NTNU
Kunnskap for en bedre verden

Department of Mechanical and Industrial Engineering

TMM4540 - Industrial ICT Engineering Design and Materials, Specialization Project

Latency for Remote Control in ATO a Scoping Review

Author:
Einar Thingstad Myhrvold

Student number:
543951

Supervisor:
Nils O.E. Olsson

Date: 16.12.2025

Preface

This article was written for the subject TMM4540 - Industrial ICT Engineering Design And Materials, Specialization Project.

The development of ATO at NTNU has been in progress for multiple years and is now in the remote operation phase. Master students at the Department of Mechanical and Industrial Engineering has steadily worked on this project. In the fall of 2024 and spring of 2025 it was from Emilia Kozarevic, and I will continue her and NTNU's research on the topic.

I have had no previous experience about ATO. But subject & knowledge listed below has been used.

- TBA4225 - Railway Engineering 2
- IT2810 - Web Development
- TDT4240 - Software Architecture
- TDT4120 - Algorithms and Data Structures

I would like to end the preface by thanking Nils Olsson for the work of supervising me during this semester and the effort he has put in to help me write this paper. I also would like to thank Patrick Urassa (NTNU) and Tomas Rosberg from VTI for discussing and the insight in the project.

Abstract

Remote train operation is an essential function for higher Grades of Automation, serving both as a development tool and as a safety fallback when automated systems degrade or fail. A critical requirement for remote operation is low and predictable latency, particularly for video transmission and control. This paper presents a scoping review of how latency is evaluated, measured, and interpreted in remote train operation systems, with supporting insights drawn from related domains including car, drone, and crane remote operation. The study reviews existing railway research, frameworks, and demo projects before comparing them with established methods and findings from other industries where remote control has been in development longer. Particular emphasis is placed on latency measurement techniques such as glass-to-glass, end-to-end, and round-trip transmission delay, as well as on the influence of video encoding, streaming protocols, network technologies, and human factors. The results show that while railway environments benefit from more predictable operation conditions than road or aerial vehicles, current latency thresholds are largely inherited from other domains and lack specific validation for trains. Human performance is shown to be sensitive not only to complete latency but also to latency variability, video quality, and system stability. The paper concludes that dedicated railway experiments, standardized latency measurement frameworks, and systematic testing of video streaming protocols are necessary to define reliable operational limits and support the safe deployment of remote train operation within future systems.

——————— Norwegian ———————

Fjernstyring av tog spiller en viktig rolle i systemer med høy grad av automatisering. Løsningen brukes både i utvikling, test og som en sikkerhetsløsning hvor automatiserte funksjoner ikke fungerer som planlagt. For at fjernstyring skal være optimal, må forsinkelsen i kommunikasjonen være lav og stabil. Dette gjelder spesielt for overføring av video og styringskommandoer mellom tog og operatør. Denne artikkelen gir en oversikt over hvordan forsinkelse i kommunikasjon analyseres og vurderes i løsninger for fjernstyrt togdrift. Fremstillingen bygger på gjennomgang av eksisterende jernbane relatert forskning, supplementert med erfaringer fra andre domener der fjernstyring har vært i bruk over lengre tid, blant annet innen bil, drone og kranoperasjon. Målet er å finne hvilke metoder som brukes for å måle og analysere forsinkelse, og hvordan disse anvendes i ulike sammenhenger. Videre gjennomgås sentrale tilnærminger til måling av forsinkelse, som måling fra kamera til skjerm, samlet forsinkelse gjennom hele systemet og fram og tilbake i kommunikasjonssystemet. Artikkelen diskuterer også hvordan valg av videokoding, strømmeteknologi, nettverksløsninger og operatøren påvirker både den faktiske og den opplevde ytelsen. Gjennomgangen diskuterer hvordan togdrift foregår under mer stabile og forutsigbare forhold enn mange andre former for fjernstyrte kjøretøy, spesielt biler og droner. Likevel er mange av kravene angående forsinkelse som benyttes i dag hentet fra nettopp bil og drone, uten at de er tilstrekkelig dokumentert eller verifisert for bruk i jernbanesystemer. Resultatene indikerer også at operatørens prestasjon påvirkes av flere forhold enn bare total forsinkelse, blant annet variasjoner i forsinkelse, bildekvalitet og systemets stabilitet over tid. Til slutt konkluderes det med behovet for målrettede eksperimenter tilpasset jernbane, felles retningslinjer for hvordan forsinkelse skal måles, samt grundigere testing av strømmeløsninger for video. Slike tiltak er nødvendige for å kunne fastsette realistiske og sikre driftsgrenser og for å legge til rette for trygg innføring av fjernstyrt togdrift i fremtidige systemer.

Table of Contents

List of Figures

List of Tables

1 Introduction

The chapter of Introduction will contain text about purpose, research questions and a small bulk about history and the level of digitalisation we have reached right now.

1.1 Purpose

The purpose of this paper is to establish a structured foundation for understanding, measuring, and evaluating latency in remote train operation (RTO) systems. As the railway sector transitions toward higher Grades of Automation (GoA), particularly GoA 3 and GoA 4, remote control becomes an essential operational component. Previous research highlights that RTO is required for safely deploying and validating autonomous capabilities, as it enables controlled testing and secure intervention when automated functions cannot perform as intended [1]. In this context, RTO functions as a mandatory fallback layer within the ATO architecture, ensuring continuity of safe operation during degraded modes or for unexpected track conditions [2].

A important requirement for such fallback control is the availability of reliable low-latency video transmission. Studies on remote train supervision emphasise that successful human intervention depends heavily on the stability and responsiveness of video streams when the physical control room is replaced by a remote workstation [3]. Industry demonstrations show that remote operation can supports practical tasks such as shunting, depot movements, and vehicle preparation and thereby improving efficiency while maintaining safety [4]. At the same time, the downgraded mode scenarios described in deliveries from Europe's Rail [5] show how responsibility will transition between automated systems and human operators. And it is therefore important to identify the operational limits imposed by latency, uncertainty, and reduced information quality.

An equally important aspect is understanding how latency affects human perception and reaction time. Human factors research provides insight into how video quality, bitrate, and delay influence a remote operator's ability to detect signals, interpret scene dynamics, and perform corrective actions under time constraints. These findings contribute to defining minimum perceptual requirements for safe fallback operation and support the formulation of latency thresholds for high automation environments [6].

1.1.1 Research Questions

To achieve the purpose of this paper, the research questions listed below was created to help.

1. RQ1: How is latency evaluated for railway and other industries?
2. RQ2: What parameters influence latency in remote train operation systems?
3. RQ3: How can we test different video streaming protocols and their impact on real-time performance of remote train operation systems?

1.2 Digitalisation in Railway

Railway digitalization has progressed from early computer assisted signalling and centralized traffic control systems to modern Automatic Train Protection (ATP), ERTMS/ETCS deployments, and integrated traffic management platforms. Recent developments include communication-based train control (CBTC) in metros, ETCS rollout on mainlines, and the integration of predictive maintenance and data analytics tools. Projects such as national ETCS rollouts, the UK East Coast Digital Programme, and research initiatives like ATO-Cargo exemplify a shift from isolated automation pilots to system wide modernization that combines ATO, interoperability standards (TSIs), and remote supervision concepts [7, 8].

The current digitalisation also shapes how future railway operations are organised. Projects across Europe show that automation and improved connectivity all depend on reliable digital systems that support both automated driving and human supervision. Remote control is becoming a standard element in this transition because it allows operators to intervene safely when automated functions cannot handle a situation. It also supports new operational concepts such as remote supervision of rural lines and more efficient traffic management [1]. At the same time, the move toward FRMCS and 5G aims to replace GSM-R and provide the higher bandwidth and lower delay needed for real-time video and control during remote operation [3].

Digitalisation also increases the system's dependency on stable communication and sensor data. Research shows that issues such as reduced bandwidth, sensor faults, or missing information can directly affect whether remote operation is possible, making clear procedures for degraded modes essential [5]. Human factor studies further underline that the quality of digital perception. Especially video streams must be high enough to replace traditional visibility when operators supervise the train remotely [6].

2 Method

2.1 Research Methodology

The following sections detail the methodological approach and structured review process used to address the research questions of this study. This methodology is designed to ensure a robust foundation for the evaluation of latency for remote train operations.

As of a scoping review described by Arksey [9] the process followed the steps of deciding research questions before moving over to identifying relevant studies. The research questions has been slightly modify to adapt to the information available and what was discovered during the research phase. All the papers and studies was also selected before charting data, comparing, summarizing and reporting on the results. As mentioned by Arksey [9], one point of a scope review is to identify research gaps in the existing literature. Which was very useful and ended up being most of the focus for this paper.

This study uses a Mixed Methods approach. Which includes combining two types of information: qualitative (ideas, experiences, and opinions) and quantitative (numbers and statistics). This method is chosen because using both together gives us a complete picture, where relying on only one type of data would not be enough [10].

The study utilizes convergent parallel design for this. Which entails collecting and analysing the numbers and the opinions at the same time. Then comparing the results from both to see how they fit together and explain our final findings [11]. The results from technical performance testing (quantitative) are thus validated and enriched by the practical feedback received from user trials (qualitative), leading to more actionable conclusions.

2.2 Literature Review Strategy

The literature review was conducted to establish a comprehensive theoretical and informative basis for the research. The process was guided by the principle of Evidence-Based Standards to ensure methodological rigour and focus [12].

2.2.1 Search Strategy and Databases

A multi-platform search approach was utilized to retrieve a wide array of high-quality sources. The primary databases included:

- Scopus: For retrieving peer-reviewed, impactful scientific articles.
- Google Scholar: For broader academic and institutional literature.

While maintaining a strong reliance on peer-reviewed scientific material, relevant non-academic reports and industry publications were also considered to provide a comprehensive perspective, with all information traced to reliable sources. Especially theoretical part of the background research has benefited from non-academic sources.

2.2.2 Inclusion Criteria

To ensure the study is based on the relevant and current information, specific criteria were applied to filter the search results. Given the rapid pace of technological change, a focus was placed on recent publications. Where modern papers have referenced to older papers, I have include some as to say and show where certain numbers are coming from and how they have influenced results.

Table 1: Inclusion Criteria for Literature Review

ID	Inclusion Criteria (IC)
IC1	Publication Date: Between 2020-2025
IC2	Language: Written in English
IC3	Document Type: Primarily "Article", "Conference Paper" or "Thesis"

I did not include a specific keyword criteria because of the exploratory nature of the research questions. Instead, broad search terms were used initially, with relevance determined through title and abstract screening against the inclusion criteria.

2.2.3 Search Queries and Refinement

Initial broad queries were executed and subsequently refined to focus on specific research gaps, such as the intersection of video communication and system latency. For example, Table 2 illustrates a query targeting the core technological elements of the study. "Filtered" are based on the inclusion criteria from Table 1 while "Selected" has amount of reports used from each search by the method in Figure 1. Note that some selected overlap with each other, especially in Table 3.

Table 2: Summary of Refined Search Queries.

Search Query	Initial Hits	Filtered	Selected
"Remote" "Train operation" OR "Train control"	12 900	5 380	4
"Automatic" "Train operation" OR "Train control"	17 600	11 900	3
"Vehicle" "streaming" "protocols" "real-time" "performance"	36 800	17 400	4
"Ethics" "Remote" Control" "Vehicle"	298 000	34 200	1
"Cybersecurity" "Remote" "Control" "Vehicle"	42 600	21 600	1
"Latency awareness" "Remote" "Control"	460	294	2

For more detailed searches and for finding reports to compare with, the following queries in Table 3 stands for the majority of papers.

Table 3: Summary of Refined Search Queries.

Search Query	Initial Hits	Filtered	Selected
"Remote" "Latency" "Operation" OR "Control"	841 000	83 000	14
"Remote" "Latency" "Operation" OR "Control" "Railway"	11 800	6 580	6
"Remote" "Latency" "Threshold" "Operation" OR "Control"	245 000	33 700	8
"Remote" "Latency" "Threshold" "Operation" OR "Control" "Railway"	4 350	2 530	3

The results of the query clearly show a minor part of the total remote control "community" specialize in railway operation. There is also possible to see a trend showing that it has become increasingly more popular in recent years as around half of the papers are from the last 5 years.

To be able to choose from multiple thousands papers down to the ones chosen, a filtration method has been used as you can see in Figure 1. Numbers are based on the example of 33 700 for *"Remote" "Latency" "Threshold" "Operation" OR "Control"*. Where "Headline filtration" was an estimate on how relevant it appears to be for this research based on headline, amount of citations and year produce where newer papers was judges easily on citations. "Abstract reading" filtered looking for relevant context of the papers. "Full reading" confirmed value for the review and research, looking for useful information that could contribute to this scoping review. Finally "Selection" was papers adding new perspective not already discussed by another paper, or papers reconfirming other papers.

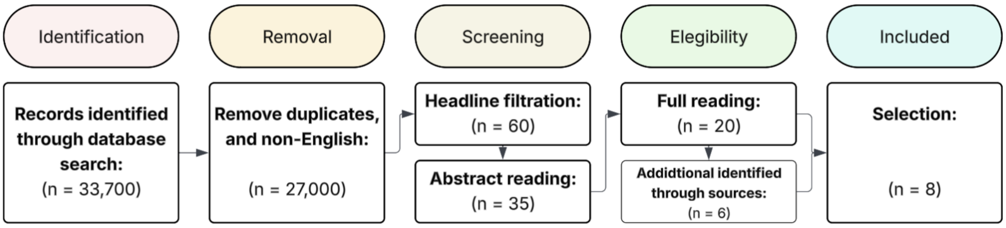

Figure 1: PRISMA flow chart for query: "Remote" "Latency" "Threshold" "Operation" OR "Control"

2.3 Supporting Tools

Several digital tools as mentioned below were used for this paper. The main purpose of the usage was to enhance clarity, assisting in latex, and in general increase quality. Specifics include writing reference list in correct format. Structuring sentences and paragraphs. Checking for typos and grammatical errors.

- OpenAI's ChatGPT: Employed as a helpful resource for LaTeX formatting suggestions, generating structured content (tables and lists), translating between Norwegian and English, and reviewing text for synonyms and restructuring ideas to ensure arguments were effectively communicated.
- Google Gemini: Used for similar fields as ChatGPT, but also for general proofreading, and refining sentence structure and tone to maintain a high standard of academic writing.
- Visual Studio Code: was used as the main LaTeX editor because of its powerful extensions for LaTeX support, syntax highlighting, and version control integration.

2.4 Validity

In "Reliability and Validity in a Nutshell", Bannigan and Watson define validity as the extent to which a measurement accurately measures what it is intended to measure [13]. When identifying relevant papers, I have tried different setups and keywords to make sure they give the wanted output. An example is to use just "Train" instead of "Railway" will give a lot more results as it will includes all reports talking about training personnel or computers. Or if you restrict to "Train Control" the results drop significantly, loosing important papers for this topic. It is difficult to determine if the aspects I have focused on is the best and correct way to compare and find a good solution. However, as I will also discuss in Chapter 5 "Results and Discussion", many of the other vehicle industries have decided to focus on the same aspects.

2.5 Reliability

Bannigan and Watson describe reliability as the degree to which a measurement is consistent and dependable [13]. In this report, reliability is addressed by drawing on a wide range of studies from different industries before making comparisons. This approach is intended to increase consistency and dependability, such that similar results could be achieved if the same research were conducted by others. Based on the number of papers reviewed and included, this report aims to identify general patterns related to latency and threshold. These patterns are expected to be replicable, even if future studies use different source materials. As this is a broad examination of the topic, the overall findings should therefore be reproducible. In contrast, conclusions related to the testing and evaluation of tools and protocols may vary between researchers. These assessments rely more heavily on specific test scenarios and individual examples, which can reasonably lead to different interpretations and outcomes.

3 Theory

3.1 ATO

Automatic Train Operation (ATO) describes systems that automate driving tasks normally performed by a human driver. ATO implementation range from assisting the driver with speed guidance, optimized speed profiles and other information to fully unattended operation where starting, cruising, stopping and door control are automatic. The primary goals are improved punctuality, energy efficiency and safe, repeatable performance [14]. In freight-specific research such as the ATO-Cargo project, ATO is combined with existing train protection systems (for example ETCS Level 2) and a Remote Supervision and Control Centre (RSC) to allow remote human oversight and intervention during degraded operation or faults [8].

3.2 GoA

The Grade of Automation (GoA) classifies how much of the train operation is automated. Standards such as IEC 62290 and industry reports [15] describe the commonly used levels from GoA 0 to GoA 4. Table 4 below summarises the practical meaning of each level.

GoA	**Meaning / operator role**
GoA 0	On-sight, manual operation without automatic protection.
GoA 1	Automatic Train Protection (ATP) Manual driving with assisted protection routines. Human driver performs traction, braking and door tasks while safety limits are done automatic. That includes track speed, safe routing and safe separation.
GoA 2	Semi-automated (STO). ATO handles start/stop and trajectory control between stations; a driver remains onboard for door operation, obstacle response and degraded mode handling.
GoA 3	Driverless (DTO). No driver needed for normal operation. Staff may be on board for passenger assistance and emergencies. ATO handle operational tasks including avoiding collision with obstacles and persons.
GoA 4	Unattended Train Operation (UTO). Fully automated operation without staff onboard. Remote supervision and controls are required for special incidents.

Table 4: Grades of Automation (GoA), summary based on IEC and industry sources.

Remote Train Operations (RTO), as discussed in this paper, is a part of stage GoA 3 as a disruption management handling system and as a safety system for GoA 4 if it were to fail [6].

GoA also changes and updates regularly because of new technology added that shifts the definition. Figure 2 below is UITP's definition of how GoA is graded.

Grade of Automation	Type of train operation	Setting train in motion	Stopping train	Door closure	Operation in event of disruption
GoA1	ATP* with driver	Driver	Driver	Driver	Driver
GoA2	ATP and ATO* with driver	Automatic	Automatic	Driver	Driver
GoA3	Driverless	Automatic	Automatic	Train attendant	Train attendant
GoA4	UTO	Automatic	Automatic	Automatic	Automatic

**ATP - Automatic Train Protection; ATO - Automatic Train Operation*

Figure 2: UITP's simple definition of GoA [16]

The formal definitions and required functions per level are described in IEC 62290 and discussed in CBTC Solutions as well as UITP [15–17].

3.3 ERTMS

The European Rail Traffic Management System (ERTMS) is the standardized signalling and control framework developed to improve interoperability, safety, and capacity across European railways. Its introduction replaces diverse national signalling systems with a unified architecture consisting primarily of the European Train Control System (ETCS) and a digital radio communication layer (GSM-R, and in the future FRMCS). According to the European Union Agency for Railways, ERTMS is designed to harmonize train control and supervision across borders while enabling higher automation levels and modern traffic management solutions [18].

The shift toward ERTMS is the result of decades of signalling evolution from mechanical and relay-based interlocking to computer-based systems. Early interlocking systems (mechanical, relay, and later electronic) require local signal interpretation and fixed block sections, limiting both network capacity and flexibility [19].

ERTMS is built around ETCS, which operates in multiple "levels" depending on how information is transmitted.

1. **Level 1:** Uses Eurobalises for spot transmission and may coexist with lineside signals.
2. **Level 2:** Adds a continuous data connection via GSM-R to a Radio Block Centre (RBC), removing dependence on optical signals and enabling real-time supervision.
3. **Level 3:** Removes track-based train detection and relies on train integrity confirmation, reducing infrastructure needs.

Norway's national program adopts ERTMS Level 2 across the entire network, using axle counters as train detection and preparing for future migration to Level 3. The system architecture in the Bane NOR program material shows a fully digital backbone connecting interlockings, RBCs, the Traffic Management System (TMS), and onboard ETCS equipment [20].

3.4 ETCS

The European Train Control System (ETCS) is a signalling and train protection element developed and included in the ERTMS initiative. ETCS provides movement authorities and intermittently or continuously supervises train speed and braking to ensure safe train separation. It replaces or complements national trackside signals by delivering standardized information to onboard equipment, enabling safer and more interoperable cross-border operation [18, 21].

3.4.1 Levels

ETCS is commonly described with levels that express how information is exchanged [21]:

- **Level 0:** Applies to trains equipped for ETCS, but there is no ETCS trackside. Effectively going back to regular control and legacy signalling.
- **Level STM:** Applies to trains equipped for ETCS, but runs on tracks with national system with ATP. Allowing ETCS to interface for the ATP.
- **Level 1:** Spot transmission with Eurobalises providing intermittent movement authorities and speed control while legacy signalling remain in place as shown in Figure 3.

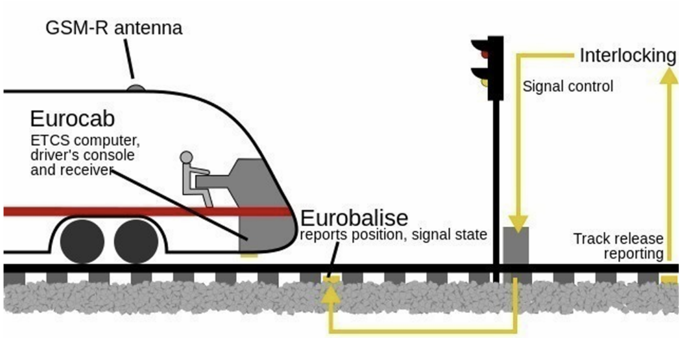

Figure 3: Train following regular signal with assisted speed and position with balises [21]

- **Level 2:** Continuous radio exchange to Radio Block Centre (RBC), typically via GSM-R or a successor. Movement authority is provided by the RBC. Eurobalises, if used, are primarily for precise positioning. Legacy signalling system are no longer needed and optional as shown in Figure 4.

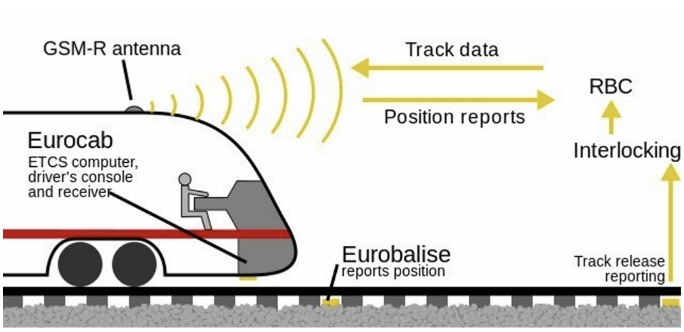

Figure 4: Train operating speed and signal with RBC with GSM-R signal [21]

3.4.2 Modes

ETCS defines different operational modes such as Full Supervision, On-Sight, Staff Responsible, Shunting, and Automatic Driving. Modes determine how the onboard equipment supervises movement authority and interacts with ATO where present. Automatic Driving mode is used when conditions for ATO are satisfied and ETCS provides the required movement and track data while Full Supervision is when ETCS is supplied with all possible train and track data needed are onboard [21].

3.5 Cellular communications

A very import part of latency for operating a train remotely will be cellular communication. It is extremely vital to have a stable and good source of communication between the remote operator and the train.

3.5.1 5G URLLC

5G is built up by OFDM, Orthogonal Frequency Division Multiplexing which divides "Spectrum" into small subcarriers. 5G is designed to support three different service categories, one of which is Ultra-Reliable and Low-Latency Communications (URLLC). URLLC is a communication service characterized by the need to successfully deliver packets with strict requirements in terms of availability, latency, and reliability. This capability is essential for providing connectivity to new services and applications from vertical domains, such as autonomous driving and factory automation [22, 23].

URLLC could be crucial for supporting emerging applications like wireless control and automation in industrial factory environments. For the specific use case of remote control, the required key performance indicators are an end to end latency of 5 ms and a reliability of 99.999%. Remote control applications often involve closed-loop control applications, like the use of collaborative robots in a factory, which require URLLC services [22, 23].

3.5.2 GSM-R

The Global System for Mobile Communications - Railway (GSM-R) serves as the unified radio communication platform for ERTMS, designed to take over from incompatible legacy analog systems with a standardized digital solution. Its operation is crucial for the railway system as it functions as the essential data carrier for ETCS, facilitating the continuous and real-time supervision of the train movement [24, 25].

GSM-R is built upon the commercial GSM standard, the system operates within a dedicated harmonized frequency band, with 876-880 MHz uplink and 921-925 MHz downlink, and utilizes a linear network of base stations to ensure continuous connectivity for trains travelling at speeds up to 500 km/h. GSM-R employs Time Division Multiple Access (TDMA) to organize radio resources and primarily relies on circuit-switched connections to guarantee dedicated bandwidth for critical transmissions [24, 25].

3.5.3 FRMCS

FRMCS or Future Railway Mobile Communication System is a product of International Union of Railways (UIC) to find a successor for GSM-R as it was getting older. FRMCS is built upon 3GPP standards with 5G URLLC and other more modern technology. FRMCS is set to replace GSM-R as the unified radio communication for ERTMS and will be a huge part of the future of railway technology [26, 27].

3GPP Rel-16 define video quality for railway operations. And is already used for setup of remote control by CAF. Mejias mention that the minimum requirements mentioned in the report is of H.264 codec, 320x240 resolution at 10 fps, but a recommended 1920x1080 at 30 FPS [3]. This will evolve as we get closer to acquiring FRMCS in the system used today.

3.5.4 Wi-Fi

Wi-Fi is a wireless communication technology following the standards of IEEE 802.11 that enables data transmission across networks. It operates across specific frequency bands such as 2.4 GHz and 5 GHz and the technology has evolved to support high speed data transfer and robust connectivity through advanced methods like beamforming, which strengthens wireless connections by targeting specific devices rather than broadcasting signals [28, 29].

In the domain of vehicular control, Wi-Fi functions by establishing Local Area Networks (LAN) that allow for the secure transmission of real-time control commands and audiovisual feeds, demonstrating in some specific single-access point configurations a lower median latency compared to LTE and 5G cellular networks. Although primarily associated with general data networking, the principles of Wi-Fi signal extension are relevant to railway logistics with the concept of relay chains. Wireless networks can employ repeating nodes to extend the operational range of remote monitoring and control systems to trains far away from the operator [28, 29].

3.6 ERTMS/ATO

Because of the international rollout of ERTMS, ATO will have to follow along. ERA has designed a solution to have an integrated ERTMS/ATO solution as you can see in Figure 5. It showcase how ETCS can be independent, but for ATO to be operational it is crucial for ETCS to guarantee safety.

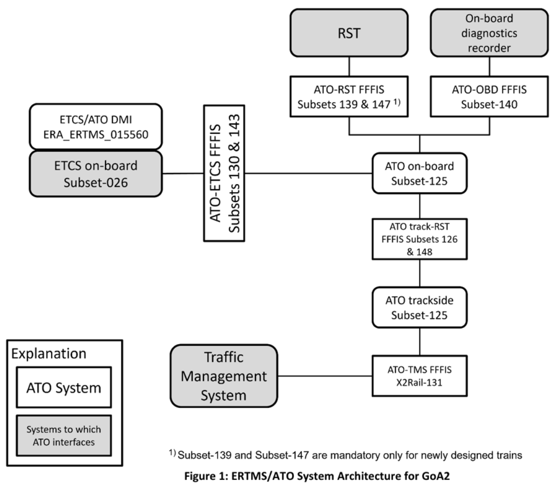

Figure 5: Architecture of ERA's ERTMS/ATO solution [18]

Figure 5 is showcasing different ATO working together with different responsibility and how they communicate with each other [18].

3.7 Latency

Latency is the time delay between when a command or data packet is sent and when it is received or acted upon. In railway automation, latency is critical for control loops, including brake initiation following an emergency command, and for ensuring the onboard ATO and remote systems remain synchronized.

The impact of latency, such as excessive one-way delay or variable delay can degrade braking calculations, slow down alarm propagation, and reduce the margin available for safe intervention. Real-time safety commands systems are designed with strict latency and reliability budgets and use prioritized and redundant communication channels.

3.7.1 Latency Measurements

There is many ways to measure and structure latency. Some of the most commonly referred to are:

- **Round-Trip Time (RTT):** Refers to the total time it takes for a data packet to travel from the source to the destination and back again to the source. It is commonly used to assess the overall latency of a network connection.
- **One-Way Delay (OWD):** Refers to the time it takes for a data packet to travel from the source to the destination in one direction. This measurement is particularly important for applications requiring real-time responsiveness, such as train control systems.
- **Glass to Glass (G2G):** Refers to the total latency from the moment a signal is generated until the corresponding action is observed. This measurement includes image capture, encoding, transmission, decoding and displaying.
- **End to End (E2E):** E2E can vary depending on papers, but most refers to the total latency from the initial source of a command or data packet to its final destination, including all intermediate processing and transmission steps. This measurement includes G2G as well as operator response time, sending control signals, transmission time and system processing delays.

Following these measurement structures, different protocols and tools exist to obtain one-way, two-way or path-level latency. Many of these originate from Request for Comments (RFC) publications maintained by the Internet Engineering Task Force (IETF), and several have been formalised further by the International Telecommunication Union (ITU) and the Institute of Electrical and Electronics Engineers (IEEE) [30, 31].

To present these methods clearly, these points summarises many relevant protocols for latency testing in remote operation research.

- **Internet Control Message Protocol (ICMP):** Supports latency measurement through Ping and Traceroute. Ping reports Round-Trip Time (RTT) by sending packets to a destination and measuring return time, while Traceroute identifies slow network segments by tracking each hop. These tools offer quick basic insights but cannot measure one-way delay due to lack of clock synchronisation.
- **One-Way Active Measurement Protocol (OWAMP):** Measures one-way latency using timestamped packets and synchronised clocks. It isolates directional delay, which is essential in remote operation where upstream delay (camera to operator) often dominates system performance.
- **Two-Way Active Measurement Protocol (TWAMP):** Extends OWAMP by measuring both directions independently. TWAMP provides structured two-way delay measurements with greater granularity than ICMP RTT and is widely used for validating network performance guarantees.
- **Real-Time Transport Protocol (RTP) timestamps:** Embeds timestamps into media packets, enabling measurement of latency from packet capture to display when combined with synchronisation (e.g., NTP or PTP). This makes RTP well suited for analysing video latency in teleoperation systems.
- **Network Time Protocol (NTP) / Precision Time Protocol (PTP):** These are synchronisation mechanisms used to enable accurate latency measurement rather than measurement tools themselves. NTP provides millisecond accuracy, while PTP enables microsecond-level precision, both of which are required to compute reliable one-way delays.

- **Time-Sensitive Networking (TSN) timing:** Provides deterministic timing behaviour in packet-switched networks. TSN ensures predictable delay and minimal jitter, improving the reliability of one-way and two-way latency measurements and making it relevant for real-time control environments.
- **Real-Time Streaming Protocol (RTSP):** A session-control protocol commonly used for video streaming. RTSP allows clients to request, pause, or resume video streams and is valued for predictable buffering behaviour, making delay easier to analyse.
- **Web Real-Time Communication (WebRTC):** A browser-based framework for low-delay peer-to-peer media transport. Uses dynamic path negotiation and congestion control, which can reduce delay but also introduce variability. Useful for real-time applications where minimal setup time is needed.
- **Real-Time Transport Control Protocol (RTCP):** Works alongside RTP to provide feedback on network conditions such as jitter, packet loss, and delay. Enables adaptive streaming and is essential when evaluating how delay evolves under degraded network performance.
- **Transmission Control Protocol (TCP):** A reliable, connection-oriented protocol that ensures packets arrive in order through retransmissions. TCP is not ideal for delay-sensitive control due to its congestion-control behaviour, but its stability makes it useful for evaluating baseline delay or control-signal reliability.
- **User Datagram Protocol (UDP):** A connectionless protocol with no retransmissions or ordering guarantees. UDP provides minimal transport overhead, making it suitable for real-time video and control testing. Its lack of reliability mechanisms also reveals how systems behave under packet loss.

3.8 Regulations

Railway automation must satisfy national and EU regulatory frameworks. In the EU, the *Technical Specifications for Interoperability* (TSIs) notably the Control-Command and Signalling (CCS) TSI define safety and interoperability requirements for ETCS, ATO interfaces and signalling subsystems. The European Union Agency for Railways (ERA) provides technical guidance, variables coordination and ERTMS documentation [7, 18].

National authorities, for example the Norwegian Railway Directorate and the Norwegian Railway Authority, implement national legislation, issue national safety rules, and specify how EU TSIs map to national processes where operators must demonstrate compliance with both national rules and applicable TSIs for approval and operation [32, 33]. For freight ATO trials (such as ATO-Cargo), project teams must prepare evidence on safety, human factors (remote supervision ergonomics), communication performance and conformity with the CCS TSI and national rules before trials are permitted [7, 8].

From ERA's [18] rules about during ATO operation, it shall be possible to:

1. Supervise train location by monitoring trains automatically using train identification and status (including delay information) to recognise deviations from normal operation as soon as possible;
2. Input the journey profile from the planning system;
3. Start the journey profile;
4. Dynamically modify the journey profile in real time to take account of changes in operating conditions including: disruption management; re-routing; re-timing.
5. Adapt the trail's journey profile to meet any update of the operational timetable;
6. Regulate trains to avoid bunching of trains and to reduce delays to trains in the case of disturbances;

7. Dispatch ATO trains to harmonise the starting of ATO trains, corresponding to results of train regulation and ensuring connecting services;
8. Operate both ATO and non-ATO trains simultaneously;

A way to guarantee these rules are followed and maintained is to implement a safety for when ATO is not working like intended, demanding a downgrade either to manual and physical steering or a option of remote control.

3.8.1 Spesific validation criteria

To complement the technical analysis of latency and communication performance, selected deliverables are included to illustrate how remote operation is currently being validated and how future requirements may be formalised. The documents WP41 and 3GPP TS 22.289 provide structured validation criteria and performance targets that are directly relevant for assessing remote train operation in safety-critical contexts.

WP41 - Validation Criteria for Tramways is a deliverable by FP2R2DATO which is a collection of test cases and validation criteria for remote driving of tramways. It is delivered by Sporveien Trikken AS with the assistance of CAF, UITP and FP2R2Dato in the Europe's Rail program and contains detailed plan and procedures for testing remote control of tramways. Which includes functionality of electric circuits implemented, TCMS, CERES actions, and safety supervision. Its a standardisation for defining Key Performance indicators and how they should be measured including: Solution QoE (1-5), Image Quality (1-5), Image latency (ms), Availability (Failure rate). Since remote control will come in the light and collaboration with Europe's Rail program. A similar design validation criteria should be expected and produced for railway infrastructure as well, which makes the WP41 a good resource on possible future regulations for RTO. [34]

3GPP TS 22.289 is a technical specification report. Completed by 3GPP in the works with UIC and European Telecommunications Standards Institute (ETSI), which among other list performance requirements for main lines[26, 35]. These requirements are set to ensure safe and good communications for the future of railway, but as mentioned in the report it is of now only target values. GoA3 and GoA4 is mentioned with the necessary high level of communications needed for operation. Here they also mention video-based remote control as a part of GoA3/GoA4 and put them in the same category for "Very Critical Video Communication".

Relevant information from table in 3GPP TS 22.289 about requirements for FRMCS can be seen in Table 5

Table 5: Regulations from 3GPP TS 22.289 Rel-17 [35]

Scenario	**End2End Latency**	**Reliability**	**Speed limit**	**Data rate**	**Traffic density**
Very Critical Video Communication	$\leq$ 100 ms	99,9%	$\leq$ 500 km/h	10 Mbps up to 20 Mbps	Up to 1 Gbps/km
	$\leq$ 10 ms	99,9%	$\leq$ 40 km/h	10 Mbps up to 30 Mbps	Up to 1 Gbps/km

In Table 5 we can see end to end latency being mentioned, and they have defined it as the same as G2G [35], measured at the communication interface, from the moment it is transmitted by the source to the moment it is successfully received at the destination.

The work of 3GPP, TS 22.289, is not yet considered in ERA's TSI CCS and therefore not completely applied. Yet it is still to be expected to be the regulations of future work and something to keep in mind under future development [7].

4 Related Work

Experiences from other similar fields, remote controlled cars, drones, cranes and other remote controlled railway vehicles. These experiences and papers tests viable aspects for the project of ATO, mainly focusing on the remote controlled

4.1 Railway Control

As showed in Table 3 there is only a handful of papers connected to Railway control compared to the rest. Therefore have I selected a few of the papers discussing latency, ATO and remote control of exactly trains.

4.1.1 ATO-Cargo Project

The ATO-Cargo project, led by the German Aerospace Center (DLR) in cooperation with DB Cargo AG, Digitale Schiene Deutschland (DSD), and ProRail B.V., focuses on developing and testing highly automated technologies for freight trains. The goal is to enhance rail freight efficiency by optimizing speed profiles, improving route utilization, and increasing competitiveness with road transport [8].

A key component of the project is the integration of an Automatic Train Operation (ATO) unit on locomotives in combination with the European Train Control System (ETCS) Level 2. This setup allows for real-time automation while maintaining human oversight. In case of system malfunctions or degraded operation, human operators at a Remote Supervision and Control Centre (RSC) can take over tasks such as remote monitoring, diagnosis, and manual control.

The project also emphasizes human factors engineering, ensuring that the RSC is ergonomically designed for operator efficiency and safety. For this project, researchers have employ virtual reality tools to simulate realistic control room environments and train personnel for future remote supervision tasks. Tests are being conducted on Betuweroute, a freight only railway, linking Rotterdam and the Ruhr region to validate the technical and operational readiness of this automation concept. The ultimate goal is to establish a European reference model for automated freight train operation [8].

4.1.2 Jürgensen 2025

The report by Jürgensen focuses on developing a functional remote control system for autonomous rail vehicles, using the REAKTOR project as the primary use case. Its main purpose is to enable safe human intervention during intermediate automation stages and to act as a fallback solution during fully autonomous GoA 4 operations. The report frames remote control as a crucial safety mechanism, both during prototype development and as an operational safeguard in future autonomous traffic scenarios [1].

The work develops a complete client-server architecture for remote driving, consisting of a Web Client for user interaction, a server coordinating data exchange, and a vehicle-side component interfacing with the autonomous controller. The system enables control over speed, and braking while providing a livestream from onboard cameras. The thesis additionally evaluates the type of cellular network needed for stable operations and includes a case study of network quality along the Malente - Lütjenburg track [1].

Overall, the paper concludes that a remote control system is essential for autonomous rail, but its reliable operation depends heavily on network latency and coverage. Full-scale deployment requires improved cellular infrastructure, ideally 5G, to ensure continuous low-latency transmission [1].

4.1.3 Kozarevic 2025

The report by Kozarevic aims to develop a remote control fallback system for autonomous rail vehicles, with a strong focus on identifying what functionality a remote operator must have in order to safely intervene when autonomy becomes unavailable. The main purpose of the work is to ensure that a human operator can reliably take over control in a downgraded mode. The thesis explores this by narrowing down required features and by analysing user needs, system safety conditions, and operational constraints for future automated railway environments.[2]

The work proceeds by starting from a broad set of functions that a driver normally performs and reducing them to the subset that is strictly necessary for remote operation. The thesis focuses on determining which controls, displays, and sensory feedback elements a remote fallback system must include and evaluates the role of human factors, describing how situational awareness degrades when control is exercised remotely. The thesis ends talking about how remote fallback operation is feasible but requires a carefully reduced function set tailored to remote conditions. [2]

4.1.4 Mejías 2024

The paper by Mejías et al. [3] examines how video streaming can support remote railway driving, focusing on keeping latency low enough for safe operation. The authors compare RTSP and WebRTC in a test setup that imitates an onboard camera system sending video over a 5G network to a remote operator. A key part of the work is a method for measuring true end-to-end latency by embedding timestamps directly into each frame. They also develop a simple adaptive bitrate mechanism that reacts to jitter and packet loss to keep the video usable under degrading network conditions. Tests in two laboratory setups show how latency grows as bandwidth drops and how bitrate adaptation delays freezing and pixelation under poor network conditions.

4.1.5 FP2R2Dato, EuropesRail 2024 D41.2

This deliverable D41.2 from Europe's Rail and FP2R2Dato project presents the testing results and assessments from the first remote driving demo for tramways. Its main purpose is to document how the system behaved during the initial demonstration phase, where auxiliary, static, and dynamic tests were performed on a modified SL18 tram in Oslo. The work focuses on verifying whether the new remote driving and command functions were correctly integrated, safe and aligned with the defined use cases. All tests were done with a structured procedure with reports, and acceptance criteria where each test case was evaluated as compliant or not. The document also includes KPI assessments such as image quality, latency, operator experience, availability, and productivity impact. Overall, the results show that the demonstrator reached the target and that the remote driving solution works reliably across the tested scenarios [4].

4.1.6 FP2R2Dato, EuropesRail 2023 D5.4 - Chapter 12

Chapter 12 in the deliverable D5.4 describes how remote driving should operate when faults occur in the remote control chain. The purpose is to define use cases for degraded remote control conditions, such as poor visibility, weak communication links, sensor failures, or loss of track information. These use cases outline how a remote driver should react and what operational limits apply when the supporting systems no longer provide full information. The chapter does not aim to solve the failures but to describe expected driver actions and system behaviour in each degraded mode [5].

The deliverable also describes allowed latency by comparing it to a previous drone test, where it does calculations of speed and creates a allowed speed limit depending on latency experienced [5].

4.1.7 Brandernburger 2023

Brandenburger et al. [6] present exploratory studies on how video quality affects a remote driver's ability to perceive information and react in time. The purpose of the work is to understand how bitrate, frame rate, and stimulus type influence perception accuracy and reaction speed when supervising trains through video feeds. The authors ran two small-scale studies where participants identified either light signals or distance markers in short video clips of varying quality. Results showed a clear trend: higher bitrate generally improved accuracy and reduced response time, while frame rate showed no meaningful effect. Distance markers were consistently identified faster and more reliably than light signals. Overall, the work provides early evidence that bitrate is the more critical parameter for supporting remote train operation.

4.2 Car Control

Remotely operation of cars has been thoroughly researched as it is one of the most common vehicles in the market. Still, remote operating them bring a lot of challenges and demands which makes it a ever continuing research objective.

4.2.1 Ouden 2022

The paper [36] investigates how remote driving performs over commercial 4G and 5G networks. Its purpose is to evaluate whether current mobile networks can support safe remote control of vehicles at low speeds. The authors design a complete remote driving architecture with vehicles, remote stations, fleet management, and a 5G mobile setup. They test the system both in Hardware in the Loop simulations and in real field trials using straight line braking and slalom manoeuvres. The study measures position accuracy, reaction time effects, and communication latency, comparing 4G and 5G performance. Results show that 5G halves network latency compared to 4G and that overall network performance is sufficient for remote driving below 40 km/h, although total reaction delay is still dominated by video processing and human reaction time rather than the network itself.

4.2.2 Jernberg 2024

Jernberg [37] investigates how latency, driving speed, and task type affect remote operation performance in a controlled simulator environment. The study exposes participants to rural and urban driving scenarios containing several hazards while systematically varying latency between baseline, +100 ms, and +200 ms. The aim is to understand how these factors influence reaction times, safety margins, perceived control, and driving behaviour.

The study measures several quantitative performance indicators such as reaction time from hazard onset to brake input, speed variation across the hazard segment, minimal distance to the hazard object, and post-encroachment time. These measurements are logged automatically in the simulator with fixed measurement windows before and after each hazard event. Hazards and environments are fully balanced so that each participant encounters all combinations of latency, speed, and task. The study also collects subjective ratings between each condition to capture perceived control, workload, and comfort. Comparisons between conditions are performed using mixed-design ANOVAs, allowing the authors to analyse how each factor independently and jointly influences the measurement values.

4.2.3 Kaknjo 2018

Kaknjo et al. [38] present a study on how to measure real-time video latency between a robot and a remote control station. The main purpose is to build a method that reliably captures one-way video delay with higher accuracy than existing tools, especially for applications that require low latency video feedback. The authors design a measurement setup using dedicated NTP servers, synchronized nodes, and software that generates precise visual events. Latency is measured by comparing timestamps from the event generator with the detection time in the received video frames. They test several configurations including web camera scenarios, PPS LED triggering, and network streaming over LAN and the Internet, while separating capture and display delay from pure network transport delay.

Their method enables accurate measurement of both one-way and total video latency in different streaming conditions, providing insights into how camera settings, encoding, and network type influence delay. They show that latency can vary significantly depending on protocol, encoding, and connection type, and that reliable measurement requires tight time synchronization.

4.2.4 Neumeier 2019

Teleoperation: The Holy Grail to Solve Problems of Automated Driving? Sure, but Latency Matters of Neumeier et al. [39], which is a highly cited and used report, investigate how latency affects human performance in teleoperated driving. The purpose of the study is to evaluate how different fixed and varying latency levels influence controllability, workload, and driving behaviour when a human remotely operates a simulated vehicle. The authors implement a static driving simulator with three large displays and inject controlled input and output delays to simulate realistic network conditions. Participants drive several predefined scenarios including following a lane, slalom, parking, and a long track with changing latency while the system continuously logs steering, speed and lane deviation.

Data collection focuses on quantifying behaviour through objective metrics such as Mean Lateral Position, Standard Deviation of Lateral Position, maximum steering angle, speed, and acceleration patterns. Latency conditions are compared using statistical methods to understand how increasing delay changes performance. The study also includes subjective workload measurements and questionnaires. In general, the work shows that higher latency increases workload and negatively impacts control, while smaller delays are easier to compensate for.

4.2.5 Kang 2018

Kang et al. [40] explore how remote control can act as a fallback when a self-driving system encounters situations it cannot interpret. The paper outlines typical failure cases, such as unclear road signs, malfunctioning traffic lights, or confusing construction layouts, and argues that a remote human operator can manage these situations when autonomy fails. Supporting the claim that remote operation is necessary for full automation. To investigate feasibility, the authors build a real-time video streaming test and measure how current LTE and Wi-Fi networks affect frame latency. In the report they vary resolution and bitrate across multiple settings while streaming compressed video from a mobile device to a remote server and back. Measurements include two-way frame delay, frame-loss rate, and how latency scales with frame size. The study provides reference values for achievable streaming performance under real conditions and highlights technical challenges that future remote driving systems must account for.

4.3 Drone Control

Drones are only remotely operated as it is a unmanned aerial vehicle (UAV) / unmanned aircraft system (UAS). Making research on latency or remote operating extremly relatable to RTO. Even though drones are controlled very differently and operated in a completely different environment, there is much to learn from the research done for remote operation.

4.3.1 González 2023

González et al. [41] study how live video for first person view (FPV) drone control behaves over LTE and Wi-Fi networks. The purpose is to assemble a full test and measure how telemetry, control, and video traffic perform under different connection setups. They build a quad-rotor platform, connect it to LTE through an interface module, and test three schemes: LTE server-based, LTE direct, and Wi-Fi direct. The authors record packet-level data, video quality, and latency to understand how each aspect influences FPV usability. They then define a QoE model combining VMAF video scores with frame-delay measurements, and use controlled packet loss and delay to evaluate how image quality and latency affect operator experience. This provides a structured method for assessing FPV performance in cellular networks.

4.3.2 Larsen 2022

Larsen et al. [42] examine whether a 5G network can support safe drone control when very low latency is required. The report looks at how delay is created as data travels from a drone, through the radio link, into the 5G transport network, and finally to a ground control station. To understand this, the authors break the network into segments and study how distance, number of switches, message size, and link speed each add delay. They then use these factors to calculate how far a drone can realistically fly from the control station while still meeting strict real-time requirements. The work also compares autonomous flight and manual remote control to show how their communication needs are different. In general, the report provides a framework for deciding how strong and fast the network must be, depending on how far drones are expected to operate and how much data they send.

4.3.3 Böhmer 2020

The report done by Böhmer et al. [43] investigate how open communication stacks can be used for remote drone control while still achieving low and predictable latency. The authors build a full test system around the Crazyflie nano-drone by extending it with a Raspberry Pi to enable Wi-Fi communication and by developing a bridge that forwards control messages between the drone and a ground station. They implement two different bridges, one in Python and one in Rust before they analyse how each contributes to overall delay in the control loop. Using detailed timestamp logging and cross-layer analysis, the paper breaks down latency into processing steps on the Raspberry Pi, UART transfer times, and the wireless path. This allows them to compare their open protocol stack with the drone's original proprietary radio link and determine whether the open solution is viable for control tasks.

4.4 Crane Control

Cranes are very different to railway, yet it has it similarities. As cranes are often also operated with joysticks and have certain set of possibilities in motion, it can be compared to trains in the aspect of how humans are effected mentally by remote operation.

4.4.1 Brunnström 2020

Brunnström et al. [44] study how latency affects the user experience when operating a forestry crane in a VR simulator. The goal is to understand how delays in the visual display and in the joystick controls influence task performance and comfort. The authors design three structured user studies where participants load logs in a VR environment using real crane joysticks and a head-mounted display. They systematically add different amounts of delay to the video update and to the joystick signals, and collect both performance metrics such as number of logs handled and subjective ratings of picture quality, responsiveness, comfort, immersion, and overall experience.

Simulator Sickness Questionnaire (SSQ) scores are also gathered before and after each session to evaluate discomfort. The work provides a controlled methodology for evaluating latency sensitivity in immersive remote-operation tasks.

5 Results and discussion

This chapter covers the result from the different research papers, and summarize some of the most relevant and important findings, as well as discuss their relevancy and application towards remote control for railway applications.

5.1 Human factors results

Human performance is a decisive part of remote train operation, because every control action is performed by a human operator who must interpret a video feed, react to system feedback, and compensate for delays. The results in this subsection focus on how operators respond to different levels and types of latency, and how stable versus unstable delay affects their ability to perform remote driving tasks. By comparing findings from railway, car, drone and crane studies, insight are gained into which latency ranges operators can adapt to, which conditions cause performance loss, and how these observations influence acceptable thresholds for remote train control. These insights constructs the basis for evaluating how much latency the system can tolerate before performance is affected.

In Table 6 is some of the previously introduced papers sorted and compared. Method and added effect show how the research has decided to set up the tests, like Sim = Simulation or IRL = In real life, while task define what the participants had to do. The results were many and varying for each research, but the number in the table is a specific number the paper decides represents a performance downgrade.

Table 6: Human evaluation of Signal Attributes and Latency

Vehicle	**Paper**	**Published**	**Method**	**Added Effect**	**Task**	**Slight performance downgrade**	**Major performance downgrade**	**Participants**
Railway	Brandernburger [6]	2023	Sim, Different bitrate and FPS	1, 6, 24 Mbps and 5, 15, 25 FPS	React to light signals or distance marker	6 Mbps	1 Mbps	31 subjects, Novice train drivers. Age 42 SD 19
Car	Ouden [36]	2022	Sim, Artificial latency	$0, 10 \pm 3$ ms, 20 ± 5 ms and packet loss	Drive slalom and parking	Packet loss at 1%	Packet loss at 2%	-
	Jernberg [37]	2024	Sim, Artificial latency	+100, +200 on base 89 ms	Adapt to Hazards	289 ms for lane bypassing	-	30 subjects, Experience drivers or gamers. Age 35 SD 11
	Neumeier [39]	2019	Sim, Artificial latency	+100, +300, +500 on base 97 ms	Drive slalom and parking	300 ms	500 ms	28 subjects, Experienced drivers. Age 27 SD 7
Drone	N. González [41]	2023	IRL, Artificial latency	+25, +50, +100, +200, +300, +500 on base 53 ms and packet loss	Fly drone	0.2% packet loss or 250 ms	0.3% packet loss or 300 ms	-
Crane	Brunnström [44]	2020	IRL, Artificial latency	+10, +20, +50, +100, +200, +400, +800 on base 80 ms	Offload lumber from truck	480 ms	880 ms	18 subjects, Experienced log lifters and novices. Age 40 SD 9

The columns of slight performance downgrade that was extracted for each paper are chosen from different values of evaluation. The research from Brandenburger [6] focus primarily on the results, measuring how fast and how accurate did the participants react to the stimulus of light signals and distance marker. Interesting founds from the report was that FPS seem to not affect the participant, at least inside the range tested, yet the bitrate was decisive for the results. Showing both higher accuracy and reaction time for higher bitrate. Another interesting result showed that the participant reacted better for the distance marker. This will be interesting forward with ERTMS as the only physical signalling on the track will be markers [20]. Jernberg [37] evaluate the participants by results like speed and deviation from lane, but also a self evaluation from the driver on their own performance when performing five different hazards:

- H1/P1: Car pulling over into your lane.
- H2/P2: Car crossing from opposing lane through your lane.
- H3/P3: Car with "yield" does not stop in crossing.
- H4/P4: Child runs into traffic from behind a bus.
- H5/P5: Bicycle in lane that driver needs to pass in opposing lane.

A finding in Jernberg was that reaction time in H1 increased more than latency added. And with the negative results from self evaluation the paper propose that higher latency reduce awareness and put a extra mental workload on the operator. But even with this reduction, a latency of almost 300 ms did not show any more problems and the final results was that with all the latency tested, that only H5 at 289 ms caused performance reduction in lateral lane position. A very similar result was confirmed by Neumeier et al. [39] where driving slalom results in participants leaving the car lane significantly more with higher latency, even with stable high latency. However as Neumeier writes "In the Parking scenario, even no differences for whatever latency could be revealed" [39]. Even Ouden et al. [36], with extremely low added latency compared to the other two car trials, did only notice a difference in the slalom driving, with then packet loss instead of latency. Ouden did not find reasonable difference between the scenarios with latency and packet loss in the straight acceleration test.

When comparing these result to railway operation, especially train control where there are no "Cars to pull over into your lane", is the values and results found viable? According to Jernberg [37], when performing a study with more naturalistic driving scenarios, speed and type of task is significant for results as well as latency. A realistic driving scenario for RTO would probably not include the tests and scenarios that resulted in the threshold for performance downgrade.

The other vehicles also include interesting results form their research. Both doing a thorough test with respectively eight and six levels of added latency and packet loss. Doing so many test is thorough but needs to be done carefully and planned to avoid the participants getting eased into the added latency, unless that is the purpose. Both the crane test done by Brunnström et al. [44] and drone test by González et al. [41] contained a evaluation form called Quality of Experience (QoE) which the papers used to evaluate remote control. Here the participants rate not their own performance but how they experienced the quality of control. Even with the same form for evaluation, the results vary between the different vehicles. For the remote crane operation a latency of 480 ms gave a slight reduction in both QoE and effect of the work at hand, but at 880 ms it was a major decrease in effect and operability [44]. While the drone evaluation of performance found added latency of around 180 ms start causing lower performance score and QoE, and steady deterioration with more latency, and ending up with the lowest grade at around 460 ms. Yet the drone test concludes with a E2E of 250 ms is viable for service usability, but draws the line at 300 ms. The same scores and QoE was measured to be around 0.2% and 0.3% packet loss.

5.1.1 Adaptation to latency

A topic for discussion in multiple of the research papers around remote control is the human ability to adapt to latency. Neumeier et al. [39] points this out to be one of the research questions for their paper. Even with a lot of test with different levels of mixed or stable latency and multiple runs, could they not confirm or deny any adaptation resulting in a stable latency being better than varied. The research by Jernberg et al. [37] also discuss stable latency, mentioning that the papers of Davis et al. (2010) and Gnatzig et al. (2013) found that a stable latency even as high as 500 ms was functional [37]. A observation was that remote operators adapt to circumstances, for example driving with safety margins to reduce risks. Yet Jernberg et al. could not confirm and conclude that this was adaption to latency alone [37]. As there was a lot of parameters tested and only up to 289 ms, the results of the test could also be credited to the differences in the individual participant, setting like rural or urban, or type of hazard.

5.2 Measured Latency for Different Vehicles

Different researchers have approached remote operation with their own measurement methods, system architectures and performance criteria, but the challenges of latency remain comparable. This subsection examines how latency has been measured across trains, cars, and drones, what numerical values these studies reported, and which thresholds they used to classify acceptable and unacceptable delay. By reviewing the technical measurement approaches, such as synchronized timestamping, RTT, G2G and E2E evaluation, the results illustrate how various industries establish practical latency measurements. The comparison in Table 7 provides a reference point for railway applications, showing how today's measurements are compared to thresholds and discussing how more research on a dedicated railway threshold is still needed.

Table 7 contains the papers including own latency measurements of remote control operation. Which method they use to measure, the threshold they decided to compare it to and where that threshold has its origin.

Table 7: Different Latency Measured and Thresholds used

Vehicle	Paper	Release year	Method	Latency found	Threshold used	Threshold Source
Railway	Jürgensen [1]	2025	RTT with traceroute	ave: 63 - 87 ms , spikes > 1000 ms	300 ms - 1000 ms	Ouden [36]
	Kozarevic	2025	G2G with screenshot of GPS synchronized digital clocks	ave: 212 ms, spikes < 400 ms	300 - 500 ms	Neumeier [39] and Jernberg [37]
	FP2R2DATO D5.4 [5]	2023	Analysation and approximation of numbers	-	500 ms	Drone
	FP2R2DATO D41.2 [4]	2024	G2G with synchronized phone clocks	ave: 340 - 380 ms	-	-
Car	Ouden [36]	2022	E2E, Video and Control, 5G. Time synchronized with a GPS-PPP source.	G2G: 95 ms average, 26 ms (95%ile) max: 500 ms	300 - 1000 ms	Neumeier [39]
	Jernberg [37]	2024	G2G with Voysys system	89 ms	300 ms	Neumeier [39]
Drone	González [41]	2023	E2E, Packet level	4 ms, 42 ms, 500 ms	100 ms - 500 ms	3GPP UAV requirements
	Böhmer [43]	2020	RTT with timestamps	4 ms, 9 ms, 18 ms for commands	-	-

Latency measurements reported in the reviewed studies vary considerably, not only because different vehicles and communication technologies were tested, but also because the research community does not use a single standardised method for reporting delay. This creates challenges when comparing results across studies, as terms such as E2E, G2G, and RTT latency capture fundamentally different parts of the video transmission.

Several studies report End-to-End (E2E) latency, but the definition of E2E differs. In some cases, E2E refers to only video transmission from the sender to receiver, while other include also image capturing and image display, making it the same as G2G. While in others it also includes operator reaction time, vehicle actuation delays, or additional processing within control software. Because the boundary of the measurement is inconsistent, two studies reporting "E2E latency" may in practice have measured processes of very different scope and it is therefore important to be precise and thorough when performing and describing the results. This is seen clearly in car research such as Ouden [36], where E2E was separated into video latency and control latency, compared to the drone studies where E2E typically includes only camera to display time, G2G. The G2G approach was utilized by Railway testing in D41.2 [4] with synchronized devices and recorded values of 340 - 380 ms. While G2G gives a clear representation of how the video stream is experienced by a human operator, it does not indicate the contribution of individual components such as encoding, network transport, or decoding. As a result, G2G measurements are useful for assessing operational feasibility but less suitable for diagnosing technical bottlenecks.

Round-Trip Time (RTT) methods, often used in robotics and drone studies such as Kaknjo [38] and Böhmer [43], measure the time for a packet to travel from source to destination and back again. RTT is simple to compute but does not represent how a remote operator perceives delay. Video transmission is not normally round-trip, and RTT values can overestimate or underestimate the latency relevant for remote driving. RTT also struggles with revealing the individual components effect.

As some of the papers has specified, they have taken in use synchronized clocks. These are essential for accurate measurements as even our computers deviate by some milliseconds [2]. Therefore more recent papers like Kozarevic [2], Ouden [36] Jernberg [37] and more have introduced this as to not produce inaccurate results of the their setups.

Because of these methodological differences, the numerical values reported across railway, car, drone, and crane studies cannot be directly compared without understanding the measurement technique used. For example, a 300 ms G2G latency during tram remote control represents a complete perception delay for an operator, while a 300 ms RTT measurement in a drone test may correspond to less than half that value in actual one-way visual delay. Similarly, E2E results that include human reaction time cannot be compared to E2E results describing only transmission delay.

5.2.1 Threshold for acceptable latency

Across the papers about remote control there are a vast amount of results for acceptable latency. The lowest discussed as a threshold was 0 ms for drone and 500 ms for crane and car. Yet the majority ends up with a threshold around the middle of those. After reading through research papers on the topic of how latency threshold are set, it became apparent that very few have done enhanced research on this very topic for specifically railway remote control. Papers such as Jürgensen [1] and Kozarevic [2], who mentions the maximum latency measured before loosing performance reference sources of other vehicles and their tests of remote control. At the same time, Neumeier observed that some tasks, particularly predictable or hazard-free ones were not significantly affected even at higher latency [39]. Which as discussed earlier is tasks expected of railway. Neumeier's research directly influenced later studies as Jernberg [37] explicitly chose +100 ms and +200 ms latency increases because Neumeier had identified 300 ms as a potential upper range for manageable control. However, Jernberg also found that even smaller increases could negatively influence operator performance in complex hazard situations, indicating that thresholds are task-dependent rather than universal. Even Jürgensen [1] decided to evaluate their results based on the results from research for remote car operation, Ouden et al. [36] which also had chosen its limits from Neumeier et al. [39].

In the release of FP2R2DATO D5.4 in Chapter 12 [5] about *Degraded modes specific to remote control*, a graph about speed threshold related to video latency for RTO is proposed as you can see in Figure 6.

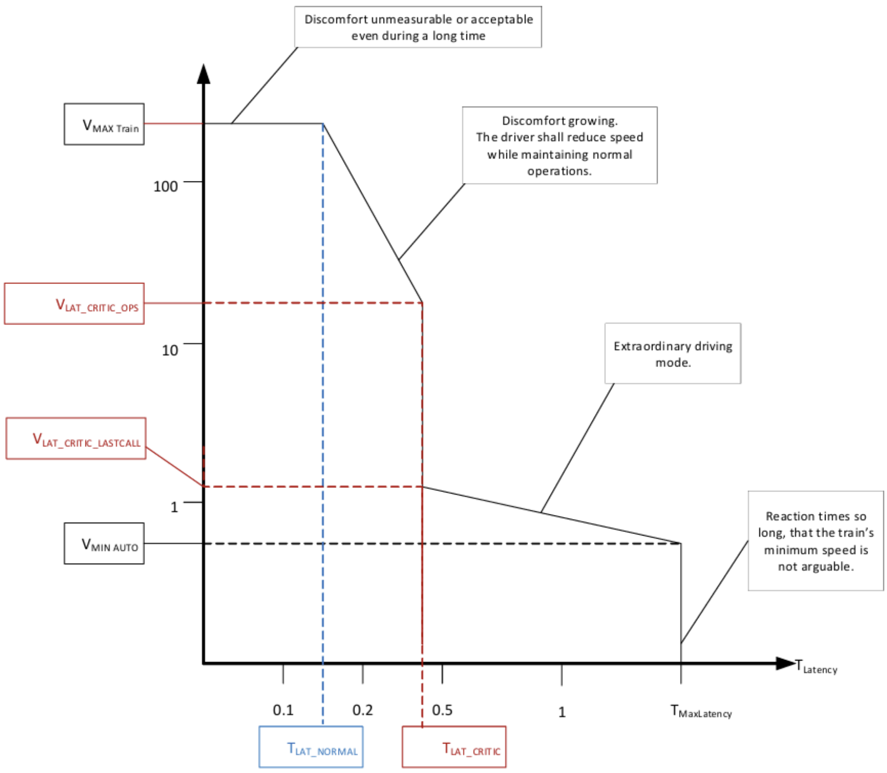

Figure 6: Maximum speed function of latency [5]

Where as the Y-axis is not directly speed in km/h but a scaled measure of the trains max speed accredited. Figure 6 suggest a steep reduction in speed at 150 ms, before at 400 ms that the train goes to the absolute minium speed possible [5]. These numbers are not directly calculated, but rather chosen from the understanding of the author with the influence of Table 7 that show driven distance at different speeds, but also result of latency threshold from drone operation.

Train speed (km/h)	Driven distance (m)									
360	10	20	30	40	50	60	70	80	90	100
250	6,94	13,89	20,83	27,78	34,72	41,67	48,61	55,56	62,5	69,44
160	4,44	8,89	13,33	17,78	22,22	26,67	31,11	35,56	40	44,44
100	2,78	5,56	8,33	11,11	13,89	16,67	19,44	22,22	25	27,78
50	1,39	2,78	4,17	5,56	6,94	8,33	9,72	11,11	12,5	13,89
30	0,83	1,67	2,5	3,33	4,17	5	5,83	6,67	7,5	8,33
10	0,28	0,56	0,83	1,11	1,39	1,67	1,94	2,22	2,5	2,78
5	0,14	0,28	0,42	0,56	0,69	0,83	0,97	1,11	1,25	1,39
4	0,11	0,22	0,33	0,44	0,56	0,67	0,78	0,89	1	1,11
3	0,08	0,17	0,25	0,33	0,42	0,5	0,58	0,67	0,75	0,83
2	0,06	0,11	0,17	0,22	0,28	0,33	0,39	0,44	0,5	0,56
1	0,03	0,06	0,08	0,11	0,14	0,17	0,19	0,22	0,25	0,28
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
	time (s)									

Figure 7: Driven distance function of speed [5]

This way to decide a threshold seem problematic. This is from Europe's Rail delivery but give no scientific grounds for the choices. Use Norway as an example where the max allowed speed for passing through a station is 40 km/h [45]. Easy formula from Norwegian railway operators Bane

NOR [46]:

$$b = \frac{v^2}{2r} + s$$

given the conservative example values of r (retardation) = 1 m/s and even ignoring the s (safety) distance, it results in 61.7 m of break length. While 150 ms of latency in video control will result in 1.7 m added length which is only 2.7% of total length. Without discussing this much further, this graph and recommendation of speed, seem very conservative, and set a maybe unrealistic and harsh threshold for railway operation.

5.2.2 Calculation of latency

The measurements and therefore the calculations of latency is not easily done. Since its a fluctuating measurement that varies a lot depending on the components in the system used, the way to calculate is very different depending on what the research aims to figure out.

Some studies, such as Larsen [42], estimate latency analytically by decomposing it into propagation delay, serialization delay, queueing, and processing. These models help explain why latency changes over long distances or across different network architectures, but they require detailed knowledge of the network path and still do not capture camera or display-related delays. Larsen express total communication delay as:

$$L_{\mathrm{e2e}} = nL_{\mathrm{proc}} + (n+1)L_{\mathrm{ser}} + L_{\mathrm{prop}} + L_Q,$$

with processing, serialization, propagation and queueing terms depending on distance, data size, and link rate. (n = switches along the network, n+1 = number of links, LQ = queuing latency)

A more simplistic method of E2E, which can be categorized as G2G, is shown in Kaknjo [38], who measures real-time video latency by timestamping the visual events in front of the camera (T_1) and event detected on the receiving end (T_2). The start of frame processing is denoted as (T_3):

$$L_{\mathrm{video}} = (T_2 - T_1) - (T_2 - T_3) = T_3 - T_1$$

This isolates video-processing delay but requires accurate timestamp extraction.

The common approach G2G latency is used for example in the tramway demo in D41.2 [4]. In that report G2G was measured by the time from light hitting the camera sensor to the moment the corresponding frame appears on the operator display. In its simplest form:

$$L_{\mathrm{G2G}} = T_{\mathrm{display}} - T_{\mathrm{capture}}.$$

Another widely used method is Round-Trip Time (RTT), which does not require synchronized clocks. Böhmer [43] computes RTT by sending a packet and measuring the time until the response arrives:

$$L_{\mathrm{RTT}} = t_{r1} - t_{p1}.$$

RTT is simple to measure but does not fully represent visual delay, since video streams are one-way.

All one-way methods depend on clock synchronisation because inaccurate clocks introduce measurement error. For this reason, studies such as D41.2 [4] synchronised devices with atomic references, while others rely on GPS time or avoid one-way measurement entirely.

Overall, because G2G, E2E, RTT, and video-latency measurements capture different parts of the signal chain, latency values from different studies cannot be directly compared without understanding which method was used. To complete a fully detailed study, the values of delay in transmission, capture and display should be measured. This will allow to compare to other test, but also to find bottlenecks in the system.

5.3 Tools and protocols for remote control

Remote operation depends not only on network performance but also on the tools and protocols used to capture, encode, transmit and display video. This subsection outlines the main technical options applied across the reviewed studies as well as encoding formats and measurement frameworks.This is done for the purpose of showing how different protocols influence latency, stability and video quality, and to compare their practical strengths and limitations for real-time control. By highlighting the variety of available tools and the results they produced in different vehicle domains, the subsection provides a foundation for identifying which approaches are most suitable for RTO.

Table 8: Tools evaluation of Latency

Vehicle	Paper	Published	Method	Tools	Major findings
Railway	Mejías [3]	2024	One-way, with NTP-synchronized timestamps	Adaptive Bitrate, RTSP vs WebRTC	RTSP $>$ WebRTC for multi-viewer setup. Adaptive bitrate allow lower bandwidth, but get higher latency
	Kozarevic [2]	2025	G2G via synchronized clocks + screenshot/OCR latency measurement	Voysys video system	Stable G2G $<$ 400 ms with Voysys
Car	Kaknjo [38]	2018	G2G latency with dedicated stratum-1 NTP-synchronized	TCP vs UDP, H.264 vs MJPEG	latency $<$ MJPEG, Bandwidth $<$ H.264 and UDP $<$ TCP for latency
	Kang [40]	2018	Two way video latency with timestamps	LTE vs WiFi, H.264 vs MPEG4	median latency: 50 ms -WiFi, 100 ms -LTE
	Ouden [36]	2022	One-way, time GPS-PPP + NTP sync, packets with timestamps	4G vs 5G	5G halves average latency
Drone	N. González [41]	2023	G2G, Video in front of camera	LTE server, LTEDirect WiFiDirect	500 ms for LTEs and 42 ms for LTEd, 4 ms for WiFi
	Larsen [42]	2022	One way, and change ethernet link rates and hops	TSN	Need high link rates, TSN reduce queuing delay
	Böhmer [43]	2020	RTT with inter packet timestamps	Crazyradio, Rust bridge vs Python bridge	RTT: Crazyradio $>$ Rust $>$ Python

Across the reviewed papers, remote-control video pipelines are implemented by combining a camera, an encoder, a transport protocol, and a timing or synchronization method. The choice of encoding format and protocol strongly affects latency and robustness, especially in cases where bandwidth becomes limited. A central distinction between the studies lies in whether the streaming system prioritises compression efficiency or minimal buffering. This is particularly relevant for remote train operation where predictable latency and visual clarity are required simultaneously.

One common finding is that H.264 remains the preferred codec for real-time operation due to its balance between compression efficiency and latency. This is reflected in studies on drones, cars, and trains, where H.264 is consistently used as the reference encoding configuration [3, 38, 41, 42]. The codec reduces bandwidth demands under constrained network conditions and maintains an acceptable video quality at relatively low bit rates. In contrast, MJPEG prioritises speed by encoding each frame independently. This reduces algorithmic delay but produces significantly higher bit rates. Kaknjo found that MJPEG achieved latencies up to 300 ms lower than H.264, yet required between 4.6 and 5 Mbps compared to 50 - 380 kbps for H.264 [38]. Although MJPEG offered the lowest encoding delay, the increased bandwidth made it more vulnerable to clogging the network suggesting that MJPEG is beneficial only under conditions with big and stable bandwidth, which may not be the case in rural railway environments. But is something to remember if that is to change in the future.

Several studies compared transport protocols intended for real-time applications. The analysis by Mejías evaluated Real-Time Streaming Protocol (RTSP) and Web Real-Time Communication (WebRTC), both of which use the Real-time Transport Protocol (RTP) as the underlying media carrier [3]. The paper introduced a detailed measurement procedure for E2E latency using Network Time Protocol (NTP) to synchronize the encoder and decoder clocks. The sender captured the timestamp at the moment of image acquisition, embedded it into RTP headers and preserved this metadata during encoding. The study showed that both protocols functioned well when bandwidth was sufficient, but once the link quality degraded their behaviour diverged. WebRTC delivered the image to the display faster but with greater visual instability, including short freezes and sections of pixelation. RTSP introduced slightly more delay, yet maintained a steadier image with fewer artefacts.

Mejías also demonstrated that bitrate adaptation driven by Real-Time Transport Control Protocol (RTCP) feedback prevented complete video collapse during bandwidth drops and enabled gradual recovery afterwards [3]. These results indicate that raw speed is less important than controlled degradation. For RTO where coverage gaps and fluctuating throughput are expected, protocols that degrade controlled and recover predictably may offer safer operational conditions than those optimised only for minimal delay.

Larsen investigated delay sources in video transmission over 5G URLLC systems and showed that most variation came from serialization and queueing in the network path rather than the application layer [42]. The work also demonstrated how Time-Sensitive Networking (TSN) reduces this variation by enforcing deterministic forwarding. For RTO this could mean that predictable network level is as important as an efficient streaming tool, since unstable forwarding alone can make control unreliable even when bandwidth is sufficient.

Kang evaluated how resolution, bitrate and network choice influence practical delay under Long Term Evolution (LTE) and Wi-Fi [40]. Higher resolutions consistently increased delay, and LTE produced more fluctuation than Wi-Fi. González also compared LTE and Wi-Fi, more specifically LTE server routing, LTE direct mode and Wi-Fi [41]. The study found large differences between the server-routed LTE who produced delays near 500 ms, direct LTE with 10 ms, while Wi-Fi was 4 ms. For RTO this shows that routing architecture, not only protocol will have an major impact on latency.

Kozarevic and Böhmer both relied on products developed by others, but in different ways. Kozarevic used the commercial Voysys teleoperation platform to enable full-scale train tests early in the project [2]. This allowed rapid iteration but limited insight into internal timing behaviour, which can make it difficult to identify where delay originated or to adjust components for railway. Böhmer instead used the open Crazyradio ecosystem with Crazy RTP and Predictably Reliable Real-Time Transport (PRRT), gaining full visibility and the ability to tune transport behaviour

[43]. For RTO, the comparison show that commercial platforms speed up deployment but reduce optimisation possibilities, while open research tools support detailed analysis but require significant redevelopment before they can meet railway safety and reliability requirements.

6 Conclusion

This study explored how latency can be measured and evaluated for remote train operation systems, and how acceptable performance thresholds can be established. By reviewing research across railway, car, drone and crane control, and by analysing human responses alongside technical measurements, the study provides a clearer picture of what aspects of latency matter most and how they should be assessed.

6.1 Evaluation of latency

The findings show that latency evaluation in train operation might need different methods than in other industries. To try answer RQ1, remote operation research on cars, drones and cranes often have different tests requiring movements, such as driving slalom or avoiding hazards, that train operation do not acquire. Making many of the latency's found in the research of other industries invalid or irrelevant. Only a few of the completed test to evaluate the latency can be comparable, these contain the setup similar to what a railway operator might encounter in different scenarios, like acceleration, keeping speed limits, deceleration and stops. Which is supported in Chapter 5 when discussing that the most comparable remote operation scenarios such as controlled braking, maintaining speed, and responding to visual cues behave consistently across different vehicle types, and can most likely be done with higher latency. Studies on car, drone, and crane control all indicate that human operators begin adjusting their behaviour once latency approaches 250 - 300 ms, even when they do not consciously perceive the delay. While the measured railway results, where G2G delays around 340 - 380 ms, were still sufficient for stable supervision tasks. Still there is many ways RTO research can learn from these industries. By examining setups, experiment formats and attributes considered when evaluating latency there is possibilities to replicate or adjust for RTO research. Also, the findings emphasise that performance is more sensitive to variability in latency than to the absolute value alone. This distinction is important for RTO, where predictable communication may outweigh the pursuit of the lowest possible delay.

6.2 Important aspects of latency

This study identifies the parameters that most strongly influence latency in remote train operation, addressing RQ2. Technical factors include encoding and decoding time, bitrate, packet loss, fps and network technology used. By being aware of the parameters affecting the total latency and how they interact together makes it possible to find individual limitations, identifying bottlenecks and ensuring progress by adjusting focus to problematic parameters. The results also show that the influence of these parameters is cumulative rather than isolated. For example, human factor studies demonstrated that lower bitrate and unstable video conditions had a stronger negative impact on perception and reaction time than latency alone, even when the latency remained within acceptable ranges. In a similar way, protocol comparisons revealed that mechanisms such as rate control, and encoding efficiency can significantly improve stability without altering the network itself. These findings reinforce that latency in RTO is not a single constraint but an interaction between network behaviour, video integrity, and operator workload. As a result, controlling these surrounding parameters becomes just as important as reducing latency, especially in critical tasks where clear visual information and stable feedback are essential.

6.3 Measuring latency

Just as other remote operation studies measure latency, train operation can also divide latency into one-way, RTT and G2G latency, and rely on synchronized timestamping to obtain accurate E2E measurements, in relation to RQ3. By doing this, train operations can explore different setups, tools and protocols. The best way to accomplish a detailed and specific research on latency is to have a standardised setup with interchanging hardware and protocols to compare each aspect of the full remote operation transaction. To be able to decide for an approach it is necessary to have a complete understanding. Since remote train operation is a project for expansion, the research needs to be detailed enough to be able to predict how extra trains, extra distance, and extra disturbances will effect the system. The measurements presented in Chapter 5 illustrate why consistent methodology is necessary for meaningful evaluation. The reviewed studies used timestamp based G2G and E2E techniques to separate camera processing, encoding, transmission, and decoding. This exposed how much of the total delay originates from video handling rather than the network itself. An insight that was confirmed by both railway and non-railway experiments. The comparison of RTSP, WebRTC, MJPEG, and H.264 further showed that measurement results can shift substantially depending on the chosen toolchain, even when the same network is used. This demonstrates the need for a unified measurement framework in RTO, where different setups can be tested under repeatable conditions and compared on equal terms. Establishing such a framework will be essential for future threshold validation and for ensuring that operation remain robust as new hardware, codecs, and communication systems are introduced.

Overall, the study shows that remote train operation faces similar latency challenges to other remote systems but benefits from the more predictable environment of railway infrastructure. While the literature provides useful initial thresholds, dedicated studies with trained railway personnel are still needed to define precise limits for safe train operation.

6.4 Future work

The results of this scoping review highlight several areas where further research is necessary to establish reliable operation and technical requirements for remote train operation. Although many experiences can be transferred from car, drone, and crane operation, the discussion shows that railway specific validation is still missing. The next steps should therefore focus on structured experimentation, standardised measurement frameworks, and system evaluations under realistic conditions.

A first priority is to conduct remote controlled experiments with trained railway personnel. Current thresholds are largely borrowed from cars and UAV studies, whose task demands differ significantly from train operation. Dedicated experiments using realistic RTO scenarios such as supervised cruising, signal interpretation, degraded mode interventions, and braking at defined reference points are needed to determine where latency begins to influence railway specific performance.

A second important step is to develop a standardised, repeatable measurement methodology for latency in RTO. Results from Chapter 5 demonstrate that differences in tools, encoding pipelines, and measurement procedures can heavily influence reported latency values. A unified framework covering G2G latency, E2E system latency, and bandwidth degradation would make it possible to evaluate different technologies on equal terms. Such a framework should be designed with future FRMCS and 5G URLLC requirements in mind, ensuring compatibility with evolving communication systems.

Further work is also needed on protocol and codec evaluation under relevant conditions. While current tests compare RTSP, WebRTC, H.264, and MJPEG, deeper analysis is required to understand how these tools behave over long distances, under varying load, or in the presence of packet loss typical for cellular networks along railway corridors. Adaptive bitrate strategies, and redundant streaming could play an important role in maintaining stable operator perception, and should therefore be tested systematically.

Finally, the practical implementation of RTO will require integration studies across the full ATO/ETCS/FRMCS stack. This includes latency contributions from onboard equipment, interlocking communication, server-based processing, and remote control interfaces. Understanding how these elements interact in a complete system is essential for determining whether the overall architecture can meet the target values from regulations, and where additional optimisation is needed.

Together, these future research directions will provide the foundation needed to validate latency thresholds, compare technical solutions, and support the safe deployment of remote train operation for future railway systems.

References

[1] Simon Jürgensen. 'Remote Control for Rail Vehicles'. PhD thesis. University of Kiel, 2025. URL: https://rtsys.informatik.uni-kiel.de/~biblio/downloads/theses/sij-bt.pdf (visited on 15th Oct. 2025).

[2] Emilia Salvesen Kozarevic. 'Developing Remote Control as a Fallback System: An Iterative Approach'. PhD thesis. NTNU, 2025.

[3] Daniel Mejías et al. *Towards Railways Remote Driving: Analysis of Video Streaming Latency and Adaptive Rate Control*. 2024. URL: https://ieeexplore.ieee.org/document/10597107 (visited on 6th Dec. 2025).

[4] Nacho Celaya Vela et al. *D41.2 - Testing reports and assessment*. CAF et al., 2024. URL: https://rail-research.europa.eu/wp-content/uploads/2025/02/D41.2-Testing-reports-assessment.pdf (visited on 7th Dec. 2025).

[5] Inzirillo Francesco et al. *D5.4 - Documentation of use cases for Remote Driving*. FP2R2DATO and Europe's Rail, 2023. URL: https://rail-research.europa.eu/wp-content/uploads/2025/07/D5.4-%E2%80%93-Documentation-of-use-cases-for-Remote-Driving.pdf (visited on 8th Dec. 2025).

[6] Niels Brandenburger, Friedrich Maximilian Strauß and Anja Naumann. *Effects of Video Quality on Perception and Reaction Time in Video-based Remote Train Control*. DLR, 2023. URL: https://elib.dlr.de/194929/1/PresentationHFES2023_DLR.pdf (visited on 5th Dec. 2025).

[7] European Union Agency for Railways (ERA). *Technical Specifications for Interoperability (TSIs)*. European Union Agency for Railways (ERA), 2016. URL: https://www.era.europa.eu/domains/technical-specifications-interoperability_en (visited on 20th Oct. 2025).

[8] German Aerospace Center (DLR). *ATO-Cargo – Automatic Train Operation Technologies for Cargo*. Institute of Transportation Systems, German Aerospace Center (DLR), 2025. URL: https://www.dlr.de/en/ts/research-transfer/projects/ato-cargo (visited on 8th Oct. 2025).

[9] Hilary Arksey and Lisa O'Malley. 'Scoping studies: towards a methodological framework'. In: *International Journal of Social Research Methodology* (2005). URL: https://www.tandfonline.com/doi/abs/10.1080/1364557032000119616 (visited on 9th Dec. 2025).

[10] John W. Creswell. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 4th. SAGE Publications, 2014. URL: https://www.ucg.ac.me/skladiste/blog_609332/objava_105202/fajlovi/Creswell.pdf (visited on 15th Oct. 2025).

[11] Tegan George. *Mixed Methods Research — Definition, Guide & Examples*. 2021. URL: https://www.scribbr.com/methodology/mixed-methods-research/ (visited on 15th Oct. 2025).

[12] Wendy Lim et al. 'Evidence-Based Guidelines — An Introduction'. In: *Hematology. American Society of Hematology Education Program* 2008 (2008), pp. 26–30. URL: https://www.ashpublications.org/hematology/article/2008/1/26/95823/Evidence-Based-Guidelines-An-Introduction (visited on 15th Oct. 2025).

[13] K. Bannigan and R. Watson. *Reliability and validity in a nutshell*. 2019. URL: https://doi.org/10.1111/j.1365-2702.2009.02939.x (visited on 14th Dec. 2025).

[14] I. P. Milroy. 'Aspects of Automatic Train Control'. PhD thesis. Loughborough University, 1980. URL: https://repository.lboro.ac.uk/articles/thesis/Aspects_of_automatic_train_control/9537395/1/files/17166803.pdf (visited on 27th Sept. 2025).

[15] CBTC Solutions Inc. *Grades of Automation (GoA) Explained*. 2023. URL: https://www.cbtcsolutions.ca/goa (visited on 1st Oct. 2025).

[16] International Association of Public Transport (UITP). *World Report on Metro Automation*. UITP, 2018. URL: https://www.uitp.org/publications/world-report-on-metro-automation/ (visited on 15th Oct. 2025).

[17] International Electrotechnical Commission (IEC). *IEC 62290-1:2025 -Railway applications - Urban guided transport management and command/control systems - Part 3: System requirements specification*. IEC, 2025. URL: https://online.standard.no/nb/nek-iec-62290-3-2025.

[18] European Union Agency for Railways (ERA). *ERTMS*. Tech. rep. European Union Agency for Railways (ERA). URL: https://www.era.europa.eu/domains/infrastructure/european-rail-traffic-management-system-ertms_en (visited on 9th Oct. 2025).

[19] Kjell Holter. *ERTMS and the History of Signalling, Part 1*. 2025. URL: https://www.banenor.no (visited on 10th Dec. 2025).

[20] Kjell Holter. *ERTMS and the Digital Railway, Part 2*. 2025. URL: https://www.banenor.no (visited on 10th Dec. 2025).

[21] Directorate-General for Mobility European Commission and Transport. *ETCS Levels and Modes*. 2023. URL: https://transport.ec.europa.eu/transport-modes/rail/ertms/what-ertms-and-how-does-it-work/etcs-levels-and-modes_en (visited on 9th Oct. 2025).

[22] S. Akhila and Hemavathi. '5G Ultra-Reliable Low-Latency Communication: Use Cases, Concepts and Challenges'. In: (2023). URL: https://ieeexplore.ieee.org/document/10112312 (visited on 22nd Nov. 2025).

[23] Zexian Li et al. '5G URLLC: Design challenges and system concepts'. In: (2018). URL: https://ieeexplore.ieee.org/document/8491078 (visited on 15th Oct. 2025).

[24] Aleksander Sniady and Jose Soler. 'An overview of GSM-R technology and its shortcomings'. In: (2012). URL: https://ieeexplore.ieee.org/document/6425256 (visited on 23rd Nov. 2025).

[25] Ruisi He et al. 'High-Speed Railway Communications: From GSM-R to LTE-R'. In: (2016). URL: https://ieeexplore.ieee.org/document/7553613 (visited on 23rd Nov. 2025).

[26] UIC. 'FRMCS Future Railway Mobile Communication System Telecoms & Signalling'. In: (2025). URL: https://uic.org/rail-system/telecoms-signalling/frmcs.

[27] Dan Mandoc. 'FRMCS Definition Specification and Standardization Activities ERA UIC FRMCS Definition'. In: (2019). URL: https://www.era.europa.eu/system/files/2022-11/4-3_dan_mandoc_era_uic_frmcs_definition_16102019_en.pdf.

[28] Oscar Amador, Maytheewat Aramrattana and Alexey Vinel. 'A Survey on Remote Operation of Road Vehicles'. In: (2022). URL: https://ieeexplore.ieee.org/document/9984654 (visited on 23rd Nov. 2025).

[29] Arvind Kumar Pandey and Warish Patel. 'A Smart Vehicle Control Remotely using Wifi'. In: (2022). URL: https://ieeexplore.ieee.org/document/10047780 (visited on 23rd Nov. 2025).

[30] International Telecommunication Union (ITU). *ITU-T G.1051 – Latency Measurement and Methodologies*. Tech. rep. International Telecommunication Union, 2023. URL: https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-G.1051-202303-I!!PDF-E&type=items (visited on 20th Oct. 2025).

[31] TechTarget. *Request for Comments (RFC) Definition*. 2024. URL: https://www.techtarget.com/whatis/definition/Request-for-Comments-RFC (visited on 9th Nov. 2025).

[32] Norwegian Railway Authority (Statens jernbanetilsyn). *Regulations on Vehicles on the National Railway Network (Unofficial Translation)*. 2013. URL: https://www.sjt.no/globalassets/00_generell/english/pdf-files/unofficial-translation-kjoretoyforskrift-en.pdf (visited on 20th Oct. 2025).

[33] Bane NOR. *Network Statement / Access Conditions 2025*. 2025. URL: https://oppslagsverk.banenor.no/en/network-statement/2025/access-conditions/ (visited on 20th Oct. 2025).

[34] Dusan Patrick Klago, Daria Kuzmina and Nacho Celaya Vela. *D41.1 - Collection of test cases and validation criteria*. 2023. URL: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e508af73fc&appId=PPGMS (visited on 8th Dec. 2025).

[35] ETSI. *ETSI TS 122 289 V17.0.0 LTE 5G Mobile communication system for railways 3GPP TS 22.289 version 17.0.0 Release 17*. Tech. rep. 2022. URL: https://www.etsi.org/deliver/etsi_ts/122200_122299/122289/17.00.00_60/ts_122289v170000p.pdf.

[36] J. Ouden, V. Ho and T. Smagt. 'Design and Evaluation of Remote Driving Architecture on 4G and 5G Mobile Networks'. In: (2022). URL: https://www.frontiersin.org/journals/future-transportation/articles/10.3389/ffutr.2021.801567/full (visited on 5th Dec. 2025).

[37] C. Jernberg, A. Sjöberg and P. Karlsson. 'The effect of latency, speed and task on remote operation'. In: *Transportation Engineering* 15 (2024). URL: https://www.sciencedirect.com/science/article/pii/S2590198224001386 (visited on 15th Oct. 2025).

[38] Admir Kaknjo et al. 'Real-Time Video Latency Measurement between a Robot and Its Remote Control Station: Causes and Mitigation'. In: *Wireless Communications and Mobile Computing* (2018). URL: https://www.researchgate.net/publication/329369713_Real-Time_Video_Latency_Measurement_between_a_Robot_and_Its_Remote_Control_Station_Causes_and_Mitigation (visited on 15th Oct. 2025).

[39] Stefan Neumeier et al. 'Teleoperation: The Holy Grail to Solve Problems of Automated Driving? Sure, but Latency Matters'. In: (2019). URL: https://dl.acm.org/doi/pdf/10.1145/3342197.3344534 (visited on 15th Oct. 2025).

[40] Lei Kang et al. 'Augmenting Self-Driving with Remote Control: Challenges and Directions'. In: (2018). URL: https://dl.acm.org/doi/pdf/10.1145/3177102.3177104 (visited on 5th Dec. 2025).

[41] N. González et al. *A quality of experience model for live video in first-person-view drone control in cellular networks*. 2023. URL: https://www.sciencedirect.com/science/article/pii/S1389128623005340 (visited on 8th Dec. 2025).

[42] Larsen et al. *Xhaul Latency Dimensioning of 5G Drone Control*. 2022. URL: https://ieeexplore.ieee.org/abstract/document/9836059 (visited on 8th Dec. 2025).

[43] Marlene Böhmer et al. *Latency-aware and -predictable Communication with Open Protocol Stacks for Remote Drone Control*. 2020. URL: https://ieeexplore.ieee.org/abstract/document/9183573 (visited on 8th Dec. 2025).

[44] Kjell Brunnströma et al. *Latency impact on Quality of Experience in a virtual reality simulator for remote control of machines*. 2020. URL: https://www.sciencedirect.com/science/article/pii/S0923596520301648 (visited on 7th Dec. 2025).

[45] SNL. *Jernbane*. URL: https://snl.no/jernbane (visited on 13th Dec. 2025).

[46] Bane NOR. *"Kapaistet på jernbane" or "Hastighet" or "Bremser"*. URL: https://oppslagsverk.banenor.no/ (visited on 13th Dec. 2025).