Deliverable 2.1

Moving Block Signalling System Test Strategy

Project acronym:	MOVINGRAIL
Starting date:	1/12/18
Duration (in months):	25
Call (part) identifier:	H2020-S2R-OC-IP2-2018
Grant agreement no:	826347
Due date of deliverable:	01/01/2020
Actual submission date:	29/02/2020
Responsible/Author:	Gemma Nicholson (UoB)
Dissemination level:	PU
Status:	Issued

Reviewed: yes

Document history		
Revision	*Date*	*Description*
1	30/09/2019	First draft
2	19/12/19	Second draft / UoB contribution
3	21/02/2020	Final draft incl Park contribution
4	29/02/2020	Final layout and quality check

Report contributors		
Name	*Beneficiary Short Name*	*Details of contribution*
Achila Mazini	UoB	Introductory and all testing literature review sections, gap analysis
Steve Mills	UoB	Legislation and safety sections
Marcelo Blumenfeld	UoB	Stakeholder workshop write-up and operational concept sections, document discussion
Gemma Nicholson	UoB	Document content and structure planning, discussion, review and revision
Bill Redfern	PARK	Operational concept and testing strategy
John Chaddock	PARK	Operational concept and testing strategy
Rob Goverde	TUD	Quality check and final editing

Funding
This project has received funding from the Shift2Rail Joint Undertaking (JU) under grant agreement No 826347. The JU receives support from the European Union's Horizon 2020 research and innovation programme and the Shift2Rail JU members other than the Union.

Table of Contents

Executive Summary

This document is D2.1 Moving Block Signalling System Test Strategy, the first of two deliverables of Work Package 2 of MOVINGRAIL, which is working to further the development of train-centric signalling systems, involving the investigation of both moving block and virtual coupling schemes. Work Package 2 addresses the important subject of testing approaches for such systems. The focus is placed upon developing testing methodologies for moving block signalling systems due to the significantly higher readiness level of moving block versus virtual coupling.

Work Package 2 can be divided into two parts based on its two deliverables. This first deliverable concentrates on presenting information on up-to-date stakeholder requirements for moving block systems testing and current capability, and on the state-of-the-art in control system testing. This information is used to develop an operational concept and high level testing strategy for migrating current railway control system testing capability towards the requirements for testing moving block systems, and is also presented in this document. The operational concept and testing strategy will be used to inform the development of a testing architecture, its interfaces and automated testing routines. These aspects are to be presented in the MOVINGRAIL WP2's second deliverable D2.2 Moving Block Signalling System Test Methods.

At a workshop held in May 2019 discussion facilitation sessions with stakeholders elicited requirements for a moving block system testing system. There were three structured topics: current testing process and the gap to zero-on-site testing, the format of a realistic moving block testing strategy, and ideas about how to bridge the gap between current testing capability and moving block testing. Attendees highlighted requirements relating to testing of additional safety requirements due to the closer running of trains, the need to develop a large number of testing scenarios in the pursuit of zero on site testing, the desire for automated testing routines and requirements for assurance of the safety of the final system.

The review of the literature covers mainly rail-specific control system testing capability, as well as some details for other transport modes' safety and conformity testing procedures. The legislative requirements for and relevant to rail testing procedures are covered. General software testing methodologies, categorised into testing techniques, testing levels and testing types are given as well as the generally accepted testing lifecycle, along with its specifics for existing ETCS systems. State-of-the-art testing tools and methodologies are reviewed, which includes an examination of relevant control system simulation models and methods. Gaps in current capability are identified and presented as a summary of the literature review and reflections of stakeholder requirements obtained in the workshop.

The operational concept for a moving block testing system is a high level description of what the testing system should do and why. The developed operational concept identifies the stakeholders and their roles in the testing process. The purpose of the system is formally set out, i.e. generation of system acceptance and trust to allow the new signalling system to be put into service. A list of functional aspects specifies what the system should do in more detail alongside success criteria and considerations for the test environment. The testing strategy outlines the general phases before specifying the specific implications for testing moving block and virtual coupling signalling systems. Finally, consideration is given to the automation of testing routines.

Abbreviations and Acronyms

Abbreviation/Acronym	Description
ATP	Automatic Train Protection System
BTM	Balise Transmission Module
DMI	Drive Machine Interface
ERTMS	European Rail Traffic Management System
ETCS	European Train Control System
EVC	European Vital Computer
HIL	Hardware-in-the-loop
IXS	Interlocking
KMC	Key Management Centre
LTM	Loop Transmission Module
RBC	Radio Block Centre
RIU	Radio Infill Unit
SUT	System Under Test
TCS	Train Control System
TIU	Train Interface Unit
BMC	Bounded Model Checking
CIT	Combinatorial Interaction Test case
CAV	Controlled Autonomous Vehicle
DNN	Deep Neural Network
GCG	Ground Communication Gateway
MCG	Mobile Communication Gateway
RBT	Requirements Based Testing
SRS	Software Requirements Specifications
TCMS	Train Control & Monitoring system
ZOST	Zero Onsite Testing

1. Introduction

This document is Deliverable 2.1 (D2.1) of MOVINGRAIL, and is part of WP2: Approaches for Testing Moving Block Signalling. The aim of MOVINGRAIL is the further development of train-centric signalling systems by introducing and applying a multidimensional analysis framework to assess train-centric signalling from the operational, technological and business perspectives. WP2 addresses the testing aspects necessary for the implementation of moving block signalling. The objective is to define improved strategies and methods for testing moving block signalling systems, including trade-offs between laboratory and on-site testing.

This document is divided into three parts. Part A elicits key stakeholder functional requirements, drawing from engagement with experts during a workshop; Part B contains a literature review of testing methods for relevant safety critical systems – which for the purposes of MOVINGRAIL are moving block signalling systems and precursors, specifically, ETCS systems, and positive control systems in complementary industries e.g. automotive and aerospace; Part C uses the knowledge acquired from the previous parts to develop an operational concept and a system strategy for testing moving block systems. The work reported in D2.2, the document following this one, will be of further investigation towards the definition of extensible architecture, interface definitions and automated testing routines for moving block.

1.1 Scope

Train control system (TCS) is a broad concept that can be described as a group of railway hardware devices and subsystems that monitor and manage the train based on its location, headway, speed and route, in order to prevent collisions and derailment [1]. Many types of train control systems have been developed and introduced to date [2][3][4][5]. These systems vary depending on the configuration they have been developed, and the type of safety solution used and can range from a simple automatic warning system to the most advanced communication-based train control system.

The European Rail Traffic Management System, also referred to as ERTMS/ETCS, was specified by the International Union of Railways (UIC) and several manufacturers and offers a uniform signalling system for seamless cross-border operations for high speed and conventional trains as well as freight hauling [6]. At level 3 ETCS may be a moving block system.

This review covers literature concerning testing of ETCS systems, i.e. the state of the art rail control system testing methods, alongside testing of control systems is relevant industries, which include aerospace, defence, marine, automotive, oil and gas, agriculture, mining and manufacturing. Hardware-in-the-loop (HIL) will form an essential part of ETCS testing; therefore this review also covers HIL architectures and testing in rail and in relevant industries. Legislation and regulations for testing train and other control systems shows which requirements must be satisfied and certified, and therefore is reviewed here alongside safety management requirements for control systems. Research that is working towards more efficient testing, in particular zero on site testing, is also covered.

The operational concept and testing strategy (Part C) are intended to be applicable to moving block systems in general, and have been developed with extensibility to virtual coupling in mind.

2. ETCS Background

Nowadays most worldwide implementations of ERTMS are of ETCS level 1 and level 2 systems [7]. Of these, ETCS level 2 is the one which can provide greatest improvements in capacity. However, to get further improvements from ETCS level 2 it means increasing the amount of trackside train detection systems, increasing the capital costs.

Conversely, ETCS level 3 can improve the capacity and flexibility of railways and reduce the infrastructure and maintenance cost with its potential to eliminate the need for signals and track-based train detection. It also allows an increase in capacity by replacing fixed-block with moving block signalling, which moves with the train and allows the train to proceed as long as it receives radio messages ensuring the safety of the track ahead [8].

The distinguishing characteristic of ETCS level 3 is the fact that train detection no longer necessarily relies on trackside equipment. Instead, the safety-critical task of determining whether a track block is occupied is performed by the on-board computer and the Radio Block Centre (RBC).

Currently, there are different implementations of ETCS level 3 depending on levels of maturity in terms of definition and development. The types of ETCS level 3 are summarised below [7].

Table 1. Types of ETCS L3

Type of level 3	Infrastructure	Benefits and Challenges
Overlay	Signals (Class B system); Trackside train detection retained; Use of Virtual blocks	Moderate increase in capacity for trains compared with operation using the Class B system. Solution needs to be found to allow ETCS L3 train to pass a lineside signal showing a stop aspect.
Hybrid (virtual blocks)	No signals; Trackside train detection retained; Use of Virtual Blocks	Increase in capacity for trains with Train Integrity Monitoring without adding trackside detection; Increase reliability because of redundancy in train localisation.
Hybrid (moving blocks)	No signals; Limited trackside train detection; Use of Moving Blocks	Increase capacity by adapting the size of the virtual blocks in software database; Impact on traffic management systems and operation impact (two trains at the same block) to be considered.
Virtual Block (without train detection)	No signals; No need for trackside train detection; Use of Virtual Blocks	Increase capacity by adapting the size of the virtual blocks in software database; Reduction of cost and increase in reliability due to the removal of trackside equipment. Solutions for trains without radio connection and degraded situation have to be found.
Moving Block (without train detection)	No signals; No need for trackside train detection; Use of Moving Blocks	Maximised capacity on the available infrastructure; Reduction of costs due to the removal of trackside equipment. Solutions for trains without radio connection and degraded situation have to be found.

Regardless of the different solution types for train control systems, all of them have the ultimate goal to improve the safety and train integrity, increase efficiency, capacity and provide a standardized and interoperable railway. Testing plays an important role to guarantee that the

quality and reliability of the system are been fulfilled.

However, applying testing techniques to train control systems can be considered a difficult and time-consuming task due to the system characteristics of high complexity and non-determinism [9]. Despite many approaches to testing available in the railway industry, up to now, ETCS testing and validation is a long process before being systems can be put into service [10].

One of the reasons is the reliance on costly and time-consuming field-testing, in addition to the fact that existing testing applications in TCS usually rely on manual test case generation and execution [11] [12].

3. Key stakeholder functional requirements

Moving block operations (ETCS L3 and equivalent) are inherently more complex than previous signalling paradigms, and thus can benefit substantially from the systems engineering approach taken by the MOVINGRAIL project. As an integral part of systems engineering processes, an operational concept (see Section 12 for a definition) is highly dependent on a good understanding of stakeholders' needs and requirements.

To obtain such an understanding, the consortium organised a workshop with key experts from industry and academia who are involved in railway signalling systems, including design and development, testing, and regulation. The workshop took place in London on 23rd May 2019, with invitations being sent a month in advance. As such, the workshop was attended by 40 specialists, and was divided into two parts. During the morning, MOVINGRAIL partners presented and discussed initial findings from other Work Packages (WPs) with the delegates. In the afternoon, experts were divided in groups to further discuss the functional requirements for testing ETCS L3 systems. Subsequently, delegates were given a questionnaire that they filled out individually.

Of the 40 participants, 23 fully completed the fifteen questions of the questionnaire, regarding (i) their current testing process and how far they are from zero-on-site testing; (ii) what a realistic ETCS L3 moving block testing strategy would look like; and (iii) how to close the gap between ETCS L2 and ETCS L3 testing. The full questionnaire and aggregated, anonymised results can be found at https://beardatashare.bham.ac.uk/getlink/fiNYac39GLAxPfS7s5WWRvi9/.

3.1 On current ETCS testing

The first part of the survey focused on identifying the current state-of-the-art in ETCS L2 testing systems. The first question was designed to understand the type of organisation the respondents worked for. Figure 1 shows that the more than half the participants (55.6%) work directly with some aspects of ETCS systems. Of those, approximately half are involved in systems integration activities. Almost one fifth of the participants answered 'other', meaning that they may not work directly with testing ETCS systems, but their area has an interface with the signalling system.

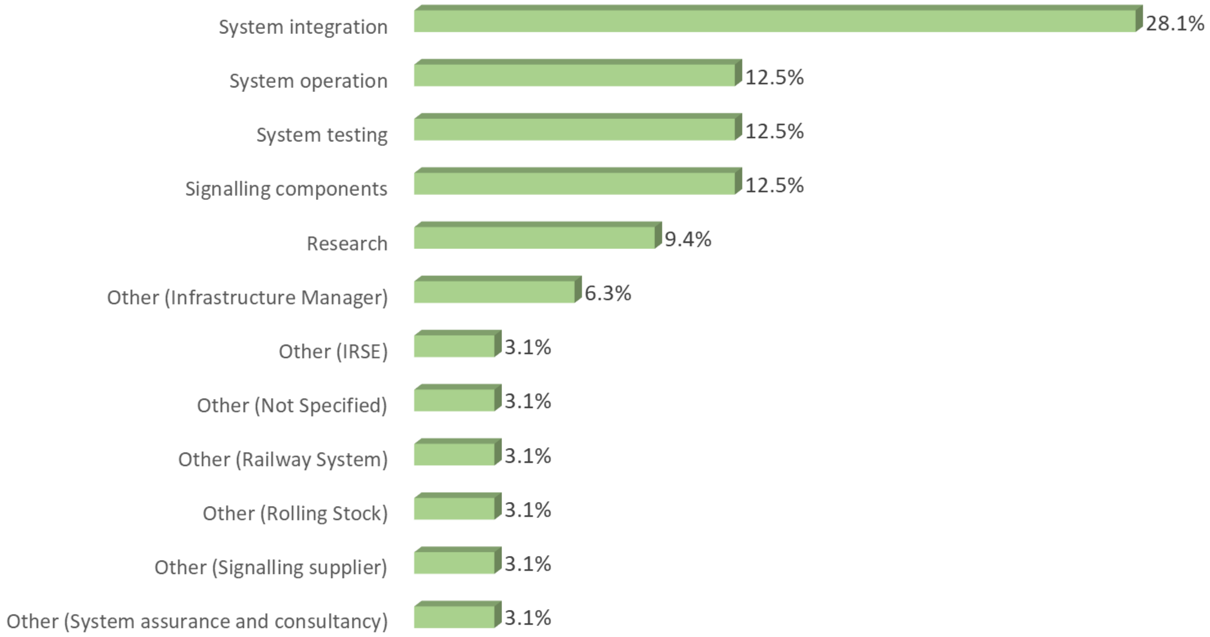

Figure 1. Type of organisation where respondents work

The second question (Figure 2) looked at the respondents' involvement with testing procedures. A slight majority (56.5%) work directly with testing procedures. These values align with the objectives of the workshop to engage with a wider set of stakeholders, while also having a relevant group of experts to discuss technical testing matters in further detail.

Figure 2. Participants directly involved in testing procedures

From there, the questionnaire changed the focus from the respondents to their respective organisations. The initial question of that section asked about the type of testing standards that their organisation adopted, and how established from on an ad hoc basis (project standard) to highly established (international standards). Figure 3 shows an approximately equal distribution between company, national, and international standards. It can be observed that the use of ad hoc standards is the least common practice among the respondents' organisations. While 24% of respondents' organisations use their own testing standards, almost 60% of them adopt standards that have been decided externally, either at national or international level.

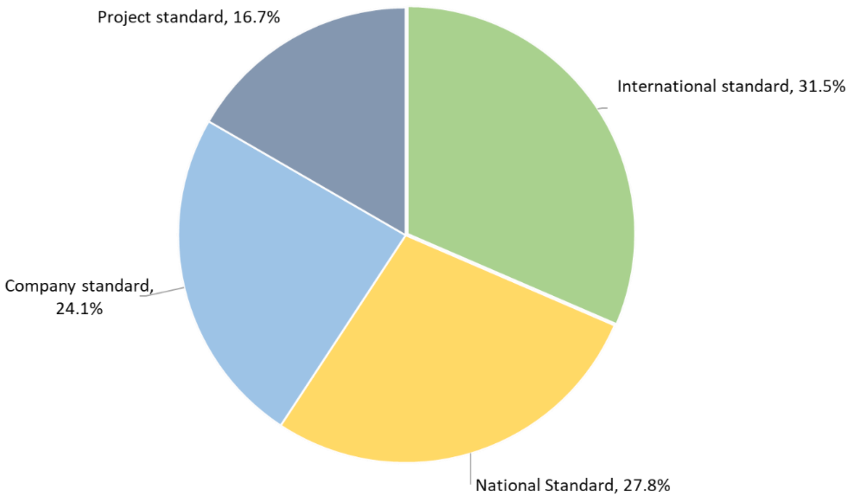

Figure 3. Types of testing standards adopted by the respondent's organisation

The following question delved into how test plans are defined within their organisations. More specifically, participants were asked about the objectives, types, and levels of their test plans. Around 42% of organisations have these plans prepared by the contractors, while close to 35% of them use test plans defined by the supplier. A smaller share, close to 15%, uses interactive or ad hoc processes with their clients.

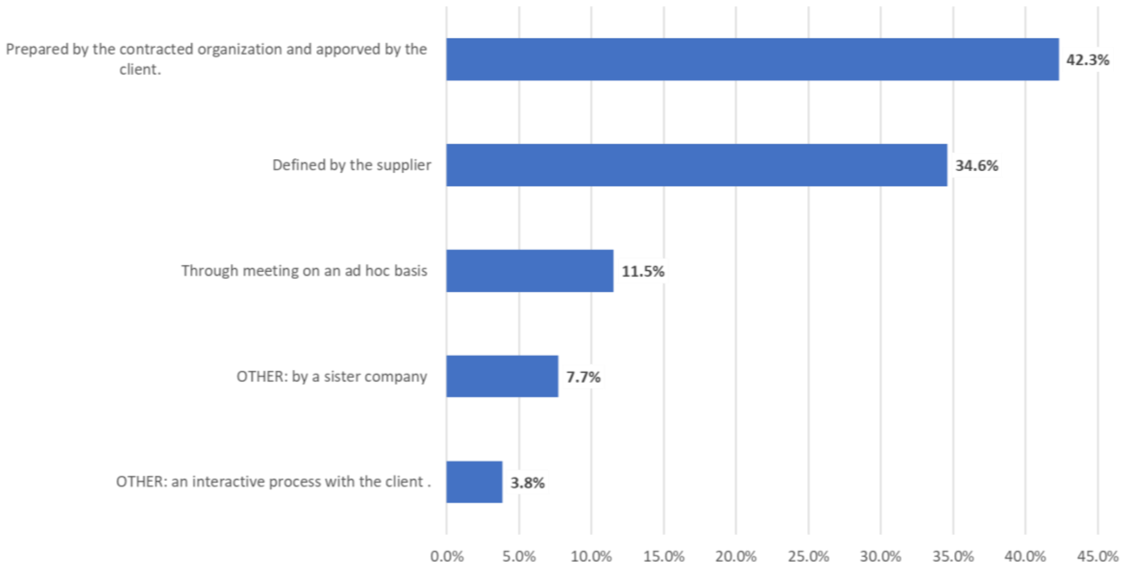

Figure 4. Responsibility for definition of test plans in respondents' organisation

Subsequently, when asked about the generation of test cases, the majority of respondents showed that the processes are still mostly defined manually (62.1%). Automation is in use, but at a smaller proportion (27.6%). A smaller share uses a combination of both, yet the underlying message is that test case generation for railway signalling systems is still mostly a manual activity.

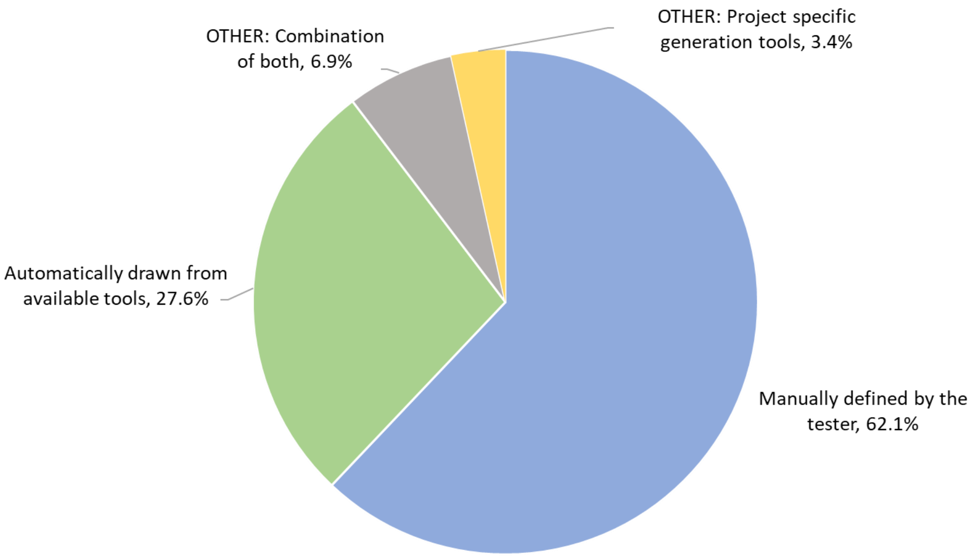

Figure 5. Generation of test cases in respondents' organisations

Question six looked at automation in test execution, and was divided in two parts. Firstly, experts were asked about the level of automation that their organisations use for test execution (Figure 6). Contrary to the findings in the generation of test cases, test execution involves some level of automation in the vast majority of respondents' organisations. While 9% execute tests manually, none do so in a fully automated manner.

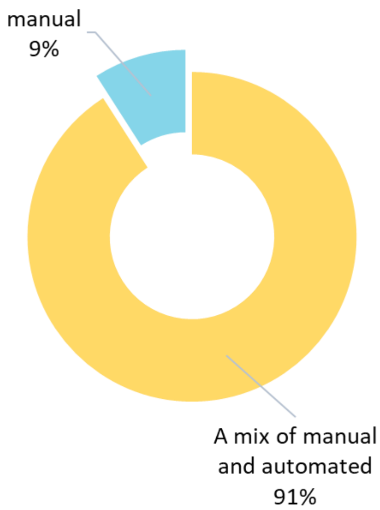

Figure 6. Level of automation in test execution

The second part of the question asked respondents about the tools used for testing, the majority (62%) uses those they have developed in-house (Figure 7). The smaller portion of COTS (commercial off-the-shelf) tools used might correlate with the greater proportion of manual execution.

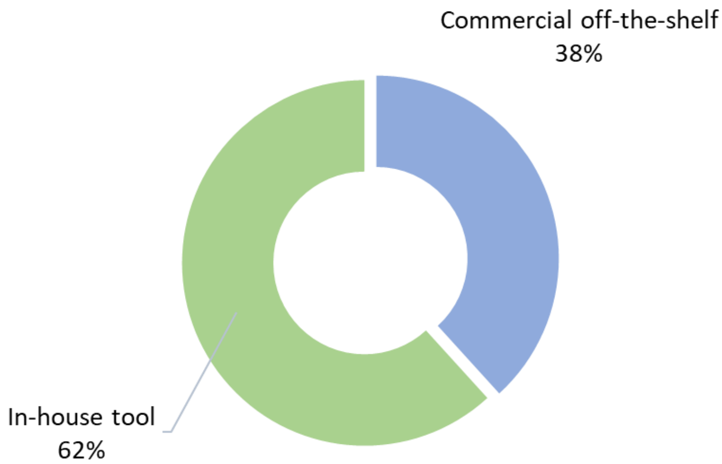

Figure 7. Type of testing tools used

The next group of questions looked at the analysis of test results. Firstly, respondents were asked on the level of automation used to analyse test results. To avoid possible confusion regarding the terms, result analysis was defined as the comparison between observed and expected results. Similar to previous results, the majority of the analysis processes are still conducted manually (73.3%) when it relates to the definition given (Figure 8).

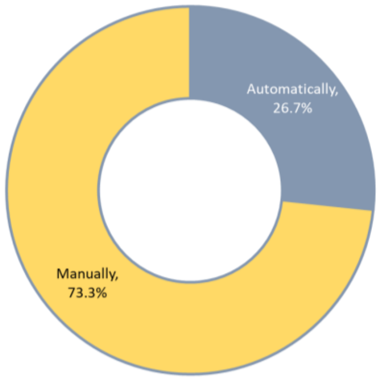

Figure 8. Level of automation used for the analysis of test results

Unsurprisingly, the process of documenting testing evidence when a test fails is also mostly done manually (Figure 9). Almost 58% of the respondents stated that their organisations send evidence as a manual bug report. Automated bug reports are used by only 3.8% of the organisations represented at the workshop. Furthermore, a proportion of evidence is not sent at all, but kept as in-house records. This is the case in approximately 35% of the organisations.

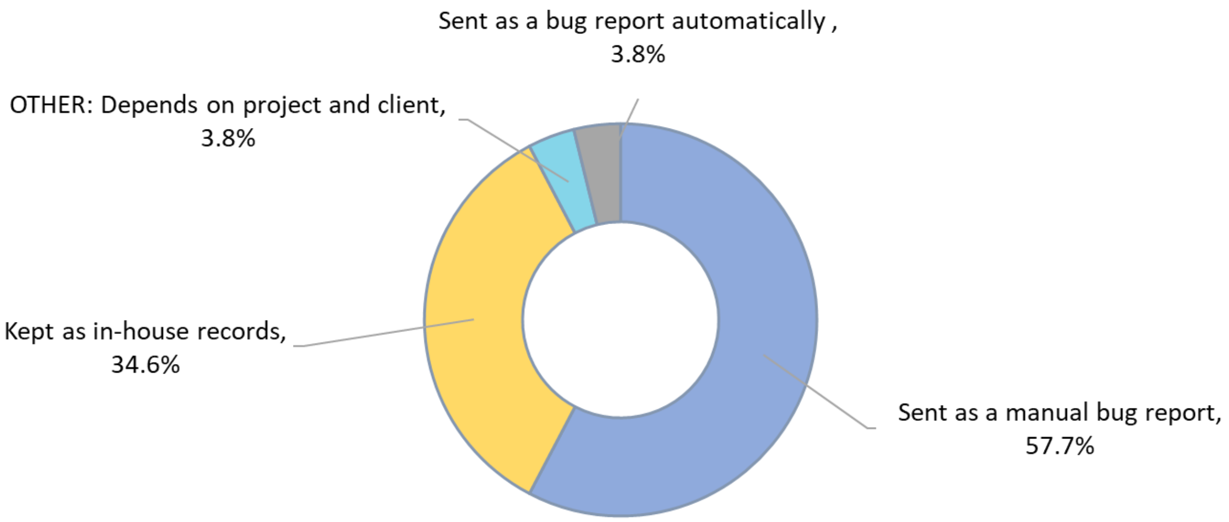

Figure 9. Conduct with testing evidence in case of failure

On a more general level, experts were then asked about what their organisation does with all testing artefacts and reports after the process is completed (Figure 10). In tune with the procedures for failures, where they are mostly sent as reports, the artefacts are also more commonly sent to customer, supplier, or regulating bodies (in approximately 48% of the cases). These are divided between those who send them to regulating/approval bodies (27%), sent to suppliers (16.2%), and a smaller share that sends to client or customer (5.4%). Nonetheless, the margin in this case is smaller than the previous, as an important 43% of the organisations keep the artefacts and records in-house. A very small minority sends summaries only to the client (2.7%).

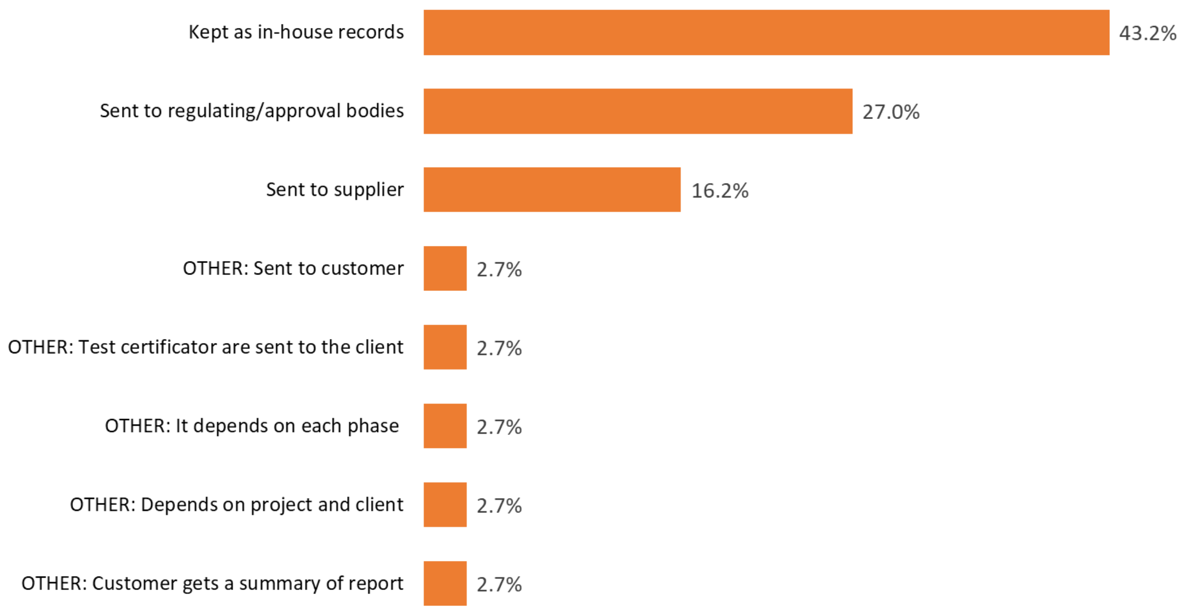

Figure 10. Conduct with testing artefacts and reports after the process is complete

The final question on current testing processes and procedures looked into the metrics used to assess the results (Figure 11). While generally divided, a slight majority (57.1%) of the organisations uses only pass/fail metrics to assess results. 42.9% of respondents alleged using performance metrics, yet further details on those were not collected due to time constraints. A

third option ‘other’ was given, but not chosen by any participant.

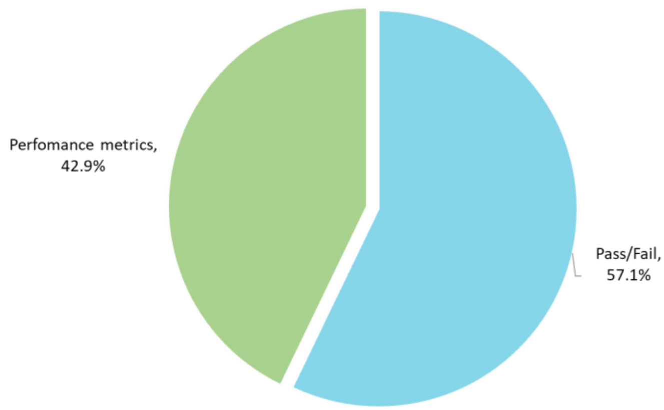

Figure 11. Metrics used for assessment of testing results

3.2 On ETCS L3

The second section of the questionnaire shifted the focus from the current systems in use to the experts’ view on future requirements for ETCS L3 / moving block. The first question asked regarded the type of ETCS L3 their organisations are thinking of adopting (see Table 1 for a description of each type). The results show that organisations are taking a conservative approach to the technology. Sixty percent of the organisations represented are adopting systems that include trackside equipment for train detection. Similarly, the use of virtual fixed blocks rather than full moving block seems to be prevalent among the respondents that have adopted one system (52.5% against 40% using moving block). One in eight organisations is planning to adopt an overlay solution which is the most conservative approach. Only around a quarter of respondents (22.5%) are working towards full moving block systems without any trackside train detection component.

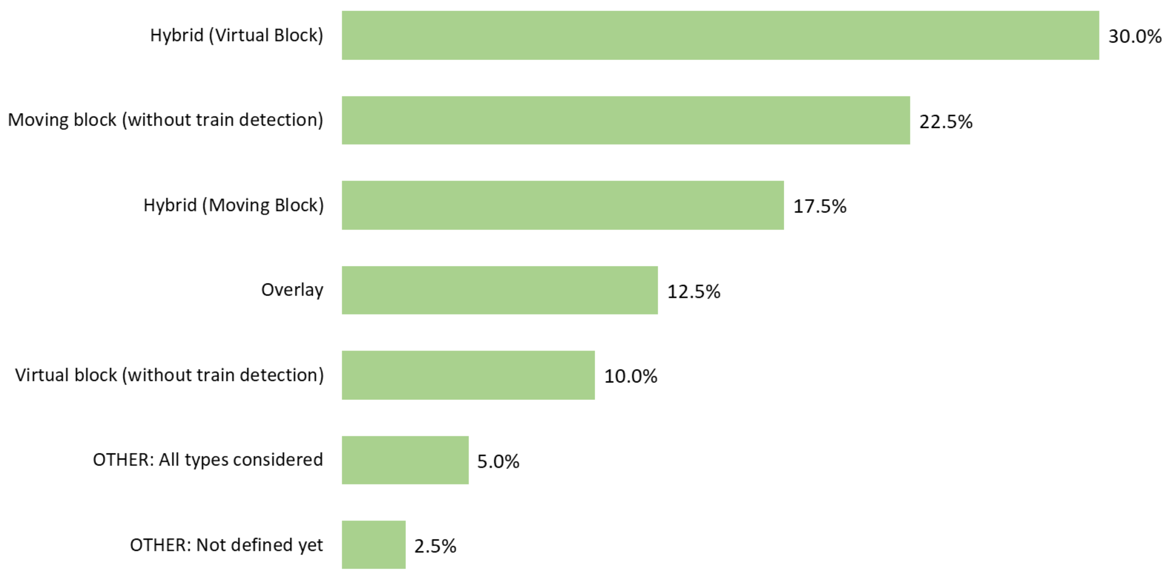

Figure 12. Types of moving block under consideration for implementation

The following part involved a qualitative analysis of the critical aspects of testing systems for ETCS L3. Two open-ended questions were asked to understand the organisations' plans for adopting and testing ETCS L3 systems. Firstly, participants were questioned over the solutions their organisations are planning to simulate ETCS L3 in their current test environments. Five main options could be derived from the answers, ranging in the use of new components and the share of off-site testing involved:

1. Current ETCS L2, with real equipment
2. Same as L2, focusing on checking the states of track occupancy on RBC and interaction with interlocking logic and TMS
3. Creating a joint test laboratory with the suppliers, using supplier's simulation software and third party consultants for controls
4. Lab system with real hardware, RBC, OBU embedded in a simulation rig
5. Hardware-in-the-loop using in house developed software

Experts were then asked about the critical aspects of testing moving block systems that require changes or adaptation in current test environments and procedures. Being an open-ended question, responses raised a wide array of issues that highlight various requirements for testing ETCS L3, but also included diverging views. For instance, one respondent stated that testing for moving block systems "requires a complete re-think of how it is done". On the other hand, another respondent views that there are "None or few [new environments and procedures] compared to CBTC environment and procedures".

In between these views, experts pinpointed some specific changes that should be put in place. Some functional requirements elicited mentioned safety and reliability assurance due to the closer running of trains, such as:

(i) The need to prove the system does not ever do what it should not rather than just proving that it will do what it should
(ii) Ensure safety also in degraded situations
(iii) Test reliability also with communication loss
(iv) How fast to recover from failures

Other answers highlight that the different operations under moving block create a much larger number of test scenarios. This would increase engineering efforts and costs, so it was unsurprising that respondents also mentioned the need to move towards zero on-site testing. Under these circumstances, experts raised several requirements such as:

(i) Negative testing is vital. Process need to be developed to convert as many "unknown unknowns" into at least "known unknowns".
(ii) Timing analysis is a key point of testing. Systems can work well on 'nominal' timings but then fail on small changes.
(iii) Similarly use of "parametric' rather than 'worst case' or 'nominal' test cases.
(iv) Testing TIM in different operational scenarios and different compositions.
(v) Black box testing of situations that even experienced testers may not think of.

Finally, participants were divided into groups to discuss two main questions on the way forward. The first asked participants to discuss what a realistic moving block testing strategy would look like, and how far are we from zero on-site testing. Some answers expressed a certain optimism in that ETCS L3 testing would not differ much from current testing, with the addition of some details related to new components. On the other hand, several answers emphasised specific aspects regarding safety and assurance. Comments state that test facilities should be able to prove that the system is safe, and there should be an emphasis on building confidence in lab-test compared to on-site testing based on body of knowledge. Moreover, the challenge of a much greater number of scenarios to be tested should focus on testing critical points of the system: (1) time for sending and receiving messages; (2) trackside equipment; and (3) outdoor systems such as radio communications. Participants were then asked to identify the gaps in testing requirements between current systems and ETCS L3. Figure 13 illustrates experts' perceptions during the workshop.

Requirements	Validation	Certification
(i) What do you do if no longer know if information is still valid? (ii) There is no non-volatile information anymore (iii) Needs to consider degraded operation	(i) How do you make sure that derivative requirements for i.e. SW match the requirements for the whole system? Including degraded operational scenarios.	(i) The hard thing is to make sure that you covered everything
(i) Failure mode scenarios arising from different operational approach.	-	(i) CBTC and GOA3-4 testing procedures and certificates could be used as a blue print for certification pathway.
(i) More interfaces, not mature compare to other.	(i) Confidence in testing data; (ii) Make sure test environment allows to test system realistically.	(i) CCS TSI, NOBO, ASBO, NSA ISA at product level. (ii) More emphasis on SW.
(i) Definition of stds; (ii) Ability to distinguish between "edge-cases" and common cases defined by previous experience and learning; (iii) Develop a "criticality model" to understand what OST is most critical.	(i) Need to test signaling and rolling stock; (ii) Being able to identify "edge cases in OST; (iii) Focus on V&V not just testing e.g. Analysis Dens, acceptable evidence..	(i) Being able to understand all configuration; (ii) Safety assurance and configuration management (iii) Simulator of simulator

Figure 13. Gaps in testing requirements between current systems and ETCS L3

As a conclusion, the main results from the survey highlight three important aspects for moving block testing. Firstly, responses emphasised the greater complexity of moving block systems, meaning that a wider array of possible scenarios need to be tested for the system to be deemed reliable. That leads to the second priority, where testing should be automated and conducted off-site whenever possible to avoid escalating costs. With it, standardisation and modularity should be also striven for. Finally, respondents recognised the greater safety criticality of moving block systems compared to their predecessors. Operating with shorter safety margins inevitably carries higher risks, so test strategies need to cover as many scenarios of degraded operation as possible.

PART B

This part is a literature review of testing methods, regulation, legislation, and safety management for moving block signalling systems and their precursors, in particular all levels of ETCS, and positive control systems in other industries.

4. Legislation and regulation for positive control systems

4.1 Rail legislation

Legislation covers ERTMS technical description and the process for putting it into service. There is additional legislation relating to the use of ERTMS.

The high level essential requirements are defined in the Interoperability Directive (2008/57/EC) [13]. The essential requirements are in force covering safety, reliability and availability, health, environmental protection and technical compatibility along with others specific to certain sub-systems. The essential requirements in summary comprise:

1. To ensure ERTMS equipment complies with the relevant specifications and to prevent additional requirements from undermining its interoperability, the EU Agency for Railways (ERA), should act as the ERTMS system authority. It is in charge of assessing the technical solutions envisaged before procurement of trackside equipment commences to confirm that the technical solutions are compliant with relevant TSIs (technical specifications for interoperability) and are fully interoperable.
2. Article 18 details the essential requirements for authorisation for placing in service of fixed installations. In the case of GB, the ORR (Office of Rail and Road) is the National Safety Authority (NSA) that verifies the completeness and relevance of the technical file and compliance of any proposed ERTMS system with the ERA decision. Once verified the NSA issues authorisation or advises rejection.
3. For renewal or upgrading of existing systems, the NSA in close cooperation with the ERA, shall examine the file and based on set criteria, decide if a new authorisation is required for placing in service.
4. Article 19 details the harmonised implementation of ERTMS in the European Union. This details the specific steps that are followed for an applicant wanting to implement a trackside control-command and signalling subsystem involving ETCS and/or GSM-R equipment.
5. Article 30 details the requirements for conformity assessment bodies (CAB) for assessing and certifying systems and sub-systems. A CAB that is notified for trackside and/or on-board control-command and signalling systems shall participate in, or ensure its assessment personnel are informed of, the activities of the ERTMS group (Article 29 of Regulation (EU) 2016/796 [14]).

EEC recommendation 2014/897/EC on matters related to the placing in service and use of structural subsystems and vehicles under directives 2008/57/EC [13] and 2004/49/EC [15] explains the procedure for putting ERTMS into service.

4.2 Rail regulation

Commission regulation (EU) 2016/919 [16]on the technical specification for interoperability to the 'control-command and signalling' subsystems of the rail system in the EU applies to all new, upgraded or renewed 'trackside control-command and signalling' and 'on-board control-command and signalling' subsystems as defined in Annex II of Directive 2008/57/EC [13].

In general, this applies to all rolling stock types with a driving cab that operate directly on the rail track on the trans-European rail network. It describes:

1. The functions that are essential for the safe control of railway traffic, and that are essential for its operation, including those required for degraded modes,
2. The interfaces,
3. The level of performance required to meet the essential requirements (safety, reliability and availability, health, environmental protection, and technical compatibility.

The control-command and signalling subsystems as defined include the following parts:

1. Train protection,
2. Voice radio communication,
3. Data radio communication,
4. Train detection.

The (Command Control and Signalling TSI) CCS TSI [16] provides an overview of the requirements, defines the parts of the control-command and signalling subsystems, and specifies the functional and technical specifications of the subsystems of ETCS. The safety characteristics relevant to interoperability are defined.

It is from this that the requirements to enable satisfactory assessment and certification are drawn. It is how the system and subsystem suppliers confirm compliance by providing objective and/or subjective evidence (through risk assessment using the Common Safety Method [17]) for these requirements that enables the Assessment Body, Designated Body and Notified Body to be satisfied.

Section 4 of the CCS TSI defines the characterisation of the subsystems, Section 5 defines the interoperability constituents, Section 6 the assessment requirements, and Section 7 implementing the CCS TSI.

In Section 5, Tables 5.1.a, 5.1.b, and 5.2.2.b define the basic interoperability constituents and the groups of interoperability constituents. These provide the requirements for assessment.

In Section 6, 6.1.2 outlines the principles for testing ETCS and GSM-R, 6.2.1 the assessment procedures for CCS interoperability constituents, 6.2.2 the assessment modules (options), and 6.2.3 the assessment requirements. Note that 6.2.4.1 details mandatory tests for on-board ETCS.

With their tables, Section 6.3.3 defines the assessment requirements for on-board subsystems, and Section 6.3.4 defines the assessment requirements for trackside subsystems. Also, it notes that in certain cases, some of the essential requirements may be met by national rules.

Regarding failure rates (which might need to be factored in for test simulation), Section 4.2.1.1 states for the hazard 'exceeding speed and/or distance limits advised to ETCS' the tolerable rate (THR) is 10^{-9} h^{-1} for random failures, for on-board ETCS and for trackside ETCS.

In 4.2.1.2 for availability/reliability, compliance with calculated values shall be ensured for:

1. Mean time between failure (MTBF) in hours of operation of a CCS on-board subsystem requiring the isolation of the train protection functions: (open point).
2. MTBF of a CCS on-board subsystem preventing radio voice communication between traffic control and the train driver: (open point).

To allow infrastructure managers and railway undertakings to monitor the level of risk in terms of the level of risk and the respect of the availability/reliability values used to define degraded situation management, the requirements for maintenance in Section 4.5 shall be used. Noting that there are no values specified, this is assumed to be left to the design authority, equipment manufacturers and system integration authority to specify. This, as well as open points, highlights the conflict that can arise between the EU member states in agreeing specific TSI requirements.

Notwithstanding this, UK HSE Guidance on Control Systems [18](see Subsection 5.1) defines MBTF safety integrity level figures which could be applied here.

4.3 Other rail regulation

Directive 2004/49 [15]creates a common European regulatory framework for safety and maintenance of safety management systems (SMS).

Regulation 402/2013 [17]is on common safety methods that provides a harmonised framework for the risk assessment process through hazard identification, risk analysis, and risk evaluation.

4.4 General Legislation and Regulation

4.4.1 Health and Safety at Work Act 1974

The UK Health and Safety at Work Act (HASAW) Act [19]places duties on employers and others to secure the health, safety and welfare of persons at work, and to protect the public from risks arising from work activities. Specific responsibilities are placed on employers, the self-employed, and employees. The act places duties on designers, manufacturers, importers and suppliers to ensure that equipment for use at work is designed and manufactured so as to be, so far as is reasonably practicable, safe and without risk to health when used, cleaned, and maintained, etc.

The duties laid down by this act have been amplified by more specific legislation, often in the form of regulations made under the act: some of these have also served to replace much of the pre-1974 legislation. The impetus behind some new health and safety legislation has been provided by the need to implement European Community Directives.

4.4.2 European Rules

The European Commission (EC) has introduced a number of directives on health and safety matters. Some of these lay down minimum requirements, which are intended to form the basis of harmonised workplace health and safety laws throughout the Member States of the EU. New regulations have been introduced in the UK to implement these directives, including: The

Management of Health and Safety at Work Regulations 1999 [20]; The Provision and use of Work Equipment Regulations 1998 [21]; and The Health and Safety (Display Screen Equipment) Regulations 1992 [22].

Other EC directives, sometimes known as New Approach Directives, aim to remove barriers to trade that may arise from different design and manufacturing standards among member states. The most significant of these is the Machinery Directive [23], which the Department of Trade and Industry has implemented in the UK as the Supply of Machinery (Safety) Regulations 2008 (as amended in 2011) [24].

4.4.3 Provision and Use of Work Equipment Regulations (PUWER)

The main objective of the UK Provision and Use of Work Equipment Regulations (PUWER) is to ensure the provision of safe work equipment throughout the lifetime of its use, regardless of its condition, age or origin.

The regulations require that machinery provided for use at work is:

- Suitable for its intended use,
- Safe for use – including keeping it maintained in a safe condition with regular inspections to ensure it is installed correctly and that its level of safety doesn't subsequently decline,
- Used only by people who have received adequate training, instruction and information,
- Accompanied by suitable health and safety measures, such as protective controls and devices,
- Used in accordance with specific requirements – mobile work equipment and power presses.

PUWER is comprised of 37 regulations and is split into 6 parts. Importantly, this regulation states that every employer shall ensure, so far as is reasonably practicable, that all control systems of work equipment are safe, and are chosen making due allowance for the failures, faults and constraints to be expected in the planned circumstances of use.

4.4.4 Connected and Autonomous Vehicles or Connected and Automated Mobility

There does not appear to yet be specific regulation or legislation in this area. This is largely assumed to be because this area is an area of research and development. However, the UK Centre of Connected and Autonomous Vehicles Code of Practice for Connected Automotive Testing [25] does refer. The Code is primarily intended to be used by organisations or individuals planning to trial or pilot automated vehicle technologies and services.

The European Commission Connected and Mobility Strategy [26] has in development policies and roadmaps, and is developing standards at a European level. What this means is that although there is not yet any specific legislation or regulation, then current road traffic laws and health and safety law will apply.

For control systems on-board existing 'standard' road vehicles such as cruise control, traffic lane monitoring, these are closed loop requiring intervention by the driver should any perturbation occur during a journey.

4.4.5 European Aviation Safety Agency (EASA)

Unmanned aircraft (drones) development has opened new opportunities in aerospace. Although unmanned aircraft offer a wide range of possibilities that manned aviation struggle with, the absence of a clear regulatory framework at EU level does not currently allow the exploitation of such opportunities. To benefit, the EU 2015 EU Aviation Strategy [27]creates a risk-based framework for all types of drone operations. This framework will ensure the safe use of drones in civil airspace and will create legal certainty for the industry.

The implementation is to ensure that future rules for unmanned aircraft are proportionate to the risk involved and to ensure that new developments are not hampered by unnecessarily heavy and costly rules and procedures. This means that the exploitation of such technology in civil aviation is at the early stages and is reliant on the technologies and regulations developed for defence applications including positive control systems.

4.4.6 European Maritime Safety Agency (EMSA)

The European Commission's objective is to protect Europe with very strict safety rules preventing sub-standard shipping, reducing the risk of serious maritime accidents and minimising the environmental impact of maritime transport; and works actively against piracy and terrorism threats.

An existing control system is the Maritime Information and Exchange System [28] (SafeSeaNet) providing information on what is happening at sea. It is a platform for maritime data sharing, linking together authorities from across Europe, to be the core of all relevant maritime information tools. It is composed of a network of national systems in the Member States and a central system acting as a nodal point, hosted and operated by the European Maritime Safety Agency (EMSA). Apart from enhancing maritime safety, port and maritime security, environmental protection and pollution preparedness, the system allows for the exchange and sharing of additional information facilitating efficient maritime traffic and maritime transport.

5. Safety Management requirements for positive control systems

5.1 HSE Guidance on Control Systems

This UK Health and Safety Executive Technical Measures document [29]provides relevant information about: control and safety related systems, system integrity, safety integrity levels, alarm systems and the human interface, protection systems, process control systems, and expert systems.

The key feature of this guidance is the integral safety management of a control system (such as ERTMS) and how to achieve this.

5.2 EU (Rail) Regulation 402/2013 Common Safety Method Risk Assessment

The EU (Rail) Regulation 402/2013 Common Safety Method Risk Assessment (CSM RA) [17] harmonises processes for risk evaluation and assessment and the evidence and documentation produced from this work. Application by one EU Member State enables acceptance in another with the minimum of further work under mutual recognition.

The CSM RA applies when any technical, operational or organisation change is being proposed to the railway system. A person making the change (known as 'the proposer') needs to firstly consider if a change has an impact on safety. If there is no impact on safety, the risk management process in the CSM RA need not be applied and the proposer must keep a record of how it arrived at its decision. This will not be the case for ERTMS.

ERTMS implementation has an impact and should be considered significant if the criteria in the CSM RA is applied correctly. The proposer must apply the risk management process. An assessment body then must carry out an independent assessment of how the risk management process is applied and the results from the risk management process. The assessment body must meet criteria set out in Annex II of the CSM RA. This includes meeting the requirements of the ISO/IEC 17020:2012 [30] and being accredited or recognised.

Notwithstanding this, if the proposer does not require the significant change to be mutually recognised in one or more other EU Member State the CSM RA allows the proposer to appoint an assessment body meeting relaxed criteria agreed by the national safety authority. This largely depends on whether the ERTMS system or constituent points are uniquely applied at the point of use or whether they can be exported elsewhere.

5.3 Aviation and Maritime Safety Management Systems

The Civil Aviation Authority document CAP 795 Safety Management Systems (SMS) guidance [31] for organisations refers. This defines SMS implementation, key components, policy and objectives, safety risk management, safety assurance and safety promotion.

It provides sufficient understanding of SMS concepts and the development of management policies and processes to implement and maintain an effective SMS. It applies to Air Operator’s Certificate (AOC) holders, continuing airworthiness management organisations, maintenance organisations, air navigation service providers, aerodromes and approved training organisations. This does not include buyers or suppliers of aircraft hardware or software. Safety requirements

are explicitly defined in the technical regulations and specifications for these.

6. Safety approvals for positive control systems in Europe

6.1 Rail

Document 2010/713/EU is Commission [32]decision on modules for the procedures for assessment of conformity, suitability for use and EC verification to be used in the technical specifications for interoperability adopted under Directive 2008/57/EC of the European Parliament [13]and of the Council refers for railway applications.

It details the various assessment modules that an applicant (contracting entity or manufacturer) can specify for a CAB to provide an assessment and certification for a subsystem or Interoperability Constituent such as ERTMS. The modules suggested in the CCS TSI 2016/919 [16]are:

- Either the type-examination procedure (Module CB) for the design and development phase in combination with the production quality management system procedure (Module CD) for the production phase; or
- The type-examination procedure (Module CB) for the design and development phase in combination with the product verification procedure (Module CF); or
- The full quality management system with design examination procedure (Module CH1).

In addition, for checking the SIM card Interoperability Constituent, the manufacturer or his representative may choose module CA.

Note that it is an independent, competent, Conformance Assessment Body that carries out the work.

6.2 Aviation

For newly developed aircraft model entering into operation, it must obtain a type certificate from the responsible aviation regulatory authority. Since 2003, EASA is responsible for the certification of aircraft in the EU and for some European non-EU Countries. This certificate testifies that the type of aircraft meets the safety requirements set by the European Union. [33]

The 4 steps of the type-certification process comprise Technical Familiarisation and Certification Basis, establishment of the certification programme, compliance demonstration, and technical disclosure and issue of approval.

Starting with the aircraft manufacturer presenting the project to EASA when it is considered to have reached a sufficient degree of maturity, the EASA certification team and the set of rules that will apply for the certification of this specific aircraft type are being established (Certification Basis). Definition and agreement is reached on the means to demonstrate compliance.

This compliance demonstration is done by analysis during ground testing (such as tests on the structure to withstand bird strikes, fatigue tests and tests in simulators) and also by means of tests during flight. EASA experts perform a detailed examination of this compliance demonstration, by means of document reviews and by attending some of these compliance demonstrations (test witnessing). This is the longest phase of the type-certification process. In the case of large aircraft, the period to complete the compliance demonstration is set at five

years and can be extended, if necessary.

If technically satisfied with the compliance demonstration by the manufacturer, EASA closes the investigation and issues the certificate. It is considered that the specific requirements for control systems for aerospace are detailed in relevant standards and regulations. Notwithstanding this, the technical knowhow is the intellectual property of the aerospace manufacturers and is not publicly available.

6.3 Automotive

UK government has published guidance on the required standards for vehicle safety in the form of Vehicle safety standards information sheets. [34]

These cover a plethora of safety requirements none of which cover positive control systems. This is covered in the Connected and Autonomous Vehicles Code of Practice (see above).

It is considered that the specific requirements for control systems are covered by automotive industry standards and regulations such that appropriate certification for vehicle roadworthiness can be issued. Notwithstanding this, the technical knowhow is the intellectual property of the automotive manufacturers and is not publicly available.

6.4 Maritime

The International Maritime Office and European Marine Safety Agency (EMSA) do not appear to have anything specific on positive control systems.

EMSA does have European Directives:

- Directive 1999/35/EC [35],
- Directive 2001/96/EC [36],
- Directive 2003/25/EC [37], and
- Directive 2009/45/EC [38].

7. Testing Methodologies and Techniques

There are several definitions for software testing in literature. According to [39] testing can be broadly defined as the process of executing a program with the intent of finding faults. Furthermore, a much more controlled definition suggested by [40], defined testing as a formal process, normally carried out by a specialized team in which a System Under Test (SUT) is examined by performing approved test procedures on approved test cases.

Although test methodologies and techniques may vary when taking into consideration different fields and purposes, it should be noted that testing can be organized into three main groups [40] [41]: (i) Testing technique; (ii) Testing Level; (iii) Testing Type.

According to the IEEE Standard for Software and System Test [41]; a testing technique or testing approach can be defined as a particular method that will be employed to pick the testing values. Table 2 below categorizes some of the general testing techniques.

Table 2. Testing approaches definitions

S.No.	Testing approach	Definition
1	White Box	This involves testing of a line of code, flow, logic, loop, structure, functions, class communication testing, etc. [42].
2	Black Box	Test how the system behaves for different inputs. It tests only for the main requirement of the product. It will verify the outputs for various types of inputs and checks if the system fails for any typical input [43].
3	Grey Box	Grey box testing stands between Black box testing and White box testing.
4	Positive Testing	Verify that a system conforms to its stated requirements. Typically, test cases are designed by analysis of the requirements specification document [40].
5	Negative Testing	Often used to test aspects that have not been documented or that have been poorly or incompletely documented in the specification [40].

Moreover, after the testing approach is defined, the testing level is categorized based on the stage of the SUT development lifecycle where testing takes place and each level has its specific testing objectives [44]. Some of the most common testing levels are explained in brief in Table 3.

Table 3. Testing levels

S.No.	Testing Level	Definition
1	Unit	Developers use this approach to test the SUT. Unit testing is the lowest test level and comes under white box testing.
3	Integration	The main goal is to verify and see how all software modules work when two or more modules are combined together. [43]
4	System	The overall system is tested to ensure that it is behaving as intended and as specified in the requirement document. [42]
5	User Acceptance	Focus on overall system features and functionality that are visible and reviewable by the customer. [42]
6	Smoke Testing	The initial testing process where the SUT is checked for its test-ready status [45]

Aside from the aforementioned testing levels, testing can also be categorized according to the different testing types that are used to test the software under different conditions. Table 4 categorizes some of the general testing types.

Table 4. Testing Type

S.No.	Testing Type	Definition
1	Regression Testing	Ensure that nothing is broken in the system after fixing bugs and testing bugs. [42]
2	Functional Testing	Verify that every functionality of the system is working as stated in the requirements. [43]
3	Usability testing	Determine the extent to which the SUT is easy to learn and operate [40].
4	Performance Testing	Verify how fast some aspect of a system performs under a particular workload [42].
5	Load Testing	Conducted to understand the behaviour of the system under a specific load which is mentioned in the requirement document. [46]
6	Security Testing	Check whether the system protects the information data and maintains functionality as intended [42].

7.1 Testing methodologies review and the potential of automation across the testing process

In order to maintain and increase certain levels of quality and safety in a system, testing is the most critical and complex activity in the development lifecycle of a system or product. Different development models are used for uncountable types of applications and usually, each company modifies the technique used in the testing process to conform with their specific need.

Many of these test methodologies can be applied to train control system testing and while they may differ in many ways, all of them follow a generic testing process, where the system under test goes through various phases that start with the requirement analysis and testing planning; followed by the test environment setup; test case design and selection; test execution; output evaluation and finally to the test closure, as shown in Figure 14.

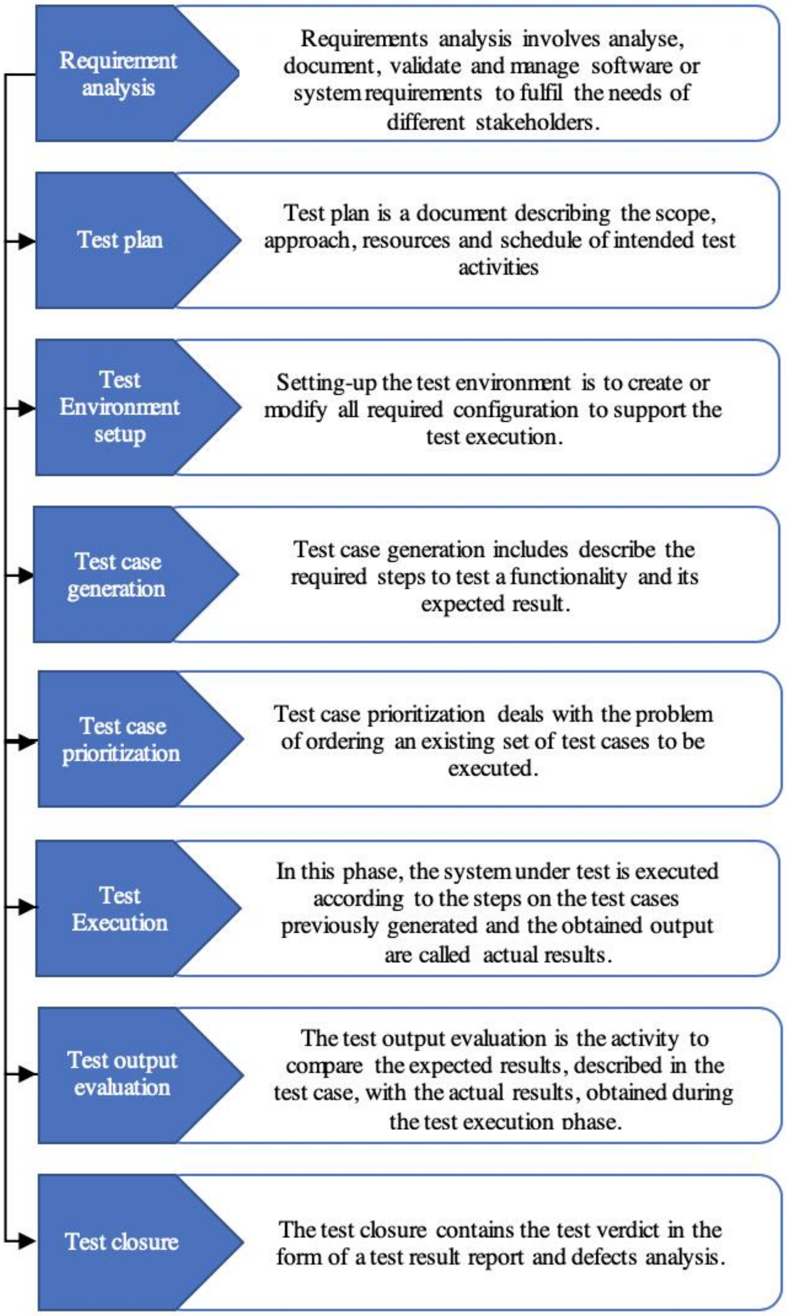

Figure 14. Test Lifecycle

The testing process can be analysed based in the methodologies used, but also regarding its level of automation. Automated solutions can be introduced across the testing process in order to make it more agile and efficient. An array of tools can be implemented to generate and select testing cases, control the execution of tests, perform the comparison of actual outcomes to predicted outcomes, perform the configuration of test preconditions and other test control and test reporting activities.

7.1.1 ETCS testing strategy

The ETCS standards SUBSET-110 [47] and SUBSET-76 [48] are used as reference to determine the guidelines to the ETCS test process and test environment. Many approaches for testing are available in the industry, but, ETCS testing and validation is a long process before acceptance into

service, relying on lab simulation in addition to costly and time-consuming field-testing. ETCS Level 3is still in its development phase and consequently, the documents [47] [48] only cover ETCS level 2. Although some guidelines, configurations and specifications through the entire test process will remain the same, some will need different levels of adaptation to fulfil the peculiarities of ETCS level 3. This transition will require changes in requirements for testing methods, especially moving towards robust zero on-site processes.

7.1.2 Requirement Analysis and Test Planning

The test preparation phase is the preliminary phase of the testing process and can be split into two main activities: (i) The requirement analysis; and (ii) The test planning.

During the analysis of functional and non-functional requirements, the testing team review and verify the requirements documents, in order to identify the expectations of the stakeholders and defined the priorities to serve as the foundation to the test plan [42]. On the other hand, the test plan outlines the scope, objectives, features to be tested, features not to be tested, types of testing to be performed, roles and responsibilities, risks and assumptions [49].

7.1.3 Test environment setup

Except for regression testing, which took place in a previous configured environment, normally during operation or maintenance, testing requires different levels of change applied to the test environment, which could take a lot of time to execute even a single test case. With this in mind, many authors define different solutions to automate the test environment configuration, installation and update.

Jahnsen et al., [7] developed two scripts that guarantee the system pre and post configuration. Firstly, prior to the test execution, a pre-installation script is used to validate the environment state before performing a new installation or update (i.e. verify disk space and software versions). Secondly, simple tasks are executed as required during the installation process (i.e. performing backup, moving files, creating new directories). The last step is to execute the installation, followed by a post-installation script which performs the necessary actions after the application installation has been completed. (i.e. reboot systems, de-fragmentation of the HD, delete old files, etc).

Kenner et al., [50] and Eirich [51] have also employed different installation and update tools which can be installed in windows embedded systems. Both tools are able to manage and monitor a testing environment and advise the user about newer versions available to update, verify the system state prior the installation and can perform the automatic download and installation of the required software into the system, ending by finally performing the necessary actions post-installation.

According to SUBSET-111 [52] the ETCS testing environment is represented into two domains: The test execution domain and the constituents domains. The former contains the necessary functionalities to simulate, control and monitor the constituents' environment during test execution. The latter contains the actual equipment under test (e.g. RBC, RIU and OBU) and the related adaptors units between the constituents and the test execution domain. The test architecture structure, adapted from [52], is represented in Figure 15.

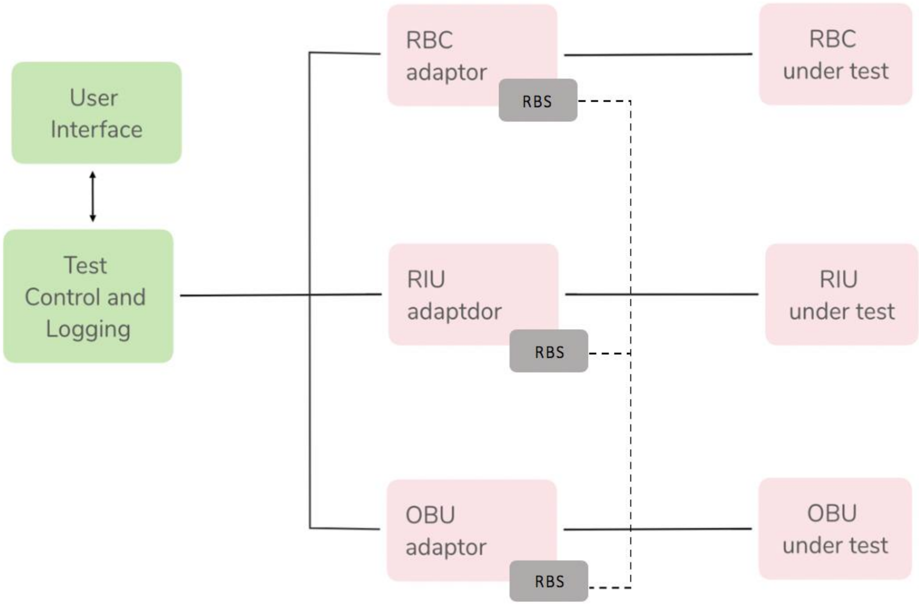

Figure 15. Generic test environment overview

The pink boxes indicate the constituent domains (e.g. RBC, RIU, OBU and respective adaptors) and although the figure contains a single representation of the constituents, it has the potential to be extended to include more than one constituent of each type. The green colour boxes represent the test execution domain by the boxes of the Test Control and Logging unit and the User interface. Finally, the RBS (Radio Block System), shown in the figure in grey boxes, represent the communication infrastructure between the constituents.

Many projects have discussed the use of Hardware-in-the-loop (HIL) system testing as an efficient solution to enable the test execution of software and hardware of complex systems such as Marine control systems [53] [54], Electric Vehicles [55] [56], Avionic Systems [45] [57] and Railway systems [58] [59]. HIL it is a well-proven solution utilise mainly as a type of black-box testing and can be defined by the operation of integrated real and simulated components. Many HIL simulations are in-house developments and the structure of its design and testing strategies vary according to the field and application purpose.

7.1.4 Test case preparation

A test case can be defined as a document that describes the required steps to test a functionality comparing its expected and actual result. In any testing process, the test case preparation includes the test case generation, test case selection and prioritization.

Although most testing teams are likely to use automated test execution tools, many still generate their test cases manually [60]. With this in mind, Gaurosi and Eleberzhager [61] and Thummalapenta et al. [62] presented similar solutions that utilize the existing and manually generated testing cases, in order to create scripts to enable the automation in the subsequent phases of the testing process. Conventionally, another commonly implemented solution is record-and-play, using one of many commercial and even open-source tools available (e.g. Selenium [63] and Gauge [64]), to record the tester interactions while performing the test case steps, in order to automate the test execution [65].

However, converting manually generated test cases into automated execution is still a time-

consuming task, In [66] Flemstrom et al., it is stated that automating the set of test cases in a totally random order increases the average manual effort required. Thus, the authors developed a similarity-based approach focusing on ordering the test case automation. Their tool considers predefined parameters to prioritize existing test cases, written in natural language, in order to convert them automatically into a computer language to achieve automated test execution.

However, those partial automation approaches rely on existing test cases and can still be time-consuming, require specialized skills, and produce fragile scripts [62]. To address these limitations many researchers suggested different solutions to automate the test case generation.

In order to create test cases from use cases automatically, Chen and Li [67] proposed a state-based model, called Interaction Finite Automation (IFA), derived from classification of interactions between actors and a system without knowing internal structures of systems. In their approach, the use case scenarios can be specified with natural language, tables or diagrams. Based on the pieces of integration presented the IFA is considered as the input to generate test cases and the prototype software called ATCGT (Automated Test Case Generation Tool) can be used to generate the test cases.

Conversely, Keränen and Räty [9] used Model-Based Technique (MBT) to automate more than just test execution. In this approach, the test cases are derived from an abstract model describing the functional behaviour of the SUT. Usually, there is an infinite number of possible tests that can be generated from a model, in this approach testers are able to choose test generation directives or select criteria to limit the number of generated tests to a small finite number by e.g. selecting highest-priority tests or by ensuring specific coverage of model structures or of other artefacts such as requirement documents.

Faucogney [68], also utilised the Model-Based Testing (MBT) method to generate system test cases into a train control system for both hardware and software. In this work, the test cases were generated according to a model based on the SUT functional requirements. However, the MBT approach proves to be limited to test complex SUTs, due to the computer computational limit when presented to some model complexity.

To overcome the difficulties of applying MBT methods to test complex TCSs, Wang [11] developed an effective simulation combined MBT platform, presenting an integrated testing platform for automating train control system functional testing in an Hardware-in-the-loop (HIL) environment. As a result, the testing platform proved to be able to detect errors contained in the SUT with a better coverage performance than existing methods.

Generally, after test case design the main concern is in guaranteeing efficiency and quality when running the test suites given limited resources and time [69], highlighting the need of strategies able to ensure the selection of a suitable test cases subset to define the balance between cost, time and accuracy.

Test case prioritization deals with the problem of ordering an existing set of test cases to be executed, based on their historical failure data [70]. Therefore, by running the test cases considered more important earlier in the test process, can help to identified failures in initial stages of testing and provide feedback to allow the necessary fix earlier.

With this in mind, Tahvili et al [71] used the result of executed test cases as a metric

call Dependency Degree to prioritize the test cases using the Fuzzy Analytic Hierarchy Process technique (FAHP). The results of the industrial case study showed that the concept of 'fail based on fail' is applicable and can reduce test execution time. Consequently, their approach will enable higher test execution efficiency by identifying and avoiding test redundancies.

In [72], Tung and Aldiwan described an algorithm for a parametric test case generation tool applying a combinatorial design approach to the selection of candidate test cases. The Test Case Generation (TCG) algorithm helped to achieve the specified degree of test coverage with a near-minimal number of test cases.

In [73] Srivastava developed the Echelon, a test prioritization system, which prioritizes the application's given set of tests based on what changes have been made to the program code. Echelon is also able to operate on millions of lines of source code and within a few minutes, it orders the given tests to cover the affected program so that defects can be are be found quickly and unnecessary costs can be prevented.

Staats et al. [74] proposed a technique for prioritizing test cases by implementing test oracles on the effectiveness of testing. Their technique operates by first capturing the flow of information from variable assignments for each test case, and then prioritizing to cover variables using the shortest paths possible.

7.1.5 Test Execution

In this phase, the system under test is executed according to the steps on the test cases previously generated. The level of automation adopted during the previous phases directly affects the test execution. For example, if all tests cases were developed as a manual test set, test execution must also be performed manually. In contrast, if all test cases are developed as scripts or executable test suites, test execution can be fully automated.

In [75] Winkler et al., presented a framework for automating software test processes based on Unified Modeling Language (UML) that enables early testing approaches. The framework consists of a test runner that manages the test suite execution and generates the test report based on the analysis of individual tests, at unit and system level.

Anbunathan and Basu [76] developed a tool to test Android mobile applications at system level based on Model-Based Testing (MBT). Their approach used a sequence of diagrams to capture input scenarios and stored them into an XML utilised to create an Android application (APK) to execute test cases in Android mobile, which injects events and user actions into Android phone. Additionally, a test schedule was implemented to executed multiple APKs in sequence. Although only a few functionalities of Android mobile are testable with this method, this approach can efficiently adoptable for Windows or Linux based applications and other embedded systems.

Many projects have discussed the use of HIL system testing as an efficient solution to enable the test execution of software and hardware of complex systems such as marine control systems [53] [54], electric vehicles [55] [56], avionics systems [45] [57] and railway systems [58] [59]. HIL is a well-proven solution utilised mainly as a type of black-box testing and can be defined by the operation of integrated real and simulated components. Many HIL simulations are in-house developments and the structure of their design and testing strategies vary according to the field and application purpose.

In [59], Huibing et al. developed a real-time HIL simulation railway system to perform testing on the system level to evaluate the performance of balise transmission module (BTM). The authors achieved a realistic environment for BTM test and research, with actual up-link signal, simulated trains, automated balises telegram and easily configured parameters facilitating the use for single balise, balise group and balise sequence simulation. Although the environment proved to be capable to provide feedback and input automatically, the authors choose to generate and execute the test case manually.

In [56] Ma et al,. proposed a HIL testing for connected automate vehicle (CAV) applications. The environment involved traffic signal controllers, communication devices, and a traffic simulator that enables physical test-vehicles to interact with virtual vehicles from traffic simulation models. The HIL developed was successfully used to perform partial automated system testing with one physical test CAV and other virtual vehicle considering common traffic situations in small scale and low-risk situations.

In [54], Pivano et al., implemented an automated HIL testing to verify and validate marine control systems in a laboratory with the actual targeted system hardware. The authors introduced a virtual test operator, responsible to operate the target control system through its Human-Machine Interface (HMI) as well as operate the HIL simulators. The automated HIL testing was proved to be successfully implemented into the factory acceptance test and able to verify that the functionality and performance according to the requirements, and after sea trials as an extra software regression check and verification.

In [45], Abdullah et al., created an approach to test automation implementation for an avionics systems on the system level at the System Integration Laboratory (SIL); the environment for avionics system integration and test facilities. To reduce human interaction during testing, various test agents were developed to send predefined commands and messages units in the test environment. In addition, keyboard/mouse macros were implemented to operate windows-based-software components, while in the meantime, a test drive PC runs automated system level test script in the SIL environment, leading their verification process to be completed up to 3–4 months earlier in comparison to the original manual verification schedule.

According to the SUBSET-111 [52], the maximum level of automation possible is expected to be implemented throughout ETCS testing process. Considering the desired test environment to test ETCS L3 (moving block), the existing testing execution process will need to be adapted to fulfil the new system needs, especially taking into account the presence of the maximum number of HIL systems.

7.1.6 Output evaluation

Evaluation of the test case results is vital to provide evidence and feedback on the effectiveness and quality of the SUT. In literature, the results obtained when the SUT is tested are called actual results and the result produced when the SUT satisfies its intended behaviour are called expected results.

Usually, the output evaluation can be carried out manually, where a test engineer makes the judgement between the expected and actual results. Conversely, automated test evaluation can be adopted considering different levels of intelligence to evaluate a test execution results, from the basic use of a hard-code test evaluation to the development of test oracles using machine

learning and artificial intelligence (AI).

In [77], Hamlet proposed a compiler as part of the program code using a supplied list of input/output pairs to verify the test results. Whereas these solutions require that the correlation between the input/output pair were generated manually, which may be difficult and time-consumed, Peters and Parnas [78] described an oracle, to be used in unit and system testing, based on the automated conversion of the requirement documentation into a high-level computer language. Aside from the input-output evaluation, the test oracle generator can also be used to ensure that documentation is kept up to date and as a result, increase the documentation consistency and reduce the cost of testing and ensure that errors that occur during testing are detected.

In [79] the authors applied decision tables, on unit and integration testing, of web-based applications. In this approach, a model of software verification is presented using a unique combination of conditions between the software's input state and output expected results.

The study in [80] presented Artificial Neural Networks (ANN) and Info Fuzzy Networks (IFN) as automated test oracles to black-box testing. To date, many IFN-based oracles were proved to be more efficient in regression testing than functional testing and were unable to deal with complex applications. Conversely, an ANN-based oracle with the improved of multi-networks architecture was able to handle different types of testing analysis and even more complex applications.

Just and Schweiggert [81] combined oracles from different applications in an integrated environment to validate the test results. The applied partial oracles approach was assessed with regard to their capability to reveal faults in the individual parts of the subsystem and proved their suitability for integration and unit testing.

7.1.7 Test closure

Test closure is a critical phase in the test process, it reports the test verdicts in the form of a test result report and defects to the developers. This activity can be performed manually or automatically using one of the many techniques and frameworks [82] [83] [84] available.

8. State of the art train control systems

ETCS operates through the communication of two subsystems: (i) Trackside Subsystem; (ii) Onboard subsystem. Based on ERA Subset-026, Chapter 2.5 [85] the general architecture of ERTMS/ETCS is shown in Figure 16.

The onboard subsystem (OB), represented in blue boxes, is installed in the train and it manages the train movements by using the information received from the trackside. In the onboard ERTMS/ETCS subsystem we consider the following modules:

- European Vital Computer (EVC) is the core of the onboard equipment and process all the train borne functions and exchange data between the OB and trackside systems.
- Train Interface Unit (TIU) is the interface that enables the ETCS to exchange and issue commands to the rolling stock.
- Juridical data records and stored the data coming from the onboard unit (OB).
- Balise Transmission Module (BTM) and Loop Transmission Module (LTM) manage messages between the train and the Eurobalise and the Euroloop respectively.
- Driver Machine Interface (DMI) is the interface between the driver and ETCS.
- Odometry provides the train position based on information from speed sensors.
- Euroradio applies to the Euroradio protocol to encode the message sent by the RBC and decode the messages received from it.

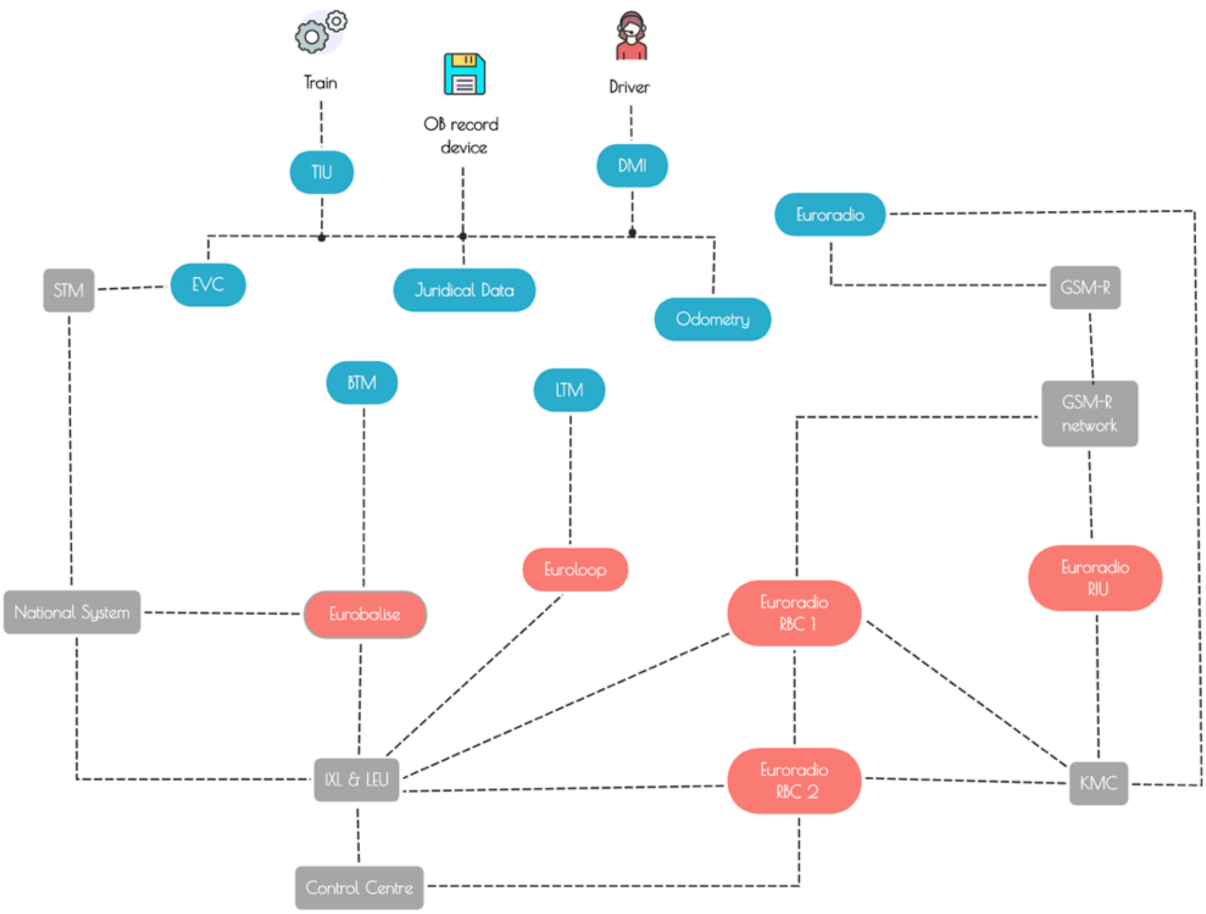

Figure 16. ERTMS/ETCS general architecture

On the other hand, the trackside system, represented in the above figure in orange boxes, acquires all the necessary information to ensure the train routing. The trackside ERTMS/ETCS subsystem is composed of:

- Eurobalise is a passive device that lays on the track and ensures the transmission of the data to the train and RBC when the train passes over it.
- Euroloop allows the transmission of additional data.
- Radio Infill Unit (RIU) is applicable to ETCS level 1. It sends additional data using radio channel.
- Radio Block Centre (RBC) is a computer-based system that manages the exchange of data between OB and trackside in its area of governance.

GSM-R system and other optional and/or adjacent modules, represented in the above figure in grey boxes, play an important role in the operation and safety of the ERTMS, some of them are defined below:

- GSM-R Trackside radio is used to ensure the message exchange between the onboard and the RBC.
- Key Management Center (KMC) manages the cryptographic keys indispensable to ensure the communication between the onboard and trackside
- GSM-R is a mobile communication system that allows data exchange between the OB and RBC.
- Interlocking (IXS) is not an ERTMS component. But it can play an important role in some ERTMS structures.
- Lineside Electronic Unit (LEU) is the interface between the Eurobalise and interlocking

9. Train control system test model and simulation

The railway industry is continuously looking at different ways to improve the safety of its operation. Recent advances in simulation methodologies have made simulation an efficient and cost-effective means to carry out system behaviour analysis and characterisation [86]. According to [87] a simulation is a technique used to experimentally apply physical scenarios by the observation of the performance, over time, of a dynamic model of the system. In many applications, simulation can involve testing; comparing alternative designs; validating; and supporting to study recommendations and proof of concept.

In order to facilitate and enable the simulation and assessment of a system, a representation, or so-called model is constructed. As a wide-ranging concept applied in many fields, models are simplified abstractions and can be defined as an entity that is used to represent a system for some defined purpose [88]. The potential to detect defects in a SUT relies on the accuracy of the specification model, the mechanism used to determine the test outcome and the test generation strategies chosen.

Train Control Systems are highly integrated systems with a large number of different functions which are reflected in the complexity in generating models to represent the system. The approach to test TCS needs to consider the combination of simulation technology and testing technology. The simulation creates a representative environment the emulates in the most accurate way the production environment for a SUT, while the testing technology takes into account the testing process as the coordination of the simulation and its sequential execution.

In order to describe a complex system and more specifically railways, two elements are important: the system architecture and the prediction of its behaviour. The system architecture should define the SUT components hierarchical decomposition and the system functions and requirements. The prediction of its behaviour is shown as a model that is able to illustrate and represent the full list of operational scenarios.

Up to now, several researchers [89] [90] [11] have developed and validated different modelling methods taking into account different interfaces. Those approaches presented some differences and the appropriate modelling method should be select according to the specific testing purpose. In order to define a modelling language that fits ETCS Level 3 requirements, it is necessary to understand TCS design theory, and what are the limitations with the existing methods.

In [11] a formal modelling method of simulation, combined MBT, is specified. In order to test CBTC (communication based train control) the author introduced the SCTIOTS method that is not limited by the complexity of the SUT as it models the SUT into parallel abstract and simulation models, where the abstract model is developed using a formal modelling tool, and the simulation model is built by a simulation tool. This combination can significantly reduce the formal model size to avoid state explosion. Although SCTIOTS does require a more profound understanding of formal modelling and the SUT operating principle, it can be applied to more scenarios because of its flexibility, high degree of efficiency and interoperability.

Traditionally, off-line testing methods are used to determine whether the CBTC system complies with the desired specification, however, these testing methods are becoming insufficient due to the growing complexity and non-deterministic behaviour of the system. With this in mind, the

authors in [91] proposed an on-line testing method, based on a micro railway simulator and the testing tool UPPAAL-TRON. The main benefits of this method make it possible to model the complex timed system and test the system completely automatically by computer. It also makes repeated or periodic tests more efficient due to its on-line characteristic and it proved that the UPPAAL based on TA theory is capable of describing the CBTC systematic behaviours and this testing method is ready to be further applied in the field of CBTC system testing, such as hardware-in-the-loop testing in the future.

In [89] the authors presented a first coloured Petri-net model for the communication system management of the signalling system ERTMS level 2. Petri nets have gained much attention and success due to their simplicity and become one of the most popular used tools for modelling discrete event systems for different phenomena. The Petri net model represents the various connections between entities of the process. For each configuration (discrete state), a set of equations can be integrated and associated to a place to model continuous phenomenon.

In [90] the authors proposed modelling the ERTMS level 2 signalling systems as a System-of-Systems (SoS) using Unified Modelling Language statecharts. They seek to evaluate its dependability parameters by considering the unavailability of the whole SoS as an emergent property. In addition, human factors, network failures, Common-Cause Failures (CCFs), and imprecise failure and repair rates are taken into account in the proposed model. ERTMS level 2 was proven suitable to be considered as a SoS because the SoS approach provides a global view of this signalling system. Furthermore, the emergent properties of the signalling system can be easily analysed at the SoS level. Because of their use of notions of hierarchy, concurrency, and synchronization, state charts have proved to be suitable for modelling the behaviour of ERTMS.

In [92] a composable modelling method for the generic test platform for CBTC system based on the port object was presented. This method defines the port object (PO) model as the basic component for composable modelling, verifies its port behaviour and generates its compositional properties. Based on the port description and the test environment description, it builds port sets and environment port cluster, respectively. Then it analyses and extracts possible crosscutting concerns, and finally generates a variable PO component library. It takes the modelling of block port objects in line simulation of generic test platform for CBTC systems as an example to verify the feasibility of the method.

In [93] the authors adopted a model-driven development (MDD) to test the CBTC system by categorizing the requirements into use cases. In this approach, the TCS was modelled on UML notation utilising a computer-aided software engineering (CASE) tool to generate the model, which enabled the fast iteration between development and testing cycles and showed good applicability in developing the CBTC model to test the system effectively.

In [94] [95] the authors proposed an approach for system formalization and safety analyses called ScOLA *(*a Scenario Oriented Language) built to understand and to formalize the specifications based on core concepts applied to CBTC*.* ScOLA is a framework where system engineers can define and validate a system formally, starting from informal and textual descriptions of the system. They presented how, starting from formal but explicit models of the system, safety analysis could be performed more efficiently and introduced some identified safety techniques and a case study.

In [96] the authors defined a highly automated approach based on ETCS L2 model using the

Model-checking technique to perform exhaustive analysis of the system behaviour. A formal model explicit the various modes characterized by their respective active functions, as well as the numerous combinations of conditions for switching between modes. The various steps guiding the translation of the specifications into a formal model are explained and shown through different examples and the obtained model proved to be a convenient basis to check safety, interoperability and liveness properties.

In [97] the authors introduce a novel class of hierarchical state machines, called Dynamic STate Machines (DSTMs), and proposes an approach for modelling and validating railway control systems, based on the new specification language. The formal semantics of DSTM allows for the definition of verification and validation methodologies supported by automated tools. The paper also described how DSTM specifications may be mapped to models in order to achieve automated generation of test cases by model checking and Spin. The language and the proposed approach are illustrated and motivated by applying them to a specific functionality of the Radio Block Centre, the vital core of the ERTMS/ETCS control system.

In [98] the authors focus on the use of the FORMOSE approach for modelling the hybrid ERTMS/ETCS L3. The requirements are specified with SysML/KAOS goal diagrams and are automatically translated into B System specifications, in order to obtain the architecture of the formal specification. Their automatic translation completes the structural part of the formal specification. The only part of the specification, which must be manually completed, is the body of events. The construction is incremental, based on the refinement mechanisms existing within the involved methods. The Rodin tool was used to verify and validate the formal specification, especially to prove the safety invariants and the refinement logic, after the completion of the body of events. Work in progress aims at improving the representation of domain predicates to make them more user-friendly and to integrate the approach within the open-source platform Openflexo which federates the various contributions of FORMOSE project partners.

10. Testing methodologies towards zero on-site testing

The zero on-site testing (ZOST) concept consists of a set of software and hardware tools to create a test environment as similar as possible as the system in production, to allow validation and verification activities in the laboratory avoiding time-consumed and costly trials on real sites. There exists little literature for ZOST for positive control systems; however, the following methods for software and hardware ZOST are available.

In Tian et al. [99], the authors designed and evaluated DeepTest, a tool for automated testing of Deep Neural Networks a.k.a. DNN-driven autonomous cars. Their tool can automatically generate test cases leveraging real-world changes in driving conditions and systematically explore different parts of the DNN logic by generating test inputs. DeepTest uses domain-specific metamorphic relations and found more than a thousand erroneous behaviours under different realistic driving conditions (e.g., blurring, rain, fog, etc.) many of which could lead to potentially fatal crashes.

In Sun et al., [100] the authors demonstrated the feasibility of automatic test case generation from functional requirements, targeting software testing in avionics. They applied Bounded Model Checking (BMC) to formal low-level requirements, in order to generate tests automatically to replace existing labour-intensive test writing procedures while maintaining independence from implementation artefacts. They then formulate apply the prototype in case studies with industrial partners. As a result, the method developed is demonstrated to significantly reduce the human effort for the qualification of software products under DO-178 guidance.

Sarla and Ramadoss [101] described an innovative approach towards automation of Combinatorial Interaction Test case (CIT) generation and execution for Requirements Based Testing (RBT) of complex avionics systems for achieving test adequacy in a highly time efficient and cost efficient manner. By combining methods from requirements engineering and software testing, this testing methodology provides a set of quality assurance activities and management tools that enable getting requirements right from the outset. The procedure involves generation of test cases by development of the simulated reference model for the functionality under test based on the detailed SRS, aiming to detailing the software requirement specifications without any ambiguity and hence the combinatorial test cases designed using this method will generate the most optimal test cases. Execution of these test cases shall not only provide the benefits of CIT but also provide benefits of RBT. In Ma et al.[56] a HIL system was developed to test a representative early deployment CAV application: queue-aware signalized intersection approach and departure (Q-SIAD). The involved software and hardware include a physical CAV controlled in real time, a traffic signal controller, communication devices, and a traffic simulator. The author utilised a Q-SIAD algorithm to generate recommended speed profiles based on the vehicle's status and system constraints and parameters (e.g., maximum acceleration and deceleration). The algorithm also considers the status of other vehicles in designing the speed profiles. The experiment successfully demonstrated its functionality with one test CAV driving through one intersection controlled by a fixed-timing traffic signal under various simulated traffic conditions. According to the authors, there are still limitations of the developed HIL system that need to be addressed in future studies. Future studies should consider more realistic testing scenarios, including high-speed traffic or complex vehicular movements, on signalized arterials.

De la Cruz et al., [102] presented the implementation of a hardware-in-the-loop real time

simulation test bench for air-to-water-heat-pumps (AWHP). The new test bench proposed uses a virtual model to set real weather conditions and simulates building's and occupants' response. This test bench allows R&D departments and manufacturers to test several heating systems in parallel as well as different control options. Optimized control parameters throughout time are proposed in view of reaching lower consumption and better performances. The analysis highlights control issues related to the heating curve, the domestic hot water generation management and the defrost process of AWHP.

This case study shows the potential of the HIL test benches for manufacturers to optimize the development of their products controls. It may also limit the need for field testing and has the potential to minimise the cost and time required in the development-to-market process.

Bouaziz et al [103] presented two complementary strategies to simulate and emulate the traffic exchanges between TCMS and MCG & GCG through LTE for a zero on-site testing approach. Although this work considers LTE technology, it can be adapted if different system models are available. The goal of their approach consisted of verifying the possibility of connecting physical emulation tools with simulation software tools in an end-to-end communication testbed in order to examine the effect of handovers and traffic delays. The experimental results presented by the authors reveal successfully primary testing and as a proof of concept of the test environment, the authors presented results in terms of delay and delivery rates considering TCMS traffic and various railway scenarios such as train speed and wireless network load.

Aguado et al [104] built an integrated simulation framework for advanced railroad traffic management and control systems that includes functional and communication layers. The authors used the railroad network geographic coordinates and real train graph schedule for high speed lines. By modelling the train control service, they were able to measure how failures or degradation in the communication link affects railroad network availability and operational performance indicators. In their demonstrator, the railroad communication network was tested in different ways and the behaviour of different standardized telecom approaches was verified.

PART C

11. Summary of requirements and gap analysis

This section describes a set of high level requirements for testing moving block systems derived from Tasks 2.1 (stakeholder functional requirements, reported in Section 0) and 2.2 (review of previous work, Sections 4 - 0). Feeding into the requirements are also the definitions of moving block signalling provided in MOVINGRAIL T1.1 – Basic signalling terminology and concepts. The baseline definitions of moving block operational rules have used ETCS Baseline 3 Release 2 [BL3 R2], as adopted in X2RAIL-1, and deliverables D5.1, D5.2, and D5.3 from that project as a starting point.

The main assumption of moving block systems, in contrast to their predecessors, is that blocks are not defined by pairs of adjacent fixed points on the line, usually at signals. Rather, their establishment is defined virtually, and movement authority is secured via radio-based communication (RBC). The calculation of movement authority under moving block systems is done dynamically and as a function of speed, performance, and length of the leading train in a pair. The challenge is then to guarantee that the system fulfils the safety requirements, especially in situations regarding identification of loose carriages or wagons (i.e. train integrity monitoring) in particular on tracks that pose a significant safety hazard by operating trains closely to each other. It logically follows that such a paradigm significantly **increases the operational complexity since it creates an infinite number of movement authority allocation scenarios in contrast to the current fixed block case**.

An effective way to approach this challenge is to apply improved strategies and methods for testing moving block signalling systems, focusing on a trade-off between laboratory and on-site testing.

As determined from stakeholder interviews and reviews of the literature, the key point to achieve a robust testing solution for moving block systems relies on **increasing the reliability in current laboratory test practices and providing a test environment that is able to reproduce relevant operational scenarios and conditions as realistically as possible.**

However, there are considerable challenges and technical limitations to consider when moving towards zero onsite testing. Even though challenges can be overcome, there are operational and system characteristics intrinsic to ETCS and to moving block that are impractical to reproduce digitally or simulate and cannot be performed in a laboratory, such as those depending on special equipment only available in the field (e.g. interlocking), or tests directly related to geographical or environmental factors, or ones related to real time and dynamic behaviour.

These premises form the basic assumptions that should be set to investigate the requirements and strategies for testing moving block systems. The key recommendations derived from those premises have been organised into seven key principles, as follows.

i. General

ID	Recommendation
R-G01	The testing system and strategy shall be a single integrated system that is capable of supporting testing by different stakeholders and with different objectives
R-G02	The testing system and strategy shall be applicable for ETCS L0, L1, L2, L3 (moving block) configurations
R-G03	The testing system and strategy shall be compliant with Subset-110 [48], Subset-111 [105] and Subset-112 [106]
R-G04	The testing system and strategy shall be capable of upgrade to support virtual coupling and future versions of ERTMS/ETCS
R-G05	Laboratory tests shall be planned to minimise costs / complexity in logistics, management and communication processes compared to on-site testing

ii. Capability

ID	Recommendation
R-C01	The testing system shall be capable of simulating a complete railway system including simulated train movements
R-C02	The simulation models shall cover the complete railway system including railway infrastructure, rolling stock, timetables and railway operation rules

iii. Automation

ID	Recommendation
R-A01	Testing should be automated as much as possible
R-A02	The testing system shall be capable of running routines of different levels of automation (e.g. manual, fully automated)

iv. Data

ID	Recommendation
R-D01	The testing system and strategy shall be adaptable to facilitate the testing requirements of different stakeholders and suppliers
R-D02	The testing system shall be capable of performing simulations that take the inputs and give the outputs mimicking those of all ETCS sub-systems
R-D03	The testing system shall be able to provide simulation data representing trackside systems when applicable
R-D04	Real data, recorded under real operation conditions, shall be used as often as appropriate as test input to increase the quality of the tests
R-D05	Prior to its deployment, validation of the testing system should be performed to compare laboratory data with on-site data outputs to ensure the behaviour of the system under test (SUT) in the laboratory matches the behaviour on-site

v. Telecommunications

ID	Recommendation
R-T01	Testing systems shall simulate telecommunication protocols and GSM-R in order to reproduce on-site signals, connection error ratio, connection loss, radio holes, and network registration delay
R-T02	The testing system shall be adaptable so that it can interface with existing telecommunications networks and testing systems

vi. Processes and Procedures

ID	Recommendation
R-P01	The testing process shall consider a wide range of operational scenarios, to include (1) functionalities and requirements (main objective); (2) potentially safety critical instances; (3) common failures in previous systems
R-P02	Appropriate processes and procedures shall be in place to manage system configuration and configuration data at a sufficient level of integrity
R-C04	Processes and procedures shall include a systematic and transparent execution and documentation considering traceability, test coverage and hazards

vii. Components

ID	Recommendation
R-M01	Any necessary reconfiguration arrangements between suppliers' systems for testing shall not require supplier involvement
R-M02	Testing systems shall use as many real components as possible as an option to minimise on-site testing whenever possible (HIL)
R-M03	Testing shall envision different grades of automation (GoA) in operations, including those with human drivers and manual operation
R-M04	The testing system shall support RBC to RBC handover testing and simulation
R-M04	The testing system shall be adaptable to test assemblies and components from different projects or suppliers
R-M05	The testing system shall be adaptable to test ETCS L0, L1, L2, L3 (moving block) configurations

12. Operational concept for moving block testing

This section presents an operational concept for testing moving block, relating to the first part of Task 2.3. The main objective is to develop the characteristics of the proposed system from the stakeholder viewpoint and highlight the necessary improvements and changes in current test strategies. This can be used in specifying the strategy (see Section 13) for the migration of current testing towards moving block compatible testing procedures in a ZOST environment.

12.1 What is an operational concept?

Operational Concepts (OpsCon) and Concept of Operations (ConOps) are terms frequently used and applied in the early stages of a project. Although some sources may use them interchangeably, the two usually differ in the level of description they adopt.

For this task we adopt the following definition: an Operational Concept is a document that describes what the system will do (not how it will do it) and why (the rationale behind it) [107] - refer to the top two levels of Figure 17. It can include text and/or models to illustrate the functional aspects of a system through stakeholder perspective. It also defines high-level requirements and objectives, pointing to critical aspects of the operations.

According to the standard ISO/EEC/IEEE 29148:2018 [108], an operational concept description is used to communicate overall quantitative and qualitative system characteristics to the acquirer, user, supplier and other organizational elements.

Operational Concepts are also an integral part of Capability Systems Engineering processes that deal with achieving or improving the outcomes in complex systems [107]. The use of operational concepts allows engineers to find ways to address the mission and objectives of systems under development. This in turn opens a wider array of possible solutions that fulfil general requirements through a traceable decision making process from purpose to sub-system requirements and specifications (Figure 17).

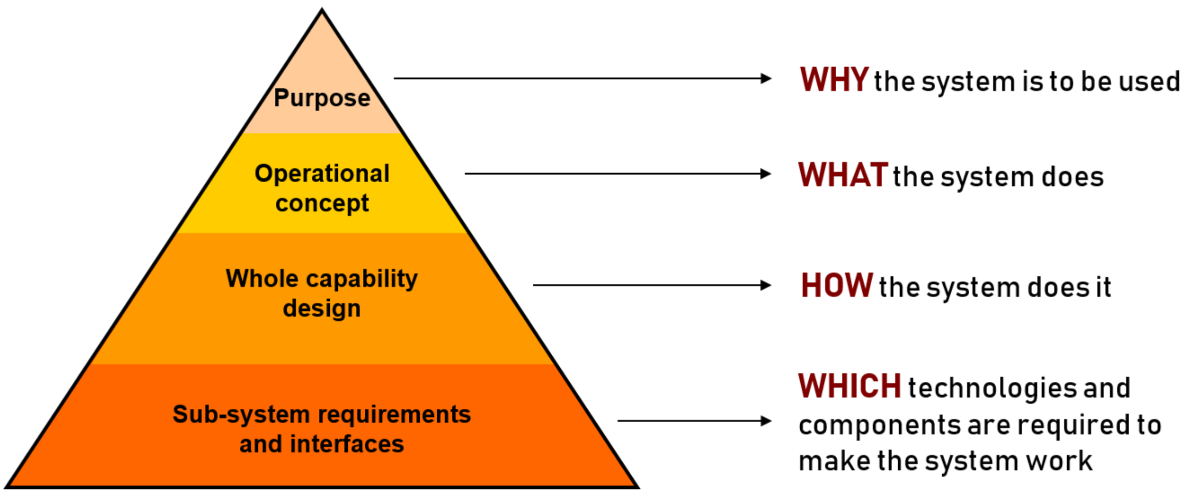

Figure 17: The capability systems engineering process (adapted from [107])

12.2 Stakeholders

Experience with past ETCS and CBTC projects shows that there are many stakeholders involved in the successful delivery of such projects. This is particularly the case where a project interfaces with other signalling systems, or even with other lines managed by different infrastructure

managers.

There is no reason to presume that the types of stakeholders for future projects including moving block and virtual coupling functionality will be any different. The stakeholders for a project may typically include:

- Signalling system/equipment supplier(s)
- Telecoms system delivery company(s)
    - Fixed telecoms
    - Mobile telecoms
    - Train-to-train telecoms (virtual coupling)
- Traffic Management System supplier(s)
- Test bench suppliers
- Project delivery companies
- Suppliers of existing signalling on the line being equipped (modifications for stageworks, etc.)
- Railway infrastructure managers (IMs)
    - Infrastructure operators
    - Maintainers
- Train Operating Companies (TOCs) – may be several on one project
    - Passenger TOCs
    - Freight TOCs
    - Infrastructure Maintainer TOCs (engineering plant)
- National safety authorities (NSAs)
- Standards bodies
    - International
    - National
    - Stakeholder
- Specialist test houses
- Consultants
    - Client's engineer
    - System integrator
    - Safety consultants (ISAs, etc.)
    - Certification bodies (NoBos, etc.)
- Stakeholders involved in adjoining sections of railway
    - Signalling system suppliers
    - Telecoms delivery companies
    - Traffic management suppliers
    - Infrastructure managers
    - Safety authorities.

N.B. There may be a varying mix of stakeholders in each of the above categories for each separate project. For example:

- The same signalling system supplier may work with different telecoms delivery companies, TOCs, etc. for the same IM and NSA.
- An IM may use different signalling system suppliers and telecoms delivery companies on different projects.

12.3 Stakeholders role in the testing process

Specifiers of functional testing:

- EU Commission /ERA
  - ETCS requirements
  - Interoperability
  - Common Safety Method (CSM)
- Standards bodies.

Specifiers of requested testing relating to the specific application:

- National Safety Authorities (NSAs)
  - TSIs
- Railway infrastructure managers
  - National rules
  - Scheme specific requirements.

Test implementers:

- Signalling system/equipment suppliers
- Project delivery companies
- Specialist test houses
- Client appointed independent testers.

12.4 Purpose of New Testing System

At the highest level, the principal purposes of a testing system are:

- To integrate separate subsystems, often from different suppliers, into a full working system,
- To gain formal acceptance of the system by the project client,
- To allow the system to be put into service,
- To minimise the testing time on site (especially if the project is on an existing working railway).

12.5 Outcomes

The successful outcomes of a testing process may be summarised as follows:

- Proving of the required level of interoperability with other stakeholders
- Gaining the necessary confidence, documentation and certification to obtain sign-off by all necessary stakeholders
- To deliver reliable in-service performance
- To satisfy the project client
- To gain, retain and grow confidence in the testing strategy and test systems, in order to maximise the reliance that can be placed on them on future projects.

12.6 Functional aspects - what the test system does

The test system must:

- Allow testing of the complete system, including interfaces with all stakeholders (e.g. DMI, maintainer terminals, traffic management systems, etc.)

- Allow all levels of ETCS operation, including L0, L1, L2, L3 (fixed and moving block, with and without virtual coupling), hybrid systems (e.g. L2 + L3), ATO of various GoAs
- Allow testing compliant with Subsets 110, 111, 112 to be achieved, enabling the inclusion of elements from different stakeholders and suppliers, and the exchange of elements between tests
- Allow maximum coverage of system testing, in order to minimise the testing required on-site
- Simulate train movements entering, running through and leaving the controlled area, as well as trains powering up and down, and trains developing faults (e.g. stopping communicating)
- Allow testing of all the possible changes of Level and Mode of Operation encountered in service
- Facilitate the use of pre-prepared test scripts to automate both testing and the recording of test results
- Allow manual intervention in testing to aid the testing of complex or unexpected events
- Make clear distinction between generic application software, capable of verification and certification (and therefore re-use on multiple projects), and project-specific configuration data (effectively part of the SUT)
- Use of recorded data from site (*)
- Proving the test system replicates behaviour seen on-site (*)
- Include fixed and mobile telecoms protocols and interfaces
- Simulate expected and abnormal degradations of fixed and mobile telecoms system performance
- Interface to target telecoms systems
- Allow fixed elements and project-specific configuration data to be placed under strong configuration control, to ensure the validity of test results over an extended test campaign
- Allow system boundaries to adjacent areas of the railway network to be tested
- Allow the testing of different levels of train position resolution, and different MA update rates
- Allow communications performance parameters (bandwidth, latency, etc.) to be varied,
- Allow efficient testing of stageworks, including regression from the final stage to partial commissioning stages, etc.
- Allow the efficient re-testing of particular functions following changes or error correction
- Allow the efficient re-testing of the system following a change of track layout, speed limit, other functionality, or following software evolution or re-platforming of a system element (e.g. RBC).

(*) It is not a primary purpose of the functional testing to verify data recorded on-site, but it should be possible to use such data to assist in replicating unexpected behaviour seen on-site in a lab environment, for further study. For example:

- If a balise group is found to be located in the wrong place, its position would normally be corrected on-site, rather than moving it in the infrastructure database. However, if there is a reason why it cannot be in the intended position, the data used in both the test system and the target system may have to be updated

- If the distance between two locations on the track was found to be different from that provided in the infrastructure data, it would be a cause for concern about the robustness of the infrastructure data collection process, rather than a matter to be adjusted in the data used by the trackside system
- If unexpected behaviour of a train is seen during on-site testing, data recorded by the on-board equipment's JRU or Diagnostic Recorder Unit (DRU) may be analysed in the lab, and compared with data recorded in the JRU/DRU of the on-board unit under test on the test bench in the same scenario

12.7 Test environment

This section of the document considers WHAT the test system will do, not HOW it does it.

Section 13 gives more details of the test environment, and the impact of Level 3 moving block and virtual coupling.

12.7.1 Phases of Testing

Several distinct phases of testing can be identified for the elements of a moving block signalling system. These include:

- Functional and non-functional testing and certification of each of the products forming the target system. This testing work is usually the responsibility of the individual product suppliers
- Sub-system integration and testing, e.g. IXL testing, RBC testing, balise telegram verification, through testing of point control, etc.
- System testing using a Test Bench. This can involve distinct phases, including:
    - All sub-systems and telecoms simulated, using target configuration data
    - Some sub-systems and telecoms simulated, other systems using target hardware/software
    - All systems using target hardware/software
    - Telecoms using actual systems
- Installation testing on-site (e.g. through testing from IXL to signalling objects)
- Testing using actual train (in a depot) linking via telecoms to RBC
- Shadow running (active on-board equipment, but not interfaced to the train)
- Test running in possessions
- Reliability testing.

12.7.2 Scope for Innovative Test Tools

It is assumed that the first two phases of testing listed above are already well covered by the in-house test facilities of the relevant equipment suppliers, so these are not described further here.

The major scope for the development of innovative testing tools comes with the System Testing, both on Test Benches and later On-site testing phases.

12.7.3 Scope of Test Bench

Because of the specification approach adopted for ERTMS/ETCS, whereby the behaviour of the on-board equipment for 'go anywhere' trains is fully specified, each design of on-board equipment is capable of certification as 'correct'. Once this certification has been achieved, the on-board equipment is strictly speaking not under test; the system testing process is focused on

testing the following aspects of the trackside sub-system:

- The correct co-ordinated operation of several trackside sub-systems (TMS, IXL, RBC, communications, etc.), perhaps from different suppliers,
- The correct implementation of the functional behaviour specified by stakeholders, in each instance of different operational scenarios found within the network section covered by the system,
- Potentially, in Level 3 and Virtual Coupling systems, the dynamic interaction of trains running in the system may need to be investigated by the test tools. This is not generally an issue with Level 2 ETCS system testing, as the behaviour of the train under test is dependent only on the occupied/clear status of train detection sections, and a finite number of bits of logical information output by the IXL. However, when the behaviour of the train under test depends on a quasi-infinite number of possible locations of the preceding train, dynamic effects may occur.

The first of the above areas of testing is concerned with de-bugging system integration, the second with functional behaviour and the third with dynamic behaviour. It is the second and third areas which are considered in more detail here.

At the simplest level, the Test Bench must permit a certified set of on-board equipment to be provided with all the necessary stimuli that would occur in actual running in the scenario being tested, at the specific location within the system control area, so that the on-board equipment's behaviour can be proved to be as expected.

As location-dependent issues (e.g. two different operational scenarios happening to occur close together on the track) may adversely affect the on-board behaviour (particularly when dynamic effects are also considered), this testing cannot use generic test data, but must use the infrastructure data to be installed in the delivered system, as well as geographical configuration data generated by the application tools for the sub-systems (IXL, RBC, etc.) forming the system.

12.7.4 Scope of On-site Testing Tools

Fundamentally, on-site testing is concerned with verifying the behaviour of actual equipment on the railway. However, some opportunities for innovative testing tools can be foreseen, particularly as systems move towards moving block, hybrid, and virtual coupling.

These new functionalities are aimed at two possible goals:

- To reduce the whole-life cost of train control systems, and/or
- To increase the usable capacity of a railway network.

The second of these goals requires special tools to facilitate the testing of very close headway operation of trains prior to formal sign-off of the train control system, without introducing a significant risk of an accident due to an equipment malfunction or application error. This is not a problem that railway engineers have traditionally had to face.

The most obviously useful tool would simulate, in an installed train control system, the presence of one or more 'virtual trains', which could closely follow actual trains around the system or which the actual trains could closely follow. Such possibilities are discussed further in Section 13 below.

The same tools might well prove to be useful in improving the cost-effectiveness of the testing of a wider range of train control systems using Level 3.

13. Moving block testing strategy

13.1 General

The testing strategy for any modern train control system can be split into several phases, typically including the following:

- Generic product testing (Type Testing)
- Generic sub-system testing
- Proof-of-concept testing of operational scenarios using simulators
- Proof-of-concept testing of operational scenarios on a test line
- Functional testing of sub-systems configured with project-specific data
- Project functional testing in a generic test laboratory
- Testing on the project's target hardware in a factory test environment
- On-site testing, following installation of the project's target hardware.

A common-sense aim is to maximise the coverage of testing in the earlier phases listed above. The reasons for this are that the early phases can be done once on an example of each product or sub-system, and do not need to be repeated for every item manufactured, and that testing on-site is very much more expensive to carry out, wasteful of time, and potentially disruptive to the operation of an existing railway line.

13.1.1 Generic Product and Sub-system Testing

These testing activities are usually carried out by equipment suppliers, either using their own resources, or those of an external test house.

Examples of each product or sub-system are subjected to a series of functional and non-functional (e.g. environmental) tests.

The outcome of successful tests is certification and client acceptance of the generic products (e.g. balises) or sub-systems (e.g. interlockings) that will be used later in commercial projects.

13.1.2 Proof-of-Concept Testing of Operational Scenarios using Simulators

With complex train control systems, great care is needed to ensure that each aspect of day-to-day operation of the system achieves the behaviour that the stakeholders desire, and operates in a manner that will later be readily accepted by all operations staff in the client organisation, in the acceptance stages of a commercial project.

The required operational scenarios needed for a project are formulated. Typically, many will already exist in libraries of scenarios – a Reference Solution – maintained by the train control supplier and/or the client railway organisation, whilst a few may be new additions for the particular project. The scenarios can be simulated using an industry-standard simulator system (possibly supplied by a third-party supplier) configured with test data. Using the simulator, they can be demonstrated to client engineers and operators, and generic acceptance obtained.

13.1.3 Proof-of-Concept Testing of Operational Scenarios on a Test Line

Practical testing of operational scenarios on a test section of railway is an important step in developing any radical approach. Testing at the ERTMS National Integration Facility (ENIF) [109] in the UK has demonstrated the value of such a test facility, prior to introduction of the

technology on a commercial line.

A further role for such a test line is to facilitate ‘type testing’ of new types of train fitted with ETCS, ATO and virtual coupling. The correct behaviour of each new class of train can be confirmed in such an environment, ensuring that when testing is later carried out on the operating railway it is entirely focused on the behaviour of the infrastructure system, not on the on-board systems.

13.1.4 Functional Testing of Sub-systems using Project-specific Configuration Data

In advance of the full integration testing of the final system, individual sub-systems can be tested with the project-specific configuration data, to prove many of the aspects of their functional behaviour before any attempt is made to integrate with other sub-systems. This provides an opportunity to detect and correct errors at the earliest possible stage.

This sub-system functional testing may be carried out on actual hardware (target project hardware or identical hardware in a testing laboratory), or it may be carried out on some form of computer simulator – perhaps the actual application software running on an office computer rather than a deliverable safety-critical hardware platform.

13.1.5 Project Functional Testing in a Generic Test Laboratory

The principal phase of system testing involves setting up a set of sub-systems, each configured with project-specific data, and carrying out comprehensive functional testing, to ensure that all the required functionality is implemented as intended.

This testing involves the use of a test bench, consisting of a number of inter-connected simulators. Some later phases of testing may involve a mix of simulators and target hardware.

Several elements of a typical test bench can be identified as elements which will also be found in the delivered system, including:

- Interlocking and Radio Block Centre. For ETCS Level 2 system testing, these are usually separate sub-systems, but as the technology progresses towards ETCS Level 3 and Virtual Coupling, the conventional boundaries may change
- Communications network, simulated or real
- At least one on-board equipment, including DMI and basic driver controls. This equipment may be provided by one or more suppliers, in order to prove the operation of the system with all expected on-board equipment
- Where Automatic Train Operation (ATO) is part of the project, there may also be on-board ATO simulators, modelling both the ATO system and the train’s dynamic response to ATO commands.

Other elements model the real-world environment, providing the necessary stimuli to allow the above items to function as they will in the final installed system. These typically include:

- A data model of the track network, with balises, train detection boundaries and all other location-specific items stored,
- A computer system able to ‘move’ trains – either virtual trains created in the computer model, or trains represented by the on-board equipment (see above) – around the network in accordance with inputs received from the interlocking (e.g. point position) and

inputs received from 'trains' (see below). The position of each train in the network is tracked by the system, all the necessary outputs to on-board tachometry interfaces, balise reader interfaces, etc., are generated, in order to make the on-board simulator behave as it would on the actual railway. So, balise telegrams are generated at the correct locations, and odometry data corresponding to the correct distances of travel between balise locations are generated,

- A computer system able to represent certain aspects of the behaviour of trains, in response to inputs received from the on-board equipment (e.g. brake interventions) and the train driver (e.g. power and braking controls). Where ATO functionality is involved, the train simulation may need to be more comprehensive; where a manual tester is 'driving' the train, dynamic behaviour of the train is less critical.

It can be seen from the above that three kinds of 'train' may be involved in such a test bench:

- Train objects created within, and moved around the network by, the above computer systems
- Trains represented by a faithful functional model of the functional behaviour of the on-board equipment, but running on a lab computer
- Trains represented by sets of actual on-board equipment, connected via testing interfaces.

The first category of train is purely to model the movement of other trains in the network, whereas the latter two categories are used by testers to run testing scenarios, and record the resulting behaviour of trains.

This phase of testing may start with most systems simulated, and later progress to testing using representative hardware.

13.1.6 Factory Test

After the previous stage of testing in a laboratory environment, the supplier or the client may wish to carry out a full factory test, to prove that the entire target system (possibly multiple interlockings, multiple RBCs, etc.) behaves as intended, before shipment to site.

Certain aspects of system behaviour, such as timing of commands through the system, may only be truly testable at this stage.

In the best case, no problems will be discovered during factory testing, but any problems that do emerge will repay the cost of setting up the test.

13.1.7 On-site Testing

As individual sub-systems are installed on site, they can be commissioned and tested individually. Some testing is essential at this stage, for example through testing of interlocking functions with signalling objects, to satisfy the requirements of traditional signalling testers, who must verify that (for example) a command to reverse a particular set of points does result in the correct points moving in the correct direction.

Once the entire system is installed on site, and tested for basic function, some testing will still be needed, to verify aspects of system behaviour that cannot be fully ascertained in the test lab.

This site testing may typically include:

Verification that key critical system behaviour are replicated in the real world, for example, proving that a train that appeared to stop before an EoA during lab testing actually stops before the real Block Marker.

Many of these on-site tests involve running a test train, which can be difficult and disruptive on a working railway.

Areas to be tested practically on site include:

- The Quality of Service achieved by mobile telecoms between the infrastructure and trains
- The sequence of operations at system boundaries, where trains are establishing and relinquishing communication sessions with RBCs
- The behaviour of trains entering and leaving fitted areas, where timings may be a critical factor.

13.2 The Implications of ATO, Moving Block and Virtual Coupling on Testing Methods

In this section, the possible impacts of the introduction of ATO, ETCS Level 3 (i.e. train location reported by the trains themselves, rather than by infrastructure-based train detection systems such as track circuits or axle counters), and virtual coupling (i.e. running at less than full braking distance behind other trains) are considered.

13.2.1 Implications of ATO on Testing Methods

It is important to understand the implications of ATO on system testing methods. There is a risk that the behaviour of an ATO system, controlling the train's speed in accordance with additional running commands (such as reductions below the maximum safe speed for traffic regulation purposes) could mask a flaw in the Automatic Train Protection (ATP) functionality, if ATO were to be used during system testing.

Therefore, complete testing of the system without ATO is essential to give confidence in ATP system behaviour. Only when that is achieved, can ATO be switched on, and further testing carried out to verify the performance of the ATO sub-system.

13.2.2 Changes to Interfaces

The reliance on train-based position reporting, rather than track-based reporting via the interlocking, reduces the number of different interfaces involved in dynamic train testing. But, in practice, many applications will retain some track-based train detection, to aid recovery from disruptions and failure situations, so the full benefit may not be realised initially.

13.2.3 Testing Complexity

The testing of 'hybrid' systems, involving the possibility of running in either ETCS Level 2 with fixed block track-based train detection, and ETCS Level 3, with train-based position reporting (either 'virtual fixed block' or 'moving block') inevitably increases the complexity of system testing significantly. Two or more separate modes of operation must each be thoroughly tested, and (possibly more significantly), all possible scenarios leading to a change of operating mode, in normal or abnormal conditions, must be tested.

Experience from testing early ETCS Level 2 applications, where a 'fallback' to Level 1 operation was provided for in the system design, showed that the cost and time implication of including alternative operating modes is considerable.

As discussed above, the inclusion of ATO in the system leads to two phases of testing, one with ATO switched off, and a further phase with ATO active. There will also be further scenarios for switching between non-ATO and ATO modes to test, in both expected and unexpected situations.

The addition of Virtual Coupling introduces another variable into the system. ATO may well be considered to be an essential part of the control of the coupling and uncoupling processes between trains. But even trains with ATO active may have a choice (perhaps made by the driver) of implementing virtual coupling, or not implementing it. For example, it might be decided that ATO or virtual coupling cannot be used in extreme weather conditions, due to the possibility of unpredictable reductions in wheel-rail adhesion.

In summary, each time a choice of operating mode is added, there is an increase in (potentially to more than double) the amount of testing required. This should be borne in mind when specifying a new train control system.

13.2.4 Implications of moving block on Testing Methods

Level 3 ETCS does not demand moving block capability.

Little is currently defined with regard to the RBC implementations of moving block or virtual coupling. It is important that these implementations take into account the foreseen testing processes. It is quite possible that these modes can be based on mathematical operations (equations) with less reliance on configuration bespoke to each application. It is the testing of the configuration against each scheme that is currently the most time-consuming testing element which must be kept at a manageable level when introducing these new modes.

What we have in mind here is that there may be a cut-off point. Practical implementations taking advantage of vehicle position-based train detection as an evolution from current block systems will in fact use ever decreasing block sizes (real or virtual). There is then a potential step change to achieve 'pure' moving block and virtual coupling.

13.2.5 Testing with Virtual Trains

Once all the reporting of train location is by data over the airgap, a new testing technique becomes possible.

A real train, running on the line, may follow a 'virtual train' created by a simulator connected to the RBC. This virtual train can be arranged to run at constant speed, to accelerate, or to brake, to suit the particular test, and the actual behaviour of the following 'real train' can be monitored and recorded. Particularly when virtual coupling is being tested, the idea of following a 'virtual train' at less than braking distance, rather than a real train, will be particularly attractive to railway authorities.

This technique may be further extended, by arranging for the safety margin behind a 'real train' running ahead of the test train to be artificially extended (e.g. by telling the system that the train is significantly longer than it really is), enabling a real train to demonstrate scenarios such as

Coupling, Coupled Running, Intentional Decoupling, and Unintentional Decoupling, without risk of a malfunction resulting in a collision.

Most of the comprehensive testing of moving block and virtual coupling will have been carried out on test lines prior to the implementation of a commercial project, but the techniques may well have value when carrying out confirmatory testing of installed systems on a commercial application of the new technology.

13.3 Automation of System Testing

This section discusses feasible targets for automation of system testing for commercial projects involving hybrid systems, moving block and virtual coupling.

13.3.1 Setting Up and Running Tests

Setting up and running tests on a test bench involves the following steps:

- Defining and storing a set of test scenarios which together allow all the necessary behaviour to be checked. These scenarios could involve
  - Boundary testing – trying to create near-worst case situations to test the system response in each instance of each scenario, and/or
  - Path testing – trying to create tests that exercise each route through a flow chart for the response to each instance of each situation (e.g. driver acknowledging/not acknowledging, train speed being above/below limit, train ahead accelerating/decelerating, etc.)
- Loading the data to create each test scenario on to all the simulators/actual equipment comprising the test bench
- Running the test, ideally with the driver actions automatically generated,
- Logging the test results
- Presenting the test results to human analysts in a way which allows rapid checking, and rapid identification of unexpected results
- Comparing results obtained from successive runs of the same test, to allow regression testing of system changes
- Comparing results from a run of a test with results obtained from the same test performed on-site with an actual train.

13.3.2 Automation Concept

It seems unlikely that the first step above is capable of automation. A skilled person must define the level of testing required, and the list of tests to be carried out.

The remaining steps are potentially capable of being automated, although the final three steps will involve a person understanding the significance of the results presented by the automated tool.

The ideal concept would be to record 'test scripts' in a tool, and for that tool to upload the necessary test data to all other parts of the test bench.

It is unclear, without more detail being developed, whether 'generic test scripts' could be generated, and then applied automatically to each instance of the particular scenario present in the system control area. If that was not possible, the details of each instance of every scenario

would have to be entered into the tool to create a full set of tests.

13.3.3 Feasibility

The requirements for automation of such a test bench would appear to be very complex, particularly when the test bench itself is reconfigurable to involve simulations or actual target equipment for some of its elements, and possibly may include equivalent items from different suppliers, as well.

References

[1] X. Zeng, C. Tao, Z. Niu, and K. Zhang, 'The study of railway control system model', in *2010 5th IEEE Conference on Industrial Electronics and Applications*, 2010, pp. 1424–1428, doi: 10.1109/ICIEA.2010.5514848.

[2] S. Oh, Y. Yoon, M. Kim, and Y. Kim, 'ATP functional allocation for Korean radio based train control system', in *2012 12th International Conference on Control, Automation and Systems*, 2012, pp. 1157–1160.

[3] Y. Wei, H. Lu, and Z. He, 'Research of the Digital Communication System for CBTC Based on 802.11', in *2011 Third International Conference on Multimedia Information Networking and Security*, 2011, pp. 95–99, doi: 10.1109/MINES.2011.26.

[4] Z. Zhang, X. Liu, and K. Holt, 'Positive Train Control (PTC) for railway safety in the United States: Policy developments and critical issues', *Utilities Policy*, vol. 51, pp. 33–40, Apr. 2018, doi: 10.1016/j.jup.2018.03.002.

[5] A. Janhsen, K. Lemmer, B. Ptok, and E. Schnieder, 'Formal Specifications of the European Train Control System', *IFAC Proceedings Volumes*, vol. 30, no. 8, pp. 1139–1144, Jun. 1997, doi: 10.1016/S1474-6670(17)43974-7.

[6] 'White Paper European Railway Traffic Management System - ERTMS'. [Online]. Available: http://www.climateaction.org/images/uploads/documents/European_Railway_Traffic_Management_System.pdf. [Accessed: 30-Jan-2019].

[7] N. Furness, H. van Houten, L. Arenas, M. Bartholomeus. 'ERTMS Level 3: the Game-Changer', *IRSE News*, vol. 232, pp. 2-9, April 2017.

[8] M. Biagi, L. Carnevali, M. Paolieri, and E. Vicario, 'Performability evaluation of the ERTMS/ETCS – Level 3', *Transportation Research Part C: Emerging Technologies*, vol. 82, pp. 314–336, Sep. 2017, doi: 10.1016/j.trc.2017.07.002.

[9] J. S. Keränen and T. D. Räty, 'Model-based testing of embedded systems in hardware in the loop environment', *IET Software*, vol. 6, no. 4, p. 364, 2012, doi: 10.1049/iet-sen.2011.0111.

[10] Gonzalo Solasa,*, Jaizki Mendizabala, Leonardo Valdiviaa, Javier Añorgaa, Iñigo Adina, Adam Podhorskia, Stanislas Pinteb, Luis Gerardo Marcosb, Jesús Ma Gonzálezc, Francisco Cosínc, 'Development of an advanced laboratory for ETCS applications', 04-Feb-2019. [Online]. Available: https://www.usenix.net/legacy/publications/library/proceedings/lisa94/full_papers/eirich.ps. [Accessed: 04-Feb-2019].

[11] Y. Wang, L. Chen, D. Kirkwood, P. Fu, J. Lv, and C. Roberts, 'Hybrid Online Model-Based Testing for Communication-Based Train Control Systems', *IEEE Intelligent Transportation Systems Magazine*, vol. 10, no. 3, pp. 35–47, 2018, doi: 10.1109/MITS.2018.2842230.

[12] D. M. Rafi, K. R. K. Moses, K. Petersen, and M. V. Mäntylä, 'Benefits and Limitations of Automated Software Testing: Systematic Literature Review and Practitioner Survey', in

Proceedings of the 7th International Workshop on Automation of Software Test, Piscataway, NJ, USA, 2012, pp. 36–42.

[13] European Commission, 'Directive 2008/57/EC of the European Parliament and of the Council on the interoperability of the rail system within the Community'. .

[14] European Commission, 'Regulation (EU) 2016/796 of the European Parliament and of the Council on the European Union Agency for Railways and repealing Regulation (EC) No 881/2004'. .

[15] European Commission, 'Directive 2004/49/EC of the European Parliament and of the Council on safety on the Community's railways and amending Council Directive 95/18/EC on the licensing of railway undertakings and Directive 2001/14/EC on the allocation of railway infrastructure capacity and the levying of charges for the use of railway infrastructure and safety certification'. .

[16] uropean Commission, 'Regulation 2016/919 on the technical specification for interoperability relating to the "control-command and signalling" subsystems of the rail system in the European Union'. .

[17] European Commission, 'Implementing regulation 402/2013on the common safety method for risk evaluation and assessment and repealing Regulation (EC) No 352/2009'. .

[18] Health and Safety Executive, 'http://www.hse.gov.uk/comah/sragtech/techmeascontsyst.htm'. .

[19] Health and Safety Executive, 'http://www.hse.gov.uk/legislation/hswa.htm'. .

[20] Health and Safety Executive, 'Health and Safety Regulation hsc13'. .

[21] Health and Safety Executive, 'http://www.hse.gov.uk/work-equipment-machinery/puwer.htm'. .

[22] Health and Safety Executive, 'https://www.hse.gov.uk/msd/dse/'. .

[23] European Commission, 'Directive 2006/42/EC of the European Parliament and of the Council on machinery, and amending Directive 95/16/EC (recast)'. .

[24] Health and Safety Executive, 'http://www.hse.gov.uk/work-equipment-machinery/new-machinery.htm'. .

[25] Centre for Connected and Autonomous Vehicles, 'Code of Practice: Automated vehicle trialling'. .

[26] European Commission, 'https://ec.europa.eu/transport/modes/road/news/2018-05-17-europe-on-the-move-3_en'. .

[27] European Commission, 'An Aviation Strategy for Europe COM (2015) 598 final'. .

[28] European Commission, ‘https://ec.europa.eu/transport/modes/maritime/digital-services/safeseanet_en’. .

[29] Health and Safety Executive, ‘http://www.hse.gov.uk/comah/sragtech/techmeascontsyst.htm’. .

[30] International Standards Organisation, ‘ISO/IEC 17020:2012 Conformity assessment — Requirements for the operation of various types of bodies performing inspection’. .

[31] Civil Aviation Authority, ‘Safety Management Systems (SMS) guidance for organisations CAP 795’. .

[32] European Commission, ‘Decision on modules for the procedures for assessment of conformity, suitability for use and EC verification to be used in the technical specifications for interoperability adopted under Directive 2008/57/EC of the European Parliament and of the Council’. .

[33] European Commission, ‘https://www.easa.europa.eu/easa-and-you/aircraft-products/aircraft-certification’. .

[34] UK Government, ‘https://www.gov.uk/government/collections/vehicle-safety-standards-information-sheets’. .

[35] ‘Directive 1999/35/EC on a system of mandatory surveys for the safe operation of regular ro-ro ferry and high-speed passenger craft services’,. .

[36] European Commission, ‘Directive 2001/96/EC establishing harmonised requirements and procedures for the safe loading and unloading of a bulk carriers’. .

[37] European Commission, ‘Directive 2003/25/EC on specific stability requirements for ro-ro passenger ships’. .

[38] European Commission, ‘Directive 2009/45/EC (former 98/18/EC) on safety standards and rules for passenger ships (on domestic voyages)’. .

[39] ‘Introduction (Chapter 1) - Software Testing’. [Online]. Available: https://www.cambridge.org/core/books/software-testing/introduction/A171D03CE38EC74FD89FD6A715A053D1. [Accessed: 29-Jan-2019].

[40] J. Watkins and S. Mills, *Testing It: An Off-the-Shelf Software Testing Process*, 2nd ed. Cambridge: Cambridge University Press, 2010.

[41] ‘IEEE Std 829-2008, IEEE Standard for Software and System Test Documentation’, p. 132.

[42] I. Hooda and R. Singh Chhillar, ‘Software Test Process, Testing Types and Techniques’, *International Journal of Computer Applications*, vol. 111, no. 13, pp. 10–14, Feb. 2015, doi: 10.5120/19597-1433.

[43] K. Sneha and G. M. Malle, ‘Research on software testing techniques and software automation testing tools’, in *2017 International Conference on Energy, Communication,*

Data Analytics and Soft Computing (ICECDS), 2017, pp. 77–81, doi: 10.1109/ICECDS.2017.8389562.

[44] Y. Singh, *Software Testing*. Cambridge: Cambridge University Press, 2011.

[45] A. S. Yildirim, E. Berker, and M. E. Kayakesen, 'System Level Test Automation in UAV Development', in *2018 IEEE AUTOTESTCON*, 2018, pp. 1–6, doi: 10.1109/AUTEST.2018.8532551.

[46] P. Mahajan, 'Different Types of Testing in Software Testing', vol. 03, no. 04, p. 4.

[47] SUBSET-110: UNISIG Interoperability test - Guidelines, UNISIG, rev 3.5.0, 17/02/2016.

[48] SUBSET-076: Scope of the test specifications, UNISIG, rev 3.2.0, 17/08/2017.

[49] R. S. Pressman, *Software engineering: a practitioner's approach*, 7th ed. New York: McGraw-Hill Higher Education, 2010.

[50] B. Kenner, K. W. Colby, and L. J. Brownell, 'System and method for automated identification, retrieval, and installation of multimedia software components', US6314565B1, 06-Nov-2001.

[51] 'EIRICH, Beam: A Tool for Flexible Software Update'. [Online]. Available: https://www.usenix.org/legacy/publications/library/proceedings/lisa94/full_papers/eirich.a . [Accessed: 22-Nov-2018].

[52] SUBSET-111: Interoperability Test Environment Definition UNISIG, rev 3.5.0, 17/02/2016.

[53] R. Skjetne and O. Egeland, 'Hardware-in-the-loop testing of marine control system', *Modeling, Identification and Control: A Norwegian Research Bulletin*, vol. 27, no. 4, pp. 239–258, 2006, doi: 10.4173/mic.2006.4.3.

[54] L. Pivano, N. Husteli, J. Mikalsen, and P. Liset, 'Automated Hardware-In-the-Loop Testing – Experience from Onboard Remote Testing with CyberSea Signature', p. 13, 2015.

[55] D. Ramaswamy *et al.*, 'A Case Study in Hardware-In-the-Loop Testing: Development of an ECU for a Hybrid Electric Vehicle', presented at the SAE 2004 World Congress & Exhibition, 2004, doi: 10.4271/2004-01-0303.

[56] Ma, Zhou, Huang, Melson, James, and Zhang, 'HARDWARE-IN-THE-LOOP TESTING OF CONNECTED AND AUTOMATED VEHICLE APPLICATIONS: A USE CASE FOR QUEUE-AWARE SIGNALIZED INTERSECTION APPROACH AND DEPARTURE'. [Online]. Available: https://www.researchgate.net/profile/Jiaqi_Ma/publication/322852153_Hardware-In-The-Loop_Testing_Of_Connected_and_Automated_Vehicle_Applications_A_Use_Case_for_Queue-Aware_Signalized_Intersection_Approach_and_Departure/links/5a7295dd0f7e9b20d48e18c1/Hardware-In-The-Loop-Testing-Of-Connected-and-Automated-Vehicle-Applications-A-Use-Case-for-Queue-Aware-Signalized-Intersection-Approach-and-Departure.pdf. [Accessed: 18-Mar-2019].

[57] T. Peleska, Jan Aliki, 'Automated Integration Testing for Avionics Systems'. [Online]. Available: http://www.cs.uni-bremen.de/agbs/jp/papers/icstest2002_peleska_tsiolakis-abstract.pdf. [Accessed: 15-Mar-2019].

[58] P. Terwiesch, T. Keller, and E. Scheiben, 'Rail vehicle control system integration testing using digital hardware-in-the-loop simulation', *IEEE Transactions on Control Systems Technology*, vol. 7, no. 3, pp. 352–362, May 1999, doi: 10.1109/87.761055.

[59] Z. Huibing, L. Huan, and P. Dongming, 'Study on Hardware-In-the-Loop Simulation System During Design and Testing of Intermittent Track-to-Train Data Transmission Equipment: BTM', in *2007 8th International Conference on Electronic Measurement and Instruments*, 2007, pp. 4-781-4–785, doi: 10.1109/ICEMI.2007.4351259.

[60] J. Gardiner, 'Delayed Failures in Software Using High Volume Automated Testing', in *Testing: Academic Industrial Conference - Practice And Research Techniques (TAIC PART'06)*, 2006, pp. 193–196, doi: 10.1109/TAIC-PART.2006.6.

[61] V. Garousi and F. Elberzhager, 'Test Automation: Not Just for Test Execution', *IEEE Software*, vol. 34, pp. 90–96, Mar. 2017, doi: 10.1109/MS.2017.34.

[62] S. Thummalapenta, S. Sinha, N. Singhania, and S. Chandra, 'Automating test automation', p. 11.

[63] 'Selenium - Web Browser Automation'. [Online]. Available: https://www.seleniumhq.org/. [Accessed: 05-Feb-2019].

[64] 'Open Source Test Automation Framework | Gauge'. [Online]. Available: https://gauge.org/index.html. [Accessed: 05-Feb-2019].

[65] S. R. Shahamiri and W. M. Nasir, 'Intelligent and Automated Software Testing Methods Classification', p. 6.

[66] D. Flemström, P. Potena, D. Sundmark, W. Afzal, and M. Bohlin, 'Similarity-based prioritization of test case automation', *Software Quality Journal*, vol. 26, no. 4, pp. 1421–1449, Dec. 2018, doi: 10.1007/s11219-017-9401-7.

[67] L. Chen and Q. Li, 'Automated test case generation from use case: A model based approach', in *2010 3rd International Conference on Computer Science and Information Technology*, 2010, vol. 1, pp. 372–377, doi: 10.1109/ICCSIT.2010.5563772.

[68] M. S. Mathieu Steiner Anthony Faucogney, 'Model based testing on a railway project (CBTC)'. [Online]. Available: https://www.all4tec.net/sites/default/files/model-based_testing_on_a_railway_project_all4tec.pdf. [Accessed: 04-Mar-2019].

[69] X. Qu, M. B. Cohen, and K. M. Woolf, 'Combinatorial Interaction Regression Testing: A Study of Test Case Generation and Prioritization', in *2007 IEEE International Conference on Software Maintenance*, 2007, pp. 255–264, doi: 10.1109/ICSM.2007.4362638.

[70] D. V. Kumar and M. Kumar, 'Test Case Prioritization Using Fault Severity', vol. 1, no. 1, p. 5, 2010.

[71] S. Tahvili, M. Saadatmand, S. Larsson, W. Afzal, M. Bohlin, and D. Sundmark, ‘Dynamic Integration Test Selection Based on Test Case Dependencies’, in *2016 IEEE Ninth International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, Chicago, IL, USA, 2016, pp. 277–286, doi: 10.1109/ICSTW.2016.14.

[72] Y.-W. Tung and W. S. Aldiwan, ‘Automating test case generation for the new generation mission software system’, in *2000 IEEE Aerospace Conference. Proceedings (Cat. No.00TH8484)*, 2000, vol. 1, pp. 431–437 vol.1, doi: 10.1109/AERO.2000.879426.

[73] A. Srivastava, ‘Effectively prioritizing tests in development environment’, in *In Proceedings of the International Symposium on Software Testing and Analysis*, 2002, pp. 97–106.

[74] M. Staats, P. Loyola, and G. Rothermel, ‘Oracle-Centric Test Case Prioritization’, in *2012 IEEE 23rd International Symposium on Software Reliability Engineering*, 2012, pp. 311–320, doi: 10.1109/ISSRE.2012.13.

[75] D. Winkler, R. Hametner, T. Östreicher, and S. Biffl, ‘A framework for automated testing of automation systems’, in *2010 IEEE 15th Conference on Emerging Technologies Factory Automation (ETFA 2010)*, 2010, pp. 1–4, doi: 10.1109/ETFA.2010.5641264.

[76] R. Anbunathan and A. Basu, ‘Automation framework for test script generation for Android mobile’, in *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT)*, 2017, pp. 1914–1918, doi: 10.1109/RTEICT.2017.8256930.

[77] R. G. Hamlet, ‘Testing Programs with the Aid of a Compiler’, *IEEE Transactions on Software Engineering*, vol. SE-3, no. 4, pp. 279–290, Jul. 1977, doi: 10.1109/TSE.1977.231145.

[78] D. Peters and D. L. Parnas, ‘Generating a test oracle from program documentation’, in *Proceedings of the 1994 International Symposium on Software Testing and Analysis (ISSTA*, 1994, pp. 58–65.

[79] A. R. F. Giuseppe A. Di Lucca, ‘Testing Web-based application’.

[80] M. E. Yousif, S. R. Shahamiri, and M. B. Mustafa, ‘Test oracles based on artificial neural networks and info fuzzy networks: A comparative study’, in *2015 IEEE 10th Conference on Industrial Electronics and Applications (ICIEA)*, 2015, pp. 467–471, doi: 10.1109/ICIEA.2015.7334158.

[81] R. Just and F. Schweiggert, ‘Automating Software Tests with Partial Oracles in Integrated Environments’, in *Proceedings of the 5th Workshop on Automation of Software Test*, New York, NY, USA, 2010, pp. 91–94, doi: 10.1145/1808266.1808280.

[82] A. B. Kolluri, ‘Effective Bug Tracking Systems: Theories and Implementation’, *IOSR Journal of Computer Engineering*, vol. 4, no. 6, pp. 31–36, 2012, doi: 10.9790/0661-0463136.

[83] N. Bettenburg, S. Just, A. Schröter, C. Weiss, R. Premraj, and T. Zimmermann, ‘What Makes a Good Bug Report?’, in *Proceedings of the 16th ACM SIGSOFT International Symposium on Foundations of Software Engineering*, New York, NY, USA, 2008, pp. 308–318, doi: 10.1145/1453101.1453146.

[84] N. Jalbert and W. Weimer, ‘Automated duplicate detection for bug tracking systems’, in *2008 IEEE International Conference on Dependable Systems and Networks With FTCS and DCC (DSN)*, 2008, pp. 52–61, doi: 10.1109/DSN.2008.4630070.

[85] SUBSET-026-1: System Requirement Specification, UNISIG, rev 3.6.0, 13/05/2016.

[86] C. J. Goodman, L. K. Siu, and T. K. Ho, ‘A review of simulation models for railway systems’, in *1998 International Conference on Developments in Mass Transit Systems Conf. Publ. No. 453)*, 1998, pp. 80–85, doi: 10.1049/cp:19980101.

[87] V.P Singh, *System Modeling and Simulation*. .

[88] P. White and R. Ingalls, ‘Introduction to Simulation.’, presented at the Proceedings - Winter Simulation Conference, 2009, pp. 12–23, doi: 10.1109/WSC.2009.5429315.

[89] A. Amraoui and K. Mesghouni, ‘Colored Petri Net Model for Discrete System Communication Management on the European Rail Traffic Management System (ERTMS) Level 2’, presented at the Proceedings - UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, UKSim 2014, 2014, doi: 10.1109/UKSim.2014.110.

[90] S. Qiu, M. Sallak, W. Schon, and Z. Cherfi-Boulanger, ‘Modeling of ERTMS Level 2 as an SoS and Evaluation of its Dependability Parameters Using Statecharts’, *IEEE Systems Journal*, vol. 8, no. 4, pp. 1169–1181, Dec. 2014, doi: 10.1109/JSYST.2013.2297751.

[91] Y. Wang *et al.*, ‘On-line conformance testing of the Communication-Based Train Control (CBTC) system’, in *2016 IEEE International Conference on Intelligent Rail Transportation (ICIRT)*, 2016, pp. 328–333, doi: 10.1109/ICIRT.2016.7588751.

[92] W. Yongbing, W. Daqing, and M. Meng, ‘Composable Modeling Method for Generic Test Platform for Cbtc System Based on the Port Object’, *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 12, 2015, doi: 10.14569/IJACSA.2015.061230.

[93] C.S. Yang, J.S. Lim, J.K. Um, J.M. Han, Y. Bang, H.H. Kim, Y.H. Yun, C.J. Kim, Y.G. Cho, ‘Developing CBTC Software Using Model-Driven Development Approach’. [Online]. Available: http://www.railway-research.org/IMG/pdf/o.3.3.5.2.pdf. [Accessed: 17-Apr-2019].

[94] M. Issad, L. Koul, and A. Rauzy, ‘A contribution to safety analysis of railway CBTC systems using Scola’, 2015, doi: 10.1201/b19094-64.

[95] M. Issad, L. Kloul, and A. Rauzy, ‘Scenario-oriented reverse engineering of complex railway system specifications SCOLA’, *Systems Engineering*, vol. 21, no. 2, pp. 91–104, Mar. 2018, doi: 10.1002/sys.21413.

[96] M. Ghazel, ‘Formalizing a subset of ERTMS/ETCS specifications for verification purposes’, *Transportation Research Part C: Emerging Technologies*, vol. 42, pp. 60–75, May 2014, doi: 10.1016/j.trc.2014.02.002.

[97] M. Benerecetti *et al.*, 'Dynamic state machines for modelling railway control systems', *Science of Computer Programming*, vol. 133, pp. 116–153, Jan. 2017, doi: 10.1016/j.scico.2016.09.002.

[98] S. J. Tueno Fotso, M. Frappier, R. Laleau, and A. Mammar, 'Modeling the Hybrid ERTMS/ETCS Level 3 Standard Using a Formal Requirements Engineering Approach', in *Abstract State Machines, Alloy, B, TLA, VDM, and Z*, vol. 10817, M. Butler, A. Raschke, T. S. Hoang, and K. Reichl, Eds. Cham: Springer International Publishing, 2018, pp. 262–276.

[99] Y. Tian, K. Pei, S. Jana, and B. Ray, 'DeepTest: Automated Testing of Deep-Neural-Network-driven Autonomous Cars', *arXiv:1708.08559 [cs]*, Aug. 2017.

[100] Y. Sun *et al.*, 'Functional Requirements-Based Automated Testing for Avionics', *arXiv:1707.01466 [cs]*, Jul. 2017.

[101] P. Sarla and B. Ramadoss, 'Automation of Combinatorial Interaction Test (CIT) Case Generation and Execution for Requirements based Testing (RBT) of Complex Avionics Systems', 2018, doi: 10.14569/ijacsa.2018.091217.

[102] 'Hardware in the loop test bench using Modelica: A platform to test and improve the control of heating systems | Elsevier Enhanced Reader'. [Online]. Available: https://reader.elsevier.com/reader/sd/pii/S0306261916317238?token=A76050C9347396F770766A7C8CC0862E6215F9780D2C6A8A430E640CEAAE776F27B2F580358CFC0B63527817CC0881C5. [Accessed: 25-Sep-2019].

[103] M. Bouaziz, Y. Yan, J. Soler, M. Kassab, and M. Berbineau, 'Zero On-Site Testing Strategies for Wireless TCMS', *IEEE Communications Magazine*, vol. 57, no. 9, pp. 64–69, Sep. 2019, doi: 10.1109/MCOM.001.1800994.

[104] M. Aguado *et al.*, 'Towards zero on-site testing: Advanced traffic management amp; control systems simulation framework including communication KPIs and response to failure events', in *2014 IEEE 6th International Symposium on Wireless Vehicular Communications (WiVeC 2014)*, 2014, pp. 1–2, doi: 10.1109/WIVEC.2014.6953265.

[105] SUBSET-111: Interoperability Test Environment Definition UNISIG, rev 3.5.0, 17/02/2016.

[106] SUBSET-112: Basics for Interoperability Test Scenario Specifications, UNISIG, rev 3.5.0, 17/02/2016.

[107] 'Capability Systems Engineering Guide', INCOSE, Ilminster, UK, 2014.

[108] 'International Standard - Systems and software engineering -- Life cycle processes -- Requirements engineering', IEEE, IEEE 29148-2018-ISO/IEC/IEEE, 2018.

[109] King, K., 'Systems engineering framework for railway control and safety systems', *IRSE News*, 2016.