This project has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement no. 826250 Assets4Rail).

Deliverable D 6.1

State-of-the-art study and requirement specifications for railway measuring and monitoring systems

Project acronym:	Assets4Rail
Starting date:	01/12/2018
Duration (in months):	
Call (part) identifier:	H2020-S2RJU-2018
Grant agreement no:	826250
Due date of deliverable:	Month 6
Actual submission date:	07/10/2019
Responsible/Author:	SENER
Dissemination level:	CO
Status:	Issued

Reviewed: (yes)

Document history		
Revision	Date	Description
1		First issue

Report contributors		
Name	Beneficiary Short Name	Details of contribution
Fundacio Eurecat	EUT	5.5, 5.6
Asociacion de Investigacion Metalurgica del Noroeste	AIMEN	5.4, 5.5, 5.6, 5.7, 8.1.1,
Università degli Studi di Roma La Sapienza	DICEA	5.1, 6.1, 6.1.1, 6.4, 6.5, 7.1, 8.1.2,
Sener Ingenieria Y Sistemas S.A.	SENER	General Review Chapter 7 and 8
Technische Universitat Berlin	TUB	5.2, 5.3, 5.8, 6.2
Oltis Group As	OLT	5.8
Witt Industrieelektronik Gmbh	WITT	7.2, 8.1.3
Ferrovie dello Stato Italiane S.p.a.	FSI	6.3
Ab Lietuvos Gelezinkeliai	LiRa	5.2

Table of contents

1. Executive Summary

This Deliverable is the result of WP6 and describes the state of the art regarding drivers, benchmarks and emerging solutions of the following topics, that this project will develop:

- Wayside Train Monitoring Systems;
- Sensor system to support track geometry monitoring;
- Data collection and diagnostics of Safety Critical Systems.

For this purpose, we have also considered some relevant findings of other EU project (e.g. IN2SMART, IN2RAIL).

In addition, this Deliverable shows the requirements and specifications of the three above systems, paving the way for the definition of the individual design specifications and architecture design in each of the subsequent workpackages.

In particular, **Section** ¡Error! No se encuentra el origen de la referencia. deals with the state of the art of **Wayside Train Monitoring System**. The system to develop should consider current track degradation models and rolling stock failures that can affect the infrastructure; the assessment methodologies have been also considered. Regarding the possible technologies, we have considered various techniques for image acquisition, processing and storage. Finally, the RFID technologies for vehicle detection and tracking have been investigated.

Section ¡Error! No se encuentra el origen de la referencia. deals with sensor system to support **Track Geometry Monitoring**. In particular, technologies for measuring track geometry parameters and monitoring solutions for assessment of track conditions have been described. Finally, sensors for measuring wheel/rail contact characteristics have been investigated in order to develop, within the subsequent WP8, a measurement concept to be applied on in-service rail vehicles, which must include the measurement of the transversal position of the wheel in relation to the rail.

Further **Section** ¡Error! No se encuentra el origen de la referencia. introduces the relevant technologies and methods for data collection/sensing, communication interfaces, data format and transmission when **Safety Critical** items are involved (i.e. signalling and train control systems). The section considers the inputs from the IN2SMART Shift2Rail Project.

Finally, the objective of **Section 8** is to show a **System Breakdown Structure (SBS)** developed considering the candidate technologies identified within the previous sections. These requirements have been classified into functional, operational, performance and safety and will be the basis for WP7, WP8 and WP9, which will serve as a validation for them.

2. Abbreviations and acronyms

Abbreviation / Acronyms	Description
AF-TC	Audio Frequency Track Circuit
ANN	Artificial neural networks
ANS	American National Standard
ARAT	Active Reader Passive Tag
ARPT	Active Reader Active Tag
ASN	Abstract Syntax Notation One
ATC	Automatic Train Control
BAP	Battery-Assisted Passive
CCD	Charge-Coupled Device
CMOS	Complementary Metal–Oxide–Semiconductor
CNN	Convolutional Neural Networks
COCO	Common Objects in Context
COG	Center of Gravity
CUDA	Compute Unified Device Architecture
DBMS	Database Management System
DBS	Distance-Based Sampling
DIC	Digital Image Correlation
DLP	Digital Light Projector
DNN	Deep Learning ANN
EAS	Electronic Article Surveillance
EBCDIC	Extended Binary Coded Decimal Interchange Code
EEPROM	Electrically Erasable Programmable Read-Only Memory
EMD	Electrical Medium Distance
EPC	Electronic Product Code
ERP	Effective Radiated Power
ERTMS	European Rail Traffic Management System
ESD	Electrical Short Distance
ETCS	European Train Control System
EVN	European Vehicle Number
EVN	European Vehicle Number
FMECA	Failure Mode And Effect Analysis
FPGA	Field-Programmable Gate Array
FTA	Fault-Tree Analysis
GIAI	Global Individual Asset Identifier
GPIO	General Purpose Input/Output
GSM-R	Global System for Mobile Communications – Railway
GTIN	Global Trade Item Number
HF	High Frequency
HOG	Histogram of Oriented Gradients

HVAC	heating ventilation and air conditioning
IM	Infrastructure Manager
IoT	Internet of Things
IP	Internet Protocol
IT	Information and Technology
KDD	Knowledge Discovery in Databases
LAN	Local Area Network
LDA	Linear Discriminant Analysis
LF	Low Frequency Band
LLPA	Transmission of Long Line Public Address
LPWAN	Low-Power Wide-Area Network
LVDT	Linear Variable Differential Transformer
M&E	Maintenance & Engineering
MDD	Multi-Depth-Deflectometer
MEM	Micro Electrical Mechanical device
MGTs	Million Gross Tonnes
MNIST	Modified National Institute of Standards and Technology
MPLS	Multiprotocol Label Switching
MRO	Maintenance, Repair and Overhaul
OCR	Optical Characters Recognition
OGF	Optical Glass Fiber
OMAP	Open Multimedia Application Platform
OSI	Open Systems Interconnection
PDC	Portable Data Capture
PMD	Photonic Mixer Device
PSD	Position Sensitive Device
PSO	Particle swarm optimization
RAMS	Reliability, Availability, Maintainability and Safety
RCF	Rolling Contact Fatigue
RFID	Radio-Frequency IDentification
RTF	Readers talks first
RU	Railway Undertaking
SBS	System Breakdown Structure
SDH	Synchronous Digital Hierarchy
SGD	Stochastic gradient descent
SIFT	Scale-invariant feature transform
SOA	State of the art
SURF	Speeded Up Robust Features
SVM	Support Vector Machine
TCN	Train Communication Network
TCP	Transmission Control Protocol
TDMA	Time division multiple access

TID	Tag Identification
TOF	Time Of Flight
TRC	Track Recording Coach
TRV	Track Recording Vehicle
TTF	Tag talks first
UGMS	Unattended Geometry Measurement Systems
UHF	Ultra High Frequency
VGCS	Voice Group Call Service
VLAN	Virtual Local Area Network
WORM	Write Once Read Many
WSN	Wireless sensor network
WTB	Wire Train Bus
WTB	Wire Train Bus
WTMS	Wayside Train Monitoring System

3. Background

The aim of the work package 6 (WP6) is to provide a review of the state-of-the-art technology for the following technologies:

- Train monitoring systems (WP7);
- Sensor systems for track geometry monitoring (WP8);
- Data collection for diagnostics from signalling components (WP9).

This state of the art review will consider previous EU/national funded projects, industry solutions and research papers, to identify drivers, benchmarks and emerging solutions, not only within the railway sector. The WP6 will also assess the potential applicable technologies and solutions that the subsequent work packages will consider.

4. Objective/Aim

Considering the developments in Assets4Rail WS2 are expected partly as a complement to the ongoing S2R projects, WP6 has investigated in collaboration with the complementary projects on their approaches and achievements so far in the first step, avoiding any duplications and ensuring the compatibility and integrability of the Assets4Rail developments.

The state-of-the-art study has reviewed previous EU/national funded projects, industry solutions and research papers to identify drivers, benchmarks and emerging solutions of the topics related to the subsequent developments.

In addition, WP6 has defined the requirement specifications of the three systems to be developed, paving the way for the definition of the individual design specifications and architecture design.

5. Wayside Train Monitoring System

One of the objectives of Assets4Rail is to develop a **Wayside Train Monitoring System (WTMS)**. This WTMS consists of a trackside stereo computer vision system and an underframe image monitoring system, incorporating the RFID technology for train identification (considering the whole train as an asset).

The developments will start with the assessment of rolling stock failures and how they impact on the infrastructure. This is based on the fact that the component positions and the features of the failure modes have significant influence on the design and configurations of computer vision systems, such as position and the amount of the required cameras and flash units, distances to track, etc.

Firstly, the common rolling stock failure modes which cause direct damages on infrastructure can be identified based on the knowledge and historical data provided by Railway Undertaking (RU) and Infrastructure Manager (IM). With this information, we can develop a **Failure Mode, Effects and Criticality Analysis (FMECA)** to prioritise and assess the critical failures and defects and, identify the key components from the infrastructure point of the view. As a case study, the effect of a critical rolling stock failure on infrastructure, especially track, will be analysed in-depth and quantified.

In order to inspect automatically the prioritised failures and components, two (2) **computer vision systems**, with specific hardware configurations and algorithms will be developed: A trackside stereo imaging system will be used for obtaining a 3D analysis of the rolling stock and an underframe imaging system will be used to measure the critical vehicle components. To link the monitoring results to the monitored vehicles, an identification system is necessary. In the Assets4Rail project, **RFID technology** will be used for vehicle identification. Apart from this function, we will investigate other potential usages of RFID technology in context of railway.

In the following subsections, the foundations of the aforementioned developments are introduced, including as-is situation, theories, methodologies, off-the-shelf available products, potential applicable algorithms, etc. To complete the required context for this activity, we will summarise the main achievements and developments already available from the complementary project **IN2SMART**.

5.1 Relevant works in IN2SMART

Deliverable 6.2 of **IN2SMART** (https://projects.shift2rail.org/s2r_ip3_n.aspx?p=IN2SMART) describes two interesting technologies regarding vehicle inspection. The first one is video based technology to detect vehicle overall defects and to identify and to characterise vehicles; the second one is the RFID technology for rail vehicle identification.

5.1.1 Video inspection of vehicle defects

The vehicle defects detection solution proposed makes use of a single linear camera for vehicle side, together with a well-defined linear illuminator. The linear camera is used for precise measurement of specific features, such as friction wedge or suspension spring height. In order to make this possible, the camera has to be first of all calibrated with reference to the position from where it acquires the passing trains. This activity is performed only once and its results are to be considered as valid as long as the camera position and field of view orientation does not change.

The main advantages related to a linear camera are the high frequency rate, allowing the acquisition of trains running at full speed, and the possibility to better focus and control illumination, being this concentrated only on the vertical plane acquired by the linear camera. Data is collected in a real railway line, with trains travelling at speeds up to 180km/h, to be used as inputs for the overall proposed solution.

After the camera has been calibrated it could be used to acquire images from trains and process these single vehicle 2D images.

A post-processing step is necessary first of all to correctly weight the vehicle images depending on the vehicle speed at the measurement point and secondly to enhance image quality for further feature extraction.

The feature detection step uses different data processing algorithms to detect automatically the position of relevant features such as suspension spring height or friction wedges. After this, the detected feature measure in number of pixels is first obtained and camera calibration configuration is applied to get the same measure in real world units, such as millimetres.

Finally, if configured, a defect alarm could be automatically detected. The methodology will end with a test on many real acquired images to prove the feasibility of the solution.

First results, based on the measurement of the brake shoe thickness, demonstrate the feasibility of the proposed solution and its robustness in relation to some common injected disturbances, such as horizontal blurring and small modifications to camera tilt angle.

5.1.2 Vehicle identification and characterisation

Two concepts for Vehicle identification and characterisation will be evaluated. One is using image recognition which will be compared with the current RFID technologies standard where a self-sufficient installation has been evaluated.

5.1.2.1 Vehicle identification using image recognition

The solution proposed is based on video identification of vehicle id number, resolving all these situations in which vehicles are not equipped with an RFID tag. The same acquisition components and setup for the overall defect detection are used, exploiting the advantages of a linear camera and illuminator. The overall approach proposed for this use case is based on the usage of a supervised learning model to detect different characters. The model is trained with real vehicles images. Best solutions are investigated for text detection and segmentation in separate characters. Knowledge from the railway domain, such as UIC number format and normal vehicle ID position are used to make the solution more robust. The first results obtained demonstrate the feasibility of the proposed solution with reference to the segmentation and recognition of a few freight vehicles.

5.1.2.2 Vehicle identification using RFID

RFID for rail vehicle identification is implemented in a European standard and used in several countries. This use case is aimed at evaluating this standard and finding a good solution for a self-sufficient RFID-reader installation. The current installations will be compared to design a good set up for the purpose. Results show that a good RFID-installation can detect almost 100% of train passages in full speed, together with axle-sensors it can also separate unidentified vehicles and match the identity with belonging axles. RFID is a reliable technique to identify vehicles but the standard needs to be more accurate and stable for achieving full efficiency.

5.2 Track degradation models

Track degradation is a failure process, which causes negative effects on passenger's comfort, vehicle behaviour, resident's conditions (track generates additional noise at the vehicle pass through the defect) and the track itself.

From the perspective of the vehicle-track system, the large impact force induced by the deterioration of tracks will further exacerbate the deterioration of vehicle components (bearing, axles, wheelsets, etc.) and infrastructures (rails and sleepers, etc.), greatly shortening their service life and increasing the risk of derailment. Besides, track degradation will increase energy consumption and maintenance costs. It is necessary to develop an effective track degradation model.

According to Oxford Economics, global infrastructure spending will exceed $9 trillion per year by 2025. The European countries are reported to allocate 15 - 25 billion EUR annually on maintenance and renewals for a railway system consisting of about 300 000 km of track, half of which is electrified, giving an average of 70 000 EUR per km track and year (Lidén, 2015). Investments will be targeted not only towards new infrastructure, but also towards renewal of existing assets. Once an asset degrades beyond a certain point, it needs to be renewed or replaced. In asset management, degradation models are developed to infer and predict how and when assets degrade. Degradation can be described by a number of variables (condition, performance, reliability, etc.) that change over time (or usage).

Public roads, railway or other infrastructure has its own life cycle. Also country by country have different infrastructure life cycles. Railway, like other industry infrastructure consists of many different components, who also have varying life cycles of each part.

On many railways, synthetic measures are used to determine the condition of the railway superstructure, mainly in terms of its geometry. A very simple indicator was used by Japanese railways. It expressed the ratio of the track length, on which the deviations of vertical irregularities exceeded ± 3 mm l±3, to the total length of the assessed track (Onodera, 2005). This indicator, being, in fact, the defectiveness of only one irregularity of the track, has recently been supplemented by the standard deviation of this irregularity and its maximum amplitude. In the repair planning support system, these three measures are used, assuming that they provide a more complete picture of the track condition. [4]. Knowing, for example, that the increase in irregularity of the track after track renewal is 0.36 mm/100 days, it is possible to predict when they will reach the limit (Baluch, 2018).

On English railways, the role of the synthetic indicator, to a certain extent, played by the standard deviation of vertical irregularities referred to the wavelength of 35 m. At train speeds between 115 and 125 mph (185 – 201 km/h), the limit value of this deviation is 4.7 mm (Baluch, 2018).

On the Malmbanan railway line in Sweden, which is used to transport ore and where the maintenance of the railway superstructure is particularly difficult, the basic parameters for assessing the railway superstructure condition are the standard deviation of vertical irregularities and the amplitude of these irregularities. A standard deviation exceeding 2.5 mm or an amplitude exceeding 10 mm qualifies the track for repair (Baluch, 2018).

In Poland, for almost 30 years, the synthetic indicator of the track condition has been used as the basic parameter of the geometrical condition of the railway superstructure. It is based on the standard deviations of four geometrical values: standard deviations of vertical and horizontal irregularities, twist and track gauge (Baluch, 2018).

A number of factors influence to the degradation of rail tracks, including the sleepers, fastenings, ballast, age of rails and axle load, speed, traffic density, traffic type, rail-wheel interaction, Million Gross Tonnes (MGTs), rail size, track curvature, rail track elevation, rail profile, and rail track construction, rail track super-elevation and rail welding, and rail lubrication (Elkhoury, Moridpour, et al., 2008).

Like its described above, there is lot of evaluation methods and they vary country by country. Currently, the track degradation models can be mainly classified into the following four categories (see in Fig. 1): mechanistic model, statistic model, mechanical-empirical model and artificial intelligence (AI) model. Each of those models have its own strengths and weaknesses. However, none of them considers the factor of rolling stock failures. As ton-kilometer of defect railway vehicles is negligible in comparison to normal vehicles, the general effects of rolling stock failures on track degradation are thus not taken into account in these models.

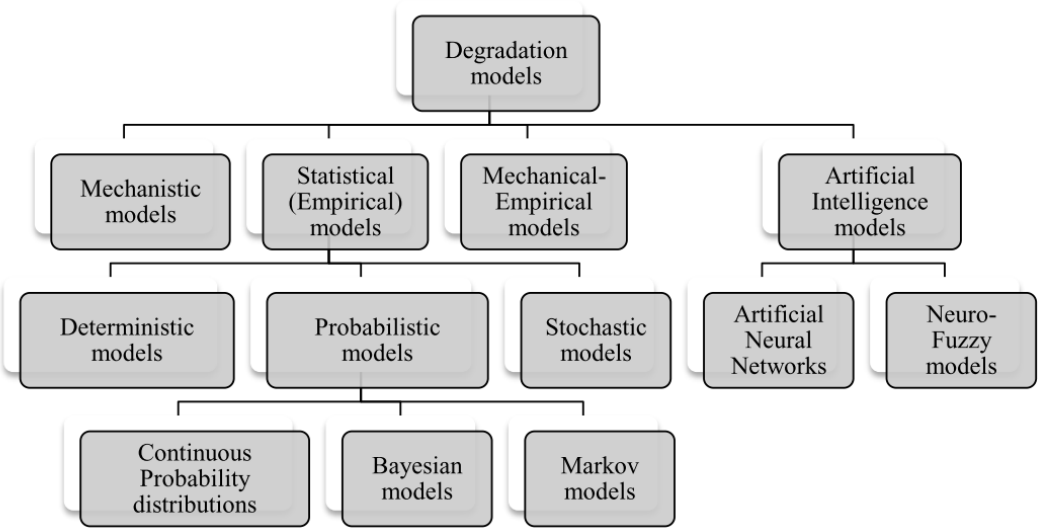


Figure 5.1 Classification of rail degradation models (Elkhoury, Moridpour, et al., 2008).

5.3 Rolling stock failures and assessment methodologies

A rolling stock failure may cause delays and disruptions to transport services or even result in catastrophic derailment accidents. There are several tools and techniques that are currently used to determine and evaluate the risk of failures occurring in engineering systems throughout their entire life cycle: from the design to production, operation and maintenance.

This document will focus on the rolling stock failures, which cause damage on infrastructure such as track degradation. Some possible rail degradations, include worn out rails internal defects, weld problems and rolling contact fatigue (RCF) originating such as surface cracks, head checks, squats, spalling and shelling. Other effects of rolling stock failures on track degradations, have been detected in Sweden, where railway operators are introducing low-cost rolling stock to reduce costs, what increase track degradation.

Critical rolling stock failures that result in catastrophic impacts on infrastructure cannot be tolerated and can be detected by numbers of existing monitoring systems. Non-safety-related failures can accelerate the degradation of railway infrastructure. However, the impacts of such failures can hardly be assessed. In the Assets4Rail project, a methodology is defined quantify the dynamic impacts of non-safety-related vehicle failures on railway infrastructure.

5.3.1 FMECA methodology

One of the widely used techniques in this regard is the failure mode, effects and criticality analysis (FMECA) which is an extended version of the failure mode and effect analysis (FMEA) method. FMECA methodology in combination with statistical analysis will be used for the evaluation of the impact of rolling stock failures on the infrastructure (e.g. tracks).

FMECA methodology steps are:

Step 1: Select a rolling stock component for the study.

The main rolling stock components that can be considered in a risk analysis study include: door unit, scroll compressor, bogie, pantograph, coupling system, braking unit, air spring suspension and heating ventilation and air conditioning (HVAC).

Another important issue is reducing defects originated mainly from rolling contact fatigue (RCF). Analysis and modelling of RCF initiated defects have been done by many researchers, [4–9] to find out ways of reducing the initiation and propagation of these defects. In this sense, ball and rolling element are used to support load and allow relative motion inherent in the mechanism and have been recognized as one of the main modes of failure by RCF.

Step 2: Collect the component function information.

After the selection of the components involved in track degradation, the next step will be the identification of the functions of each component in rolling stock. The component function information can be collected by answering the following questions, (e.g.: taking the wheels component as an example of rolling stock component):

- What functions do wheels perform?
- Can rolling stock operate without this component?
- Do the wheels contain redundancies or backups?
- Will rolling stock fail if the wheels fail?
- In which ways will the wheels affect the other components or the overall system?

Step 3: Determine potential failure modes that can cause damage to the component through reviewing past failures.

The identification of potential failure modes is an important part of the risk analysis studies. In the present study will be determined by:

- Literature review
- Data collection from:
    - Reviewing past failures
    - Inspection records

Step 4: Identify root causes that contribute to failure of the rolling stock component through interviewing experts from various fields

The failure root causes will be determined by interviewing experts including designers, train operators, inspectors, maintenance technicians, etc. A Fault-Tree Analysis (FTA) will be performed as top-down and deductive failure analysis method through which all undesired events that may lead to system failure will be analysed.

Step 5 and 6: Assign: a) a likelihood rating and b) a severity (consequence) rating, to each failure mode of the rolling stock component

The failure data will be analysed using statistical techniques to create models for estimation of rolling stock defects. The result of these statistical techniques will be the assignment of two different ratios:

a. Likelihood (occurrence) rating: evaluated as total number of failures by duration of time operation
b. A severity (consequence) rating: evaluated as failures which has the following impacts on: safety, social, economic or in the environment.

Both ratios from each failure mode will be assigned on 10-point scale:

- For likelihood (occurrence) rating scale: 1 point represent "remote" and 10 point indicates "almost certain".
- For severity (consequence) rating scale: 1 point represent "no effects" and 10 point indicates "dangerous without warning".

The aim of both scales will allow determining the implication/importance of these failures of the rolling stock component.

Step 7: Evaluate the critically level of a rolling stock failure and prioritise the failure modes in descending order

The risk factors obtained (as the product of likelihood rating by impact rating), for all failure modes are prioritized in descending order and the most critical ones with respect to both reliability and damage severity are identified.

Therefore, this analysis will allow identify the most critical failure modes as the ones occurring most frequently and leading to largest losses.

Step 8: Categorise the failure modes into five classes of critically and proposal potential protective measures to prevent recurrences

The failure modes according to the level of their criticality are categorised into five classes, namely very low, low, medium, high and very high critical. These classes of failure criticality and the associated improvement actions are described in Table 5.1.

Table 5.1 Five classes of failure criticality and the associated improvement actions.

Criticality level	Risk Factor (R)	Recommendation
Very low	$1 \leq R \leq 4$	Almost unnecessary to take the improvement actions
Low	$5 \leq R \leq 9$	Minor priority to take the improvement actions
Medium	$10 \leq R \leq 25$	Moderate priority to take the improvement actions
High	$26 \leq R \leq 49$	High priority to take the improvement actions
Very high	$50 \leq R \leq 100$	Absolute necessary to take the improvement actions.

Figure 5.2 shows the critically matrix, which provides a graphical portrayal of the risk factors obtained from the analysis. For that purpose, both rating scales: occurrence and severity, are represented in each matrix

axis, and different regions of the matrix represent different levels of criticality for rolling stock components, where red cells represent "Very high critical region", whilst the green cells represent "very low critical region".

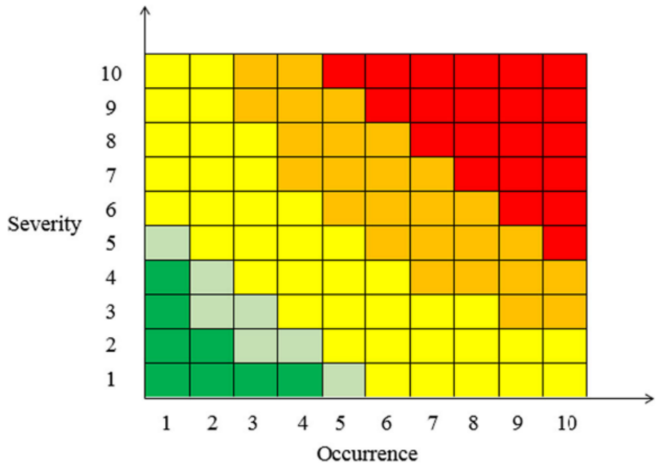


Figure 5.2 A critically matrix for rolling stock failures.

5.3.2 Quantification of failure effects

Once selected the most critical rolling stock failures on infrastructure, a case study will be carried out to quantify the effects of the selected failure. To date, the track degradation mechanisms are mainly classified into four categories: track settlement, track component fatigue, rail wear and rolling contact fatigue (RCF). Although, as mentioned, the current track degradation models do not directly take into account rolling stock failures, there are quantitative methods to link the vehicle characteristics and exerted wheel-rail forces to each category of the track degradation mechanisms (Öberg & Andersson 2009; Smith et al. 2016). On the other hand, the rolling stock failures of interest cause abnormality in wheel-rail contact force. Therefore, the effects of rolling stock failures can be quantified through investigation on the abnormal wheel-rail forces. This concept is illustrated in Figure 5.3. This concept will be implemented through multibody dynamic simulations in different scenarios such us straight track, curves with different radius or rail stiffness.

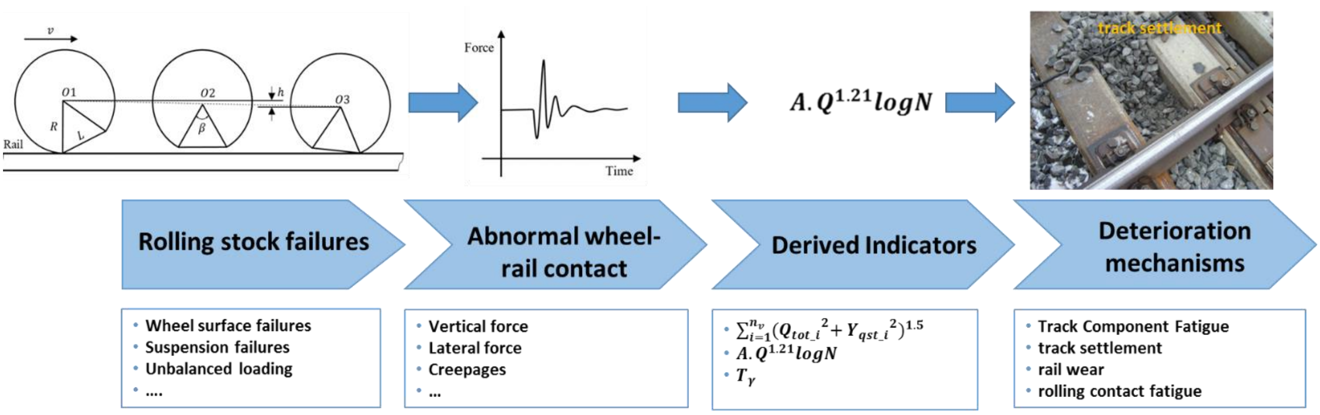


Figure 5.3 Concept for quantification of effects of rolling stock failure on track degradation.

5.4 State of the art computer vision based WTMS

Computer vision based WTMS have been well deployed for train inspection over twenty years, especially in USA and China, where they are mostly used for fault diagnosis of freight wagons, covering geometrical measurements of wheel and brake shoes, failure detection of bogie components and couplers as well as scanning of the whole wagon body. The standard solution is to use a matrix of 2D (video) cameras for specific inspection objectives. The arrangement of the camera matrix, such as illumination devices, position of cameras, number of cameras, etc., depends on the inspection objectives. Some examples of SOA computer vision based WTMS are presented in the following. To be noticed, the examples are mostly three-piece bogie of freight wagons, which are used in USA.

5.4.1 Trackside brake pad inspection

The brake pad is designed to provide friction between the wheel tread and the brake pad (shoe), but the wear rate of the brake is much greater than that of the wheel so that the inexpensive brake wears out before the expensive wheel. The challenge for the rolling stock maintenance manager is scheduling the change of the brake pad early enough so that it does not present a safety risk before the next inspection but late enough to maximize the life of the component and minimize costs.

A typical CV system can measure brake Pad thickness from top to bottom and also check the presence of the securing key and brake pad. 90% of brake pads are correctly processed automatically in one pass. The accuracy could reach ±2mm, while repeatability 1.5 mm.

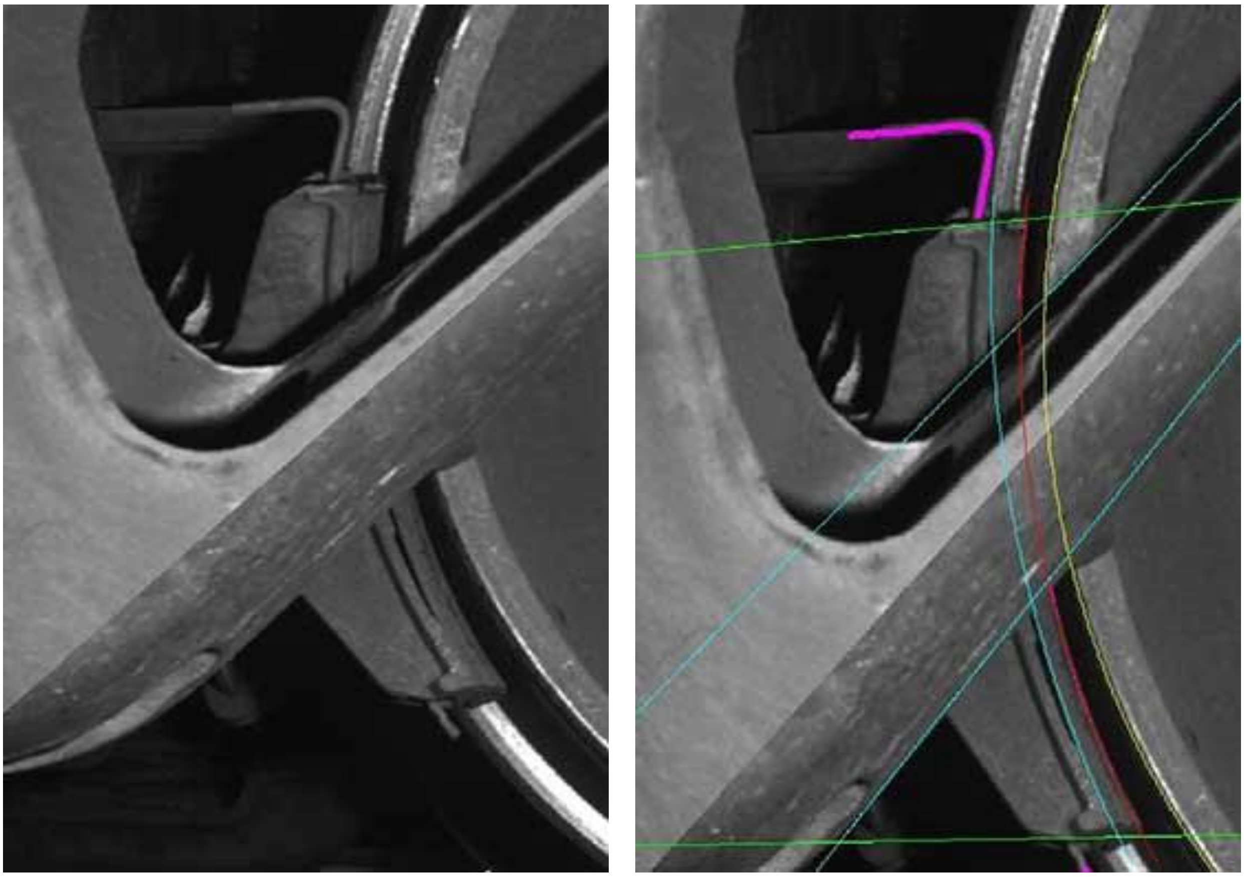


Figure 5.4 Sample Processing Results from Beenavision Automatic Wayside Brake Shoe Inspection system.

5.4.2 Trackside bogie Side-Frame inspection

Bogie frame defects are less common than other components such as brakes and wheels, however, the results of undetected failed components can have catastrophic consequences for the rail operator. Broken, dislodged or missing springs, broken adaptors, missing bearing end-caps and dislodged friction wedges are the common defects that are visible on the side-frame of the bogie.

CV systems analyse the images and identify broken or missing components that are visible on the side-frame of the bogie. Typical failures could be broken or missing springs, missing bearing end-caps, dislodged friction wedges and damaged adaptors.

This project has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement no. 826250 Assets4Rail).

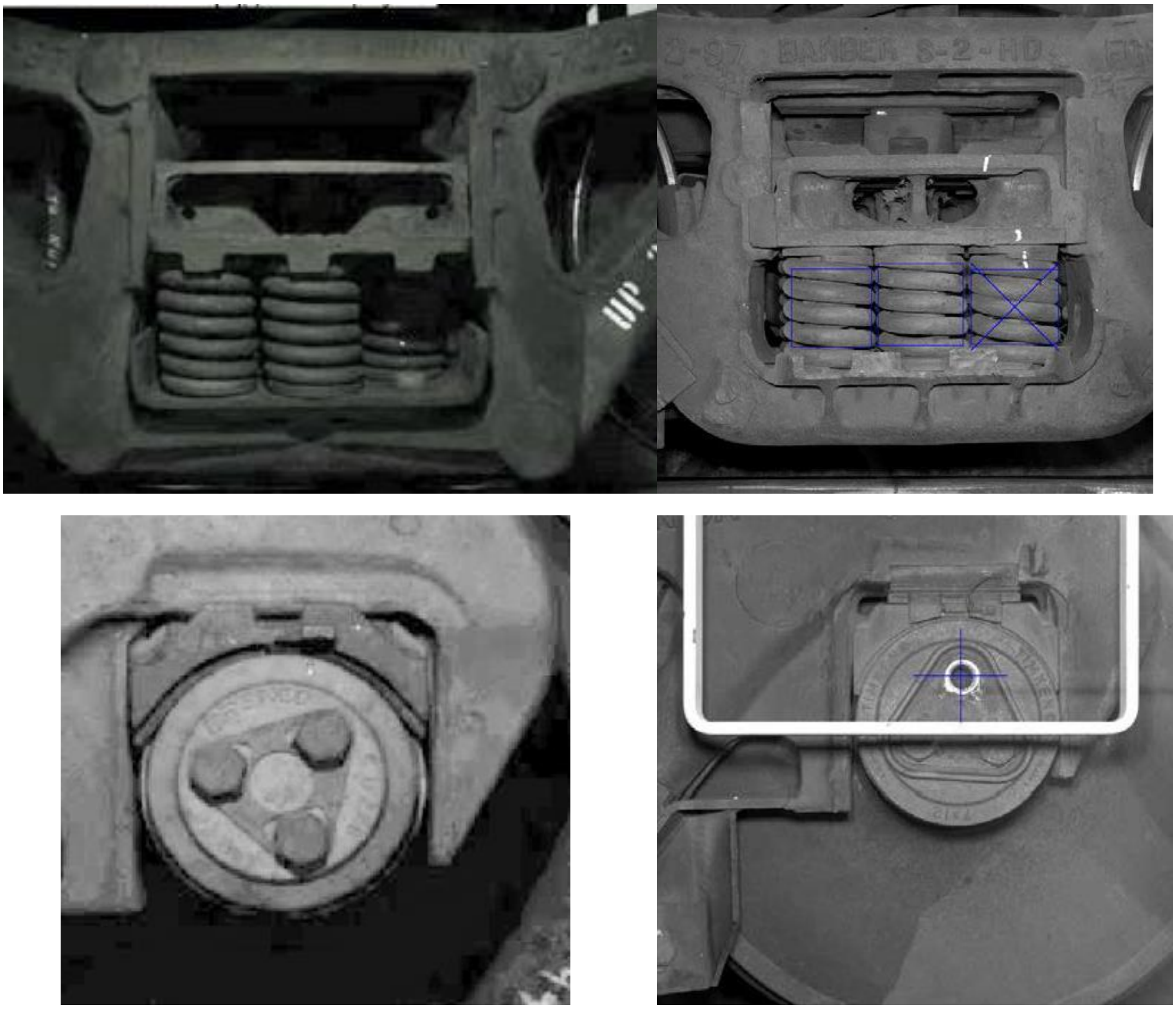


Figure 5.5 Sample Images from Using automated wayside systems for condition-based maintenance and critical events reporting (K. Kilian et al., 2011) and Bogie Side-frame Monitor (Lynxrail, 2014).

5.4.3 Underframe wheel inspection

Underframe wheel inspection systems are designed to measure wheel profiles of moving trains and inspect wheel tread damages, usually combing laser scanning and optical imaging. For instance, The Lynxrail Wheel Profile Monitor is the ideal tool for rolling stock maintenance managers for the efficient and effective management of rolling stock wheels. The WPM consists of six image capture units (three for each wheel) mounted below rail level, each focused on a different part of the wheel. The ICU use a shutter speed of 1/64,000th of a second and have sophisticated electronics to precisely trigger the shutter and store images in a buffer. The flash unit generates an equivalent output power of 30kW for 33 microseconds to provide ample light in the range of interest but keeping input power needs low. The captured images are processed in a server using machine vision algorithms to measure the key wheel parameters.

This project has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement no. 826250 Assets4Rail).

Table 5.2 Accuracy and repeatability of the measured wheel profile parameters.

Parameter	Accuracy	Repeatability
Flange Width	±1.0 mm	0.5mm
Flange Height	±1.0 mm	0.5mm
Wheel diameter	±3.5 mm	
Hollowing	±0.5 mm	
Back to Back	±2.0 mm	
Inner Rim Thickness	±2.0 mm	
Outer Rim Thickness	±2.0 mm	
Wheel Wear of Rim	±1.0 mm	
qR Measurement	±1.0 mm	

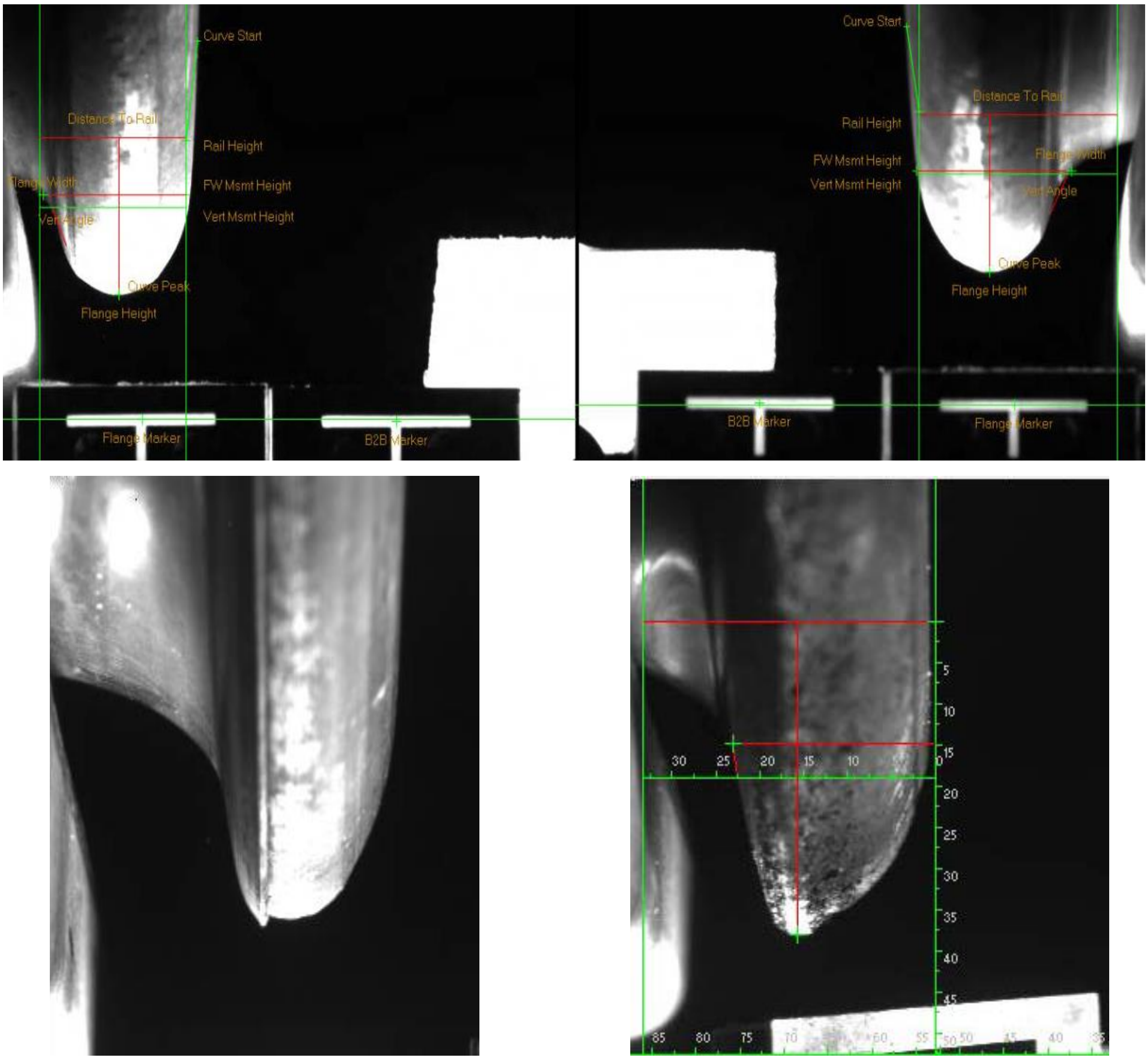


Figure 5.6 Sample Images from Wheel Profile Monitor (Lynxrail, 2014).

5.4.4 Other underframe inspection

CV systems produces high quality images of the structural components of bogie and wagon body in bottom view, such as brake rigging components, coupler components, jacking plate, draft gear carrier, etc. For specific components, some CV systems could automate the inspection and abnormality detection. For other cases, the inspection has to be manually done.

5.5 Hardware products for computer vision systems

The hardware components and potential applicable commercial products are introduced below.

5.5.1 Stereo matrix-scan

Stereo imaging is the process of taking two or more images and estimating a 3D model of the scene by finding matching pixels in the images and converting their 2D positions into 3D depths.

The most common application in stereo correspondence is using two frontal cameras to take photos from the same scene and then, aligning the images from the cameras, obtaining a perception of the depth in the scene.

The human vision works with a similar principle, called motion parallax. According to it, the amount of horizontal motion between the two frames (disparity) is inversely proportional to the distance from the observer, and this information can be used to calculate a 3D model.

The main principles can be explained considering a point X in the real 3D space, being projected simultaneously in two image points x and x' through two camera projection centers C and C'. The points X, C, C', x and x' lie in a plane called epipolar plane. The projection of the rays through x and X, and x' and X are called the epipolar lines of x and x', respectively. This projection is described by epipolar geometry.

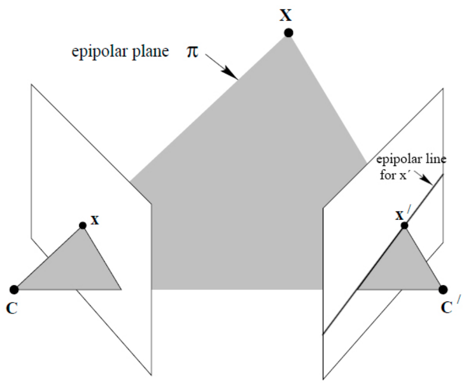


Figure 5.7 Epipolar geometry.

Epipolar geometry used in stereo imaging by comparing these two images, in a process called stereo matching that obtains relative depth information from image differences, usually computed and represented in the form of a disparity map. In this map, objects that are closer to the stereo camera system will have a larger disparity than those that are further away.

In practice, the key problems to solve in stereo vision are:

1. Identify which pixels in multiple images match the same world feature. This is known as the correspondence problem.
2. Identify for each pixel in the image the corresponding ray in 3D space. This is known as the calibration problem, and requires accurate calibration of the camera optical parameters and physical location.

The correspondence problem is solved through image processing software. Depending on the application, the algorithm used may solve correspondences for only a sparse set of features in the image (feature-based algorithms), or attempt to find correspondences for every pixel in the image (dense stereo algorithms).

Camera calibration is done before using the stereo rig. For dense stereo algorithms, typically the images from the cameras must be remapped to an image that fits a pin-hole camera model. This remapped image is called the rectified image. For lenses with barrel distortion, straight lines in the world will appear curved in the image. In the rectified image, however, barrel distortion is removed and straight lines will appear straight. Because of this, most manufacturers that supply stereo systems offer some form of calibration charts and/or software to perform this task.

Most common stereo systems employ area array image sensors based on CCD and CMOS technology. Matrix-scan stereo systems can either build such systems using reference designs or purchase ready-made systems.

Reference designs can be obtained from companies such as e-con Systems (www.e-consystems.com). e-con Systems Capella stereo vision camera reference design, for example, is based on OMAP/DM37x processors from Texas Instruments (www.ti.com) and a stereo camera daughterboard. To develop stereo applications, the design is supplied with and SDK and sample applications for synchronous stereo image capture and depth measurement.

Ready-made stereo systems can be obtained from companies such as FLIR (https://www.flir.com). FLIR 3d solution is the Bumblebee®2 stereo vision camera that includes a GPIO connector for external trigger and strobe functionality. Bumblebee family include cameras from 0.3 MP 648x488, 48 FPS to 0.8 MP 1032x776 20 FPS. Every FLIR Stereo Vision camera system includes a free copy of the FlyCapture SDK, which is used for image acquisition and camera control, and the Triclops SDK, which performs image rectification and stereo processing.

Figure 5.8 Stereo Vision camera.

5.5.2 Stereo line-scan

Stereo imaging techniques are now being be applied to line-scan cameras. Such camera systems are especially useful in applications that require 3D data to be obtained from the surface of moving objects. In the design of its 3D-Pixa stereo line-scan camera, for example, Chromasens (www.chromasens.de), has developed a novel method of imaging 3D surfaces using a 7500x3 (RGB) line scan imager.

To obtain the two images required to create a disparity map, the camera system employs two lenses that are used to focus light onto both the right and left hand sides of the line-scan imager. Raw image data is then transferred from the camera over a Medium Camera link interface at 22kHz.

After images are captured, the system includes CUDA-based software running on an NVIDIA graphics processor that can be used to produce the rectified image and the disparity map generated. Also, 3D-Pixa stereo line-scan supports the libraries: HALCON (MVTec), MIL (Matrox), LabVIEW (National Instruments), Coake (SAC).

Specifications of 3D Line Scan Camera 3DPIXA dual 200µm HR:

- Optical resolution (µm/pixel) 200;
- Field of view (mm) 1400;
- Number of pixel 7000;
- Height resolution (µm) 60;
- Height range (mm) 400;
- Free working distance (mm) 1646.3;
- Maximum speed (mm/s) 4240;
- Line frequency (kHz) 21.2;
- Dimensions LxWxH (mm) 220.3 x 463 x 98.5.

Figure 5.9 Stereo line-scan.

5.5.3 Laser light

While such passive systems use stereo images to create depth map information, other active 3D techniques that employ structured lasers and single cameras can also be used. Such structured laser based systems are often used in web inspection systems to generate a 3D profile of the object being imaged.

In operation, a structured laser light is first projected across the web and reflected light captured by a camera. By measuring the displacement of the reflected laser line across the image a depth map can be calculated. Depending on the application, these systems can either be configured using off-the shelf structured laser light sources with a separate camera or using 3D smart cameras.

In these cameras laser line profile is sampled by the camera system as a line of points with a finite width. To obtain the correct height information from the 2D reflected laser line profile, the center of the reflected Gaussian-like curve must be determined. This can be accomplished in a number of different ways, either by determining the peak pixel intensity across the line, thresholding the Gaussian and computing an average and determining the center of gravity of the Gaussian. Alternatively, as in the case of the SAL-3D shape analysis library software from AQSense (www.aqsense.com), the point of maximum intensity of the Gaussian can be determined using nonlinear interpolation techniques, a method the company says is more accurate than COG methods.

Figure 5.10 Laser light camera system.

Speeding the computation of calculating the height of the laser profile can be accomplished using FPGAs. Indeed, a number of companies such as Allied Vision Technologies (www.alliedvisiontec.com) and Photonfocus (www.photonfocus.com) implement this in hardware in the FPGA of the camera. Alternatively, frame grabber companies such as Silicon Software (www.silicon-software.com) allow this task to be performed in the FPGA on a frame grabber. One such COG implementation has been developed by the company using its Smart Applets software and embedded in an FPGA on Silicon Software's microEnable IV V-series of frame grabbers. Once computed, height profile data can then be used to generate a 3D depth map of the object being imaged.

An example of these cameras is the Gocator from LMI Technologies (www.lmi3d.com) that integrate both structured laser light sources and image detectors. Gocator uses blue-LED structured light technology and wide field of view for full field inspection of large targets.

Specs of Gocator® 3504:

- Fast scan rate (6 Hz w/ acceleration);
- 6.7 µm XY resolution for micron-level inspection;
- 0.2 µm Z repeatability for reliable measurement;
- Measurement range = 7 mm;
- Field of view = 12.1 x 13.2 - 13.0 x 15.0 mm.

Figure 5.11 Laser light camera Gocator.

Gocator integrates both structured laser light sources and image detectors into a single unit.

Another example of cameras with these technology is the Ensenso 3D from IDS (https://en.ids-imaging.com). Each model of Ensenso 3D has two integrated CMOS sensors and a projector that casts a high-contrast textures onto the object to be captured by using a pattern mask. These cameras come with lockable GPIO connectors for trigger and flash and GigE connector, has two monochrome CMOS sensors and a projector. They come with HALCON, C, C++ and C# sample programs with source code. The projector is available either with blue lights in the visible area (465 nm) or with infrared lights (850 nm). Via Power over Ethernet a data transfer and power supply are possible with long cable lengths. The 3D camera meets the requirements for protection code IP30 and is therefore protected against dirt, dust, water splashes or cleaning agents.

N35: Global Shutter, 1280x1024 pixels. Suitable for 3D acquisition of still objects and for working distances up to 3,000 mm. Available with focal lengths from 6 to 16 mm.

N30: Global Shutter, 1280x1024 pixels. Suitable for 3D acquisition of moving objects and working distances up to 3,000 mm and are available with focal lengths from 6 to 16 mm.

N10: Global Shutter, 752x480 pixels. Suitable for 3D acquisition of moving objects and working distances up to 2,000 mm and are available with focal lengths from 3,6 to 16 mm.

Figure 5.12 Laser light camera Ensenso 3D.

X30: For working distances up to 5 meters and the acquisition of objects with volumes of several cubic meters. 3D camera system consists of a projector unit, two GigE uEye cameras either with 1.3 MP or 5 MP sensors (CMOS, monochrome), mounting brackets and adjustment angles, three lenses as well as sync. and patch cables to connect the camera with the projector unit. You can choose between variants with protection class IP65/67. The latter also include, in addition to special cables, lens tubes for the cameras and the projector. The 5 MP cameras use the GigE Vision standard to communicate with the pattern projector. This eliminates the need for additional installation of the IDS Software Suite.

Figure 5.13 Laser light camera Laser light camera Ensenso X30.

5.5.4 Fringe patterns

While such structured light-based systems are useful, they require that the camera system and laser light projector or more commonly the object to be imaged to be moved across the field of view of the system. In addition, where shiny specular surfaces need to be imaged, such systems may not prove effective. To overcome these limitations, a technique known as fringe pattern analysis can be implemented.

In this technique, a series of intensity patterns with shifting periods is projected across the object to be imaged. Phase-shifted images reflected from the object are then captured by a camera and the relative phase map or measurement of the local slope at every point in the object calculated. From this phase map, 3D coordinate information can be determined.

To illuminate the object to be imaged, a number of different methods can be used. In the design of a system to inspect large specular auto body panels, for example, Stefan Werling at the Fraunhofer Institute of Optronics (www.iosb.fraunhofer.de) use a 42-in. diagonal LCD diffuse display as a pattern generator to illuminate the object while 3D Dynamics' (www.3ddynamics.eu) Mephisto scanner uses a digital light projector (DLP) from InFocus (www.infocus.com). Other companies such as Numetrix Technologies (www.numetrix.ca) employ a combination of split-spectrum light projection and dual-CCD camera technology in its NX3D series of scanners.

5.5.4.1 Time of flight

While traditional methods such as stereo vision, structured light and fringe pattern projection all present the systems developer with different price/performance tradeoffs, a faster, albeit lower resolution method known as time of flight (TOF) imaging can be used in applications such as obstacle avoidance and object tracking.

The measurement principle is based on the time the light needs to travel from the light source to the object and back to the camera; the further the distance, the longer the time taken. Both light source and image acquisition are synchronized in such a way that the distances can be extracted and calculated from the image data.

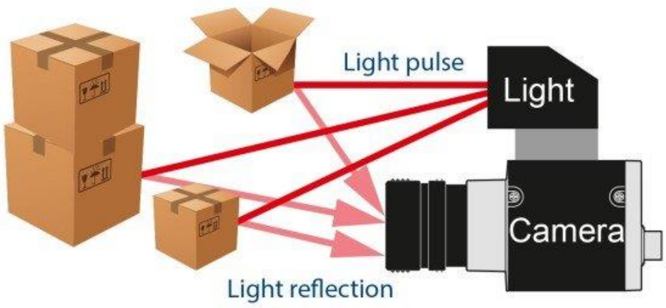


Figure 5.14 The Time-of-Flight Camera system.

Of the numerous systems now available that perform TOF imaging, both pulsed and continuous wave techniques can be employed these are some of the most important:

Odos imaging (www.odos-imaging.com) uses pulsed wave techniques in the design of its 3D Cam 3D imaging. In operation, short pulses of 905nm light are emitted from a laser diode to illuminate the scene. Pulses reflected by objects within the scene and are detected by a 1280x1024 CMOS imager with a global shutter. Proprietary algorithms then convert the detected pulses into a distance measurement. Simultaneously, a conventional intensity image of the scene is captured so that each pixel on the image sensor provides both distance and intensity information. Specs of this camera are:

- Resolution 640x480 pixels;
- Frame Rate up to 44 fps;
- Operating Range 0.5 - 6 meters
- Precision ~ 1 cm (typical, varies with return signal level);
- Field-of-View 43° x 33° (H x V);
- Illumination 7x LEDs @ 850 nm;
- Output Data Options xyz point cloud (via SDK), range, active infrared;
- Interface GigEVision and GenICam compatible;
- Power 12 VDC / 20 W (typical);
- Software StarStream-Swift Demo Software (Windows);
- SDK .NET (Windows) / C++ (documented with examples) 7/8/10 / Linux (i686, x86_64, armhf, AArch64).

Figure 5.15 real.iZ VS-1000 camera.

The Basler Time-of-Flight Camera (https://www.baslerweb.com) is an industrial 3D camera operating on the pulsed Time-of-Flight principle. It is equipped with eight high-power LEDs working in the NIR range (850nm), and generates 2D and 3D data in one shot with a multipart image, comprised of range, intensity and confidence maps. The Basler ToF Camera stands out in particular for its combination of high resolution (VGA) and powerful features at a very attractive price. Specs of this camera are:

- Resolution: 640x480 pixels (NIR);
- Frame Rate: 20 fps;
- Working Range: 0 m to 13 m;
- Accuracy: +/- 1cm (Scene dependent), document on accuracy;
- Interface: Gigabit Ethernet, GigE Vision and GenICam compliant;
- Lens: 57°h x 43°v;
- Software: Windows and Linux compatible;
- Easy to integrate and use;
- Total system cost reduction.

Figure 5.16 The Basler Time-of-Flight Camera.

IFM (www.ifm.com) offers products such as the efector pmd 3d, a stand-alone 3D system. With these specs:

- Operating distance [mm] 300...8000 (with reflectivity of 18 % and object size of 200 mm x 200 mm);
- Max. measuring range [m] 30 (depending on settings and reflectivity, typically up to 5000 mm);
- Resolution pixels [pixel] 352 x 264;
- Angle of aperture [°] 60 x 45 (nominal value without lens distortion correction);
- Image repetition rate max. [Hz] 25;
- Software API C, C++, Halcon.

Figure 5.17 Efector pmd 3d camera.

PMD Technologies (www.pmdtec.com) use continuous wave techniques. Here, phase differences between emitted and reflected signals from the sensor system are measured and the 3D depth information computed. PMD Technologies offers a depth sensor reference design, the CamBoard nano a 37x30x25 mm board that incorporates both a light source and image sensor with these specs:

- Resolution: 160x120 pixel image sensor;
- Frame Rate: of up to 90 fps;
- Field of View: 90°;
- Output: Depth map and grey-scale image data.

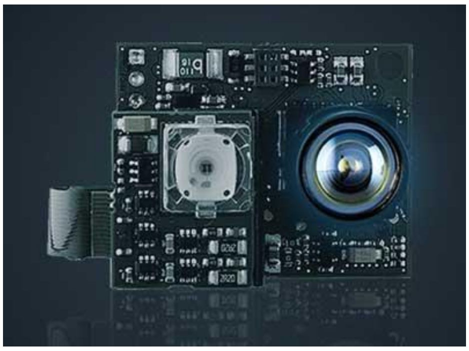


Figure 5.18 PMD Technologies camera.

5.6 Object recognition algorithms

Recognition is the higher level task in Computer Vision, and therefore the most challenging one.

The objective of recognition is to interpret the image, extracting knowledge about the objects present in the image, and therefore it is related to the Artificial Intelligence topic of learning.

The task of recognition can be classified as follows:

- Object Detection: When we are looking for a specific object. This technique involves scanning an image to determine where a match may occur. If the object we are trying to recognize is rigid, the problem can be solved searching for characteristic feature points and verifying that they align in a geometrically plausible way. In the non-rigid case, deformation models and complex representations of the objects may be required. Also in a simplistic scenario, object detection can be considered like a segmentation problem (separation from object and background) and in a general case, the problem can be approached like a binary class recognition problem.
- Class Recognition: This problem involves recognizing instances of extremely varied classes. To solve this problem some techniques, rely purely on the presence of features and their relative positions, while other techniques are based on finding the constituent parts of the objects and measuring their geometric relationships. In many instances, recognition depends heavily on the context of surrounding objects and scene elements.

The general case of recognition is still far from a solution and there is not even any consensus among researchers about when a significant level of performance might be achieved.

A general recognition problem can be usually separated mentioned in the following paragraphs.

5.6.1 Segmentation

Image segmentation is the process of partitioning a digital image into multiple segments. The goal of segmentation is to transform the representation of an image into a more meaningful or simpler one. Image segmentation is typically used to locate objects and boundaries in images, assigning a label to every pixel such that pixels with the same label share certain characteristics.

The result of image segmentation is a set of regions that cover the entire image. Each of the pixels in a region are similar with respect to some characteristic or property, such as colour, intensity, or texture.

Several general-purpose algorithms and techniques have been developed for image segmentation. These techniques must typically be combined with a domain's specific knowledge to solve the segmentation problem.

- Colour/intensity/texture thresholding;
- Statistical clustering methods;
- Edge detection;
- Motion based segmentation;
- Depth based segmentation;
- Compression-based methods;
- Histogram-based methods;
- Region-growing methods;
- Graph partitioning methods.

5.6.2 Representation

Involves describing the detected objects using meaningful and robust descriptors called features, which may be compared to other ones. Usual types of features include: scale-invariant features, graphs, lines, contours, shape or texture descriptors, Gabor filters, inverted indexes, frequency vectors and others. The problem of representation involves two different tasks:

- Feature extraction. It creates new features from functions of the original raw data;
- Feature selection. Select a subset of the features that represent the data without losing relevant information for the problem domain.

5.6.2.1 Feature extraction

Feature extraction is based on transforming the computational representation of the object of interest, from the raw representation to a representation meaningful to the domain. It addresses two fundamental issues:

- Obtaining an informative and non-redundant representation. Facilitating the subsequent learning and generalization steps, and leading to a better interpretation.
- Reducing the amount of resources required to describe a large set of data. When performing analysis of complex data one of the major problems derives from the number of variables involved. Since a large number of variables may require a large amount of memory and computation, and may cause the classification algorithm to overfit (poor generalization to new samples due to the algorithm being too specific for the training set).

Feature extraction is a general term for methods of constructing combinations of the variables while still describing the data with sufficient accuracy. Properly optimized feature extraction is one of the key factors in effective model construction.

The first of the problems related with feature extraction is based on the construction sets of application-dependent features. This process called feature engineering usually requires the participation of an expert. In addition, it constitutes in practice, one of the bottlenecks in Machine Learning. Recently, deep learning techniques perform extraction application-dependent features as part of the model construction process.

The first problem is usually achieved by general dimensionality reduction techniques such as:

- Independent component analysis;
- Isomap;
- Kernel PCA;
- Latent semantic analysis;
- Partial least squares;
- Principal component analysis;
- Multifactor dimensionality reduction;
- Nonlinear dimensionality reduction;
- Multilinear Principal Component Analysis;
- Multilinear subspace learning;
- Semidefinite embedding;
- Autoencoder.

5.6.2.2 Feature selection

Feature selection, also known as variable selection, attribute selection or variable subset selection, is the process of selecting a subset of relevant features for use in model construction.

Feature selection techniques are used for four reasons: simplification of models, shorter training times, reduce overfitting. The central premise when using a feature selection technique is that the data contains some features that are either redundant or irrelevant, and can thus be removed without incurring much loss of information.

A feature selection algorithm can be seen as the combination of a search technique for proposing new feature subsets, along with a metric that scores the different feature subsets.

Search techniques

Search techniques iteratively evaluate a candidate subset of features, and estimate if the new subset is an improvement over the old. Evaluation of the subsets requires a scoring metric. Search approaches include:

- Exhaustive search;
- Heuristic search;
- Random search methods;
- Best first;
- Genetic algorithms;
- Greedy forward selection;
- Greedy backward elimination;
- Particle swarm optimization;
- Scatter Search;
- Variable Neighbourhoods Search.

Metrics scores

To evaluate the set of features two main strategies can be used. On the one hand, the model itself can be evaluated on those features, however, this approach is usually very impractical, since it will be simply an exponential grow of complexity in computation time, to evaluate every model for every possible set of features. To solve this problem, it is usually employed a metric to score how well the data is described by a subset of selected features. In most cases this metric consists on a measure of accuracy penalised by the number of features selected (e.g. the Bayesian information criterion) or using any statistical function that measures the relationships between all the variables with the aim of minimizing redundancy and maximizing relevance.

Classification of feature selection algorithms

Feature selection algorithms can be broken up into three classes, Wrappers, Filters and Embedded techniques:

- Wrappers use a search algorithm to search through the space of possible features and evaluate each one by running a model on the subset. Wrappers can be computationally expensive and have a risk of overfitting to the model.
- Filter methods use a simpler, less expensive metric. Instead of evaluating a feature subset against a model, features are selected based on their scores in various statistical tests for their correlation with the output variable. Common measures include Linear discriminant analysis, ANOVA analysis, Chi-Square method, the mutual information, Pearson correlation coefficient, inter/intra class distances, test scores of significance for class/feature combinations… etc. Filter methods are generally used as a pre-processing step.
- Embedded techniques perform feature selection as part of the model construction process and are specific to a model. Typical embedded techniques are decision tree algorithms.

5.6.3 Recognition

Recognition is the problem of determining if a detected object belongs to a class or a set of classes of interest. Visual object recognition is not a single mechanism, and given the extremely rich and complex nature of this topic, several techniques are used, ranging from machine-learning techniques (such as boosting or neural networks) to support vector machines or Bayesian approaches.

Object recognition presents many challenges, being the most important ones:

- Changes in lightning or colour;
- Changes in the background (scene) that affect the object;
- Occlusions;
- Changes in pose or in the relative perspective of the viewer to the camera;
- Deformation in some objects;
- Changes in size or distance.

5.6.3.1 Object recognition and Matching

Matching can be seen as an underlying concept of object recognition, though it can also be applied in a broader sense avoiding the concepts of object or class.

Using this first perspective, matching in image processing consists of finding a correspondence between an objective and a sample based on comparing features (edges, colour, texture, gradients, frequencies… etc.) in the images.

An example of matching can be applied on an image region for estimating the correlation with a predefined template using a local search strategy. Though this is usually not considered like an object recognition technique, we must note that the same principles apply, and problems such as 3D reconstruction, tracking, pose estimation, optical flow, pattern recognition… and others overlap significantly and also the techniques and mathematical models to solve them.

5.6.3.2 Classification of recognition techniques

Object recognition techniques can be classified in several ways. There are three main classes of object recognition methods. These classes however, are not strictly independent or separated from each other and can be combined.

Hypothesize and test

The main idea is to hypothesize a correspondence between a collection of image features and a collection of the model features. Then use this to generate a hypothesis about the projection from the object coordinate frame to the image frame. Then, use this projection hypothesis to generate a rendering of the object. This step is usually known as backprojection. Compare the rendering to the image, and, if the two are sufficiently similar, accept the hypothesis. A specific example of these techniques is:

- Alignment methods. This set of methods assume that the objects in the scene are the result of a transformation from a stored object model. For each model, a set of allowable transformations can be defined and the main goal is to find the correct transformation to maximise a measure of fit between object and model. Usually with these techniques, transformations are explicitly applied to the stored model as a part of the algorithm. Alignment methods are based on the principle of:

Invariant properties methods

This set of methods is based on using a set of features that are invariant to camera transformations, light variations or other type of variation expected in the image, and use these features to find a correspondence of the model in the image. Examples of these techniques are:

- Scale-invariant feature transform (SIFT). In this technique, an object is recognized in a new image by individually comparing each feature from the new image to this database and finding candidate matching features based on Euclidean distance of their feature vectors.
- Speeded Up Robust Features (SURF). It can be seen as an optimization of the idea behind SIFT based on the use of sums of approximated 2D Haar wavelet responses and made efficient use of integral images.
- Histogram of Oriented Gradients (HOG). The technique counts occurrences of gradient orientation in localized portions of an image. It is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization. The HOG descriptor is invariant to geometric and photometric transformations, except for object orientation.

Part decomposition methods.

This set of techniques is based on the fact that many objects seem to contain natural parts, that can be detected and classified independently. These techniques are useful in cases where the parts are simpler to detect than the whole object, or where the variability of object views is due to variability of structure, and structure can be detected by connectivity between parts. Part decomposition methods work usually with the following workflow:

1. Description of objects in terms of constituent parts;
2. Location of constituent parts in the image;
3. Classification of constituent parts;
4. Analyse the relationships between parts;
5. Select objects for which structure matches detected relationships best.

5.6.3.3 Machine learning methods

Machine learning methods are a special kind of object recognition methods that do not rely on an algorithmic approach to solve the problem. They try instead to find data regularities based on inductive observation of data. Machine learning techniques work with a set of labelled data (also called training data) to create an adaptive model that is able to extrapolate the data relationships from the presented examples to new cases (This is called supervised learning).

Optimisation

Supervised machine learning techniques are therefore based on data samples with known object correspondences. These techniques use this set of known samples to find a function from the input domain (features) to the output domain (classes) that matches the examples, inferring the mapping implied by the data. It works using a cost function that relates the mismatch between the mapping and the data and uses an optimization strategy to minimize this cost function. Optimization strategies include:

- Stochastic gradient descent (SGD);
- Evolutionary methods;
- Particle swarm optimization (PSO).

Classification of Machine Learning techniques

Machine learning techniques in object recognition scenarios:

- Linear classifiers. This is a family of techniques that look for linear combinations of variables to model the difference between the classes of data. An example of these techniques is linear discriminant analysis (LDA).
- Naive Bayes. These classifiers assume each feature to contribute independently to the probability tha one object belongs to a particular class, regardless of any possible correlations between features (Breiman, 1996b).
- K-Means. This Is a non-parametric method that uses the k closest training examples in the feature space to determine the class membership of the input object (Domingos & Pazzani, 1997).
- Support Vector Machine (SVM). This model builds a hyperplane or set of hyperplanes in a high-dimensional space defined by the feature space, for classification. Classification is achieved by the hyperplane that has the largest distance to the nearest training-data point of any class (so-called functional margin) (Huang, Kecman, & Kopriva, 2006).
- Decision tree. This is a predictive model used to go from observations about an item (represented in the branches) to conclusions about the item's target value (represented in the leaves). Algorithms for constructing decision trees usually work top-down, by choosing a variable at each step that best splits the set of items. Different algorithms use different metrics for measuring this criterion. These generally measure the homogeneity of the target variable within the subsets (Quinlan, 1986).
- Artificial neural networks (ANN). ANN are computing systems inspired by the biological neural networks that constitute animal brains. An ANN is based on a collection of connected units called artificial neurons. Each connection can transmit a signal from one neuron to another. A neuron that receives a signal can process it and then sends the result as the input to additional neurons connected to it.
    - Deep Learning ANN (DNN). It is a modern set of techniques based on classic ANN architectures to which a cascade of multiple layers of nonlinear processing units are added

for feature extraction and transformation. Each of these layers learns to transform its input data into a slightly more abstract and composite representation. In image recognition, the raw input may be a matrix of pixels; the first representational layer may abstract the pixels and encode edges; the second layer may compose and encode arrangements of edges; and so on. The main advantage of these techniques is that they obviate the problem of feature selection and feature extraction engineering, by translating the data into compact intermediate representations. Its drawback is their inherent computational complexity and their need for great amounts of training data. (Schmidhuber, 2015).

The application of an ANN to an image through convolutional kernels has brought about convolutional neural networks (CNN). Recently CNN have experienced a tremendous popularisation amongst computer vision researchers thanks to the availability of big image datasets and the employment of GPUs on this methodology. In fact, CNN have become the state of the art (SOA) in numerous fields of image processing, surpassing by far previous breakthroughs. One of the main applications of CNNs to image segmentation is semantic segmentation, where a CNN categorises every pixel in the image into meaningful concepts, so that pixels are grouped in sets that cover different categories in the image, but without distinguishing between different instances of the same category. This limitation is overcome in object detection, wherein a CNN is trained to find and localise different instances of the same category in the same image.

Object detection

A CNN can address object detection with various approaches. In Faster R-CNN (S. Ren et al, 2016) the image is provided as an input to a CNN whose output is used by a separate network to propose regions where may potentially be objects. These regions proposals are then classified into object categories. YOLO (J. Redmon and A. Farhadi, 2018) divides the image into a grid, which defines a set of bounding boxes. For each bounding box, a CNN outputs a class probability and offset values for the bounding box. SSD (C. Szegedy et al, 2016) works in a similar manner as YOLO, but it makes predictions based on feature maps taken at different stages of the convolutional network. YOLO and SSD achieve better performance (~50 FPS) than Faster R-CNN (~7 FPS) but less accuracy in object detection.

Object detection can be combined with other applications from CNNs, resulting in Dense captioning (object detection along with captioning), or Instance segmentation, which is a hybrid between semantic segmentation and object detection: it detects instances of a category and tags the pixels that belong to the object (Mask R-CNN, He et al. 2017).

Meta Algorithms

Meta algorithms or meta models, can be seen like a set of techniques based on the construction, assembly, aggregation or combination of underlying model sets. These meta algorithms include:

- Bootstrap aggregating. Also called bagging, it is designed to improve the stability and accuracy of machine learning algorithms, reducing variance and helping to avoid overfitting. This technique works generating new training sets by sampling with replacement from the original set. This sampling is known as bootstrapping. The models are fitted using bootstrap samples and combined by voting (Breiman, 1996a).
- Random forest. These techniques work by constructing a multitude of decision trees and outputting the class that is the mode of the classes of the individual trees. Random forest is in reality a modification of a Bootstrap aggregating strategy for trees, with the addition that it selects, at each candidate tree split, a random subset of the features. (Ho, 1995).

- Boosting. This is a meta-algorithm that iteratively learns weak classifiers and adds them to a final strong classifier. When they are added, they are typically weighted related to each learners' accuracy (Breiman, 1996b).
- Adaptive Boosting (AdaBoost). It is a meta-algorithm similar to boosting that combines different models into a weighted sum that represents the final output of the boosted classifier. This technique selects only those features known to improve the predictive power of the model, reducing dimensionality and potentially improving execution time. (Freund & Schapire, 1997).

5.7 Database of referenced images

Despite defect inspection in railway infrastructures is an active field in computer vision research (Qing He et al, 2012; Aliza Raza Rizvi et al, 2017; Yongzhi Min et al, 2018), the lack of public datasets of the railway realm prevents the research from approaching this problem via CNNs. Deep learning techniques have showed promising results on railway defect inspection but at the expense of private datasets (S. Faghih-Roohi et al, 2016; Siamak Hajizadeh et al, 2016). However, it is possible to take advantage of specific properties of DNNs and apply them to this problem. These properties exploit the modularity of DNNs and can be listed as follows:

- Fine-tuning: It consists of further training a pre-trained DNN (usually with data features that differ from the original domain).
- Transfer learning: It consists of concatenating one or several additional layers to a DNN that has been trained for a given problem so that the resulting DNN can be trained with relatively little data for a new but similar problem.
- Domain adaptation: Given a DNN that has been trained for a domain A, domain adaptation consists of further training this DNN with a small amount of data from a different domain B in order to apply the adapted DNN to this new domain B.

By applying these properties to publicly available image datasets, we take advantage of the thorough and extensive testing, balancing and verification checks that have been applied over these datasets.

The amount of image datasets existent in the literature is vast, and they cover a great range of topics and categories. There are datasets made out of pictures of the real world (natural), datasets composed of synthetic images (artificial), datasets of letters (MNIST), clothes (fashion MNIST), human poses (DensePose, Surreal), faces, pets, etc. The number and depth of categories per dataset varies, as well as the purpose. We consider that the two datasets that generalise best for CNN modelling are ImageNet and COCO.

ImageNet is designed for use in visual object recognition methods. It includes more than 14 million images in which the appearing objects have been annotated by hand. In at least one million of the images, bounding boxes for these objects are also provided. ImageNet contains more than 20.000 categories. A CNN trained with ImageNet is an excellent starting point to apply transfer learning.

COCO is a large-scale object detection, segmentation, and captioning dataset developed by Microsoft. It is oriented to object segmentation, recognition in context and superpixel segmentation. It contains 330K images, with more than 200K labelled. It has 1.5 million object instances and 80 object categories. It also has five captions per average image. Additionally, there are 250K people annotated with keypoints that have proved useful to the building of human pose datasets (DensePose). Its level of granularity in masking object categories has made of it one of the most used image datasets in Deep Learning training.

Many others datasets are available (CIFAR, CALTECH, PASCAL VOC, labelme, Berkeley segmentation dataset, OpenImages dataset, etc.) from which railway-related categories can be extracted. Images with labelled defects can be obtained by data augmentation of railway images, like in S. Ritika et al, 2018. With these strategies, a comprehensive dataset for railway defect inspection can be generated. Subsequently, this dataset can be applied to pre-trained CNNs on ImageNet or COCO by means of fine-tuning, transfer learning and domain adaptation methods. This will result in a CNN able to perform defect inspection on railways.

5.8 Big data solutions (collection, processing, storage, etc.)

In the last years, the data available have increased exponentially up to 10x every five years, due to the last Information and Technology (IT) advances that allows creating, not only to users but also to devices following an Internet of Things (IoT) approach, a large amount of Gigabytes of data per day.

On the other hand, the storage and computing capacity of the general purpose computers is increased at a rate of 58% every year, which generates an issue known as Big Data.

Big Data issue have different definitions, Chen et al define Big Data as "the datasets that could not be perceived, acquired, managed, and processed by traditional IT and software/hardware tools within a tolerable time" which means that Big Data is an ambiguous term that depends on the quantity of data and the current software/hardware capacity.

The challenges of Big Data management are not only related with the amount of data but also with the structure. The datasets have data from different sources that can be structured (databases) but also unstructured (video, image), which requires more computational power to extracting knowledge from them. The challenges identified by Khan et al. for big data are:

- Heterogeneity: Big Data applications usually have to deal with multiple data coming from different sources with different formats between them;
- Scalability: another challenge in these systems comes from the fact that the volume of information that must be dealt with increases at a much faster rate than the processing power of computers;
- Complexity: traditionally data mining algorithms had to deal with structured data with similar formats and predefined lengths. In Big Data, the information of unstructured data, such as text or multimedia content, must also be processed;
- Accuracy: the veracity of the results extracted in Big Data contexts is negatively affected by the heterogeneity and complexity of the source data;
- Storing/Sharing/Publishing: Even with the huge amount of data these applications must deal with, Big Data systems must store it in an accessible and reliable way. This is translated in distributed systems that have to ensure the consistency and availability of the data even if parts of the system experiment failures;
- Security: Data leakage is a major risk in Big Data systems. Furthermore, these systems need to deal not only with this possibility but also take into account data integrity to, for example, avoid the possibility of fraud;
- Retrieve/Reuse/Discover: After being stored, the possibility of retrieving and reusing the data allows to discover new valuable information. To ensure reusability the semantics of the stored data must be determined and with the inherent heterogeneity of Big Data systems this is not always a trivial process.

The Big Data management tools are based on distributing the computational power and storage to manage datasets and also to extract knowledge from them. Several technologies are available for data management such as NoSQL, Google Big Table or Voldemort, but in many cases the companies still have to develop their own tools to compute and extract the knowledge from this data.

In 2006, an Apache project named Hadoop propose a new solution to both manage and compute Big Data in a distributed way, it is currently being used by approximately the 63% organizations that manage huge number of unstructured logs and events (Sys.con Media, 2011). Hadoop is developed in Java and is composed by several components that manage certain tasks but the two main components are HDFS and MapReduce.

The Hadoop Distributed File System (HDFS) is a distributed file system based on a master/slave architecture that distribute and replicate all the data stored in the filesystem mapped using DataNodes, which are blocks that contain files and manage the links between them.

Map Reduce is a parallel programming model for large data computing and is integrated with Hadoop. This paradigm is based on mapping clusters of data using a tree structure for later computing and reducing (with reduce operation) each leaf of the tree.

Hadoop have still some limitations such as SQL support or inefficient execution for querying, which have been improved in the last updates but is still an issue.

5.9 RFID technologies

5.9.1 General Introduction of RFID

Radio Frequency Identification (RFID) is a wireless technology that is widely used for the unidirectional transmission of the pre-written information such as identification numbers. The basic components are RFID tag, RFID reader, antenna and so-called middleware. The RFID tag itself consists of a silicon microchip attached to a radio antenna. Figure 1 shows that bi-directional communication takes place between RFID tag and RFID reader, between RFID reader and RFID software, and between RFID software and application. Energy (electromagnetic waves) and timing proceed from the reader to the tag, allowing information transmission. The timing module of RFID tag enables the user, upon interrogating the RFID tag, to determine the precise length of time from the previous charge of the RFID tag.

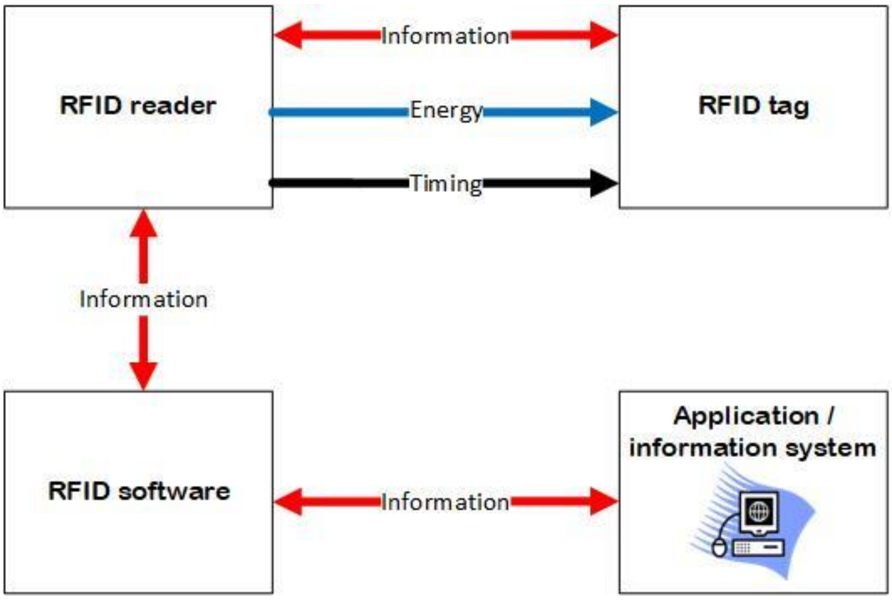


Figure 5.19 Principle of RFID system.

5.9.1.1 RFID Frequency Bands

RFID technology can work in different frequency bands, specifically in:

- Low Frequency Band (LF);
- High Frequency (HF);
- Ultra High Frequency (UHF);
- Microwaves.

The low frequency band ranges are from 125 to 134 kHz with a read range of up to 0.5 m. RFID technology utilizing this band is used, for example, to control access to administrative buildings or production areas, to identify animals or to identify metal products (beer casks, etc.). The benefits include greater interference immunity and the possibility of mounting near water or a metal underlay. On the contrary short read range, small communication speed and financial demand of tag design are some of the disadvantages.

The HF band communicate on frequency 13.56 MHz and the read range is about 1 m. The main disadvantage of the HF band is the fact that metal underlays and water significantly reduce read range and interrupt communication. On the other hand, the LF band allows higher communication speed, greater read range, reading of 10 to 40 tags per second without collisions and the frequency is standardized worldwide. HF tags are used for labelling of luggage, non-contact payment or smart electronic labelling.

The UHF band ranges from 860 to 960 MHz. In Europe frequency band is 868 MHz. In the USA and Canada the frequency band of 915 MHz is used. Read range is up to 10 m. The UHF band has the advantages such as possibility of remote reading, high transmission speed and cheap production. The disadvantages are the non-unified global frequency, difficult reading through metals and liquids. Due to the non-unified global frequency, multifrequency variants using the frequency range of 865-928 MHz are available, but they are more expensive than one-frequency UHF tags. UHF tags have a wide range of uses, such as parking cards or electronic tolls, group package tracking and warehousing.

The last frequency band is a 2.4 GHz microwave band with a read range of up to 2 m. The microwave frequency band is also used within electronic toll systems, for the identification of luggage in air transport or wireless recording. Despite the benefits of high communication speed (up to 2 Mbps) and small tag sizes, microwave tags are expensive and complicated, having a smaller range than the UHF bands. There is also a great influence of interference from liquids and metals.

5.9.1.2 Standardization of RFID

RFID technology uses the standard for product labelling. This is the Electronic Product Code (EPC), i.e. products' electronic code. The primary code holder is the tag. The organization that deals with RFID standardization is, for example, GS1. The EPC code is a unique data structure (unique number) stored in an RFID tag that has 96 bits (24 hexadecimal characters). The commonly used standard protocol is defined in EPC Radio-Frequency Identity Protocols Generation-2 UHF RFID, referred to as EPC Gen2 Class 1 UHF standard which is also approved as the standard ISO 18000-6C.

The structure of EPC code is:

- 8 bits header, EPC number of version;
- 28 bits information about company (max. 268 million companies);
- 24 bits product class, 16 million classes in each company;
- 36 bits unique (serial) product number, 68 billion numbers in each class.

The EPC code is assigned to each particular piece of goods and on the central level to producers. So far, there are no theoretical prerequisites for using such numbers. If it is necessary, it is possible to change the EPC code to the length 128 bit or longer up to 204 bit. The code structure is shown in Figure 5.20.

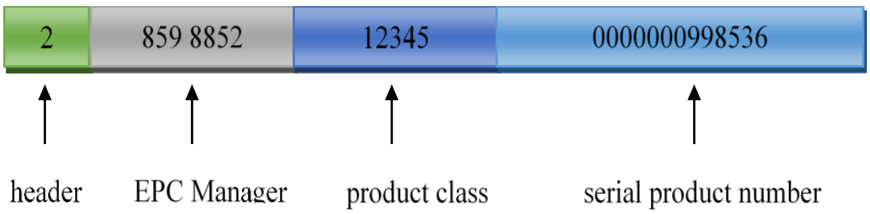


Figure 5.20 Structure of EPC code.

The EPC code header defines the EPC structure, the EPC version and the code size. The EPC Manager number indicates the manufacturer or company identification number. The product class defines the type of product of the manufacturer. The serial product number refers to the particular product.

GS1 standards are the most common open standards for RFID applications in Rail in Europe. It was developed by GS1 in Europe with the help of key rail actors in Europe including railway infrastructure managers, train operators, components manufacturers, rail industry bodies and solution providers involved in RFID in Rail implementations. So far there are four GS1 standards for vehicle identification (GS1, 2012), identification of MRO parts/components (GS1, 2018), data exchange for rail vehicles (GS1, 2015) as well as exchange of MRO component/part lifecycle data (GS1, 2018) respectively, detailing how to identify rail assets using GS1 Identification Keys and EPC Gen 2 UHF (ISO 18000-63) tags as well as exchange data using EPCIS within the framework of the GS1 System.

For the identification of rail assets in Europe, the Global Individual Asset Identifier (GIAI) is recommended.

This project has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement no. 826250 Assets4Rail).

Furthermore, it is recommended to use GIAI-96 encoding and not the GIAI-202 encoding. The GIAI includes the 12 digit European Vehicle Number (EVN), while the Vehicle Keeper Marking (alphabetic) shall not be encoded in the data carrier. The data structure is presented in Figure 5.21 (GS1, 2015).

Company Prefix	Individual Asset Reference	
	Vehicle end/side Indicator	EVN
735999271	**1, 2 or 3****	**917400000019**
Entity responsible for maintaining the individual asset reference (vehicle owner, train operator, transport agency etc)	To determine side indicator, please see below.	12 digit European Vehicle Number (EVN)

Figure 5.21 Structure of the Global Individual Asset identifier encoding the EVN (GS1, 2015).

In terms of tag location, two tags are required. Identifying which of the tag was read can show actual orientation of the vehicle, e.g. enabling actual parameters from Wayside Train Monitoring System (WTMS) to be matched with correct wheel / axle. A filter value in tag data was allocated, providing a means to read desired tags in an environment where there may be other tags present. Figure 5.22 shows the recommended tag location, being consistent with TSI CR WAG:2006 (GS1, 2015).

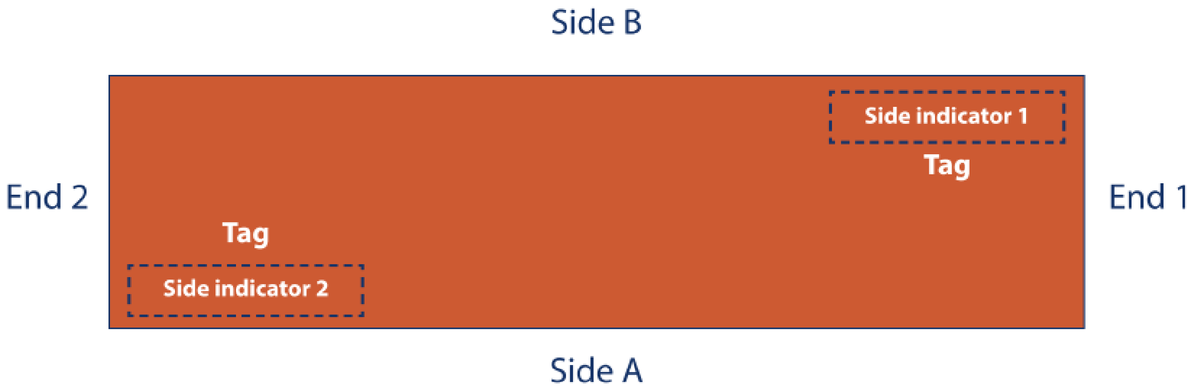


Figure 5.22 Vehicle end/side Indicator (GS1, 2015).

For identification of MRO part/component, it is recommended to use a GIAI or a serialized Global Trade Item Number (GTIN) in the case the manufacturer/supplier allocates an identifier to the part. When both the vehicle and the MROs are identified with a GIAI, it is currently not possible to select / separate tags before interrogation meaning that a number of GIAIs may appear to the trackside reader (GS1, 2018).

The RU standard EN 17230:2018 (draft) aims at describing the implementation of the European Vehicle Number (EVN) of the railway rolling stock in an electronic format via the ISO/IEC 18000-63 UHF Radio Frequency Identification (RFID) technology in order to enable a consistent approach for an interoperable implementation (EN17230, 2018). This standard is to specify the RFID tag location, tag data content and functional requirements in relation to the RFID tag track side reading performance for the purpose of vehicle identification. The tag location is consistent with GS1 standard. Besides GS1 data structure, EN17230 also recommends data structure using American National Standard (ANS) data identifier. Regarding tag performance requirements, Tag shall be fully operating with an ISO/IEC 18000-63 standard compliant reader with an output power from the antenna (ERP, effective radiated power) transmitting power of 33dBm and a radiation pattern of 45 degrees with following constrains (environmental conditions in line with the EN 50125-3:2003 value T2 for temperature, rain and humidity) (EN17230, 2018):

- Minimum read ranges 5m (tag reader typical installation distance is 3.1-5m from rail);
- Maximum read range 15m (should define also maximum range to limit cross/over reading);
- Minimum tag backscatter signal strength (power) of −70dBm (at distance of 5m);
- RX Data transmission with 62,5 Kbit/s is to be supported;
- The assumption is that the tag and antenna are on the same height and their normals are coaxial.

EN 17230 also mentions that times for reading 96 bits tags have been measured with 11,5 ms to 12 ms, and for 240 bit tags with 13 ms to 13,5 ms. However, when addressing other memory banks in addition to the memory bank for the unique identification the tag read performance degrades very significantly. Therefore, it is not recommended to read other memory banks than EPC when the vehicles are operating at higher speeds. Additional information that should be readable when the vehicles are moving at higher speeds should not be encoded into the user memory (EN17230, 2018).

5.9.1.3 Types of tags

There are two basic types of RFID tags, passive and active.

- Passive tags (RTFs) do not have a transmitter, so they reflect energy (radio waves) back to the reader/antenna that broadcasts radio waves. Read distance is 0.5 - 10m for UHF tags. Memory size is between 64 - 256 bits. In Europe and Africa, they use the frequency bands 865 - 869MHz (Region 1). For the Czech Republic is reserved the band 865.5-867.6 MHz. Passive tags have lower acquisition costs;
- Active tags (TTFs) have a transmitter and power supply. This source can be a battery that lasts for about 1-5 years and can have operating restrictions in relation to operating temperatures. Furthermore, a photovoltaic panel or other external source may be the power source. This type of RFID tag is used, for example, to identify containers, railway wagons and large tanks, which in practice need to be identified from a large distance. The read distance ranges from 20 m to 100 m. Active tags have higher acquisition costs, however, are suitable for reuse. Memory size can be up to 2Mb.

There are also following two hybrid types of tags:

- Semi-Active tag – is an active tag that remains in so-called sleep mode if it is not prompted to communicate with the RFID reader signal. It has a longer lifetime than the active tag. However, it is not suitable for fast loading or multiple tag reading at the same time;
- Semi-passive tag – is a passive tag that uses a battery as a source but does not generate a back-up signal. This type of tag theoretically offers higher performance compared to passive RFID tag (reading distance).

The basic types of tag memories are mentioned below:

- Read only (RO) – the tag has saved only a serial number that was encoded in its production. It is similar to a barcode. It is for reading-only and can no longer be changed;
- Read Write (RW) – it can store a large amount of data, e.g. an active tag from 16Kb to 2Mb. Data written to a tag can be deleted and rewritten up to a thousand times;

This project has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement no. 826250 Assets4Rail).

- Write Once Read Many (WORM) – is read-only tag. However, the tag is not programmed in production, but at a dealer.

In addition to tag types stated above, some tags may contain both the RO and the RW part at the same time. For example, a tag attached to a palette may be labelled with a sequential pallet number in the RO part and thus remain the same throughout the pallet life. The RW part can be used to indicate the pallet content at the current time. At the moment when goods were unloaded from the pallet and new goods loaded, this would be written to the RW memory.

According to EPCglobal, i.e. a joint venture between GS1 abd GS1 US, tags can be categorised into following classes, depending on their functionality.

- EPC Class 0 Tags – Generation 1, factory pre-programmed read-only passive tag;
- EPC Class 1 Tags – Generation 1 and 2, Read-only passive tag similar to class 0 and has one-time field programmability;
- EPC Class 2 Tags / EPC Class 1 Gen 2 Tags – EPC Class 2 tags are enhanced Gen 2 Class 1 tags, which are passive tags with read-write capability, working in UHF area. It is the standard tag type used in railway applications according to GS1 standards. All Gen 2 tags contain the same basic memory features, consisting of 96 bit EPC number support, 32-64 bit tag identifier (TID), 32 bit kill password to permanently disable the tag and 32 bit access password to lock the read write characteristics of the tag. Some tags include user memory which facilitates reading and writing of additional data in the tag;
- EPC Class 3 Tags – Semi-passive or active tag with read-write memory, on-board sensor and an incorporated battery to provide increased coverage; capable of recording parameters like temperature, pressure, and motion;
- EPC Class 4 Tags – Read-write active tag with integrated transmitter for communication using the battery on-board; can communicate with other tags and readers;
- EPC Class 5 Tags – Class 4 tags that provide additional circuitry to communicate with or provide power to other tags.

In terms of packaging, tags could be in form of labels or encapsulated inside a hard case. In the case of a label (paper or plastic self-adhesive label), the chip with the antenna is attached to the underlay forming the label. If the label is printed with a barcode (one of the label layers is paper), it is a so-called smart label.

The encapsulated variant may be in the form of a disc, plaque, keychain, bracelet, card, or glass stick. The encapsulation allows placing a tag on the metal. It can also be specially encapsulated according to specific customer requirements, so it can withstand big shocks (e.g. laundry washing in the washing machine) or it can withstand temperatures from -40 ° C to + 300 ° C. For a better overview, see Figure 5.23 below.

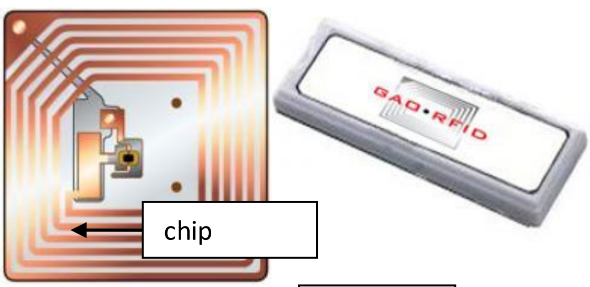


Figure 5.23 Label (left) an... d (right) tag.

5.9.1.4 Readers

A RFID reader is a device that emits radio waves through one or more antennas (in case of a stationary reader) and receives a scattered signal from tags, in which data is encoded. Then, the reader decodes the data and send it to a host device, e.g. PC or microcontroller. The range of readers is from centimetres to tens or hundreds of meters. The read range is related to the transmitting power and the frequency band used. Readers can be stationary (RFID gateway) or mobile (data terminal or handheld). Most stationary readers support multiple protocols for transponder reader communication and can operate in standalone and in networking mode. They could have multiple antenna connections for connecting more than one antenna to the reader, allowing the user to achieve greater and diverse radiation patterns of the reader's interrogation area. They use power supply from 12V DC up to 24V, weigh from 1.5 kg up to 5kg and can achieve reading ranges up to 300m. Mobile readers have built in antennas and usually do not have connectors for additional antennas. They are battery powered and are light weighted (from 82g up to 700g) and can achieve shorter reading ranges than fixed readers (up to 100 meters).

In case of selecting a reader, it depends on number of tags to be read per second and the distance between tags and antennas. In addition, the interface that connects reader to the network and the power supply is also necessary take into consideration. Figure 5.24 shows commercial examples of UHF fixed readers, while Figure 5.25 compares stationary and mobile readers.

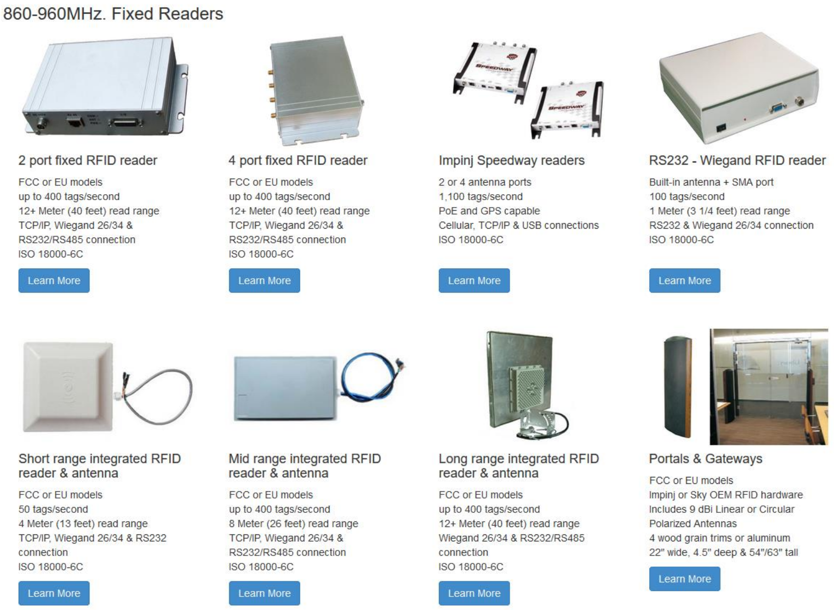


Figure 5.24 Typical UHF fixed readers

Figure 5.25 Stationary and mobile readers.

5.9.2 State of the art for the proposed application scenarios

Under certain conditions, RFID technology is an alternative to Barcode Technology or Optical Characters Recognition (OCR). The main difference from barcode technology is that the radio frequency tag does not have to be visible for retrieval of data and allows multiple tags to be scanned at once. RFID technology is usable in many fields (logistics, retail, automotive, postal sector, transport, health). In general, the applications of RFID technology can be categorised into the following groups in terms of its implementation.

- EAS (Electronic Article Surveillance) systems – currently one of the most widely used RFID applications in the world. It is a security system of products (usually smaller and more expensive) located in stores. Customers can be touched by the system very often. The readers are stationary and located at the entry of the stores. When a customer purchases a product tagged with an RFID tag, the tag is discarded or deactivated upon transfer through the cashier. However, if a customer attempts to remove a product with an active tag, a signalling device is activated to alert the operator;
- PDC (Portable Data Capture) systems – a system used to monitor product or person movement. RFID readers are being implemented in observed space to collect information from each tag. This information is then transferred to the central system where it is processed;
- Networked systems – the purpose of use is similar to PDC, but stationary readers are used to scan information from RFID tags when the product or people cross within the reader range. This information is sent back to the central system. Networked systems are used, for example, in automated warehouses at production lines;
- Positioning systems – identify the exact position of the item. To identify vehicles when crossing checkpoints, such as state borders or nodes where tolls are collected.

In the railway sector, Table 5.3 lists the potential applications of RFID technology.

Table 5.3 Potential RFID applications in railway system.

General process	Process in freight transport	Process in passenger transport
Vehicle identification for diagnostic systems	Weighing of wagon, i.e. customer stationary wagon weighing and dynamic train weighing	Support for automated cleaning machine for vehicles

Tracking objects, e.g. vehicles, trains and loading units, on networks	Support for automated marshalling yard command and control systems, such as generating train composition listing for arrival trains, parking positions on track, etc.	Control of level crossing safety system
Generating operational lists of vehicles, such as train composition list, border transfer list, etc.	Tracking of goods	Safety related process, such as train integrity and exact location of particular train in station. (according to Czech traffic rules D3)
Vehicle performance tracking, such as mileage tracking, time of sidings occupation and time of foreign wagon accommodation	Consignment clearance	Validation of tickets – determination of tariff zone
Support for maintenance by tracking MRO components		Control of station lighting
		Detection of arrival and departure of vehicles for Traffic and Passenger Information Display System

Most of applications listed above use the basic function of RFID technology, namely wireless transmission of simple information. In the Assets4Rail project, four specific application scenarios for railways are proposed with different expected TRL as below.

1. Reading the train identity information and interact with the developed image-based monitoring system. This scenario relates to the basic RFID reading function and will be eventually validated along with the developed monitoring systems at TRL 6;
2. Identification at the MRO component level by attaching RFID tags on wheelsets. In this case, the resistance against the violent removal breaking away of the tag from the wheelsets will be tested, and this use case will be tested at TRL 4;
3. Reading more pre-written information on tags, such as failure records, required maintenance procedures, etc., over the pass-by time. This use case will be tested at TRL 3;
4. Writing the information of the detected failures by the monitoring system to on-board tags over the pass-by time. This use case will be tested at TRL 3.

The following subsections provide a survey of current RFID applications in the proposed scenarios.

5.9.2.1 Scenario 1: Vehicle identification for wayside monitoring systems

RFID for vehicle identification, interacting with wayside train monitoring systems (WTMS), has been applied on the main line for daily operations, especially in north Europe such as Finland and Sweden. Mäkitupa (2015) presented the deployment of RFID readers and WTMS in Finland and in Sweden in GS1 Event 2015. There were 122 WTMS in total in Finland, while 120 sites were jointed with RFID readers, see Figure 5.26.

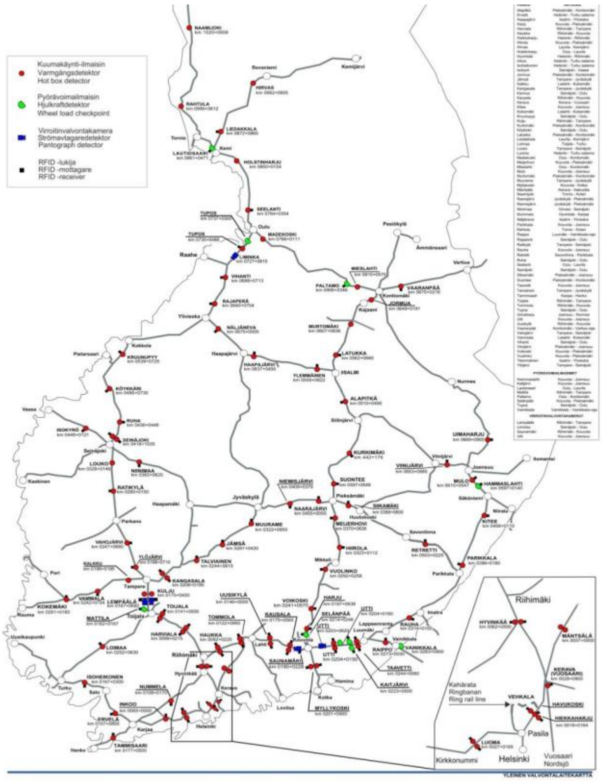


Figure 5.26 Sites of wayside monitoring systems and RFID units in Finland (Mäkitupa, 2015).

In Sweden, 183 RFID readers were installed, while 176 readers were matched with WTMS, see Figure 5.27. More than 3500 railway vehicles are equipped with RFID tags (Mäkitupa, 2015). Trafikverket planned to deploy up to 600 RFID readers since 2012. The common GS1 standard for Rail was applied so that all tags fulfilled UHF Gen2 Class1 (Ivansson, 2012). It was tested that RFID readers can be placed 3 m away from the track and read the vehicle information, i.e. company prefix and vehicle number, at speeds of up to 160 km/h.

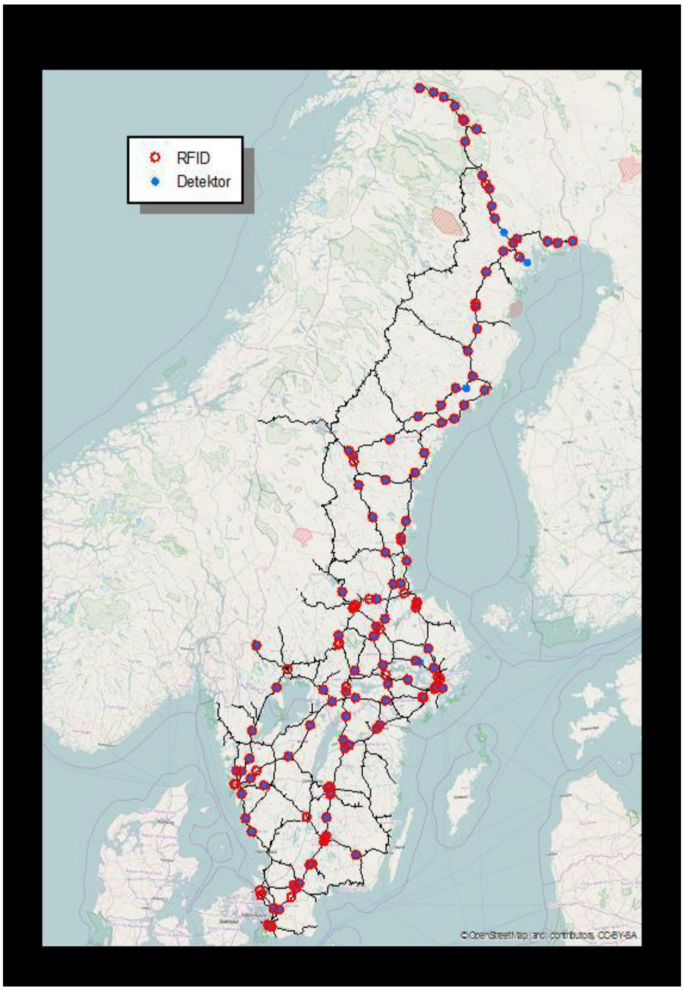


Figure 5.27 Deployment of RFID readers and WTMS in Sweden (Mäkitupa, 2015).

Since 2017, DB Netz established 7 RFID stations, joined with WTMS, providing diagnostics information for railway undertakings in Germany. These RFID stations also met the GS1 standard (Krüger, 2018).

In Switzerland, SBB Cargo has equipped all of their mainline locomotives and 1375 freight wagons with RFID tags (fulfilling GS1 standard) by June 2017 (Thalmann, 2017). SBB planned to install RFID readers on all sites of wheel load checkpoints by 2020.

In the Czech Republic, RFID technology was tested, for example, in the laboratory conditions (research centre) of VŠLG o.p.s. There were tested the application of GS1 standards in the railway transport model. The Impinj Speedway Revolution R420 is used to read the RFID tag in railway transport. RFID tags were placed on rolling stock according to the standard, see Figure 5.28.

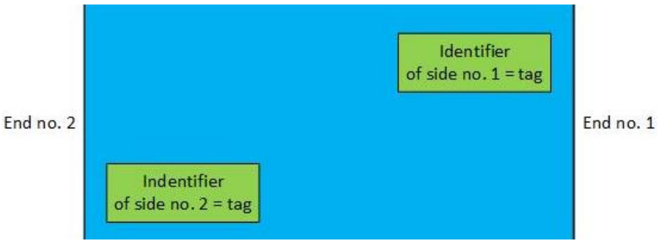


Figure 5.28 Tag positions on the vehicle according to the GS1 standard.

RFID technology was tested under the LOGIGATE project in a real environment on České dráhy, a. s. (Czech Railways) passenger electric units class 471, see Figure 11. (OLTIS, 2011)

Figure 5.29 Tagged unit class 471 with reader.

Various inventors seek to locate RFID tags and RFID readers both on the track and in stations and terminals and in workshops. The SelectraVision WheelCheck measuring system checks the profile of the train wheels in order to prevent train derailment. The system is able to measure the profile of each passing wheel at very high speed and with an accuracy of 0.5 mm. It is currently working from 0 to 150 km/h. Based on lasers and on high speed cameras, the wheel measuring system checks wear condition, diameter and parameters of the wheel tread profile comparing them to the standard measurements.

The trend of the profile of every wheel and all measured parameters of the tread profile are shown in the interface. In case of problems on a wheel it is indicated as alarmed. In the web user interface every wheel is associated to a train that can be identified by RFID tags, plate of cars (detecting by OCR optical character recognition) and information given by the central control unit. The wheel measuring system is currently part of a diagnostic portal situated in Milan. Besides measuring every wheel, it checks at very high speed pantographs and profile of trains giving immediate alarms in case of problems. Thanks to the web user interface, it is possible to check a single train in every part and control every single passage of it through the diagnostic portal. In particular, the Wheel Profile & Diameter measuring system is able to measure:

- Flange thickness;
- Flange height;
- QR Code;
- Rim nominal width;
- Wheel diameter;
- Back to back wheel gauge at very high speed. It was tested at 150 km/h. (Selectra Vision, 2019).

Train detector systems are used to identify trains and wagons at specific locations within a railway network, often at the location of so called wayside train monitoring systems (WTMS) – e.g. hotbox detectors, wheel set condition monitoring and noise emission measurements or similar – or at the gates of ports and terminals to record the sequence of wagons and their licence number in a freight train. They consist of a wheel sensor interface, a RFID detector module, a low power microprocessor and communication interfaces. Eventually also further sensors are added.

The WTMS and/or RFID information must be accurately allocated to a specific wagon and/or its component(s) (a bogie, wheel or axle). The allocation is based on the timestamp sequence of detected axles giving the "axle scheme" of a train (axle by time). As speed and distances are calculated from the wheel sensor signal, the accuracy of that timing is critical. An adaptive algorithm identifies each single wagon of the train based on axle distances. The variation of the distances must be less than 5 cm based on a measured axle distance of 3 m (2 axle bogie) at a speed range from 5 km/h up to 230 km/h or more. The RFID reading module of a train detector reads the wagon number stored in the RFID transponder at the moment of passage. This information is used to pinpoint the wagon of the train and also to identify the running orientation of the wagon. This enables the proper allocation of the axle oriented WTMS-measurements axle to the wagon and/or its axle or wheel. (Novak, 2017)

Assessment and Recommendation

RFID technology is already mature for Scenario 1 in railway operational environment. There are numbers of commercial products applicable for this scenario. It is recommended to select and adapt the off-the-shelf available products to the proposed application on the testing site, where the developed image-based monitoring system will be deployed, and define the communication interface.

5.9.2.2 Scenario 2: Identification of MRO components

Based on the available information, RFID tags are used to identification of components on production line and storage prior to handover to customers. This is commonly used in the automotive industry. In the railway industry, where the fleet generally belongs to a single carrier, it offers the possibility to be used for the correct recording of vehicle components such as chassis, wheelsets, bumpers and other large equipment or components. At present, labels, stickers, or inscriptions wearing a part number designation are most commonly used for identification. In the railway environment, in the case of identifying wheelsets, metal strips placed on the axle carrying the wheelset number are still used. The disadvantage of these strips is the risk that they can be loosen due to wrong installation and subsequent shocks and vibrations. Loosen strip can irreversibly damage the axle by turning around its axis and gradually grinding the groove. Metal strip is in the picture below.

Figure 5.30 Metal strip.

An alternative to metal strips are, for example, stickers with barcodes or QR codes or inscriptions. OCR technology could be also considered. However, similarly to numerical labels, etc., checking the correct placement of a component on a car is often dependent on the cleanliness and accessibility of the relevant identifier. The RFID technology does not have this limitation except of case of placing an RFID tag by the axle manufacturer. The disadvantage is that it is impossible to read while in operation on the railway line because the RFID tag is hidden behind the axle box.

Figure 5.31 Stickers with barcodes.

Labelling of component by RFID tags to ensure their identification is a common practice in airspace industry. However, there it is not necessary to read the RFID tags during operation (in flight) but only on the ground or directly in the workshops. RFID technology is being deployed across the aerospace industry for applications that run the gamut from maintenance tracking to inventory management, and from asset management to materials monitoring. These varied applications have different RFID requirements based on applicable standards, operating environment, application requirements, the material being tagged, and the processes to which the tags will be exposed. (XERAFY, 2019)

The Maintenance & Engineering (M&E) functions rely on knowing exactly which item was installed at which location on which aircraft by exact serial number – not just a part number. RFID can quickly, easily, and correctly identify that serial number from a tagged part with no human mistakes to invalidate the data. Such data is currently stored in the airline's official System of Record, but with the advent of industry standards and high memory, RFID tags this data, which can also be stored on the part as a portable traceability record. (Skybrary, 2017).

Today, critical parts are being affixed with an RFID tag that can store the following need-to-know information:

- Birth record, including serial number and manufacturing information;
- Current record, including reconfiguration details;
- Maintenance history, including changes in custody/location and utilization/consumption;
- Additional notes added by the technician.

Technicians are able to read and update the tags using a handheld scanner, with collected data being automatically added to a centralized database and real-time reports generated. Unlike manual or even barcode-based systems, RFID simplifies this process by allowing tags to be scanned without a direct view of the label and read multiple tags at once, which is particularly useful for hard-to-reach components. Innovative tag suppliers are also optimizing aerospace RFID tags to have a smaller footprint, to allow for longer read ranges and to handle the effects of metal surfaces and harsh environmental conditions. (Vizinex, 2017)

Assessment and Recommendation

In the industrial and aviation sector, applications of RFID for tracing and managing MRO components can be found. However, this application has not been found in the railway sector within the survey, although the relevant GS1 standards exist (GS1, 2018). GS1 standard concerns the general aspects and the data format of EPC code, rather than the requirements of RFID components in terms of technical specifications and installation positions on rolling stock, which is however extremely important in practice. This will be investigated in the following task. Apart from this, the back-end information system which connects RFID readers and maintenance management system should also be investigated to ensure that the acquired MRO information can be integrated in the maintenance management system.

5.9.2.3 Scenario 3: Reading more pre-written information on tags

RFID tags mostly store an EPC of the physical objects. More details of the physical object are stored in database files that can be accessed via a centralized network. The EPC from the tag is used to retrieve more details about the physical object from the network database. This system of using data from an RFID tag in conjunction with data on the enterprise network is called as data-on-network approach (Pais and Symonds, 2011). For instance, a typical data-on-network approach is presented in Figure 5.32.

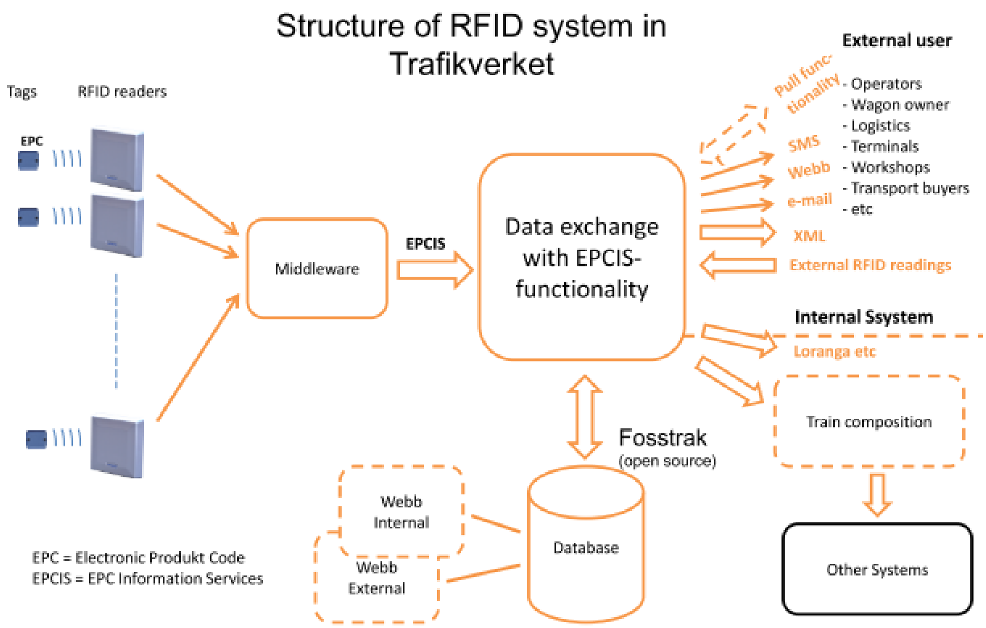


Figure 5.32 Structure of RFID system used by Trafiverket (Ivansson, 2012).

Due to the enhanced data storage capabilities, RFID tags can be as decentralised data storage. RFID tags come with additional user memory apart from that used for storing the EPC. This user memory on the tag is commonly of Electrically Erasable Programmable Read-Only Memory (EEPROM) type, ranging from 256 bits up to 64kB. This approach is known as a data-on-tag approach. In a data-on-tag approach, data is stored on the tag, with lesser reliance on the centralised network database (Pais and Symonds, 2011). The application of data-on-tag approach is not popular because not enough is known about how to store data in a very small space and how to process that data with the limited capability of an RFID reader (Pais and Symonds, 2011).

In the Assets4Rail application, the RFID technology is used in conjunction with WTMS. RFID readers and WTMS belong to infrastructure managers (IM), the diagnostics data can bring benefits for operators/ keepers of railway vehicles. Considering operators/keepers and IM usually have separate systems/ networks, the data exchange between them could be a challenge in practise. In this sense, data-on-tag approach provides a way of data exchange between different stakeholders. Scenario 3 and Scenario 4 address this approach in terms of read and write function respectively.

The concept of data-on-tag is not new. It has been used in production control and maintenance documentation in manufacturing, tracking patients and medical equipment in healthcare (Paris, 2010). Velandia et al. (2017) proposed a RFID based automatic manufacturing and assembly of crankshafts. The proposed solution involves the attachment of bolts with embedded RFID functionality by fitting a reader antenna reader to an overhead gantry that spans the production line and reads and writes production data to the tags. The manufacturing, assembly and service data captured through RFID tags and stored on a local server, could further be integrated with higher-level business applications facilitating seamless integration within the factory. In this case Gen 2 UHF passive tags were applied. The RFID reader reads and writes process information in user memory. Wang et al. (2017) proposed hybrid-data-on-tag approach to solve slow data reading and writing problems. The on-tag data contains the basic machining information and the index of further machining information which is stored in the backend database. Based on the hybrid-data-on-tag approach, the control node model of decentralized control system and the corresponding cyber-physical systems architecture of flexible manufacturing shop floors are presented.

Apart from storing production information, another application of data-on-tag approach is RFID sensor tags. The Deliverable D1.1 of the INNOWAG project (i.e. a Shift2Rail IP5 project) (INNOWAG D1.1, 2017) reviews different types of RFID based sensors and current researches. For railway application, a passive UHF tag was used in conjunction with an embedded temperature sensor for the train axle temperature measurement (Qian et al., 2011). The passive sensor tag was dedicated designed to achieve the total sequence time of 8.75 ms for temperature reader, allowing the pass-by vehicle speed of 400 km/h. The trackside reader obtained the information from the tags, that were attached to each train axle, when the train passes by the reader. This system was successfully tested at a speed of up to 400 km/h (Malakar and Roy, 2014).

In the INNOWAG project, commercial various available UHF RFID readers, extern antennae, generic tags and sensor tags were tested, see Figure 5.33, Figure 5.34, Figure 5.35. It was learned from the INNOWAG project, the read range of readers with extern antennae for sensor tags is a little lower than for generic tags. However, response time for sensor tags is much longer than for generic tags. The energy harvesting module has to accumulate energy from successive read attempts by the RFID reader. When sufficient energy is available the tag responds with sensor data after a delay of typically 5-10 seconds (INNOWAG D2.3, 2018). In a simulated environment of a trackside reader reading a tag attached on a vehicle, the maximum read range is 2.7 m for a generic tag and 2.5 m for a sensor tag. Time to first success read is <0.2 s for a generic tag and 3 s for a sensor tag. As explained, this delay is caused by accumulating power for the sensor. If the data is pre-written in user memory, there should not be so much delay.

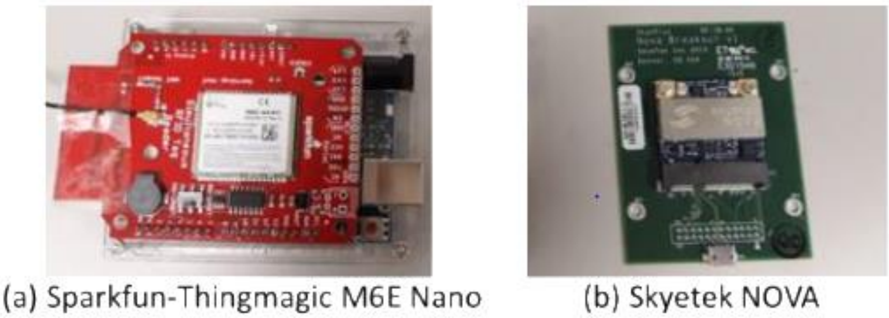


Figure 5.33 Tested RFID readers in the INNOWAG project (INNOWAG D2.3, 2018).

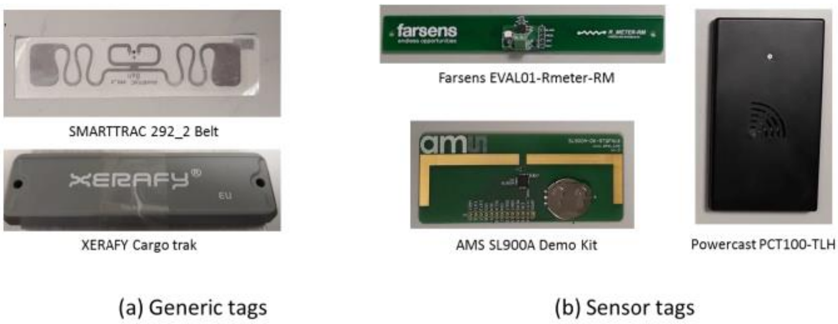


Figure 5.34 Tags in the INNOWAG project (INNOWAG D2.3, 2018).

This project has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement no. 826250 Assets4Rail).

Figure 5.35 Antennae in the INNOWAG project (INNOWAG D2.3, 2018).

In the field test in a shunting yard, the SL900A RFID tag and Sparkfun RFID reader were used. The testing is divided into static tests and dynamic tests. Static tests were carried out with a static wagon to show the system characteristics and the possible reading range of trackside readers. The dynamic tests were carried out with a moving locomotive/wagon to see the feasibility of the concept and whether the system works under a dynamic condition (INNOWAG D2.4, 2019). In both tests, it was programmed to generate sensor data in the user memory of the onboard RFID tag. In the static tests, when the RFID reader was placed in open air, the reading range can reach 7.5 m. In the dynamic tests, the maximum allowed running speed of the shunting locomotive is 25 km/h. The tests were carried out under the max. speed. The reading range can reach 3 m. However, when the read distance is more than 2 m, some data was missing and not every read was successful. An important reason is that the reader was placed between two wagons, which has negative effects on RFID communication.

Figure 5.36 setup of dynamic (left) and static tests (right) of trackside RFID solutions (INNOWAG D2.4, 2019).

Assessment and Recommendation

Reading additional data from RFID tags that is apart from EPC has been applied for industrial applications with the wave of Industry 4.0. The production information such as machine information, process status, etc., can be read and written during processes to achieve automatic control and production. Apart from this, RFID based sensors are also a specific application of data-on-tag approach to achieve condition monitoring without involving batteries for sensor nodes.

In railway sector, there are few applications in practise. Most RFID applications are based on data-on-network approach. There were scientific researches investigating RFID based sensors for monitoring of safety-relevant components of rolling stock. This was also addressed by the INNOWAG project, from which the following things can be learned.

- Surrounding conditions of the RFID reader is significant for communication performance. The RFID should be placed in open-air environment without anything being close to it.
- A part of data could be missing despite of a successful reading. A suitable read distance should be identified without any information missing.
- Given a read distance, the maximum pass-by speed can be estimated. The test should be performed at higher speeds (more than 25 km/h). The test environment should allow the access to track as well as back-and-forth operating of the vehicle. This test environment will be very difficult to find.

5.9.2.4 Scenario 4: Writing information on tags

In principle, RFID readers can read and write additional data on tags at the same time. In the application of automatic manufacturing and assembly of crankshafts (Velandia et al., 2017), the production data was read and written to the tags at the same time. However, considering the specific scenario where wayside readers interact with tags attached to trains over the pass-by time, the time for read/write is very limited, depending on configuration of RFID readers, distance between track and reader as well as types of tags. In addition, the write range, depending on the type and size of the tag, can range between 40 to 80% of the read range. Therefore, it can hardly write information on tags over the pass-by time. In the survey, no comparable applications have been found.

Assessment and Recommendation

The Assets4Rail project will investigate if it is possible to write diagnostic results given by wayside monitoring systems on vehicle-installed tags over the pass-by time in order to provide maintenance related information to operators/keepers, which is data-on-tag approach for data exchange between IM and operators/keepers.

5.10 Conclusions

This section reviews track degradation models and introduces the approaches for the impact assessment of rolling stock failures on infrastructure, especially focusing on track. Since few previous work has been done for this purpose, Assets4Rail proposes a twofold approach to identify the relevant rolling stock failures and quantify the impacts of the specific failures.

In terms of SOA image based WTMS, commercial products of hardware, existing solutions of data processing and handling have been reviewed. There are many image based solutions for train monitoring systems, nowadays stereo imaging has gained relevance in the industry. For the purpose of train monitoring, stereo-line cameras are a novel technology that increases the potential of predictive-maintenance due to its effectiveness and its cost. This stereo-line imaging system combined with another imaging system such as a multiple set of 2D cameras could obtain a depth view of the vehicle's parts. This approach will be used in this project, using state-of-the-art recognition and machine learning algorithms to detect prioritised rolling stock failures.

In addition, RFID technologies and applications in rail have been reviewed. RFID is a mature technology and well applied in diverse sectors for various applications, such as healthcare, retail, transport, etc. In rail, RFID is mostly used for vehicle identification. This basic functionality enable extensive high-level applications through the information systems at back office. This data-on-network approach will be used in the proposed Scenario 1 and 2. In Scenario 3 and 4, the data-on-tag approach will be investigated, where data is directly written and read on the tags. The specific assessment and recommendation for each scenario have been given in Sub-section 5.8.2. Due to the nature of RFID technologies, there will be technical limitations of RFID applications in rail, especially when RFID systems are applied in metal environment. These limitations should be identified, and the recommendations to tackle these limitations should be given in the future investigation.

6. Sensor system to support track geometry monitoring (Task 6.1.2)

6.1 Relevant works in IN2SMART and IN2RAIL

The IN2RAIL Project "*Deliverable D2.3 Embedded & Integrated Sensors: Systems Design Hierarchy*" and "*Deliverable D2.4 Embedded & integrated sensors - System Design Selection*" give an overview and the describe methods and sensors for measuring cinematic and dynamic parameters to determine and monitoring track geometry, applied on switches and crossing. An overview of these is given in the paragraphs 6.1.1. "*Deliverable D4.1 Report on track/switch parameters and problem zones*" specifies requirement specifications of track geometry measurements, which is integrated in subsection 8.3 Requirements Register.

Concerning the project IN2SMART , WP4 is to develop cost effective and robust solutions for monitoring track geometry and rail defects by measuring the accelerations of commercial in-service trains along multiple axis, in continuation of In2Rail WP5.

Areas of interest for research/development regarding track using accelerometers are geometry (lateral alignment and twist), rolling contact fatigue, fastening systems, switches and crossings.

Monitoring of track geometry with in-service trains is more and more of interest for the infrastructure managers and railway undertakers.

Monitoring of the longitudinal levelling of track is a straight forward method because there is a strong relation between the vehicle reaction, like vertical acceleration of axle boxes, and the vertical track defects. Therefore, monitoring of track geometry with accelerometers is mainly done only for this track geometry parameter like in the ICE2 for several years [DB2003].

Looking at other parameters of the track geometry such as twist or lateral alignment, which are measured by inspection cars and which are also essential for maintenance and safety issues, it is much more complicated to monitor these parameters with in-service trains using accelerometers.

In the market unattended systems are installed on in-service trains that measure all track parameters but they use expensive inertial platforms (e.g. Velaro in Russia or Mermec).

Beside this, an accurate position of the measurements is important to verify the reproducibility and to provide sufficient information for trend analysis to predict the degeneration and to identify the root cause.

The concept developed by IN2SMART for monitoring/estimating the lateral alignment from an onboard monitoring system (measurement of accelerations) is showed in the Figure 6.1.

This project has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement no. 826250 Assets4Rail).

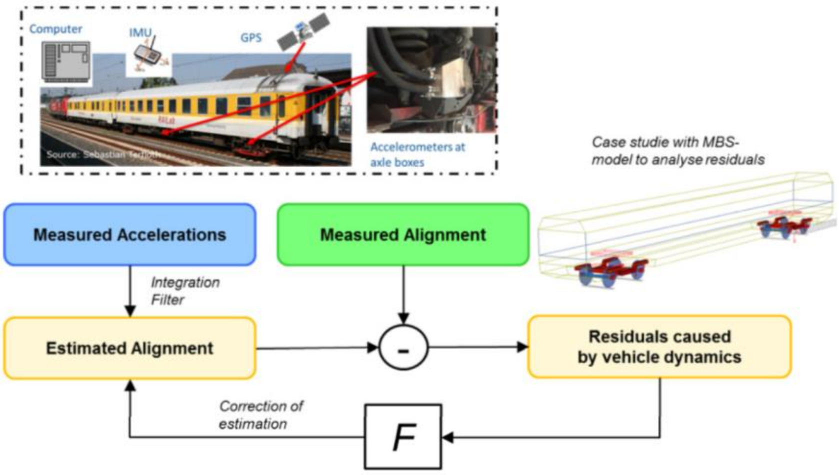


Figure 6.1 Concept for estimating the lateral alignment developed by IN2SMART.

6.1.1 Sensors and methods for track geometry measurements

In IN2RAIL Project are described the principal sensors used and methods to determine the degradation of track geometry over time, summarised in Table 6.1 and listed below:

- accelerometers;
- geophones;
- video camera followed by digital analysis of the images captured (digital image correlation, DIC);
- laser systems with Position Sensitive Devices (PSD) mounted on sleepers/rails to detect the position of the laser;
- multi-depth-deflectometers (MDD).

Table 6.1 Summary of methods for measuring dynamic track displacement.

Instrument and parameter measured	Attributes, advantages (+) and disadvantages (-)
Geophone: velocity [A]	**+Easy to deploy** **+Requires only one stage of integration and filtering** -Train speed must be above geophone natural frequency **-Signal processing requires skill and care**
Accelerometer: acceleration [B]	+Micro Electrical Mechanical devices (MEMs) are low cost **+Easy to deploy** **-Requires two stages of integration and filtering** -Signal processing requires skill and care
Digital Image Correlation (DIC) of high speed filming: displacement [C]	**+Can be used for any realistic speed of train** +Accurate at lower speeds where accelerometers, gyroscope and geophones tend to be less reliable -Susceptible to vibration at the camera location (ground borne and wind), although methods to correct for this are available **-Line of sight may be problematic**
Laser based systems: displacement [D]	As for DIC although differing processing methods may result in relative differences in accuracy.

Instrument and parameter measured	Attributes, advantages (+) and disadvantages (-)
Multi depth deflectometer: displacement [E]	LVDT (linear variable differential transformer) based system installed in shallow vertical borehole in trackbed +Will give an absolute measure with no zero shift and will in principle measure permanent settlements **-Requires fixed datum at depth** **-Difficult or problematic to install**

References and further reading:

A. (Bowness et al., 2007), (Priest et al., 2013), (Le Pen et al., 2014), (Le Pen et al., 2016)
B. (Lamas-Lopez et al., 2014)
C. (Bowness et al., 2007), (Le Pen et al., 2014), (Murray et al., 2014)
D. (Paixão et al., 2014), (Kim et al., 2014)
E. (Gräbe & Shaw, 2010), (Priest et al., 2010), (Mishra et al., 2014)

Moreover, excessive forces appear more frequent in S&C than in plain line. Forces and accelerations can be measured to predict wear and fatigue.

New measurement methods could increase the precision of the estimation of the rolling contact fatigue and of the rail wear.

Optical fibers attached to the rail can also be used to detect broken rails, or be used as strain sensors to measure the forces at the wheel-rail contact (Kourousis, 2015). Finally, the measurement of the noise produced by the wheel-rail contact may inform the infrastructure manager about the degradation of the rail surface.

Table 6.2 Summary of methods for measuring rail impact and temperature stress.

Instrument and parameter measured	Attributes, advantages (+) and disadvantages (-)
Strain gauges (electrical)	+ Not depended on train speed + Low cost - Relative measurement as strain gauges is not stable over time - Signal processing requires skill and care
Fiber optical strain gauges	+ Absolute measurement as they are stable over time - Strain is temperature dependent - High cost
Piezo-electrical force transducers	
Inductive sensor	
Microphone	+ Easy to deploy - Requires filtering and spectral analysis - Signal processing requires skill and care - Dependence on train speed
Ultrasonic	
Acoustic Emission	

Rail profile and gauge can be measured by laser cameras mounted on passing vehicles. Embedded measurement systems could be used, but to form a valid business case the failure development time must be shorter than the interval possible for vehicle mounted inspections.

Monitoring of the check rail gauge and blade gauge is motivated as an enlargement or a narrowing of these gauges could make the train derail. Distance sensors could be installed on the rails to monitor the gauge. Most common distance sensors are laser sensors using time of flight technology.

Table 6.3 Summary of methods for measuring rail impact and temperature stress.

Instrument and parameter measured	Attributes, advantages (+) and disadvantages (-)
Laser profile measurements	+ Several commercial systems available for wheel profiles + Several commercial embedded systems available to measure profiles - Sensitive to environment (rain, dirt) - Measuring rail profile is only possible at short distance and placing outside the track will not give good precisions. Automatically moving laser scanner into the track area by robot arm or drones is unlikely to be realized and will be costly, but is mentioned as a possible solution
Proximity sensor	+ Established and reliable technology + High precision at short distances - Normally just measure a specific distance - Limit in distance
Distance sensor laser, time of flight	+ Can measure all distance + High precisions - Measuring one point
Time of flight camera	+Take a whole picture - Low resolution

Finally, in the following Table 6.4 are summarised the monitored parameters with embedded sensors.

Table 6.4 Information asked for and possibility to measure by condition monitoring with embedded sensors

Quantity to be measured	Measurement technology	Data used for	Information sought/purpose
Vertical displacement and acceleration Noise	Accelerometer Displacement sensors Microphone	Management Maintenance Research & Development	Structural displacement and impact force in crossing panel. The impact force doesn't follow directly from the acceleration, but can be estimated.
Strain Contact area Lateral wheel position at transition Forces in checkrail	Strain gauge Ultrasonic probe Radar, lidar Piezo-electrical sensors	Research & Development	Crossing deformation Switch rail deformation Contact zone This information can be used to estimate contact loads and the fatigue life of the crossing and switch rail. It might also be possible to estimate where a particular wheel makes the transition.
Rail profile Kink of crossing	3D-laser-scanner	Maintenance Research & Development	Switch panel geometry Crossing running surface geometry Repeatedly profile measurement will give a better understanding of the wear and deformations.
Distance	Laser	Maintenance	Check rail gauge
Forces in rail-sleeper	Piezo-electrical load cells	Research &	How the load is distributed and transferred from rails

Quantity to be measured	Measurement technology	Data used for	Information sought/purpose
connections		Development	to sleepers.
Sleeper displacements and accelerations along their length Ballast displacement	Accelerometer Displacement sensors	Research & Development	Structural deformation. With information on sleeper deformations and accelerations it should be possible to obtain information about how the impact load of the crossing propagates down through the track structure.
Sleeper-ballast contact pressure	Pressure sensors	Research & Development	Information on the sleeper-ballast contact pressure is the most important parameter when it comes to determining ballast crushing and track settlement.
Cracks	Ultrasonic probe (using long wave), eddy current and acoustic emission	Maintenance Research & Development	Local crossing damage (e.g. RCF, squats, cracks etc.); Type and location of damage for the validation of damage models.
Track geometry (vehicle based) 3D laser scanner on site	Track geometry car/In service vehicle with accelerometers	Maintenance Research & Development	Long term development of track irregularities/settlements
Strain Rail temperature	Strain gauge Temperature sensor	Operation	Longitudinal stresses
Sleeper strength	Acoustic emission	Management Maintenance	Sleeper condition, remaining useful life
Energy need over time for the movement	Electric current, force, position, time of movement	Operation Maintenance Management	Friction or poorly adjusted position detectors
Form of switch blade	Fiber-optic probes Strain Machine vision 3D scanner	Research & Development	Switch blade form during movement and final position
Picture to identify objects	Machine vision	Operation	Switch blade obstruction
Vibration in different type of bars	Accelerometers	Operation Maintenance Management	Motor, control and stress bar condition
Outdoor temperature Rail temperature Wind Moisture Precipitation	Temperature sensor Anemometer Hygroscope	Operation	Local weather prognosis

6.2 Technologies for measuring track geometry parameters

Infrastructure managers (IM) commonly deploys a dedicated track recording vehicle (TRV) or hauled track recording coach (TRC) running around the rail network gathering track geometry data for the purpose of inspection/measurement (Weston et al., 2015). The measurement by a TRV/TRC is mature technology, standardized in the standard EN 13848 series. The EN 13848 series covers several aspects concerning characterization of track geometry and measurement devices, including measurement methods. The following track geometry parameters are covered: track gauge, longitudinal level, cross level, alignment and twist. These parameters can be measured by using either an inertial system or a versine system, which is also the measurement principle of TRV/TRC. (RIVAS D2.5, 2013) reviewed TRC operating in Switzerland, Sweden and the UK describing the mentioned measurement methods.

Table 6.5 standard EN 13848 series for track geometry quality.

EN 13848-1:2016	**Characterisation of track geometry**
EN 13848-2:2006	Part 2: Measuring systems - Track recording vehicles
EN 13848-3:2009	Part 3: Measuring systems - Track construction and maintenance machines
EN 13848-4:2011	Part 4: Measuring systems - Manual and lightweight devices
EN 13848-5. 2017	Part 5: Geometric quality levels – Plain line, switches and crossings
EN 13848-6. 2014	Part 6: Characterisation of track geometry quality

In recent years, IMs attempt to deploy unattended geometry measurement systems (UGMS) on in-service vehicles without interrupting the normal traffic (Weston et al., 2015). For instance, Mermec provides both inertial based and optical versine based UGMS for measuring full track geometrical parameters and full rail profile at high speeds, see Figure 6.2. UGMS is compact and lightweight so that it can be installed on in-service vehicles. London Undergroud was the world's first metro to introduce "unattended" measuring technology for the monitoring of its 408 km of track (Mermec, 2012).

Figure 6.2 Mermec inertial based (upper photo) and optical versine based (bottom photo) UGMS (Mermec, 2019).

Track geometry monitoring is a different thing. Monitoring aims at indicating track geometry quality and detecting the local track geometry deviations, rather than reconstruction of track geometry parameters, that inspection/measurement systems do. Despite the push for UGMS on in-service vehicles, many track condition monitoring systems that do not give full geometry have been developed, and a few commercialised. (Weston et al., 2015) reviewed the relevant monitoring systems used by IMs as well as in academic and experimental researches.

It can be learned that acceleration measurement along with advance data processing techniques is the most popular option due to the robustness of accelerometers. In comparison, optical sensors, such as laser based, camera based, etc., have to be cleaned regularly to keep it working and thus needs special treatment to avoid getting dirt when applied on an in-service vehicle. Robustness is the most important factor for the applications on in-service vehicles, as the monitoring system should not cause additional maintenance, affecting the reliability and availability of the vehicle itself.

If the reconstruction of track geometry parameters is not required, the indicators can be defined to reveal the quality of track geometry. For instance, prEN 13848-1: 2016 recommends calculation and analysis of mean to peak and/or peak to peak values of axlebox accelerations in the given frequency range which are linked to dynamic wheel-rail forces and to isolated defects. Standard deviation over a specified distance and a given frequency range can be used for assessing corrugation and / or density of short geometric defects of the rail. Accelerations on bogie and carbody can also be used to represent isolated defects by comparing the amplitude from the mean value to the peak value or from zero to the peak value with the defined value by the IM. As a good example of commercial applications, Perpetuum's wireless sensor network measures axlebox acclerations and define the dedicated indicators for track condition, see Figure 6.3. Based the indicator exceeds the calibrated threshold, the IM can be informed that some track sections need maintenance.

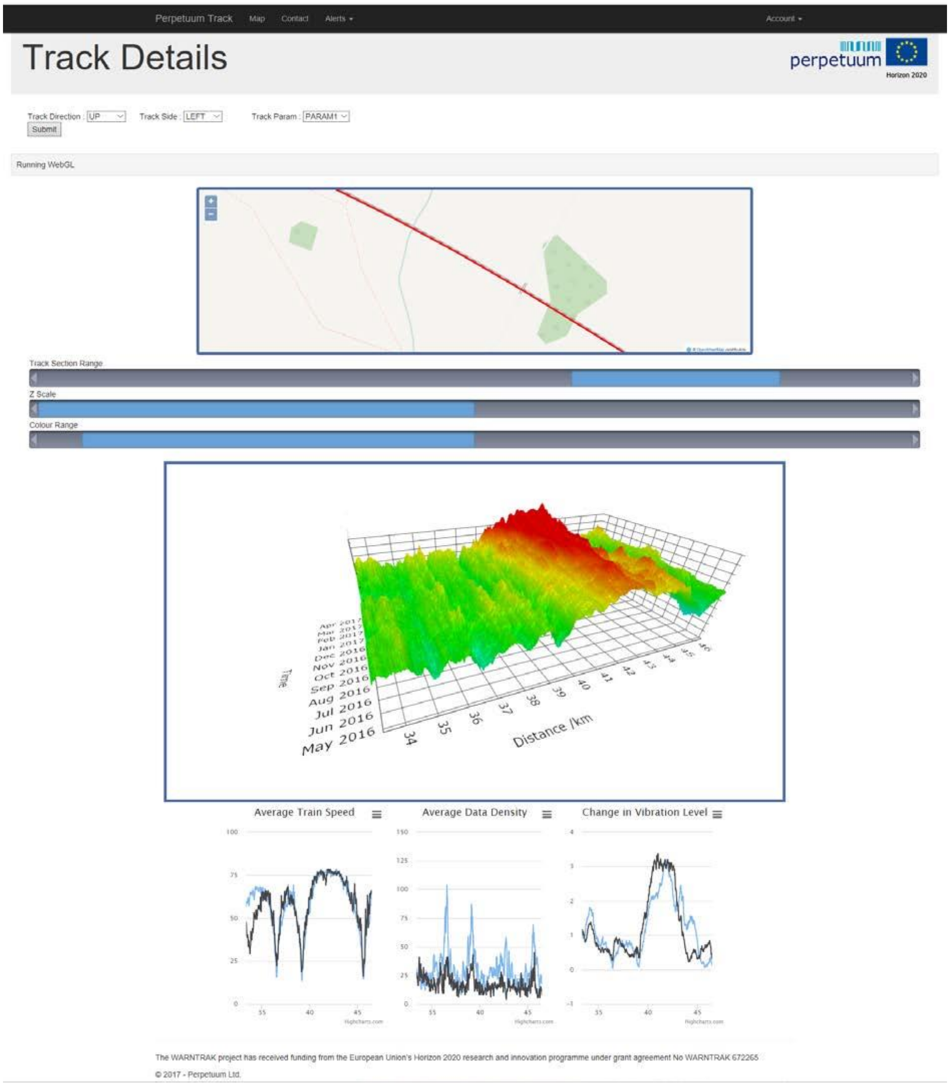


Figure 6.3 Long-term localised trends in track condition (PER, 2016).

If the reconstruction of track geometry parameters is required, specific data handling is needed. Longitudinal level and vertical rail profile can be obtained by double integration of vertical accelerations. In terms of accuracy, it is probably best measured with acceleration on the bogie over the axlebox, and the vertical displacement to the axlebox from the accelerometer, despite displacement transducers are more vulnerable and expensive than accelerometers (Weston et al., 2015). To avoid involving displacement transducers, axlebox mounted accelerometers in conjunction with dedicated signal processing techniques can also be used. RAIDARSS-3 installed on shinkansen train sets (Tsunashima et al., 2012)

Cross-level and twist can be obtained by sensor fusion of a yaw rate gyro, roll rate gyro, laterally sensing accelerometer, and vehicle speed (Lewis, 1988).

However, Lateral alignment is hard to measure accurately without optical sensors. Lateral acceleration at the bogie plus lateral displacement from the bogie to the wheelback has been tried, but the displacement component was found to be small and could be done without. The lateral alignment is then the lateral alignment according to the plan view trajectory taken by the wheelsets, which is not the same as the lateral

This project has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement no. 826250 Assets4Rail).

alignment of the track. Determining track lateral alignment and gauge without optical sensors is not practical (Weston et al., 2015).

To address the measurement of track lateral alignment, Weston (2007) proposed to use an inverse dynamics model to calculate the bogie lateral movement on the rail by bogie yaw rate gyro data or lateral acceleration data so that an approximation of 35 and 70 m lateral track alignment can be obtained. However, this method is not straightforward and cannot measures the alignment with short wavelength (ranging from 3m to 35m).

6.3 Monitoring solutions for assessment of track conditions

Track condition monitoring is an essential activity to ensure the safety of railways operation. Nowadays, Infrastructure Managers use diagnostic vehicles as probes to analyse track condition and detect in real-time potential problems at an early stage. The aim of these diagnostic activities is the improvement of maintenance strategies but also the infrastructure management cost reduction.

The principal monitoring solutions for assessment of track condition can be summarize in Table 6.6).

Table 6.6 Monitoring solutions for assessment of track conditions.

N.	System solutions	Monitored parameters	Technologies	References
1	Track geometry	Rail profile geometry Accelerations	Optical / Laser profilometer Inertial Measuring Unit	MERMEC GROUP www.mermecgroup.com DMA www.dmatorino.it ESIM www.esimgroup.com AMG www.amg-tech.it MARINI IMPIANTI www.mariniimpianti.it
2	Wear monitoring	Rail profile geometry Rail surface Equivalent conicity	Optical / Laser profilometer Video camera	
3	Vehicle dynamic behaviour	Accelerations (axle box, wheelset, bogies and vehicle body) Wheel-rail contact forces (rail strain, wheelset strain, wheel strain)	Electrical strain gauge Accelerometer	
4	Non-conformity of the track	Missing fastenings, sleepers, track bed, etc. joint status, crack, scratch, etc.	Video camera	
5	Rail temperature	Temperature	Thermal probe Thermal camera	

<u>Track geometry system</u>

The track geometry system is based on non-contact optical/inertial technology (Figure 6.4). Optical sensors are used for the measurements of the rail profile and the rail location, while the inertial unit make available the linear and angular accelerations. The combination of optical and inertial data allows the determination of the rail geometric parameters (gauge, cant, cross-level, twist, alignment).

This project has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement no. 826250 Assets4Rail).

Figure 6.4 Track geometry system.

The optical measurements are based on the application of the technique called "light sectioning", capable of producing a section profile of the railhead from which extract the points named "at the top" and "at the gauge" of the rail (Figure 6.5), useful to evaluate the geometric parameter.

Figure 6.5 Top and gauge points.

The calculation is realized by applying the "triangulation" principle, based on the combined action of a laser and a high-speed camera (Figure 6.6), without any mechanical contact with the rail. Each camera is combined with a special high selectivity optical filter, set on the specific laser emission frequency: this makes the system immune to interference of the sunlight or other light sources, allowing the system to operate in any light condition.

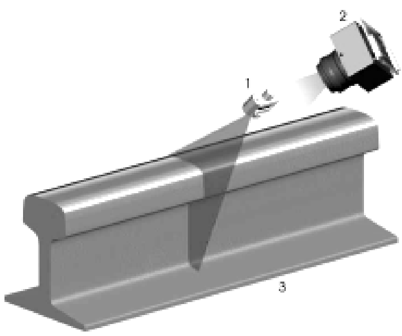


Figure 6.6 Triangulation principle (1 laser, 2 camera, 3 rail).

The information obtained from the optical sensors and inertial unit refer to a common coordinate system. Optical sensors allow to measure the distances between the optical boxes and the track, while the inertial sensors make it possible to establish the position of the optical groups in the real space. The system is provided also of an inclinometer. In this way the geometry measurements of the track can be obtained referring to an absolute reference system, compensating the vehicle movements respect to the rolling plane.

Usually, a track geometry system includes at least two optical boxes (laser + camera) and an inertial unit. The system is installed on the bogie near the axles. As said above, the track parameters derive from the combination of the data related to the optical box and the inertial unit. Through a elaboration process, the images acquired by the camera are converted into millimetres, then the points "at the top" and "at the gauge" (usually detected -14 mm below the rolling plane) are calculated. These points form the basis for calculating the track parameters.

Data relating to optical and inertial sensors are sampled in the space domain. Eventually, the track parameters are connected to the railway line information (localization).

The European standard for track measurement systems is EN 13848-2. This European Standard specifies the minimum requirements for measuring principles and systems in order to produce comparable results. It applies to all measuring equipment fitted on dedicated recording vehicles, or on vehicles specifically modified for the same purpose.

The correct calibration of the system is essential to obtain accurate performance and reliable measurement data. The inertial unit is usually calibrated in the factory while the optical boxes are calibrated after being mounted on the bogie. The optical boxes must be re-calibrated every time the system is unmounted for the vehicle maintenance and when there are hardware or software changes to the whole system. The optical calibration is also done using manual calliper to check the presence of offset error on the gauge and the cant (Figure 6.7).

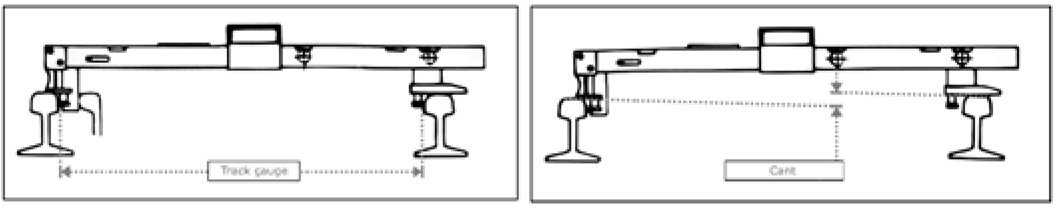


Figure 6.7 Calibration of a track geometry system.

2) Wear monitor system

The wear monitor system uses no-contact optical measurements of the rail profile geometry and railhead surface recorded with cameras and lasers. The monitored track parameters are vertical, horizontal and 45° wear of the rail profile, rail corrugation and equivalent conicity. Usually, a wear monitoring system includes two optical boxes (laser + camera) and is installed under the train coach.

The system uses the same "light sectioning" technique of the track geometry system.

The triangulation principle, applied for measuring the track geometry, is also valid for rail wear measurements. The corresponding processing consists in the overlap of the rail profile acquired by each camera on a reference rail section in order to identify the inclination and wear levels of the railhead.

The reference rail profile can be defined manually by the operators or automatically recognized based on a database of rail profiles. The maximum accuracy is achieved when there is the complete visibility of the rail profile from the railhead to the rail foot. When the rail foot is not visible (ex. due to the ballast) the calculation of the wear is still possible only if the rail web is visible. If this condition is not possible, the wear could not evaluate.

The calibration process is the same of the optical boxes of a track geometry system.

3) Vehicle dynamic behaviour system

Vehicle dynamic behaviour system evaluate the running characteristics by means the measurement of the accelerations on the axle boxes, bogies frame and vehicle body, and by the evaluation of wheel-rail contact forces. The magnitude of the accelerations is directly connected with the track condition (track geometry, rail wear, etc.), while the contact forces are monitored starting from the strain measurement of wheels, wheelsets and rails. For the on-board solution, the system uses electrical strain gauges installed on the wheels or wheelsets.

Wheel-rail contact forces

Usually a wheel-rail interaction system is based on the measure of the wheel disc strains using electrical strain gauges. In fact, the wheel disc is sensitive to the resultant force acting on the wheel-rail contact zone.

Thank this method, the system can directly measure the magnitude of the lateral force Y and vertical force Q, while can indirectly measure the derailment ratio Y/Q, the sum of lateral forces ΣY and the contact point distance d (Figure 6.8). The calibration of the system is done with finite element simulations and laboratory tests to evaluate the matrix of the proportionality coefficients between forces and strains.

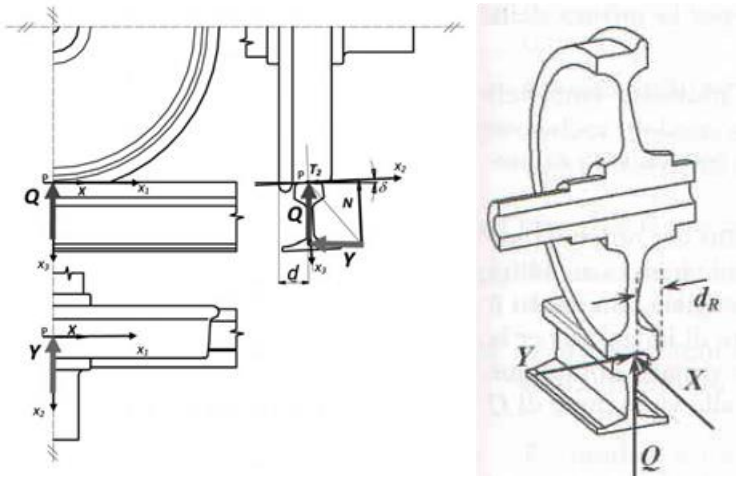


Figure 6.8 Wheel rail contact forces.

The complete definition of the measuring chain requires choosing radial and angular strain gauges installation positions. While the choice of radial position makes it possible to measure the components of the contact force, a proper angular position selection is useful to reduce as much as possible the influence of wheel rotation on acquired signals. There are various strategies aimed to the determination of the radial position of the strain gauges (Figure 6.9). They can be placed on points on the web in which the sensitivity to one of the components is zero, on points with the same sensitivity to contact forces to be properly combined in post-processing or close to holes specifically drilled on the web.

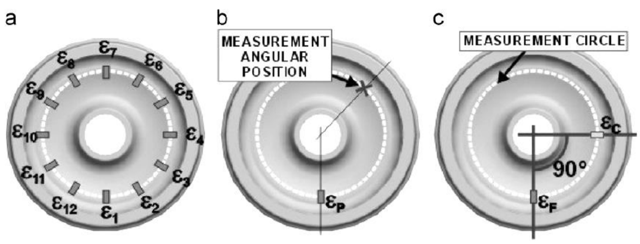


Figure 6.9 Strain gauges position.

Other solutions are based on the measure of the strains on the axle of the bogie. The axle is fitted with electrical strain gauges in order to provide signal correlated to the strain state generate on the axle by the contact forces. In this way, it is possible measure the three contact force components: lateral force Y, vertical force Q and longitudinal force X.

Writing and inverting the linear relationships linking applied forces and bending moments it is possible to estimate vertical, lateral and longitudinal forces starting from strains measured on the axle (Figure 6.10).

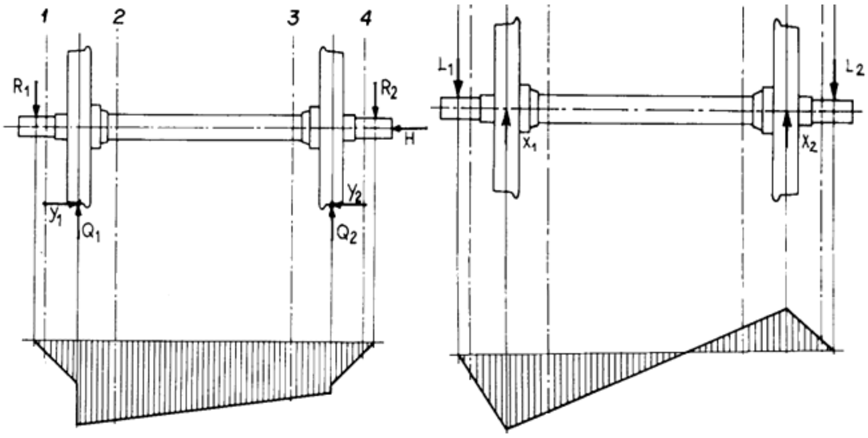


Figure 6.10 Bending stress diagrams in the vertical (left) and in the horizontal (right) planes.

Strain due to torsion can be used to calculate only longitudinal forces X. Strain due to bending can be used to determine X, Y and Q. The bending moment diagram is reconstructed by means of a sufficient number of bridges in the classical bending measurement configuration. This means that it is possible to determine the coefficients in the equilibrium equations of the axle. The inversion of the system of equations gives rise to the applied forces.

The system is usually calibrated in the laboratory to determine the stress-strain proportionality coefficients.

Methods based on axle instrumentation are affected by many "sources of error" mainly due to the wheel-rail contact position change, worsened by the fact that inertia effects of the mass between the contact point and strain gauges are neglected. Signals of strain gauges mounted on the axle also are affected by the rotation of the axle resulting in signals that are amplitude modulated with a purely sinusoidal disturbance in phase with the axle rotation.

Measure of the accelerations

The measure of the accelerations on the on the axle boxes, bogies frame and vehicle body is done by classic mono or multi axial accelerometer (Figure 6.11). Usually each accelerometer is calibrated within a specific measurement range in order to guarantee constant accuracy within the entire frequency band of interest. For example, the interest frequency of the acceleration on the bogies are an order of magnitude greater than of vehicle body, which is isolated from the ground by the primary and secondary suspension. Usually, data are sampled in the time domain at specific frequencies (~2kHz).

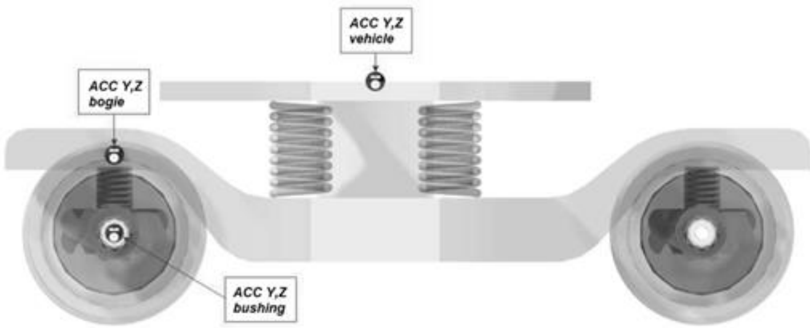


Figure 6.11 Accelerometer on the on the axle boxes (bushing), bogies frame and vehicle body.

The accelerations recorded on the bushing are exclusively referable to track anomalies because of they have not any suspension damping stage. The anomalies are punctual defects with a very short surface wavelength (joint defects, alignment anomalies, etc.). These accelerations can be used to evaluate the hunting oscillation of the vehicles.

The acceleration recorded on the bogies are directly connected to the damping stage of the primary suspensions. The information obtained by these accelerations are concerning the safety of running based on the identification of defects of short-wavelength tracks.

On the other hand, vehicle body accelerations are influenced by the damping stages of the primary and secondary suspensions. They allow to obtain information concerning the running quality, based on the identification of vertical and transversal defects of long-wavelength tracks.

The evaluation of the running characteristics of railway vehicles (measure of the acceleration) must to be compliant to the EN 14363 standard.

4) Non-conformity of the track

The non-conformity of the track is automatically identified by video inspection system. Among the different information, the system can recognise several rail defect or non-conformities like missing fastenings, sleepers and track bed. It also is able to identify joint status, the presence of crack, scratch, etc. The system is based on dedicated algorithms able to process the video images.

Usually the system consists of several subsystems installed on the carriage. Each subsystem has a specific task: the monitoring of rolling plane, the integrity of the rail and the verification of the fundamental elements of the track. The system consists of two optical boxes installed under the vehicle body (Figure 6.12). Grey scale camera and infrared LED illuminator compose the optical boxes.

Figure 6.12 Optical boxes under the vehicle body.

This project has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement no. 826250 Assets4Rail).

The essential objective of the system is to allow automatic recognition of non-conformities such as surface defects, broken joints, missing bolts or fastening, cracks, lack of ballast, presence of foreign elements, etc. (Figure 6.13,Figure 6.14, Figure 6.15 and Figure 6.16).

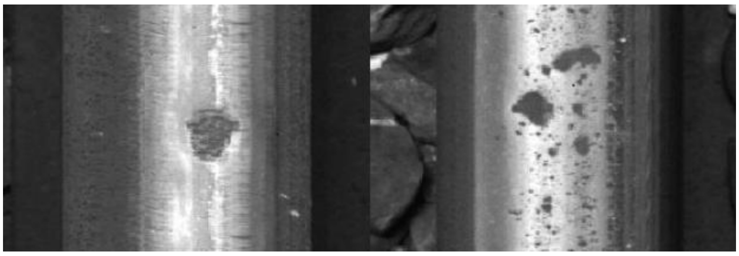


Figure 6.13 Surface defects.

Figure 6.14 Broken joints.

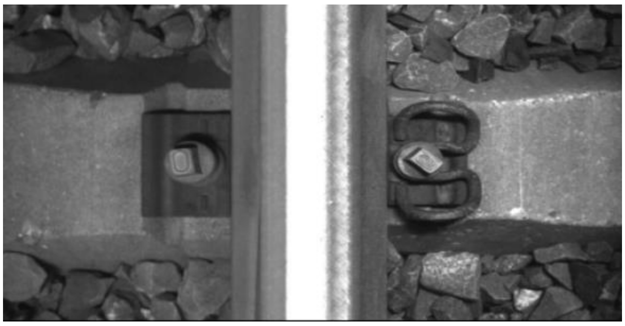


Figure 6.15 Missing fastening.

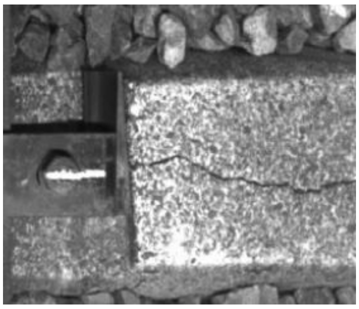


Figure 6.16 Cracks.

5) Rail temperature

The rail temperature system continually measures the temperature of a rail track at specific points of the railway line. The system uses thermal probes installed inside the rail web or thermal cameras mounted on the vehicles.

6.4 Sensors for measuring wheel/rail contact characteristics

6.4.1 Measurement of wheel–rail contact forces

Measurement of wheel–rail contact forces is very important, but it is not executed so often because of its difficulties.

To measure the contact forces, there are two kinds of measuring method, i.e., "on-board measuring" and "wayside measuring". "On-board measuring" can measure along whole route where test train runs, and "wayside measuring" can measure whole vehicles that run through measuring point because no equipment required on vehicles, but only at specified point.

Today on-board measuring uses "specially designed wheelsets". Strain gauges are attached on the wheel web, and the contact forces are transduced from the measured changes of strain of the wheel web. This method is popularly used to measure wheel–rail contact forces on commercial line and test stand on board. It consists of strain gauges attached on the wheel web. Strain gauges on both surfaces of web are gauges to measure lateral contact force, and strain gauges in the centre of holes (neutral axis of web) are gauges to measure normal contact force, i.e., vertical wheel load.

For instance, an innovative solution in which "special wheelset" is lightly equipped is proposed in RUN2RAIL project.

The proposed configuration consists of a Strain Gauge system to monitor in-service components P and Q of the contact forces on at least one axle of a train. The gauge configuration with the best accuracy at lowest number of measurement channels is proposed for testing. The system is capable of providing two samples/revolution for all three force components P, Q, with gauges only on the inner side of the wheel. The system is conceived with electrical resistance strain gauges applied to a wheel for high-speed tilting trains, but other gauge types are possible. The general applicability to other wheel types would need to be studied. The DBS Distance-Based Sampling technique is good for keeping number of channels and processing algorithm complexity low. The algorithm for sample acquisition has already been proven in service but the use of the rain flow algorithm could also be possible through further study.

The four Strain Gauge bridges on the wheel conceptually plug into a conditioning systems. The use initially identified for the system is for in-service load spectra. In this way the processing is kept simple – with data storage on the wheelset till download (no telemetry). With the same setup also predictive maintenance applications could be possible by examining the evolution of signals over time. Telemetry would be required for applications in corrective maintenance, safety monitoring and control signals for active systems.

Another method uses no sensors on rolling part, i.e., wheelset, and no transmission equipment, such as slip rings and telemeters and can realize constant monitoring of wheel/rail contact forces, i.e., derailment coefficient, even from a passenger service car.

In order to realize constant monitoring, the method use non-contact sensors attached on non-rolling parts of the bogie. In conventional method lateral forces are measured by strain values detected by strain gauges attached on the wheel web. In the new method such distortion of wheel is detected by sensors attached on the bogie frame, which is not rotating. The wheel distortion (Figure 6.17) is calculated from the displacement of the wheel rim (4a) which is measured by the non-contact gap sensor (3) attached on the bearing box (2) through the sensor base (7).

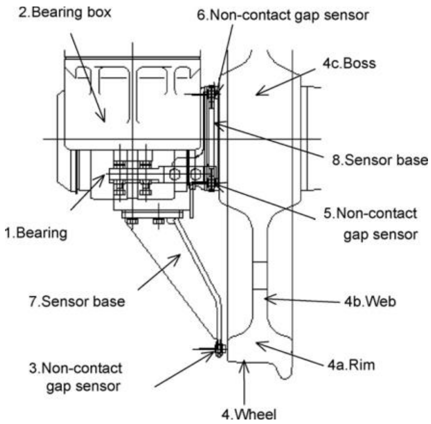


Figure 6.17 Layout of sensors in new method for measuring lateral force.

Because accuracy less than 0.01mm is required, inductive displacement sensors are chosen. In order to get highest accuracy, the gap sensor had better be installed at wheel "rim" position; if it is not possible, another position inside "vehicle gauge", i.e., the out edge of wheel "web", can be chosen. The positions of these sensors are shown in Figure 6.18.

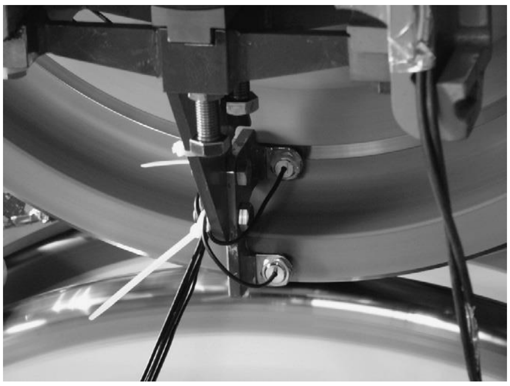


Figure 6.18 Non-contact gap sensors for lateral force measuring—Bogie stand test.

As the measured values by gap sensors are small, the movement of the wheelset cannot be neglected. For compensation of the wheelset movement, 2 gap sensors (5, 6) are attached on the bearing box. In order to compensate, 2 corrections are necessary; one is an axial movement correction δ_4, which is produced by the thrust clearance of bearings, and the other is an inclination correction δ_5, which is produced by the relative inclination of the wheelset against bearing box. δ_4, δ_5 and the wheel lateral distortion after compensation δ can be obtained using the following equations. δ_1, δ_2, δ_3 are the measured gap shown in Figure 6.19.

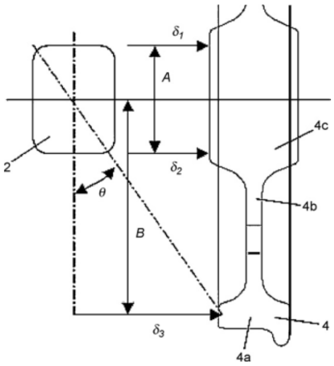


Figure 6.19 Principle of compensation against wheelset movement.

As the normal contact forces, i.e., wheel loads, are important for railway vehicle dynamics, constant monitoring is also necessary. In order to avoid using slip rings and telemeters, measurements from non-rotating parts of the bogie are chosen. Trial measuring are strain of the side frame of the bogie and deflection of primary suspension.

The longitudinal contact forces are also important for the safety and the performance of the railway vehicles especially in curving conditions, but these are rarely measured up to now because of difficulty of measuring. In order to measure them, the measurement of the strain of axle-box support link levers is chosen (Figure 6.20).

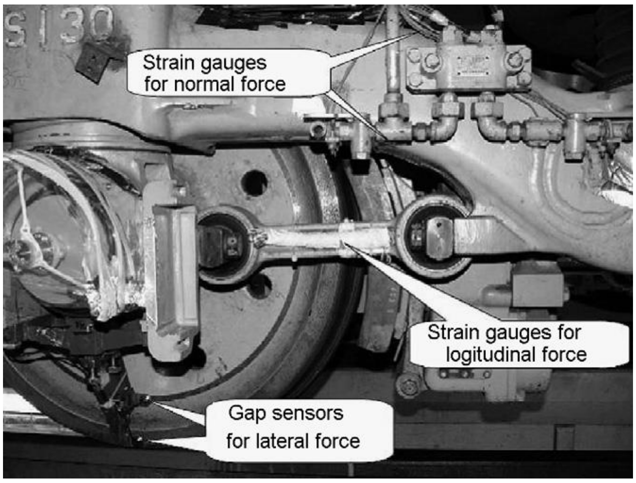


Figure 6.20 Sensors and gauges attached on test bogie—Train running test.

6.4.2 Detection of the wheel/rail contact point

Corazza et al., 1999 proposed a concept for the detection of the contact point between the wheel and the rail.

According to the proposed scheme a wheel engages the rail with an angle α and a variable transversal distance (see Figure 6.20).

In order to detect the wheel-rail relative position, an appropriately shaped arm is applied to the axle-box, which carries at the two ends M and N two displacement sensors which detect the distances m and n.

The two reference points are aligned on the trace of the plane defined by the inner lateral surface of the wheel-tire. Therefore, the average distance (m + n) / 2 corresponds to the distance a, to which the thickness of the wheel-flange k must be subtracted to have the wheel-rail distance, and consequently the position of the contact point.

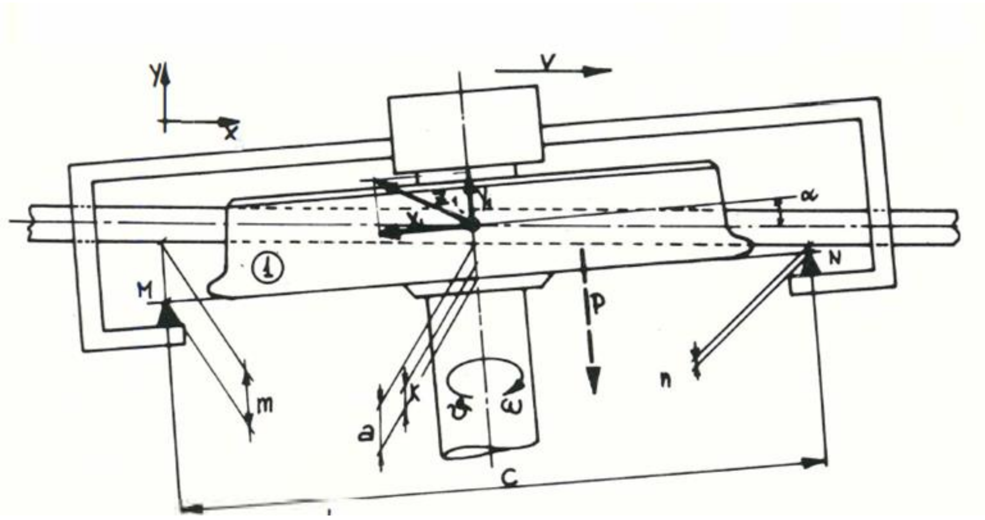


Figure 6.21 Concept for the detection of the wheel/rail contact point.

An example of application of the proposed measurement scheme is that used to measure the relative position of the bogie and the track on line B of the Rome Metro within an experimental campaign carried out in 2002 in order to investigate the wheel/rail contact forces.

The aim of the measure system was to detect the position of the bogie with respect to the track, in an indirect way measuring the two distances showed by the two red arrows in Figure 6.22.

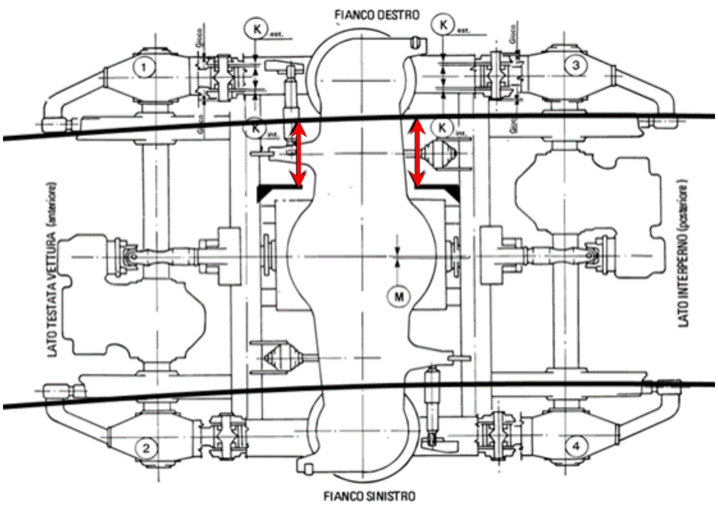


Figure 6.22 Distances detected by the proposed measurement system.

In this case the measurement was carried out using lasers and optical sensors, the laser beam is emitted in the vertical direction downwards, through a mirror placed at 45° it is reflected so as to illuminate the inner edge of the rail (see Figure 6.23). Through the optical triangulation, the sensor creates a signal correlated to the total distance covered by the beam.

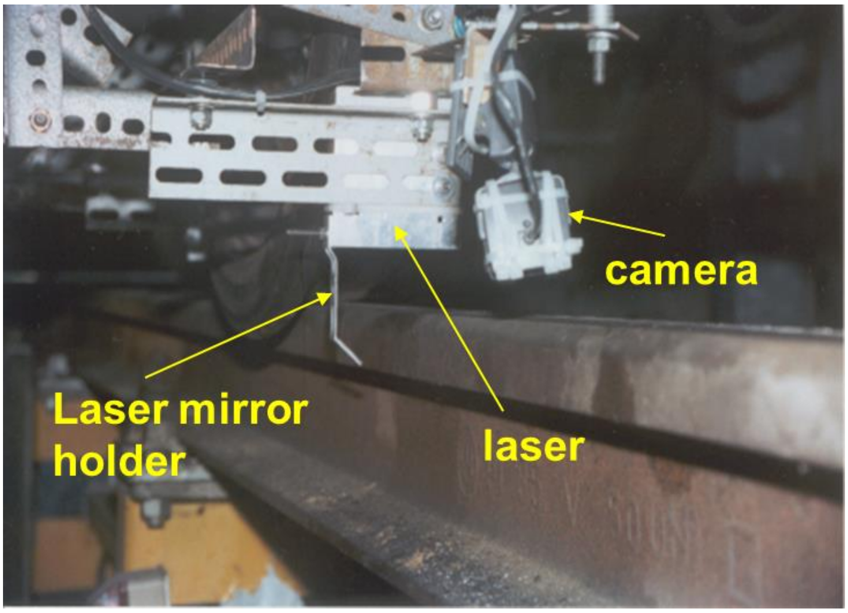


Figure 6.23 Picture of the measurement system mounted on the bogie of Metro B vehicle.

In Figure 6.24 an example of the images provided by the measurement system is shown.

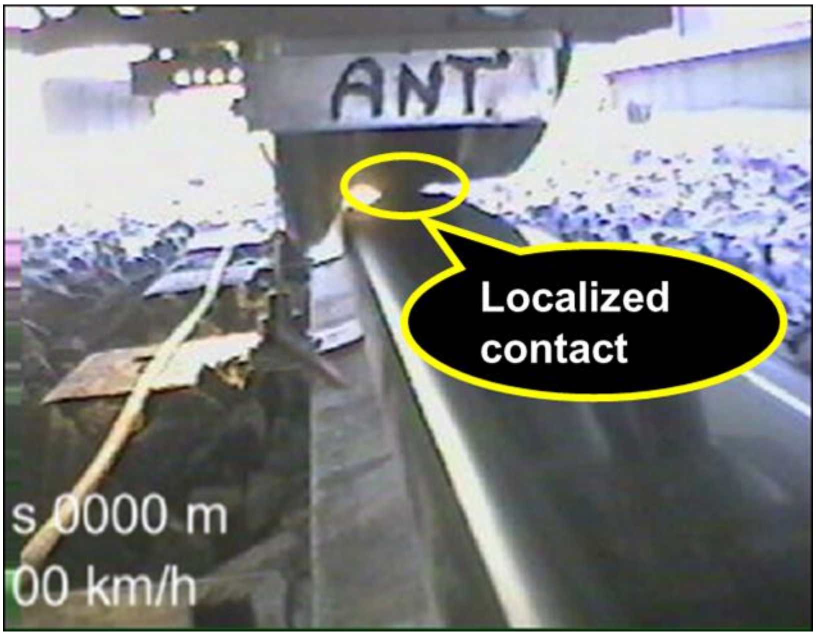


Figure 6.24 Picture taken from the measurement campaign on Rome Metro B.

A critical element of the system is the choose of the size of the mirror: if it is too small, it can cause the loss of the linearity between signal and distance, if it is too large it can interfere with the Gabarit. In any case, the system worked satisfactorily in the described application.

Moreover, the system is delicate due to the mirror and requires frequent maintenance. For this reason, it appears suitable for a measurement campaign but much less to be mounted on a vehicle in commercial operation. Therefore, for the application to commercial vehicles the concept would require to be realized with a different technology.

Burstow et al., 2011 described some tests which were undertaken using a thermal imaging camera to determine whether the forces developed in the wheel/rail contact patch resulted in a change in temperature on the rail head which could be detected and therefore used to indicate the location of the contact patch and possibly infer information about the magnitude of the wheel/rail forces.

A thermal imaging camera was mounted on a railway vehicle which forms part of Network Rail's track measurement fleet and thermal images were recorded on a journey from Crewe to London (157 miles), with much of the running at (or near) line speed. The camera, a FLIR A320G which was capable of imaging data at 60 frames per second, was mounted on the vehicle body directed towards the wheel/rail interface behind the trailing wheelset of the bogie, as shown in Figure 6.25.

Although this mounting position was convenient for the test, it was not ideal as the highest wheel/rail forces would be expected to be generated at the leading wheelset of the bogie. However, detection of temperature changes from the lower forces developed at the trailing wheelset and the interpretation of the behaviour of this wheelset would give confidence that this method would yield useful results if applied to the leading wheelset.

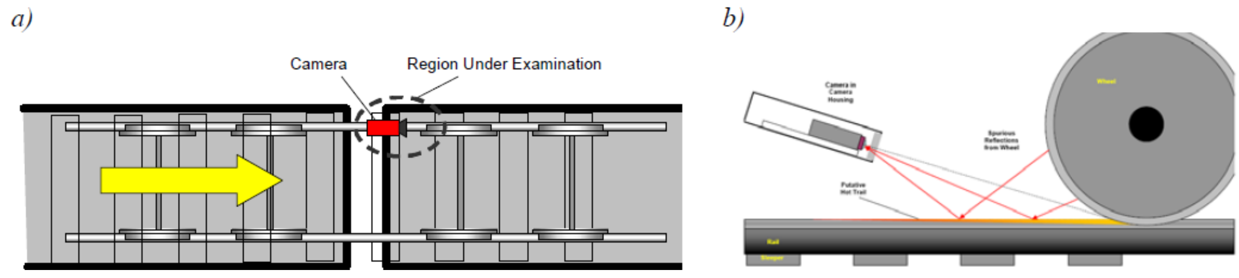


Figure 6.25 Location of the thermal imaging camera showing a) the plan-view position of the camera mounted from the vehicle body, directed towards the trailing wheelset of the trailing bogie, b) the camera directed towards the head of the rail just behind the wheel.

A thermal image of a wheel running on a rail is very complex because of a number of factors:

- Image blurring: the relatively low frame rate of 60fps results in 'blurring' of the image on the rail head when the vehicle is running at speed.
- The low emissivity of the rail: it is estimated that 10-20% of the thermal radiation reaching the camera comes from the rail. The rest of the radiation is reflected from the environment.
- The variable emissivity of the wheel and rail: some parts of the rail or wheel (for example the flange and wheel chamfer) are dull or dirty and have a higher emissivity, which makes them appear brighter in a thermal image.
- Reflections from the 'polished' rail and wheel surfaces: reflected radiation is not diffuse but can show distorted images of other objects. For example, a reflection of the wheel can appear on the rail.

However, despite these complexities in the region on the rail where the contact patch would be expected images showed a clear temperature increase on the rail emerging from the wheel which was not evident in the image of the rail before it enters the wheel, indicating that some temperature rise from the forces in the contact patch was observable. Therefore, the measurements were able to show interesting and useful information about wheel/rail interaction, and the observed changes in temperature agreed reasonably well with theoretical calculations.

By improving the visualization method presented in the Burstow's study, Yamamoto, 2019 proposes a new technique for locating the wheel/rail contact point more precisely, i.e. by installing a thermographic camera under the truck frame close to the rail head, and locating the contact point on thermal images through computed analysis. Running tests were carried out to verify the validity of the proposed technique on the RTRI test line. Tests were carried out at three speeds: 10 km/h, 20 km/h, and 30 km/h.

As the space for installing the camera under the truck frame was small, the compact and light-weight "AIR32 Professional" made by IR System Co., Ltd., with a similar performance to the one used by Burstow et al. was used. To make it easier to measure the frictional temperature of the wheel/rail, the measurement range of this thermographic camera was customized to be from 20 to 40 degrees Centigrade.

In order to locate the contact position of wheel/rail accurately with a simple measurement device, the thermographic camera was mounted under one of the cross beams of the truck frame, close to the rail head, and viewing the target wheel from behind (see Figure 6.26). Depending on the type of the axle box suspension, the running wheelset can be displaced in the yawing direction due to bending of the axle spring, however, since this displacement is very small, it could be ignored.

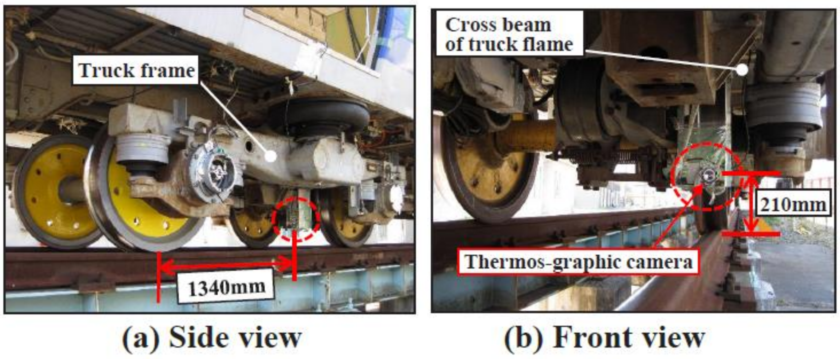


Figure 6.26 Thermographic camera mounted on the truck frame

In order to understand wheelset motion, two techniques were proposed. One technique was to digitalize thermal images based on the relationship between a metal comb type scale and the number of pixels, after which the wheel/rail contact point could be accurately located through calculation. Another technique was to understand wheelset motion by setting two search points: one on the wheel flange and the other on the edge of the wheel tread touching the outside face of the wheel.

As a result of the running tests the wheel/rail contact point could be accurately located and insight could also be gained into wheelset motion, when the experiment was carried out under the following three conditions:

- minimal influence from direct sunlight;
- low surface temperatures of the earth, track facilities, and wheel/rail interface, and little fluctuation in temperature;
- high wheel/rail flash temperature.

To conclude, the above results indicate that it is possible to use to the proposed technique to locate the wheel/rail contact point accurately if the photographing is done in suitable environmental conditions, i.e. performing test runs at night when temperatures are low and do not fluctuate much.

6.5 Conclusions

In the field of measuring the geometric parameters of the track there are and are commonly applied many sensors suitable for detecting the track geometrical characteristics, applicable to vehicles in normal commercial service. While, for the location of the measured measurements, inertial platforms (e.g. Velaro in Russia or Mermec) are used, which are generally quite expensive. Therefore, these systems lend themselves to use on special diagnostic trains, rather than to extensive use on the commercial fleet, which is composed of numerous vehicles.

For this reason, this project intends to focus on systems that allow the detection of the wheel track position that can avoid the inertial platforms to localize the geometric parameter measures.

In the previous sections a possible concept for the detection of the contact point between the wheel and the rail, proposed by Corazza et al. is described.

This concept was realized by means of lasers, optical sensors and a mirror and gave good results during a measurement campaign, but the system is too much delicate due to the mirror and requires frequent maintenance. For this reason, it appears not suitable to be mounted on a vehicle in commercial operation. Therefore, for the application to commercial vehicles the concept would require to be realized with a different technology.

Some candidate technologies have been examined based both on direct measurements:

- lasers;
- high speed cameras;
- stereo cameras;
- thermographic cameras;

and on indirect measurements:

- accelerations;
- ultrasonic reflection.

However, the most promising technologies, to be taken into consideration for subsequent developments, seem to be stereo cameras and thermographic cameras.

In particular, thermographic cameras have already been used in the two experiments described in section 6.4.2 giving good results, even if not for all the speeds and atmospheric conditions that may occur during normal operation of a commercial train.

7. Data collection for diagnostics from fail-safe systems (Task 6.1.3)

In the Assets4Rail project, a modular data collection system, reading and/or generating diagnostics information from fail-safe system, especially focusing on signalling systems, will be developed. The modular design allows a flexible adaption to different signalling components.

The data collection module depends on the target signalling component to be read. The signalling component could be a relay-based analog system that does not provide diagnostic information. In this case, the essential parameters thereof need to be identified so that diagnostic information can be generated by sensing and processing these parameters. The signalling component could be a computerised digital system that provides log data, status information or diagnostic code. In this case, the diagnostics information can be transformed into a defined format for the further usage.

The collected information will be forwarded by the data transmission module in a defined format by using converter proxy developed in the IN2SMART project.

The following subsections introduce the relevant technologies and methods for data collection/sensing, communication interfaces and data transmission. Apart from this, the relevant works done in the complementary project IN2SMART is briefed.

7.1 Relevant work in IN2SMART

In the project IN2SMART is described among the case studies one concerning optimization of maintenance of track circuits through the use of data recorded by track circuits units. The aim is to increase track circuits monitoring and optimize their calibration, management and maintenance.

The Audio Frequency ASTS Track Circuit (AF-TC) is one of the components of the ASTS computer based interlocking system. One of the main issue for track circuits is therefore the false track occupancy, which means that the track is erroneously considered in occupied state due to natural variations of the current levels in the track circuits. These variations, that may affect the correct behaviour of the Track Circuit producing malfunctions, are caused by degradation phenomena and by the influences of the surrounding environmental conditions. The mitigation of these critical events is possible by increasing track circuits monitoring and by optimizing their maintenance.

Consequently, it was developing an anomaly detection data-driven model for the detection of anomalies of track circuits and a data-driven method for predicting the occurrences of false track occupancies.

The data used by these models are:

- *Track Circuits currents*: ASTS collects currents data for track circuits, by exploiting existing diagnostic systems and using them to collect data regularly.
- *Traffic Management data*: data related to the usage of the tracks, and consequently of the trains passed over the track circuits, collected from the Traffic Management System.
- *Maintenance data*: data on the past failures and interventions. ASTS is responsible for full maintenance of some of the plants in which its products have been installed.

These models make possible to identify anomalies, to have an assessment of track circuits status and a prediction of the future status, driving the recalibration and maintenance of track circuits.

This upgrade of the monitoring system is applied in order to implement predictive maintenance functionalities using data-driven techniques. The general idea is to provide a tool to maintainers in order to assess in advance whether an AF-TC needs a recalibration so to prevent false track occupancies. In this scenario, the anomaly is represented by a particular TC status in which the asset is characterized by some malfunction but still not affected by false track occupancy phenomena. Thus, the proposed anomaly detection system should be able to allow to predict in reasonable advance if a TC will go in a false occupied state.

7.2 Data collection/sensing solutions for diagnostics

Railway signalling systems are mainly divided into two groups on a railway block basis. In fixed-block signalling systems, namely, conventional railway signaling systems, railway lines are partitioned into blocks with fixed length, and in moving-block signaling systems, the sum of the length of train and its braking distance is considered as a moving block (Durmus, 2014). The fixed-block signalling systems mainly consist of traffic control center, signalling system control software (i.e. interlocking system), signals, level-crossing, point machines and railway blocks (i.e. track circuits or axle counters). In terms of moving-block signalling systems, track circuits or axle counters are not needed. Apart from this, there are trackside equipment of train protection system, such as balise for ETCS and coded track circuits or cable loops for national train protection system.

Considering signalling comprises a wide range of heterogeneous subsystems, the existing diagnostic approaches are usually defined by the equipment suppliers for a specific product. These diagnostic approaches are integrated in modern signalling equipment. Either the equipment has the defined interface for providing data to extern diagnostic systems or the diagnostic module is embedded in the equipment so that an alarm or diagnostic code can be generated when a failure occurs. For instance, the point machine provided by Voestalpine Signaling Zeltweg GmbH, e.g. AH950, can provide its operational current that is needed to switch the tongue to their diagnostic system RoadMaster. RoadMaster combines the data of current with other external sensor data for diagnosis (Gerald, 2016). Another diagnostic system called DIANA used by DB has a similar scheme. One or more drive motors are attached to the switch. When they set the switch in motion, electricity flows. In the signal box, sensors record the power consumption and transmit the data to DIANA (Figure 7.1). The computer program compares the results with the set-point. If the switch is missing nothing, then the curve corresponds to the reference current curve. If the program detects any deviations, DIANA will sound alarm (Wingler, 2018).

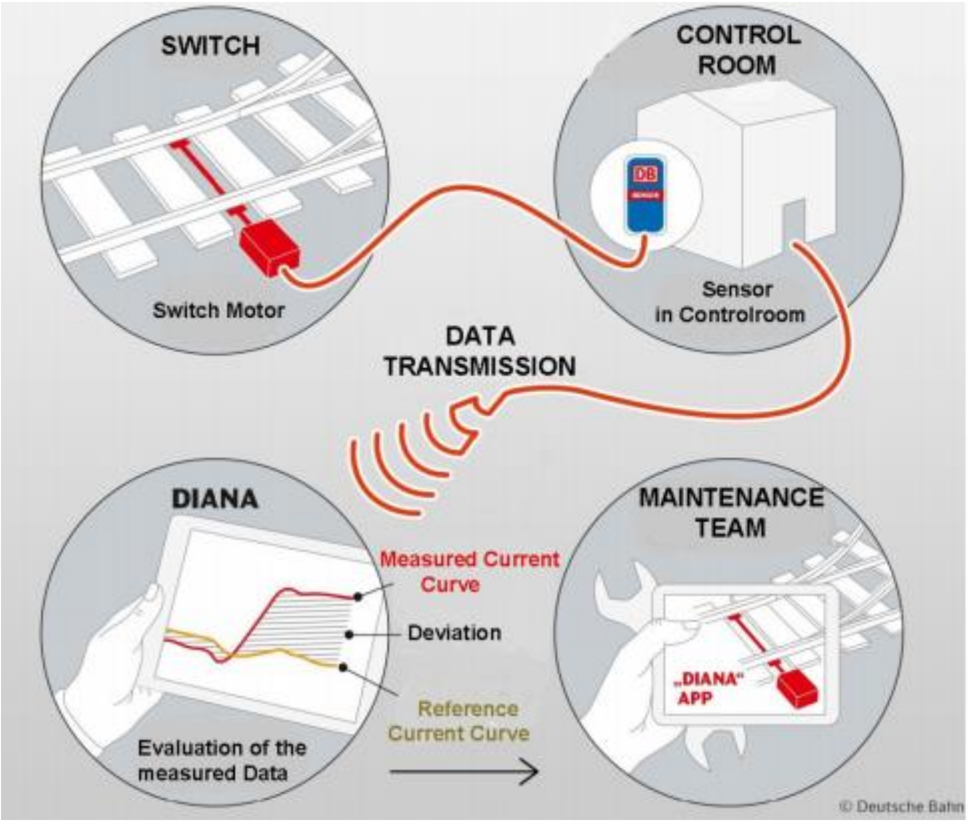


Figure 7.1 Diagnostic principle of DIANA (Wingler, 2018).

PINTSCH level crossing control systems have an embedded diagnosis concept. PINTSCH integrate the diagnosis software in the control system with LEDs on the front panels to permit immediate detection. The diagnosis software can be used for local diagnosis by reading the stored operating, functional data, system configuration and statistical data from the level-crossing control system. It is also allowed remote polling of all data via GSM and GSM-R network (Pintsch, 2019). In the research area, researchers attempt to use the data from the control unit for the purpose of diagnosis, which allows self-diagnostics of the system. (Meuitis, 2016) used the high-power outputs of the object controller of a computer-based interlocking system to develop an expanded set of diagnostic tools.

As described above, the diagnostics-related information in the computer-based signalling systems is usually defined in the specific format that can only be by the software provided by the supplier. In case that an IM operates signalling systems from different suppliers, the IM have to handle with different diagnostic systems, making it difficult for the centralised monitoring and maintenance management at components/subsystems level.

For the systems in which a diagnosis concept is not foreseen such as mechanic or relay-based analog systems, additional no-invasive sensors are needed for condition monitoring and diagnostics. Figure 7.2 presents the essential parameters of the signalling subsystems that are usually measured for diagnostics.

RoadMaster for track circuit monitoring uses a non-invasive Hall-effect sensor monitors energy levels at the DC track circuit receiver (Voestalpine, 2019).

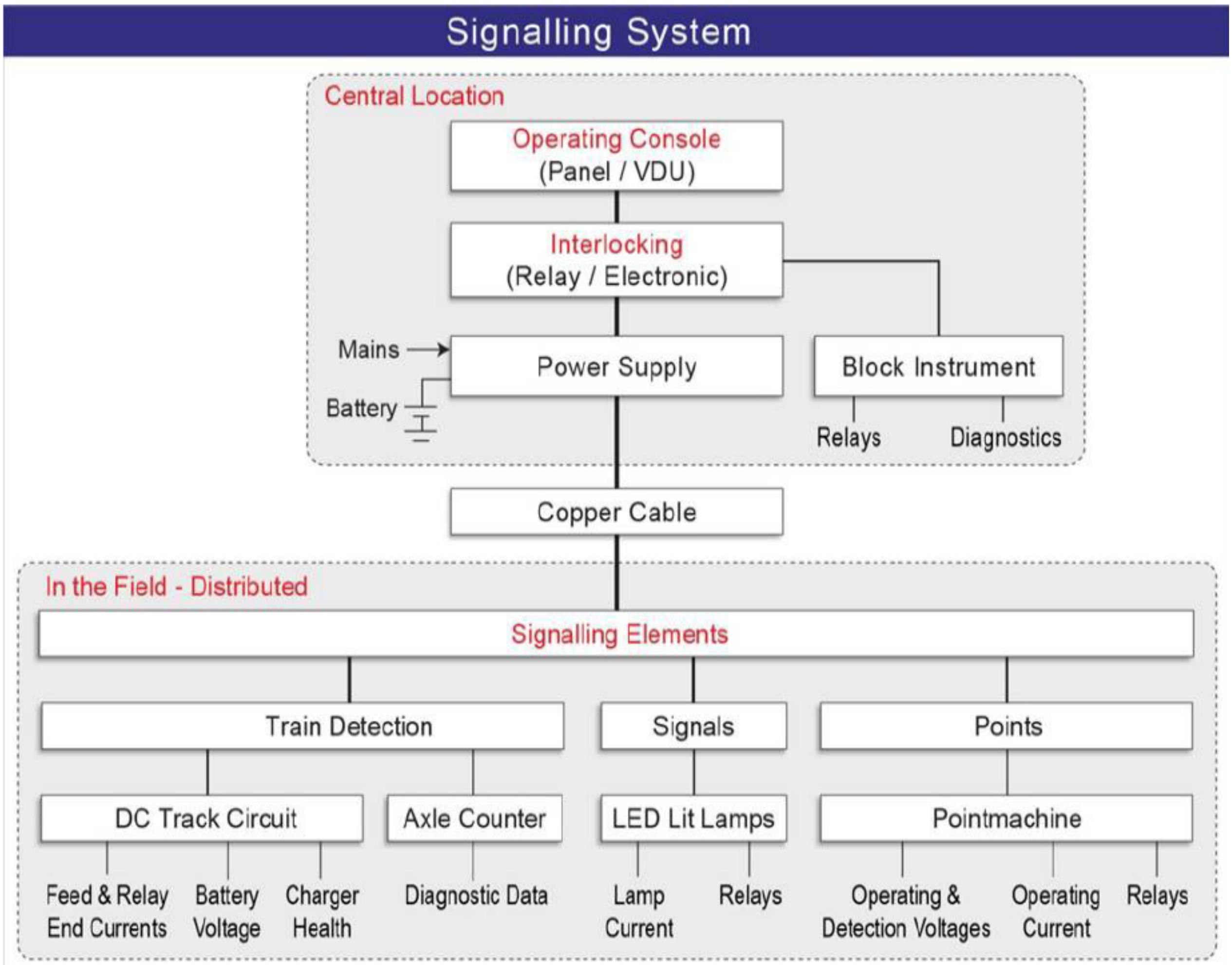


Figure 7.2 Signalling system diagram with the essential parameters for condition monitoring (Saksena, 2019).

7.3 Communication interfaces and data format

7.3.1 Communication interfaces

As per the technical description of the Assest4Rail project, communication interfaces need to be EN 50159 compliant safe communication interfaces. The reference standard divides the transmission systems into three categories depending on their preconditions and specifies different mandatory protections for each of the categories. Category 1 is the less vulnerable and requires the lowest number of measures while Category 3 covers the most vulnerable systems where all measures defined in EN 50159 are required.

In the context of Assets4Rail there is a variety of scenarios and interfaces that, on first approach, correspond to all defined categories:

- Cabled interfaces between the Assest4Rail related on-board systems belong to the closed networks Category 1. The number connections and the characteristics of the transmission system (basically, train bus) are known and fixed. Risk of unauthorised access is negligible.

- Wireless interfaces between on-board systems, between trackside systems and between trackside and on-board belong to the open networks Category 3, where risk of unauthorised access is not negligible.
- Cabled trackside interfaces belong to open network Category 2 or Category 3, depending on the type of equipment involved. For trackside equipment that is located in controlled technical buildings, the unauthorised access risk is negligible (Category 2). On the contrary, unauthorised users can easily access field trackside equipment (Category 3). Category 1 does not apply to these interfaces because of the uncertainty with respect to the performance/characteristics of the transmission system used for trackside equipment connections. In case that field and building located equipment are integrated into the same network, the Category 3 worst-case assumption would apply.

EN 50159 suggests a layered design for safety related communication interfaces. This type of design is not mandatory, but it is in line on state-of-the-art communication protocol implementations. This design principle in based on dividing the tasks of the protocol is smaller blocks or layers. Each layer is responsible for one of more sub-functions and interacts with its adjacent layers. The relations between all layers of the protocol define the so-called protocol stack. By layering, overall design is simpler which implies easier implementation, debugging, testing, incident investigation etc. Figure 7.3 shows a schematic diagram of a generic layered safe communications protocol.

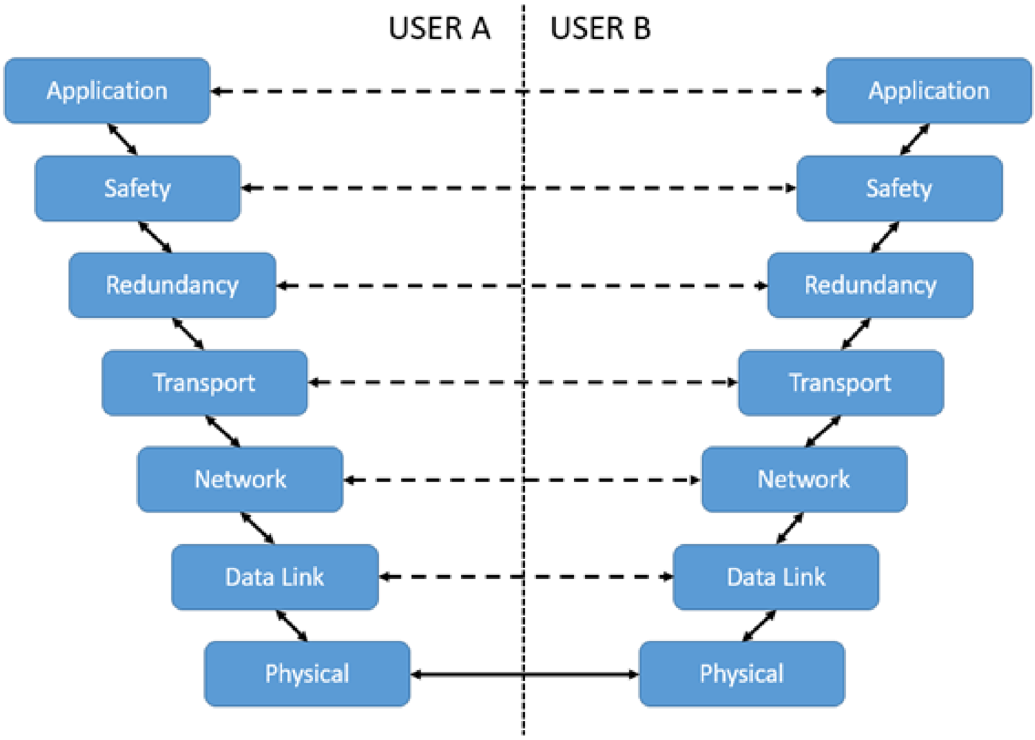


Figure 7.3 Generic design of a layered safe communications protocol.

In a layered protocol, each layer in each side of the interface directly communicates with its lower and upper layers (solid lines in Figure 1). The dashed lines represent the logical interface between layers at the same level at the different sides. The solid line between the physical layers represents the physical data exchanges between the equipment at both sides of the interface through the related transmission system. The safety related software producing and or consuming the user data to be transmitted with the communications protocol is located upstream. Depending on the unidirectional or bidirectional nature of the communication, the data flow represented by the dashed line between the application layers will be in only one or in both directions respectively.

The lower layers of the protocols below the redundancy layer of Figure 1 can be considered purely transmission layers and are analysed in more detail in the following section of this document. Regarding the EN 50159 recommended structure, only de safety and application layers in the upper part of the protocol are necessary. The use of a redundancy layer in the Assests4Rail project is subject to a more detailed study of the availability requirements for the high-level functionality. Redundancy implies the use of independent transmission channels to exchange identical information. The redundancy layer manages transmission/reception by such redundant channels.

For each layer, the modules in both sides analyse the related layer information in order to implement the functionality allocated to the layer. A *header* + *payload* structure is commonly used. The *header* contains the layer related information. The *payload* of the message is forwarded unchanged to the upper/lower layer and contains the relevant information for these adjacent layers. The following Figure 7.4 shows a basic representation of the *header* + *payload* structure.

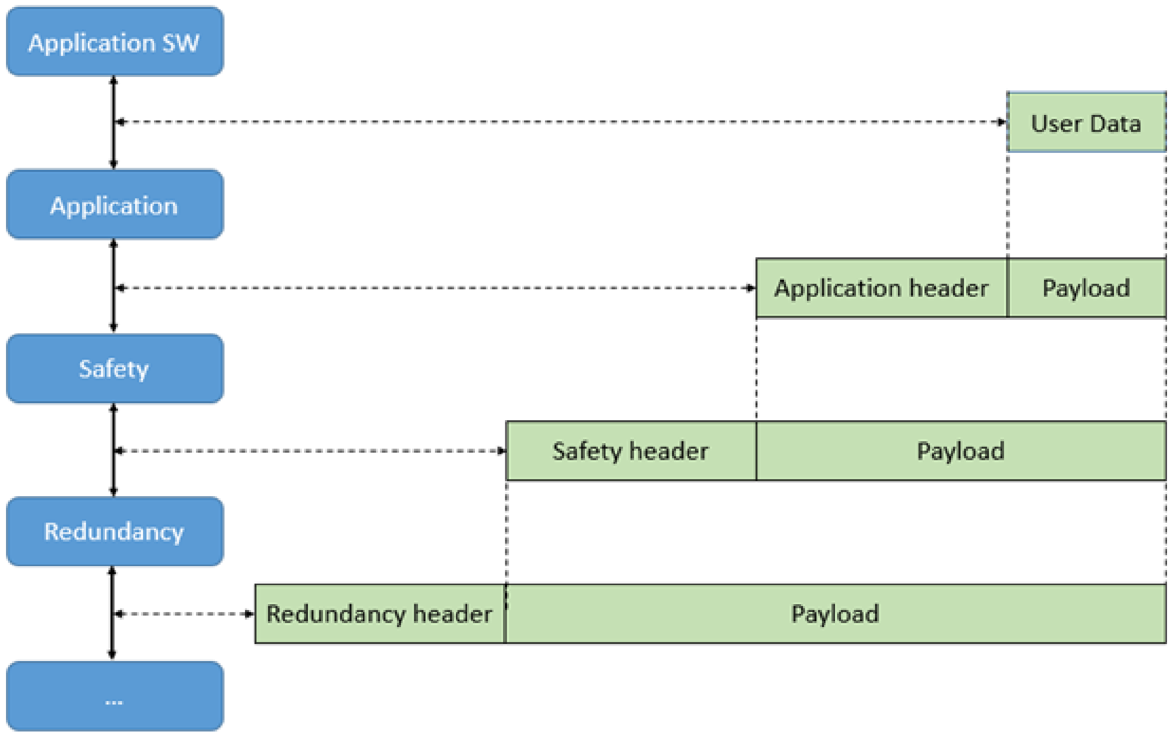


Figure 7.4 Header/payload structure of a layered communications protocol.

Typical examples of the fields in the safety header are:

- Timestamp of the messages in order to manage the intrinsic latency of the communication process.
- Sequence number in order to control message flow.

- CRC of the message to detect possible corruption of the message content.

Increasing the number of layers and of the size of information in headers/payload has an impact in the performance requirements for the communication equipment and the transmission media (network). This impact is not only on the required bandwidth but also in the computing capabilities of both sizes of the interface that need to manage all layer related functionality (timestamping, CRC, encryption, etc.). Consequently, demanding requirements to the communication protocol could imply restrictions to the type of technology used as transmission media and to the power supply for the communication equipment. In the context of Assets4Rail where some of the monitored equipment will be in remote locations with limited access to power supply, this effect could become a critical issue. For example, implementing a complex protocol could imply discarding long range/low power consumption systems such as, for example, Lora, which is in principle a good choice for remote track size equipment. Therefore, it is recommended to carefully evaluate the RAMS requirements for communication equipment/protocols in the Assest4Rail Project.

Once the main characteristics of the communication protocol are agreed, there are at least three approaches to choosing a protocol specification/implementation:

- Proprietary/commercial protocols. Signalling suppliers have a wide range of proprietary protocols for with different requirements and for a wide variety of applications in the signalling field. These protocols have been demonstrated to be reliable, efficient and robust by their use in a significant amount of projects and environments. However, using this option for Assest4Rail may proof difficult. Even after successful election of one or several protocols, the owner of the protocol should need to agree opening the protocol to other suppliers. The other suppliers should also agree to implement a protocol whose specification is owned and maintained by another company. These difficulties seem to rule out this approach.
- Specific Assets4Rail protocol(s). This approach is based in specifying an own Assests4Rail set of protocols. In case that compliance to EN 50159 is required, also justification of compliance must also be provided. This activity requires a significant amount of effort, which is probably out of the scope/budget of the Assets4Rail project.
- Open safety protocols for the railway environment. This is probably the most feasible approach. A number of protocols with open specifications are available which do not imply any additional costs for their implementation (in terms of rights of use). For example, ERTMS specifies a safe protocol for trackside/on-board communication whose lower layers are already implemented by a significant amount of signalling providers. Its use in the Assest4Rail context would only require the definition/implementation of the specific application data. For trackside/trackside communication, the protocols specified in the context of the Eulynx project based on RasTa may also become a good choice. Eulynx is a european infrastructure managers initiative that aims at standardization of all interfaces around interlockings.

7.3.2 Data format

At this stage of the Assets4Rail project, the nature and amount of the data to be transmitted is not known. Thus, it is not possible to define the exact data structures to be used for each communication pair. Only general high-level characteristics can be outlined.

Based on usual industry practice, we assume a basic architecture in which a centralised equipment is the consumer of the data generated by field equipment (trackside and on-board). It could also be necessary to use intermediate equipment similar to object controllers in the context of interlockings. Assuming that the performance of the transmission system at field level could be relatively low, the intermediate equipment would collect data from field equipment in its area and retransmit this information to the centralised equipment using a network with higher capacity. This preliminary analysis focusses on the logical interface between the centralised and the field equipment without taking into account the final architecture of the system. Three types of basic messages are considered:

- Periodic status data. The field equipment sends status data to the centralised equipment periodically. Each type of field equipment would transmit all of its specific data. The required periodicity depends on the type of field equipment and the related functionality. Commonly, this kind of periodical transmission is also used as *lifesign* of the application software at field level.
- Event driven status data. Field equipment sends real time information of relevant events whenever they are detected. For example, if a command to move a point fails, it might be necessary for the monitoring equipment to immediately sned this information to the centralised equipment without waiting for the periodical status transmission.
- Request for information. The centralised equipment sends requests to field equipment to transmit its status data whenever they are required. This represents a complementary approach to the periodical status transmission model. On one hand, the *request for information* model requires bidirectional application layer communication (contrary to a pure periodical + event transmission model). On the other hand, processing/communications load of the centralised equipment can be more easily controlled since the centralised equipment would only request information when it is certain that it can manage it.
- Regarding the high-level structure of the transmitted data, common present day practice is to use a header + information packets structure. In the railway context, this structure is used for example in the specification of the ETCS application data. Such structure provides high scalability in terms of possible future additions of new types of field equipment and/or new types of user data. Usually, both headers and packets are defined using a common set of variables with defined bit lengths.
- Headers contain information about the type of equipment transmitting the data, identification of the individual transmitting equipment and about the type of the transmitted message.
- Information packets contain the relevant user data transmitted by the field equipment. Normally, one type of packet is used only for one type of information. For example, data from a signal could be defined in at least two packets: a measured current packet for periodical/requested status messages and failure information packet to inform in real time of failure to show a determined aspect. A message can contain none, one or more packets depending on the nature of the transmitting equipment.

7.4 Data transmission

One of the most important aspects of a massive data collection process is its transport. In a railway environment where the scope is to collect data from signalling components for its monitoring, very different sources can be found. In general, they will be treated as sensors for the acquisition of essential parameters. These sensors will be located all around the trackside or even on-board the rolling stock. In spite of its possible differences the set of sensors will share some basic characteristics, these are:

- Power consumption constraints for electronic devices using batteries or energy harvesting;
- Some mobility of nodes;

- Heterogeneity;
- Small dimensions;
- Harsh environmental conditions;
- Low data to transmit.

These characteristics are coupled with the fact that the sensors can be located along the entire railway environment, so they can be spaced up to the order of kilometres, so the solution must be scalable. There can be several solutions for the communication and transmission of this data, but given the characteristics previously mentioned the most indicated could be to establish a wireless sensor network.

Wireless sensor network (WSN) refers to a group of spatially dispersed and dedicated sensors for monitoring and recording the physical conditions of the environment and organizing the collected data at a central location.

The WSN is built of sensor nodes, and there are several wireless standards and solutions for its connectivity. A wireless solution suitable is a low-power wide-area network (LPWAN) which is a type of wireless telecommunication wide area network designed to allow long-range communications at a low bit rate among things such as sensors operated on a battery. The LPWAN data rate ranges from 0.3 kbit/s to 50 kbit/s per channel (Figure 7.5). This kind of network includes different wireless technologies like Sigfox, LoRa, NB-IoT or Cat-M1 (also known as LTE-M).

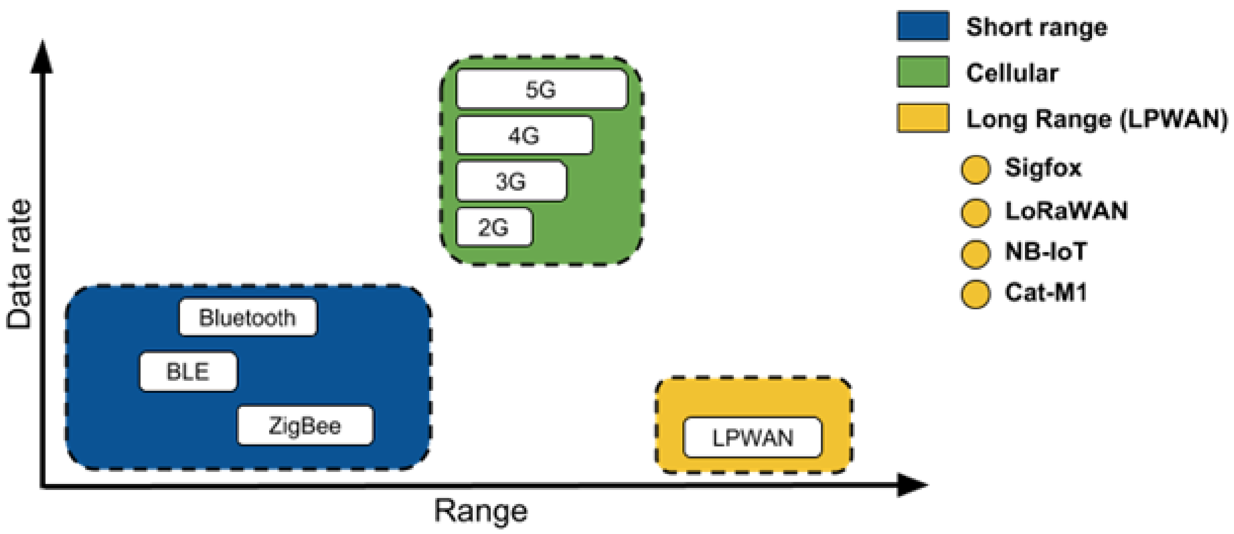


Figure 7.5 The range of LPWAN data rate.

Monitoring, as it was known traditionally reinvents itself to satisfy security and quality needs and policies productive more demanding. The modern processes of monitoring should include not only the ability to receive data about critical points of the process or chain of services, but also the possibility of predicting the possible problems that they could put in danger the operation of a system or even the physical integrity of the people.

In the railway, there are two differentiated components:

- On-board systems;
- Trackside systems.

Onboard systems include all kind of devices or sensors installed in the train, and the trackside systems are set in the track.

The network that they will create between them must be analyzed differently.

7.4.1 On-Board systems

The large number of manufactures of equipment for trains, each with its own wiring and its protocol, and the greater use of international trains, formed by railway cars from different countries, generates a problem of compatibility in the communication between the onboard equipment.

7.4.1.1 TCN

Railway operators and equipment manufacturers have specified a communications network called TCN (Train Communication Network), whose main objective is to allow communication between devices of different manufacturers installed in the same or different train cars. This network is composed of a vehicle bus that interconnect locally the equipment boarded in a vehicle, and by a train bus, which connects the vehicles of the bus.

The Train Communication Network is a hierarchical combination of two fieldbus for data transmission within trains. It consists of the Multifunction Vehicle Bus (MVB) inside each vehicle and of the Wire Train Bus (WTB) to connect the different vehicles. The TCN components have been standardized in IEC 61375.

This kind of network could be different in the locomotives and the coaches.

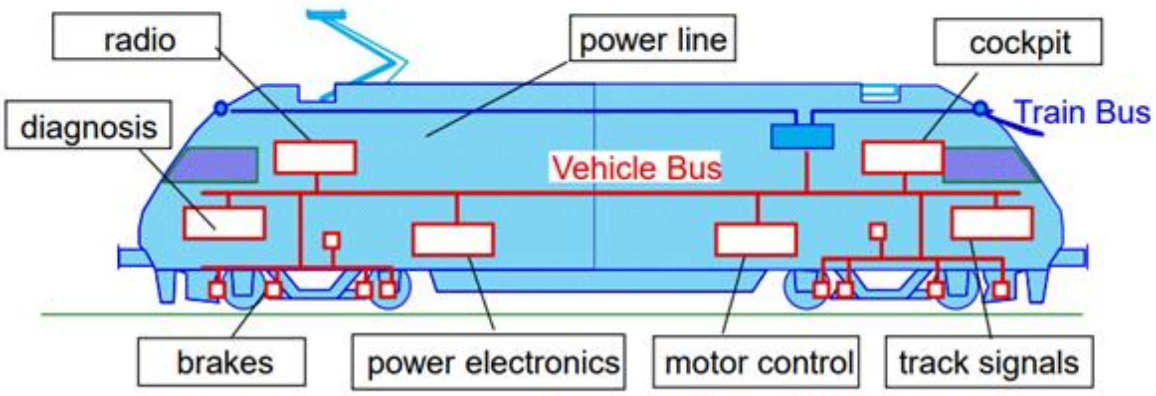


Figure 7.6 MVB in locomotives.

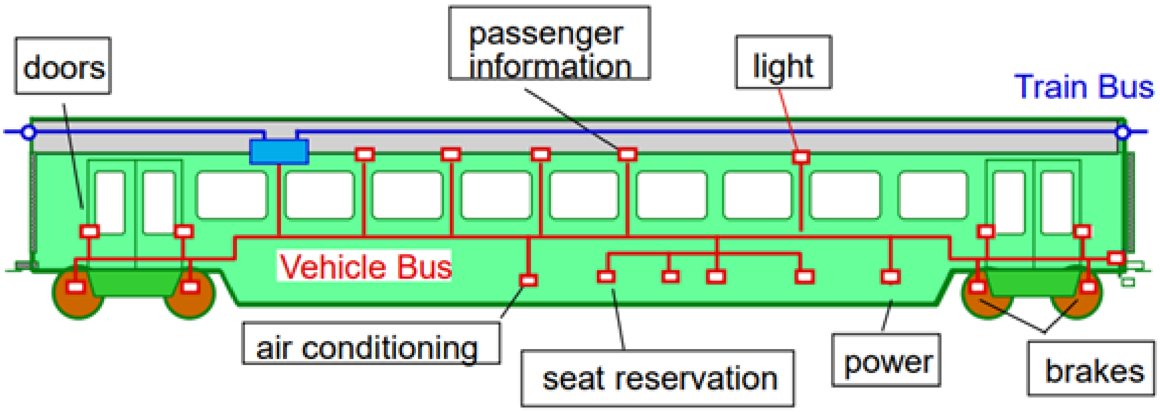


Figure 7.7 MVB in coaches.

This protocol has into account three different physical layers:

- ESD: Electrical Short Distance. Interconnects devices over short distances (20m) without galvanic separation. Based on proven RS-485 technology (Profibus)
- EMD: Electrical Medium Distance. Connects up to 32 devices over distances of 200m;
- OGF: Optical Glass Fiber. Covers up to 2000m.

Advantages of this type of network:

- Modular and distributed system;
- Flexible system;
- Interoperability.

Disadvantages:

- Low capacity of data transmission. Between 1 – 2 Mbps.

7.4.1.2 LAN NETWORK

A Local Area Network (LAN) is a group of computers and associated devices that share a common communications line or wireless link to a server. Ethernet and Wi-Fi are the two primary ways to enable LAN connections (Figure 7.8). Ethernet is a specification that enables computers to communicate with each other. The rise of virtualization has fueled the development of virtual LANs (VLAN), which allows network administrators to logically group network nodes and partition their networks without the need for major infrastructures changes.

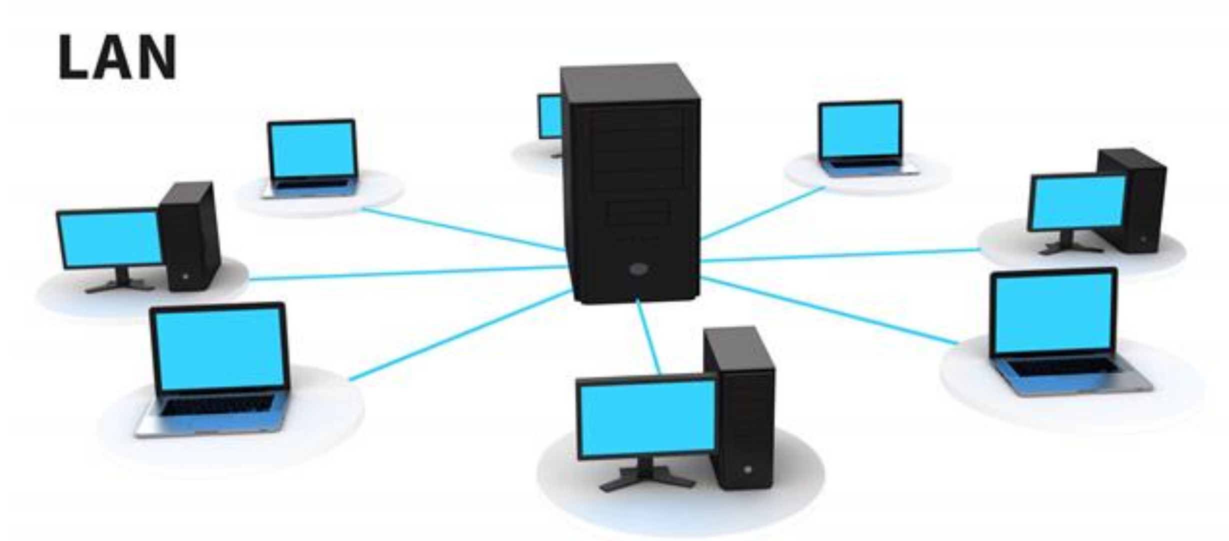


Figure 7.8 Local Area Network (LAN).

Main characteristics:

- Broadcast technology with the shared transmission medium.
- Transmission capacity between 1Mbps and 1Gbps.
- Maximum length not exceeding 3 Km.
- Use of a private media.
- The simplicity of the transmission medium:
    - Coaxial cable.
    - Telephone cables
    - Fiber optic.
- Wide variety and number of connected devices.
- Possibility of connection with other networks.
- Ease to change software and hardware.

Two basic categories of network topologies exist, physical and logical topologies.

- The transmission medium layout used to link devices is the physical topology of the network. For conductive or fiber optical mediums, this refers to the layout of cabling, the locations of nodes, and the links between the nodes and the cabling. The physical topology of a network is determined by the capabilities of the network access devices and media, the level of control or fault tolerance desired, and the cost associated with cabling or telecommunication circuits.
- Logical topology is the way that the signals act on the network media, or the way that the data passes through the network from one device to the next without regard to the physical interconnection of the devices. A network's logical topology is not necessarily the same as its physical topology.

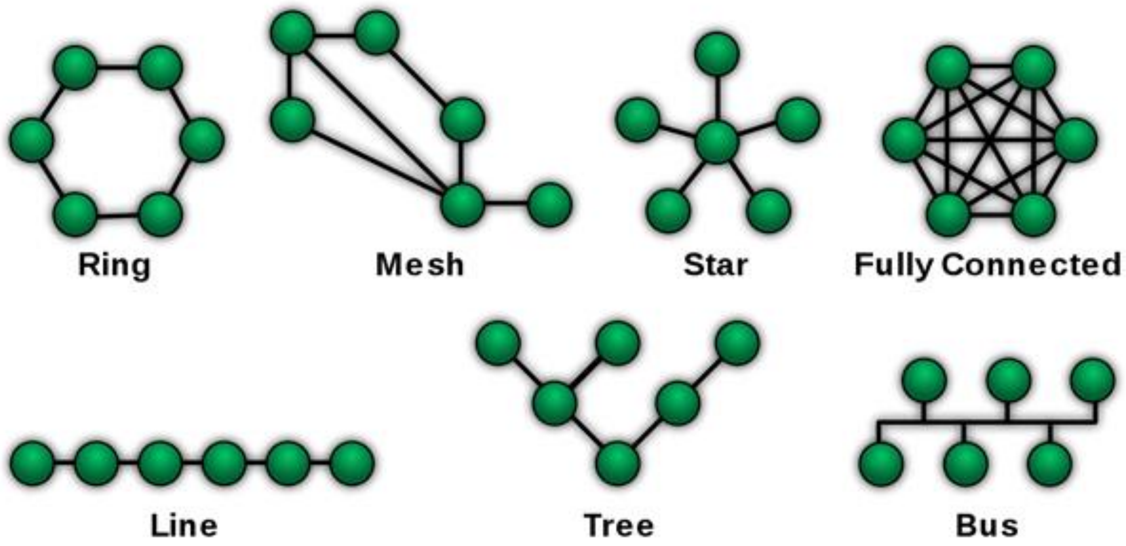


Figure 7.9 Different network topologies.

7.4.2 Trackside systems

Rail transport, due to its multiple activities and its operational and management complexity, needs to exchange a large volume of information.

In order to record all the data of the trackside devices, it will be necessary a network between all of this equipment.

7.4.2.1 GSM-R

GSM-R (Global System for Mobile Communications – Railway) is an international wireless communications standard for railway communication and applications.

GSM-R is based on the cellular GSM technology, with further enhancements specific to the requirements of railroad operation, such as train control.

In general, GSM-R uses characteristics that are identical or similar to those of the GSM system, such as frequency spacing (200 kHz), modulation (Gaussian minimum shift keying, GMSK) and access type (TDMA, TDD/FDD). The frequencies of GSM-R are extended below the frequencies of the GSM-900 standard.

The general packet radio services (GPRS) for data communications up to 14.4 kbit/s is supported by GSM-R for data transport in the same way as with the regular GSM system. It serves customized applications such as automatic train control (ATC) and electronic train control system (ETCS) for remote or automatic control of train movements and monitoring as well as functional addressing (calling a user by an assigned function rather than by a fixed number) and location-dependent addressing.

Additional railway-specific characteristics include advanced speech call items (ASCI) such as voice group call service (VGCS) and voice broadcast service (VBS) to communicate to a group of handsets simultaneously. GSM-R mobile phones are similar to cellular mobile phones but are more robust. They also offer extended functionality and a different user interface to utilize the additional applications.

GSM-R permits new services and applications for mobile communications in several domains:

- Transmission of Long Line Public Address (LLPA) announcements to remote stations down the line;
- Control and protection (Automatic Train Control/ETCS) and ERTMS);
- Communication between train driver and regulation center;
- Communication of on-board working people;
- Information sending for ETCS;
- Communication between train stations, classification yard and rail tracks.

The new monitoring devices could adapt to the existing GSM-R network, to communicate to the central server for data storage, with a GSM interface. This is an advantage in order not to use fiber cable for long distances.

7.4.2.2 Fiber Optic

The fiber optic is the medium used to propagate the bits of information. According to the requirement of the rail service, two network topologies are deployed, one based on circuit switching and another on packet switching.

The first one, obtained through a SDH (Synchronous Digital Hierarchy) transmission network, offers the advantage of having a transmission in dedicated circuits with reserved capacity. The second one, a Data Network based on IP (Internet Protocol) and MPLS (Multiprotocol Label Switching) technology.

7.4.2.3 Circuit Network

The Synchronous Digital Hierarchy allows the establishment of dedicated circuits for point-to-point communications between two sites. They are networks with synchronism of high quality and without loss of information. When the circuit is established, there are no temporary losses in the calculation and decision making for the selection of the appropriate way in the intermediate nodes. Each of these has a single way for the incoming and outgoing packets of the specific session.

An external GPS antenna and high-performance synchronization equipment allow the generation of a 2 MHz clock signal that is transmitted through all the nodes of the SDH network. This signal is taken as a reference to maintain perfect synchronization in its transmissions and receptions. This is required by the services that uses it as a medium of transport.

Dedicated circuits are established with the inconvenience when a failure occurs in one of the intermediate nodes, all communications are broken down and is necessary to restore all connections from the beginning. However, as in a packet networks, connectivity is maintained using link protection mechanism with protocols such as SNCP (Subnetwork Connection Protection Ring).

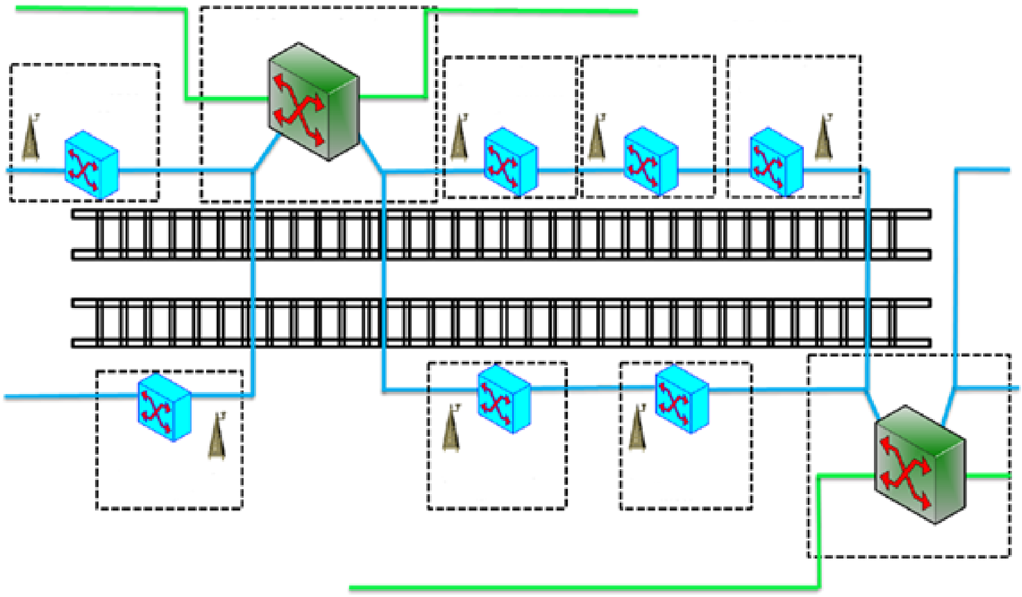


Figure 7.10 Circuit network.

7.4.2.4 Data Network

The lower level is the Data Access Network:

- Sufficient transmission capacity for the demands of required services (1Gbps);
- Depending on the configuration of the services, its operation is based on level 2 link protocols, on level 3 IP protocol, or including MPLS technology in the access;
- The top level is the Data Core Network;
- Bandwidth with interfaces in general 1 Gbps to support the aggregation of information from the access network;
- Supported by MPLS technology together with routing protocols.

The access nodes collect the information from the sites. Depending on the destination of the communication, they flow into the access ring or are transmitted to the node of the core network. This network, with greater switching capacity, is responsible for making connections over long distances and avoid overloading other access rings with traffic.

This schema represents the typical architecture of a railway data network. A Core Network that have the backbone through all the information passes, and set of access rings, of a lower level, whose mission is to give access to the network to all the locations of the line.

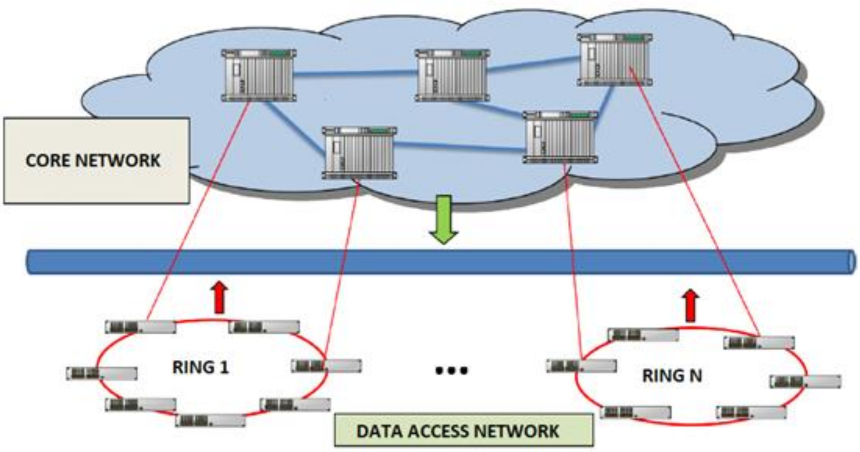


Figure 7.11 Data network.

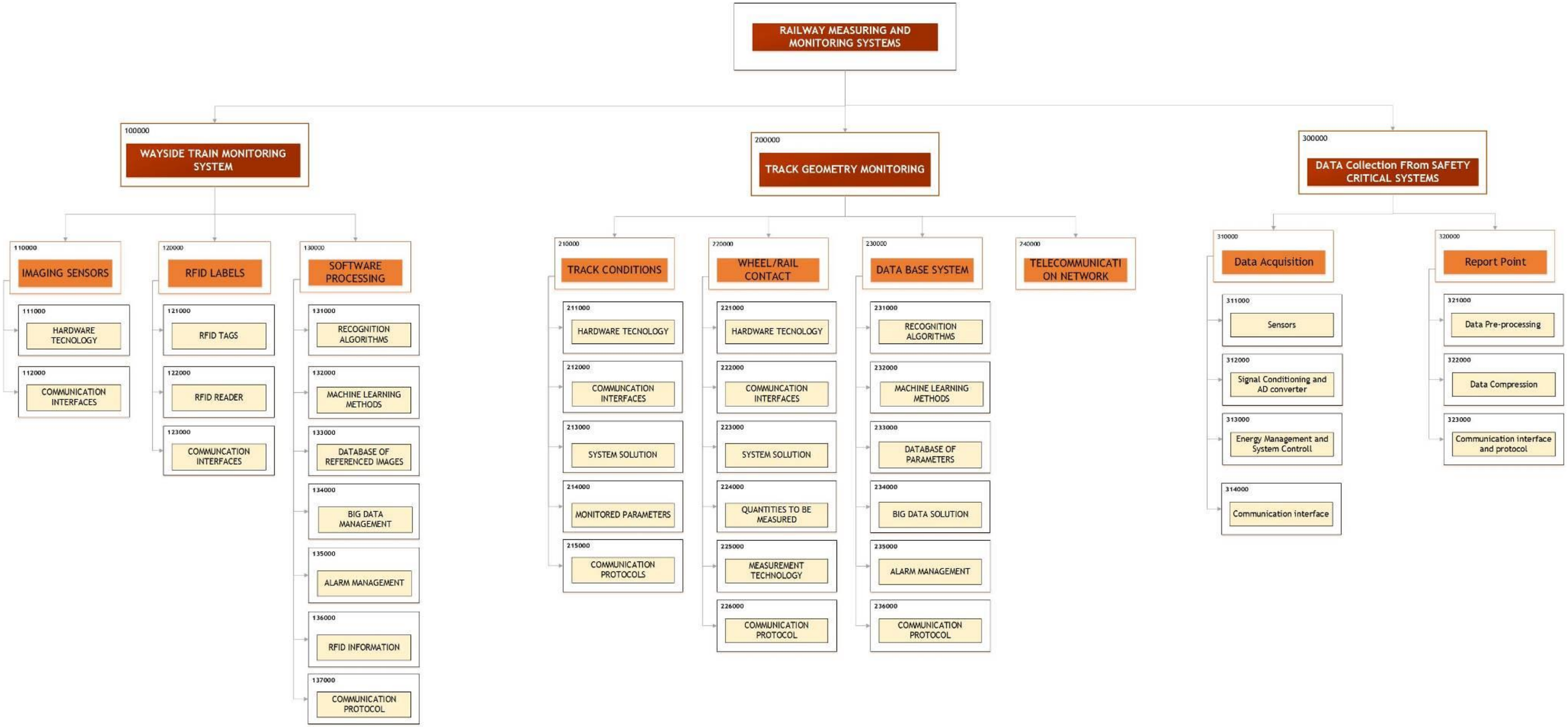
7.5 Conclusions

There are existing condition monitoring systems for the individual signalling systems. The sensing parameters for different signalling systems are different. Therefore, the data collection system to be developed should has a modular architecture, which allows easy adaptation of sensors for the specific application scenarios. In addition, there should be digital inputs for the computerized signalling systems, that can output the diagnosis-related information in a standard industrial protocol. Regarding data communication and transmission, the existing railway and industrial standards should be taken into consideration to ensure the interoperability. The existing communication networks are mostly used for safety-related purpose. For instance, MVB and WTB are used for train control and management. The data collection system could read the digital data provided by the existing network infrastructure, where interface and protocol are available. However, the data collection system should not affect the existing network and the monitored objects.

8. Specifications of requirements (Task 6.2)

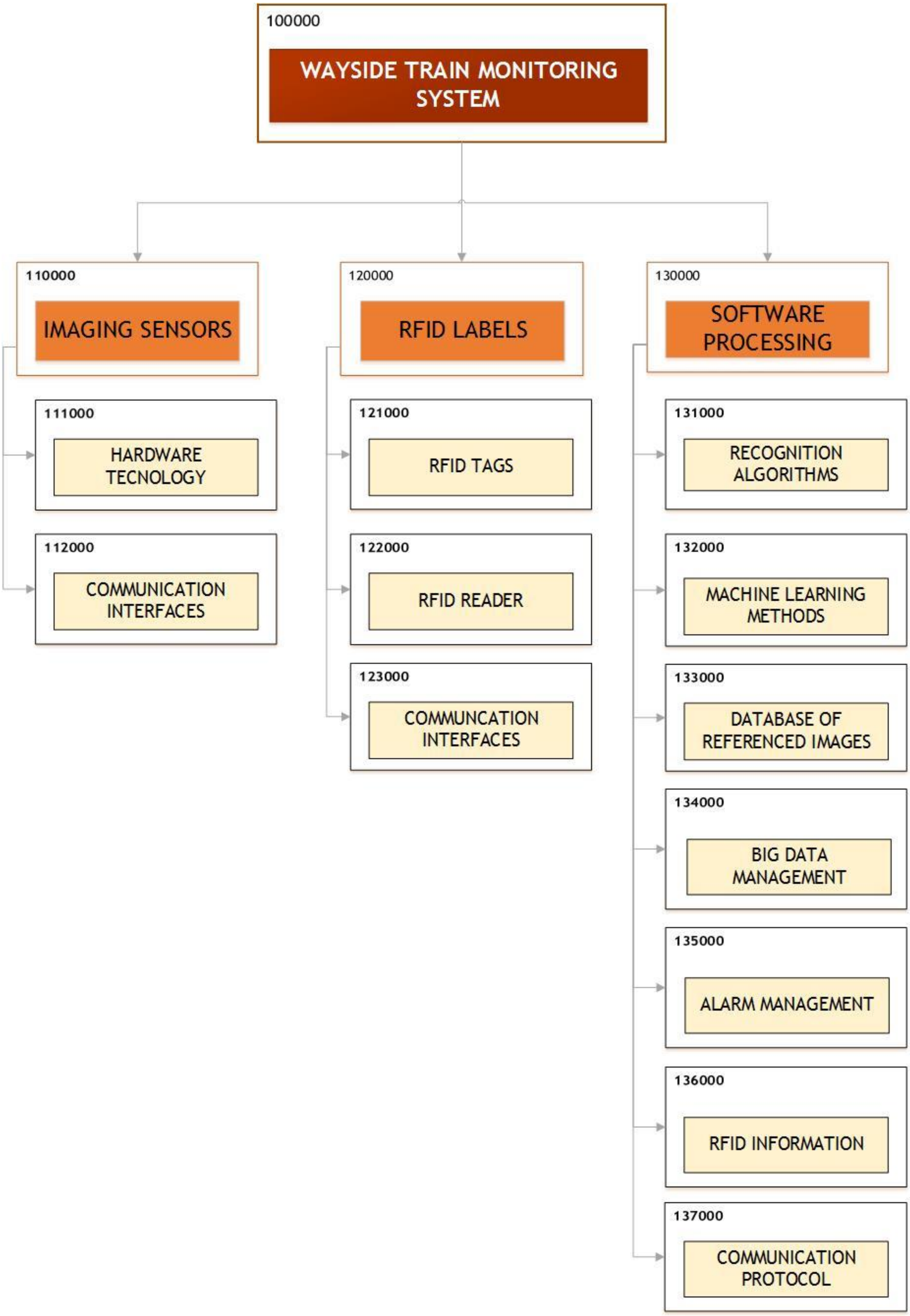
8.1 System Breakdown Structure and High Level Requirements

The objective of this document is to develop a System Breakdown Structure (SBS) considering the candidate technologies identified in Task 6.1. These requirements will be classified into functional, operational, performance and safety and will be the basis for WP7, WP8 and WP9, which will serve as a validation for them.



RAILWAY MEASURING AND MONITORING SYSTEMS
100000
WAYSIDE TRAIN MONITORING SYSTEM
110000
IMAGING SENSORS
111000
HARDWARE TECNOLOGY
112000
COMMUNICATION INTERFACES
120000
RFID LABELS
121000
RFID TAGS
122000
RFID READER
123000
COMMUNCATION INTERFACES
130000
SOFTWARE PROCESSING
131000
RECOGNITION ALGORITHMS
132000
MACHINE LEARNING METHODS
133000
DATABASE OF REFERENCED IMAGES
134000
BIG DATA MANAGEMENT
135000
ALARM MANAGEMENT
136000
RFID INFORMATION
137000
COMMUNICATION PROTOCOL
200000
TRACK GEOMETRY MONITORING
210000
TRACK CONDITIONS
211000
HARDWARE TECNOLOGY
212000
COMMUNCATION INTERFACES
213000
SYSTEM SOLUTION
214000
MONITORED PARAMETERS
215000
COMMUNICATION PROTOCOLS
220000
WHEEL/RAIL CONTACT
221000
HARDWARE TECNOLOGY
222000
COMMUNCATION INTERFACES
223000
SYSTEM SOLUTION
224000
QUANTITIES TO BE MEASURED
225000
MEASUREMENT TECHNOLOGY
226000
COMMUNICATION PROTOCOL
230000
DATA BASE SYSTEM
231000
RECOGNITION ALGORITHMS
232000
MACHINE LEARNING METHODS
233000
DATABASE OF PARAMETERS
234000
BIG DATA SOLUTION
235000
ALARM MANAGEMENT
236000
COMMUNICATION PROTOCOL
240000
TELECOMMUNICATI ON NETWORK
300000
DATA Collection FRom SAFETY CRITICAL SYSTEMS
310000
Data Acquisition
311000
Sensors
312000
Signal Conditioning and AD converter
313000
Energy Management and System Controll
314000
Communication interface
320000
Report Point
321000
Data Pre-processing
322000
Data Compression
323000
Communication interface and protocol

8.1.1 Wayside Train Monitoring System



8.1.1.1 Imaging sensors

The train monitoring sensors are splitted in some modules.

HARDWARE TECHNOLOGY

- The train monitoring system shall be composed of:
- Multiple sets of 2D cameras. This array of 2D cameras uses high power illumination resources, high power processing unit and high level IP housing;
- Stereo Line – Scan. This stereo-line camera uses high power illumination resources, high power processing unit and high level IP housing. This camera will be used to obtain 3D images of the wayside of the train;
- Synchronization system. This system will detect passing trains to synchronize the acquisition of the WTMS;
- Database processing system. This system will receive and incorporate the information from every part of the WTMS.

COMMUNICATION INTERFACES

The communication interfaces are the physical ways that the sensor has to communicate with the network or with other devices.

- RS.232;
- Coaxial cable;
- Camera Link

COMMUNICATION PROTOCOL

The communication protocol is the protocol used to communicate with the server or the database. It could be a safe protocol based in the UNE 50159 or a public protocol like TCP.

The protocol has to define all the layers.

8.1.1.2 RFID labels

A RFID system basically consists of the RFID tag and the RFID reader with the antenna. For specific applications, the RFID system should have middleware and information system. In the application of WTMS, the RFID system has a communication interface with image monitoring systems.

RFID technology can work in different frequency bands, i.e. Low Frequency Band (LF), High Frequency (HF), Ultra High Frequency (UHF) and Microwaves. Due to the required communication range, UHF RFID systems are commonly used in rail.

RFID TAG

According to EPCglobal, tags can be categorised into following classes, depending on their functionality.

- EPC Class 0 Tags – Generation 1, factory pre-programmed read-only passive tag
- EPC Class 1 Tags – Generation 1 and 2, Read-only passive tag similar to class 0 and has one-time field programmability
- EPC Class 2 Tags / EPC Class 1 Gen 2 Tags – EPC Class 2 tags are enhanced Gen 2 Class 1 tags, which are passive tags with read-write capability, working in UHF area. It is the standard tag type used in railway applications according to GS1 standards. All Gen 2 tags contain the same basic memory features, consisting of 96 bit EPC number support, 32-64 bit tag identifier (TID), 32 bit kill password to permanently disable the tag and 32 bit access password to lock the read write characteristics of the tag. Some tags include user memory which facilitates reading and writing of additional data in the tag.

- EPC Class 3 Tags – Semi-passive or active tag with read-write memory, on-board sensor and an incorporated battery to provide increased coverage; capable of recording parameters like temperature, pressure, and motion.
- EPC Class 4 Tags – Read-write active tag with integrated transmitter for communication using the battery on-board; can communicate with other tags and readers.
- EPC Class 5 Tags – Class 4 tags that provide additional circuitry to communicate with or provide power to other tags.

The commercial products of RFID tags are usually designed for specific applications in order to meet the specific requirements.

RFID READER

Readers can be stationary or mobile. They could have integrated PCB antenna or external antenna connections, allowing the user to achieve greater and diverse radiation patterns. The reader output power is the essential specification which affects reading performance.

Communication interface

Communication interface refers to physical interface and communication protocol. The standard interface and protocol should be considered at first.

8.1.1.3 Software processing

RECOGNITION ALGORITHM

- The WTMS must implement recognition algorithms in order to identify the prioritised rolling stock failures. The WTMS shall segment the point cloud and the images of the 2D array in the pre-selected rolling stock components of interest. These recognition algorithms will identify defects in the segmented components of the rolling stock.

MACHINE LEARNING METHOD

- The WTMS will implement machine learning methods in order to improve the defects recognition.

DATABASE OF REFERENCED IMAGES

- The wayside monitoring system will implement a database of referenced images.

BIG DATA SOLUTIONS

- The wayside monitoring system should manage large data from different sources and shall be able to extract knowledge from that data.

ALARM MANAGEMENT

The fundamental purpose of alarm annunciation is to alert the operator to deviations from normal operating conditions, i.e. abnormal operating situations. The ultimate objective is to prevent, or at least minimise, physical and economic loss through operator intervention in response to the condition that was alarmed. For most digital control system users, losses can result from situations that threaten environmental safety, personnel safety, equipment integrity, economy of operation, and product quality control as well as plant throughput. A key factor in operator response effectiveness is the speed and accuracy with which the operator can identify the alarms that require immediate action.

RFID INFORMATION

The information of vehicle identification obtained by the RFID system should be used to stamp the measurement data captured by the camera system.

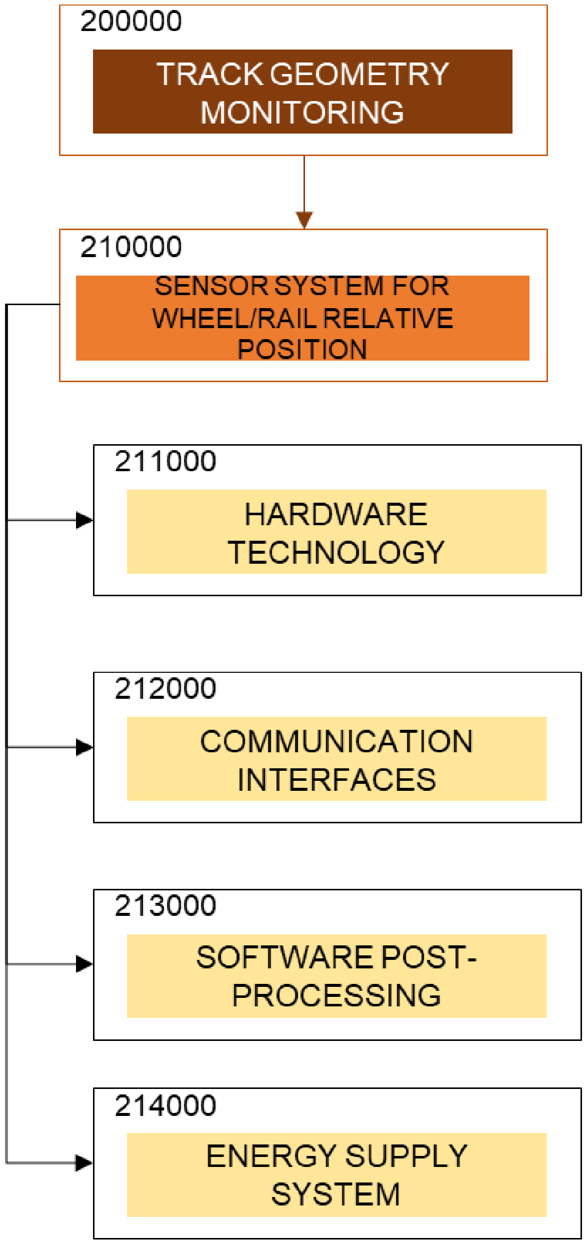
COMMUNICATION PROTOCOL

The communication protocol is the protocol used to communicate with the server or the database. It could be a safe protocol based in the UNE 50159 or a public protocol like TCP.

The protocol has to define all the layers.

8.1.2 Sensor System for Track Geometry Monitoring

In the field of measuring the geometric parameters of the track there are and are commonly applied many sensors suitable for detecting the track geometrical characteristics, applicable to vehicles in normal commercial service. While, for the location of the measured measurements, inertial platforms are used, which are generally quite expensive. Therefore, these systems lend themselves to use on special diagnostic trains, rather than to extensive use on the commercial fleet, which is composed of numerous vehicles. For this reason, this project intends to focus on systems that allow the detection of the wheel track position that can avoid the inertial platforms to localize the geometric parameter measures.



8.1.2.1 Wheel/rail relative position

The track conditions sensors are splitted in some modules.

HARDWARE TECHNOLOGY

The system for wheel rail contact position monitoring shall determine continuously the transversal position of the wheel in relation to the rail, must be integrated and synchronized with the other on-board systems for detecting the track geometry, should be able to be installed on different typologies of bogies. Moreover, the system should be able to work on in-service trains at 60-200 km/h, it should be able to be operated under a large set of weather and environmental conditions (dust, rain, snow, etc.), it must be able to withstand the stresses due to vibrations in the entire operation speed range. Finally, the system should be easily maintained, it must not interfere with the vehicle's gabarit and shall be powered by the low-voltage power supply system of the wagon.

In the following table, there are some kind of hardware for wheel rail contact measuring systems:

Instrument and parameter measured	Attributes, advantages	Disadvantages
Digital Image Correlation (DIC) of high speed filming: displacement	**+Can be used for any realistic speed of train** +Accurate at lower speeds where accelerometers, gyroscope and geophones tend to be less reliable	-Susceptible to vibration at the camera location (groundborne and wind), although methods to correct for this are available **-Line of sight may be problematic**
Laser based systems: displacement	As for DIC although differing processing methods may result in relative differences in accuracy.	
Multi depth deflectometer: displacement	LVDT (linear variable differential transformer) based system installed in shallow vertical borehole in trackbed +Will give an absolute measure with no zero shift and will in principle measure permanent settlements	**-Requires fixed datum at depth** **-Difficult or problematic to install**
Laser profile measurements	+ Several commercial systems available for wheel profiles + Several commercial embedded systems available to measure profiles	- Sensitive to environment (rain, dirt) - Measuring rail profile is only possible at short distance and placing outside the track will not give good precisions. Automatically moving laser scanner into the track area by robot arm or drones is unlikely to be realized and will be costly, but is mentioned as a possible solution
Distance sensor laser, time of flight	+ Can measure all distance + High precisions	- Measuring one point
Time of flight camera	+Take a whole picture	- Low resolution
Thermal Vision Technologies	Can be used without additional requirements for the lighting components	Additional software is required The technology is sensitive to ambient conditions that may influence in the temperature measures of the components
Stereo Visio Technologies	Complete imaging is available and higher resolution	Additional software is required The technology is sensitive to ambient conditions that may influence in the temperature measures of the components Additional requirements for light control and stability

COMMUNICATION INTERFACES

The communication interfaces are the physical ways that the sensor has to communicate with the network or with other devices.

- RS.232
- Optical Fiber.
- Coaxial cable.
- Bus.

COMMUNICATION PROTOCOL

The communication protocol is the protocol used to communicate with the server or the database. It could be a safe protocol based in the UNE 50159 or a public protocol like TCP.

The protocol has to define all the layers.

SOFTWARE POST PROCESSING

Data processing shall be based on algorithms for determining the transversal position of the wheelset in relation to the rail by determining the wheel/rail contact points and the angle of attack (yaw).

in particular, an algorithm will be necessary to correct the relative displacement of the wheelset with respect to the bogie frame on which the other sensors for track monitoring are mounted.

ENERGY SUPPLY SYSTEM

The energy supply system shall be constituted by dedicated batteries by powered both by means of sustainable and permanent technologies (e.g. harvesting systems using the vibration of bogies) and/or by the low-voltage power supply system of the wagon.

8.1.2.2 Database Systems

The Database system is the device that is going to recollect all the data from the sensors and manage the information in order to generate alarms or advices to improve the maintenance.

RECOGNITION ALGORITHM

Pattern recognition is the automated recognition of patterns and regularities in data. Pattern recognition is closely related to artificial intelligence and machine learning, together with applications such as data mining and knowledge discovery in databases (KDD), and is often used interchangeably with these terms. However, these are distinguished: machine learning is one approach to pattern recognition, while other approaches include hand-crafted (not learned) rules or heuristics; and pattern recognition is one approach to artificial intelligence, while other approaches include symbolic artificial intelligence.

MACHINE LEARNING METHOD

Machine learning methods are a special kind of object recognition methods that no rely on an algorithmic approach to solve the problem. They try instead to find data regularities based on inductive observation data.

Classification of Machine learning techniques:

- Linear Classifiers.
- Naive Bayes.
- K-Means.
- Support Vector Machine.
- Decision tree.
- Artificial Neural Networks (ANN).

- Deep Learning ANN (DNN).

DATABASE OF PARAMETERS AND IMAGES

Deep learning techniques have showed promising results on railway defect inspection but at the expense of private datasets. However, it is possible to take advantage of specific properties of DNNs and apply them to this problem:

- Fine-tuning.
- Transfer learning.
- Domain adaptation.

BIG DATA SOLUTIONS

The datasets have data from different sources that can be structured (databases) but also unstructured (video, image), which requires more computational power to extracting knowledge from them. The challenges are:

- Heterogeneity.
- Scalability.
- Complexity.
- Accuracy.
- Storing/sharing/publishing.
- Security
- Retrieve/reuse/discover.

ALARM MANAGEMENT

The fundamental purpose of alarm annunciation is to alert the operator to deviations from normal operating conditions, i.e. abnormal operating situations. The ultimate objective is to prevent, or at least minimise, physical and economic loss through operator intervention in response to the condition that was alarmed. For most digital control system users, losses can result from situations that threaten environmental safety, personnel safety, equipment integrity, economy of operation, and product quality control as well as plant throughput. A key factor in operator response effectiveness is the speed and accuracy with which the operator can identify the alarms that require immediate action.

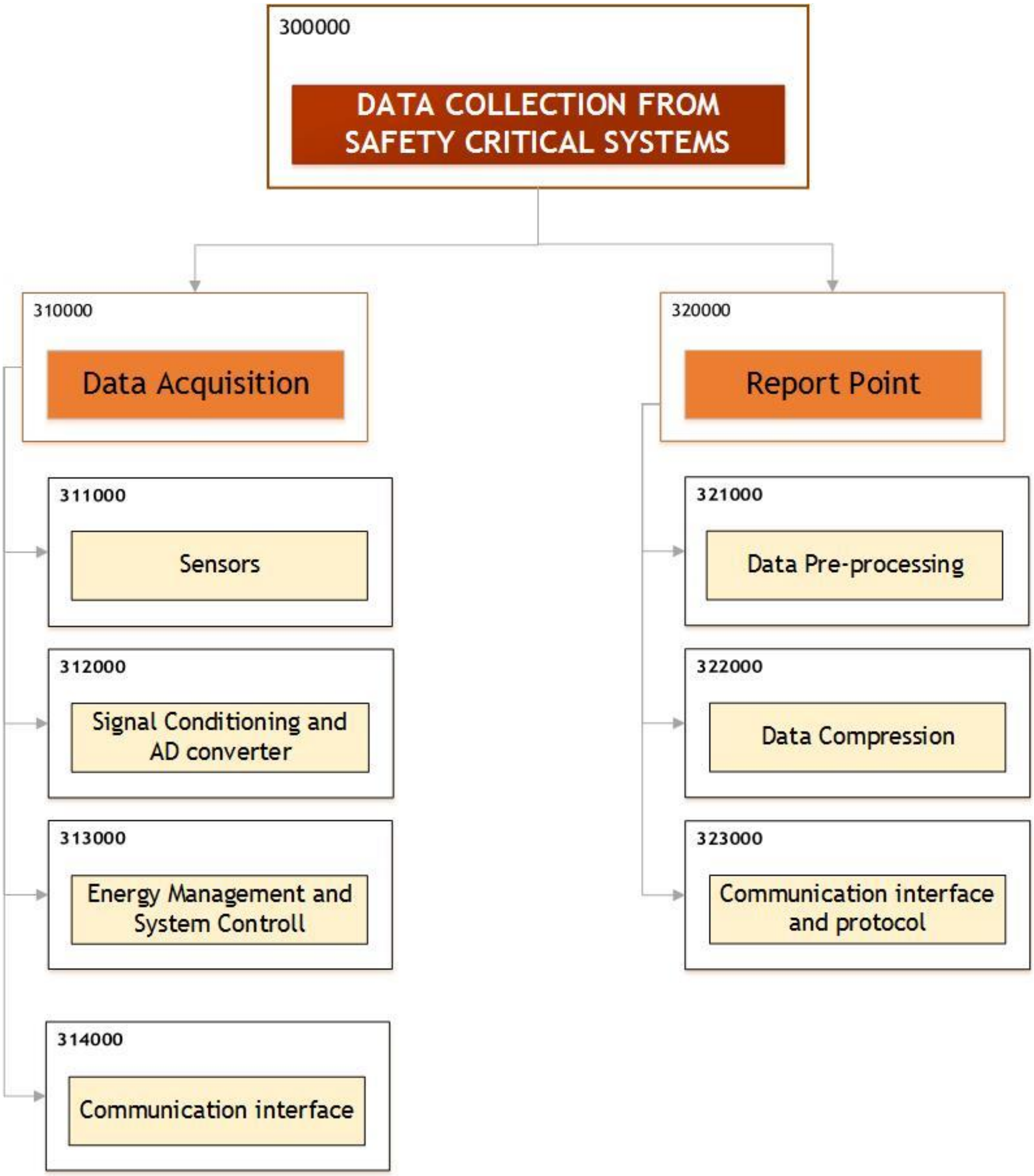
COMMUNICATION PROTOCOL

The communication protocol is the protocol used to communicate with the server or the database. It could be a safe protocol based in the UNE 50159 or a public protocol like TCP.

The protocol has to define all the layers.

8.1.3 Data collection from fail-safe Systems

The data collection system is to be a middleware, incl. hardware and software between sensors and a data platform for asset management.



8.1.3.1 Data Acquisition

Sensors

The data collection allows connection to various sensors for sensing essential physical parameters. The sensors could be either digital sensors or analog ones. The sensor type depends on application scenarios. In case of switch monitoring, sensors could be vibration, temperature, displacement, voltage and current sensors.

Signal Conditioning and AD conversion

For the applications of analog sensors, signal conditioning and Analog-to-Digital (AD) conversion is required. Signal conditioning includes voltage limiting (e.g. -10 - +10 V) and anti-aliasing filtering. Afterwards, AD conversion is performed to obtain the digital data.

Energy Management and System Control

The central control unit manages the whole device and also optimizes the energy consumption, which is in particular necessary when grid power is not available.

Communication Interfaces

The communication interfaces are the analog and digital input interfaces to a wide range of standard industrial sensors and equipment and offer numbers of input channels, such as 12 or 16-bit analog inputs, 8-channel modules, sampling rates up to 50 kHz. The physical interface for input should be either wired or wireless communication interface such as industrial buses Ethernet, RS232, RS485, Bluetooth, etc.

8.1.3.2 Report point

Data Pre-processing

Data Pre-processing includes intermediate storage of raw data, processing of raw data to the key information and storage of the key information.

Data Compression

The data should be compressed to save the storage space and the transmission volume. The common solution is to store the data in a binary format.

Communication interface and protocol

The standard industrial protocol should be considered such as IEC 61850. The physical interface for output should be either wired or wireless communication interface such as industrial buses Ethernet, RS232, RS485, LTE, etc.

8.1.3.3 Database System

DATA PROCESSING

Data processing is the conversion of data into usable and desired form. This conversion or “processing” is carried out using a predefined sequence of operations either manually or automatically. Most of the data processing is done by using computers and thus done automatically. The output or “processed” data can be obtained in different forms like image, graph, table, vector file, audio, charts or any other desired format depending on the software or method of data processing used.

DATABASE

Database is a systematic collection of data. Databases support storage and manipulation of data.

The database management system (DBMS) is the software that interacts with end users, applications, and the database itself to capture and analyze the data.

Existing DBMSs provide various functions that allow management of a database and its data which can be classified into four main functional groups:

- Data definition – Creation, modification and removal of definitions that define the organization of the data.
- Update – Insertion, modification, and deletion of the actual data.
- Retrieval – Providing information in a form directly usable or for further processing by other applications. The retrieved data may be made available in a form basically the same as it is stored in the database or in a new form obtained by altering or combining existing data from the database.

- Administration – Registering and monitoring users, enforcing data security, monitoring performance, maintaining data integrity, dealing with concurrency control, and recovering information that has been corrupted by some event such as an unexpected system failure.

ALARM MANAGEMENT

The fundamental purpose of alarm annunciation is to alert the operator to deviations from normal operating conditions, i.e. abnormal operating situations. The ultimate objective is to prevent, or at least minimise, physical and economic loss through operator intervention in response to the condition that was alarmed. For most digital control system users, losses can result from situations that threaten environmental safety, personnel safety, equipment integrity, economy of operation, and product quality control as well as plant throughput. A key factor in operator response effectiveness is the speed and accuracy with which the operator can identify the alarms that require immediate action.

COMMUNICATION PROTOCOL

The communication protocol is the protocol used to communicate with the server or the database. It could be a safe protocol based in the UNE 50159 or a public protocol like TCP.

The protocol has to define all the layers.

8.2 Requirements Management Methodology

The key elements on the Assets4Rail approach to requirements development are summarised below.

All the technical members and project managers of the different workpackages within Work Stream Two (WS2) will carry out four stages of requirements development, throughout the Assests4Rail project lifecycle:

- Requirements Elicitation;
- Requirements Analysis;
- Recording the requirements (related documents and register);
- Requirements Validation (within the Validation Workpackages).

Requirements Elicitation is the process of capturing all the data needed in order to be able to produce a complete set of requirements.

Requirements Analysis is the process of reviewing all the data gathered at the elicitation stage.

Recording the Requirements

- Requirements will be recorded directly into a spreadsheet register by the team and by nominated requirements authors within the Assets4Rail project.
- The requirements will be exported from the spreadsheet into documents at appropriate stages through the lifecycle of the project to support the design, development and validation stages.

Requirements Validation will be an iterative process which takes place throughout the Verification and Valisation workpackages (namely WP10 for WS2).

High quality requirements reduce the likelihood of errors, gaps and conflicts in what is to be delivered, hence will help to reduce project errors or overspends.

The language used when writing requirements is extremely important.

A requirement must be unambiguous and “SMART” i.e. Specific, Measurable, Attainable, Realistic and Traceable.

A consistent style helps to improve the appearance of the printed document and assists automated processing of requirements documents.

To apply a consistent approach for requirements management for the Assets4Rail project, there will be three separate modules for each WP:

- WP7 Wayside Train Monitoring;
- WP8 Track Geometry Monitoring;
- WP9 Data Collection for Signalling Systems.

Within the different Requirements Teams in the WPs will establish a set of metrics to illustrate the quality and status of the WS2 Assets4Rail requirements. The statistics will be available as a minimum at WorkStream, WorkPackage, System and Subsystem Level.

The metrics will cover the full requirements life cycle, including Verification and Validation statistics, and include as a minimum:

- Total number of requirements;
- Requirements without an identified source (identifies potential over or under specified work);
- Requirements compliance status statistics;
- Requirements without agreed acceptance criteria;
- Requirements not apportioned to a lower level in the requirements structure (identifies potential gaps);
- Requirements pending an approved change;
- Requirements still to be finalised and approved;
- Requirements that do not meet requirement quality guidelines.

These metrics will be available once this deliverable is approved and will serve as a guideline for the evolution of the entire project, mainly WP7, WP8, WP9 and WP10.

8.2.1 Overview of Requirements Engineering

8.2.1.1 Introduction

The purpose of this section is to give guidance of the application of Requirements Engineering processes to the Assets4Rail Project.

Assets4Rail must demonstrate satisfactory completion of the project objectives to the Shift2Rail Programme Offices. This requires Assets4Rail to demonstrate that all relevant requirements have been elicited from the state-of-the-art reports, captured, analysed, documented and tracked through design to validation stage.

The term "Requirements" in engineering is most easily defined as: "A set of prioritised needs elicited from all the stakeholders that together cover the functionality required for the system or service to be developed and deployed."

Defining requirements is an iterative process that gives the Assets4Rail project participants the opportunity to explore, discuss, clarify, define and agree upon what the different Assets4Rail WS2 systems will be.

It is unrealistic to assume that once identified that the requirements will be a fixed target until the Assets4Rail programme is delivered. The requirements will change as issues arise and the Assets4Rail programme develops during its lifecycle. These changes must be managed properly, with a robust change control process.

Requirements can be categorised in different ways, the most common are:

1. Functional Requirements are a list of objectives that the Assets4Rail WS2 Workpackages must meet. They identify the individual tasks that the system will carry out in order to meet the project needs;

2. Non-Functional Requirements generally relate to the way in which the systems behave in its environment i.e. integrity, ease of use, reliability, maintainability, availability, scalability and so on.

When defining requirements, they must be defined so that they are consistent, feasible and verifiable. Requirements need to be written in a manner that does not address how the goal is going to be achieved but only the goal, i.e. the requirement must not define the solution. The requirement defines the WHAT not the HOW.

Assets4Rail requirements must as a minimum follow the 'SMART' principle:

- Specific – an effective requirement specifies one requirement only and is unique;
- Measurable – objectively verifiable (by test, analysis, inspection, or demonstration) to prove compliance; the requirement should be written in a manner where the means of verification is clearly understood;
- Attainable – an effective requirement is technically feasible, affordable, and fits within the project schedule and other constraints;
- Realistic – realisable given the known constraints and resources;
- Traceable – a requirement which is not traceable can lead to the design and construction leading away from.

A good requirement has the following characteristics:

- Necessary: Needed to deliver a need that is not covered by other requirements or standards.
- Unambiguous: Concise and clear, expressing objective facts leaving no doubt in the readers' mind as to what is required.
- Singular: Addresses only one need such that the requirement cannot be split into two or more requirements at the same level.
- Stand-alone: Must be understandable on its own when separated from other requirements and context in the document.

8.2.1.2 Requirements Engineering Processes Overview

There are two segments that run concurrently:

- Requirements Development;
- Requirements Management.

Requirements Development is carried out to produce and analyse the requirements, and comprises of four processes as described below:

- Requirements Elicitation – this is the process that captures the data needed in order to be able to produce a set of requirements.
- Requirements Analysis – this process involves reviewing all of the data gathered at the elicitation stage. Any conflicts must be resolved, including completeness and traceability prior to moving on.
- Requirements Documents – this is the translation of all the data gathering and interpretation of data into a set of coherent requirements documents.
- Requirements Validation – this requires that all the requirements identified in the requirements document accurately reflect all the Assets4Rail needs and are an accurate reflection of the different WS2 systems to be developed.

Requirements Management is the control function for requirements engineering activities. It is the control framework that governs the requirement development process. Requirements will need to be developed and updated during the life of the Assets4Rail programme.

- One of the goals of the requirements management process is to ensure that there are no inconsistencies between the requirements that are identified and the product/system that is developed.
- All requirements are fully, defined and prioritised between the work package leaders.

- A coherent and complete requirements document(s) is issued, agreed upon and kept up to date during the lifecycle of the Assets4Rail programme.
- Any changes to requirements during the lifecycle are reviewed, verified, approved, implemented, are fully tracked and traceable.

It is important to understand how the requirements are to be complied with. **Acceptance Criteria** are used to define conditions that must be met before the project is considered completed and the associated deliverables can be accepted

The Acceptance Criterion should be defined when developing the requirement, this will help to establish if the requirement is unambiguous and can be verified.

Having clearly defined Acceptance Criteria helps sets the expectations regarding the WP objectives as well as to measure the achievement towards the completion of each task.

The Acceptance Criteria must be defined for each requirement specified in the requirements document.

Requirements validation is an iterative process which should take place throughout the lifecycle of the Assets4Rail project. During elicitation and analysis, the data should be constantly questioned and clarified in order to check its validity. This will ensure that the requirement design that is produced is complete, consistent and ready for the formal requirements validation process.

During the formal requirements validation process, the aim is to ensure that the requirement document is complete, consistent, verifiable and traceable.

Requirements validation ensures that the specified requirements are complete, correct, feasible, necessary, prioritised, unambiguous and verifiable.

The role of the requirements management process is to oversee the requirement development process through to delivery and operation of the system to ensure that the requirements baseline is what is delivered as a result of the Assets4Rail project.

However, complete the requirements set is, it is only a snapshot at a given point in time during the Assets4Rail programme lifecycle. There are many reasons why requirements might change during the lifecycle. Even if some of the requested changes to requirements can be rejected, they still need to be analysed first to determine whether they are acceptable changes or not.

Assets4Rail WS2 will implement the following key steps to apply a robust requirements management process:

- Requirements Version Control:
  - Requirements are not static and neither is the project environment. Once the initial set of requirements are agreed these must be version controlled by having a base line to be applied by the Task owner.
  - A requirements baseline is an approved set of requirements agreed by all the relevant task participants at a given point of time during the lifecycle and is normally associated with a formally issued set of documentation.
- Requirements Change Control:
  - Change control in requirements management is the process by which any changes needed to the requirements within the applied baseline are managed. Whether the change effects the scope, is simple, complex or editorial in nature they still need to follow a change control process.
  - The Requirements management must ensure that all proposals for change to requirements are formally documented.
  - The Requirements management must ensure that each requirement change has an impact assessment carried out:
- Requirements Attributes:

  - Attributes define the specific characteristics that need to be tracked and monitored throughout the lifecycle of the Assets4Rail programme as part of the requirements engineering process.
  - The key requirements attribute that Assets4Rail programme will track are:
    - Requirement Identifier – the identification of the specific requirement this must be a unique.
    - Source of the Requirement – where the requirement originated.
    - Requirement Text – the actual requirement text, formatted in accordance with the guidance in Appendix J.
    - Acceptance Criteria – statement of the high level requirements verification or validation activities that will demonstrate that the requirement has been met i.e. to understand how the requirement is to be complied with.
    - Supporting Information – additional information that can be provided to enable the reader to understand unambiguously the need specified in the requirement text.
    - Requirement Maturity – candidate, obsolete, agreed etc.
    - Compliance Status – whether the requirements have been complied with at a given stage in the Assets4Rail project lifecycle.
    - Comments – it is always useful to have a place for comments
    - Requirement Owner – the role of the person that raised the requirement
    - Requirement Justification – the reason for the requirement to exist
    - WBS – the WBS to which the requirement will need to be allocated
- Requirements Traceability
  - By ensuring that each requirement can be traced to its origins, unnecessary development is avoided. It is also the only way to make sure that no requirements are dropped in error or that new ones creep in without approval or awareness of Assets4Rail members.
  - Bidirectional traceability can only be achieved if there is a defined requirements hierarchy in place
  - With a traceability matrix, requirements can be traced back to their source and also traced forward to the lower level requirement which defines the detail. Requirements are often developed at different layers; hence it is difficult to trace from one end to another.

8.2.2 Requirements Quality and Style

The purpose of this section is to provide guidelines for syntax, quality and consistency of style for use in developing requirements across Assets4Rail.

A few common mistakes in requirement writing can mean ineffective communications that can lead to large increases in the overall time money invested in the Assets4Rail programme.

High quality requirements reduce the likelihood of misunderstandings, gaps and conflicts in what is to be developed. A consistent style also helps to improve the appearance of the printed document and to assist automated processing of requirements documents.

What is a good requirement?

A good requirement states something that is necessary, verifiable, feasible and concise. Even if it is verifiable, attainable, and correctly written, if it is not necessary, then it is not a good requirement. To be verifiable, the requirement must state something that can be verified by examination, analysis, demonstration, inspection or test.

Requirement statements that are subjective, or that contain subjective words, such as may, are not good requirements as they cannot be verified. A good requirement is unambiguous.

When writing requirements, the author must consider the following:

- Is it necessary? If there is a doubt about the necessity of a requirement, then ask: “What is the worst thing that could happen if this requirement was not included?” If an answer cannot be found of any consequence, then the requirement is probably not needed;
- Can it be verified? As the requirement is developed, determine how it will be verified, write down and agree the Acceptance Criteria with the requirement owner. This step will ensure that the requirement is verifiable;
- Is it feasible? The requirement must be technically feasible and fit within budget, schedule, and other constraints;
- Is it concise? Each requirement should express a single need, be concise, and simple. It is important that the requirement is not misunderstood - it must be unambiguous. Simple sentences will most often suffice for a good requirement.

There are terms to be avoided and terms that must be used in a very specific manner. Authors need to understand the use of shall, may, will, and should:

- Requirements use shall;
- Optional requirement use may;
- Statements of fact or future events use will;
- Goals use should.

These are standard usage of these terms in government agencies and in general industry. All shall statement (requirements) must be verifiable, otherwise, compliance cannot be demonstrated. Terms such as are, is, was, and must do not belong in a requirement.

Requirements documents state WHAT is needed, not HOW it is to be provided. Unfortunately, this is a common mistake made by a lot of authors, most do not intend to state implementation; they simply do not know how to state the need correctly. To avoid stating implementation, ask WHY the requirement is needed.

8.2.2.1 Writing Requirements

To ensure consistency and readability throughout WS2 Assets4Rail requirements documents all requirements must be written in a standard style and format. An example of requirement format is:

The <subject> shall <verb><adverb><condition or constraint>

Words and phrases that must be avoided when writing requirements are:

- Words that join sentences together e.g. and, or, with, also. Use of these may lead to multiple requirement sentences;
- Let-outs e.g. ought to, consider, could, perhaps, probably, if, when, except, unless, always and although;
- Long rambling sentences which could lead to misinterpretation, confusion and error;
- Speculation, where generalisation words can create wish lists e.g. usually, generally, often, normally, typically;
- Indefinable terms, such as desirable qualities that are vague e.g. user-friendly, versatile, flexible, approximately, as possible, efficient, improved, high performance, modern;
- Use of subjective words such as efficient, sufficiently and relevant within a requirement is likely to mean that the requirement is not clearly defined, is ambiguous and may not be capable of being verified or validated with absolute certainty.

A list of potentially ambiguous and subjective words that should be avoided and only used with care can be found in the table below:

adequate	effective	maximised	safely
adequately	effectively	maximum	same
adverse	efficient	may	satisfactory
adversely	efficiently	min	should
allowable	equivalent	minimal	significant
appropriate	especially	minimise	significantly
appropriately	etc.	minimised	similar
approximate	excellent	minimum	smaller
approximately	fair	moreover	smallest
aspires	good	most	so as
aspiration	great	must	substantial
bad	her	nearly	substantially
badly	his	possibilities	sufficient
because	however	possibility	sufficiently
below	ideal	possible	suitable
between	ideally	possibly	suitably
can	large	quality	timely
cannot	larger	rapid	typical
clearly	largest	readily	typically
compatible	lesser	reasonable	user friendly
completely	like	relevant	whether
consider	likely	necessary	worse
could	low	needed	worsen
down to	lowest	nominally	
easily	max	normally	
easy	maximise	safe	

8.2.2.2 Requirements Style

Requirements must have correct English spellings and grammar (e.g. “authorisation” rather than “authorization”).

Requirement text should not contain lists as the requirements are expected to be singular.

The following guidelines must be applied when using tables:

- Tables must not contain requirements;
- Tables must not contain diagrams;
- Tables should not be large - maximum of two A4 size pages when printed;
- Large amounts of text in tables must be avoided;
- Tables should be used for presenting data, parameters, conditions or constraints to the previous requirement leading the table.

All abbreviations and definitions of terms used in requirements documents must meet the following criteria:

- Are defined in the requirements document glossary;
- Abbreviations are defined in full the first time it appears in the document: e.g. Radio Frequency Identification (RFID);
- Full stops must not be used in abbreviations. The only exceptions to this are “e.g.” and “i.e.”.

When capturing and developing requirements, there may be a value or other aspects of the requirement that is initially unknown or requires further confirmation.

As the requirement matures through development and reviews, these values will be confirmed. It is important to have a consistent way of identifying these values to enable the requirements to be tracked and managed.

The characters “TBC” (To Be Confirmed) should be used and must be capitalised and enclosed in brackets.

Requirements developed must meet minimum quality criteria to ensure a set of robust requirements are produced.

Characteristic	Description
Concise	The requirement is stated in language that is easy to read and only states what must be done.
Design Independent	The requirement must not contain implicit, or explicit, statements or assumptions about the nature of the design solution to satisfy the requirement.
Feasible	The stated requirement can be achieved by one or more developed system concepts within budget and timescales.
Necessary	The stated requirement is an essential capability, physical characteristic, or quality factor of the business need.
Singular (Atomic)	The sentence / paragraph cannot be split into two or more requirements.
Standalone	Each requirement is capable of standing alone without losing its meaning when separated from other requirements.
Traceable	The necessity and sufficiency of each requirement is demonstrated by it being linked to a parent requirement or source.
Unambiguous	Each requirement must be stated to have one and only one interpretation and must not leave a doubt in the reader’s mind as to what needs to be done.
Understandable	Terms and language used in the requirement description can be understood by all stakeholders who need to read it.
Verifiable	The stated requirement is not vague or general but is quantified in a manner that can be verified.

8.3 Requirements Register

[Excel Document Embedded]

9. Conclusions

This Deliverable has described the state of the art regarding drivers, benchmarks and emerging solutions of the topics, that will be developed within Assets4rail project:

- Wayside Train Monitoring System;
- Sensor system to support track geometry monitoring;
- Data collection for diagnostics from signalling systems.

In particular, the main findings regarding the "Wayside Train Monitoring System" are related to the following subjects:

- Track degradation models;
- Rolling stock failures and assessment methodologies (e.g. FMECA);
- Hardware products for computer vision systems (e.g. stereo matrix-scan);
- Object recognition algorithms (segmentation, representation, recognition);
- Database of referenced images;
- Big data solutions (collection, processing, storage, etc.);
- RFID technologies.

Concerning "Data collection for diagnostics from signalling systems" the main findings regard the subjects as follows:

- Relevant works in IN2SMART and IN2RAIL;
- Sensors and methods for track geometry measurements;
- Technologies for measuring track geometry parameters;
- Monitoring solutions for assessment of track conditions;
- Sensors for measuring wheel/rail contact characteristics.

In the field of "Sensor system to support track geometry monitoring" the main findings regard the following subjects:

- Relevant work in IN2SMART;
- Data collection/sensing solutions for diagnostics;
- Communication interfaces and data format;
- Data transmission (On-Board and Trackside systems).

Finally, this Deliverable shows a System Breakdown Structure (SBS) for the assignment of requirement specifications to the three systems to be developed, considering the candidate technologies identified in Task 6.1. These requirements are classified into functional, operational, performance and safety and will be the basis for WP7, WP8 and WP9, which will serve as a validation for them.

10. References

- AGARWAL, Vivek, 2001. Assessing the benefits of Auto-ID Technology in the Consumer Goods Industry. AUTO-ID LABS [online], [cit. 2014-12-10]
- Available from: http://www.intechopen.com/books/reliability-and-safety-in-railway/condition-monitoring-ofrailway-track-using-in-service-vehicle
- B. Malakar, B.K.Roy (2014), Survey of RFID Applications in Railway Industry, 2014 First International Conference on Automation, Control, Energy and Systems (ACES), DOI: 10.1109/ACES.2014.6807999
- BALOG, Michal a Lucia KNAPČÍKOVÁ. Smart techniques for technical conditions monitoring of railway wagons. In: Proceedings of the 3rd EAI International Conference on Management of Manufacturing Systems [online]. EAI, 2018, 2018, s. - [cit. 2019-01-13]. DOI: 10.4108/eai.6-11-2018.2279349. ISBN 978-1-63190-167-6. Dostupné z: http://eudl.eu/doi/10.4108/eai.6-11-2018.2279349
- Baluch, H. (2018). Improwement of the methods for planning railway superstructure repair. Railw. Transp. Ukr., vol. 4, no. 129, pp. 20–26
- BANKS, Jerry. RFID applied. Hoboken, N.J.: John Wiley, c2007. ISBN 978-047-1793-656
- BARTOLŠIC, Martin, 2012. RFID se prosazuje v různých oborech. Computerworld. 2012(19). ISSN 1210-9924
- BOTTANI, Eleonora a Antonio RIZZI, 2007. Economical assessment of the impact of RFID technology and EPC system on the fast-moving consumer goods supply chain. International Journal of Production Economics [online], [cit. 2014-12-10]
- Breiman, L. (1996a). Bagging predictors. Machine Learning, 24(2), 123–140
- Breiman, L. (1996b). Bias, variance, and arcing classifiers
- Burstow, M., de Podesta, M., and Pearce, J. (2011). "Understanding wheel/rail interaction with thermographic imaging, Proceedings of the 22nd IAVSD Symposium, August 14-19, paper No.105, 2011, Manchester
- Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. Mobile networks and applications, 19(2), 171-209.
- Chowdhury, R., Adhikari, S. (2012). Fuzzy parametric uncertainty analysis of linear dynamical systems: A surrogate modeling approach. Mechanical Systems and Signal Processing, 32, 5–17
- Chuang Wang, Pingyu Jiang and Kai Ding (2017), A hybrid-data-on-tag–enabled decentralized control system for flexible smart workpiece manufacturing shop floors, Proc IMechE Part C: J Mechanical Engineering Science 2017, Vol. 231(4) 764–782, DOI: 10.1177/0954406215620452
- COLLINS, Jonathan, 2006b. Post Danmark to Tag Its Roll Cages. RFID Journal [online], [cit. 2014-12-15]. Dostupné z:http://www.rfidjournal.com/articles/view?2281
- Cooper, G.F.; Herskovits, E. A Bayesian Method for the Induction of Probabilistic Networks from Data. *Mach. Learn.* **1992**, *9*, 309–347
- Corazza G.R., Licciardello R., Malavasi G:, Marcone M:, "La ruota come sensore di interazione ruota-rotaia", Ingegneria Ferroviaria, Marzo 1999; 119-131
- Dinmohammadi, F.; Alkali, B.; Shafiee, M.; Bérenguer, C.; Labib, A. Risk Evaluation of Railway Rolling Stock Failures Using FMECA Technique: A Case Study of Passenger Door System. *Urban Rail Transit* **2016**, *2*, 128–145
- Domingos, P., & Pazzani, M. (1997). On the optimality of the simple Bayesian classifier under zero-one loss. Machine Learning, 29(2–3), 103–130
- Dostupné z: http://cocoa.ethz.ch/downloads/2014/06/None_CAM-WH-003.pdf
- Dostupné z: http://www.kodys.cz/rfid.html
- Dostupné z: http://www.sciencedirect.com/science/article/pii/S092552730700196X
- Dostupné z:http://www.emeraldinsight.com/doi/pdfplus/10.1108/EUM0000000004267
- Dostupné z:http://www.emeraldinsight.com/doi/pdfplus/10.1108/EUM0000000004267
- Dostupné z:http://www.sciencedirect.com/science/article/pii/S092552730700196X
- Dostupné z:http://www.sciencedirect.com/science/article/pii/S092552730700196X

- Durmus (2014), Control and Fault Diagnosis of Railway Signaling Systems: A Discrete Event Systems Approach, Doctoral Dissertation, Graduate School of Engineering, Osaka University
- Ekberg, A.; Kabo, E.; Andersson, H. An engineering model for prediction of rolling contact fatigue of railway wheels. *Fatigue Fract. Eng. Mater. Struct.* **2002**, *25*, 899–909
- Elkhoury, N., Moridpour, S., Robert, D.J. (2008). Degradation prediction of rail tracks: a review of the existing literature. The Open Transportation Journal, 12, 88–104
- EN 17239 (2018), EN standard prEN 17230:2018, Information technology – RFID in rail
- ETSI – Evropský ústav pro telekomunikační normy, ETSI EN 302 208 V3.2.0 (2018-02)
- FERNIE, John, Ruggero GOLINI a Matteo KALCHSCHMIDT, 1994. Quick Response: An International Perspective. International Journal of Physical Distribution [online], 24(6), s. 231-247 [cit. 2014-12-10]. DOI: 10.1007/978-1-84996-525-5_16.
- Fletcher, D.I.; Beynon, J.H. The effect of contact load reduction on the fatigue life of pearlitic rail steel in lubricated rolling-sliding contact. *Fatigue Fract. Eng. Mater. Struct.* **2000**, *23*, 639–650
- Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences, 55(1), 119–139
- Gerald, Poier: ROADMASTER – Hardware Betriebshandbuch, voestalpine SIGNALING Zeltweg GmbH, Abteilung HDF, 2016, Zeltweg
- GS1 (2012), AutoID in Rail (RFID in Rail)
- GS1 (2015), GS1 EPCIS for Rail Vehicle Visibility Application Standard
- GS1 (2018), Exchange of component/part lifecycle data in the rail industry Application Standard
- GS1 (2018), Identification of Components and Parts in the Rail Industry - Application Standard
- Gunnar Ivansson, Lennart Andersson (2012), RFID in Rail: From pilot to implementation, presentation, 2012
- Hakan, G. (2014). Prediction of railway track geometry deterioration using artificial neural networks: a case study for Turkish state railways. Structure and Infrastructure Engineering, 10, 614-626
- Hecht, M., Mitusch, K. (2017). Lärm des Schienengüterverkehrs - wie weiter nach Einführung der Verbundbremssohle? ZEVrail, 141, 294-300
- Hendrickson S., Getting Started with Hadoop with Amazon's Elastic MapReduce, EMR, 2010.
- Hilbert M. and López P., "The world's technological capacity to store, communicate, and compute information," Science, vol. 332, no. 6025, pp. 60–65, 2011.
- Ho, T. K. (1995). Random decision forests. In Document analysis and recognition, 1995., proceedings of the third international conference on (Vol. 1, pp. 278–282)
- Huang, T.-M., Kecman, V., & Kopriva, I. (2006). Kernel based algorithms for mining huge data sets (Vol. 1). Springer
- Khan, N., Yaqoob, I., Hashem, I. A. T., Inayat, Z., Ali, M., Kamaleldin, W., & Gani, A. (2014). Big data: survey, technologies, opportunities, and challenges.The Scientific World Journal, 2014.
- INNOWAG D1.1 (2017), Deliverable D1.1 Benchmark and market drivers for an integrated intelligent and lightweight wagon solution, Shift2Rail IP5 project INNOWAG
- INNOWAG D2.3 (2018), Deliverable D2.3 Wireless Data Communication Concept for Cargo Condition Monitoring System, Shift2Rail IP5 project INNOWAG
- INNOWAG D2.4 (2019), Deliverable D2.4 Validation of Cargo Condition Monitoring System, Shift2Rail IP5 project INNOWAG
- Ishida, M.; Akama, M.; Kashiwaya, K.; Kapoor, A. The current status of theory and practice on rail integrity in Japanese railways-rolling contact fatigue and corrugations. *Fatigue Fract. Eng. Mater. Struct.* **2003**, *26*, 909–919
- J.Q. Qian, C. Zhang, L.J. Wu, X.J. Zhao, D.G. Wei, Z.H. Jiang and Y.h. He, (2011), A Passive UHF Tag for RFID-based Train Axle Temperature Measurement System, 2011 IEEE Custom Integrated Circuits Conference (CICC), DOI: 10.1109/CICC.2011.6055420
- Jeong, D. Analytical modelling of rail defects and its applications to rail defect management. 2003
- JONES, Peter, Colin CLARKE-HILL, Peter SHEARS, Daphne COMFORT a David HILLIER, 2004. Radio frequency identification in the UK: opportunities and challenges. International Journal of Retail

[online], 32(3), s. 164-171 [cit. 2014-12-10]. DOI: 10.1108/09590550410524957. Dostupné z: http://www.sciencedirect.com/science/article/pii/S092552730700196X
- KADLEC, Jaroslav, Radek KUCHTA a Radovan NOVOTNÝ, 2013. Přehled identifikačních technologií pro "Internet věcí" = koncept sítě mezi nejrůznějšími objekty. Slaboproudý obzor [online]. 69(1): 14-20 [cit. 2015-10-07]. Dostupné z: http://www.slaboproudyobzor.cz/files/20130104.pdf
- KÄRKKÄINEN, Mikko, A. LUI, C.K.Y. LO, Alp USTUNDAG a Mehmet Serdar KILINC, 2003. Increasing efficiency in the supply chain for short shelf life goods using RFID tagging: The Tagging Cost Sharing Approach. International Journal of Retail [online], 31(10), s. 69-77 [cit. 2014-12-10]. DOI: 10.1007/978-1-4471-4345-1_6
- KÄRKKÄINEN, Mikko, Jan HOLMSTRÖM a Matteo KALCHSCHMIDT, 2002. Wireless product identification: enabler for handling efficiency, customisation and information sharing. Supply Chain Management: An International Journal [online], 7(4), s. 242-252 [cit. 2014-12-10]. DOI: 10.1108/13598540210438971
- Karl Åkerlund, Seppo Mäkitupa (2015), The Value Proposition for GS1 ID Keys and RFID in Rail, presentation, Transport & Logistics Workshop GS1 Event 2015
- Kodys. RFID - radiofrekvenční identifikace. [online]. [cit. 2015-04-02]
- Kumar, S. A Study of the Rail Degradation Process to Predict Rail Breaks, Lulea University of Technology, 2006
- Larsson, D. A Study of the Track Degradation Process Related to Changes in Railway Traffic; 2004
- Lewis RB, Richards AN. A compensated accelerometer for measurement of railway track crosslevel. IEEE Trans Veh Technol. 1988;37(9):174–178.
- Lidén, T. (2015). Railway infrastructure maintenance - A survey of planning problems and conducted research. Transp. Res. Procedia, vol. 10, no. July, pp. 574–583
- LOŠŤÁK, Michal, 2009. RFID. EULOG: Informační logistický portál [online]. [cit. 2015-10-01]. Dostupné z: http://www.eulog.cz/clanky/rfid/?mt=&id=1958&m=a08
- Matthias Krüger (2018), Smarte Infrastuktur – Strategien und Anwendungen, presentation, VPI Hamburg 09.01.2018
- Mermec (2012), London Underground chose MERMEC Unattended Geometry Measuring Systems (UGMS) for line inspection, Online resource: http://www.mermecgroup.com/de/pageview2.php?i=730&sl=1 , accessed on 18.03.2019
- Mermec (2019), trackmeasurement - Track Geometry, online resource: http://www.mermecgroup.com/inspect/track-measurement/185/track-geometry.php , accessed on 18.03
- Mezitis et al. (2016), Object Controller of Electric Drive of Point Machine with Expanded Set of Diagnostic Tools, ICTE 2016, December 2016, Riga, Latvia
- Nie, Y., Tang, Z., Liu, F., Chang, J., Zhang, J. (2017). A data-driven dynamics simulation framework for railway vehicles. Vehicle System Dynamics, 56, 406–427
- O´CONNOR, Mary Catherine, 2007. Spain's Post Office Improves Delivery Speed. RFID Journal [online]. [cit. 2015-09-16]. Dostupné z: http://www.rfidjournal.com/articles/view?3209/
- O´CONNOR, Mary Catherine, 2008. Spanish Postal Service Expands Its RFID Deployment. RFID Journal [online]. [cit. 2015-09-16]. Dostupné z: http://www.rfidjournal.com/articles/view?4349
- Öberg, J. & Andersson E., 2009. Determining the deterioration cost for railway tracks. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 223(2), pp.121-129
- OLTIS (2011), Unikátní testování RFID čipů. [cit. 2019-02-30]. Dostupné z: https://www.oltis.cz/unikatni-testovani-rfid-cipu/
- Onodera, T. (2005). Decision making utilizing a large data set measured by track condition-monitoring device on a commercial service car, in 16th International Conference RAIL ENGINEERING, 2015
- P. Weston, C. Roberts, G. Yeo & E. Stewart (2015) Perspectives on railway track geometry condition monitoring from in-service railway vehicles, Vehicle System Dynamics, 53:7, 1063-1091, DOI: 10.1080/00423114.2015.1034730
- PER (2016), Rail track monitoring system – Wireless Autonomous On-Board System measuring

vibration with continuous reporting to reduce maintenance costs and enhance reliability and safety, H2020 project
- Philipp Thalmann (2017), RFID als Grundlage der Automation bei SBB Cargo, Presentation, 2017
- Pintsch (2019), PINTSCH BAMAG diagnosis software for level-crossing control systems, online resource, accssed on 12th May 2019, http://pintschbamag.de/en/produkte/2064-2/railway-signaling/level-crossing-components/pb-inductive-atc-control-unit-2/
- PRATER, Edmund, Gregory V. FRAZIER a Pedro M. REYES, 2005. Future impacts of RFID on e-supply chains in grocery retailing. Supply Chain Management: An International Journal [online], 10(2), s. 134-142 [cit. 2014-12-10]. DOI: 10.1108/13598540510589205.
- PRERADOVIC, Stevan a Nemai Chandra KARMAKAR, 2012. Multiresonator-based chipless RFID: barcode of the future. New York: Springer, 170 p. ISBN 978-1-4614-2094-1.
- Quinlan, J. R. (1986). Induction of decision trees. Machine Learning, 1(1), 81–106
- Ringsberg, J.W.; Bergkvist, A. On propagation of short rolling contact fatigue cracks. Fatigue Fract. Eng. Mater. Struct. 2003, 26, 969–983
- RIVAS D2.5 (2013), Overview of Methods for Measurement of Track Irregularities Important for Ground-Borne Vibration, EU FP7 project: Railway Induced Vibration Abatement Solutions
- S. Preradovic, N. Karmakar (2006), RFID Readers – a Review, 4th International Conference on Electrical and Computer Engineering ICECE 2006, 19-21 December 2006, Dhaka, Bangladesh
- Sackfield, A., Dini, D., Hills, D. (2007). Contact of a rotating wheel with a flat. International Journal of Solids and Structures, 44, 3304–3316
- Sadeghi, F.; Jalalahmadi, B.; Slack, T.S.; Raje, N.; Arakere, N.K. A Review of Rolling Contact Fatigue. J. Tribol. 2009, 131, 041403
- Saksena (2019), Functional Requirements Specification For Condition Monitoring for Predictive Maintenance of Signaling assets using IOT & Data Analytics, RailTel Enterprises Limited, online resource: http://relindia.in/images/Draft%20FRS%20Document%20for%20Preedective%20Maintenence-Revised%201.0.pdf
- Sarita Pais (2010), Data Format for RFID Data Using a Data-on-Tag Approach, Master Thesis, Auckland University of Technology
- Sarita Pais and Judith Symonds (2011), Data Storage on a RFID Tag for a distributed System, International Journal of UbiComp (IJU), Vol.2, No.2, April 2011
- Sawley, K.; Kristan, J. Development of bainitic rail steels with potential resistance to rolling contact fatigue. Fatigue Fract. Eng. Mater. Struct. 2003, 26, 1019–1029
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural Networks, 61, 85–117
- • Steenbergen, M.J. (2008). The role of the contact geometry in wheel–rail impact due to wheel flats: Part II. Vehicle System Dynamics, 46, 713–737
- SWEDBERG, Claire, 2012. Post Danmark Boosts Mail-Collection Efficiency. RFID Journal [online], [cit. 2014-12-15]. Dostupné z:http://www.rfidjournal.com/articles/view?9257
- SWEENEY II, Patrick J, 2005. RFID for Dummies. Indiana: Wiley, 388 s. ISBN 978-0-7645-7910-3
- SZAFRAŃSKI, Bohdan, 2011. Automatická identifikace - čárový kód nebo RFID. Řízení & údržba průmyslového podniku: Trade Media International [online], [cit. 2014-12-11]. Dostupné z: http://udrzbapodniku.cz/index.php?id=47&tx_ttnews%5Btt_news%5D=3883&tx_ttnews%5BbackPid%5D=35&tx_ttnews%5BsViewPointer%5D=2&cHash=639564f2a2 track using in-service vehicle. In: Perpinya X, editor. Reliability and safety in railway. InTech; 2012
- Tsunashima H, Naganuma Y, Matsumoto A, Mizuma T, Mori H. Condition monitoring of railway
- VIOLINO, Bob, 2005a. The Basics of RFID Technology. RFID Journal [online]. [cit. 2015-09-08]. Dostupné z: http://www.rfidjournal.com/articles/view?1337
- VIOLINO, Bob, 2005b. What is RFID? RFID Journal [online]. [cit. 2015-09-30]. Dostupné z: http://www.rfidjournal.com/articles/view?1339/
- Voestalpine (2019), DIAGNOSTIC AND MONITORING TECHNOLOGIES FOR TRACK AND FIXED INFRASTRUCTURE ASSETS, online resource, accessed on 13th May 2019,

http://www.voestalpine.com/signaling/static/sites/signaling/.downloads/produkte/fiam/voestalpine_signaling_fixed-infrastructure-asset-monitoring_en.pdf

- Weston PF, Ling CS, Goodman CJ, Roberts C, Li P, Goodall RM. Monitoring lateral track irregularity from in-service railway vehicles. IMechE Pt F: J Rail Rapid Transit. 2007;221(1):89–100
- Weston PF, Ling CS, Goodman CJ, Roberts C, Li P, Goodall RM. Monitoring lateral track irregularity from in-service railway vehicles. IMechE Pt F: J Rail Rapid Transit. 2007;221(1):89–100.
- Wingler (2018), LEVERAGING DEVELOPMENTS IN MONITORING TECHNOLOGY FOR OPTIMISING TRACK MAINTENANCE -- Switch Monitoring and Diagnostic Systems, online resource, accessed on 12th May 2019 http://www.drwingler.com/wp-content/uploads/2018/01/SWITCH-MONITORING-AND-DIAGNOSTIC-PDF.pdf
- Yamamoto, D. (2019). "Improvement of Method for Locating Position of Wheel / Rail Contact by Means of Thermal Imaging". QR of RTRI (Railway Technical Research Institute), Vol. 60, No. 1, Feb. 2019

11. Appendices