	-82,8	-77
Urban	-95	-87	-87,7	-80	-100	-93	-91,6	-85

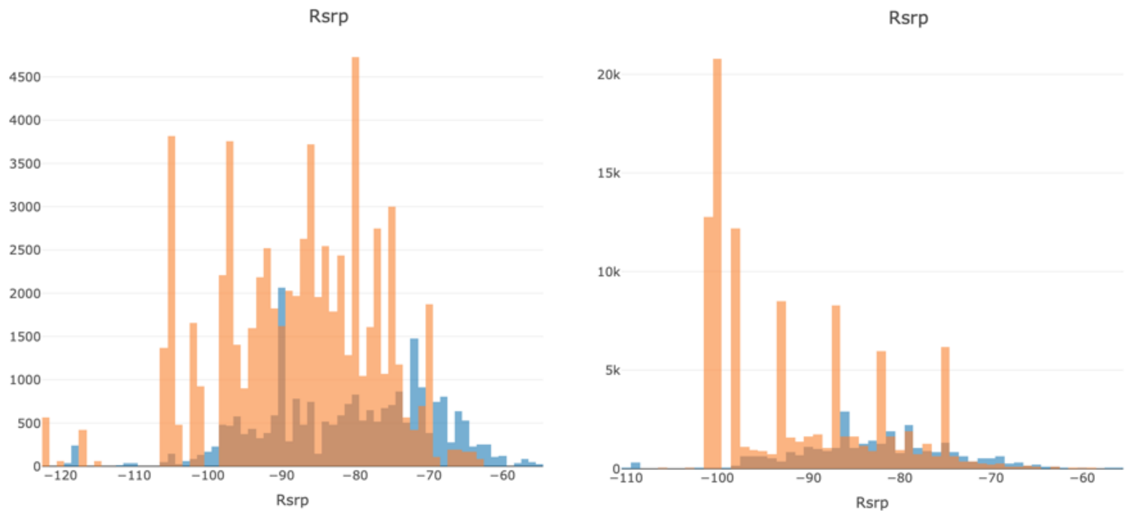


Figure 43 Histograms of 5G (left) and 4G (right) RSRP (dBm) for industrial area (in blue) and urban area (orange).

Table 34 presents statistics for 5G and 4G RSRP per usage scenario. The histogram in Figure 44 presents the diagram for 5G. The statistical indicators suggest comparable results for 5G and 4G in both usage scenarios; both show a slightly lower performance for the realistic usage scenario. The histograms here too show different distributions for the two usage scenarios.

Table 34 Comparison of RSRP (dBm) for 5G and 4G under ideal and realistic conditions.

Operational setting	5G SA URLLC RSRP				4G RSRP			
	1st Qrt	Median	Mean	3rd Qrt	1st Qrt	Median	Mean	3rd Qrt
Ideal	-92	-82	-83,0	-74	-93	-86	-86,2	-79
Realistic	-97	-89	-87,2	-76	-96	-90	-89,6	-85

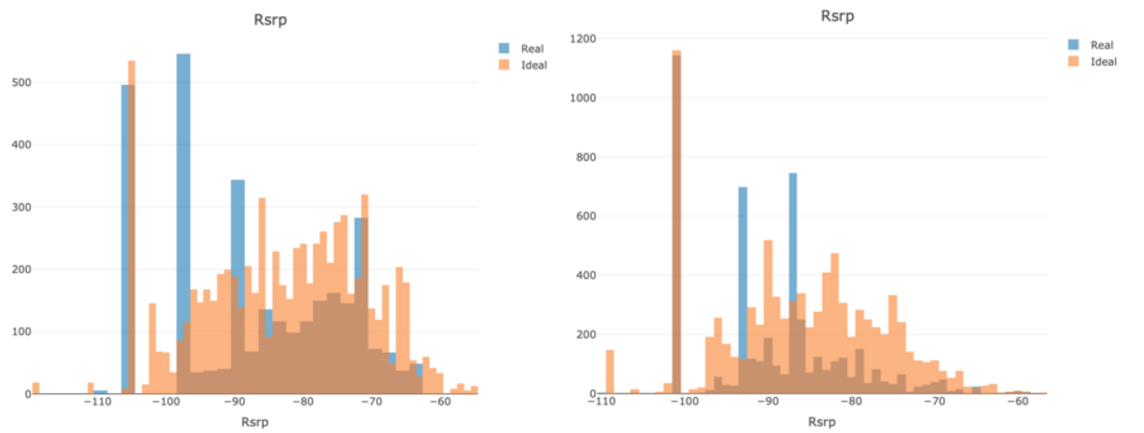


Figure 44 Histograms of 5G RSRP (left) and 4G RSRP (right) for realistic (blue) and ideal (orange) usage scenarios.

EF2-4 - Reliability

Reliability was measured by relaying messages of apps back to the originating devices and registering the number of messages that arrived.

The service requires that 98% of all sent messages per hour are delivered, which corresponds to 2.4 seconds of the warning horizon of 2 minutes (derived in D6.2, section 2.2.1). Reliability of the 5G SA network was benchmarked against 4G using Kruskal-Wallis, Wilcoxon and Dunn's Test for the different modalities, operational settings and usage scenarios. In total data from 22,957

VAM messages were exchanged and analyzed.

Table 35 Details of KPI EF2-4.

#	KPI	Definition	Target values	Trial result
EF2-4	Reliability	Number of messages made available via MQTT Broker with position of VRU, and potentially warning, per hour	98% of 34,200 messages	98%

Table 36 Reliability statistics per operational setting, modality and usage scenario. Due to a lack of handsets the combination of urban – ideal - pedestrian could not be trialed for 5G.

Operational setting	Modality	Usage scenario	5G SA URLLC Reliability				4G Reliability			
			1st Qrt	Median	Mean	3rd Qrt	1st Qrt	Median	Mean	3rd Qrt
Industrial	Pedestrian	Ideal	100%	100%	99,7%	100%	100%	100%	99,8%	100%
Industrial	Cyclist	Ideal	100%	100%	99,9%	100%	100%	100%	99,9%	100%
Industrial	Pedestrian	Realistic	100%	100%	99,8%	100%	100%	100%	99,8%	100%
Industrial	Cyclist	Realistic	100%	100%	98,7%	100%	100%	100%	100,0%	100%
Urban	Pedestrian	Ideal	100%	100%	99,3%	100%				
Urban	Cyclist	Ideal	100%	100%	96,4%	100%	100%	100%	99,8%	100%
Urban	Pedestrian	Realistic	100%	100%	97,9%	100%	100%	100%	99,7%	100%
Urban	Cyclist	Realistic	100%	100%	95,5%	100%	100%	100%	99,6%	100%

Table 36 shows that:

- In the urban environment, despite fair RSRP values, on average too many messages are lost on 5G.
- The 4G network appears more reliable in delivering messages to and from the handsets.

To further elaborate on these findings additional statistical analyses were carried out. Statistical tests were carried out on all records to determine if reliability overall differs between 5G and 4G (n=299,670). Table 37 lists general statistics on reliability. The Kruskal-Wallis (p-value of 2.2×10^{-16}) and Wilcoxon (p-value of 2×10^{-16}) tests confirm a significant overall difference in reliability between 5G and 4G in favor of 4G.

Table 37 Comparison of reliability (%) between 5G and 4G overall.

5G SA URLLC Reliability				4G Reliability			
1st Qrt	Median	Mean	3rd Qrt	1st Qrt	Median	Mean	3rd Qrt
100%	100%	98,0%	100%	100%	100%	99,6%	100%

Table 38 presents reliability statistics per operational setting, regardless of modality and usage scenario (n=262,771). It shows that only in the urban setting too many messages are lost on 5G. This is not surprising considering the urban trials were carried out using the 3.5GHz band in a test site with only one 5G node, whereas the 4G reference network has a dense configuration in this urban area. In the industrial setting, where multiple 5G nodes are available, the 5G network does provide the required reliability. Hence, it is likely that a close-knit 5G network will provide the reliability that is required.

Table 38 Comparison of reliability (%) of 5G SA and 4G per operational setting.

Operational setting	5G SA URLLC Reliability				4G Reliability			
	1st Qrt	Median	Mean	3rd Qrt	1st Qrt	Median	Mean	3rd Qrt
Industrial	100%	100%	99,0%	100%	100%	100%	99,6%	100%
Urban	100%	100%	96,8%	100%	100%	100%	99,6%	100%

Table 39 presents statistics for 5G and 4G reliability per usage scenario. It shows that only in the realistic usage scenario too many messages are lost. This could be partially remedied by deploying the service only in VRU vehicles where the placement of modem and antenna is fixed, and designed and tested diligently.

Table 39 Comparison of reliability (%) for 5G and 4G under ideal and realistic conditions.

Operational :	5G SA URLLC Reliability				4G Reliability			
	1st Qrt	Median	Mean	3rd Qrt	1st Qrt	Median	Mean	3rd Qrt
Ideal	100%	100%	98,4%	100%	100%	100%	99,9%	100%
Realistic	100%	100%	97,9%	100%	100%	100%	99,7%	100%

EF2-5 - Latency

Latency was measured by relaying messages of apps back to the originating devices and registering the roundtrip time. These roundtrip times were stored in the central data store. The service requires a maximum latency of 100ms one-way (D6.3 [4]), hence a maximum roundtrip time of 200ms. Please note that all values indicated in this, and the next section are roundtrip times in milliseconds.

Roundtrip times of the 5G SA network were benchmarked against 4G using Kruskal-Wallis, Wilcoxon and Dunn's Test for the different modalities, operational settings and usage scenarios. In total data from 22 957 messages were analyzed.

Table 40 Details of KPI EF2-5.

#	KPI	Definition	Target values	Trial result
EF2-5	Latency	Roundtrip time for messages sent to the MQTT server	<200ms	137ms median, 215ms average

Table 41 Roundtrip time (ms) statistics per operational setting, modality and usage scenario. Due to a lack of handsets the combination of urban – ideal - pedestrian could not be trialed for 4G.

Operational setting	Modality	Usage scenario	5G SA URLLC roundtrip time				4G roundtrip time			
			1st Qrt	Median	Mean	3rd Qrt	1st Qrt	Median	Mean	3rd Qrt
Industrial	Pedestrian	Ideal	98	142	153,5	174	158	172	179,6	206
Industrial	Cyclist	Ideal	94	106	135,0	131	171	205	199,5	218
Industrial	Pedestrian	Realistic	103	125	155,6	155	151	198	185,0	210
Industrial	Cyclist	Realistic	102	125	136,3	139	167	187	185,8	205
Urban	Pedestrian	Ideal	118	143	253,4	193				
Urban	Cyclist	Ideal	118	143	273,8	260	247	369	383,0	448
Urban	Pedestrian	Realistic	109	123	177,2	153	270	388	410,5	471
Urban	Cyclist	Realistic	120	140	246,6	233	264	378	403,9	454

Table 41 shows that:

- 5G clearly outperforms 4G in overall roundtrip time
- 5G exhibits a wider spread in roundtrip time than 4G
- 5G average is much higher than median roundtrip time as a result of outliers
- 5G meets the requirement in the industrial setting
- The median values for 5G indicate that 5G in general does meet the requirement, but that outliers raise the average values too much to be acceptable
- There are only minor differences in roundtrip time between the different usage scenarios

The values are application level roundtrip times. It should be noted that the exchange service was running in the cloud and hence the speed of the public internet also contributes to these values. Below a drill down of the contribution of each link in the data communication chain is carried out.

To further elaborate on the findings additional statistical analyses were carried out. Statistical tests were carried out on all records to determine if latencies overall differ between 5G and 4G (n=299,670). Table 42 lists general statistics and Figure 45 shows the corresponding histograms. The Kruskall-Wallis (p-value of $2.2x10^{-16}$) and Wilcoxon (p-value of $2x10^{-16}$) tests confirm a significant overall difference in roundtrip time between 5G and 4G.

Table 42 Comparison of roundtrip time (ms) between 5G and 4G overall.

5G SA URLLC roundtrip time				4G roundtrip time			
1st Qrt	Median	Mean	3rd Qrt	1st Qrt	Median	Mean	3rd Qrt
114	137	215,0	184	200	308	353,3	436

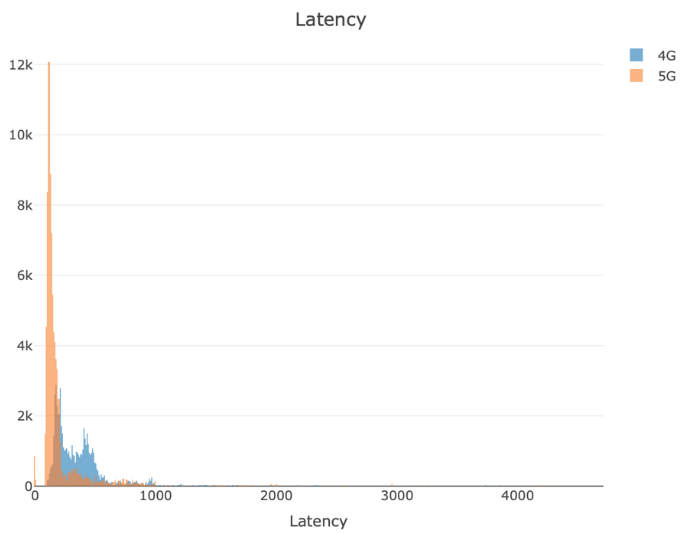


Figure 45 Histogram of roundtrip time (ms) of 4G (blue) versus 5G (orange) for all operational settings and all usage scenarios.

Table 43 presents roundtrip time statistics per operational setting, regardless of modality and usage scenario (n=262,771). Figure 46 presents histograms of 5G and 4G roundtrip time for both operational settings. It shows that both networks exhibit higher latencies in an urban setting. This is confirmed by the outcomes of the Kruskal-Wallis (p-value of $2.2x10^{-16}$) and Wilcoxon (p-value of $2x10^{-16}$) tests.

The bottom row of Table 44 indicates the difference per statistical value, of the urban setting versus the industrial setting. Interestingly, 5G performs more consistent than 4G, as the difference in roundtrip time between the two operational settings is much smaller for 5G than for 4G.

The histograms also illustrate that that roundtrip time is not normally distributed and that high outlier values affect the roundtrip time average.

Table 43 Comparison of roundtrip time (ms) of 5G SA and 4G per operational setting.

Operational setting	5G SA URLLC roundtrip time				4G roundtrip time			
	1st Qrt	Median	Mean	3rd Qrt	1st Qrt	Median	Mean	3rd Qrt
Industrial	102	119	166,5	148	171	197	233,4	232
Urban	119	148	241,5	195	255	377	402,7	462
Difference	-16,7%	-24,4%	-45,0%	-31,8%	-49,1%	-91,4%	-72,5%	-99,1%

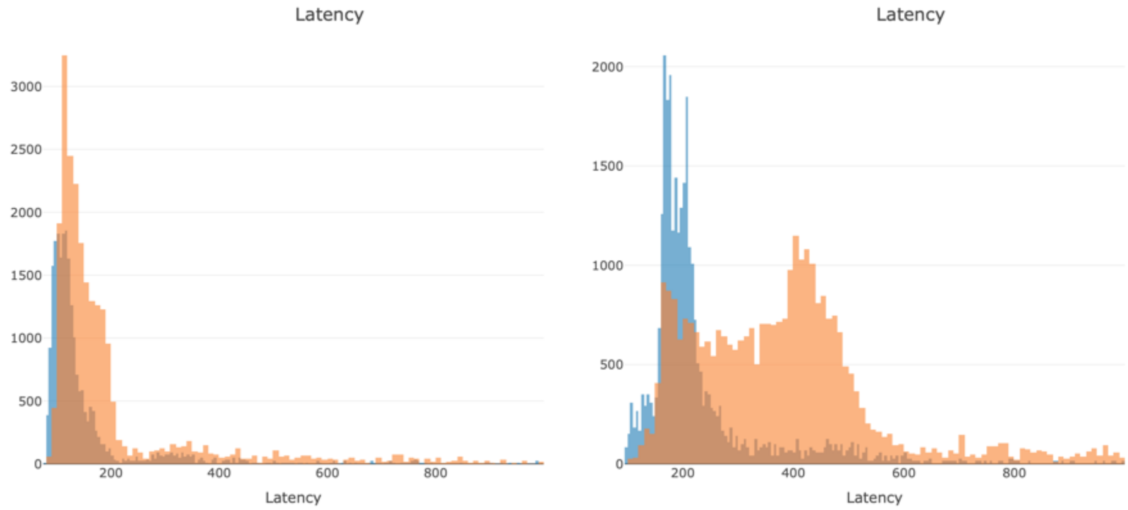


Figure 46 Histograms of 5G SA (left) and 4G (right) roundtrip time (ms) for industrial area (in blue) and urban area (orange).

Table 44 presents statistics for 5G and 4G roundtrip time per usage scenario. The histogram in Figure 47 presents the diagram for 5G. Again, 5G clearly outperforms 4G in terms of roundtrip time in both usage scenarios.

Here too, 5G appears more robust than 4G. Average roundtrip time in realistic conditions only slightly decreases compared to the roundtrip time in ideal conditions on 5G (-5%), but quite a lot on 4G (-25%).

Table 44 Comparison of roundtrip time for 5G and 4G under ideal and realistic conditions.

Usage conditions	5G SA URLLC roundtrip time				4G roundtrip time			
	1st Qrt	Median	Mean	3rd Qrt	1st Qrt	Median	Mean	3rd Qrt
Ideal	109	122	195,6	162	175	228	300,8	398
Realistic	112	132	205,1	169	216	350	374,6	445
Difference	-2,8%	-8,2%	-4,9%	-4,3%	-23,4%	-53,5%	-24,5%	-11,8%

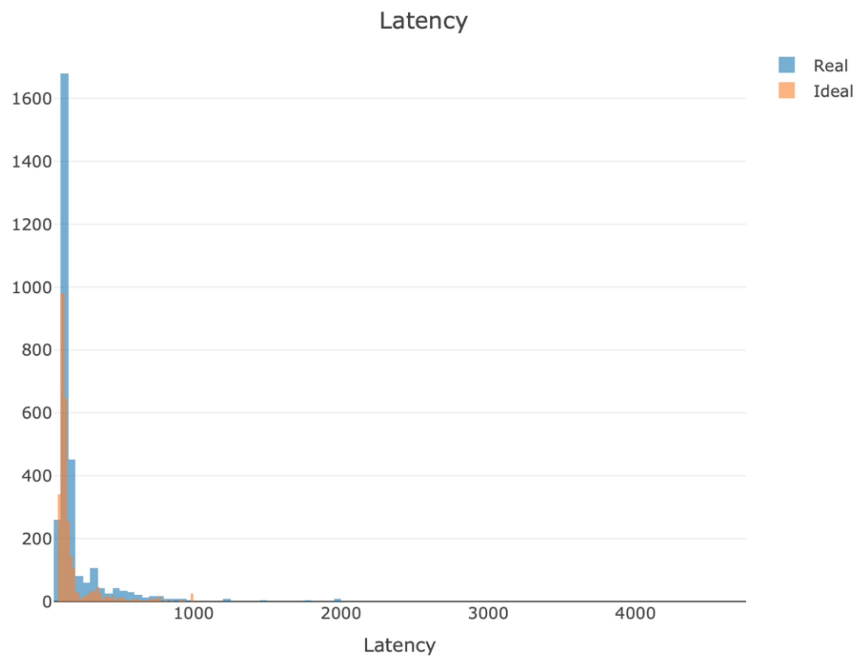


Figure 47 Histogram of 5G roundtrip time (ms) for realistic and (blue) and ideal usage scenarios.

As indicated all values presented in this section on KPI EF2-5 are roundtrip times. The median of the roundtrip time for 5G data communication recorded in the trials in the industrial setting was 119ms. To better understand where in the data communication chain latency is incurred, additional tests were carried out.

- Tests were carried out in the Locatienet back-office to determine the net processing time of the VectorDrive App and Exchange. This was done by measuring the roundtrip time from the app, to the exchange and back to the app, on a local network connection in the same way as for 4G and 5G (Table 45), hence this is an indication for the total processing time of the VectorDrive app and Exchange. Assuming up and downlink symmetry this would amount to a median one-way latency of 3.5ms.
- A ping test revealed that the median roundtrip time of the connection between the Locatienet and Telenet backend is 21.815ms. Assuming up and downlink symmetry this would amount to a median one-way latency of 10.9ms.
- In D5.4, median latency values of 14ms were reported for one-way uplink 5G SA URLLC in the Antwerp Port Area.

It should be noted that the underlying data were captured at different times, and that these were captured under optimal circumstances with optimal equipment. Also, the first two values should be considered approximations because of the underlying assumptions.

Table 45 presents these values In relation to the median latency measured on 5G in the port area under the ideal usage scenario (handset on bicycle handlebar: 106ms), ~14% of the overall end-to-end latency in the communicated chain is associated with radio network impact, ~3.5% to VectorDrive application processing, and the rest to the impact of public internet.

Table 45 Roundtrip time (ms) of backend processing of messages.

VectorDrive roundtrip time			
1st Qrt	Median	Mean	3rd Qrt
5	7	9,7	10

Conclusions:

- The VectorDrive app and exchange performed well, showing very few app crashes and no server downtime.
- 5G in the trial areas in general provided a good to excellent radio connection, performing slightly worse than 4G in terms of RSRP.
- The radio connection is slightly inhibited in dense built-up areas and in the realistic usage scenario. However, this degradation is lower than for 4G.
- In the urban setting 5G latency incidentally exceeds the latency threshold and misses too many messages, which can also be caused by the fact that the urban site hosted only one 5G node whereas multiple cells of the 4G network were used. To deploy safety critical cooperative services in an urban environment smaller cells need to be deployed to achieve adequate coverage.
- In settings that provide sufficient line-of-sight, 5G provides a reliable connection for the exchange of cooperative awareness messages with a very high frequency and low latency. This means VectorDrive could be used as a safety critical solution to avoid collisions in industrials and port yards.

4.5 Testing of 5G-connected Intelligent traffic lights

4.5.1 Overall testing procedure

The scope of the intelligent Traffic Light Controller (iTLC) or EF3 is twofold:

1. Test the conflictless crossing of intersections of teleoperated vehicles by providing a time slot for 'green-lighted passages' which will reduce the likelihood of collisions and ensure smooth navigation of the intersection for truck platoons using both the standard installed communication (e.g., 4G production) and 5G.
2. Validate if 5G can replace copper/fibre for iTLC uplink (i.e., fixed wireless access for critical infra) on the Vlissingen test site and can replace 4G at the Zelzate test site.

Due to restrictions in time and effort of extensive testing with truck platooning, the quality and stability of the communication with the iTLC is also tested over a longer period running in normal controller mode using the installed base in Vlissingen pilot site, using the 5G setup that is available on the premises.

For the duration test, the communications from and towards the iTLCs were logged, including the following C-ITS messages: Signal Phase and Timing Messages (SPAT), Cooperative Awareness Message (CAM), Signal Request Message (SREM), Signal Status Message (SSEM) and MAPem (topology of the infrastructure of the junction).

The Traffic Light Controller along the roadside, next to all the hardware on the junction, exists of TLC, ITSapp and RISmon (see explanation below) which communicate with the outside world through the RIS to the UDAP/Tlex environment. UDAP/Tlex is a centralized application in Netherlands/Belgium where all C-ITS communication is routed through. At the other side service providers and In car systems can connect to UDAP/Tlex to send and receive C-ITS messages towards the Traffic Light Controller. EF7 provider sends a time slot request through this communication path towards UDAP.

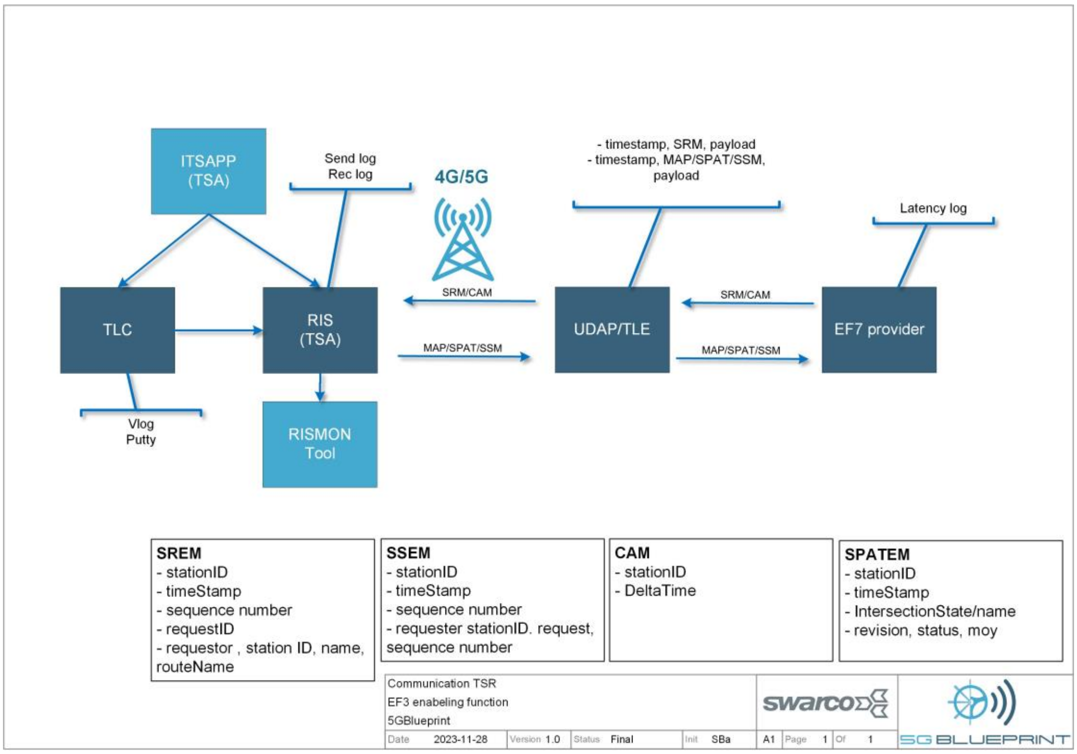


Figure 48 Test environment for EF3.

Acronyms used in Figure 48:

- TLC: Traffic Light Controller (hardware)

- ITSapp: Software for the green management of the Traffic Light Controller (software)
- RIS(TSA): Roadsite ITS station responsible for the C-ITS communication on the roadside
- RISMon: Monitoring tool enabling logging
- UDAP/Tlex: Urban Data Exchange platform (Netherlands), Traffic Live Exchange (Flanders)
- EF7 provider = Service provider for enabling TOV to connect to the C-ITS infrastructure

Duration test 4G and 5G

EF3 Time slot reservation at intersection enabling service receives a timeslot reservation request in combination with platooning information from EF7 (Estimated Time of Arrival) service tested with the 4G and 5G communication from the iTLC towards UDAP/TLE.

The Time Slot Reservation (TSR) request is sent out using the SREM message which includes vehicle ID, time slot, inbound lane and signal group. SREM message is sent with a maximum of 5 minutes before reaching the intersection but at least from the moment the first vehicle of the platoon enters the MAP specified by the MAPem topology. For this purpose, the MAP is extended from the normal 300 meters to a maximum of 1000 meters.

The iTLC will inform the vehicles of the platoon about the request status using the SSM message [**Prioritization-ResponseStatus**]. Also during the TSR request, the iTLC can propose a different time slot (never earlier than the requested time slot from the platoon).

When a vehicle of the platoon enters the MAP, CAM and SPAT messages are also exchanged. When all vehicles of the platoon are on the MAP and the requested signal group turns green within the requested timeframe of the negotiated TSR, the iTLC will send the status granted to the vehicles of the platoon and will keep the signal group green until the last vehicle passes the stop line.

Specifications will be followed from ‘Priority services for target groups’, version 1.1.0[3] Use Case: Priority for convoy, chapter 7. And the ‘Functional specification handling Time Slot Reservation (SRM SSM)’ version 0.4.

The KPIs for EF3 are shown in Table 46.

Table 46 EF3 Preconditions and relevant KPIs.

#	KPI	Definition	Target values	Measurement
EF3.1	Latency	To compare performance over 4G vs 5G we need to know the difference between both channels, we measure the time between the iTLC and the TOV.	<200 msec	*Due to separate testing of EF3, the total loop time is less relevant. Comparison between 4G and 5G is only tested*

[3] CROW. (2021). *Priority services for target groups: Functional description for Emergency Services, Public Transport and Logistics version 1.1.0* (1.1.0).

		This maximum value is given by the requirement of the government for C-ITS communication		*between iTLC and UDAP*
EF3.2	Latency	to compare performance over 4G vs 5G we need to know the difference between both channels, we measure the time between iTLC and UDAP/TLEX (Brooker)	<100 msec	Comparison of log time UDAP and log time RIS for SREM and SSEM messages
EF3.3	Availability	the traffic engineering application is operational as a cloud service in an online situation	System availability >99,9%	Check on not-received messages in RIS and UDAP (SREM and SSEM)

4.5.2 Results EF3 duration test Vlissingen

The route of the platoon starts from Verbrugge Scaldia Terminals, via Europaweg Oost, Borselssedijk turning at the roundabout on the Bernhardweg West (N666) – Lange Noordweg travelling the same route back to Verbrugge Scaldia Terminals (see Figure 49). Figure 50 and Figure 51 represent both layouts of the crossings named K0436 and K0038.

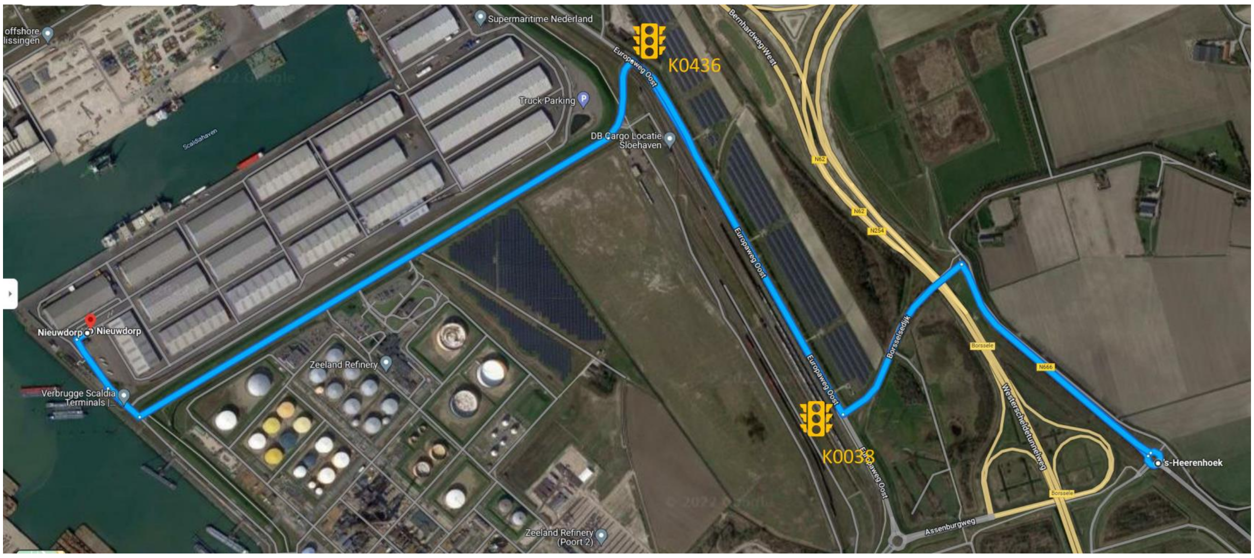


Figure 49 Vlissingen site route for EF3 testing.

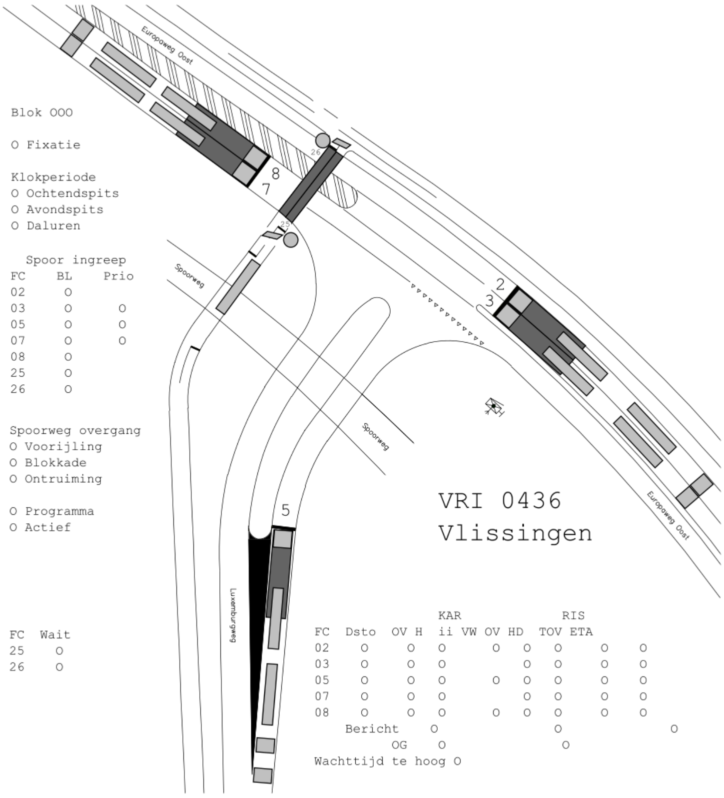


Figure 50 iTLC K0436 Crossing layout Europaweg oost - Limburgweg.

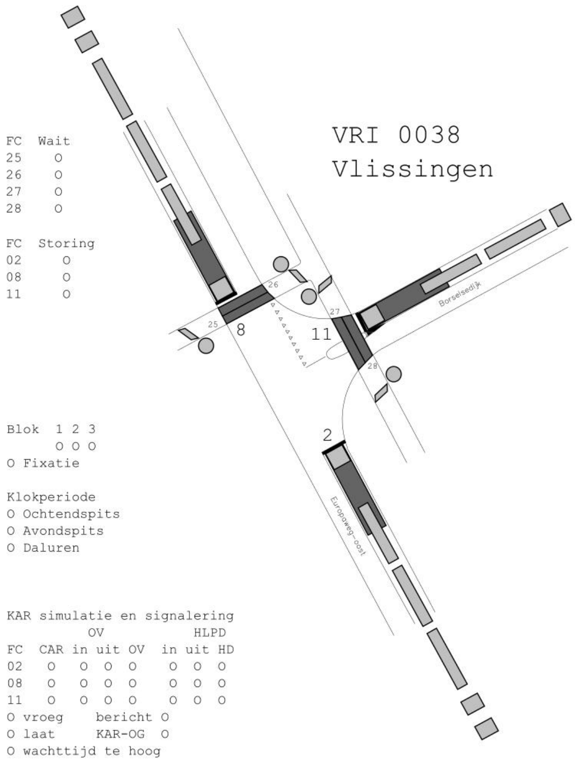


Figure 51 iTLC K0038 Crossing layout Europaweg oost - Bosrelsedijk.

As EF3 is tested separately, the EF7 function (ETA sharing) provided a simulation of platoons every two minutes on average. As the simulations were done on normal operational traffic light controllers, some parameters were changed so that the impact on the traffic flow was minimized.

The duration test was held from 17/04/2023 8:00 hr up until 21/04/2023 0:00hr. The modems started on 5GNSA and were changed to 4G only modus on 19/04/2023 at 17:22 hr. Table 47 and Table 48 show the number of gathered messages per day.

Table 47 Duration test, number of messages junction K0436.

7b36000A - K0436								
	UDAP				RIS			
	SREM	SSEM	SPAT	CAM	SREM	SSEM	SPAT	CAM
17-apr	3047	3212	25937	-	3047	3212	32230	12733
18-apr	4214	4473	29862	-	4214	4473	34774	17740
19-apr	4859	5163	30209	-	4855	5165	34918	20469
20-apr	4857	5161	31031	-	4857	5161	35660	20259
21-apr	4762	5036	29689	-	4762	5036	34496	19793

Table 48 Duration test, number of messages junction K0038.

7b360014 - K0038								
	UDAP				RIS			
	SREM	SSEM	SPAT	CAM	SREM	SSEM	SPAT	CAM
17-apr	1932	1918	17998		1928	1919	18733	3922
18-apr	2697	2680	20499		2689	2681	20501	5532
19-apr	3113	3095	19343		3105	3096	19350	6265
20-apr	3124	3111	19093		3116	3112	19095	6196
21-apr	3080	3063	19639		3070	3063	19643	5981

The log files of UDAP did not contain CAM messages, because these are not logged due to privacy constraints. So CAM messages are not used in the analyses of package loss. There is a slight delta between UDAP and RIS in the total number of sent/received messages regarding SREM and SSEM. The same goes for the number of SPAT messages per day. This will be looked into when analyzing KPI EF3.3, package loss.

Analysis of results per KPI

Latency TOV – RSU (iTLC)

Enabling Function 3 is tested separately from the platooning Use Case 4.3. To enable testing EF7 simulates TOVs that send messages towards UDAP. This part of the communication uses a normal internet connection of which the latency is not of interest to our EF3. So focus lies on the latency over 4G/5G between iTLC and UDAP.

Latency RSU – UDAP

For the communication between RSU and UDAP 4G vs 5GNSA is tested. The latency between

RSU (iTLC) and UDAP is measured based on the date-time stamps of the sent message and the logging stamp on the receiving side. UDAP itself also tracks the latency values per minute in the UDAP dashboard. This is used as a reference, using the original messages the average latency, and standard deviation 95th percentile can be calculated accompanied with the statistical p-values.

During the first analysis of the datasets for K0436, some outliers disrupted the results of 21/04/2023. SREM request Id 232 with sequence number 9 has a delta of 2 hours in the log time of the iTLC and also 12 other messages had doubtful values, the reason was not traceable but could be due to a clock synchronization issue. Based on these a filter was set for both SREM and SSEM messages for both iTLCs on -500ms and +500 ms. The results for K0436 and K0038 are shown in Table 49 and Table 50 to give an impression of the daily profiles.

Table 49 Duration test, results analysis latency per day junction K0436.

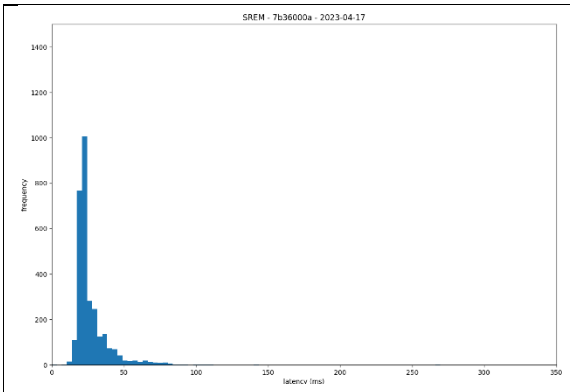
7b36000a - K0436									
SREM	(Msec)								
Date	Min	Max	Mean	Median	Count	Var	std	95%	P-value
17/04/23 5G	-43	267	27.32	22	3044	266.15	16.31	52	0
18/04/23 5G	-35	250	27.1	22	4208	181.44	13.47	53	0
19/04/23 5/4G	-10	296	29.48	23	4850	268.89	16.4	59	0
20/04/23 4G	-1	185	35.04	30	4851	304.77	17.46	67	0
21/04/23 4G	-4	258	37.44	32	4731	371.86	19.28	74	0
SSEM	(Msec)								
Date	Min	Max	Mean	Median	Count	Var	std	95%	P-value
17/04/23 5G	-15	101	23.62	21	3212	134.49	11.6	45	0
18/04/23 5G	-12	126	24.15	21	4473	151.26	12.3	46.4	0
19/04/23 5/4G	-27	125	24.4	21	5162	154.23	12.42	48	0
20/04/23 4G	-5	289	25.72	21	5161	184.07	13.57	50	0
21/04/23 4G	-12	162	26.04	22	5023	175.34	13.24	50	0

Table 50 Duration test, results analysis latency per day junction K0038.

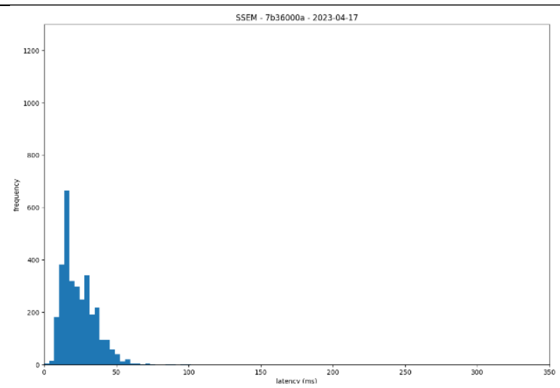
7b360014 - K0038									
SREM	(Msec)								
Date	Min	Max	Mean	Median	Count	Var	std	95%	P-value
17/04/23 5G	3	272	26.17	21	1922	283.34	16.83	52.95	0
18/04/23 5G	-33	279	25.12	20	2677	232.95	15.26	49	0
19/04/23 5/4G	-18	265	28.91	23	3093	290.47	17.04	57	0
20/04/23 4G	6	191	32.44	26	3105	357.49	18.91	65	0
21/04/23 4G	1	179	34.65	28	3048	397.61	19.94	74	0
SSEM	(Msec)								

Date	Min	Max	Mean	Median	Count	Var	std	95%	P-value
17/04/23 5G	-9	132	22.92	18	1918	145.55	12.06	45	0
18/04/23 5G	-3	171	21.57	18	2680	133.31	11.55	44	0
19/04/23 5/4G	-50	261	22.95	18	3092	225.99	15.03	45	0
20/04/23 4G	-7	105	24.92	20	3111	183.72	13.55	51	0
21/04/23 4G	-71	255	24.84	21	3061	234.17	15.3	50	0

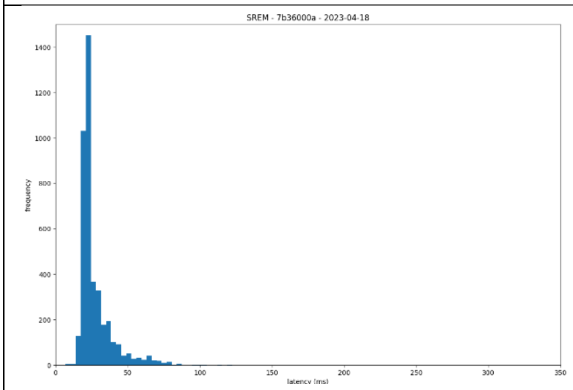
The mean values of the four testing days show a distinct but not big difference between 5G and 4G. The latency of SREM is on average higher than that of SSEM messages. SREM and SSEM messages are both rather small packages (±1250 bits and ±1160 bits). The clock deviation could be the cause but is to say with 100% certainty, see also Figure 56. Between the iTLCs, this is rather consistent. This could be an indication that there is a difference in upload and download speed both for the 5G and 4G connection of the iTLC modem or a shift in time synchronization. The median is smaller than the mean value of all days, meaning that the distribution of the measurements is right-skewed. Figure 52 and Figure 53 show the graphical representation of the latency distribution. P-values were calculated with the Shapiro-Wilk test per day and are nil due to the large datasets.



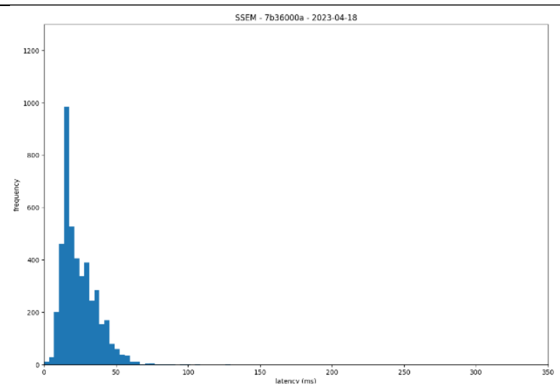
(a)



(b)



(c)



(d)

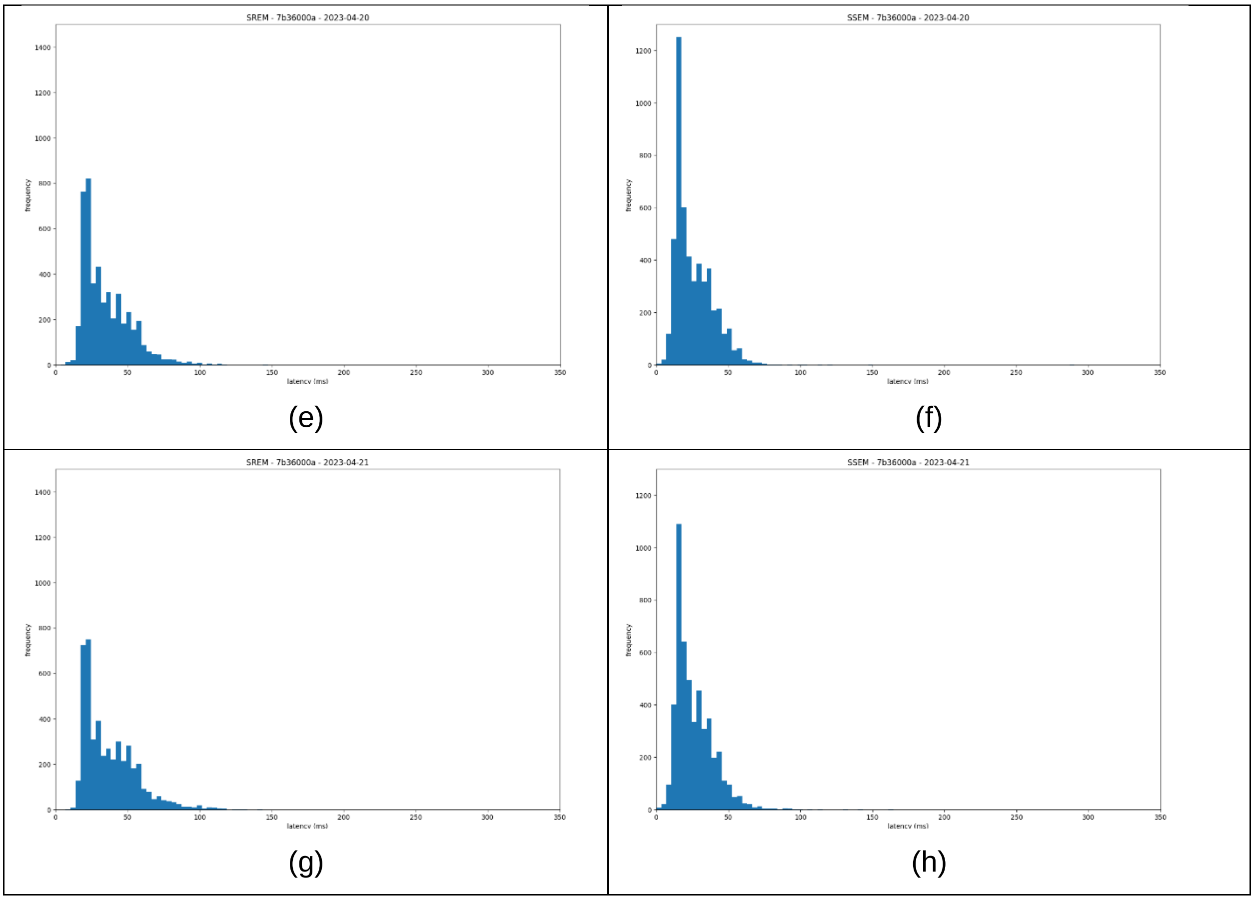
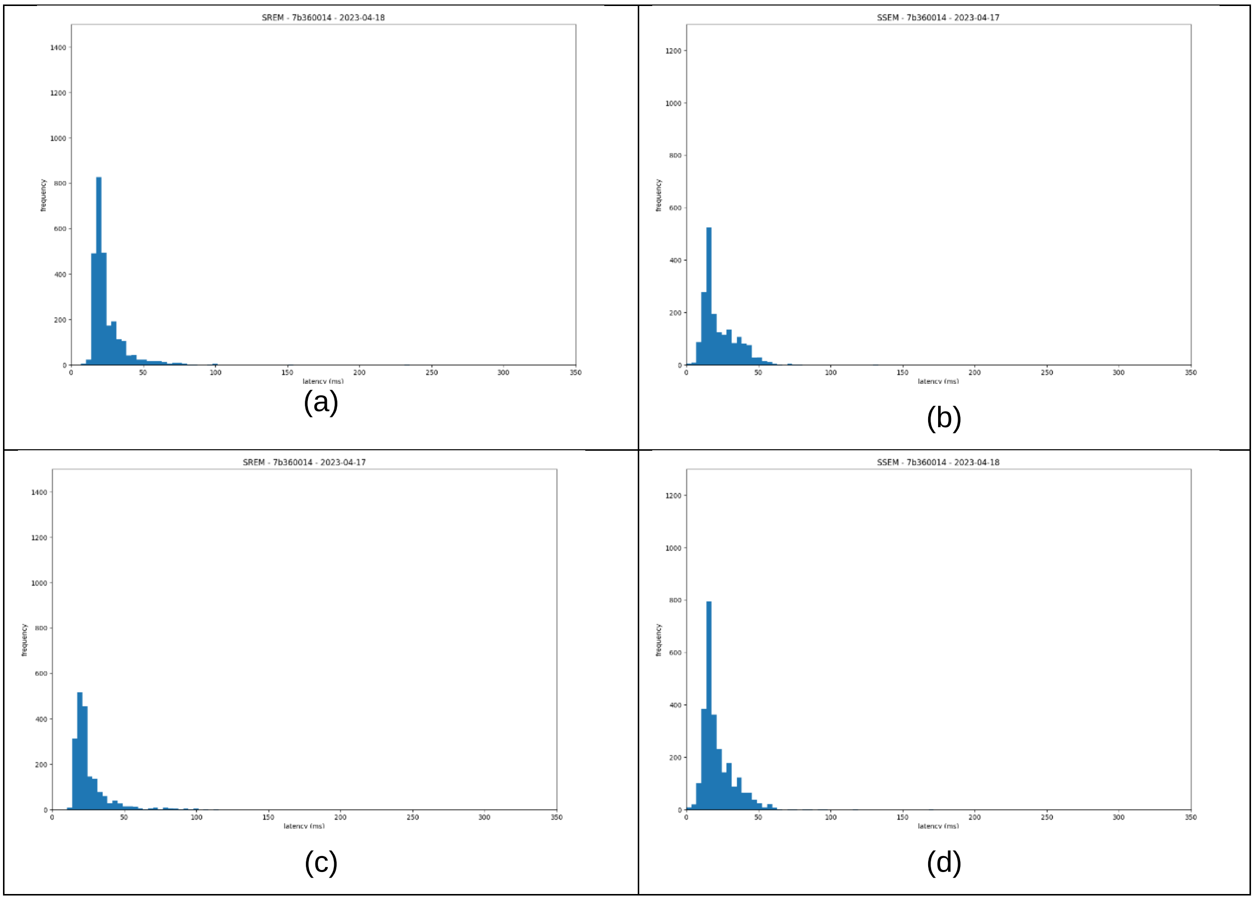


Figure 52 iTLC K0436 bar graphs latency SREM (left) SSEM (right).



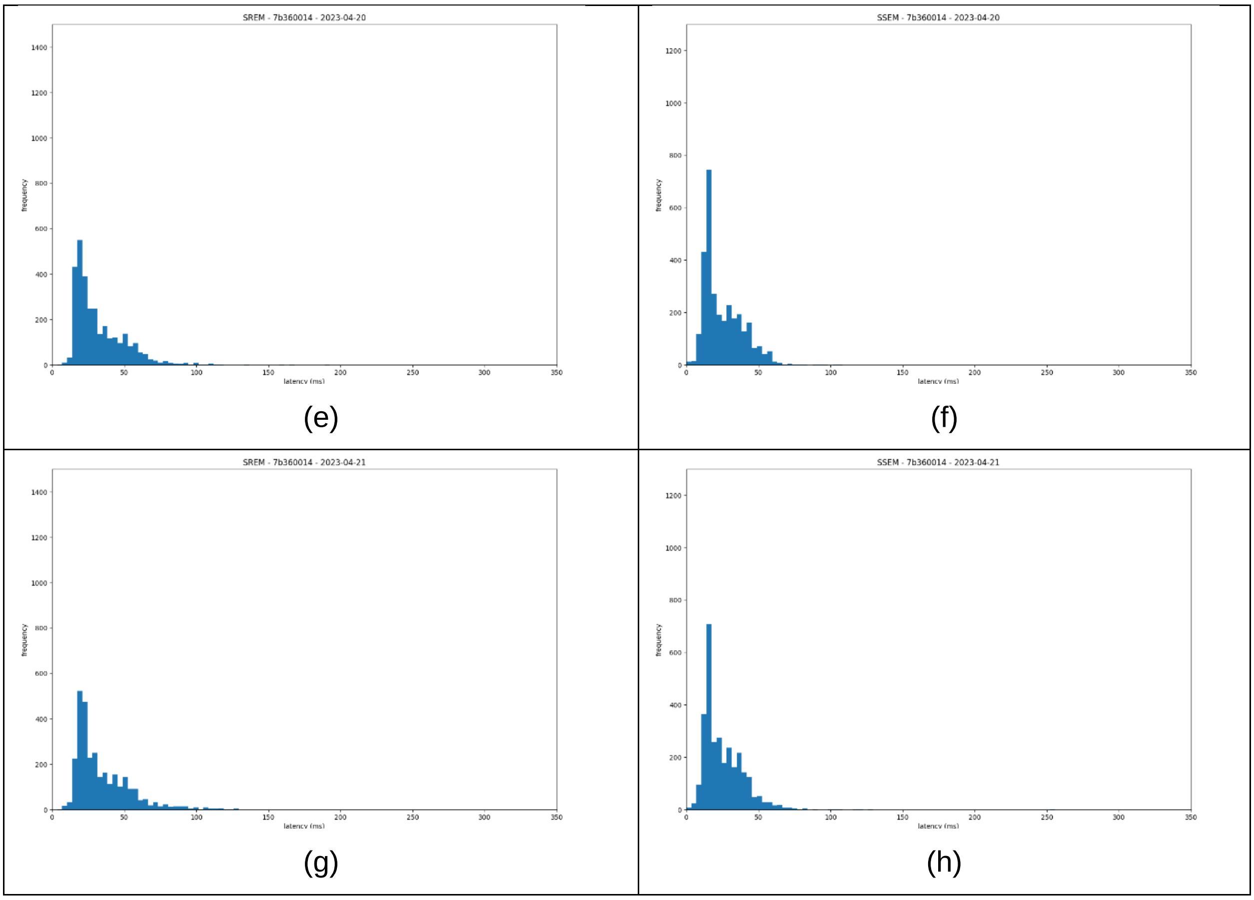


Figure 53 iTLC K0038 bar graphs latency SREM (left) SSEM (right).

For further analysis, the data gathered during 5G and 4G connections are combined into two datapools 17/04/2023 – 19/04/2023, 17:00hr (5GSA) and 19/04/2023, 17:30hr (4G only). Results are presented in Figure 54, Figure 55, Table 51, and

Table 52 (count is the number of measurements).

Table 51 Duration test, results analysis latency K0436 5G/4G.

7b36000a - K0436									
SREM	(ms)								
Date	Min	Max	Mean	Median	Count	Var	std	95%	P-value
5GNSA	-43	296	26.99	22	10646	221.05	14.87	52	0
4G only	-10	258	36.26	31	10942	332.64	18.24	70	0
SSEM	(ms)								
Date	Min	Max	Mean	Median	Count	Var	std	95%	
5GNSA	-27	126	23.62	20	11296	140.26	11.84	45	0
4G only	-12	289	26.14	22	11639	182.24	13.5	51	0

Table 52 Duration test, results analysis latency K0038 5G/4G.

7b360014 - K0038									
SREM	(ms)								
Date	Min	Max	Mean	Median	Count	Var	std	95%	P-value
5GNSA	-33	279	26.01	21	6776	255.05	15.97	50	0
4G only	1	261	33.58	27	7010	375.63	19.38	68	0
SSEM	(ms)								
Date	Min	Max	Mean	Median	Count	Var	std	95%	
5GNSA	-50	171	22.14	18	6771	140.77	11.86	44	0
4G only	-71	255	24.71	20	7033	207.93	14.42	50	0

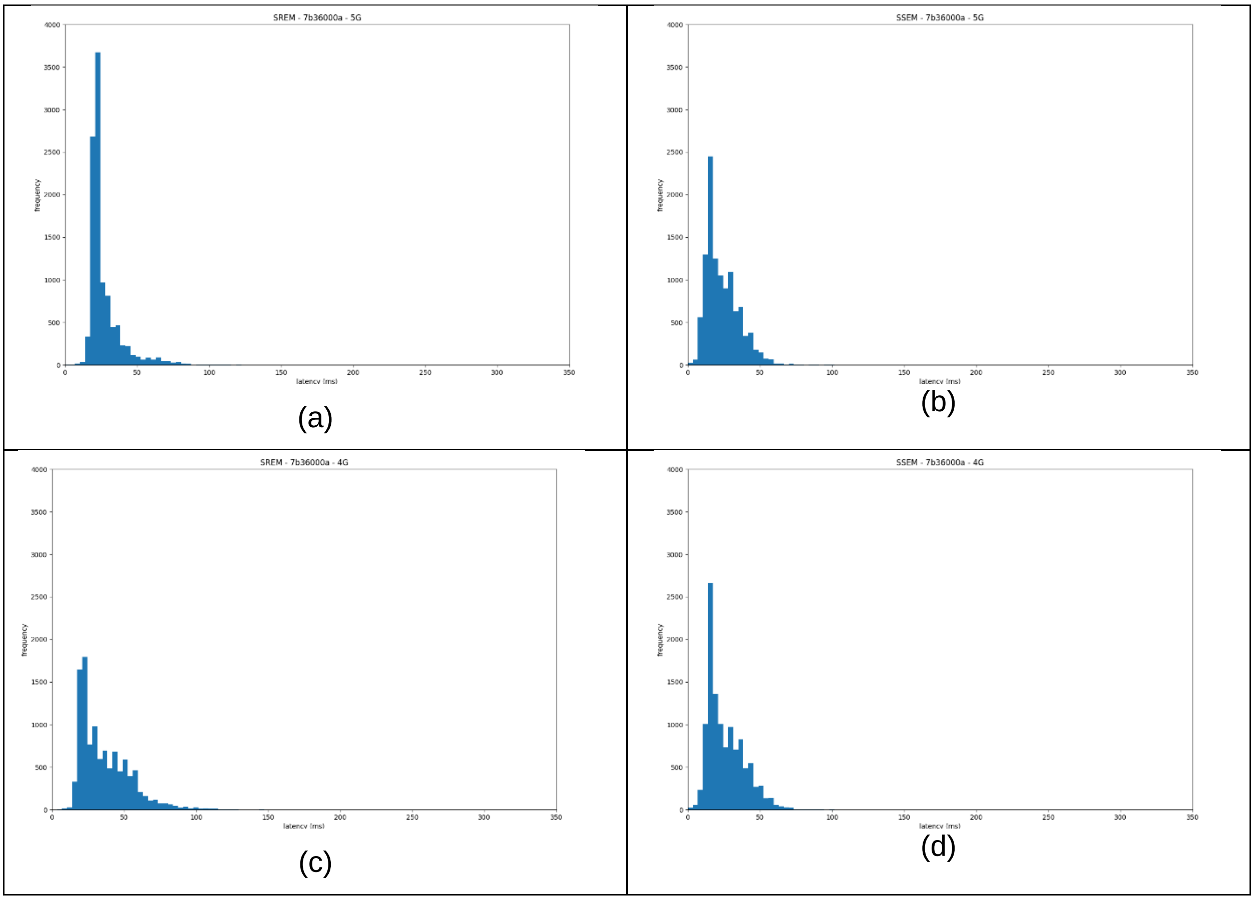


Figure 54 iTLC K0436 bar graphs latency 5G (a,b) /4G (c,d) SREM (left) SSEM (right).

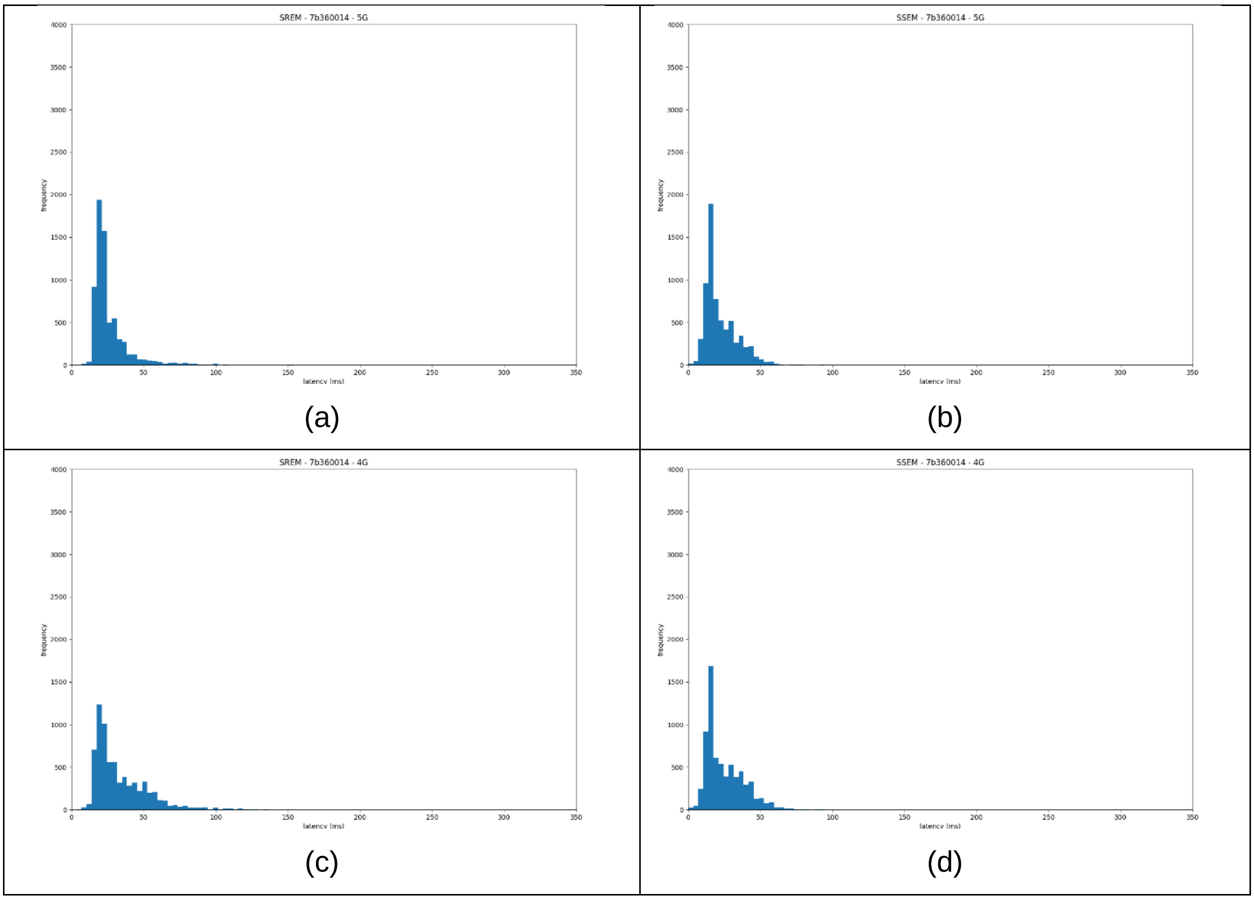
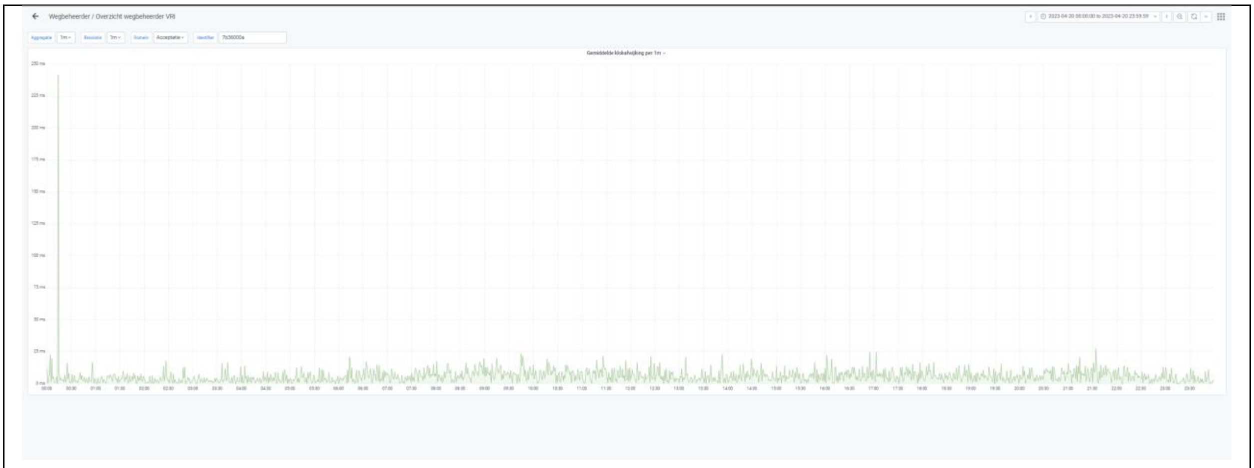


Figure 55 iTLC K0038 bar graphs latency 5G (a,b) /4G (c,d) SREM (left) SSEM (right).

For SREM on average, 5G latency is 26% (K0436) and 23% (K0038) lower. For SSEM on average, 5G latency is 10% lower for both iTLCs. For 4G the scatter (variance and standard deviation) is also a bit higher than for 5G.

The scatter in latency is not directly expected. After comparing with other iTLCs and consulting KPN the scatter in latency for both 5G and 4G are most certainly because the clock synchronization of the iTLCs (minimum of every minute) is done respectively over the 5G and 4G connection. Figure 56 shows the possible correlation between clock deviation and latency of K0436 on 20/04/23 (source UDAP dashboard).



Clock deviation K0436 20/04/23

Latency K0436

Figure 56 iTLC K0436 clock deviation and latency on average per minute 20/04/23 (source UDAP)

The conclusion is that 5GSA shows a lower latency of 10-20% and less scattering of the latency values. Due to the rural area, no problems regarding network traffic are expected.

For EF3 and the C-ITS communication a lower latency means a slightly quicker response and adaptation of the ETA requested (EF7) or proposing a new Time Slot (EF3) and adaptation of the approaching vehicles. This makes the service more accurate.

Reliability (message loss)

Analyzing the dataset used for the latency, a check is done on SREM messages registered in UDAP which were not matched with SREM messages logged in the iTLC and the same for SSEM messages sent by the RIS and not logged in UDAP.

Table 53 Duration test results analysis message loss junction K0436.

7b36000a - K0436				
SREM				
Station_id	Timestamp	Mseconds	Sequence number	Request_id
107789910	19-04-23 17:21	42719	9	115
107789910	19-04-23 17:21	44283	10	115
107789910	19-04-23 17:21	45802	11	115
107789910	19-04-23 17:21	47305	12	115
SSEM				
Station_id	Timestamp	Mseconds	Sequence number	Request_id
107789910	19-04-23 17:21	42692	4	115
107789910	19-04-23 17:21	47964	5	115
92252229	21-04-23 19:26	16197	123	221

On 19/04/23 at 17:22 both modems of the iTLCs were changed from 5G NSA to 4G only modus. This explains the package loss of K0436 around that time. On 21/04/23 there was one connection loss reported in the UDAP dashboard around 15:40 hr which does not explain the missing message with station_id 92252229.

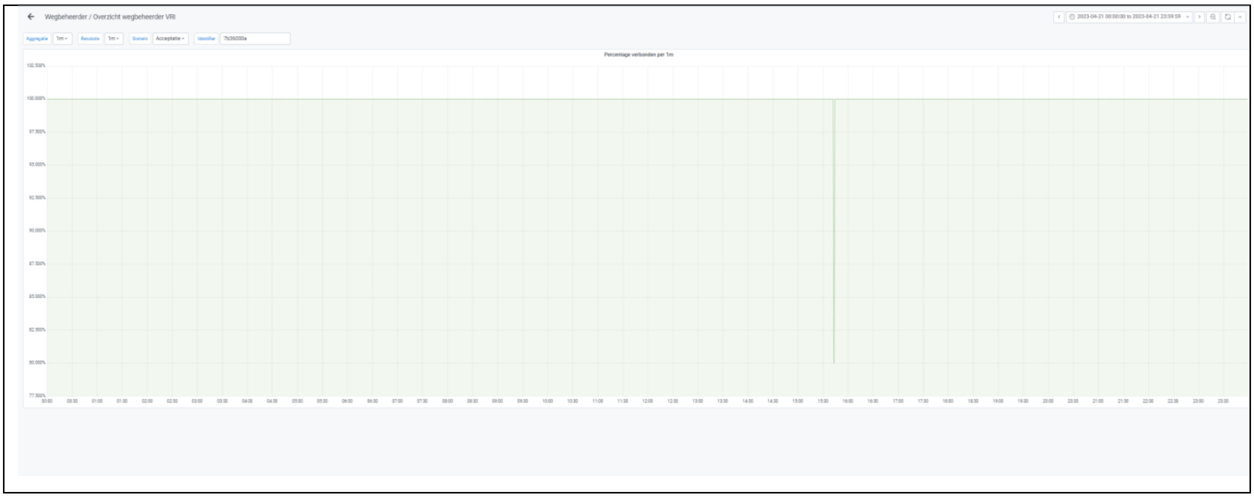


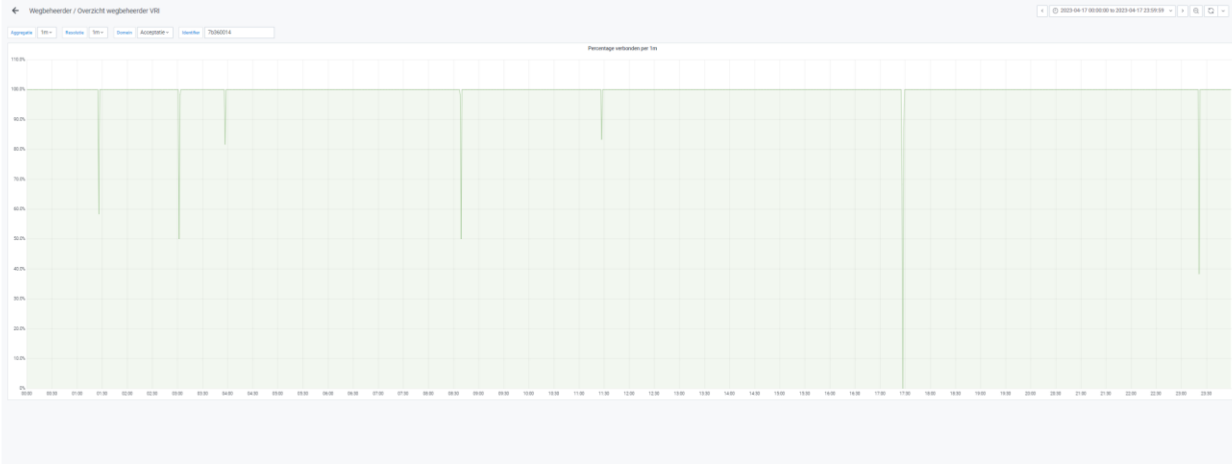
Figure 57 iTLC K0436 percentage connected on 21/04/23(source UDAP).

Table 54 Duration test, results analysis message loss junction K0038.

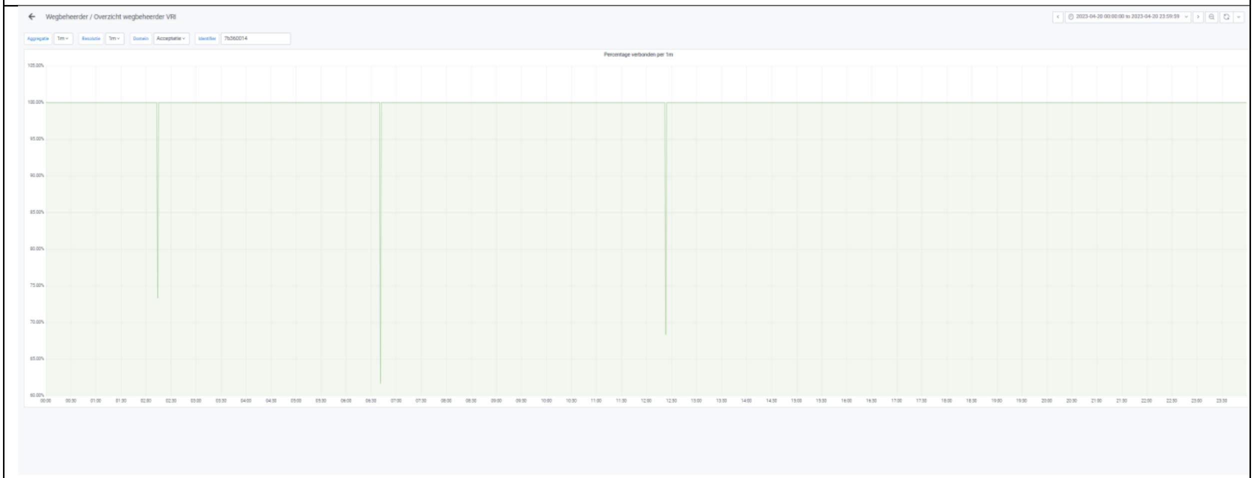
7b360014 - K0038					
SREM	*Start times*				
Station_id	Timestamp	Mseconds	Sequence number	Request_id	Conn.
79255822	17-04-23 17:26	32320	3-6 #4	146	X
90169171	18-04-23 14:49	1740	5-8 #4	242	-
121033431	18-04-23 22:31	45469	4-7 #4	56	-
98842732	19-04-23 02:07	56414	1-4 #4	115	-
123559849	19-04-23 08:35	9701	2-5 #4	219	-
109603300	20-04-23 06:41	36513	4-7 #4	67	X
126860063	20-04-23 12:23	36599	6-9 #4	160	X
88996308	21-04-23 10:47	10704	2-5 #4	89	X
90611357	21-04-23 11:20	15385	2-3 #2	98	X
117098604	21-04-23 19:31	58713	1-4 #4	219	X
SSEM					
Station_id	Timestamp	Mseconds	Sequence number	Request_id	
79255822	17-04-23 17:26	37526	59	146	X
121033431	18-04-23 22:31	50598	85	56	-
123559849	19-04-23 08:35	14332	95	219	-
126860063	20-04-23 12:23	41737	44	160	X
67493423	21-04-23 08:21	10190	102	49	?

For K0038 the number of missing SREMs is higher than for K0436 and also in comparison to the missing SSEMs. In Table 54 the SREMs messages are grouped per station_id with the start and end sequence number and the count of the number of messages. SREM and SSEM messages are sent at intervals and the sequence number is added up when the station (iTLC or a vehicle) in the communication session sents an update. The column Conn. Shows the interpretation if there was a connection loss registered in UDAP. X = ‘connection loss’, - = ‘no information available’, ? -=unknown reason.

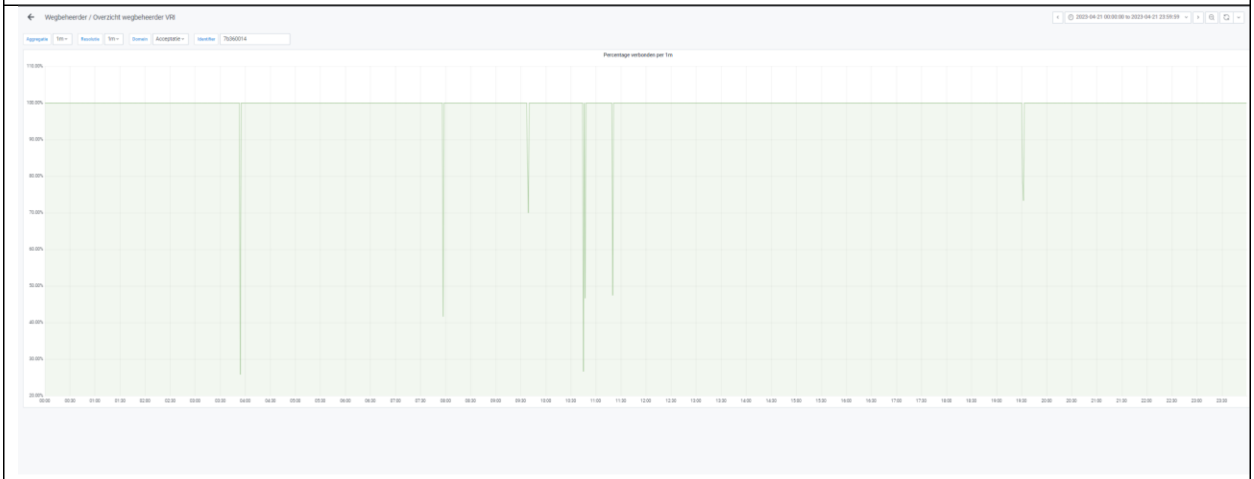
The first four missing SSEM messages have a counterpart in the SREM list around the same time. For the other SREMs, a check is done against the connection scheme of UDAP. Figure 58 contains the screenshots of the UDAP dashboard, only 18/04 and 19/04 are missing due to the retention time of the minute data of around 1,5 days (the creator missed out). Looking at the available reference of the connection drops for the available days in UDAP only station_id 67493423 cannot be matched with a connection loss and is then the only message that is not accounted for in the communication, reason was not traceable. K0038 suffers from connection loss between 3 to 7 times a day and clarifies the majority of the missing messages. The cause is probably the loss of the 4G/5G network connection due to the rural location. Due to the communication rules of UDAP a connection is closed if 10 seconds no connection could be made or messages are received. The hypothesis is that the other non-confirmed missing messages are also due to the interruption of the connection with the 5G and 4G networks (more results on network evaluation in Vlissingen provided in D5.4). Based on the number of occurrences the interruption cannot be tied to the type of connection (5G-4G) and is probably caused by the rural location of mainly K0038. Reliability only suffers from the rural coverage, especially for junction K0038



17/04/2023



20/04/2023



21/04/2023

Figure 58 iTLC K0038 percentage connected on 17-20-21/04/23 (source UDAP).

4.5.3 Discussion

Looking at both KPIs, 5G is between 10% and 25% faster than 4G with a bit less scattering of the clock deviation which is caused by the clock synchronization using the wireless connection. Using GPS synchronization instead of using the 4G/5G connection makes the clock settings more accurate. For EF3 and the C-ITS communication a lower latency means a slightly quicker response and adaptation of the ETA requested (EF7) or proposing a new Time Slot (EF3) and adaptation of the approaching vehicles. This makes the service more accurate.

Message loss is bound to connection loss of the iTLCs probably caused by coverage on the iTLC locations. This can be verified by the coverage analyses done for this project, and presented in Section 3.2 and D5.4. When no messages are received within 10 seconds the connection will be stopped and reinitialized from the iTLC end. This is due to UDAP connection rules. This potentially will disrupt the Time Slot Reservation sequence. EF7 and EF3 will try to proceed with the Time Slot reservation but the change of a match in the right (green)time windows will be much lower.

In the scope of testing further capabilities of 5G SA network on the performance of intelligent traffic light controllers, a detailed network analysis has been performed and presented in Annex (Section 6.3).

4.6 Distributed perception capabilities in the Port of Antwerp pilot site

The test demonstrates the distributed perception as Enabling Function EF4 (Figure 59), which performs 3D object detection surrounding the ego vehicle by fusing point clouds received from other cars in the platoon and the ego point clouds. The retrieval of LiDAR point clouds from the vehicles in the platoon is done via 5G Standalone, using the network infrastructure in the Port of Antwerp-Bruges (Antwerp pilot site). The 3D object detection is then displayed in the EF1 dashboard, and as such presented to the TO to increase his/her situational awareness. The process of publishing detected objects to EF1 is done via pub/sub mechanism (objects shared in a JSON format). We measured different KPIs of EF4 with reference to EF1 and EF7. In this section, we give an overview of those KPIs, and then present and discuss the results of each KPI separately.

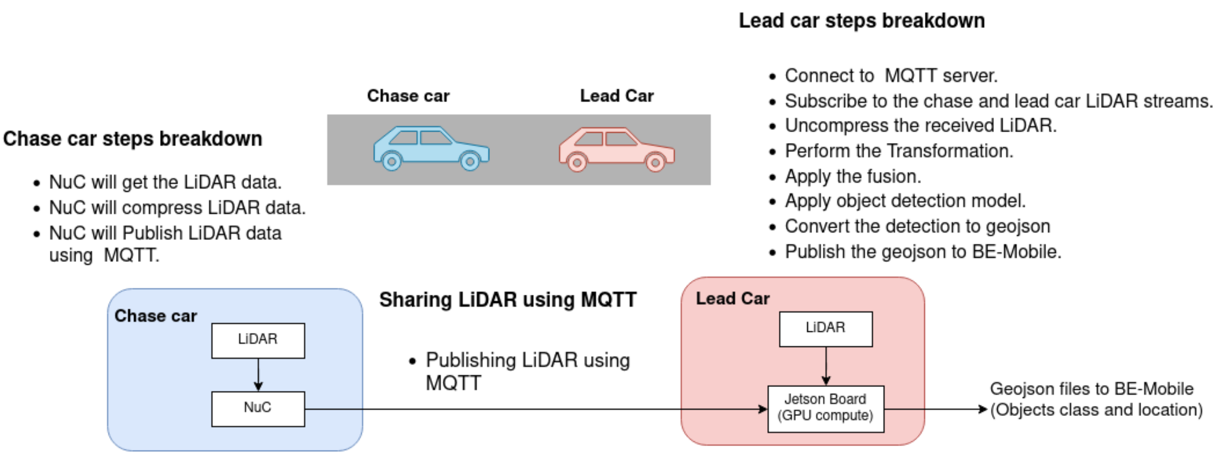


Figure 59 Overview of Distributed perception (EF4).

4.6.1 EF4 results

Table 55 Measured KPIs defined for EF4: Distributed perception.

#	KPI	Definition	Target values	Measurement	Status
KPI1 – EF4	Fusion algorithm computation time	Time required to transform and fuse the point clouds	200-400miliseconds	Based on metrics and tracing in code + processing time in response	Jan 2023 (Antwerp)

KPI2 – EF4	Object detection average precision and accuracy	Object's available being successfully detected by the algorithm	Visual comparison between available objects and detected ones	Visual	Jan 2023 (Antwerp)
KPI3 - EF4	Transmission time	The overall transmission time needed for transfering data.	Data of about ~3Mb should be transmitted successfully in real time	Frequency of point clouds data received at the edge node	Jan 2023 (Antwerp)

KPI1 – EF4

During the tests, EF4 fused LiDAR data retrieved from two different LiDAR (Livox horizon) sensors deployed on a car on two spatially different possess and heading. The pre-processing of the point clouds was first conducted on a NuC to convert the point cloud into a feature vector reducing its size by a factor of 10. The resulting feature vector is then compressed and published via MQTT servers along with the GPS coordinates, which is called the Cooperative Perception Message (CPM). EF4 used Nvidia jetson AGX orin (on-board processing unit) to aggregate and fuse the feature representation based on the GPS, and the fusion algorithm developed as part of EF4 (before July 2022). The average time to fuse the feature vector and perform object detection is 210 milliseconds which is within the range opted to reach to perform real-time and efficient distributed perception.

KPI2 – EF4

To evaluate the object detection accuracy we utilize Average Precision (AP) at an Intersection Over Union (IoU) of 0.7. Intersection over Union (IoU) is a metric commonly used to evaluate the performance of object detection. It provides a measure of the overlap between the predicted and ground truth bounding boxes or segmentation masks. The IoU is calculated as the ratio of the area of overlap between the predicted and ground truth regions to the total area covered by both regions. The IoU is expressed by the following formula:

$$IoU = \frac{Area\ of\ Intersection}{Area\ of\ Union}$$

Where "**Area of Intersection"** represents the region where the predicted bounding box or segmentation mask overlaps with the ground truth bounding box or mask. It is the common area shared by both. Whereas, "**Area of Union"** represents the total region covered by both the predicted and ground truth bounding boxes or masks. It includes the overlapping region and any areas that are unique to each. IoU is widely used in tasks such as object detection to quantitatively assess the accuracy of the algorithm's predictions. It provides a meaningful measure of how well the predicted regions align with the true regions of interest in the data. High IoU values are generally desired, indicating a better agreement between the predicted and ground truth regions.

Average Precision (AP) is a widely used metric for evaluating the performance of object detection algorithms. It combines precision and recall values at different confidence thresholds to provide a comprehensive assessment of the algorithm's ability to detect objects in an image. In our tests, we have examined the average precision at IoU of 0.7 for 9000 samples at different weather conditions (Snowing, raining, sunny) corresponding to 3 hours of testing. The AP was computed using the ground truth and the generated bounding boxes. The distributed perception algorithm developed was able to detect objects (cars, cyclists and pedestrians) within a range of 50-100m with AP of 74% at IoU of 0.7.

The current state of the art methods shares the following AP @IoU=0.7 are V2VNet [5]: 72.08% and Disconnect [6]: 72.87%. Distributed perception using intermediate representation is one of the emerging research topics, as it introduces great benefits and efficiency for the network's bandwidth usage. This methodology is currently under extensive research in the V2V and V2X research communities. Fusing intermediate representation is a challenging topic, as it deals with developing AI models that encodes the input perception data (point cloud) into a feature vector consisting of only the relevant important features, then transforming and fusing those feature vectors in one common representation with different weights given to every vehicles based on their importance to the object detection task and relevancy. Our results exceed the SOTA, and goes beyond our targeted value.

KPI3 – EF4

Network bandwidth and latency are crucial factors in distributed perception for autonomous driving due to the real-time and safety-critical nature of the tasks involved. Below we discuss why are network bandwidth and latency are important in this context.

Real-time Communication

The teleoperator relies on real-time data exchange between the teleoperated vehicle and the teleoperation control unit to make split-second decisions. Any delay in communication can impact the teleoperator's ability to react promptly to changes in the environment. Additionally, low latency ensures that information from different sensors and vehicles is quickly transmitted and processed, allowing for rapid updates to the perception of the surrounding environment.

Dynamic Environment Awareness

The teleoperator need to be aware of dynamic changes in the environment, such as the sudden appearance of pedestrians, other vehicles, or road hazards. This awareness is achieved through the exchange of timely and accurate information among vehicles. Furthermore, high bandwidth supports the transmission of large volumes of sensor data, while low latency ensures that this information reaches all relevant vehicles quickly, allowing them to adapt to changes in the environment. A robust network infrastructure with sufficient bandwidth helps ensure that there are alternative paths for data transmission in case of failures or congestion.

Transmission time – Statistical Analysis

For the network performance analysis, EF4 utilized iperf. "iperf“ is a widely used tool for measuring network performance. It allows you to test various aspects of network performance, including bandwidth, and overall transmission time. For each network type, the distributions of the measured latency values are visualized in the figures below. For the 5G SA technology, we observed low values for the latency as illustrated below.

Network type: 4G (Figure 60**)**

Average (seconds): 1.132
Standard deviation (seconds): 1.092
50% Percentile (seconds): 0.773
95% Percentile (seconds): 3.523

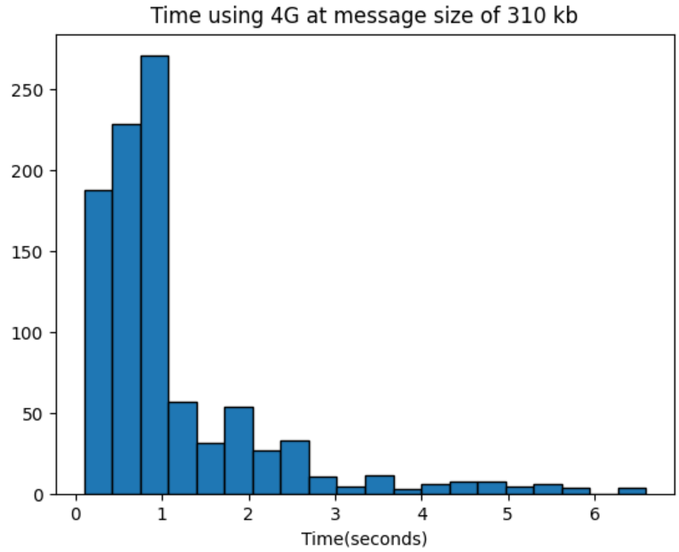


Figure 60 Time taken for message for cooperative perception message to be shared between vehicles over 4G network.

Type of network: 5G SA (Figure 61)

Average transmission time (seconds): 0.476
Standard deviation (seconds): 0.974
50% Percentile (seconds): 0.138
95% Percentile (seconds): 3.094

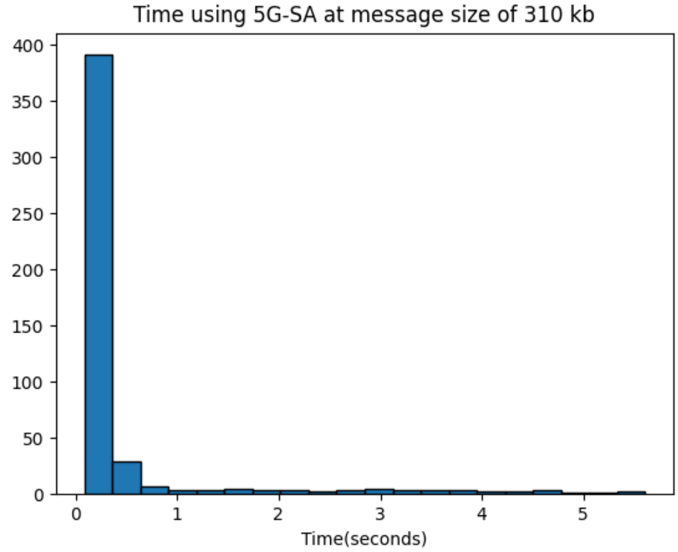


Figure 61 Time taken for message for cooperative perception message to be shared between vehicles over 5G SA network.

Network type: 5G NSA (Figure 62)

Average transmission time (seconds): 0.534
Standard deviation (seconds): 0.615
50% Percentile (seconds): 0.543
95% Percentile (seconds): 0.710

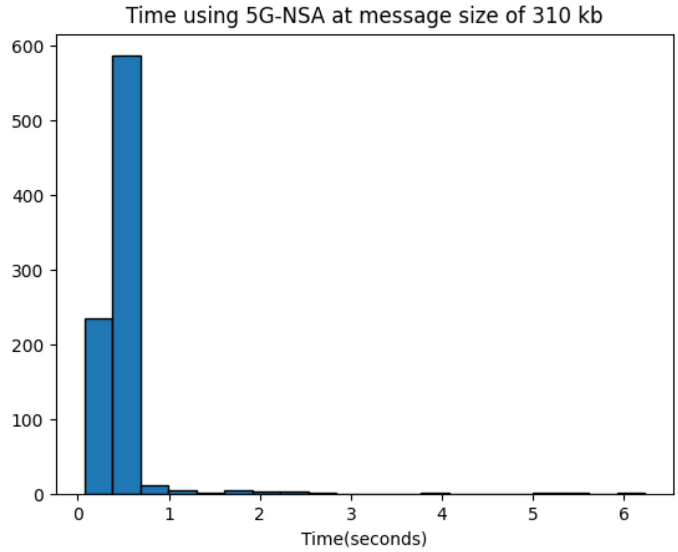


Figure 62 Time taken for message for cooperative perception message to be shared between vehicles over 5G NSA network.

Table 56 Statistical analysis of the transmission time results in case of EF4.

Transmission time (Seconds)	4G	5G NSA	5GSA
Average transmission time	1.132	0.534	0.476
50% percentile	0.773	0.543	0.138
95% percentile	3.523	0.710	3.094
Standard deviation	1.092	0.615	0.974

4.6.2 Discussion

In our methodology, we utilized Scientific Foresight Unit (STOA[4]) methodology that guarantees low computation time ensuring that the system can quickly analyze and fuse data from different sources, allowing the autonomous vehicle to respond promptly to changes in the environment. The collaboration between vehicles relies on a feedback loop where each vehicle processes its local sensor data, communicates relevant information to others, receives data from neighboring vehicles, and integrates this new information into its perception system, then broadcasts the output to the teleoperator. Therefore, fast computation time ensures that this feedback loop operates smoothly, allowing vehicles to iteratively update their understanding of the environment in near real-time. Furthermore, rapid fusion of data is crucial for responding to unexpected events, such as sudden obstacles or changes in traffic conditions. Additionally, efficient algorithms and hardware with low computation times allow the system to scale effectively, ensuring that collaborative perception remains feasible and efficient even as the number of participating vehicles grows.

Based on our tests 5G SA brings a multitude of benefits to distributed perception, fundamentally transforming the way devices and systems interact and share information. 5G SA networks offer

[4] The Scientific Foresight Unit (STOA) carries out interdisciplinary research and provides strategic advice in the field of science and technology options assessment and scientific foresight.

significantly increased data transfer speeds, allowing for the rapid exchange of large volumes of information among devices. This high bandwidth is crucial for handling the massive data streams generated by LiDARs in real-time, enhancing the accuracy and richness of distributed perception. The low transmission time over 5G SA, measured in milliseconds, ensures near-instantaneous communication between devices. This is essential for vertical applications such as autonomous vehicles, where split-second decisions are critical. Distributed perception benefits from this reduced latency, enabling more timely and synchronized responses in dynamic environments.

4.7 Video stream processing using MEC (container/rail-wagon recognition)

The testing of Enabling Function EF6 is focused on determining whether the 5G SA network is suitable for Multi-Access Edge Computing (MEC), in particular for deployments of where cameras streams are analyzed. In this case, the “edge node” is a server that resides in a datacenter from KPN. This server and underlying network nodes are configured in such a way, that the 5G modem and this server are on the same VLAN and hence, can directly communicate with each other.

A camera and 5G modem are placed locally on the Verbrugge Scaldia premises in Vlissingen. Refer to Figure 63 for an impression. The video feed is streamed 24/7 to the software, which runs in an edge node in the network (as opposed to a PC that is connected via wired network to the camera). At the edge, software is running and performing necessary analysis, in this case scanning container and rail-wagon codes.

For proper working of software running on the MEC, the connection should be stable enough so that continuous monitoring of the rail carts and containers is possible. The trains typically drive about 15-30km/h, but often such cameras are placed at a distance of only 1 to 1.5m to the container/railwagon. This implies that the container/railwagon codes are typically visible for 3 to 10 consecutive frames. Any connectivity breakdown results in lost (camera) frames, and hence increases the probability that containers/rail wagons are missed.

This testing is executed within EF6, and detection results are sent to EF1. Figure 63 gives an impression of the test system, positioned at train gate at Verbrugge premises.

Figure 63 Test system positioned at the train gate.

The system set-up including the used components are shown in Figure 64. In particular, the camera continuously streams video, a Fibocom FM160[5] is used as a 5G modem, and a UP Squared 6000 Edge[6] system is used to interface the modem.

[5] https://www.fibocom.com/en/products/5G-FM160-EAU.html

[6] https://up-board.org/up-squared-6000/

Note that the system set-up has been built as an outdoor system, so that for any test location, the same set-up is used.

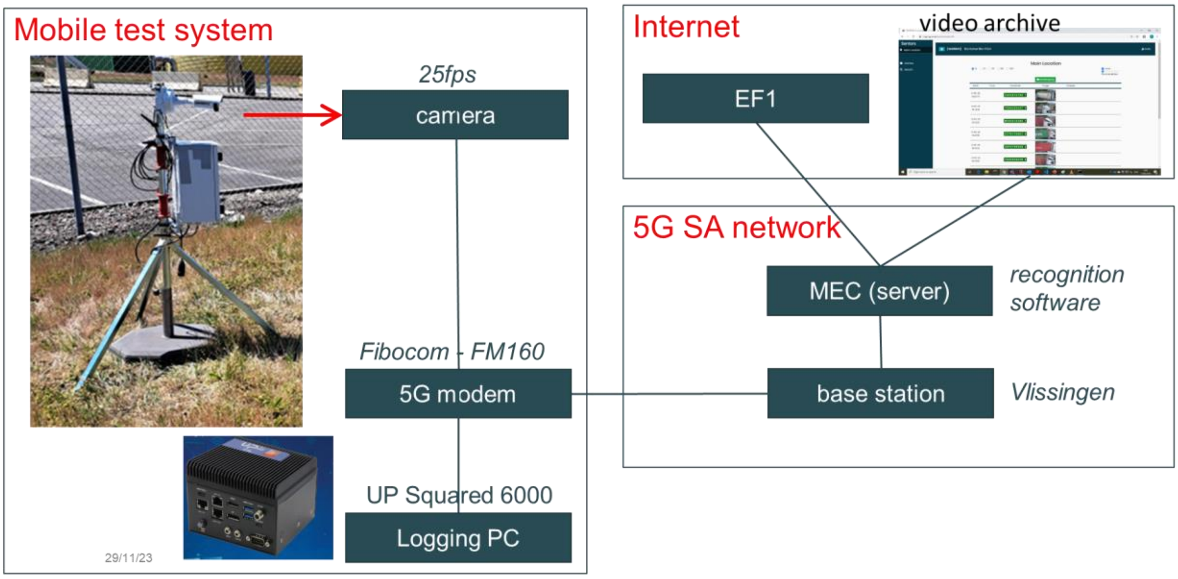


Figure 64 EF6 system setup.

The set-up has been placed at 3 different locations. Table 57 gives the overview of the 3 test locations and Figure 65 shows their locations on a map, including the position of the base station. The number of test days for each locations was 33 for Location 1, 89 for Location 2 and 3 for Location 3.

Table 57 Overview of the 3 test locations.

	Photo impression	**Characteristics**
Location 1		**Indoor** Measured from Feb 17 until May 31, 2023
Location 2		**Outdoor, next to rail gate** Measured from June 1 until November 6, 2023

Location 3		**Semi-indoor, located in a cabin** Measured from November 7 to November 30

Figure 65 Overview of the 3 test locations and 5G-SA base station.

4.7.1 EF6 KPIs

The KPIs are focused on two aspects: the 5G performance (radio quality), and the end-to-end performance of the overall system for container/rail-wagon recognition. Overall, the end-to-end performance is the most relevant for an application user such as EF6, but obviously the 5G radio performance needs to show stable results in order to achieve that.

Table 58 EF6 KPIs.

#	KPI	Definition	Target values	Measurement
1	Radio quality	radio reception as measured by the modem: RSRP, SINR, RSRQ	RSRP > -100dBm RSRQ > -20dB SINR > 0dB More elaboration about these values is provided in section 0.	Every 2 minutes
2	Bit rate	kbps as received by the edge system	Stable upload for 1 continuous camera stream, order of magnitude >=100kbps	25 times per second
3	Framerate	The number of camera frames per second that are received (or not) at the edge system	>= 25fps (frames per second)	1x per second
4	Corrupt frames	The number of camera frames that cannot be fully reconstructed	<= 1fps (frames per second)	1x per second

Table 59 shows for each location the actual number of days with logging data. Days on which the network was down or for any other reason when no logging could be performed, are not counted. Note that the radio quality was logged at the modem, and the other KPIs on the edge.

Table 59 Measurements days per test location, split per KPI.

Logging tool	Modem	Edge
KPI	Radio quality	Bitrate, framerate, corrupt frames
Location 1	103	33
Location 2	142	89
Location 3	22	3

Note that the KPI, as originally defined in Deliverable D7.2 [1], have been modified in the following way.

These KPIs have been introduced:

- The radio quality values RSRP, SINR, and RSRQ. They are included here because the focus for this EF6 has shifted to gaining more learning on the network impact, by statistically relevant logging of 5G signal quality.
- Corrupt frames. They are measured since they potentially have a large impact on the end-to-end performance. Because if frames are received, but these frames contain errors, the software still cannot use it.

At the same time, some other KPIs were initially identified, but eventually not chosen to be executed.

- Service continuity: since this is a test network, there was planned and unplanned maintenance and test actions from the network operator. During these periods, the 5G SA network was down for several hours or days, so that the connectivity was lost. Also changes in the core and routing were made, so that effectively the edge was not connected to the 5G modem. In that sense, in practice it turned out that measuring the service continuity did not add any value, as opposed to a situation in a production network.
- API delivery time: this was initially chosen to represent the end-to-end latency. However, the latency from the edge system to the public internet was judged not to be relevant, as it will differ per edge node and API endpoint, which can be optimized using commercially available options. Rather it was decided to focus on the 5G aspects, since this is the part of the chain where this project is all about. Therefore the radio quality values (RSRP, SINR, RSRQ) are logged in detail, as opposed to the API delivery time.

Radio signals

Every 2 minutes, several radio strength signals are queried on the modem[7] to measure the radio signal quality: timestamp, ss-sinr, ss-rsrp, ss-rsrq.

Signals are reported by the Fibocom FM160, the so-called "reported values"[8]. It uses the standards from 3GPP called "measurements quantity value"[9]. These values have to be converted to dB/dBm. This is done as follows:

- ss_sinr[10]: measured quantity value = -23 + reported value*0.5
- ss-rsrp[11]:measured quantity value = -156 + reported value
- ss-rsrq[12] measured quantity value = -43 + 0.5*reported value

The following sections show the measurements for all three locations.

[7] Via the command: AT+GTCCINFO? The syntax of its output is: <IsServiceCell>,<rat>,<mcc>,<mnc>,<tac>,<cellid>,<narfcn>,<physicalcellId>,<band>,<bandwidth>,<ss-sinr>,<rxlev>,<ss-rsrp>,<ss-rsrq>

[8] As a doublecheck, we know the modem values are indeed from a NR (5G modem) network, since otherwise the modem would set all these values to 255 (refer to page 169 of the Fibocom manual).

[9] 3GPP TS 38.133 version 15.16.0. release 15 on the various pages

[10] refer to page 202, Table 10.1.16.1-1

[11] refer to page 191, Table 10.1.6.1-1

[12] refer to page 197, Table 10.1.11.1-1

SINR (Signal Interference + Noise Ratio)

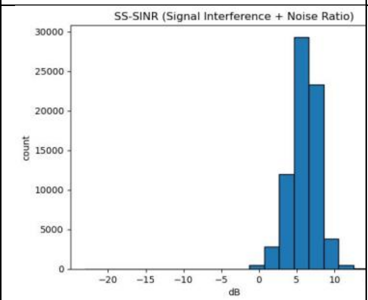
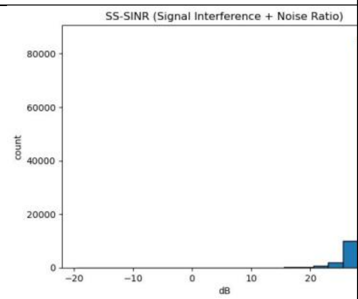
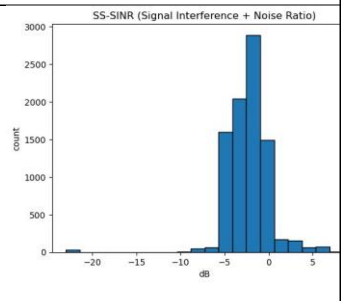
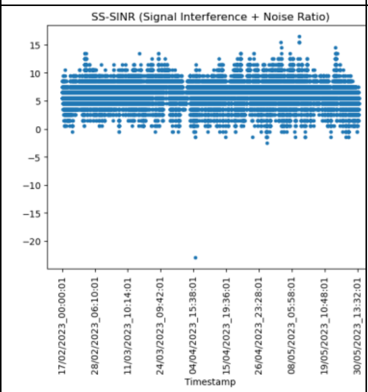
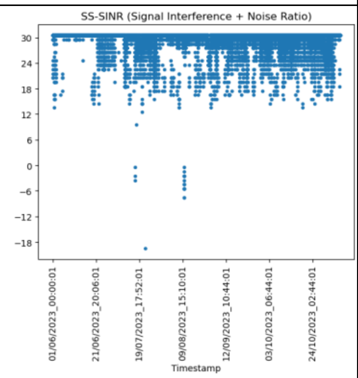
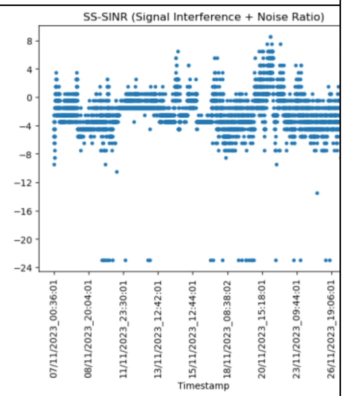
The ratio of the signal level to the noise level, is measured in dB. The higher the value, the better the signal quality. At values below 0, the connection speed will be very low, and the probability of losing a connection exists (source [7]).

Table 60 shows the measured values at the 3 locations. The top figures show the histogram, i.e., it counts how often a certain measurement level was obtained. The bottom figures show the measured value in time. Using the quality qualification as defined by Figure 66 it shows that location 1 is a typical "mid cell" location, that outdoor location 2 is "excellent", and that location 3 is at the "cell edge".

RF Conditions		RSRP (dBm)	RSRQ (dB)	SINR (dB)
	Excellent	>=-80	>=-10	>=20
	Good	-80 to -90	-10 to -15	13 to 20
	Mid Cell	-90 to -100	-15 to -20	0 to 13
	Cell Edge	<=-100	<-20	<=0

Figure 66 Qualification of RF Conditions (source [7]).

Table 60 SINR measurements at three locations.

SINR	Location 1	Location 2	Location 3
Histogram			
Time			
Typical range (dB)	Values 0 and 10	Values between 20 and 30, mostly skewed towards 25-30	Values between -10 and 5, mostly just below 0
Qualification	Mid cell	Excellent	Cell edge

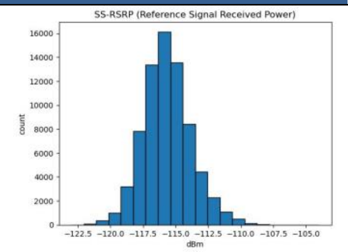
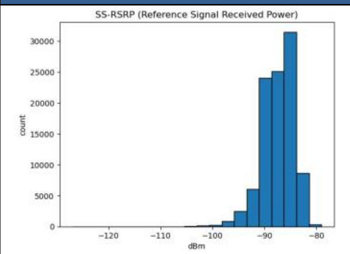
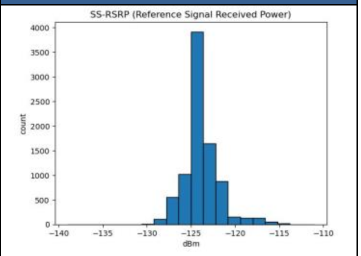
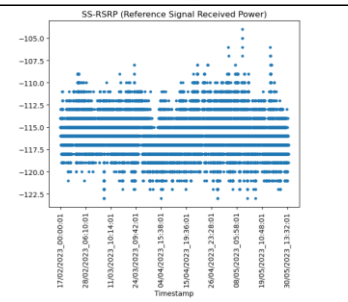
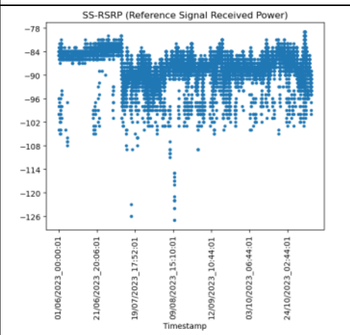
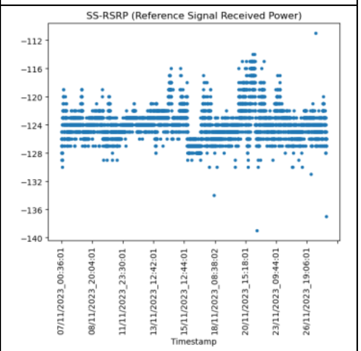
RSRP (Reference Signal Received Power)

RSRP is the average power of the received pilot signals (Reference Signal) or the level of the received signal from the Base Station. The RSRP value is measured in dBm.
With RSRP = -120 dBm and below, the connection may be unstable or not established at all (source [7]). Table 61 shows the measured values for all 3 locations.
It can be concluded that Location 2 showed a signal that can be classified as "good to excellent". The signals for Locations 1 and 3 are much less strong and are more around the "cell edge".

RF Conditions		RSRP (dBm)	RSRQ (dB)	SINR (dB)
	Excellent	>=-80	>=-10	>=20
	Good	-80 to -90	-10 to -15	13 to 20
	Mid Cell	-90 to -100	-15 to -20	0 to 13
	Cell Edge	<=-100	<-20	<=0

Figure 67 Qualification of RF Conditions (source [7]).

Table 61 RSRP measurements at 3 locations.

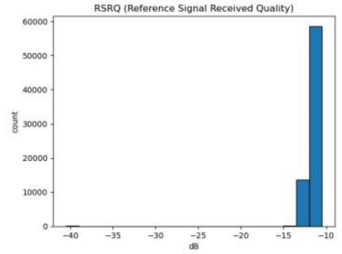
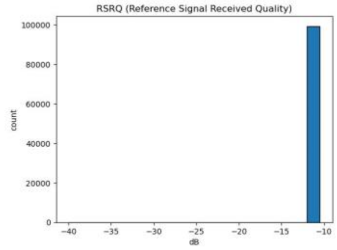
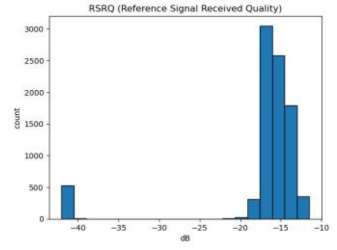
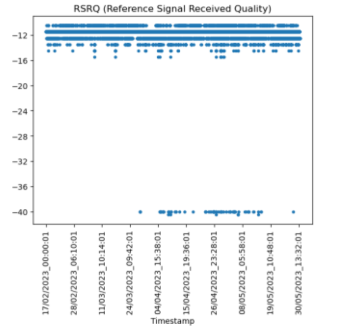
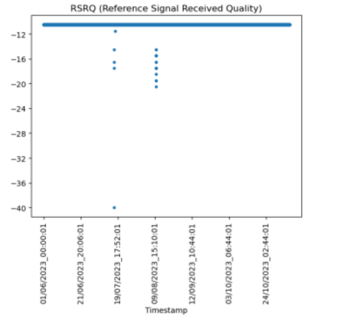
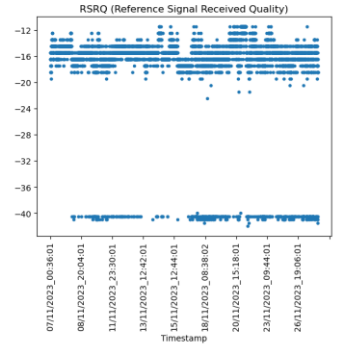
RSRP	Location 1	Location 2	Location 3
Histogram			
Time			
Typical range (dBm)	Between -122 and -107, relatively normally distributed with average around -116	Between -102 and -80, mostly skewed around -85.	Between -130 and -115, relatively normally distributed with a peak around -125
Qualification	Cell edge	Good to Excellent	Cell edge

RSRQ (Reference Signal Received Quality)
RSRQ characterizes the quality of the received pilot signals. The RSRQ value is measured in dB (dB) (source [7]).

		RSRP (dBm)	RSRQ (dB)	SINR (dB)
RF Conditions	Excellent	>=-80	>=-10	>=20
	Good	-80 to -90	-10 to -15	13 to 20
	Mid Cell	-90 to -100	-15 to -20	0 to 13
	Cell Edge	<=-100	<-20	<=0

Figure 68 Qualification of RF Conditions (source [7]).

Table 62 RSRQ measurements at 3 locations.

RSRQ	Location 1	Location 2	Location 3
Histogram			
Time			
Typical range (dB)	Between -15 and -10, mostly around -12	Almost exclusively -11	Between -20 and -12
Qualification	Good	Good	Mid cell

Bitrate

Bandwidth is the maximum rate at which bits can be transferred from a source to a destination across a given path or medium. As a rule of thumb for this particular camera (Axis P1455-LE),1 single camera has an expected bandwidth of 1.39Mbps in an indoor retail scenario, and 701kbps in an outdoor parking scenario [8].

This assumes a 1920x1080 pixel resolution at 25 fps (frames-per-second) and H265 compression. As our scenario mostly resembles a parking scenario, we assume that the camera has a ~700kbps bandwidth.

The camera will not just send 700kbps continuously. Any (IP security) camera tries to optimize the bitrate to minimize its load on the network. The **bit rate** defines the amount of data or information that is transmitted from a source during specified time period. The bitrate continuously fluctuates and depends, amongst others:

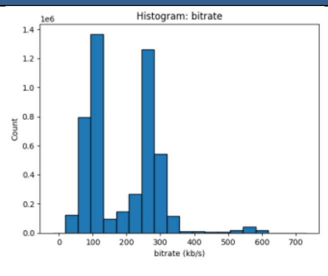
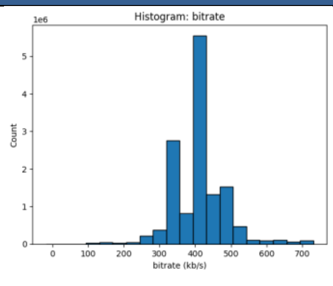
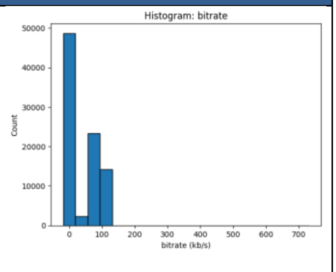
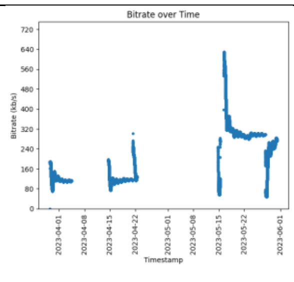
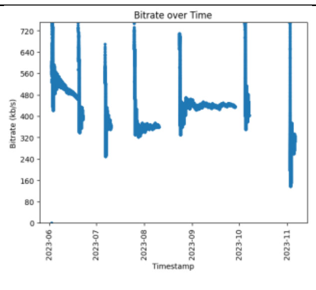
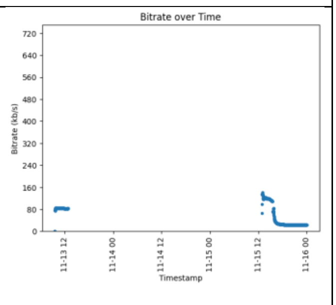
- The available end-to-end network: the RTSP protocol will capture less frames when the network cannot keep up.
- Dynamics in the scene: a completely black environment during the night has much less detail than a daylight scene with fast moving vehicles and persons
- Compression algorithm, typically MJPEG, H264 or H265. The latter 2 algorithms use variable compression: when there is less detail to capture, less data is used.
- Camera settings such as resolution, video key points, etc.

Because of the variation, it is not straightforward to simply state when a bitrate has good quality. We know the upper limit is the bandwidth of 700kbps, but without context one cannot say that a low bitrate is necessary bad. That having said, we can assume that very low values (e.g., below 50kbps) may give an indication of bad performance, since even in darker environments with no motion the typical bitrate is higher.

Table 66 gives an overview of the measurements of bitrate at the 3 different locations. The top figures give the histogram and the bottom figures the measurements in time.

It shows that the bitrate for Location 2 is "good to excellent" so this is a suitable location. Location 3 was a poor location for bitrate, but this was already clear from the radio performance. For Location 1 it's difficult to give a criterium based on the bitrate alone.

Table 63 Bitrate measurements at three locations.

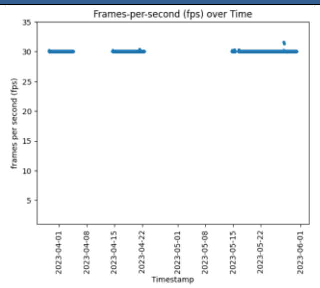
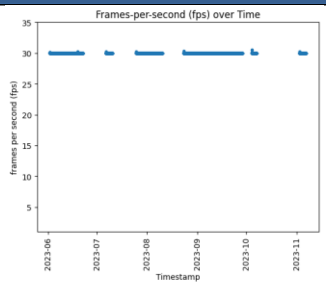
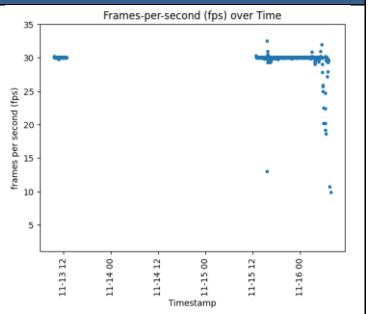
Bitrate	Location 1	Location 2	Location 3
Histogram			
Time			

Number of recorded measurements	4.821.176	13.618.941	88.599
Typical range (kbps)	Between 80 and 320, with peaks to 640	Between 160 and 560, with peaks to 720	Around 80
Context	Overall the bitrate seems okay. Also because the room had blinders, during nights the scene was almost completely dark which may explain the lower bitrates.	This looks like a healthy bitrate, the dynamics have mostly to do with day/night fluctuations, and also the amount of traffic in the background.	This basically shows that almost no stable connection could be made. Connection was only made on 2 days, and even then there was a low bitrate.
Qualification	Okay to good	Good to excellent	Poor

Framerate

Framerate is the number of frames that are obtained, usually expressed as "frames per second" or fps. In this set-up, this is logged once per second. Table 64 shows the measurements of the framerate. It shows that the framerate is kept stable at 30 frames per second. On location 3 it can be seen that the framerate varies between 10 and 30. For location 3 we've seen the poor radio reception and bitrate, so this observation confirms this situation is far from stable.

Table 64 The framerate (frames-per-second) shown in time.

Lost frames	Location 1	Location 2	Location 3
Time			
Observation	The framerate is almost exclusively around 30fps	The framerate is almost exclusively around 30fps	The framerate is mostly 30, but the last hour it varies.

Even though the framerate seems rather stable at 30fps, this only shows the framerate when there were frame coming in. It is just as interesting to see how often the framerate is not logged. And especially the times in-between, because in the cases there are frames lost. These measurements are shown in Table 65. This shows the time between 2 reported (i.e., logged) values of the framerate. When the difference is 1 second, everything is well and behaves as expected, since the fps is only reported 1x per second. If the difference is 2 seconds, one or more frame were lost - but not more than 1 second. It can be seen, that in Location 1 and 2, 99.9% of the frames was consecutive, i.e., the number of lost frames was less than 0.01%. On Location 3 the number is 99.4% and hence 0.6%. However, Location 3 already showed poor radio reception and bitrate, so more lost frames are expected.

Table 65 The time between 2 consecutive reported "fps" in the log data.

<table>
<tr><th></th><th>Consecutive reported fps</th></tr>
<tr><td>Location 1</td><td>
<table>
<tr><td>Δt [s]</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>>=5</td></tr>
<tr><td>Count</td><td>0</td><td>2510612</td><td>837</td><td>218</td><td>319</td><td>352</td></tr>
</table>
Count of time differences [seconds] between all 2512338 consecutive fps</td></tr>
<tr><td>Observation</td><td>99,931% (2510612/2512338) of the frames is logged in the next second.</td></tr>
<tr><td>Location 2</td><td>
<table>
<tr><td>Δt [s]</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>>=5</td></tr>
<tr><td>Count</td><td>0</td><td>7134773</td><td>145</td><td>5</td><td>0</td><td>8</td></tr>
</table>
Count of time differences [seconds] between all 7134931 consecutive fps</td></tr>
<tr><td></td><td>99,998% (7134773/7134931) of the frames is logged in the next second.</td></tr>
<tr><td>Location 3</td><td>
<table>
<tr><td>Δt [s]</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>>=5</td></tr>
<tr><td>Count</td><td>0</td><td>67303</td><td>2021</td><td>267</td><td>101</td><td>112</td></tr>
</table>
Count of time differences [seconds] between all 69804 consecutive fps</td></tr>
<tr><td></td><td>96,417% (67303/69804) of the frames is logged in the next second.</td></tr>
</table>

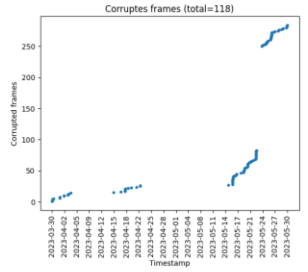
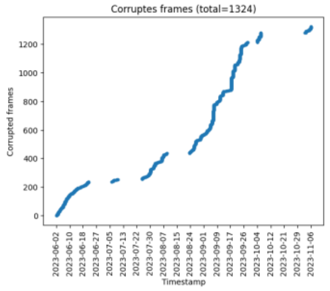
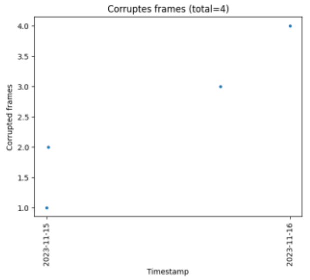
Corrupt frames

Next to the framerate, another important measure is the number of corrupt frames, which is defined as the number of camera frames that are not received properly at the edge system. This can be caused by any error in the end-to-end pipeline, in the network but also in encoder in the camera or decoder in the edge software. Note that logs for corrupt frames are only measured by the software, if data flowing is coming in – if there are no frames, there are also no corrupt frames.

Table 66 shows the related measurements at the 3 locations. The top figures show the measurements where the x-axis represents the time, and the y-axis the number of corrupt frames from the starting time. Hence, the y-axis shows the cumulative values.

The lines show mostly an (almost, not perfect) linear slope, implying that the variations were not large across the days. Also the percentage of corrupt frames (vs. valid frames) are quite similar across all 3 locations, and lie in the 0.004% - 0.009% range. In particular, we know from the radio performance, that Location 3 had a bad connection. It is noticable that when frames are submitted, still such a low percentage of these frames is corrupt. That shows the robustness of the end-to-end chain: when frames are received at the edge, the vast majority (>99.99%) is a valid frame. Hence, the measured percentage of corrupt frames is statistically neglectible.

Table 66 Corrupt frame measurements at the 3 locations.

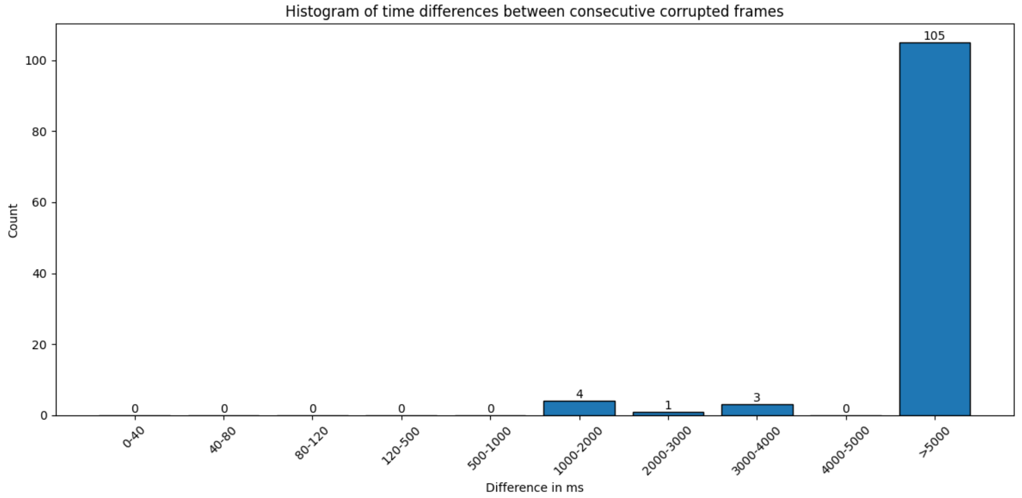
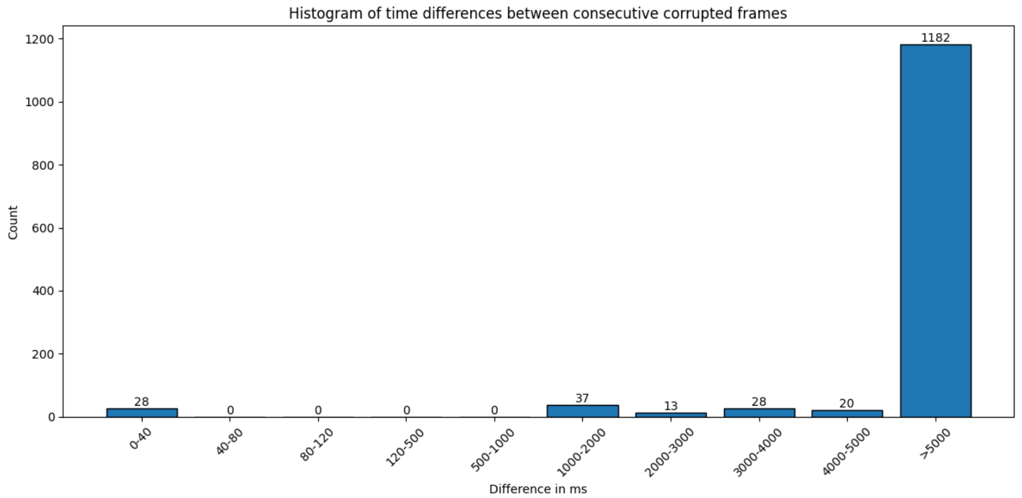
Corrupt frames	**Location 1 (*)**	**Location 2**	**Location 3**
Time			
Corrupt frames vs. total measurements (percentage)	284 vs. 4.821.176 (0.006%)	1324 vs. 13.618.941 (0.009%)	4 vs. 88.599 (0.004%)

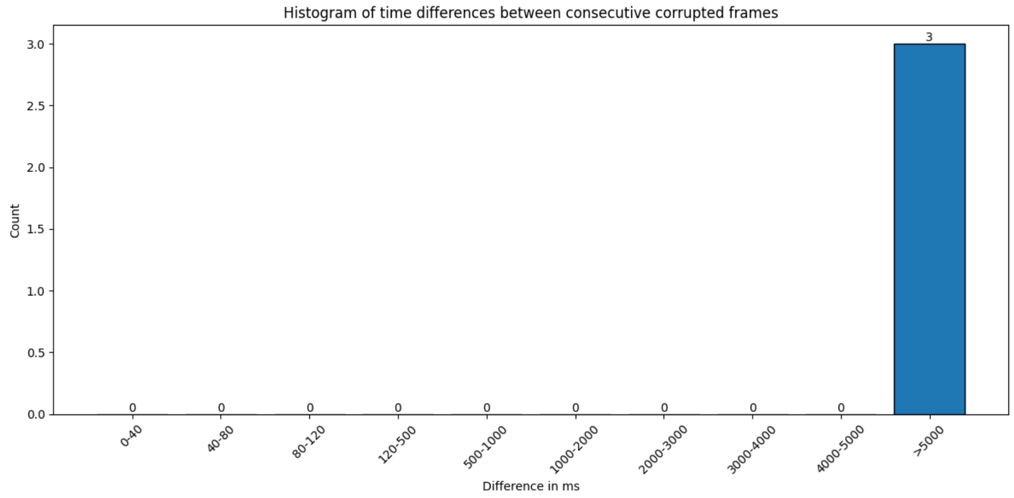
(*) In a short period on May 23 166 corrupt frames took place, that could not be explained by the telecom operator. That day is left out in this figure since we suspect a different issue (e.g., with the ffmpeg stream).

However, for EF6 the percentage of corrupt frames by itself is not the only aspect. Strictly speaking, it would be okay if 1 in 25 consecutive frames are lost, which would imply a 4% loss – but for EF6 this could still be a workable situation. However, it is not acceptable if several consecutive frames are corrupt, because then the codes on containers and railwagons would be missed. Typically a container or railwagon is about 5-10 frames visible, before it disappears. Therefore we've plotted the amount of times that corrupt frames happened in a sequence. This is shown in Table 67. At 25fps, a frame is 40ms, therefore the first few columns start with 0-40ms, 40-80ms, 80-120ms, and then bigger steps are taken. The most right column represents corrupt frames that lie more than 5 seconds after each other. For EF6 such time differences are large are not an issue, as mentioned above.

Location 1 and 3 show no consecutive corrupt frames. On location 2, there are 28 times that this happened. However, remember that location 2 was also the longest measurement period. These 28 consecutive corrupt frames happened on a total of 13.618.941 frames, which makes this statistically negligible. Also, it is worth noting that the cause of a corrupt frame can lie in the network, but also in other aspects, such as the encoder on the camera and the decoder on the edge. As such, the measured value is so low that from EF6 point of view this does not cause a concern.

Table 67 Histogram of time differences between corrupt frames.

	Correlation between corrupt frames (time differences)
Location 1	
Observation	There was always >=1 second difference (4 times), but the vast majority (105) was longer than 5 seconds
Location 2	
	There were 28 times where a corrupt frame was followed by a another corrupt frame. For the rest there was always >=1 second difference

Location 3	
	There was always >=5 second difference

4.7.2 Discussion

A real-life deployment means that a camera and 5G modem would suffice, and produce a reliable video stream towards the edge node where the software runs. Since there is no local PC, any connection drop would result in lost or corrupted frames, and hence, missing the container codes and railwagon codes, and the video footage for storage.

Table 68 Summary of KPIs.

KPI	**Location 1**	**Location 2**	**Location 3**
Measurement period	Feb 17 until May 31, 2023	June 1 until November 6, 2023	November 7 to November 30
Radio performance			
SINR	Mid cell	Excellent	Cell edge
RSRP	Cell edge	Good	Cell edge
RSRQ	Good	Good	Mid cell
End-to-end performance			
Bitrate	Okay to good	Good to excellent	Poor
Framerate	Stable 30fps	Stable 30fps	Varying fps
Framerate: consecutive loss	<0.01%	<0.01%	0.6%
Corrupt frames	Statistically neglictible and without correlation between conseuctive frames		
Conclusion			

Table 68 gives an overview of the conclusions for each KPI for all 3 locations.

Looking at the radio performance, the following can be observed:

- The performance on Location 2 looks good to excellent
- The performance on Location 1 is not ideal but working in practice
- Location 3 is not a useable location because the reception is too far towards the cell edge.

Regarding the end-to-end performance, these are the observations:

- Since bitrate is continuously varying, it is a difficult KPI to evaluate stand-alone – although it shows a correlation with the radio performance. The bitrate for Location 2 is “good to excellent”, which is clear sign that this is a suitable location. For Location 3 it is clear the bitrate is poor, and for Location 2 it seemed “okay to good” .

- For Location 1 and 2, it can be concluded that more than 99.99% frames are received, and the vast majority (>99.99%) is a valid frame. And if a frame is corrupt, there is ignorable probability that the next frame will also be corrupt.
- The measured percentage of corrupt frames is statistically neglectible. For the EF6 use case it would be acceptable if 1 in 25 consecutive frames are lost, but not if several consecutive frames are corrupt, because then the codes on containers and railwagons would be missed. The measurements have shown that this rarely happens and is statistically insignificant.

Based on these measurements, we conclude that the usage of the 5G SA network for edge computing would be suitable for this use case for Location 1 and Location 2. Location 3 had “cell edge” radio performance, and therefore its associated end-to-end performance understandably shows inferior performance.

Obviously, this is a test network with no other users, so it remains to be seen in practice how the performance will deteriorate. Nevertheless, if more users would only mean that a lower bitrate is available, and hence a lower framerate (or more compression) would be required, but further a similar end-to-end behavior, this would indeed provide a solution to process the video stream in the edge.

5 CONCLUSIONS AND LESSONS LEARNED

Having deployed both 5G Non Standalone (NSA) and Standalone (SA) networks at different pilot sites in two countries, 5G-Blueprint created a diverse environment for real-life testing and validation of teleoperation, aiming to optimize the transport & logistics operations in busy port environments. Along with teleoperation, autodocking aspects have been significantly studied and functionality developed and validated to ensure smooth interaction between two modes: teleoperation and automation. In addition, 5G-Blueprint provided a list of enabling functions whose role is to increase situational awareness, via: creating dynamic VRU and obstacle detection, using intelligent traffic light controllers, and leveraging container ID recognition services to optimize port operations. Figure 69 shows geographical locations of three pilot sites, covering both national: Antwerp (Belgium) and Vlissingen (The Netherlands), and international pilot site: Zelzate (Belgium-The Netherlands), including the indication of UCs and EFs piloted at certain locations.

Figure 69 Geographical overview of three pilot sites along with the tested UCs and EFs.

Over the project lifetime, the **pilot sites** have been **evolving**, which included either adding more locations for testing to facilitate the process and more the testing even more geographically-convenient, or increasing the coverage by deploying more gNBs. For example, during the MVP phase (reported in D7.2 [1]), Helmond site in the Netherlands has been used as a suitable location for testing basic teleoperation capabilities over 5G NSA network. Such exercise proved to be useful given the convenience of location for the partners involved in building the corresponding use cases. In the last phase of the project, Antwerp pilot site has been enriched with more gNBs on the Right bank to increase the capacity for teleoperated sailing and driving, including one more testing location in Mechelen city center as well. This location has been added to facilitate the process of debugging issues on the radio side, and for providing urban setting for testing VRU warnings (EF2).

As WP7 is a work package in which all technical functionalities developed in WP4-6 became an integral part of the **end-to-end ecosystem for 5G-enhanced teleoperation**, this deliverable also summarizes the main lessons learned stretching over network, use case, and enabling function, aspects. To provide sufficient understanding of the 5G capabilities in the national and international pilot sites, we leverage the extensive network performance analysis from D5.4 as a reference. From the results obtained in all three pilot sites, it is clear that the 5G SA network deployed in the 3.5GHz range suffers from limited range, which offers good and stable signal quality but only up to 2km away from the gNB. This signifies the importance of **proper dimensioning of 5G SA** networks, with careful gNB placement decisions, as a good signal quality is essential for uplink throughput and end-to-end latency, required for latency-sensitive applications such as teleoperation. In addition to challenges related to limited coverage, the challenging network conditions in the busy port area with many metal constructions and large trucks and ships/vessels passing by, represent a significant impact factor for network performance. Nevertheless, despite the challenging conditions, careful and extensive network evaluation resulted in measurements that are displaying promising results, showing that both SA and NSA are able to support the teleoperation requirements (5Mbps uplink throughput per sensor/camera, below 30ms end-to-end latency for remote control commands, and below 150ms interruption time during handover process). In particular, service interruption time has been measured to evaluate how much time is needed for UE to continue using the previously established session in the home network when it attaches to the visiting one. This value is specific for the cross border site and as such it needs to be minimized to ensure seamless teleoperation across country borders. The values obtained during testing show that various **optimizations in the handover procedure** significantly contribute to minimization of interruption time by proactively starting handover process (PDU session relocation prepared before handover actually happens), and minimizing the number of messages exchanged between 5G Core functions during the actual handover. The results show that both median and 95th percentile are significantly below 150ms, making service interruption time unnoticeable for cross-border teleoperation of both vehicles and barges.

During the teleoperation piloting activities in all three pilot sites, **shadow-mode testing** has been used on the public roads. In the 5G-Blueprint project, this mode of testing refers to direct control teleoperation, in which all subsystems of the teleoperation solution are active. This means that the camera streams are normally sent to the remote operator station, and the control signals created by the remote operator (steering wheel, pedals, joysticks) are normally sent to the teleoperated vehicle/barge. The specificity of the shadow mode testing is that these commands sent from the teleoperation center over 5G to the UE in the vehicle, do not obey to final translations to mechanical signals that perform the actual steering, acceleration or braking. As a result, the remote driver/skipper is not in the control of the vehicle/barge but the safety driver/skipper in the vehicle/barges. Nevertheless, all data collected during these processes are identical in both situations (remote operator in control and not in control). Shadow mode testing has proved extremely useful for testing scenarios when there is no permit to perform direct remote control, such as in the case of public roads. All results obtained during shadow mode testing allow us to fully assess if teleoperation would have been possible on these public roads or not, in normal mixed traffic, without introducing any safety risks to the surrounding traffic or infrastructure.

The tests obtained for teleoperation-based platooning in Vlissingen and Antwerp (Section 4.1) show that the controller is able to steadily control the following vehicle in the platoon over 5G network with minimal distance error, thereby validating the overall performance of the CACC system. Although the results reported the maximum speed of 60kmph, it is important to note that same quality of service is observed for higher speeds (up to 100kmph) when they were allowed on the public roads using the shadow mode testing. In addition to that, the overall CACC setup with PC5-based communication between the vehicles in the platoon showed no deactivations caused by delays imposed by 5G network, which confirms the **stability of the teleoperation over 5G**. Other service KPIs such as steering accuracy, which are relevant for teleoperation chain, exhibit values that belong to acceptable ranges, thereby reinforcing the validity of the results obtained during piloting campaigns both within and across the country boundaries.

When it comes to autodocking tests, the delay variation in relaying remote commands from the operator to the truck are usually associated with network impact. Based on the results presented in Section 4.2, it is evident that the performance of the autodocking functionality is highly reliant on the network quality. A stable network with an end-to-end latency of less than 100ms (control loop speed and TC cut-off threshold) will of course give the best results. Given the network analysis digest in Section 3, this requirement is met in all pilot sites, including the MSP Onions location where the autodocking is tested. Other service KPIs relevant for autodocking have been measured as well, such as path planning efficiency which is not directly impacted by network, but the performance of the underlying computing platform. Another one is final docking state error, which corresponds to the end position of the trailer, and if large, it means that truck trailer combination is not parked properly. As this KPI is also affected by network, obtained values of below 10cm are considered sufficient for safe autodocking process, **validating the positive impact of stable 5G connectivity**. For the purpose of localization, RTK GPS has been used and it proved as a robust and suitable method for precise localization during autodocking of full-scale trucks. However, the high prices of RTK systems motivate further studies of alternative solutions, such as localization based on cellular networks, which is out of scope of this project.

In the context of testing network connectivity for intelligent traffic light controllers (EF3) in the Zelzate pilot site, results provided in Sections 4.5 and 6.3 show interesting findings related to **slice isolation**, which is important for ensuring stable network connectivity for iTLCs. In a broader sense, the effective isolation is essential for ensuring that the performance of one slice does not impact another. Our findings show that although there is some level of isolation, the impact of high-load conditions across slices shows that more refined isolation mechanisms are needed. In the context of the iTLCs in particular, this is essential to guarantee that each slice can independently meet specific service requirements for efficient traffic control, regardless of the overall load on the network caused by other users (e.g., teleoperated vehicles approaching intelligent traffic lights at busy intersections). The current deployment of network slicing in the Zelzate city center 5G SA environment shows promising capabilities, particularly in handling diverse network demands through eMBB and URLLC slices. Nevertheless, there is still room for improvement, particularly in enhancing latency management for the eMBB slice and ensuring consistent performance and better isolation for the URLLC slice under varying network traffic load conditions.

Another interesting result is obtained during VRU Warning trialing activities in Antwerp pilot site, in both industrial and urban settings. Having network reliability as one of the key KPIs for ensuring efficient dissemination of VRU-related notifications and potential collisions, the evaluation presented in Section 4.4 shows that the urban setting resulted in many lost messages when connected on 5G SA. However, given that the urban trials were carried out using the 3.5GHz band in a test site with only one 5G node (Mechelen city center), whereas the 4G reference network has a dense configuration in this urban area, this is not a surprising result. In the respective industrial setting in the Port of Antwerp area, where multiple 5G nodes are available, the **5G network provides the required reliability** of at least 98% in real-life conditions. Such result reinforces the learning from the network evaluation test that shows essential value in proper network dimensioning at higher frequency ranges such as the one centered around 3.5GHz.

As it can be seen from all results summarized in D7.4, 5G Standalone plays an essential role for achieving strict network requirements in both network flows, i.e., uplink and downlink, and for crossing the border between two countries. With 5G SA being available at all pilot sites, the obtained results show promising future of 5G-based teleoperation in European cross-border corridors. However, with large scale deployments of remotely operated barges/trucks/cars/skid steers, it will be extremely important to dimension the network to offer higher uplink throughput for multiple parallel camera streams, and low end-to-end latency which is critical for transferring remote commands, and dissemination of safety-critical notifications to VRUs and teleoperated vehicles. Therefore, this final deliverable of WP7 provides valuable insights into realistic results obtained during extensive testing of all necessary technical elements in the 5G-enhanced teleoperation chain (network, teleoperation use cases, and enabling functions providing increased situational awareness). Such insights will further pave the way towards achieving large-scale teleoperated transport based on uninterrupted in-country and cross-border 5G connectivity.

6 ANNEX

This section presents two EFs, i.e., EF1 (Section 6.1) and EF2 (Section 6.2), which are extensively used during the piloting activities as valuable asset for displaying relevant messages for the remote drivers during teleoperation, and for calculating Estimated Time of Arrival (ETA) that is further used in the teleoperation chain. However, as the performance of these enabling functions is not directly impacted by 5G network, we include the performance analysis in the annex of this deliverable. Also, the analysis of network performance results obtained at Zelzate city center location is provided in Section 6.3. This network performance has been conducted at the city location for the purpose of testing network connectivity for intelligent traffic lights located in the city center of Zelzate.

6.1 Testing Enhanced Awareness Dashboard

The **Enhanced Awareness Dashboard** (EAD) facilitates clear and concise on-trip information about the situation on the road/waterway via a dashboard presenting a consolidated view of all safety-related information to the teleoperator (TO), increasing his/her situational awareness without creating information overload. The EAD is provided to the teleoperator on which three types of information will be displayed: speed advice, warnings, navigation & routing features.

We calculated Key Performance Indicator measurements for the enhanced awareness dashboard during the lab tests, field tests and through digital surveys, encompassing metrics such as the instant availability of route information and all safety and support functions, such as real-time display of Vulnerable Road User (VRU) information including predicted paths and potential collision risks, feedback regarding priority requests, object detection, and container identification. This input is coming from various EFs, such as EF2, EF4, and EF6, respectively.

The KPI measurements related to speed advice and the integration of warnings successfully achieved the specified target, demonstrating an average display latency of 1 second. These measurements were derived from a comprehensive analysis of metrics and tracing data. As well the Integrated yard map view and information on path, estimated time of arrival and tracking error to support automated docking are successfully tested.

6.1.1 Overview of KPIs

There are three KPIs that are measured using configured metrics, more specifically: KPI8-EF1, KPI9-EF1, KPI10-EF1. All other KPIs are measured through different digital surveys, more specifically one for each enabling function that is connected to EF1. We provided all hyperlinks to the digital surveys in the corresponding footnote[13].

At the end of each test, the tester/TO, fills in the digital survey, which contains all questions concerning the visualization, the accuracy and other related KPIs of EF1 in relation with the other enabling functions: EF2, EF3, EF4, EF8 and also Use Case 4.2.

The test results described in this document are gathered during various testing rounds of the complete set of enabling functions that are linked to EF1. For each KPI, we listed in the column

[13] All surveys are accessible via the following URLs:

EF2: https://flitsmeister.typeform.com/5GBP-survey-EF2

EF3: https://flitsmeister.typeform.com/5GBP-survey-EF3

EF4: https://flitsmeister.typeform.com/5GBP-survey-EF4

EF8: https://flitsmeister.typeform.com/5GBP-survey-EF8

UC4.2A: https://flitsmeister.typeform.com/to/lUS3ZJtU

"Status" the time of the year of the KPI is measured and at which pilot site the tests were executed. A large part of the KPIs are measured during different months and at different pilot sites. This leads to valid and robust results across the different pilot sites.

Although we succeeded to gather the needed measurements for the largest part of the KPIs leading to valid and valuable results, for some KPIs we were not able to gather the needed measurements due to the higher focus and priority that were given to the 5G communication aspects or due to some small technical interruptions of the EAD during these test rounds. Compared to the 5G technology, the specific functional aspects related to the EAD received a lower priority as the enabling functions mainly have a supporting and indirect role to measure and to demonstrate the added value of 5G. In the table below, we clearly specified for each KPI, the definition, the target values, the measurements, the status (and whether we were able to collect the needed data) and a reference to the tables or figures related to the results of the listed KPI.

Table 69 Measured KPIs defined for EF1: Enhanced awareness dashboard (EAD).

#	KPI	Definition	Target values	Measurement	Status	Ref
KPI0 - EF1	Availability of current position, speed and heading of the TOV	Basic GNSS data (coming from MQTT): current position, speed and heading.	100% available	Survey - Visual confirmation of route information on EAD.	Jan 2023 (Antwerp, EF4), February 2023 (Vlissingen, UC4.2a), April 2023 (Vlissingen, EF3)	Table 70
KPI1 - EF1	Availability of route information on EAD	Display Route information on EAD, Display Route information for specified start and end GPS position.	100% available	Survey - Visual confirmation of route information on EAD.	Jan 2023 (Antwerp, EF4), April 2023 (Vlissingen, EF3)	Table 71
KPI3- EF1	Efficiency of displaying feedback from EF3 (Timeslot Reservation at Intersections) on EF1	EAD Display result of requested priority of EF3 on EAD	100% available	Visual confirmation that when we approach intersection, priority is requested and result is displayed.	April 2023 (Vlissingen, EF3)	Table 72

KPI4-EF1	Efficiency of displaying results from EF4 (Distributed perception) on EF1	Display critical detected object for teleoperator on EF1.	100% available	Survey - Visual confirmation that all transmitted objects are shown in 3D on the road map view.	Jan 2023 (Antwerp, EF4)	Table 73
KPI6-EF1	Efficiency of displaying results from EF7 (Estimated Time of Arrival Sharing) on EF1	Display and real-time update of ETA, turn-by-turn instructions, and speed information	100% available	Visual confirmation that ETA and next turn instruction update when the TOV progresses along the route. Moreover, the max speed and speed advice must be corrected with respect to the current position of the TOV.	April 2023 (Vlissingen, EF3)	Table 74
KPI8-EF1	Frequency	Calculation of speed advice and integration into EAD	1 Hz	Based on metrics	April 2023 (Vlissingen, EF3)	Table 75
KPI11(1)-EF1	User Acceptance	Validate the driving experience of the TO with all information streams visible on the EAD.	100%	Survey	Jan 2023 (Antwerp, EF4), February 2023 (Vlissingen, UC4.2a), April 2023 (Vlissingen, EF3)	Table 76
KPI11(3) -EF1	User Acceptance	% of driver that indicate that EAD is useful for the operation of the TOV.	100%	Survey	Jan 2023 (Antwerp, EF4), February 2023 (Vlissingen, UC4.2a), April 2023 (Vlissingen, EF3)	Table 77

Summary of EF1 KPI results

Notice that the impact of 5G on the EAD is indirect. This means that the performance of the EAD and the added value of 5G are the result of the performance of other EFs (connected to the EAD) using 5G. This means that the better the other EFs work using 5G, the more accurately and timely the EAD will function and the higher the value of the EAD to the teleoperator.

The EAD shows warnings related to obstacle detection (EF4) and vulnerable road users (EF2). It is critical to receive these types of warnings in an extremely fast way. For example, for EF2 and EF4 where the positions of vulnerable road users and detected objects are gathered and shared using 5G, it is essential that the service continuity of the EAD is guaranteed and that the response time of showing this information in the EAD is consistently as small as possible. These response time values are measured in KPI8-EF1 (in terms of frequency). To make sure that the high performance of the connected EFs is preserved, the EAD should function without flaws. For this reason, there are also KPIs included that focus on the functionalities of the EAD component itself where the results are not affected by the used technology. More details about the input data collection can be found in the Sequence Diagram of EF1 in the deliverable D7.3.

We can conclude that for the following EF1 KPIs, indirect impact of 5G was evaluated as better than on 4G: the tests of EF1 with other EFs (while they are being connected to 5G) received higher ratings related to the impact (KPI11(1)-EF1) and the usefulness (KPI11(3)-EF1) compared to tests using 4G, suggesting a more positive user experience and a better perception of the technology's value.

- KPI0: Correctly displayed position, speed, and heading on the EAD, meeting expectations.
- KPI1: Correctly displayed route on the EAD, meeting expectations.
- KPI3: Successfully displayed requested priority and results of EF3, as expected.
- KPI4: Correctly displayed the transmitted object on the road map.
- KPI6: Correctly displayed ETA, turn-by-turn instructions, and speed information.
- KPI8: Response time exceeded the target value for both 5G technologies.
- KPI11(1): Enabling functions affected the driving experience; rated lower compared to 5G.
- KPI11(3): Usefulness using 4G rated lower compared to 5G technology.

KPI0 - EF1

Results

The digital survey is filled in by the tester at the end of the test. The specific question that is asked to the tester in the survey for this specific KPI is included in the header of the table below. In total we collected 12 survey responses for this question. The results show that the position was correctly displayed on the EAD for all the available technologies (while other EFs are connected to them).

Table 70 Results KPI0 - EF1.

	Question: Was the position correctly displayed (on EAD)?
4G	100% correct
5G NSA	100% correct
5G SA	100% correct

Discussion

This KPI focuses on a specific functionality of the EAD component itself. The results are not affected by the used technology. According to expectations, the position is correctly displayed on the EAD regardless of the used technology.

KPI1 - EF1

Results

The digital survey is filled in by the tester at the end of the test. This KPI focusses on correctly displaying the route on EAD. In total we collected 12 survey responses for this question. The results show that the route was correctly displayed on the EAD for all the available technologies (while other EFs are connected to them).

Table 71 Results KPI1 - EF1

	Question: Was the route correctly displayed (on EAD)?
4G	100% correct
5G NSA	100% correct
5G SA	100% correct

Discussion

According to expectations, the route is correctly displayed on the EAD regardless of the used technology. Similarly to the previous KPI, this KPI focuses on a specific functionality of the EAD component itself. We can conclude that technology used does not impact the correct visualization of the route on the EAD.

KPI3-EF1

Results

The digital survey is filled in by the tester at the end of the test. We collected in total 6 responses. This KPI focusses on correctly displaying the result of the requested priority of EF3 on EAD. The results show that the requested priority and the result were correctly displayed on the EAD, regardless of the technology.

Table 72 Results KPI3 - EF1.

Question: Was priority requested and was the result correctly displayed?
100% correct

Discussion

Similarly to the previous KPI, this KPI focuses on a specific functionality of the EAD component itself. According to expectations, the priority was successfully requested and the result is correctly displayed on the EAD.

KPI4-EF1

Results

The digital survey is filled in by the tester at the end of the test. In total we collected 5 survey responses for this question. This KPI verifies whether the transmitted object on the road map was correct. The testers confirmed that the transmitted object on the road was correctly displayed for each of the available technologies (while other EFs are connected to them).

Table 73 Results KPI4 – EF.

	Question: Was the transmitted object on the road map correct?
4G	100% correct
5G NSA	100% correct
5G SA	100% correct

Discussion

As expected, the transmitted object on the road map was correct regardless of the used technology. This KPI also focuses on a specific functionality of the EAD component itself.

KPI6-EF1

Results

The digital survey is filled in by the tester at the end of the test. We collected in total 6 survey responses. This KPI focusses on correctly displaying and updating the ETA, turn-by-turn instructions and speed information on EAD when the TOV progresses along the route. The results show that this information was correctly displayed on the EAD, regardless of the underlying network technology.

Table 74 Results KPI6 - EF1.

Question: Were the ETA and next turn instructions updated when the TOV progresses along the route?
100% correct

Discussion

According to expectations, the ETA, turn-by-turn instructions and speed information are correctly displayed and updated on the EAD when the TOV progresses. The underlying technology does not impact the correct visualisation of the route on the EAD as this KPI further focuses on a specific functionality of the EAD component itself.

KPI8-EF1

Results

For this KPI we measured the frequency related to the calculation of speed advice and integration into EAD. The target value of this KPI is set to 1 Hz, i.e., the frequency should be at least 1 Hz. This means that the calculation and integration should be updated at least every second or that the time between every update should be maximum 1 second.

This KPI was not measured yet during the MVP phase [D7.2], but only tested in lab environment. Now that the development of this EF is completed, we could measure the latency of this integration. The results are gathered on the 1st of April. However, the results are independent of time and technology used (4G, 5G NSA, 5G SA). The performance of the application will remain stable and will not be influenced by time or the type of network.

In the table below, the number of observations, the average latency, the standard deviation, the 50th percentile and the 95th are presented. The measured average latency is 1.014 seconds or 0.986 Hz. This average value is slightly below the target value of 1 Hz. The update frequency could be slightly increased to meet the target value of at least 1 Hz.

Table 75 Results KPI8-EF1.

Response time	
Number of observations	85 132
Average (s)	1.014
50th percentile	1.001
95th percentile	1.021
Standard deviation	0.146

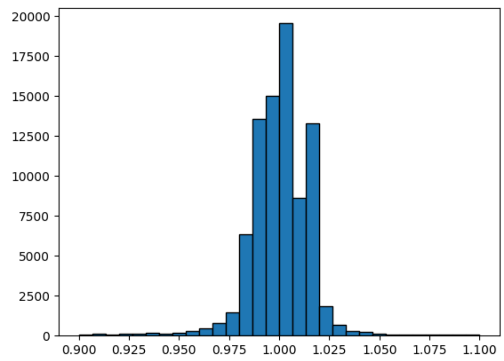


Figure 70 Histogram of the response time (s); KPI8-EF1.

Discussion

The technical KPI related to response time meets our expectation. The average value is slightly below the target value of 1 Hz. This means that if needed, the update frequency could be slightly increased to make sure that the time between two updates is never higher than 1 second or in other words the update frequency never lower than 1 Hz. However, the maximum observed time between two updates is 1.1 second and the minimum observed time between two updates is 0.9. This means that in the worst-case scenario, the time between two updates is 10% higher than the target value. But in the best-case scenario, the time between two updates is 10% lower than the target value. Only 5% of the observations of the time between two updates were equal to or larger than 1.021 second. The histogram above shows that the occurrence of the worst-case scenario of 1.1 second between two updates is exceptionally rare as well. These small variations are acceptable and the TO will not be affected by this.

KPI11(1)-EF1

Results

At the end of the test, the tester was requested to rate the enabling functions, more specifically to what extent the functions affect the driving experience. In total we collected 12 survey responses for this question. The results show that the testers rate the enabling functions highest for both 5G NSA technology and 5G SA technology. The enabling functions received a score of 8 points using 5G technology compared to a score of 6.4 using 4G technology. This KPI is an indication for user acceptance.

Table 76 Results KPI11(1) - EF1.

	Question: How did enabling functions (on EAD) affect your driving experience? (score 0-10)

4G	6.4
5G NSA	8
5G SA	8

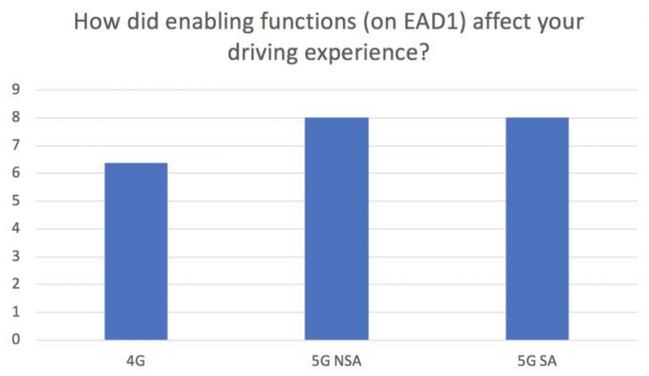


Figure 71 Results KP11(1) – EF1.

Discussion

The results show that the used technology influence the extent to which the enabling functions affect the driving experience. Using 5G technology indirectly via other EFs to which EF1 is connected, the enabling functions on the EAD affect the driving experience more. Despite the enabling functions did not receive the maximum score, the enabling functions are rated a high score for the 5G technologies. This means that the target value is reached when using 5G NSA or 5G SA. The 5G technology ensures the TOV receives all warnings on time.

KPI11(3)-EF1

Results

At the end of the test, the tester was requested to rate the usefulness of the EAD. In total we collected 12 surveys responses. The results show that the usefulness is rated the highest for the 5G SA and 5G NSA technology, more specifically an equal score of 9 points. Using the 4G technology, the usefulness is rated only 7 points. This KPI is an indication for user acceptance.

Table 77 Results KPI11(3) – EF1.

	Question: How useful is EAD? (score 0-10)
4G	7
5G NSA	9
5G SA	9

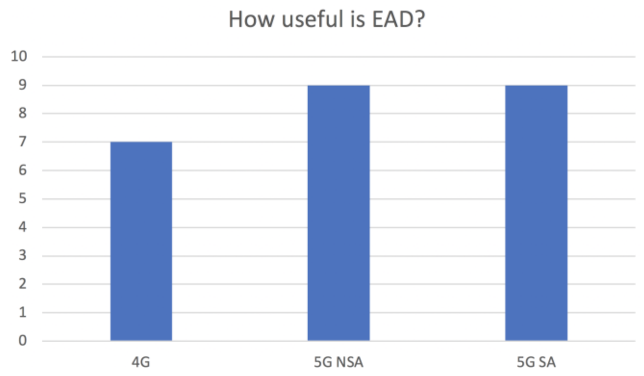


Figure 72 Results KPI11(3) - EF1.

Discussion

The usefulness of the EAD is perceived higher using 5G technology. Using 4G technology, the usefulness is rated two points lower. Despite the experienced usefulness did not receive the maximum score using 5G SA or 5G NSA technology, the target value is reached for both technologies. The testers were most satisfied with the usefulness of the EAD using 5G technology. As the EAD is connected to the other EFs using 5G, the warnings related to obstacle detection (EF4 Distributed Perception) and vulnerable road users (EF2) are captured in an extremely fast way. This makes the EAD more useful using 5G technology.

Summary results EF1 KPIs and discussions

KPI0 - Availability of current position, speed, and heading of the TOV: All technologies correctly displayed the position on the EAD. Regardless of the technology used, the position was correctly displayed, indicating that technology does not impact visualization.

KPI1 - Availability of route information on EAD - Display Route information on EAD, Display Route information for specified start and end GPS position: All technologies correctly displayed the route on the EAD. Like KPI0, the technology used did not affect the correct visualization of the route.

KPI3 - Efficiency of displaying feedback from EF3 (Timeslot Reservation at Intersections) on EAD: The priority and result are displayed correctly.

KPI4 - Efficiency of displaying results from EF4 (Distributed perception) on EAD: The transmitted object on the road map was correct for all technologies. Regardless of the technology used, the displayed object was correct.

KPI6 - Efficiency of displaying results from EF7 - Display and real-time update of ETA, turn-by-turn instructions, and speed information on EAD: ETA, turn-by-turn instructions, and speed information were correctly displayed.

KPI8 – Response time - Calculation of speed advice and integration into EAD: Average frequency slightly below the target value of 1 Hz. Thus, the response time met expectations, and minor adjustments could further improve it. Variability in response time observed but generally within an acceptable range.

KPI11(1) - User Acceptance - Validate the driving experience of the TO with all information streams visible on the EAD: Testers followed speed advice 60% of the time for 4G technology, not meeting the target value. 4G technology fell short of the target, indicating room for improvement in user acceptance.

KPI11(3) - User acceptance - % of drivers that indicate that EAD is useful for the operation of the TOV: Usefulness rated higher for 5G technologies compared to 4G. 5G technologies received higher ratings for usefulness, reaching the target value.

6.2 Sharing Estimated Time of Arrival (ETA)

The Estimated Time of Arrival sharing (EF7) provides real-time ETA and routing information to the TO and other interested parties (e.g., EFs using ETA information), as well as sets up an exchange of data with terminal systems to dynamically organize the container pick-up or drop-off time at the terminal.

High performance of the ETA module was demonstrated and measured via various KPIs, including:

- 100% uptime for the ETA calculation component and data feed.
- Efficient frequent ETA updates (60 ETA calculations per minute) leading to an accurate and no outdated ETA on the EAD.
- Positive user acceptance, with a high percentage of drivers finding the ETA in the EAD operationally useful for the TOV.
- Exceptionally low processing time of ETA requests resulting in a very responsive application with the most up to date information: the median processing time of an ETA request is 6ms.

Measured KPIs and results

We configured the required metrics to measure the number of ETA requests versus ETA responses (KPI1-EF7), the number of refreshed ETAs per minute (KPI2-EF7) and the processing time of the ETA request (KPI4-EF7). The other KPI (KPI3-EF7), concerning the correctness of the ETA, route and turn-by-turn instructions is measured via digital surveys, more specifically one for each enabling function that is connected to EF7. We provided all hyperlinks to the digital surveys in the corresponding footnote[14]. The tester / TO is asked to rate the quality of the route information using a score between 1 (low) and 10 (high) at the end of the test.

During various testing rounds of the complete set of enabling functions that are connected to EF7, we gathered the test results. Similar to the documentation of the test results of EF1, we listed in the column “Status” the time of the year of the KPI is measured and at which pilot site the tests were executed. In the table below, we clearly specified for each KPI, the definition, the target values, the measurements, the status (and whether we were able to collect the needed data) and a reference to the tables or figures related to the results of the listed KPI.

[14] All surveys are accessible via the following URLs:

EF2: https://flitsmeister.typeform.com/5GBP-survey-EF2

EF3: https://flitsmeister.typeform.com/5GBP-survey-EF3

EF4: https://flitsmeister.typeform.com/5GBP-survey-EF4

EF8: https://flitsmeister.typeform.com/5GBP-survey-EF8

Table 78 Measured KPIs defined for EF7: Estimated Time of Arrival (ETA) sharing.

#	KPI	Definition	Target values	Measurement	Status	Ref
KPI1-EF7	Service continuity	Uptime of the ETA calculation component Uptime of the (internal) ETA data feed	100% available	number of errors when ETA is requested = 0	March 2023 (Vlissingen, EF3)	Figure 73
KPI2-EF7	Frequency of ETA computations	To have an accurate and no outdated ETA, the Number of ETA calculations per hour must be sufficiently large	at least 10/min per vehicle	Based on metrics and tracing in code	April 2023 (Vlissingen, EF3)	Figure 74
KPI3-EF7	Visual confirmation of correctness	When driving, verify if the quality of the route information is correct	Correctness of ETA/route/turn-by-turn information	Visual	Jan 2023 (Antwerp) April 2023 (Vlissingen, EF3)	Table 79
KPI4-EF7	Process time of ETA request on ETA API	time it takes before eta is returned when asked for via the ETA API	less than 100ms	Based on metrics and tracing in code: Be-Mobile offers processing time in headers of API's. This allow partners to measure communication time from request coming in till arriving at partner. Be-Mobile will on top report on processing time.	April 2023 (Vlissingen, EF3)	Figure 75

Notice that EF7 is an external-facing component that provides the ETA, the fastest route and the turn-by-turn information to the EAD. The listed KPIs are all directly related to the performance of this component regardless of the technology used (4G or 5G). Nonetheless, a stable, high-performance and reliable ETA component is critical to ensure safe and smooth teleoperation realized over 5G.

For the ETA calculations, the minimum requirement of 10 ETA calculations per minute per vehicle, ensuring exceptionally accurate and up-to-date information, was exceeded. Thus, regardless of the underlying network technology, EF7 demonstrated outstanding performance in ensuring service continuity, in terms of the frequency of ETA updates, and the processing times of ETA requests.

- KPI1-EF7 (Service Continuity): Achieved 100% uptime of ETA calculation with no errors during ETA requests.
- KPI2-EF7 (Frequency of ETA Computations): Exceeded the target with an average of 60 ETA calculations per minute, ensuring up-to-date ETAs.
- KPI3-EF7 (Visual Confirmation of Correctness): Quality of route information rated at 7 out of 10.
- KPI4-EF7 (Process Time of ETA Request): Consistently achieved processing times well below the target of 100 milliseconds.

<u>KPI1-EF7</u>

Results

For this KPI we measured the uptime of the ETA calculation. The target value is set to 100%. The goal is to not have any errors when the ETA is requested. Therefore, it is important that the (internal) ETA data feed is available without any interruptions. This KPI is an indication for the service continuity. To evaluate whether the target value for this KPI is reached, we measured both the number of requests and the number of responses on March 31 2023 during testing. If the number of requests is equal to the number of responses, there were no errors when the ETA is requested. This means that the target value is reached.

The figure bellow visualizes the number of requests and the number of responses.

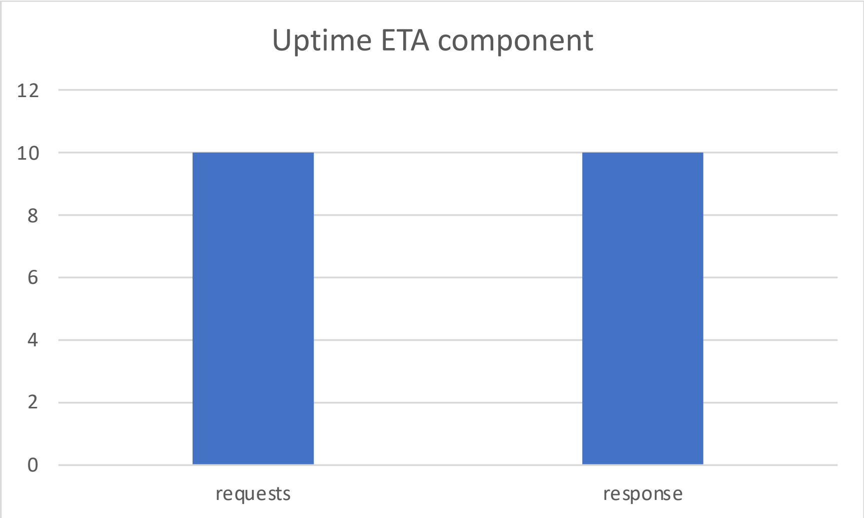


Figure 73 Uptime ETA calculation component (number requests and number responses).

The figure shows that the measured number of requests is equal to the measured number of responses. This means that the target value is reached during the performed tests. The uptime of the ETA calculation component is 100% and is independent of the technology as the ETA component is not influenced by the used technology of the TO and the remaining part of the process flow.

Discussion

The results show that the service continuity is guaranteed. This is essential for a smooth operation of the service. Although we only gathered the number of responses for 10 requests during the additional tests, we are convinced that the uptime of this component meets the target as this was also carefully tested during the implementation. Moreover, this KPI focusses more on the performance of the ETA component itself. We can conclude that the results meet our expectations.

KPI2-EF7

Results

For this KPI we measured the frequency of the ETA computations. The number of the ETA calculations per minute must be sufficiently large to make sure that there is always an accurate and no outdated ETA available in the application. For example, for EF3 an accurate ETA is needed for smooth time slot reservation at the intersection. Next to EF3, this KPI is also relevant for vulnerable road user interaction (EF2) by providing timely warnings to TOs and VRUs about potential conflicts. The target value is set to at least 10 ETA calculations per minute and per vehicle to meet the needs of both EF2 and EF3 and assure compliance between the different components. The measurement of this KPI is based on metrics and tracing in the code.

This KPI is an indication of the accuracy and usefulness of the ETA calculations. The figure bellow shows a histogram of the measurements of the time in seconds between ETA updates during the tests in Vlissingen. The horizontal axis presents the observed time in seconds between ETA updates and the vertical axis shows the frequency of the observed measurements. For example, 1 second between two occurring ETA updates was observed more than 30 000 times.

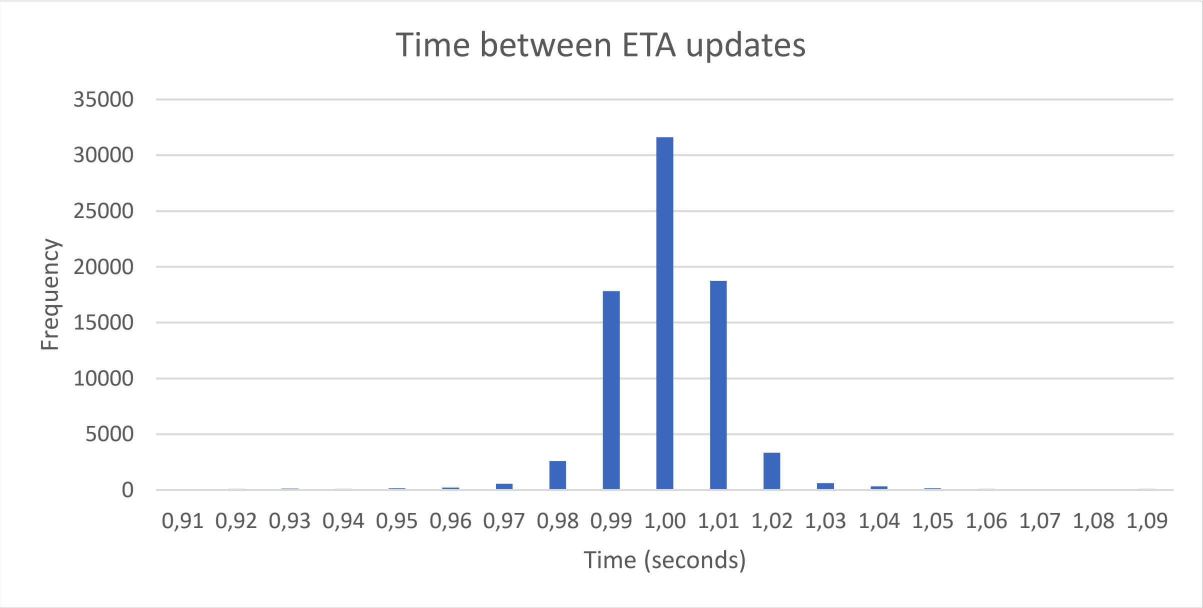


Figure 74 Histogram time (s) between ETA updates.

The results show that the average time between ETA updates is approximately 1 second. This results in approximately 60 ETA calculations per minute. The number of ETA calculations per minute largely exceeds the minimum value of 10 updates per minute defined as the target value. We can conclude that the service exceeds the minimum requirements. The results show that the target value is reached. The time between ETA updates is independent of the underlying network technology used.

Discussion

The service calculates 6 times more updates per minute than the minimum expected number of updates set as a target value. This assures us that the application always uses up to date ETA calculations leading to high usefulness. The number of updates per minute is configured in the implementation of this EF. The results show that the implementation works as expected meeting and exceeding the minimum requirements. The time between ETA updates is not impacted by the technology used.

KPI3 - EF7

Results

At the end of the test, the tester was requested to give a score to the quality of the route information. The results show that the quality of the route information is rated the 7 points out of 10 during performed tests. This KPI focusses on visual confirmation of the correctness of the route information.

Table 79 Results KPI3 – EF7.

Question: What score would you give to the quality of the route information? (score 0-10)
7

Discussion

The results show that testers were quite satisfied with the quality of the route information. We can state that the target value is reached for any underlying network technology.

KPI4-EF7

Results

For this KPI we measured the processing time of an ETA request on the ETA API. The time it takes before an ETA value is returned when requested via the ETA API must be sufficiently low to assure the application is sufficiently responsive and to make sure that the ETA updates are constantly up to date with a minimum of delay. The target value of the processing time of an ETA request on the ETA API is set to less than 100 milliseconds. The measurement of this KPI is based on metrics and tracing in the code.

The figure bellow shows a histogram of the measured processing time in seconds of ETA requests. The horizontal axis presents the observed processing time in seconds of an ETA request and the vertical axis shows the frequency (in absolute numbers) of the observed measurements.

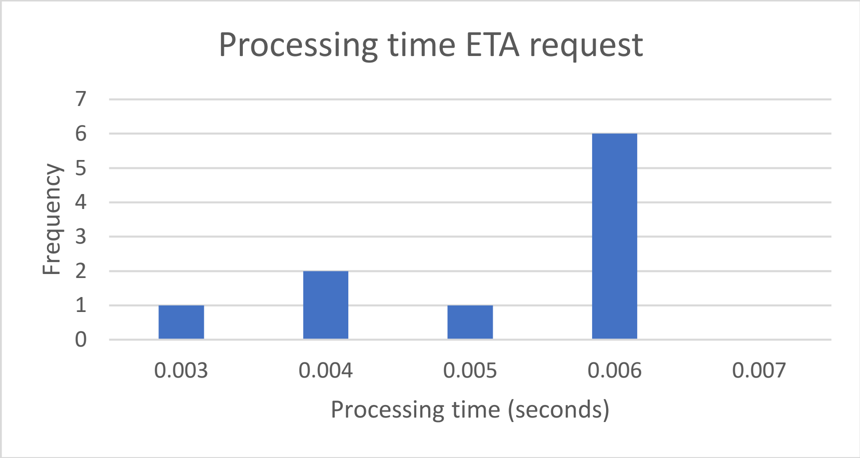


Figure 75 Histogram processing time (s) of an ETA request.

For all measured ETA requests, the processing time in seconds is always smaller than the target value of 100 milliseconds or 0.1 seconds. The median processing time is 6 milliseconds or 0.006 seconds and thus very much lower than the target value of 100 milliseconds. The processing time of an ETA request is independent of the technology used.

Discussion

In a similar way to the previous metrics related to the uptime and the frequency of ETA calculations, the processing time was also carefully tested during the implementation. The results meet our expectations. Low processing times are very beneficial. This guarantees a smooth and up to date service. Moreover, the processing time between ETA updates is not impacted by the technology used.

Summary EF7 KPI results and discussions

KPI1-EF7 - Service continuity - Uptime of the ETA calculation component - Uptime of the (internal) ETA data feed: Uptime of ETA calculation was 100%, and the number of requests was equal to the number of responses, indicating no errors during ETA requests. Service continuity is guaranteed, which is crucial for smooth operations. Results met expectations.

KPI2-EF7 - Frequency of ETA computations - To have an accurate and no outdated ETA: The service largely exceeded the minimum requirement of 10 ETA calculations per minute per vehicle, with an average time between ETA updates of approximately 1 second. The application consistently provided up-to-date ETA calculations, exceeding the minimum requirements, and this metric was independent of the technology used.

KPI3 - EF7 - Visual confirmation of correctness - Quality of route information: Quality of route information was rated 7 out of 10 during the performed tests. Testers were relatively satisfied with the quality of route information, meeting the target value.

KPI4-EF7 - Process time of ETA request on ETA API: Processing time for ETA requests via the ETA API was consistently lower than the target value of 100 milliseconds, with a median processing time of 6 milliseconds. Low processing times ensure a responsive and up-to-date service, meeting expectations. Processing time was not impacted by the technology used.

6.3 Network evaluation at Zelzate city center location

In this section, we present the results of our network evaluation for network slicing in Zelzate city center, for the purpose of validating the network performance for intelligent traffic light controllers (EF3). The primary focus of this evaluation is to evaluate TCP performance in uplink direction, with a specific emphasis on latency and throughput metrics. The goal is to evaluate the network slicing on eMBB and URLLC slices.

The evaluation not only underscored the individual strengths of each slice but also highlighted their combined efficiency and robustness in handling high-demand scenarios.

We explain how the evaluation was conducted, explaining the architecture and the methodology. Then we comment the results for uplink scenarios, with a dedicated section for each scenario.

6.3.1 Test setup

In Zelzate city center, the evaluation of network slicing was conducted to assess the capabilities and performance of different network slices. This evaluation utilized two UEs (Peplink routers), each configured with distinct SIM cards and connected to Telenet's gNB. To perform the test, two laptops were integrated into the setup, connected to the respective Peplink devices. The setup can be resume as depicted in Figure 76. The setup included two UEs, with UE1 attached to the eMBB slice and UE2 connected to the uRLLC slice. Traffic generation and latency measurement were carried out using Iperf3 and ping, respectively.

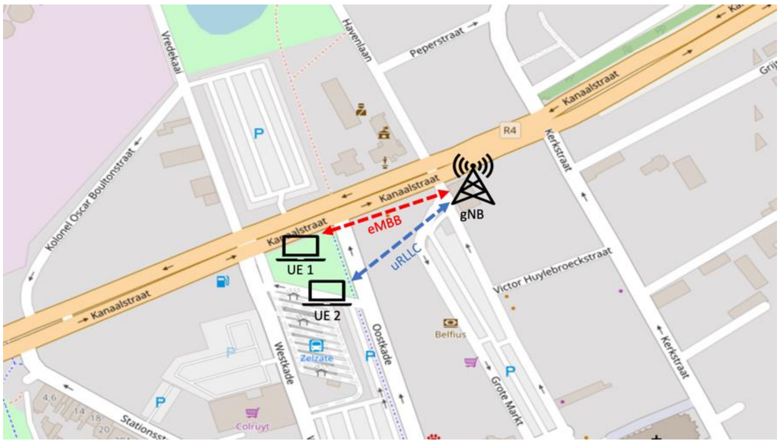


Figure 76 Testing location in Zelzate city center.

The evaluation was structured into two main phases. The first phase focused on testing each slice individually to measure KPIs, such as throughput and RTT latency. These measurements were critical in establishing benchmarks for future comparisons. During this phase, traffic was generated through Iperf3, and RTT latency was simultaneously monitored using ping tests. The KPIs obtained from this phase were categorized as "NO_impact", serving as a baseline for the network performance under standard conditions.

In the second phase, the emphasis shifted to evaluating the performance and isolation of both the eMBB and uRLLC slices when subjected to concurrent, high-load conditions. This phase aimed to mimic extreme operational scenarios to test the resilience and reliability of the network slices on the 5G SA infrastructure. The traffic of approximately 200Mbs was generated for each slice over a duration of 60 seconds to identify their respective failure points. The phase also focused on verifying the effective isolation between the slices under stress conditions. Concurrently, the impact on network performance, particularly on RTT latency, was assessed, with these results being marked as "YES_impact ".

6.3.2 Result analysis

Uplink

The results obtained on end-to-end latency while performing uplink throughput tests via iperf3 are described below.

In the case of eMBB slice**, no Impact Scenario, which is used as a benchmark, the eMBB slice exhibited stable performance with low latency and minimal standard deviation. This consistent behavior, unimpacted by external factors, is indicative of the eMBB slice's robust capability in managing latency under typical operating conditions.** On the other hand, with Impact Scenario is introducing background traffic brought about a significant change. The latency within the eMBB slice showed a marked increase, with the average latency escalating to 181.50ms. Such a notable increase in latency during high-load conditions points to a potential limitation of network slicing. While it proves efficient in regular circumstances, its performance, particularly in terms of latency, can be compromised in latency-sensitive eMBB applications under substantial load. This becomes critically important in scenarios where prompt data transmission is essential, such as in data generated from cameras in teleoperated vehicles. The increase in latency under stress conditions underscores the need for enhanced management and optimization strategies in network slicing to ensure consistent performance across varying operational demands.

In the case of URLLC slice, and no Impact Scenario, mean latencies were observed in the range of 42-44ms. This demonstrates a notably low-latency performance, which is essential for the efficiency of URLLC services. The ability to maintain such low latencies is crucial for applications that depend on real-time data transmission and processing. With Impact Scenario, even with the added strain, the Latency_URLLC_YESimpact scenario maintained low latency levels, albeit with a slight increase. This subtle rise in latency underlines the potential impact of network slicing configurations on URLLC services, especially in loaded environments.

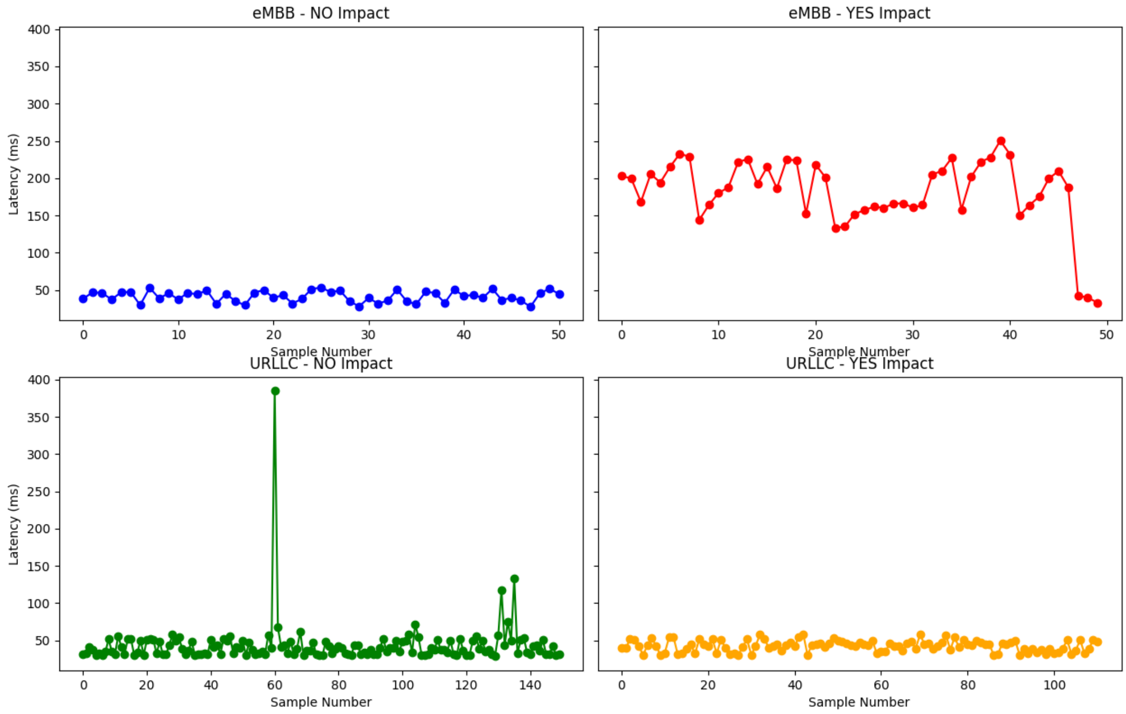


Figure 77 Latency measurements for URLLC and eMBB slices.

Table 80 Latency performance in Zelzate city center.

Scenario	Mean	Median	Standard Deviation	Minimum	Maximum	90th Percentile	95th Percentile
Latency_eMBB_NOimpact	41.53	43.3	7.35	27.5	53.3	50.8	51.5
Latency_eMBB_YESimpact	181.5	189.8	46.9	33.0	250.45	227.3	230.4
Latency_URLLC_Noimpact	43.8	38	31.3	29.5	385.8	54.4	57.8
Latency_URLLC_YESimpact	42.6	43.2	7.5	30.3	58.5	52.3	54.2

When evaluating throughput performance using iperf3, we observed distinct behaviors in the eMBB and URLLC slices under various conditions.

In the case of eMBB slice, the following results are obtained. In No Impact Scenario, the eMBB slice showcases a robust and consistent bitrate performance. The mean bitrate was recorded at 45.701613Mbps, with a median close to this value at 45.9Mbps, indicating a stable distribution of data rates. The standard deviation was 2.49836Mbps, reflecting minimal variability in bitrate. The range of bitrates spanned from a minimum of 38.8Mbps to a maximum of 52.5Mbps. The 90th and 95th percentiles, at 48.37Mbps and 49.19 Mbps respectively, further confirm the slice's capability to maintain high bitrates consistently.

In the case of scenario with Impact, the eMBB slice demonstrated a slight decrease in performance, though it remained relatively stable. The mean bitrate marginally reduced to 45.274194Mbps, and the median was 45.35Mbps, showing only a minor deviation from the No Impact Scenario. The standard deviation decreased to 1.933818Mbps, suggesting a tighter concentration of values around the mean. The minimum and maximum bitrates recorded were 39.8Mbps and 51.8Mbps, respectively. The 90th and 95th percentiles, at 46.6Mbps and 47.595Mbps, were slightly lower than in the No Impact Scenario, indicating a small but noticeable impact on the upper end of the bitrate distribution.

The measurements in the case of URLLC slice are provided as follows. In no Impact Scenario, the URLLC slice exhibited commendable performance with a mean bitrate of 12.7Mbps and a median of 12.65Mbps, suggesting consistent delivery of data rates. The standard deviation was relatively high at 4.45Mbps, indicating greater variability in the bitrates. The bitrate ranged from a minimum of 1.19Mbps to a maximum of 22Mbps. The 90th and 95th percentiles stood at 18.3Mbps and 19.45Mbps, respectively, underscoring the slice's ability to achieve higher bitrates under normal conditions

On the other hand, in the scenario with Impact, the URLLC slice experienced a decline in bitrate performance. The mean bitrate was 9.8Mbps, with the median at 9.3Mbps, showing a decrease compared to the No Impact Scenario. The standard deviation was 3.5Mbps, lower than in the No Impact Scenario, which points to a narrower spread of bitrate values under stress. The minimum and maximum bitrates were 1.37Mbps and 18.2Mbps, respectively. The 90th and 95th percentiles were 15.6Mbps and 16.6Mbps, indicating a reduction in the higher bitrate values achievable under loaded conditions.

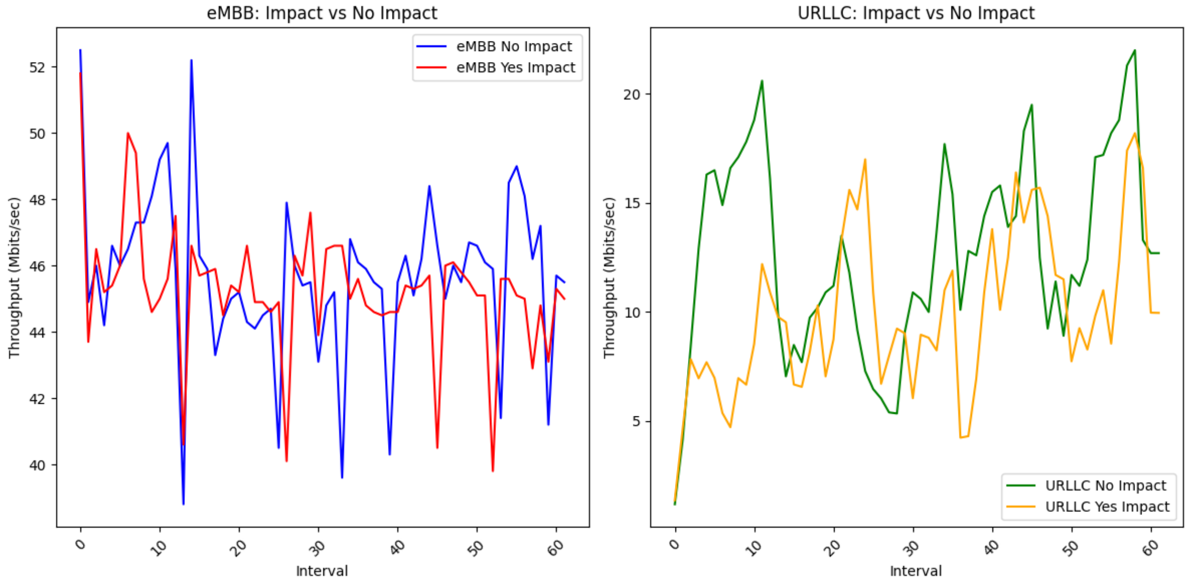


Figure 78 Uplink throughput values on uplink, Zelzate city center location.

Scenario	Mean Bitrate	Median Bitrate	Standard Deviation	Minimum Bitrate	Maximum Bitrate	90th Percentile	95th Percentile
eMBB_NOimpact	45.701613	45.9	2.49836	38.8	52.5	48.37	49.19
eMBB_YESimpact	45.274194	45.35	1.933818	39.8	51.8	46.6	47.595
URLLC_Noimpact	12.7	12.65	4.45	1.19	22	18.3	19.45
URLLC_YESimpact	9.8	9.3	3.5	1.37	18.2	15.6	16.6

6.3.3 Conclusion

In this section, we presented and evaluated the current state of network slicing development within the Telenet 5G SA environment in Zelzate, focusing on eMBB and uRLLC slices. The evaluation centered on critical performance metrics, notably latency and uplink throughput. Two metrics with great relevance in the context of intelligent traffic light controllers (EF3).

From our analysis, it is evident that network slicing in its current form exhibits both strengths and areas for improvement. The eMBB slice, tailored for high-throughput requirements, showed remarkable consistency in bitrate across various scenarios. Its performance under no impact conditions was particularly strong, showcasing the slice's capability to handle high data demands efficiently. However, under loaded conditions, we observed a noticeable impact on latency, suggesting that while the eMBB slice excels in throughput, its latency management could be further optimized, especially for applications where time-sensitive data transmission is crucial.

The URLLC slice, on the other hand, is designed for scenarios demanding ultra-low latency. Our findings indicate that in no impact scenarios, the URLLC slice successfully maintains a low-latency profile. However, under stress conditions, the slice experienced significant fluctuations in both latency and throughput. This variability, particularly in latency, raises concerns about the slice's reliability in consistently delivering the ultra-low latency required for critical applications.

The performance of network slices under loaded conditions also brings into focus the aspect of slice isolation. Effective isolation is pivotal for ensuring that the performance of one slice does not detrimentally affect another. Our evaluation suggests that while there is some level of isolation, the impact of high-load conditions across slices indicates a need for more refined isolation mechanisms. This is essential to guarantee that each slice can independently meet its specific service requirements, regardless of the overall load on the network.

In conclusion, the current development of network slicing in the Zelzate city center 5G SA environment shows promising capabilities, particularly in handling diverse network demands through specialized slices. However, our analysis highlights the need for ongoing improvements, particularly in enhancing latency management for the eMBB slice and ensuring consistent performance and better isolation for the URLLC slice under varying load conditions.

REFERENCES

[1] 5G-Blueprint, "Deliverable D7.2: Platform Integration and Minimum Viable Platform," 2022. [Online]. Available: https://www.5gblueprint.eu/wp-content/uploads/sites/62/2023/08/D7.2_Platform-Integration-and-Minimum-Viable-Platform_V1.0_31.10.2022.pdf. [Accessed December 2023].

[2] 5G-Blueprint, "D6.1 Description of enabling functions and their requirements," H2020 5GBlueprint Project Consortium.

[3] 5G-Blueprint, "D6.2 Description of functional architecture (role and functional),," H2020 5GBlueprint Project Consortium.

[4] 5G-Blueprint, "D6.3: Description of enabling functions architecture (Technical)," 2023. [Online]. Available: https://www.5gblueprint.eu/wp-content/uploads/sites/62/2023/01/D6.3_31_1_2023..pdf. [Accessed December 2023].

[5] Wang, Tsun-Hsuan, S. Manivasagam, M. Liang, B. Yang, W. Zeng and R. Urtasun, "V2vnet: Vehicle-to-vehicle communication for joint perception and prediction.," in *Computer Vision–ECCV 2020: 16th European Conference*, Glasgow, UK, 2020.

[6] Li, Yiming, S. Ren, P. Wu, S. Chen, C. Feng and W. Zhang, "Learning distilled collaboration graph for multi-agent perception," in *Advances in Neural Information Processing Systems 34 (2021)*, 2021.

[7] [Online]. Available: https://support.zyxel.eu/hc/en-us/articles/360005188999-Signal-quality-LTE-5G-LTE-and-5G-signal-quality-parameters.

[8] "Axis site designer," [Online]. Available: https://sitedesigner.axis.com/.

[9] 5G-Blueprint, "D4.1 Barge Use Case," H2020 5GBlueprint Project Consortium.

[10] 5G-Blueprint, "D4.2 Tele-operated Docking System Use Case," H2020 5GBlueprint Project Consortium.

[11] 5G-Blueprint, "D4.3 CACC-enabled Cars and Containers Use Case," H2020 5GBlueprint Project Consortium.

[12] 5G-Blueprint, "D5.1: 5G network requirements and architecture," H2020 5GBlueprint Project Consortium.

[13] 5G-Blueprint, "D5.2: Initial report on the 5G network evaluation," H2020 5GBlueprint Project Consortium.