**UNIVERSITY OF GENOVA**

**POLYTECHNIC SCHOOL**

**DIME**

**DEPARTMENT OF MECHANICAL, ENERGY, MANAGEMENT**

**AND TRANSPORTATION ENGINEERING**

**MASTERS OF THESIS IN**

**SAFETY ENGINEERING FOR TRANSPORT, LOGISTICS, AND PRODUCTION**

# PERFORMANCE ASSESSMENT AND VALIDATION OF RAILWAY MAINTENANCE OPTIMIZED PLANS VIA DISCRETE EVENT SIMULATIONS

**SUPERVISOR**:

Prof. Sacco Nicola

Prof. Alice Consilvio

**CO-SUPERVISOR**:

Dr. Emanuela Baglietto

**CANDIDATE**

Thakur Deepak Chowhan (4708904)

Karthika Suresh (4708652)

# Abstract

The purpose of this thesis is to understand, analyse and resolve the issues regarding railway maintenance planning under uncertainty by using ExtendSim. A predictive maintenance strategy is considered, aimed at planning the interventions before failure occurrence, avoiding corrective maintenance and service disruptions.

A degradation model for rail track geometry, provided in Literature, is considered to evaluate the interventions priority and the related deadlines.

Two different approaches are applied: a discrete event simulation approach and an optimization approach based on Mixed Integer Linear Programming (MILP).

Three scenarios are taken into consideration: One, in which execution of the activities are based on the position of the assets along the line; Second, execution of the maintenance interventions according to the degradation of the assets; and the third which is an optimization scenario in which a complex cost function is considered to take into account different aspects. The results of these simulations and of the optimization model are reported, highlighting the differences between the considered scenarios.

# Acknowledgements

We express our deep sense of gratitude to our respected and learned guide Prof. Sacco Nicola, Prof. Alice Consilvio and Dr. Emanuela Baglietto for their valuable help and guidance.

We are also thankful to our University “Università Degli Studi di Genova” for giving us this opportunity.

Lastly, we would like to express our deep appreciation towards our classmates and our indebtedness to our Family for providing the moral support and encouragement.

# Table of Contents

# 1.Introduction

A brief introduction to the problem which covers the underlying background and problem areas of safety, security, efficiency, and economy through the optimization of rail maintenance management. Maintenance has a great impact on rail transport performances, costs, and quality. It is indispensable to ensure service availability and safety for people and goods. Furthermore, a correct and efficient maintenance management may be a way to reduce costs and improve the quality of services.

Nevertheless, the space-distributed aspect of railway infrastructure should be considered, and such a characteristic generates significant difficulties in the organization of maintenance activities and in the management of the relevant supporting resources. Another aspect that makes railway maintenance critical is the time constraint. In fact, the available time for maintenance activities is strictly limited due to various factors, such as rail traffic, climate, fulfilment of fixed operation sequences, etc. Some of these requirements result in soft constraints, that is, violations can be tolerated if no better choices exist, and some are hard constraints that can never be violated. Therefore, any asset of a railway system needs very carefully planned maintenance activities, aiming at guaranteeing the best performance as possible in any time.

To cope with this problem, many maintenance approaches have been developed in the relevant literature, such as corrective maintenance, performed when a fault occurs, or preventive maintenance which can be subdivided into:

- **Planned Maintenance**, performed on a regular fixed time schedule. It can lead to a significant reduction of the useful life of components, due to early replacement and unnecessary, a-priori scheduled maintenance activities.
- **Condition-Based Maintenance** performed only when necessary, based on the continuously monitored asset conditions. This approach allows a better usage of infrastructure components but requires a regular and frequent monitoring of the degradation state of railway assets.
- **Predictive Maintenance** performed only when necessary, on the base of suitable model estimations.

In this context, considering the recognized relevance of the problem at the European level, the aim of this thesis is to study the planning of predictive maintenance activities in the railway sector through simulative and optimization approaches.

# 2.Context and Background

Track needs regular maintenance to remain in good order, especially when high-speed trains are involved. Inadequate maintenance may lead to a "slow order" being imposed to avoid accidents. Track maintenance was at one time hard manual labour, requiring teams of labourers, or trackmen, who used lining bars to correct irregularities in horizontal alignment of the track, and tamping and jacks to correct vertical irregularities. Currently, maintenance is facilitated by a variety of specialised machines.
The surface of the head of each of the two rails can be maintained by using a rail grinder.

Common maintenance jobs include changing sleepers, lubricating, and adjusting switches, tightening loose track components, and surfacing and lining track to keep straight sections straight and curves within maintenance limits. The process of sleeper and rail replacement can be automated by using a track renewal train.

Spraying ballast with herbicide to prevent weeds growing through and redistributing the ballast is typically done with a special weed killing train.

Over time, ballast is crushed or moved by the weight of trains passing over it, periodically requiring relevelling ("tamping") and eventually to be cleaned or replaced. If this is not done, the tracks may become uneven causing swaying, rough riding and possibly derailments. An alternative to tamping is to lift the rails and sleepers and reinsert the ballast beneath. For this, specialist "stone blower" trains are used.

Rail inspections utilize non-destructive testing methods to detect internal flaws in the rails. This is done by using specially equipped Hi Rail trucks, inspection cars, or in some cases handheld inspection devices.

Rails must be replaced before the railhead profile wears to a degree that may trigger a derailment. Worn mainline rails usually have sufficient life remaining to be used on a branch line, siding or stub afterwards and are "cascaded" to those applications.

The environmental conditions along railroad track create a unique railway ecosystem. This is particularly so in the United Kingdom where steam locomotives are only used on special services and vegetation has not been trimmed back so thoroughly. This creates a fire risk in prolonged dry weather.

In the UK, the cess is used by track repair crews to walk to a work site, and as a safe place to stand when a train is passing. This helps when doing minor work, while needing to keep trains running, by not needing a Hi-railer or transport vehicle blocking the line to transport crew to get to the site.

## 2.1 Maintenance Objectives

Since rail is an important transportation mode, proper maintenance of the existing lines, repairs and replacements carried out in time are all important to ensure efficient operation Moreover, since some failures might have a strong impact on the safety of the passengers, it is important to prevent these failures by carrying out in time and according to some predefined schedules preventive maintenance works. It is a complex infrastructure that requires of a high degree of safety and reliability. The maintenance of this system is a complicated and expensive task which represent and important share of total railway infrastructure costs.

- to ensure safe operations of rolling stocks at the scheduled speed,
- to afford conditions for the highest quality and reliability of transport,
- to maintain and increase high level of safety,
- to reduce costs, without however decreasing safety standards,
- to improve organization, materials, equipment, and staff's qualification to respond more efficiently to requirements of operation.

# 2.2 Railway maintenance planning

Railway infrastructures scarcely redundant (no or very few path alternatives) and this implies that, when a fault occurs, the system performances have a dramatic drop space distributed aspect of railway infrastructure Railway assets are often not spatially delimited to a point, and this implies difficulties in the organization of maintenance activities and resources time constraints due to rail traffic operation Therefore, any asset of a railway systems needs very carefully planned maintenance activities, aiming at guaranteeing the best performances as possible in any time In particular predictive railway maintenance is aimed at minimizing the probability of the occurrence, of the so called mission critical faults during train service, that is. those that prevents trains for circulating or can lead to accidents, while keeping maintenance costs as low as possible

Railway infrastructure maintenance works need possession of the infrastructure. The possession for maintenance can be,

- partial when maintenance and trains share the infrastructure (risk conditions!)
- privative when the maintenance takes full possession of it

The privative possession can be divided in

- Overnight possession takes place in the free of service periods (time window)
- Weekend possession makes use of the fact that train services are reduced (may be reduced, re-scheduled or re-routed) respect to labour day services.
- Daytime possession the shortage of available time windows makes this possession to be focused on operations that cannot be postponed for latter, such as corrective works.

# 2.3 Maintenance Operations

- Routine (spot) maintenance works consist of inspections and small repairs of the local irregularities carried out manually or using small machines.
  For example: switch inspections
- Project (systematic maintenance) include large amount of work. These activities are carried out with heavy track maintenance machines (e.g., tamping machines, ballast regulators, rail grinding machines, ballast cleaners)
- Renewal is done for safety reasons or when the maintenance of different track components is becoming too expensive.
- New constructions include all activities that are intended to construct completely new tracks, tunnels, bridges, stations, etc.

# 2.4 Maintenance Strategies

Four general types of maintenance philosophies can be identified, namely corrective, preventive, risk-based and condition-based maintenance.

- **Corrective Maintenance** - Maintenance is carried out following detection of an anomaly and aimed at restoring normal operating conditions. This approach is based on the firm belief that the costs sustained for downtime and repair in case of fault are lower than the investment required for a maintenance program. This strategy may be cost-effective until catastrophic faults occur.

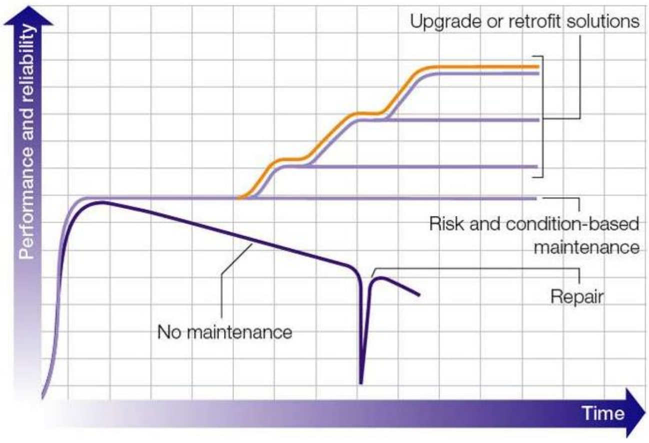


Figure 1: Maintenance Strategies

- **Risk-Based Maintenance** - Maintenance carried out by integrating analysis, measurement, and periodic test activities to standard preventive maintenance. The gathered information is viewed in the context of the environmental, operation and process condition of the equipment in the system.
  The aim is to perform the asset condition and risk assessment and define the appropriate maintenance program. All equipment displaying abnormal values is refurbished or replaced. In this way it is possible to extend the useful life and guarantee over time high levels of reliability, safety, and efficiency of the plant.
- **Condition-Based Maintenance** - Maintenance based on the equipment performance monitoring and the control of the corrective actions taken as a result.
  The real actual equipment condition is continuously assessed by the on-line detection of significant working device parameters and their automatic comparison with average values and performance. Maintenance is carried out when certain indicators give the signalling that the equipment is deteriorating, and the failure probability is increasing.
  This strategy, in the long term, allows reducing drastically the costs associated with maintenance, thereby minimizing the occurrence of serious faults, and optimizing the available economic resources management.

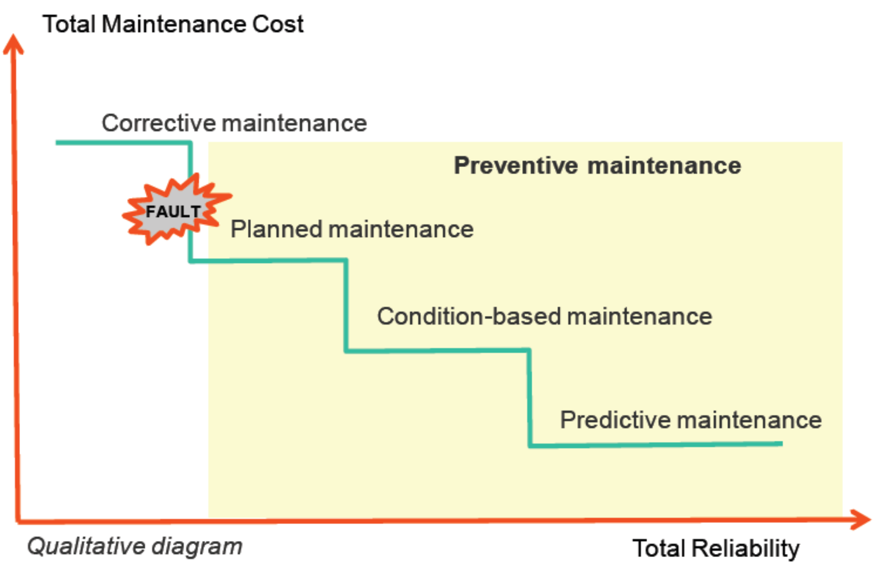

Figure 2: Comparison of Maintenance Strategies

## 2.5 Comparison of Maintenance strategies

In this subsection, the behaviour of the four presented maintenance policies will be analysed. Then, the policies will be compared to each other and some advantages and disadvantages will be shown.

In Figure 3 shown below, the maintenance policies are evaluated with respect to the planning period. As suggested in the description of the policies, reactive maintenance must be executed promptly or short-term. In condition-based maintenance policies, the planning period depends on the selected trigger and the inspection interval. In the most cases, condition-based maintenance must be planned in short-term, but if the trigger has a large buffer, the planning period can be longer. Preventive and predictive maintenance have longer planning periods. In preventive maintenance, the planning period depends on the length of the time trigger. In predictive maintenance, the planning periods depends on the credibility of the prediction models. It is assumed that predictive maintenance is implemented if and only if the prediction is good enough. Then, the planning period is medium- to long-term.

Figure 3: Evaluation with respect to planning period.

In Figure 4, the maintenance policies are evaluated with respect to risks. Reactive maintenance is classified as risky, because in the most applications it is too risky to wait until failures occurs. That does not mean, that reactive maintenance is never suitable. There are some applications, where reactive maintenance is a good option, for example when some components are redundant. Condition-based maintenance is safe, if the inspection interval is not too long or the condition trigger is not too low for the degradation rate. To evaluate the inspection interval and the condition trigger value, the degradation rate must be approximated. If deterioration is underestimated or the chosen parameters are not suitable, maintenance can be requested too late, and the risk is higher. Preventive and predictive maintenance are safe, because of the longer planning period. With it, possibly misjudgements in the parameter evaluation can be seen in time and the parameters can be adjusted.

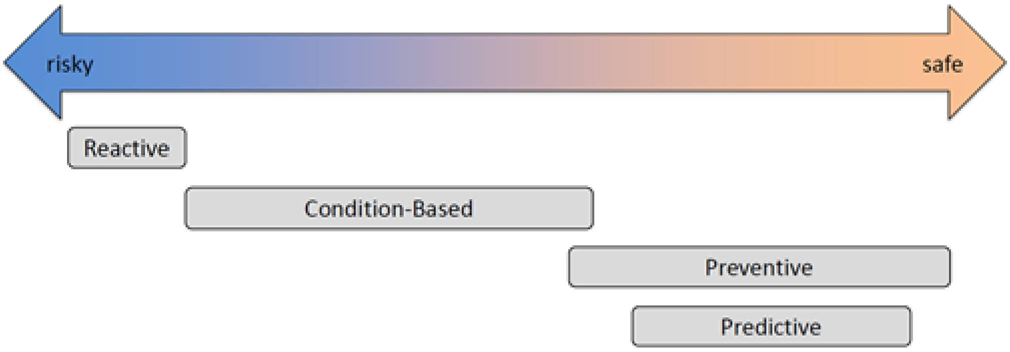


Figure 4: Evaluation with respect to risk

In Figure 5, the maintenance policies are evaluated with respect to the cost effectiveness. Reactive maintenance is in the most cases expensive because failures are not avoided by improving the asset through additional maintenance. Then, deterioration is fast, and the assets must be replaced frequently. Preventive maintenance is in the most cases expensive. To reduce risks, the maintenance intervals and usage triggers should be chosen pessimistic. With it, more maintenance is done than necessary which increases the costs. Condition-based maintenance has a strong spread of costs. It can be effective, but it can also be expensive – depending on the selected parameters, costs for inspection/monitoring and costs for maintenance. Predictive maintenance can be cost effective if monitoring is not too expensive. This results from the long-term planning of maintenance when needed. With it, the condition trigger can be chosen lower than in condition-based maintenance.

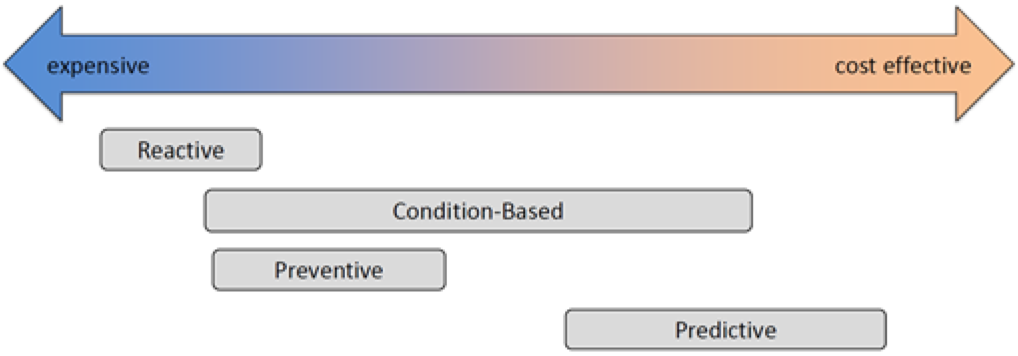


Figure 5: Evaluation with respect to cost effectiveness

In Figure 6, the maintenance policies are evaluated with respect to the amount of information the policy provides regarding the behaviour of the infrastructure system. Reactive maintenance provides no additional information since only breakdowns are observed. Also, preventive maintenance provides only few information because inspection and monitoring have a secondary role. Condition-based and predictive maintenance gives a lot of information about the infrastructure condition. To observe the deterioration process, monitoring systems are implemented, or inspections are performed. Because in predictive maintenance also information about the future condition is provided, this strategy leads to the best-informed situation.

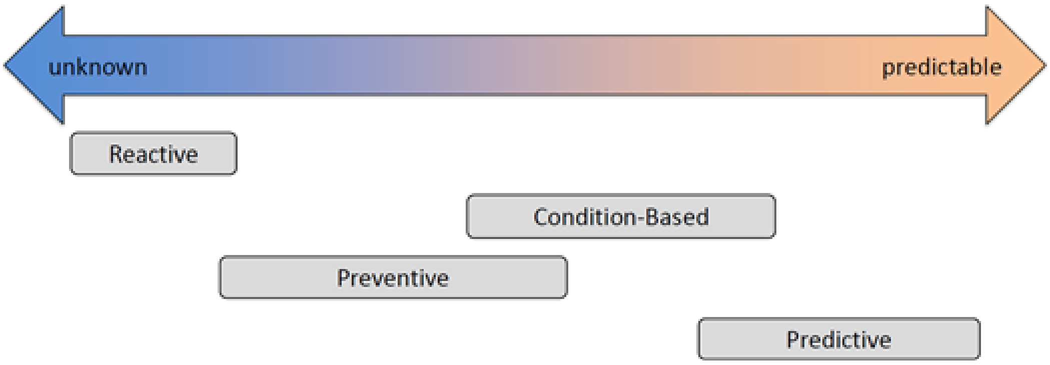


Figure 6: Evaluation with respect to the amount of information

In Figure 7, the maintenance policies are evaluated with respect to the implementation effort. Reactive maintenance can be implemented intuitive without defining trigger values and installing monitoring systems. Only inspection to obtain failures is necessary. Also, preventive maintenance has a low implementation effort, time triggers can be defined based on expert knowledge. But if the maintenance operator increases the effort for implementation, e.g., by analysing the deterioration process to derive better time triggers, the strategy can be improved in terms of costs or risks. The implementation effort for condition-based maintenance is higher because monitoring systems must be installed, or the assets have to be inspected closely. The highest implementation effort has predictive maintenance: it requires monitoring, data evaluation and expertise to derive suitable deterioration models to predict future condition with a high quality.

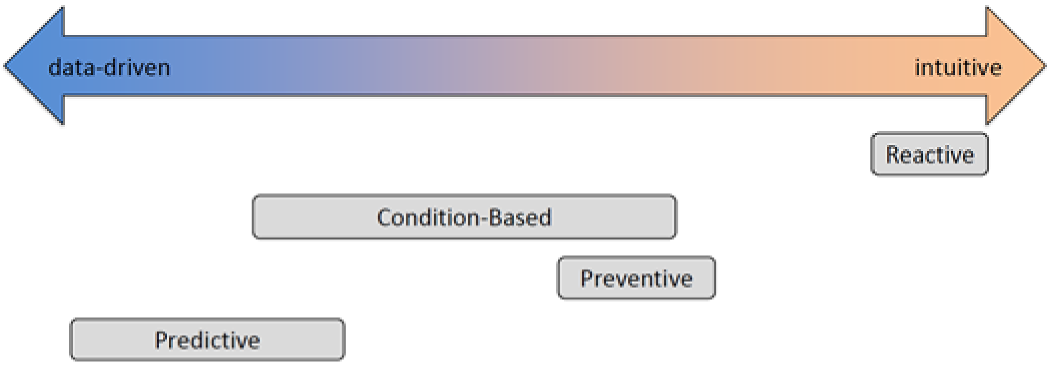


Figure 7: Evaluation with respect to the Implementation Effort

In Figure 8, the different evaluation criteria of the maintenance policies are aggregated into one diagram. The evaluation regarding,

- the length of the planning period from short-term to long-term,
- the risk awareness from risky to safe,
- cost effectiveness from expensive to cost effective,
- the amount of infrastructure quality knowledge from unknown to predictable and
- the implementation effort from data-driven to intuitive

is summarised as a radar chart.

Reactive maintenance is easy to implement, but the planning period is short, risks are not avoided and no information about condition is given. Thus, reactive maintenance can be used for components with less risk in case of break down and less replacement effort.
Preventive maintenance also has a low implementation effort, only the maintenance activities and the time or usage trigger must be defined. Therewith, maintenance activities can be planned in long-term, and risks can be avoided. But often, this approach is expensive because preventive maintenance is usual done before needed. This approach also provides only less information about infrastructure quality.

Condition-based maintenance helps to reduce risks and costs because maintenance is done when necessary. This requires closely monitoring to know the current infrastructure condition. After detecting signs of deterioration, maintenance should be executed promptly. So, the planning period is more short- up to medium-term.
Predictive maintenance combines long-term planning with condition-based maintenance. Based on failure and deterioration models, future infrastructure condition is predicted and based on it maintenance can be planned in advance. To use this approach, a depth understanding of the underlying deterioration processes and the failure models is necessary.

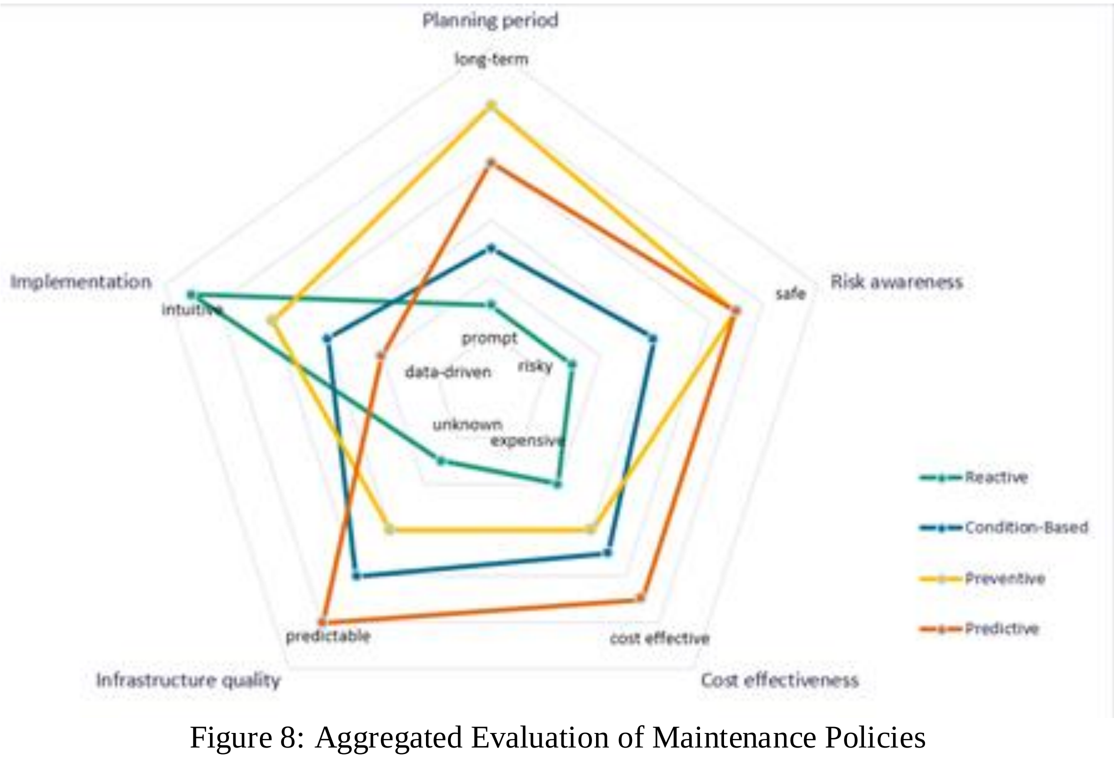


Figure 8: Aggregated Evaluation of Maintenance Policies

Table 1: Comparison of Maintenance policies

| | Pro | Contra | Possible application |
|---|---|---|---|
| Reactive Maintenance | No monitor systems are needed | Unexpected Failure. | Assets whose breakdown has a minor influence in the network only |
| Condition-Based | Operator knows a lot about network condition | Inspection or monitoring are necessary | Assets whose breakdown has an higher influence in the network, but whose deterioration is difficult to predict |
| Preventive Maintenance | No monitor systems are needed.<br><br>Activities are planned long-term | Less condition information.<br><br>Good choice of trigger value is important | Assets with an estimable deterioration where monitoring is too expensive |
| Predictive Maintenance | Activities are planned long-term.<br><br>Best maintenance time with respect to costs and risks can be chosen | Monitoring is necessary.<br><br>Deterioration has to be predictable.<br><br>Deterioration should be largely independent from external influences | Assets with monitoring systems and an analysed and recognised deterioration process |

## 2.6 Maintenance Management

Maintenance strategies must evolve to support the technological requirements of modern equipment and the challenges of a competitive and legislated environment.
Selecting a successful maintenance strategy requires a good knowledge of maintenance management principles and practices as well as knowledge of specific facility performance. There is no one correct formula for maintenance strategy selection and, often, the selection process involves a mix of different maintenance strategies to suit the specific facility performance and conditions.

There are several maintenance strategies available today that have been tried and tested throughout the years. These strategies range from optimization of existing maintenance routines to eliminating the root causes of failures altogether, to minimize maintenance requirements. Ultimately, the focus should be on improving equipment reliability while reducing cost of ownership.

An effective maintenance strategy is concerned with maximizing equipment uptime and facility performance while balancing the associated resources expended and ultimately the cost. We need to ensure that we are getting sufficient return on our investment.

Are we satisfied with the maintenance cost expended versus equipment performance and uptime? There is a balance to be had in terms of maintenance cost and facility performance. We can develop a suitable maintenance strategy to help tailor this balancing act in order to ensure the return on investment is acceptable.

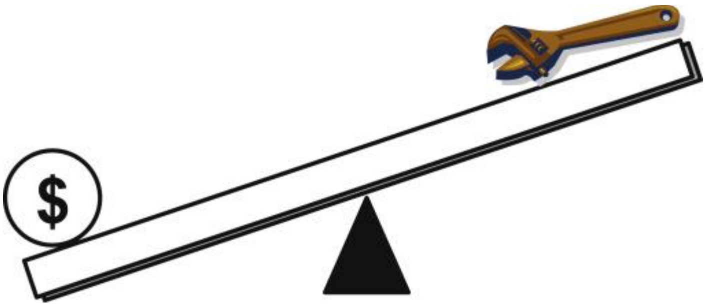


Figure 9: Maintenance Balancing Act

A maintenance strategy should be tailored specifically to meet the individual needs of a facility. The strategy is effectively dynamic and must be updated periodically as circumstances change. The strategy must include a detailed assessment of the current situation at the facility and consider the following questions:

- What is the performance history of the facility equipment and systems?
- What are the production targets, i.e., what are the mission times for facility equipment and systems?
- What is the facility shutdown targets?
- What is the current maintenance budget?

Once we have clarity on the current situation and constraints, we need to define the objectives of the maintenance plan. The objectives must align with the business objectives of the company. They must be developed by all the key facility stakeholders and be clear, concise, and realistic. There may be several components to the strategy objectives – for example: improve equipment uptime, reduce maintenance costs, reduce equipment operating costs, extend equipment life, reduce spare parts inventory, improve MTTR, etc.

An example of a maintenance strategy workflow is illustrated in figure 10. This workflow is developed to optimize and improve an existing facility maintenance program. Depending on the specific circumstances at the facility, our strategy may also take us into the direction of a step change approach to maintenance management and opt for a reliability-centered maintenance (RCM) program, which may replace our existing maintenance program. This strategy is labor and time intensive.

It is a common theme in the industry that maintenance budget and resources are very thin on the ground relative to the amount of work that needs to be done. Therefore, prioritization of maintenance resources is essential to be successful. Once we have defined our maintenance strategy objectives, we need to define facility equipment criticality. Criticality is a risk-based approach that can help us to prioritize our resources effectively. It can also help to appraise the requirement and effectiveness of maintenance tasks already populated in the MMS or CMMS.

The output of the criticality review provides an input into several maintenance strategy activities. These activities may include the following:

- Reviewing and optimizing maintenance spares holding
- Defining and reviewing equipment tasks selection, such as corrective, proactive, and run to failure
- Review and update maintenance rounds
- Define and review the condition monitoring strategy.
- Review the CMMS to assess the effectiveness of maintenance tasks (data cleansing).

The key principle behind the review and optimization of an existing maintenance strategy is an accurate and robust criticality assessment, which may impact many of the strategy objectives.

Another common theme in the industry is that many computer maintenance management systems are populated with a large proportion of preventive maintenance tasks that may be considered as superfluous and even not necessary. These tasks may consume a large proportion of the maintenance resources and time without an acceptable return on the investment made (maintenance cost). The maintenance strategy should also ensure the current data in the CMMS is value adding and therefore carry out a “cleansing” exercise. A data cleansing exercise critically reviews and appraises the current CMMS tasks and aims to eliminate the tasks that may not be adding value and therefore are superfluous. By focusing on equipment criticality, these activities can be reviewed and appraised in a logical and systematic way.

Once the equipment criticality assessment is completed and the strategy objectives have been reviewed and updated, maintenance resources can then be aligned to the strategy. The maintenance strategy objectives will dictate the resources and associated maintenance costs. The next step in the strategy development process is to update the equipment maintenance and operating plan as presented.

The EMOP is the primary record and source of maintenance and operation information of each equipment item and includes the up-to-date maintenance and operating strategies. It provides the baseline information including equipment maintenance and operating parameters. We are then able to implement the maintenance strategy on the facility.

It is important to understand the impact (and the success) of the new maintenance strategy. This is achieved by setting key performance indicators (KPIs) to assess the facility maintenance performance. This is done by first developing a benchmark data set. How is the facility currently performing? What is the cost of maintenance? What is the MTBF? What is the MTTR? What is the maintenance rework ratio? Once the current facility maintenance performance is benchmarked, we can then measure maintenance performance against this

benchmark. Maintenance performance is reviewed periodically and depending on the results, may be reviewed and updated more frequently.
If the maintenance performance is in line with business objectives, then the facility operation will continue; however, if there is any deviation in performance or change in the facility process or criticality ranking, then the maintenance strategy should be revisited.

**Reliability-Centered Maintenance**

In 1978 Stanley Nowlan and Howard F. Heap published a report aimed at determining new and more cost-effective ways of maintaining complex systems in the aviation industry. It was called "Reliability-Centered Maintenance" (RCM)
Today, reliability-centered maintenance (RCM) is used across many industries and is recognized as one of the leading practices for oil and gas and petrochemical facility maintenance. RCM acknowledges that all equipment in a facility does not have an equal importance and that there are significant advantages in prioritizing maintenance efforts on certain facility equipment.

RCM effectively provides a structured approach to the development of a maintenance program. It focuses on equipment needs and ultimately results in a well-grounded basis for facility maintenance with a high proportion of proactive maintenance. RCM addresses the basic causes of equipment and system failures. It aims to ensure that controls are in place to predict, prevent or mitigate these functional failures and hence the associated business impact. RCM is defined by a technical standard from the Society of Automotive and Aerospace Engineers (SAE), namely SAE JA1011 (1999)

**RCM Workflow**

Reliability-centered maintenance (RCM) analysis provides a structured framework for analyzing the functions and potential failures of facility equipment, such as pumps, compressors, a facility processing unit, etc. The emphasis of the analysis is to preserve system function, instead of focusing on preserving the actual equipment. The output of an RCM program is a series of scheduled maintenance plans. The RCM standard, SAE JA1011, describes the minimum criteria that a process must comply with to qualify as an RCM Process.

Although in the application of RCM there tends to be a large amount of adaptation, usually it follows the steps illustrated in the workflow in below Figure.

**Prepare for the RCM Analysis**

To ensure the RCM analysis is executed smoothly, there are several preparatory activities that should be completed in advance of an RCM analysis.

First the RCM team should be carefully assembled. The team should comprise a cross-section of facility operations, maintenance, and FI&R teams with a strong technical understanding of the equipment to be analyzed. The team should also be conversant with the RCM analysis methodology.

RCM analysis requires a large investment of time and resources. Given this, it is often necessary for the facility maintenance group conducting the analysis to focus on a selection of equipment or systems. The equipment or systems to be analyzed should be identified and boundaries drawn around the battery limits of systems. This is to ensure clear demarcation of the RCM scope so that efforts and time are directed appropriately. It is often the case that a criticality assessment is used to determine the equipment or systems selection.

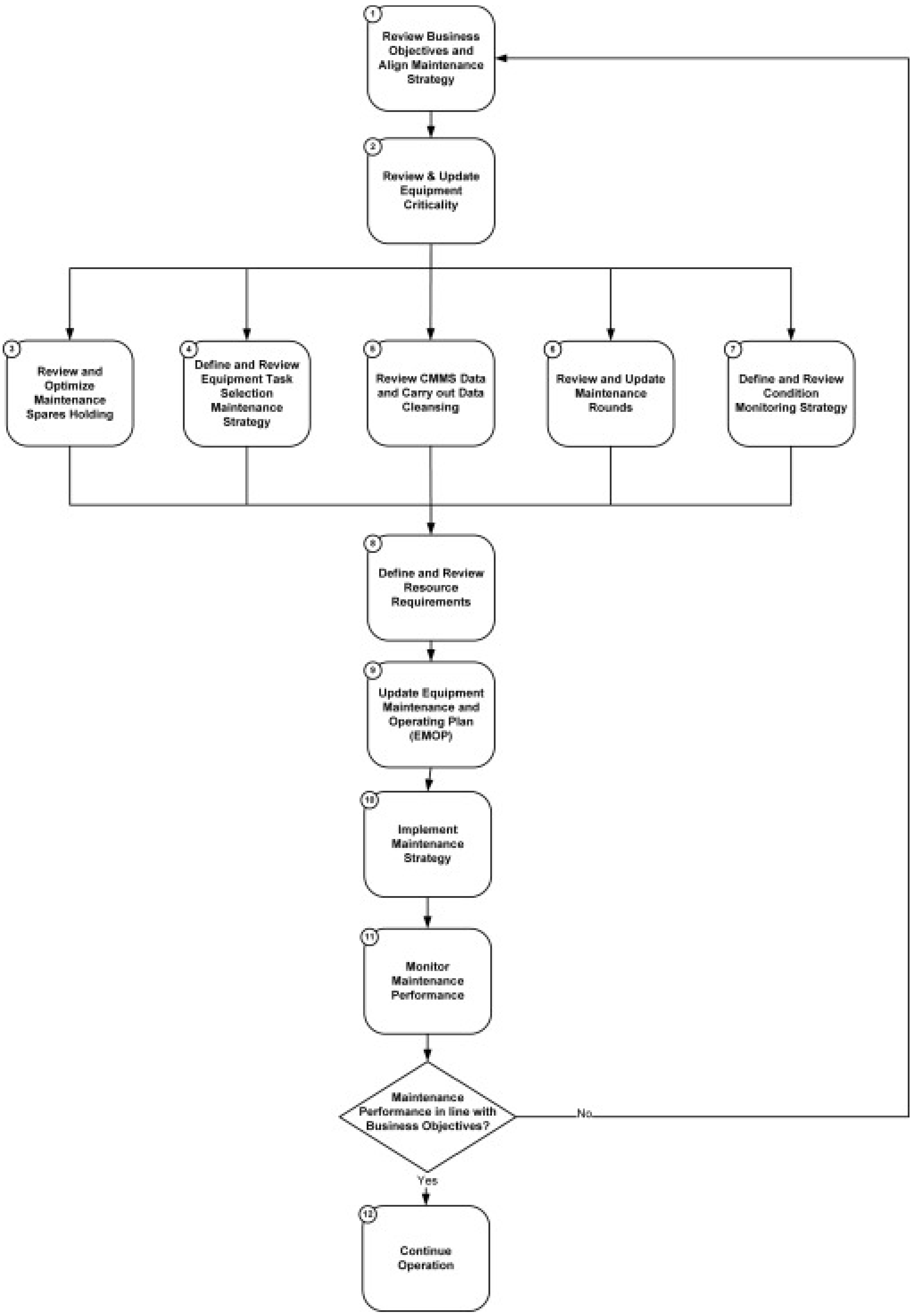

Figure 10: Maintenance Strategy Workflow

**Determine the Functions and Potential Functional Failures**

Reliability-centered maintenance focuses on preserving equipment functionality. The next step in the process is to determine the function or functions that the equipment or systems are intended to perform. Equipment functions should also be prescriptive in the definition of a function and include performance limits, for example.

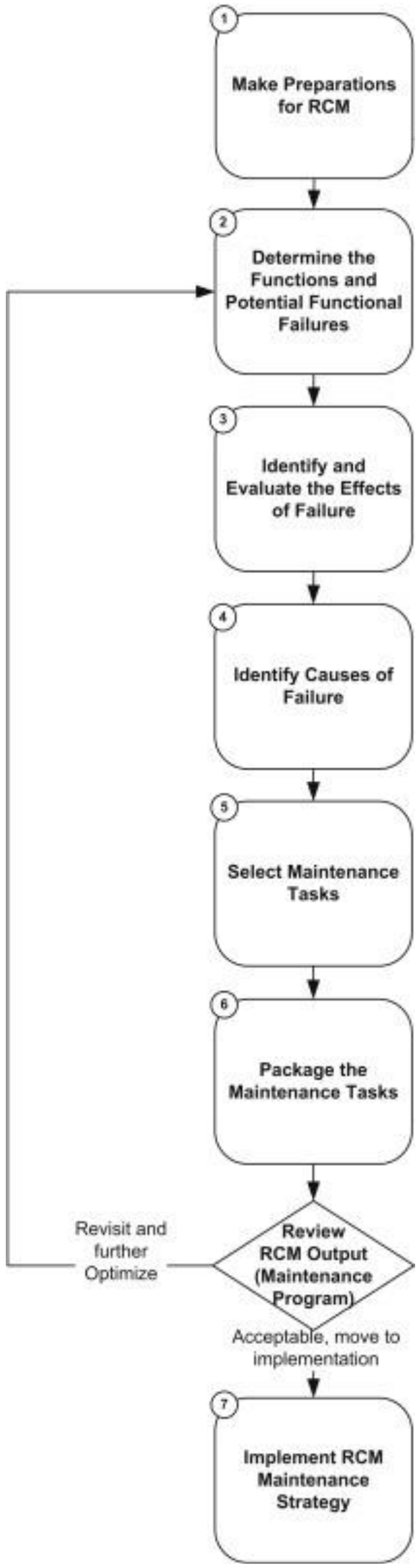


Figure 11: Reliability Centered Maintenance Workflow

Once the functions are clearly defined by the RCM team, their corresponding potential functional failures are defined. Functional failures may also include poor performance of a function or overperformance of a function as well.

**Identify and Evaluate the Effects of Failure**

The next step in the process is to identify and evaluate the effects of the equipment failure. This step enables the RCM team to prioritize and choose an appropriate maintenance strategy that can tackle the failure. It is common to employ a logic diagram to structure this part of the process to consistently evaluate and categorize the effects of failure.

**Identify Causes of Failure**

By identification of the specific cause of the failure we can understand the root cause and ultimately define an appropriate maintenance strategy that can address the failure altogether.

It is important to leverage the skills and experience of the RCM team to ensure the cause of the failure is clear and accurate. The cause of the failure should be described in sufficient detail at this stage. This is so that we can ensure the maintenance task selection step in the process is confidently and reliably completed. It may be appropriate to refer to the RCM standard, SAE JA1012, which presents useful guidance as to how to identify causes of failure.

**Select Maintenance Tasks**

At this stage in the process, we have identified the functions that equipment is intended to perform and the ways that these functions could fail. We have evaluated the effects of functional failures and identified their causes; the next step in the RCM process is to select appropriate maintenance tasks for the equipment to prevent such failures. There are a number of ways to carry out this exercise; however, the RCM team's skill set, and knowledge is the key factor.

**Package the Maintenance Tasks**

The final step in the RCM process is to package the maintenance tasks into a practical and robust maintenance program. This process involves reviewing the selected maintenance tasks and grouping them in a logical way so that they can be uploaded into the facility CMMS. The ultimate goal in packaging the RCM tasks is to arrive at a practical and efficient maintenance program.

**Implementation of RCM Maintenance Strategy**

Reliability-centered maintenance (RCM) has been in use for several years. It provides a structured and systematic framework which can result in an effective maintenance management program for facility equipment.

It is no surprise that RCM is a resource intensive and time-consuming process that can be expensive to develop and implement. There are several iterations of RCM that attempt to reduce the effort needed to develop and implement an RCM program, with varying degrees of success. It is important to maintain the key principles of RCM and not to overstretch the battery limits that were agreed on by the facility maintenance team at the start of the process. This may lead to disillusionment and frustration and eventually may result in a failed implementation effort.

The approach to the development and implementation of an RCM maintenance program must be executed with dedication and tenacity. It is also important for the facility management team and the wider facility functional groups to buy in and support the RCM implementation effort.

**Planned Maintenance Optimization**

Planned maintenance optimization (PMO) is a well-established, tried, and tested maintenance strategy, dating back to the 1990s. Around this time there was a lot of concern from the industry that RCM did not suit the requirements for facilities that had existing maintenance programs with limited resources and timescales to perform an RCM study. This is because primarily RCM is a tool that is designed for use in the design stage of the facility life cycle. PMO, on the other hand, is specifically designed to target existing maintenance programs.

The PMO process is illustrated in the workflow in Figure below. PMO identifies planned maintenance database activities from an existing facility CMMS, categorizing them into planned maintenance craft groups. The workflow then reviews each corresponding facility equipment history to determine if the planned maintenance task is necessary. These tasks are critically evaluated and ultimately optimized based on the added value. Finally, the maintenance program is updated along with the CMMS.

A PMO study may be conducted manually in a task force team or by employing commercially available software. There are numerous PMO software titles available in the market, some of which can be interfaced with a CMMS. Typically, the decision to implement a PMO strategy is made in an ad hoc fashion by the maintenance management team. It is usually driven by budget and resource constraints.

**Defect Elimination**

"Failures result from defects. Eliminating defects is the way we improve constantly and forever the system of production and service".
Equipment failures are a result of defects; therefore, by eliminating defects we can improve equipment reliability. Defect elimination is a maintenance strategy that takes us back to design. It aims to prevent defects being introduced at the early stages of the equipment life cycle, thereby removing the defects during the operational stage of the equipment life cycle.

By eliminating the defects that have potential to cause future equipment failures, maintenance requirements will also be reduced, resulting in improved equipment uptime. Defect elimination can reduce the maintenance requirements on equipment or systems and hence lower maintenance cost.

Defect elimination aims to identify failure modes and eliminate them at the outset. Each part of the equipment is taken in its component parts and corresponding defects are identified. Mitigation plans are then prepared for each defect identified to eliminate the failure mode. Control measures and quality assurance standards are developed to detect and eliminate defects before they are designed into the equipment and systems. One of the methods that could be employed in defect elimination is the FMEA tool which is based on failure mode and effects analysis.

Pre-emptive maintenance strategies such as defect elimination are very useful because they can be very cost effective and have lasting impressions on reducing maintenance requirements.

**Intentional Overdesign Selection**

In some instances, maintenance managers may decide to purposefully overdesign a particular equipment or system on the facility. The idea behind this is that these equipment items or systems are therefore able to withstand deterioration processes more and function for longer periods of time between failures.

This decision may be made when dealing with highly critical processes on the facility, such as processing toxic or hazardous materials or chemicals, or where there is a requirement to increase the reliability of a certain part of a process that may warrant additional robustness of equipment design.

This is a strategic maintenance decision intended to prolong facility equipment and systems life and therefore maintain longer periods of production. It involves increasing the design specification of equipment or systems with more robust parts, higher specification materials of construction, better surface protection coatings, etc.

Maintenance management is a continuous improvement process. The intention is to add value by improving equipment reliability while reducing cost of ownership. Clearly there is a balance to be had with cost of ownership versus additional value added, particularly with this strategy. There may be a higher cost of ownership; however, this is offset against the improvements to the production output.

**Shutdown Overhaul Maintenance**

During shutdown overhaul maintenance equipment and systems are repaired or overhauled on a set frequency that is shorter than the MTBF. By doing this we should prevent an unexpected failure.

Such work is typically done as an overhaul, where the whole of the equipment is removed from operation during a shutdown and taken to the workshop to be stripped down to its component parts and rebuilt as new.

Use of shutdown overhaul maintenance strategy is aimed at ensuring uninterrupted production for a specific period of time. By renewing or overhauling equipment regularly we remove the wear-out related stoppages. Once equipment is overhauled to manufacturer's standards we can expect as-new performance. However, we are also exposed to "infant mortality" risks due to poor quality control, mistakes during assembly, incorrect material selection and introduced damage.

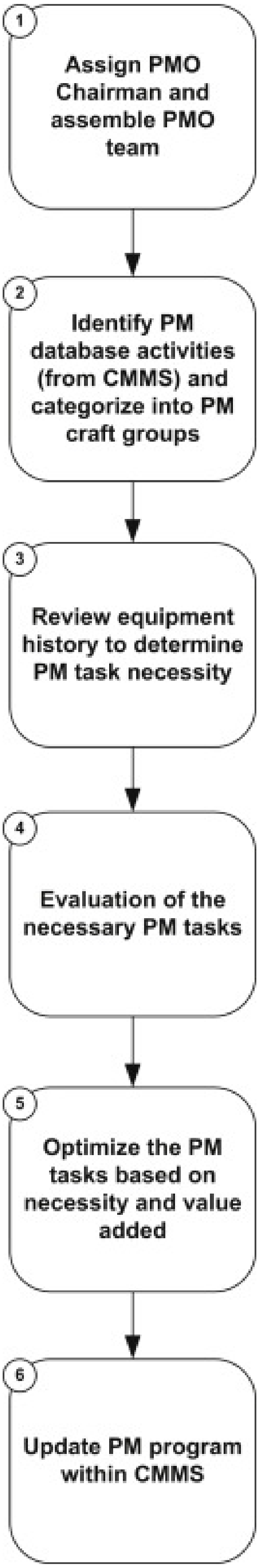


Figure 12: Planned Maintenance Optimization Workflow

# 3.Background on Railway Maintenance

A distinctive property of railway systems is that most activities are exclusive. Usually, you are unable to perform maintenance on the components and subsystems (track, power distribution, interlocking etc) currently involved in the train operation and vice versa. If there are redundancies (e.g., parallel tracks) it might be possible to perform concurrent operation and maintenance on neighbouring parts of the network, usually with some sort of restrictions (e.g., speed reductions, safety distances). This means that network services can still be offered during maintenance although the service level might be degraded (e.g., longer travel times, other routings). Rules and regulations (international, national, and company-wide) will set the limitations for how this can be done. Some countries might allow maintenance work to be carried out on a parallel track if the train traffic obeys a certain speed limit (e.g., Sweden), while others might not allow any adjacent train traffic at all (e.g., Holland). Some railway systems (e.g., certain subway or tram services) might be able to close operation completely during a couple of (night) hours while others must operate more or less continuously. Railway systems have some further complicating properties affecting both operation and maintenance, that are worth summarizing on this introductory level:

**Interdependency between infrastructure and trains.**
All rail guided transportation has a tight coupling between the fixed rail and the moving wheels, especially when having high weights and/or speed, metal-metal contact, and stiff axles (which practically all rail transportation has). The requirements and tolerances for the track are demanding, both regarding load bearing (including suspension distribution from rail via sleepers and ballast to the substructure), levelling (lateral and transversal), gauge and displacement. Furthermore, the rail surface quality has a crucial importance. All these properties affect riding comfort and degradation speed, both for the trains and the track. Equally important are the requirements and tolerances on the rolling stock (trains, locomotives, wagons, and motor units). Flat wheels, slippage, locked brakes, or bad roller bearings can cause extensive damages on the track (and thus indirectly affecting other trains). For electrified railways there are equally high demands on the power distribution, both electrically (substation capability, motoring, electric braking, disturbances etc) and mechanically (catenary wire, pantograph etc). Finally, this tight interdependency between infrastructure and trains exists for the complete infrastructure, all rolling stock and the whole transport chain, which makes railways unique when comparing to other transportation modes (shipping, air lines and road traffic).

**Geographic layout of the network and it is components.**
Equipment and crew must be transported to remote locations to perform maintenance. Some of these transportation activities must be done on the infrastructure itself and will thus consume traffic capacity. Furthermore, the different sub-systems will have different geographic layouts, e.g., signal interlocking will not always match the electrification system. Thus, a maintenance activity that requires a section of the electrification or interlocking system to be turned off will affect a larger part of the network than one that only requires a specific track or turnout to be blocked for traffic. The way the different sub-systems are partitioned will therefore greatly influence the level of serviceability and maintainability.

**Safety.**
Since trains have very long braking distances and run-on common tracks, the safety requirements must be high. Sufficient spacing and speed limitation must be guaranteed, both between trains, through turnouts and when approaching occupied or ending tracks. Similarly, safety is crucial when performing maintenance, both for guaranteeing the integrity of the work force as well as trains that can pass the work site.

**Organization and deregulation.**
Several different functional units are involved in making a rail transportation possible, including legislation, design, construction, planning, procurement, infrastructure and rolling stock maintenance, marketing, selling, operation, service, and education. Some enterprises, such as Indian Railways (employing about 1.3 million people), cover almost all these aspects - usually divided into geographical zones or traffic regions. Such large enterprises will always have organizational difficulties and cooperative problems not only due to size, but also due to conflict of interests, economic incentives etc. Primarily in Europe a far-reaching deregulation has been going on since the 1980s, with the overall purpose of opening for commercial competition in several of the working fields. We will not discuss the pros and cons of this development, only noting that new requirements and questions are raised, and that different roles and responsibilities need to be spelled out and made clear, which is usually beneficial. For infrastructure maintenance this trend has extended the use of maintenance contractors and hence the demands on contractual forms, public procurement as well as planning has increased. The organizational split between infrastructure manager (IM), railway undertakings (RU) and maintenance contractors is a common theme throughout this document.

# 3.1 Railway infrastructure maintenance

In this section we first describe the methodology used for collecting the information. Then we categorize the infrastructure maintenance activities followed by a sub-section describing the possessions, which grant access to the infrastructure in a safe way. Finally, we explain all the planning process steps, ranging from the very long term to day-of-operation.

## 3.1.1 Methodology

The material in this section is based on a series of unstructured interviews with planners, coordinators, technical experts, and managers that are involved in planning and performing infrastructure maintenance on the Swedish national railway system. Most of the reference persons comes from the Swedish Transport Administration but several contractors have also been interviewed. The meetings (real life or over telephone) did not have a fixed questionnaire but focused on the following topics: How are the tasks planned and performed, what are the preconditions and effects of the task, what type of equipment and crew is involved, how long possessions are needed/wanted/gotten, what kind of coordination’s are done, what are the costs, how large volumes of work is conducted, seasonal variations, suggestions for improvements etc. Each interview has been documented with written notes, which the reference persons has reviewed and corrected.

Rules, regulations, steering documents, and guidelines have been collected as well as some statistical material from the IT systems currently in use. Background information and basic facts have been collected from some railway literature and internet web sites. The results are biased towards the Swedish situation and a national railway system, but we try to present the results as generally as possible.

## 3.1.2 Maintenance activities

In this sub-section we describe and categorize all major maintenance activities that affects the train operation, normally by requiring possession time. We also make some notes about the contractual forms that are currently used in Sweden.

### 3.1.2.1 Categorization and overview

There are different ways of categorizing maintenance activities, based on

- If they are done before or after a fault has been detected, resulting in the classical split into preventive vs corrective maintenance. We will argue that this distinction can be hard to make in some cases and possibly not the best classification for our needs.
- What they consist of, which results in a more practically oriented categorization into diagnostic and restoring actions (further subdivided into technical systems or competencies required). This is often how the maintenance organizations think about and organize their work, e.g., having different crew groups working with inspections and corrections.
- How they are or can be planned, which results in a categorization suitable for the planning tasks. We will present one such categorization model and map the maintenance tasks into it.

**Preventive vs corrective**

The European standard EN 13306 for maintenance terminology use the terms preventive and corrective maintenance, for work taking place before and after a fault has been detected. Preventive maintenance is further divided into condition-based and predetermined maintenance, where the former uses measurements and inspections to determine when actions are needed, and the latter uses fixed maintenance intervals/schedules. In addition to these categories TrV sometimes use the term operational maintenance2 for activities that handle normal operational conditions such as snow removal, slippery rail etc (these activities may also be classified as corrective maintenance). Thus, we have the following categories:

- Preventive maintenance (before a fault has been detected)
  - o Condition-based maintenance, e.g., measurements and inspection, grinding, tamping etc.
  - o Predetermined maintenance, e.g., exchange of components (light bulbs, batteries, signaling relays etc) on specified intervals (usually specified by the manufacturer/supplier)
- Corrective maintenance (after a fault has been detected), e.g., fixing short circuits, repairing broken fasteners, welding, work after accidents etc.
- Operational maintenance, e.g., snow removal, handling slippery rail etc. Above we listed tamping and grinding as condition-based, preventive maintenance tasks. This is normally true, since they mostly are done well before any immediate action is needed. But sometimes the deterioration is faster than anticipated and corrective tamping or grinding is needed. Also, it should be noted that condition-based maintenance usually employs several intervention levels, stating the time frames for restoring actions. At the most serious level, immediate actions areneeded, and operative restrictions may be

imposed (lowered speed or train weights). Such immediate actions are considered corrective since the system is faulty (does not operate properly), although technically the components are not yet broken. These examples show that it's not always clear that one type of activity belongs to one of the above categories. In fact, this categorization is mostly used for contractual and budgetary/follow-up reasons. Diagnostic vs restoring actions the maintenance organizations tend to use a more practically oriented classification, where activities are grouped into.

- Diagnostic actions, which consists mainly of inspections and periodic measurements,
- Restoring actions, which consists of all repairs, exchanges, remedies etc.

Bundling all restoring actions together is however not very descriptive, since it will include everything from large track renewal projects to small repairs of insulation joints. Hence a further subdivision is necessary. The frequency of the diagnostic actions as well as the predetermined maintenance and some restoring actions (e.g., grinding) is determined by a set of factors, the most important being the amount of train traffic (volume, weight, and speed). Other factors can be surrounding environment, geotechnical standard, age etc. In Sweden the high-speed lines are safety inspected 6 times per year while low speed lines with low train weights may be inspected once a year or even less frequently. It should also be noted that inspection might be conducted simultaneously with preventive or prescribed maintenance, a practice that is adopted for catenary wire maintenance.

Inspections and measurements may result in remarks, which calls for some restoring action. The remark will have a time frame for when the restoring action should be performed. At TrV these time frame codes are immediate/acute (A), week (V), month (M) and before next inspection (B). Note that the time frame for the last code depends on the inspection frequency for the affected track section. Also, the time frames are somewhat flexible - for example "week" is normally considered to mean "within two weeks from inspection", "month" to mean "within one-three months" etc. Some of our reference persons consider these relative time frames a bad practice and would rather see that the inspection remark sets an absolute time limit for when the remark should be handled.

Note that the inspection remark does not specify the restoring action to be taken, just that a threshold limit has been reached and that some action is needed in order to restore the infrastructure to within the prescribed limits. If there are alternate actions, it is up to the maintenance contractor to decide which action that is appropriate. Sometimes there are more than one option, for example spot tamping with a small vehicle (cheap) or tamping a longer section with a dynamic stabilizing train (expensive). Whichever is chosen will be influenced by the contractual form and sometimes negotiations are taking place between the contractor and the IM before the final action is selected.

Table 2: Planning Horizon of each activity based on it's Possession Time

| Possession Time | Activity | Planning Horizon |
|---|---|---|
| >8h | Catenary wire replacement | 2-3 years / urgent |

| | | |
|---|---|---|
| | Track / turnout replacement | 2-3 years |
| 4-8h | Tamping of tracks | 1-2 years / 1 month |
| | Grinding | 1-2 years |
| | Switch replacement | 1-2 years |
| | Catenary inspection & maintenance | 2-3 years |
| 1-4h | Tamping of turnouts | 1-2 years / 1 month |
| | Ultra-sonic testing | 1-2 years |
| | Fasteners, joints, rail repair.. | 1-2 months |
| As Train slots | Periodic measurement | 1 year |
| | Fast grinding | 1 year |
| 0-1h | Inspection | 0-2 month |
| | Signal repair, vegetation etc | 0-2 month |
| | Slippery rail, snow removal | 1 year / 0-1 week |
| 1h- xdays | Accidents, urgent repair | none |

Capacity usage and planning horizon from a planning perspective we may also categorize the actions according to how much infrastructure capacity they consume and how long in advance they are planned. Ideally the highly disruptive actions (requiring very long and exclusive access to the track) should be planned long in advance while actions requiring less possession time can be planned in later stages. In table 1 we list the different activities according to the needed amount of possession time (per work shift) and how long in advance the planning can be done.

Problematic cases are those that require long possession time but have a short planning horizon:

- Catenary wire replacements are highly disruptive and can usually be planned well in advance. But the degradation can in some cases be very quick, especially if the pantographs of the trains are worn, which may lead to urgent need for replacement or repair.

- Tamping can normally be planned well in advance but in cases where preventive maintenance have been neglected tamping might be needed with short notice.
- Rail repair are often due to problems with cracks found by the ultra-sonic testing. In severe cases the possession time might be 4-6 hours.
- Accidents and urgent repair may result in possessions ranging from hours up to several days.

Most maintenance activities are planned and scheduled in the capacity planning processes, such that they are included in the (daily) operational timetable which is handed over to the traffic control centres. Some activities - indicated with a planning horizon of zero in table may however be carried out directly in the operative phase. Two types of activities are handled in this way:

- Accidents and urgent repair, which are triggered by external events and are managed by the traffic control center.
- Small and quick maintenance tasks, which are triggered and managed by the maintenance contractors. These tasks can either be secured by
  - a possession, given directly by the traffic dispatcher, or
  - a manual train warning procedure handled by the work force itself. In either case the tasks must be of such nature that it can be ended/cleared with a very short notice, typically within a couple of minutes (when secured with a possession) or less than a minute (when using a manual train warning procedure). The latter handling is hazardous, especially for sections with high-speed traffic, and the use of it is discouraged.

#### 3.1.2.2 Contractual forms

The Swedish railway maintenance market has undergone a quick deregulation since 2001 (Trafikverket, 2012). Today all railway maintenance work on the national infrastructure is performed in contracts signed after open competitive tendering. Three types of contracts are used:

- Re-investment projects, where parts of the infrastructure are re-established to its intended standard. Typical examples can be the replacement of catenary wires, track sections and turnouts. A contract is signed for each project.
- National maintenance contracts. These contracts are used for activities that are performed with a limited but expensive and/or highly specialized set of equipment and crew that operate over the whole infrastructure network. Typical examples are the periodic track geometry measurements, ultra-sonic testing and grinding. The contract lengths are 3-5 years with a prolongation of 2 or 2+2 years. One contract is set up for each distinctive type of activity.
- Regional maintenance contracts. These contracts are used for all remaining activities including inspection, predetermined activities, corrective maintenance etc. The contract length is currently 5 years with a prolongation of 2 or 1+1 years. The Swedish infrastructure network is divided into 34 such contracts. TrVreevaluate and revise the contractual forms continuously, but the tendency is to move more responsibility to the contractors. This requires revising the specifications from a detailed component/technical level to a system/functional level.

### 3.1.3 Possessions

All activities that require secure access to the railway infrastructure must obtain a (work) possession.

- A possession shall guarantee that no trains will run on the designated area - usually coinciding with a signaling stretch such as a line blocking, one or more tracks on a station or a complete station between the entry signals. In addition, the possession may impose restrictions on neighboring tracks, usually such that passing trains can only run past the work site with a reduced speed. Depending on the capabilities of the train control system and how temporary speed restrictions are communicated such speed limits may have to be imposed for a longer time period (e.g., one or more days) than the actual possession lasts. In addition, there might be speed restrictions after a work has been carried out until a "burn-in/settlement" period has elapsed. Note also that all temporary speed restrictions affect the train operation and should be communicated/negotiated with the train operators if the timetable has to be altered
- Each possession is given a unique work id, in the same way as train paths (slots) receive unique train numbers. Just as with train numbers the possession can span several days.

The possession area may be loosely given in the early planning stages. As an example, a budgeted track renewal project may have a possession between stations A and D for 8 hours every night for two weeks, but the exact work location for each night cannot be detailed until the project has been established and planned - some months before the actual work. In fact, this is a challenge for the IM - to dimension appropriate possessions for the early capacity planning applications, which will constrain the contractors that plan and perform the actual work. Effectively the "possession design" will impact the project cost for investment and re-investment projects. Furthermore, the exact effect on the train traffic can be hard to envisage for such loosely specified possessions (also when the future timetable is unknown). Simulation is a technique that can be used to quantify the capacity restrictions and find suitable runtime margins and slacks that should be added for train slots affected by the possession. Some possessions are due to civil engineering work performed in the vicinity of the railway, for example on bridges, tunnels, or buildings such that train traffic is not possible. These cases offer an excellent possibility to perform railway maintenance in parallel with the surrounding engineering work.

As noted previously, the sectioning of the signalling system and the power distribution network will determine how much of the infrastructure that will be affected by a possession. If the power must be turned off, no electrified traffic will be possible within the same power section. On traffic lines, this is of little importance since the power sections usually follow the signalling. But on stations and marshalling yards several parallel tracks will belong to the same power section. In addition, all tracks that become dead-ended due to the possession will have very little use for the train traffic (other than close to the start and end time of the possession) - also giving an opportunity for coordinated maintenance work.

Hence the possession should not only describe the work area needed (tracks, signals etc) but the whole affected traffic area. Ideally the maintenance contractor should only need to consider themselves with what components/objects they need access to (and the time needed), while the planning system would keep track of all surrounding objects that should be included in the possession as well as the restrictions imposed on the adjacent traffic tracks.

## 3.2 Planning process

Here we describe the complete planning process for obtaining possessions and train operation slots as it works in Sweden today. The work process follows EU guidelines and will probably be quite similar throughout the European countries up till the publication of the yearly timetable. The subsequent steps might differ more between different countries. The process can be divided in the following steps:

- Freight corridor planning, where so called prearranged paths (PaP's) for the international freight trains are established and coordinated with the major possessions (large infrastructure maintenance activities).
- Preparation and publishing of the network statement, which shall contain all major possessions that the train operating companies should adhere to.
- Yearly timetable planning, where the regular timetable for all train paths is planned together with the major work possessions.
- Timetable revision planning, where all dated timetable adjustments are made, and final coordination of train paths and possessions should be done.
- Planning of minor possessions, where plannable work which do not require any train path adjustments are scheduled.
- Operational planning and control, where the traffic control center will make operative adjustments, authorize unplanned possessions, and control all activities (train runs and work) on the railway infrastructure.

Steps 1-5 make up the capacity planning process, while step 6 is the operational phase. In steps 1-4 timetable adjustments and conflicts between different requests are handled, while in step 5 only requests for "spare" capacity should be handled. The handover between step 5 and 6 happens one day before the operational day at TrV.

We will now spell these six steps out in greater detail, describing the process as it is intended to work according to guidelines and regulations, while noting any known deviations from the target process. We make a distinction between:

- Major possessions, which will (or is likely to) be in conflict with one or more train paths and hence requires coordination (handled in steps 1-4)
- Minor possessions, which do not affect the published train paths (handled in steps 5-6)

Whether a possession is major, or minor depends on several factors, such as the possession area and its duration, the time-of-day, the train traffic patterns and whether a published timetable exists or not. A very short possession can be considered major as soon as there is a conflict with a (wanted or scheduled) train path. Conversely a possession of several hours could be considered minor if no train paths will run on that part of the infrastructure the same day. The IM will use rule-of-thumb or a specified criterion for which possession to consider in which planning step. Day-time possessions are usually more severe than on the night or over the weekend, while the work cost follows the opposite pattern. Freight corridor planning This process should follow the guidelines given in RailNetEurope (2013b). Several rail freight corridors (RFC) exist in Europe and to secure a stable rail freight service across them, the traffic is run on so called pre-arranged paths (PaP). These PaP belong to the RFC and have priority over the regular (national) timetable. They are planned roughly one year ahead of the yearly timetable planning and the RFCs should organize two coordination meetings per year: in November and in May, where all concerned IM's ad RU's shall participate. Bigger possessions should be coordinated early while lesser ones may be handled in the later stages of the process.

## 3.3 Maintenance of Railway Track

The maintenance of railway track is a complicated and expensive task which represent an important share of total railway infrastructure costs. As a matter of fact, the maintenance of the track represents around 40% of the total maintenance cost of the railway system.
The state of track depends on many factors such as the characteristics and age of the elements, the track geometry, topography and geology, weather conditions and supporting loads. Track maintenance is still a very little automated process, relied on the skills of specialised human operators and based on rules established a long time ago (preventive maintenance) complemented by the execution of on-call corrective tasks whenever there are faults in the system.

Furthermore, the saturation of the capacity of the track sections because of increased load of rail services requires intensified maintenance and the planning and coordination of the rail activities, to accommodate maintenance tasks to the availability of time windows needed to guarantee technical regulatory levels. Thus, as railway uses increases so does the need for maintenance while the availability of the track for maintenance decreases. Therefore, the work is mostly carried out outside of daylight conditions and under pressure, increasing the risk of staff accidents. Finally, the maintenance management based on cyclical preventive works and on corrective maintenance entails high costs in both resources' reliability and availability of infrastructure.

This situation requires the streamlining of the maintenance management based on monitoring the track condition, automating the planning management, and especially monitoring the evolution of the parameters that determine the track condition for predictive maintenance and risk analysis. This schema would allow evolving the maintenance management model based on corrective/preventive maintenance into a model based on conditions/predictions, helping those responsible for making decisions to achieve optimal maintenance plans that minimise the maintenance costs, ensure a satisfactory safety margin and prevent quick degradation of track quality.

As regards railway structure, it is possible to distinguish the superstructure and the subgrade.
The superstructure, which supports and distributes train loads and is subject to periodical and maintenance and replacement, consists of:

- The track,
- The track bed

The track consists of:

- The rails, which support and guide the train wheels.
- The sleepers (also called ties, mainly in North America), which distribute the loads applied to the rails and keep them at a constant spacing.
- The fastenings which ensure the rail-sleeper connection.
- The switches and crossing.

The track bed consists of:

- The ballast, usually consisting of crushed stone and only in exceptional cases of gravel. The ballast should ensure the damping of most of the train vibrations, adequate load distribution and fast drainage of rainwater.

- The sub-ballast, consisting of gravel and sand. It protects the upper layer of the subgrade from the penetration of ballast stones, while at the same time contributes to further distributing external loads and ensuring the quick drainage of rainwater.

The subgrade, on which the train loads, after adequate distribution in the superstructure, are transferred and which in principle should not be subjected to interventions during periodical maintenance of the railway track, consists of:

- The base, which in the case of the track laid along a cut consists of onsite soil, while in the case of an embankment is composed of soil transported to the site.
- The formation layer used whenever the base soil material is not of appropriate quality.

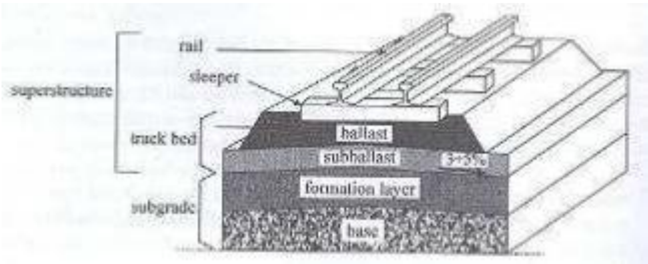


Figure 13: Superstructure and Subgrade

The track usually lies on ballast which provides a flexible support. It is referred as ballasted track.

However, it is possible, that the track is supported by a concrete slab, instead of ballast. In this case, the support is inflexible, and it is called slab track. Although a slab track is used in certain railways (e.g., the Japanese and the German, among others), it is most effective when used in tunnels, because it allows a smaller cross-section and facilitates maintenance. In most of the tracks worldwide, a ballasted track is still the case, as it ensures flexibility (an important factor in the event of differential settlements) and much lower construction cost, while at the same time offering a very satisfactory transverse resistance, even at high speeds. The problem of noise, which is much greater with the track on concrete slab than with the track on ballast, should not be disregarded. When a slab track is applied (e.g., in the case of a tunnel), the sudden variation in track stiffness (felt by passengers as a jolt) is lessened by placing rubber pads of a suitable thickness along the tunnelentrance and exit.

The choice between ballasted and non-ballasted track should consider construction cost (much greater for non-ballasted track), maintenance cost (much greater for ballasted track), together with technical requirements. Both solutions have pros and cons.

### 3.3.1 Track Defects

This section describes most representative defects and failures that can be found in tracks. Track defects is defined as the deviations of the actual form theoretical values of the track's geometrical characteristics. Track defects are consequence of train traffic, they are of a macroscopic and geometric nature and usually they are rectified by track maintenance.

Geometry track defects include:

- Longitudinal defect

- Transverse defect
- Horizontal defect
- Gauge deviations
- Track twist

The longitudinal defect (LD), is defined as the difference between the theoretical and the real value of track elevation and is given by the equation:
$LD=Z^{th}(T,x)-Z(T,x)$

The longitudinal defect is the most reliable in illustrating the effect of the vertical loads on track quality and it is the principal factor (together with the transverse defect, see below, which accompanies longitudinal defects) in determining the magnitude of the track maintenance expenses.

The transverse defect (TD), is defined as the difference between the theoretical and the real value of cant:
$TD=(Z_{int}-Z_{ext})th-(Z_{int}-Z_{ext})$

For rectilinear parts of track layouts, where curvature is zero, the transverse defect is the difference of elevation between internal and external rail: $Z_{int}-Z_{ext}$

**Horizontal defect**
The horizontal defect (HD) is defined as the horizontal deviation of real position of the track from its theoretical position. The horizontal defect depends on the transverse track effects (more than the two previous types of defects) and on the characteristics and particularities of the rolling stock.

**Gauge deviations**
Certain track gauge deviations, affected by the mechanical properties of track materials and the particularities of the rolling stock, are permissible and will be given below. Gauge values acceptable for standard gauge tracks are given in each track line.

**Track twist**
Along straight and circular sections (where cant is constant), four point of the track lying on two transverse sections must lie in the same plane. Track twist is defined as the deviation of one point from the plane defined by the other three.
If i and i+1 are two successive transverse sections of the track, spacers $\Delta l$ apart (e.g. at the positions of two sleepers), track twist is defined as the variations of the transverse defect per unit length.
$Tracktwist = (TD_{(i+1)}-TD_i)/\Delta l$
The risk of derailment is prevented when the real value of twist is smaller than its critical value causing derailment, which depends mainly on speed and to a lesser degree on the type of the track equipment and of rolling stock.

**3.3.1.1 Rail Defects**

The rail suffers from stresses that can cause defects and may bring it to failure. The total stresses developed in the rail are the sum of:
- Stresses at the wheel-rail contact,
- Stresses resulting from rail bending on the ballast,

- Stresses resulting from bending of the rail head on the web,
- Stresses resulting from thermal effect,
- Plastic stresses, remaining in the rail after the removal of external loads.

Those stresses gradually decrease the mechanical strength of the rail due to repeated loading. That is called fatigue. Once the fatigue limit is reach, the rail is bring to failure. The effects of these stresses could be:

- Plastic deformation. The rail support high stresses during the train circulation. If those stresses are greater than the elastic limit, a flange can appear in the rail head, because below this flange, the stress limit may be exceeded.
- Rail wear. The traffic load produces the rail wear that affect the rail profile. There are two type of rail wear:
    - Vertical wear. It reduces the rail section and consequently the rail resistance. The maximum permissible vertical wear of the rail is a function of the maximum train speed and of traffic load. There are regulations, in different countries, to limit the maximum permissible vertical wear of the rail.
    - Lateral wear. It reduces the rail section and consequently the rail resistance and affect to the gauge of the track too. As in case of vertical wear, the maximum permissible lateral wear of the rail is a function of the maximum train speed and of traffic load. There are regulations, in different countries, to limit the maximum permissible lateral wear of the rail.

Surface defects can be distinguished in:

- Short-pitch corrugations Their cause is train traffic and they consist of corrugations with a wavelength $\lambda$=3-8 cm. They can provoke many adverse effects: high frequency oscillation of the track, including resonance, and leading to higher rail stresses, concrete sleeper fatigue with cracking in the rail seat area, loosening of fastenings, accelerated wear of pads and clips, premature failure of ballast and the subgrade, and increase by 5+15 dB in noise level. This defect is detected either visually or by appropriate recording equipment. It is repaired by passage of special equipment, which grinds and smooths the rail.
- Long-pitch corrugations They have wavelengths $\lambda$=8-30 cm and occur mainly on the inside rails of curves having a radius of 600 m and smaller. This kind of defect is the most common on suburban and underground railways. Detection and repair processes are similar than those for short-pitch corrugations.

## 3.3.2 Track Maintenance Operations

This section describes most representative maintenance and renewal tasks to be performed on the track.

### 3.3.2.1 Rail Maintenance

There are four methods to eliminated rail defect:

- Rail weld recharge.
- Rail grinding.
- Rail replacement.
- Rail tamping.

**Rail weld recharge**
It is a technique for the cost-effective repair of discrete defects on the running surface of rail. The results archived are very much depending on the expertise of the operator.
Repair welding is the most cost-effective method of repair of defects of the tread and guide surfaces of railway tracks and switches.

Two methods for welding repair could be used:

- **Manual weld**. Manual or automatic welding process with consumable electrode is used to repair railway tracks. The welding technology is developed and monitored by the corporate expert welding staff. The completed works are continuously checked by the use of non-destructive test methods because the quality of work is function of the welding staff capacity. These processes are used in contact surface to repair defect at the initial stage and it is always necessary to make a grinding in the weld zone after the weld process.
- **Union welding**. This process is similar to joining a continuous welded rail. As in Manual weld, it is always necessary to make a grinding in the weld zone after the weld process and before the end of the repair works.

**Rail grinding**

Rail is the single most valuable asset of railways. The wheel/rail interface of any railroad is a sophisticated and much talked about subject. Primarily because of the cost involved in premature rail change-outs. Typical problems encountered on all railroads include shelling, spalling, side wear, plastic flow, dipped welds, corrugation and fatigue, as well as unique challenges of noise control and ride quality.

Rail grinding is considered the single most effective maintenance practice to control the effects of rolling contact fatigue, restore profile, and maximize value from the rail asset.
The substantial return on investment from rail grinding is well documented and includes:

- Extended rail life
- Fuel savings
- Reduced surfacing cycles
- Extended track component life
- Reduced wear on rolling stock
- Increased axle loads
- Increased train speeds
- Improved ride quality and passenger comfort.

Railroads everywhere are facing continued challenges of maintaining track in shortened work blocks with limited resources. A proper rail grinding program is a key component to a maintenance plan.

Rail-grinding equipment may be mounted on a single self-propelled vehicle or on a dedicated rail grinding train which, when used on an extensive network, may include crew quarters. The grinding wheels, of which there may be more than one hundred, are set at controlled angles to restore the track to its correct profile.

The machines have been in use in Europe since the early 1990s. They are made by specialist train maintenance companies who may also operate them under contract.

The rail grinding could be made in large dimensions of rail line or only in a short location in a rail line. There are different types of machine to adjust the rail profile to theory profile to short or to large grinding work.

**Rail replacement**

When the rail defects are very severe and the rail profile cannot be recovery by other process, it is necessary to replace the rail in the damage section. In this case it is always necessary to use welding to join the new profile of rail with the existing one. After that, some local grinding on the join is required to achieve appropriate rail profile.

**Rail Tamping**

Tamper machine can restore the ballast, the initial sleeper's position, the rail geometry, and the vertical rail deviation. The tampering activity obtained by introducing vibrating blades in the ballast. The tamping is performed using a train called tamper train. In addition, the restoration of the ballast may be executed by refilling stones. The speed of the tamping and ballast cleaning machines is around 0.4-1 km/h.

# 3.4 Railway Track Monitoring

This section focuses on actual measurement techniques to evaluate the track condition. If measures are within the threshold limits, then train operation is safe and complies with an adequate level of comfort. Otherwise, maintenance tasks should be carried out.
The maintenance of the track requires the inspections on the track. Such inspections are mainly visual performed by operators walking along the track, sometimes after user's alerts.
This section describes several measurements techniques used to evaluate track defects. These techniques are based on ultrasonic, laser and cameras.

The geometric state of the track is evaluated by the control of some geometrical parameters, for which the railway network regulations establish the permitted values. Analysing the measurements for these parameters and comparing them to the threshold values, it is possible to assess the geometric quality of the track, according to which corrective actions and driving restrictions are programmed.

## 3.4.1 Measurement of Rail Defects

Measurement's techniques widely used to evaluate rail profile are: Ultrasonic and Laser.
Ultrasonic inspection (UT) of rails is based on using multiple UT probes with different angles for maximum coverage of railway section. The results from each probe are stored with the respect to its position along the rail, depths, and amplitude. The map created in such case provides a pattern for each artefact laid in the rail: ball holes, drilled holes, and, for sure, the defects.

The advantages of the UT rail inspection of the high inspection speed, the very good coverage of rail cross section, and the minimum time possession due to use of high-speed inspection wagon (up to 80-100 Km/h).

The disadvantages of UT inspections are:

- relatively high cost
- the evaluation of the final results depends very much on the experience of the operator

Large amount of water is required to be used as coupling material for UT. Some of modern systems are equipped with the special wheels where UT probes are installed inside of it. Such systems require much less water consumption, though the cost of wheel and speed limitations is still an issue.

The method of laser triangulation is used for rail profiles and gauge measurements. The instruments are mounted on the bogie of a measurement train. These instruments are also increasingly used on grinder trains, to improve the process and to eliminate operators walking along the track to verify the result.

The rail profile measurement for verifying the wheel-rail contact configuration requires a much higher accuracy than the measurement for assessing the rail wear. Accurate instruments allow the computing of important parameters, like the equivalent conicity, for entire lines. These parameters are fundamental for the global modelling of the train-track system. These parameters are well defined in the UIC519 standard. The UIC518 standard, one of the most important ones to verify if a train is suitable for running, also defines limit values for these parameters.

Corrugation measurement is normally based on laser triangulation techniques. Rails corrugation is a source of vibrations and noise and must be eliminated by grinding. Different techniques have been developed to:

- Just detect the corrugation, and/or
- measure the corrugation amplitude. An amplitude measurement allows a more accurate planning of the works, because it allows a better evaluation of the necessary grinding time.

Nevertheless, the corrugation detection gives a rough evaluation of the amplitude, thus the result may not be very different.

Corrugation detectors are based on accelerometers or on the measurement of the primary suspension movement (the bogie mass is supposed to filter the corrugation out). Those accelerometers determine the position of rail defect while the laser measurement evaluates the corrugations amplitude.

## 3.4.2 Monitoring Train

To evaluate track geometry, some measurement instruments are normally mounted on two types of vehicles:

- Special (and expensive) measurement trains
- Low speed, inexpensive vehicles

Small vehicles (often two axles ones) loaded with track geometry evaluation instruments can also run along the track. More sophisticated, complex, and expensive track inspection vehicles are also owned by railway administrators to cover the rail network. These vehicles perform comprehensive inspection of the track. They run on more important lines and its schedule is planed with enough anticipation.

Measurement trains are normally managed centrally and distribute the results to a central data base and to the regional maintenance offices. On the other hand, low speed, inexpensive vehicles are normally assigned to regional maintenance office to allow prompt measurement in critical area, checking the works done by contractors, etc. It is important to recall that the measurement train could be far away, and it could have its own work plan. Changing the work plan and moving the train would be expensive and inefficient.

Therefore, most of the infrastructure inspection is currently carried out by the so called "measurement trains". A measurement train carries a wide variety of instruments. Different railways have adopted very different configurations. Some maintenance actors prefer having quite "specialized" trains, e.g., one for the track, one for the catenary, one for the signalling and telecommunications.

Others prefer integrating everything on a single train (to take better advantage of every measurement run). In any case, these trains are running on the entire network of an administration on a planned, routine, basis, delivering the data every N days (where N normally depends on the line class).

The new idea is to use trains in commercial service to perform inspections on the track is taken more adepts every day. This way to carry out maintenance would save cost based on the following facts:

- expensive measurement train would not be required.
- cost of performing the operation (crew, traction, etc.) would be saved.
- slots for maintenance operations, difficult to find on busy lines, would not be required.

Therefore, the track availability for service will increase.

Moreover, because the train would be on service, normally going up and down the same line every day, the frequency of the measurements would be high.

To perform inspections of the track using on-service trains, many of them must be equipped with the instruments to cover the network, normally one equipped train for each line. Moreover, the localization equipment becomes very critical: special techniques must be used the GPS, even where available, does not provide the necessary resolution. The cost of the instruments is more than compensated by the savings in the train and relevant operational expenses.

Obviously, such an automatic system is complicated to design, build, and manage. The reliability of every component must be much higher than the reliability normally accepted for a measurement train. A large-scale experiment on this subject started years ago in the UK (about 30 trains to be equipped), eventually it failed, probably due to unreliable components and wrong software architecture. Other experiments are known, but none is in commercial operation at this time.

Evaluating track condition using on-service trains can deliver very useful data for an improved maintenance strategy:

- Better trends (more accurate)
- Immediate verification of the works
- Early detection of unpredictable faults

Data collected by any of the above means (visual ones or by monitoring equipment) can be reported in paper or in electronic format. Severe faults are reported at once and immediate actions are taken. Other non-urgent defects are delivered to the central/regional office where the defects are evaluated by rather simple algorithms and ranked by severity indexes.

### 3.4.3 Inspection and Monitoring Technologies

The inspection and monitoring technologies can be divided in:

- Inspection technologies laid along the track.
    - Fibre optic sensors laid along the track and other infrastructure elements (such as bridges)
- Inspection technologies that could be embarked on commercial trains.
    - Hollow shaft integrated acoustic sensor system.
    - Rail monitoring sensor combining eddy current distance measurement with acceleration data.
    - Laser profiler and inertial pack to monitor the track geometry.
- Inspection technologies embarked on special testing trains.
    - Ultrasonic non-destructive fuzzy inspection techniques
    - Non-contact thermography system for rail surface monitoring.
    - Visual camera

The first four techniques above are the most promising because they allow the Railway Infrastructure Manager keeping updated information daily on the track state in a very cost-effective way. Technologies 2nd, 3rd and 4th above could be embarked on commercial trains obtaining an automated and unattended measurement system. This is a step forward in track inspection technologies which will reduce the need for expensive instrumentation trains inspecting the infrastructure during the night when there are not rail services and will increase the capacity of the rail transport. Moreover, the availability of frequent and quality data on the track condition makes possible the condition-based maintenance based on daily updated measurements.

At current state of development, technologies 5th and 6th above can only be run on trains at low speed which make them unsuitable for commercial trains.

In the following paragraphs, these technologies are described more in detail.

- **Fibre optic sensors**

In the railways field, distributed optical fibre sensors can be employed for spatially continuous monitoring of the track's temperature and deformation, as well as the monitoring of the structural integrity of infrastructures as tunnels, bridges, and embankments. To this purpose, a single-mode optical fibre cable must be attached to the track and/or the structure under test in order to detect both tensile and compressive strains. On the other hand, for temperature monitoring purposes, the optical fibre cable must be deployed in such a way to assure a good thermal contact with the structure but without allowing any strain transfer from the fibre to the structure itself. Different kinds of adhesives must be employed to match the above conditions. Laboratory and field tests must be performed to select the most suitable optical fibre cables and bonding protocols for railways monitoring purposes.

- **Acoustic Inspection Techniques**

While ultrasonic techniques are typically using the system response to active excitations, acoustic methods only “listen” to natural sound sources like the rolling noise. If hollow shafts are available (e.g., in some high-speed trains) hollow shaft integrated acoustic sensor systems can be used to detect defects in wheel sets of the rolling stock. The system could include acceleration sensors as well as structure-born sound sensors and uses wireless real-time data transmission. The acoustic part of the system detects and evaluates acoustic signals generated by the rail-wheel contact. Based on a special averaging technique periodic scattering contribution from defects inside the wheel set can be efficiently detected and evaluated.

- **Inspection using pulsed eddy currents.**

Eddy current measurements have been a standard technique for a long time for finding cracks in metals either on the surface or within the material. Especially for nonmagnetic materials it is often the only technique employed to test materials in field service operations (e.g., critical parts of airplanes are checked before every take off for cracks using portable eddy current equipment). Recently, eddy current sensors have also become a common method for rail inspection. However, the probe must be either in contact or very close to the surface (<1mm). Therefore, usage in rail inspection has for a long time been limited to hand-held system or system mounted on manually driven trolleys which are offered on the market by several companies. In recent research projects, the concept has been adapted to test and to repair vehicles without changing the basic premise of short distance between rail and sensor. To adapt the method of eddy current testing to a train borne platform mounted on a commercial train the distance between sensor and rail must increase significantly. While eddy current can achieve high resolution and also quantitative assessment of crack depth, for regular rail inspection reduced performance of the system could be tolerated while at the same time increasing the frequency of inspection which would be possible using commercial trains. Initial experiments carried out by Siemens have shown that eddy currents of lower frequency are quite able to detect larger cracks even when probe-surface distance exceeds 10mm.

- **Laser profilometer**

Profilometer is a measuring instrument used to measure a surface's profile, to quantify its roughness. Vertical resolution is usually in the nanometre level, though lateral resolution is usually poorer. An optical profilometer is a non-contact method for providing much of the same information as a stylus based profilometer. There are many different techniques which are currently being employed, such as laser triangulation, and confocal microscopy.

- **Ultrasonic inspection**

Ultrasonic techniques belong to the most commonly used non-destructive methods with a wide variety of application fields. They are mostly used in pulse-echo mode so that only one-sided access to the structure under investigation is necessary. In most cases a broadband pulse is excited by a piezoelectric transducer and is send into the structure using an appropriate coupling agent like water, oil, or viscous paste. The waves interact with interior defects and are reflected so that they can be detected by a sensor. The latter can be either the same transducer that was used for excitation or an additional sensor. Ultrasonic techniques are well known for the inspection of the rolling stock of high-speed lines in Europe and abroad.

Typical inspection systems today are capable of testing rail wheels ranging from 680 mm to 1250 mm in diameter; it takes about four minutes to inspect a wheel, while the operations are performed five days a week on three shift schedules. Test trains typically include ultrasonic and eddy current systems to automatically scan the rail during run of the train (<100 km/h). The procedure is usually organised in a three-tier inspection process. The first tier, fast mapping of the rail is performed by the inspection car traveling at high speed on the track.

Once data is recorded and stored it is analyzed offline. The analysis or processing can identify and categorize flaws in a scan. The processing step issues a report which contains a list of all suspect flaws, their location in the scan and their distance from the nearest reference points. This scan report serves the repair team who returns to the relevant section of rail and must locate the detected flaw and verify it prior to maintenance operations on the rail. Such approach reduces the amount of time the track is blocked by ultrasonic inspection process, but on the other side it relates very much on experience of analysis team, and increases the time until final results with defect classification will be available for maintenance planning.
Despite the problems mentioned above automatic test trains or test vehicles still provide useful information about the track and rail condition. One main goal is therefore to significantly improve and enhance the present techniques. In this context the coupling conditions, lift-off phenomena, and the treatment of noisy signals are among the most challenging aspects.

- **Thermographic inspection**

Infrared thermography also belongs to the well-known non-contact non-destructive techniques. The material under test is first heated by a flash lamp or an inductive technique. After that, the spatial-temporal evolution of the thermal field is monitored by an infrared camera. If defects are present, the thermal conductivity is locally decreased so that “hot spots” of higher temperature can be detected. From the temporal change of the thermal field additional information about depth and size of the defect can be determined in principle.
In earlier investigations this technique has already shown its high potential for the characterization of typical flaws in rails. It could be shown that this technique principally allows the characterisation of the rails with a high sensitivity and a high testing speed. Just like ultrasonic and electromagnetic techniques a thermography system can also be integrated in a testing train. With the current hardware an automated testing of the rails and automated defect recognition at speeds up to 20 m/s (about 70 km/h) seems to be possible.

- **Inspection using visual cameras**

The main goal for using images of the track is to eliminate, or reduce as much as possible, the visual inspection done by workers walking along the track to detect any fault, missing components, etc. The state of the art of these instruments does not yet allow a complete and safe elimination of the inspection done by humans, but helps a lot and, also, allows detecting several risky situations difficult to detect by the human eye.
Several linear cameras are mounted under the vehicle.

The linear cameras are space triggered (e.g., every 1 mm). Every image is equivalent to a single line of a normal camera. Assembling all this line in an endless sequence gives a stream image of the track: an image having:

- on the transverse (Y) axis, as many pixels as the camera,
- on the longitudinal (X) axis, as many pixels as the number of mm travelled by the train.

A colour image is normally used to allow a human inspector viewing the track as if he were walking, but with obvious advantage for safety and line capacity. Sometimes it used for the faulty fasteners automatic detection, by machine vision techniques.

A black and white image is normally used for detecting rail surface defects and for every automatic analysis by machine vision techniques (rail surface, fasteners, sleepers, joints). The automatic analysis is useful to focus the attention of the workstation operator, who then goes to examine the relevant colour image and decides the relevant actions.

# 3.5 Track Degradation Models

The railway track and infrastructure degrade with age and usage and can become unreliable due to failure. When a failure occurs, the consequences can be significant, including a high cost of railway operation, economic loss, damage to the railway asset and environment and possible loss of human lives. Unreliability may also lead to annoyance, inconvenience and a lasting customer dissatisfaction that can create serious problems for the company's position in the marketplace. An applicable and effective maintenance strategy can guarantee the achievement of reliability goals and compensate for unreliability.

Maintenance actions are used to control the degradation of the track, reduce, or eliminate the likelihood of failures, and restore a failed part to an operational state. It is necessary to model track degradation behaviour to select an applicable and effective maintenance policy but modelling and predicting the track geometry degradation is a complex task, requiring the following information:

- the interaction of different track components,
- the effect of maintenance actions on track quality,
- the heterogeneity factors e.g., environmental factors, soil type and condition.

In addition, higher demand for railway transportation creates an essential requirement for higher speed and axle load which accelerates the track aging process and negatively affects its reliability.

The increased demand and complexity dictate the need for comprehensive track degradation models.

In the two last decades, a great deal of research has been done in the field of track geometry degradation modelling. Determining an indicator to represent track quality is an essential prerequisite for modelling track degradation. Different indicators are used based on the aim of the research. The indices for representing track quality condition are demonstrated in Fig. 14.

Sadeghi et al. proposed a track geometry index uses the following track geometry parameters: alignment, profile, twist, gauge, and rail cant. Using justified coefficient, they combined the parameters to design the track geometry index.

To consider structural defects, Sadeghi et al. proposed a quantitative track structural quality index. This index is defined for each track component group, i.e., rail, sleeper, fastener, ballast.

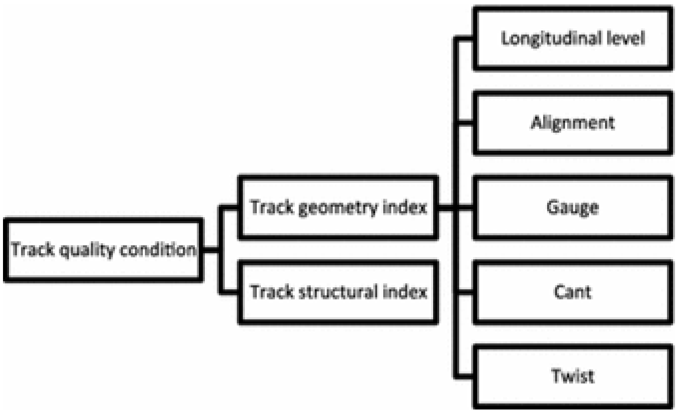

Figure 14: Track Condition Measures

Faiz et al. studied the geometry parameters used in the UK track maintenance process and applied linear regression analysis to explain their correlations. A Generalized Energy Index (GEI) instead of a Track Quality Index (TQI) for track quality evaluation is proposed by Li et al. The GEI can consider different track irregularity wavelength and speed. Haifeng et al. proposed an integral maintenance index (IMI) that considers the distribution of track geometry parameters to evaluate track condition. El-Sibaie et al. developed several track quality indices to evaluate track quality condition in relation to different track classes.
By looking to the literature, it can be observed that most of the researchers considered short wavelength longitudinal level as the crucial factor in degradation modelling. This issue can be seen in Figure.

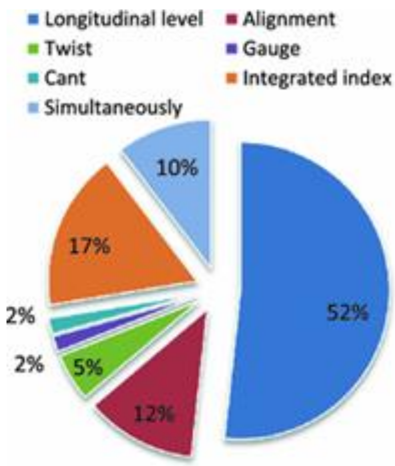

Figure 15: Distribution of Applied Track Geometry Measures

After finding the proper track quality measure, a degradation model must be constructed and the effect of different maintenance strategies on track degradation evaluated. There are two major approaches for track geometry degradation modelling, i.e., mechanistic, and statistical approaches. In this thesis, statistical approach is the focus.

Concerning mechanistic approach, several researchers tried to find the interactions among track components and their influences on track geometry degradation.

The most important models are those proposed by Shenton, Sato, Chrismer et al. Öberg et al., and Zhang et al. Dahlberg also provide an extensive review on mechanistic models applied for track geometry degradation.

Concerning statistical approaches, the most applied methods are summarised in Figure.
Andrade et al. assessed track geometry degradation and the uncertainty of degradation model parameters. They considered a linear model for track longitudinal level degradation. They performed statistical correlation analysis for each group section and fitted the log-normal distribution to the track's longitudinal level degradation. A multi-stage linear model is applied by Gou et al. to cope with different phases of degradation between two consecutive maintenance interventions and the exponential growth of track irregularity. Famurewa et al. compared the accuracy of linear, exponential, and grey models in the estimation and prediction of track geometry degradation. The comparison demonstrated the grey model has lower mean average percentage error than the linear model and an approximately equal error value with the exponential model.

In the deterioration of track quality at one specific track position is shown over a period from 2001 to 2007. The theoretical exponential function is in good accordance with the real track behaviour since the measured track indexes are well fitted.

Lyngby suggested a methodology for evaluating track degradation in terms of track geometry irregularities and proposed a multivariate regression model to demonstrate the relationship between the track degradation measure variable and influencing variables on track degradation. Since different sections of track are not identical, the track was split into homogenous sections with similar variables. He concluded:

- axle load has a nonlinear relation with degradation.
- degradation after tamping is dependent on the number of previous tampings.
- soil consisting of clay material will settle sooner than other types of soil.
- light rail tracks degrade faster than heavy rail tracks.
- harsh rainfall increases degradation rate.

Using waveform data, Liu et al. proposed a short-range prediction model to estimate any track irregularity index over a short track section length (25 m) and on a day-by-day basis. They concluded the total process of track surface change over track sections is nonlinear and different track sections have different nonlinear process.

Xu et al. proposed an approach based on historical changes in track irregularity to predict the short-term track degradation. They estimated the non-linear behaviour of track irregularity during a cycle using several short-range linear regression models.

Two degradation models to predict track alignment irregularities are proposed in the work by Kawaguchi et al. First, they developed a degradation model based on analysis of lateral track deformation to estimate mean time to maintenance of track alignment irregularities. Second, they designed another degradation model based on the exponential smoothing method to accurately predict the track alignment irregularities a maximum of 1 year in advance.

The comparison of the efficiency of the double exponential smoothing method, a generic degradation model, and an autoregressive model for track degradation prediction is addressed in the work by Quiroga et al. The three models lose their efficiency in track degradation prediction after performing several tamping procedures. After considering these issues, they developed a hybrid discrete-continuous framework based on a grey box model. After comparing these four models, they concluded the proposed hybrid model is more efficient in terms of track degradation behaviour prediction.

A stochastic approach based on Dagum distribution is developed by Vale et al. to model track longitudinal level degradation over time. The researchers classified the track longitudinal level changes into three speed classes and different inspection intervals.

The Gaussian random process is used by Zhu et al. to model track irregularities in vertical profile and alignment. They discussed power spectral density analysis and cross-level statistics about track irregularities to improve track degradation modelling.

A stochastic Markov model is used by Bai et al. to evaluate track degradation.
They considered various heterogeneous factors and argued that the existence of these factors caused two maintenance units with the same mileage to show different degradation behaviour. A Markov model is deployed by Yousefikia et al. to model tram track degradation and obtain the optimal maintenance strategy. A model by integrating the grey model and Markov chain is developed by Liu et al. to predict track quality condition.

Andrade et al. used a Bayesian approach to evaluate a track geometry degradation model and deal with the uncertainty of its parameters. They considered the track longitudinal level deviation to have a linear relationship with passing tonnage and assumed the initial longitudinal level and degradation rate would take a bivariate log-normal prior distribution. They argued that the parameter uncertainties are significant in the design stage.
Guler used artificial neural networks to model the degradation of different track geometry parameters. The model considered traffic load, velocity, curvature, gradient, cross-level, sleeper type, rail type, rail length, falling rock, land slide, snow, and flood as influencing factors. A modified grey model is developed by Chaolong et al. to analyse track irregularity time series data and obtain a medium-long term prediction of track cross levelling. They compared the stochastic linear autoregressive model, Kalman filtering model, and artificial neural network with respect to the short-term track cross levelling prediction. They observed the accuracy of the ANN model was higher than the two other models.

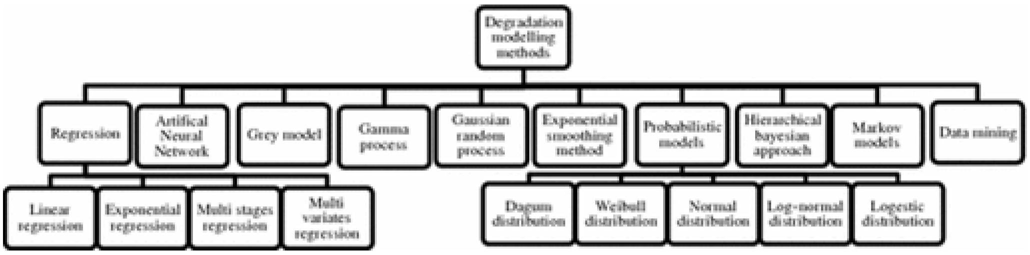

Figure 16: Track Degradation Approaches

A machine learning model based on the characteristics and inspection data of the track using a multi-stage framework is developed by Xu et al. to predict changes in track irregularity over time. They defined different stages of track changes based on maintenance thresholds and linear regression is used to predict track degradation in each stage.

Xu et al. proposed a track measures data mining model to predict railway track degradation for a short time period. Data mining and time series theories are applied by Chaolong et al. to predict track irregularity standard deviation time series data. To predict the changing trends of track irregularity, they used the linear recursive model and the linear autoregressive moving average model.

According to data mining techniques, the prediction of the asset condition can be categorised in two ways: nowcasting and forecasting. Nowcasting methods are used to identify faults that will lead to failure within a few hours; this is done for safety reasons and also to extend remaining useful life (RUL). Forecasting can be useful to assess the condition of an asset for the remaining useful life in the long run. There are three types of methods to quantify remaining useful life: data driven, symbolic and physical models. Data driven methods are purely based on the data acquired by sensors; they carry out classification and clustering techniques to identify anomalies. Symbolic methods make use of work orders and other empirical records of maintenance. Finally, physical methods exploit the physical structure of the component to analyse degradation. The combination of symbolic, data driven and physical models into hybrid models is demonstrated to be a good solution for nowcasting and forecasting of asset condition.

Models for nowcasting and forecasting predicting information about the asset condition provide not only single estimations of current or future values for relevant parameters, but they shall also be used in the alert management system to derive probability distributions (or their characteristic parameters). The future (unknown and random) degradation and defect evolution shall be described as stochastic processes, and functions for transition probabilities can be determined using the output of prognosis models.

# 4.Optimization Model

Optimization models are mathematical models that include functions that represent goals or objectives for the system being modeled. Optimization models can be analyzed to explore system trade-offs to find solutions that optimize system objectives. The model consists of three elements: the objective function, decision variables and business constraints. It is a decision tool to find the best feasible solution of the problem, in which the objective function is maximized or minimized via the variable values subjected to some constraints.
Modeling is a fundamental process in many aspects of scientific research, engineering, and business. Modeling involves the formulation of a simplified representation of a system or real-world object. These simplifications allow structured representation of knowledge about the original system that facilitates the analysis of the resulting model. Schichl notes that models are used to.

- Explain phenomena that arise in a system;
- Make predictions about future states of a system;
- Assess key factors that influence phenomena in a system;
- Identify extreme states in a system that might represent worst-case scenarios or minimal cost plans; and
- Analyze trade-offs to support human decision makers.

Additionally, the structured aspect of a model's representation facilitates communication of the knowledge associated with a model. For example, a key aspect of a model is its level of detail, which reflects the system knowledge that is needed to employ the model in an application.

Mathematics has always played a fundamental role in representing and formulating our knowledge. Mathematical modeling has become increasingly formal as new frameworks have emerged to express complex systems. The following mathematical concepts are central to modern modeling activities:

- Variables: These represent unknown or changing parts of a model (e.g., which decisions to take, or the characteristic of a system outcome).
- Parameters: These are symbolic representations for real-world data, which might vary for different problem instances or scenarios.
- Relations: These are equations, inequalities, or other mathematical relationships that define how different parts of a model are related to each other.

Optimization models are mathematical models that include functions that represent goals or objectives for the system being modeled. Optimization models can be analyzed to explore system trade-offs in order to find solutions that optimize system objectives. Consequently, these models can be used for a wide range of scientific, business, and engineering applications.

Optimization problems can be classified in terms of the nature of the objective function and the nature of the constraints. Special forms of the objective function and the constraints give rise to specialized algorithms that are more efficient. From this point of view, there are four types of optimization problems, of increasing complexity.

An Unconstrained optimization problem is an optimization problem where the objective function can be of any kind (linear or nonlinear) and there are no constraints. These types of problems are handled by the classes discussed in the earlier sections.

A linear program is an optimization problem with an objective function that is linear in the variables, and all constraints are also linear. Linear programs are implemented by the Linear Program class.

A quadratic program is an optimization problem with an objective function that is quadratic in the variables (i.e., it may contain squares and cross products of the decision variables), and all constraints are linear. A quadratic program with no squares or cross products in the objective function is a linear program. Quadratic programs are implemented by the Quadratic Program class.

A nonlinear program is an optimization problem with an objective function that is an arbitrary nonlinear function of the decision variables, and the constraints can be linear or nonlinear. Nonlinear programs are implemented by the Nonlinear Program class.

# 4.1 Basics on Mathematical Programming

Mathematical models are usually composed of relationships and variables. Relationships can be described by operators, such as algebraic operators, functions, differential operators, etc. Variables are abstractions of system parameters of interest, that can be quantified. Several classification criteria can be used for mathematical models according to their structure:

Linear vs. nonlinear: If all the operators in a mathematical model exhibit linearity, the resulting mathematical model is defined as linear. A model is considered to be nonlinear otherwise. The definition of linearity and nonlinearity is dependent on context, and linear models may have nonlinear expressions in them. For example, in a statistical linear model, it is assumed that a relationship is linear in the parameters, but it may be nonlinear in the predictor variables. Similarly, a differential equation is said to be linear if it can be written with linear differential operators, but it can still have nonlinear expressions in it. In a mathematical programming model, if the objective functions and constraints are represented entirely by linear equations, then the model is regarded as a linear model. If one or more of the objective functions or constraints are represented with a nonlinear equation, then the model is known as a nonlinear model.

Linear structure implies that a problem can be decomposed into simpler parts that can be treated independently and/or analyzed at a different scale and the results obtained will remain valid for the initial problem when recomposed and rescaled.
Nonlinearity, even in fairly simple systems, is often associated with phenomena such as chaos and irreversibility. Although there are exceptions, nonlinear systems and models tend to be more difficult to study than linear ones. A common approach to nonlinear problems is linearization, but this can be problematic if one is trying to study aspects such as irreversibility, which are strongly tied to nonlinearity.

**Static vs. dynamic:** A dynamic model accounts for time-dependent changes in the state of the system, while a static (or steady-state) model calculates the system in equilibrium, and thus is time-invariant. Dynamic models typically are represented by differential equations or difference equations.

**Explicit vs. implicit:** If all of the input parameters of the overall model are known, and the output parameters can be calculated by a finite series of computations, the model is said to be explicit. But sometimes it is the output parameters which are known, and the corresponding inputs must be solved for by an iterative procedure, such as Newton's method or Broyden's method. In such a case the model is said to be implicit. For example, a jet engine's physical properties such as turbine and nozzle throat areas can be explicitly calculated given a design thermodynamic cycle (air and fuel flow rates, pressures, and temperatures) at a specific flight condition and power setting, but the engine's operating cycles at other flight conditions and power settings cannot be explicitly calculated from the constant physical properties.

**Discrete vs. continuous:** A discrete model treats objects as discrete, such as the particles in a molecular model or the states in a statistical model; while a continuous model represents the objects in a continuous manner, such as the velocity field of fluid in pipe flows, temperatures and stresses in a solid, and electric field that applies continuously over the entire model due to a point charge.

**Deterministic vs. probabilistic (stochastic):** A deterministic model is one in which every set of variable states is uniquely determined by parameters in the model and by sets of previous states of these variables; therefore, a deterministic model always performs the same way for a given set of initial conditions. Conversely, in a stochastic model-usually called a "statistical model"-randomness is present, and variable states are not described by unique values, but rather by probability distributions.
Deductive, inductive, or floating: A deductive model is a logical structure based on a theory. An inductive model arises from empirical findings and generalization from them. The floating model rests on neither theory nor observation, but is merely the invocation of expected structure. Application of mathematics in social sciences outside of economics has been criticized for unfounded models. Application of catastrophe theory in science has been characterized as a floating model.

**Strategic vs non-strategic:** Models used in game theory are different in a sense that they model agents with incompatible incentives, such as competing species or bidders in an auction. Strategic models assume that players are autonomous decision makers who rationally choose actions that maximize their objective function. A key challenge of using strategic models is defining and computing solution concepts such as Nash equilibrium. An interesting property of strategic models is that they separate reasoning about rules of the game from reasoning about behavior of the players.

As mentioned before, the objective of this thesis is the study of a maintenance activity scheduling problem, in which not only costs, but also risk issues are considered. The problem is analyzed as a Mixed Integer Programming problem (MIP). Moreover, it has been proven to be a non-polynomial time hard problem (NP-hard). This makes the problem solvable in a reasonable time only for very tiny instance dimensions. Then, it is necessary the implementation of “heuristic” approaches so as to find good solutions in short times, even if sub-optimal. Before showing the core of the problem formulation, it seems to be interesting giving a brief exposition of the main concepts of MIPs, Machine Scheduling, heuristic and matheuristic approach.

Therefore, first, a brief overview on the theoretical background and a brief dissertation on Mathematical Programming is proposed. In particular, the focus is on Mixed Integer Linear Programming and its application in scheduling problems.

# 4.2 Mixed Integer Linear Programming

Many MP problems exist where it is necessary to restrict the decision variables to integer or binary values. Examples include cases where the decision variable represents a non-fractional entity such as people or bicycles, or where a decision variable is needed to model a logical statement (such as whether or not to assign task A to agent B). These problems are called Mixed Integer-Linear Programming (MILP) problems and are often much harder to solve than LP problems. This is because instead of having feasible solution points at the easily computed corners of the feasible region, they are instead usually internal and more difficult to locate. For example, constraining X and Y from the previous LP formulation to have integer values, the feasible solution points are shown in Figure.

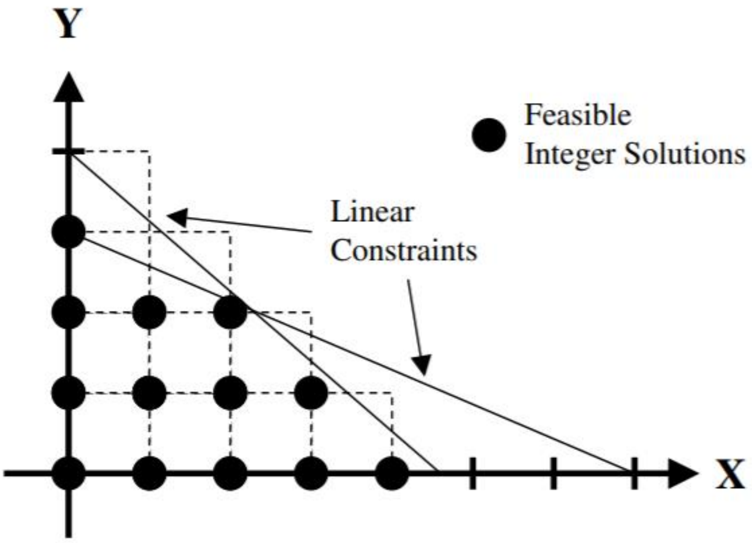


Figure 17: Mixed Integer Linear Programming problem showing all feasible Solutions.

The "classical" theory of mathematical optimisation deals with the following abstract and generalised problem: given a set of decisions, find amongst the "feasible" choices one that is "best". Formally, this is expressed as follows: find an optimum solution $x^* \in M$ such that,

$$f(x^*) \leq f(x) \qquad \forall x \in M$$

This compact representation includes all three ingredients of a mathematical optimisation problem: The degrees of freedom $x$ for which a decision has to be made, the objective function $f(x)$ to be achieved, by which the different choices can be compared to each other, and the restrictions to be met, expressed by the set of feasible solutions (or choices) $M$. The set $M$ normally is constructed by a number of constraints in the form,

$$h_i(x) \leq 0, \qquad i = 1, \dots, m$$

or $h(x) \leq 0$ where $h(x)$ is considered as the vector

$$h(x) = (h_1(x), \dots, h_m(x)).$$

The Mixed Integer Mathematical Programming is the part of mathematical optimization that make possible to formulate many complex optimization problems in which some variables are restricted to be integer.

A MIP problem is defined by a set of variables $x$ and an objective function $(minc^T x)$ subject to a set of linear constraints $(Ax = b)$ and a set of integrity constraints on part of the variables.

A typical formulation is:

$$Z = minc^T x \qquad (4.1.1)$$

subject to:

$$Ax + Gy = b \qquad (4.1.2)$$

$$x \geq 0; \qquad x \in Z \qquad (4.1.3)$$

$$y \geq 0; \qquad y \in Z \qquad (4.1.4)$$

where $A \in \mathbb{R}^{mn}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$and $I \subseteq \{1,..,n\}$.

The set $X = \{x \in \mathbb{R}^n Ax \leq b, x \geq 0, x_i \in \mathbb{Z}, \forall i \in I\}$ represents the feasibility region.

This kind of problems are also called Combinatorial Problems. Mixed integer optimization models are suitable to describe situations in which the objective is to optimize the use of no divisible resources or the choice between discrete alternatives. Frequently, the formulations present Boolean variables, so as the integer variable assume only $0-1$ values. These cases, and in particular the one proposed in this thesis, are called $0-1$ Programming. In general, binary values are used to describe the occurrence (or not) of the considered event:

$$x = \begin{cases} 1 \ iftheeventoccurs \\ 0 \ otherwise \end{cases} \qquad (4.1.5)$$

# 4.3 Scheduling Problem

Scheduling is concerned with the allocation of production resources to production orders in an optimized way. Scheduling is utilized to support production, purchasing, logistics, and so forth and plays a key role in process operations where it may yield great improvements of production performance. The research area of batch and continuous process scheduling has received great attention from both academia and industry in the past two decades. Consequently, a remarkable amount of mathematical models for the solution of these challenging problems have been developed recently.

MILP models are used in a great set of applications, but the scope of this dissertation is focused on a peculiar kind of formulation: sequencing/scheduling problems.

The basic concepts and notation of scheduling problem are introduced (Graham et al. **Error! Reference source not found.** and Blazewicz**Error! Reference source not found.**). In scheduling problem, a set of $n$ jobs/tasks$j = 1, \ldots, n$ and $M$ machines/processors $m = 1, \ldots, M$ are considered. In the scheduling literature, the terms “task” and “job” are often used interchangeably, although in some cases, tasks are decomposed into separate parts called jobs. Each machine may work on a single job at a time, and each job may be processed by a single machine at a time. The schedule is a list of jobs with the times when the jobs are processed by machines, and a feasible schedule satisfies the timing requirements as well as the fundamental assumptions described above.

In one-machine environment there is only one machine that can process one job at time.

If each job must be processed in an uninterrupted time period, the schedule environment is non-preemptive, whereas, if a job may be processed for a period of time, interrupted and continued in a later point in time, the scheduling is preemptive.

Each task $j$has the following properties (Conway et al. **Error! Reference source not found.**):

- the task has a vector of processing times with each element of the vector corresponding to the processing on a particular processor, $[p_{j1}, p_{j2}, \ldots, p_{jM}]$,
- the task has an arrival time or ready time, $r_j$,
- the task has a due date or deadline, $d_j$
- the task has a weight or priority, $w_j$,
- the task may be preemptive or non-preemptive, depending on whether preemption is allowed in the schedules (preemption is also referred to as “task splitting”),
- the task may be dependent or independent. Dependence between tasks is specified by means of a precedence tree or a more general precedence graph.

Therefore, a schedule is an assignment of processors to tasks. At each moment, at most one task is assigned to each processor, and at most one processor is assigned to each task. Each task is processed after its arrival time, and all tasks are completed.

The peculiarity of a scheduling problem consists in the way the variables are used to describe precedence/succession constraints, which are defined “disjunctive constraints”.

This definition is related to the fact that, while usually MILP formulations require the satisfaction of the whole set of constraints, in this case only a subset is needed to be respected. This is important when there are sequencing activities that cannot be processed at the same time. In fact, supposed to have $n$operations to be sequenced on a unitary capacity machine. Let $p_i$ be the processing time of the $i-$th task on the $m-$th machine, $t_i$ the starting time of the $i-$th task. So, if the $i-$th task precedes the $j-$th, then: $t_j \geq t_i + p_i$.

Vice versa, if $j$ precedes the$i$ then$t_i \geq t_j + p_j$. The representation of these constraints is possible thanks to the following binary variables:

$$x_{ij} = \begin{cases} 1 & \text{if i is processed before j} \\ 0 & \text{otherwise} \end{cases} \qquad (4.3.6)$$

Then, the constraints are formulated in this way:

$$Bx_{ij} + t_i - t_j \geq p_j \qquad 1 \leq i < j \leq n \qquad (4.3.7)$$

$$B(1 - x_{ij}) + t_i - t_j \geq p_i \qquad 1 \leq i < j \leq n \qquad (4.3.8)$$

where B is a big real number representing the infinite.

It is clear that, if $x_{ij} = 1$ then the first constraint is always satisfied because $B \gg p_j + t_j - t_i$, while the latter expresses the starting time for the $j-$th activity. Vice versa when $x_{ij} = 0$. Starting from this formulation for the activity sequencing, it is possible to construct a scheduling problem through the introduction of temporal constraints.

The performance characteristics and performance measures of individual tasks and of schedules can be defined. Each task in a schedule can have:

- a completion time which we denote as $C_j$,
- a flow time, denoted $F_j = C_j - r_j$,
- a lateness, denoted $L_j = C_j - d_j$
- a tardiness, denoted $T_j = \max(L_j, 0)$
- a unit penalty $U_j = 0$ if $C_j \leq d_j$, else 1.

These properties of schedules not only provide measures for evaluating schedules, but also provide criteria for optimization in algorithms that produce schedules.

In particular, schedules are evaluated using,

- schedule length or make span, $C_{max} = \max(C_j)$;
- mean flow time, $\overline{F} = \frac{1}{n}\sum_{j=1}^{n} F_j$
- mean weighted flow time, $\overline{F}_w = \frac{\sum_{j=1}^{n} w_j F_j}{\sum_{j=1}^{n} w_j}$
- maximum lateness, $L_{max} = \max(L_j)$;
- mean tardiness, $\overline{T} = \frac{1}{n}\sum_{j=1}^{n} T_j$
- mean weighted tardiness, $\overline{T}_w = \frac{\sum_{j=1}^{n} w_j T_j}{\sum_{j=1}^{n} w_j}$
- number of tardy tasks, $\overline{U} = \sum_{j=1}^{n} U_j$

# 4.4 Simulation Based Optimization

Simulation-based optimization (also known as simply simulation optimization) integrates optimization techniques into simulation modelling and analysis. Because of the complexity of the simulation, the objective function may become difficult and expensive to evaluate. Usually, the underlying simulation model is stochastic, so that the objective function must be estimated using statistical estimation techniques (called output analysis in simulation methodology).

Once a system is mathematically modelled, computer-based simulations provide information about its behaviour. Parametric simulation methods can be used to improve the performance of a system. In this method, the input of each variable is varied with other parameters remaining constant and the effect on the design objective is observed. This is a time-consuming method and improves the performance partially. To obtain the optimal solution with minimum computation and time, the problem is solved iteratively where in each iteration the solution moves closer to the optimum solution. Such methods are known as 'numerical optimization' or 'simulation-based optimization'.

In simulation experiment, the goal is to evaluate the effect of different values of input variables on a system. However, the interest is sometimes in finding the optimal value for input variables in terms of the system outcomes. One way could be running simulation experiments for all possible input variables. However, this approach is not always practical due to several possible situations and it just makes it intractable to run experiments for each scenario. For example, there might be too many possible values for input variables, or the simulation model might be too complicated and expensive to run for suboptimal input variable values. In these cases, the goal is to find optimal values for the input variables rather than trying all possible values. This process is called simulation optimization.

Specific simulation–based optimization methods can be chosen according to Figure based on the decision variable types.

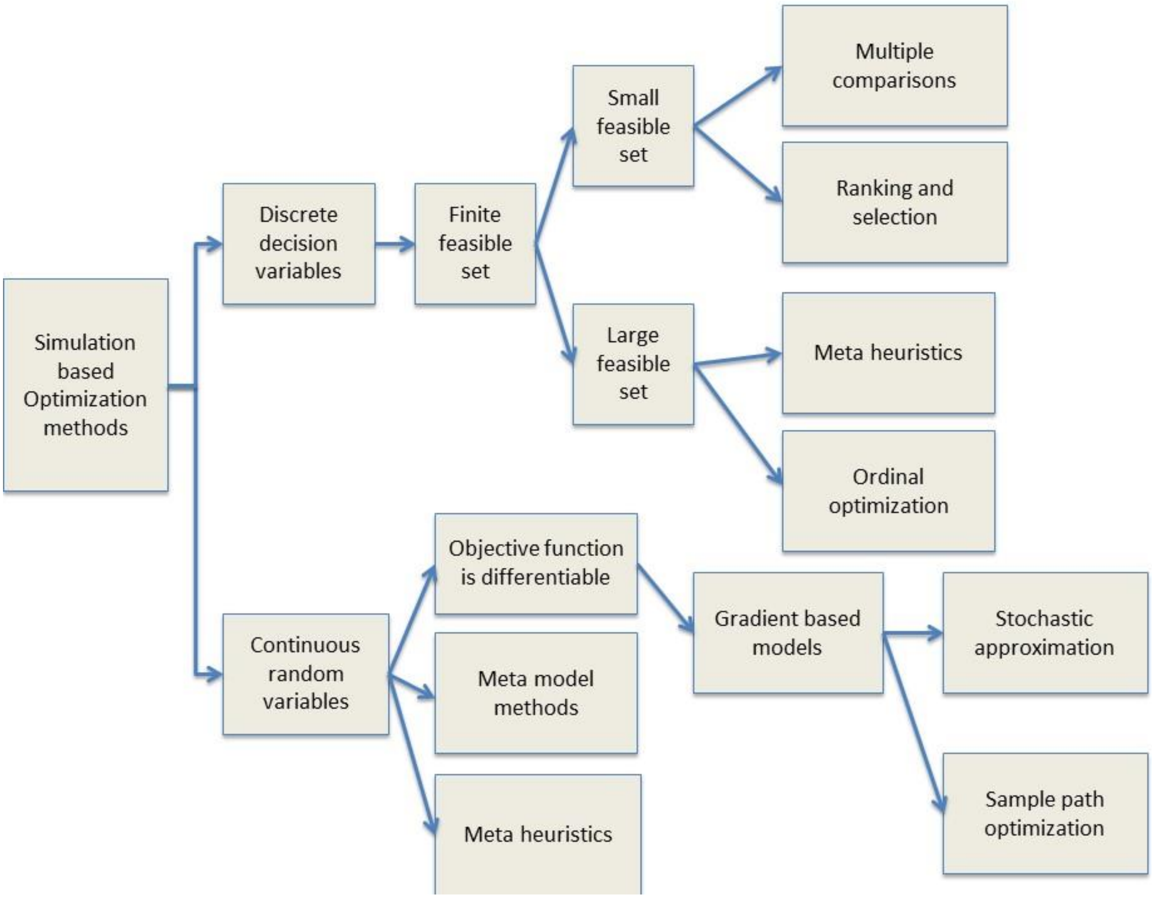


Figure 18: Classification of simulation-based optimization according to variable types

**Optimization exists in two main branches of operations research:**

**Optimization parametric (static)** – The objective is to find the values of the parameters, which are “static” for all states, with the goal of maximizing or minimizing a function. In this case, one can use mathematical programming, such as linear programming. In this scenario, simulation helps when the parameters contain noise or the evaluation of the problem would demand excessive computer time, due to its complexity.

**Optimization control (dynamic)** – This is used largely in computer science and electrical engineering. The optimal control is per state and the results change in each of them. One can use mathematical programming, as well as dynamic programming. In this scenario, simulation can generate random samples and solve complex and large-scale problems.

# 5. Simulative Approaches

A simulation is the imitation of the operation of a real-world process or system over time. Whether done by hand or on a computer, simulation involves the generation of an artificial history of a system and the observation of that artificial history to draw inferences concerning the operating characteristics of the real system. The behavior of a system as it evolves over time is studied by developing a simulation model. This model usually takes the form of a set of assumptions concerning the operation of the system. These assumptions are expressed in mathematical, logical, and symbolic relationships between the entities, or objects of interest, of the system. Once developed and validated, a model can be used to investigate a wide variety of "what if" questions about the real-world system. Potential changes to the system can first be simulated, in order to predict their impact on system performance. Simulation can also be used to study systems in the design stage before such systems are built. Thus, simulation modeling can be used both as an analysis tool for predicting the effect of changes to existing systems and as a design tool to predict the performance of new systems under varying sets of circumstances. In some instances, a model can be developed which is simple enough to be "solved" by mathematical methods. Such solutions might be found by the use of differential calculus, probability theory, algebraic methods, or other mathematical techniques. The solution usually consists of one or more numerical parameters, which are called measures of performance of the system. However, many real-world systems are so complex that models of these systems are virtually impossible to solve mathematically. In these instances, numerical, computer-based simulation can be used to imitate the behavior of the system over time. From the simulation, data are collected as if a real system were being observed. This simulation-generated data is used to estimate the measures of performance of the system.

It is recommended to use simulation when the studied system involves variables with stochastic behavior, none or minimal correlation and independent and identically distributed (IID) properties. If one of those characteristics are not met, the data should be treated, or the decision maker should consider the use of other types of modeling and optimization techniques, such as linear and non-linear optimization. Another simulation characteristic refers to how the entities change during time. If it only changes at specific points in the system, it is considered discrete (e.g., operations such as cut, weld, paint), in opposition to variables that change continuously during a period of time. Other types of modeling and simulation are based on agents' behaviors. In these cases, the agents are individuals with their behavior and rules, where the modeler can specify the condition when the rules will be executed. Agents are considered like decision makers with some level of learning and adaptation. Simulation projects often aim to answer questions related to the optimization of specific characteristics that represent "what if" scenarios to the proposed system. Optimization is defined as the minimization or maximization or both related to a one or multi-objective function that summarizes, in a mathematical form, the questions made for the system. If so, different combinations of alternatives are considered viable if it satisfies all the restrictions of the problem, or unviable if at least one restriction is not satisfied. The alternative that has the best value for the objective function is considered optimal. If the simulation has sufficient data to represent the analyzed system, the best-simulated solution can be inferred as optimal and have good chances to be implemented in a real system, performing the goal to be an excellent tool to help decision making. To find the optimal solution, a search space made from the combination of the possible values from the variables is evaluated. The size of this search space can be a problem regarding the resources necessary to perform a full search covering all the possible solutions, to find the best one. The resources, in this case, are commonly related to the computational power available to perform all the possible solutions that represent a quantity of time that the decision-making person could not have. Those types of problem are considered NP-hard. Both simulation and optimization are used for performance improvement in management, planning, control and methods for decision-making.

## 5.1. Discrete Event Simulation (DES)

A discrete-event simulation (DES) models the operation of a system as a (discrete) sequence of events in time. Each event occurs at a particular instant in time and marks a change of state in the system. Between consecutive events, no change in the system is assumed to occur; thus, the simulation time can directly jump to the occurrence time of the next event, which is called next-event time progression.

In addition to next-event time progression, there is also an alternative approach, called fixed-increment time progression, where time is broken up into small time slices and the system state is updated according to the set of events/activities happening in the time slice. Because not every time slice has to be simulated, a next-event time simulation can typically run much faster than a corresponding fixed-increment time simulation.

Both forms of DES contrast with continuous simulation in which the system state is changed continuously over time on the basis of a set of differential equations defining the rates of change of state variables.

A common exercise in learning how to build discrete-event simulations is to model a queue, such as customers arriving at a bank to be served by a teller. In this example, the system entities are Customer-queue and Tellers. The system events are Customer-Arrival and Customer-Departure. (The event of Teller-Begins-Service can be part of the logic of the arrival and departure events.) The system states, which are changed by these events, are Number-of-Customers-in-the-Queue (an integer from 0 to n) and Teller-Status (busy or idle). The random variables that need to be characterized to model this system stochastically are Customer-Interarrival-Time and Teller-Service-Time. An agent-based framework for performance modeling of an optimistic parallel discrete event simulator is another example for a discrete event simulation.

**Steps in a Simulation Study**

Set of steps to guide a model builder in a thorough and sound simulation study. Similar discussion of steps can be found in other sources. The steps in a simulation study are as follows:

**Problem formulation** Every study should begin with a statement of the problem. If the statement is provided by the policymakers or those that have the problem, the analyst must ensure that the problem being described is clearly understood. If a problem statement is being developed by the analyst, it is important that the policymakers understand and agree with formulation. Although there are occasions where the problem must be reformulated as the study progresses. In many instances, policymakers and analysts are aware that there is a problem long before the nature of the problem is known.

**Setting of objectives and overall project plan:** The objectives indicate the questions to be answered by simulation. At this point, a determination should be made concerning whether simulation is the appropriate methodology for the problem as formulated and the objectives as stated. Assuming that it is decided that simulation is appropriate, the overall project plan should include a statement of the alternative systems to be considered and of a method for evaluating the effectiveness of these alternatives. It should also include the plans for the study in terms of the number of people involved, the cost of the study, and the number of days required to accomplish each phase of the work, along with the results expected at the end of each stage.

**Model conceptualization:** The construction of a model of a system is probably as much art as science. Pritsker provides a lengthy discussion of this step ‘Although it is not possible to provide a set of instructions that will lead to building successful and appropriate models in every instance, there are some general guidelines that can be followed. The art of modeling is enhanced by an ability to abstract the essential features of a problem, to select and modify basic assumptions that characterize the system, and then to enrich and elaborate the model until a useful approximation results. Thus, it is best to start with a simple model and build toward greater complexity. However, the model complexity need not exceed that required to accomplish the purposes for which the model is intended. Violation of this principle will only add to model-building and computer expenses. It is not necessary to have a one-to-one mapping between the model and the real system. Only the essence of the real system is needed.

It is advisable to involve the model user in model conceptualization. Involving the model user will both enhance the quality of the resulting model and increase the confidence of the model user in the application of the model.

Data collection there is a constant interplay between the construction of the model and the collection of the needed input data. As the complexity of the model changes, the required data elements can also change. Also, since data collection takes such a large portion of the total time required to perform a simulation, it is necessary to begin as early as possible, usually together with the early stages of model building.

The objectives of the study dictate, in a large way, the kind of data to be collected. In the study of a bank, for example, if the desire is to learn about the length of waiting lines as the number of tellers changes, the types of data needed would be the distributions of the inter arrival times (at different times of the day), the service-time distributions for the tellers, and historic distributions on the lengths of waiting lines under varying conditions. This last type of data will be used to validate the simulation model.

Model translation most real-world systems result in models that require a great deal of information storage and computation, so the model must be entered into a computer-recognizable format. We use the term program even though it is possible, in many instances, to accomplish the desired result with little or no actual coding. The modeler must decide whether to program the model in a simulation language, such as GPSS/HTM, or to use special- purpose simulation software. For manufacturing and material handling, such software as Any Logic, Arena, Auto mod tm, Enterprise Dynamics, Extend, Flexism, ProModel, and SIMUL8. Simulation languages are powerful and flexible. However, if the problem is amenable to solution with the simulation software, the model development time is greatly reduced. Furthermore, most simulation software packages have added features that enhance their flexibility, although the amount of flexibility varies greatly.

**Verified:** Verification pertains to the computer program that has been prepared for the simulation model. Is the computer program performing properly? With complex models, it is difficult, if not impossible; to translate a model successfully in its entirety without a good deal of debugging; if the input parameters and logical structure of the model are correctly represented in the computer, verification has been completed. For the most part, common sense is used in completing this step.

**Validated:** Validation usually is achieved through the calibration of the model, an iterative process of comparing the model against actual system behavior and using the discrepancies between the two, and the insights gained, to improve the model. This process is repeated until model accuracy is judged acceptable. In the previously mentioned example of bank, data was collected concerning the length of waiting lines under current conditions. Does the simulation model replicate this system measure? This is one means of validation.

Experimental design the alternatives that are to be simulated must be determined. Often, the decision concerning which alternatives to simulate will be a function of runs that have been completed and analyzed. For each system design that is simulated, decisions need to be made concerning the length of the initialization period, the length of simulation runs, and the number of replications to be made of each run.

**Production runs and analysis:** Production runs and their subsequent analysis, are used to estimate measures of performance for the system designs that are being simulated.

**More runs?** Given the analysis of runs that have been completed, the analyst determines whether additional runs are needed and what design those additional experiments should follow.

Documentation and reporting there are two types of documentation: program and progress. Program documentation is necessary for numerous reasons. If the program is going to be used again by the same or different analysis, it could be necessary to understand how the program operates. This will create confidence in the program, so that model users and policymakers can make decisions based on the analysis. Also, if the program is to be modified by the same or a different analyst, this step can be greatly facilitated by adequate documentation. One experience with an inadequately documented program is usually enough to convince an analyst of the necessity of the important step. Another reason for documenting a programs so that model users can change parameters at will in an effort to learn the relationships between input parameters and output measures of performance or to discover the input parameters that "optimize" some output measure of performance.

Musselman{1998} discusses progress reports that provide the important, written history of a simulation project. Project reports give a chronology of work done and decisions made. This can prove to be of great value in keeping the project on course. Musselman suggests frequent reports so that even those not involved in the day- to- day operation can be kept abreast. The awareness of these others can often enhance the successful completion of the project by surfacing misunderstanding early, when the problem can be solved easily. Musselaman also suggests maintaining a project log to provide a comprehensive record of accomplishments, change requests, key decisions, and other items of importance.

On the reporting side, Musselman suggests frequent deliverables. These may or may not be the results of major accomplishments. His maxim is that "it is better to work with many intermediate milestones than with one absolute deadline. "Possibilities prior to the final report include a model specification prototype demonstration, animations, training results, intermediate analysis, program documentation, progress reports, and presentations. He suggests that these deliverables should be timed judiciously over the life of the project.

The results of all the analysis should be reported clearly and concisely in a final report. This will allow the model users to review the final formulation, the alternative systems that were addressed, the criteria by which the alternatives were compared, the results of the experiments, and the recommended solution to the problem. Furthermore, if the decisions have to be justified at a higher level, the final report should provide a vehicle of certification for the model user/decision maker and add to the credibility of the model and of the model-building process.

Implementation the success of implementation phase depends on how well the previous eleven steps have been performed. It is also contingent upon how thoroughly the analyst has involved the ultimate model user during the entire simulation process, If the model user has been involved during the entire model-building process and if the model user understands the nature of the model and its outputs, the likelihood of a vigorous implementation is enhanced Conversely, if the model and its underlying assumptions have not been properly communicated, implementation will probably suffer, regardless of the simulation model's validity.

The simulation-model building process can be broken down into four phases. The first phase, consisting of steps 1(problem formulation) and 2 {Setting of objective and overall design}, is a period of discovery or orientation. The initial statement of the problem is usually quite "fuzzy", the initial objectives will usually have to be reset, and the original project plan will usually have to be fine turned. These recalibrations and clarifications could occur in this phase or perhaps will occur after or during another phase.

The second phase is related to model building and data collection and include steps 3(Model conceptualization), 4 (Data collection), 5 (Model translation), 6 (Verification) and 7 (Validation). A continuing interplay is required among the steps. Exclusion of the model user during this phase can have implications at the time of implementation.

The third phase concerns the running of the model. It involves steps 8 (Experimental design), 9 (Production runs and analysis) and 10 (More runs). This phase must have a comprehensively is, in fact, a statistical experiment. The output variables are estimates that contain random error, and therefore a proper statistical analysis is required. Such a philosophy is in contrast to that of the analyst who makes a single run and draws an inference from that single data point.

The fourth phase, implementation, involves steps 11(Documentation and reporting) and 12 (Implementation). Successful implementation on continual involvement of the model user and on the successful completion of every step in the process. Perhaps the most crucial point in the entire process is step 7 (Validation), because an invalid model is going to lead to erroneous results, which, if implemented, could be dangerous, costly, or both.

# 6.Degradation Model for Rail Track

This section describes the risk- based approach to determining the set of rail stretches to be maintained starting from a stochastic degradation model. To this end, since a data analysis is outside the scope of the present study, the rail vertical deformation model provided by Famurewa at [41][42] is considered. This model, developed by means of data collected by measurement cars from 2011 to 2016 on a single track line in Sweden, Provides the rail deformation overtime of given rail stretches.

NOTATION FOR THE CONSIDERED DEGRADATION MODEL

| Symbol | Variable |
|---|---|
| $\mathcal{R}$ | complete set of $\|\mathcal{R}\|$ railway stretches |
| $\mathcal{A} \subseteq \mathcal{R}$ | set of $\|\mathcal{A}\|$ rail stretches needing maintenance |
| $\tau$ | generic time instant |
| $\tau_k^i$ | time instant of the last maintenance activity on the rail stretch $i$ |
| $\delta_i(\tau_k^i)$ | residual average deformation after maintenance activity in $\tau_k^i$ |
| $\delta_i(\tau, \tau_k^i)$ | predicted average deformation model |
| $\overline{\delta}$ | maximum admissible positive/negative deformation |
| $T_{\overline{\delta}}^i$ | random variable that represents the time instant at which the deformation of the rail stretch $i$ reaches the maximum deformation threshold |
| $\epsilon$ | Gaussian random variable representing the vertical deformation model's approximations and error |
| $\varphi_i$ | failure event of the rail stretch $i \in \mathcal{A}$ |
| $F_{T_{\overline{\delta}}^i}(\tau)$ | cumulative distribution function (cdf) |
| $f_{T_{\overline{\delta}}^i}(\tau)$ | probability density function (pdf) |
| $\overline{\Pr}\{\varphi_i\}$ | maximum admissible failure probability |
| $\gamma_S, \gamma_R$ | parameters defining the the soft deadline and the release time |
| $\tau_H^i$ | hard deadline for maintenance activity of rail stretch $i \in \mathcal{A}$ |
| $\tau_S^i$ | soft deadline for maintenance activity of rail stretch $i \in \mathcal{A}$ |
| $\tau_R^i$ | release times for maintenance activity of rail stretch $i \in \mathcal{A}$ |

Nevertheless, although the authors introduced the notion of deformation uncertainty, their determination of the deadlines for maintenance interventions neglected this factor and considered only the average deformation.

From the physical point of view, it is then possible that the average deformation falls within the thresholds while at some points, the real profile may exceed these limits. A sketch of this condition is depicted in Fig. 19. In this article, to evaluate the probability that the deformation exceeds the relevant threshold, the model uncertainty is explicitly considered. The relevant notation is summarized in Table I. The vertical deformation model proposed in [41], [42] for determining the *spatial average* (computed over the whole length) of the vertical rail deformation of the generic rail stretch expressed as

$$\delta_i(T, T_k^i) = \delta_i(T_k^i)\exp(\alpha_i T) + \epsilon \qquad (1)$$

and is a random variable. In addition, in (1), $\delta_i(T_k^i)$ is the average deformation along the rail stretch $i$ after the maintenance performed in$T_k^i$, and $\delta_i\ (T,\ T_k^i)$ is the average deformation in a generic instant$T > T_k^i$. In addition, $\in\in N(0, \sigma^2)$is a Gaussian random variable with null expectation modeling the deformation uncertainty due to the model approximations, measurement errors, and punctual deviation of the real value of the deformation with respect to the average predicted by (1). The random variable _ is assumed to be independent of the particular rail stretch. As mentioned, the model parameters, the initial deformation $\delta_i(T_k^i)$, the coefficient $\alpha_i$, and the variance $\sigma^2$ are determined via field data from [41], [42] and hereafter are assumed to be known.

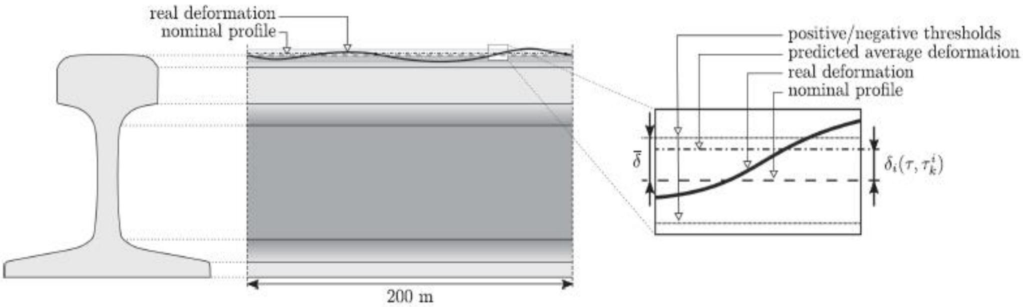


Figure 19: Rail Profile

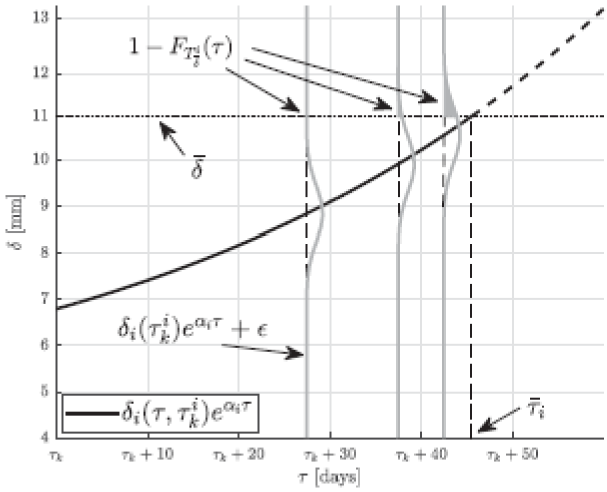


Figure 20: Gaussian stochastic process. The thick line represents the expectation of the different Gaussian random variables $\delta_i(T, T_k^i), \forall_T$.

According to the model in (1), The deformation $\delta_i(T, T_k^i)$ results in a Gaussian stochastic process with time-dependent expectation $E[\delta_i(T, T_k^i)] = \delta_i(T_k^i)\exp(\alpha_i T)$ and constant variance $\sigma^2$. As a consequence, at any time $T$, there is a nonzero probability $\Pr\{\delta_i(T, T_k^i) \geq \bar{\delta}\} = 1 - F\delta_1(\bar{\delta}, \delta_i(T, T_k^i)$ that the actual deformation is greater than the threshold$\bar{\delta}$, as respresented by the gray area in Fig.2. Therefore, the deadline for maintenance intervention, i.e., the time instant at which the maximum deformation is reached, can be computed by considering a maximum threshold of such a probability. To this end, Let $T_{\bar{\delta}}^{i}$ be a random variable representing the time instant at which the deformation of the rail stretch $i$ reaches the threshold $\bar{\delta}$, i.e.,

$$\bar{\delta} = \delta_i(T_k^i)\exp(\alpha_i T_{\bar{\delta}}^i) + \epsilon \qquad (2)$$

By means of simple manipulation (2) becomes

$$T_{\bar{\delta}}^{i} = g(\epsilon) = \frac{1}{\alpha_i} 1n \frac{\bar{\delta}-\epsilon}{\delta_i(T_k^i)} \qquad (3)$$

Which is defined if the argument of the logarithm is positive. This assumption is reasonable since, thanks to Chebyshev's inequality $\Pr\{\bar{\delta}-\epsilon \leq 0\} \leq \sigma^2/\bar{\delta}^2$, the probability that $\epsilon \geq \bar{\delta}$turns out to be negligible for realistic values of $\sigma^2$ and $\bar{\delta}$. In other words,)(3) requires that the model uncertainty $\epsilon$ be small with respect to maximum admissible $\bar{\delta}$, although it can be not negligible in general.

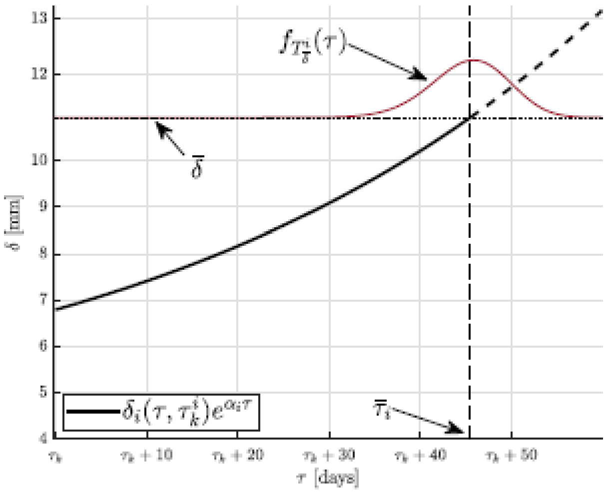


Figure 21: Shape of the pdf in (5) with $\bar{\delta}$ = 11mm and $\alpha_i$= 0.01 days[-1]

Since the function $g(\epsilon)$ in (3) is continuous and decreases monotonically, the cumulative distribution function (cdf) of $T_{\bar{\delta}}^{i}$ can be defined as

$$F_{T_{\bar{\delta}}^{i}}(T) = \Pr\{T_{\bar{\delta}}^{i} \leq T\} = \Pr\{\frac{1}{\alpha_i} 1n(\frac{\bar{\delta}-\epsilon}{\delta_i(T_k^i)} \leq T\}$$
$$= \Pr\{\epsilon \geq \bar{\delta} - \delta_i(T_K^i)\exp(\alpha_i T)\}$$
$$= 1 - F_\epsilon(\bar{\delta} - \delta_i(T_K^i)\exp(\alpha_i T)) \qquad (4)$$

Where $F_\epsilon(.)$ is the cdf of the Gaussian random variable $\epsilon$. Therefore, the relevant probability density function (pdf) is

$$f_{T_{\bar{\delta}}^{i}(T)} = \frac{dF_{T_{\bar{\delta}}^{i}(T)}}{dT} = \frac{\alpha_i \delta_i(T_K^i)\exp(\alpha_i T)}{\sqrt{2\Pi\sigma^2}} \,.\exp\left(-\frac{(\bar{\delta}-\delta_i(T_K^i)\exp(\alpha_i T))^2}{2\sigma^2}\right) \qquad (5)$$

Whose shape is reported in Fig. 3, where it is possible to note its heavy right tail. The pdf in (5) allows to determine the median, which thanks to the symmetry of the Gaussian pdf of $\in$, always coincides with the instant $\overline{T_\iota} = 1/\alpha_i 1n(\frac{\bar{\delta}}{\delta_i(T_k^i)})$ at which the expectation of $\delta_i(T_k^i, T)$reaches the threshold $\bar{\delta}$. Note that $\overline{T_\iota}$ is always greater than expectation $E[T_{\bar{\delta}1}]$, meaning that

Such a property can be easily proven by applying the Jensen's inequality to the random variable $T_{\bar{\delta}}^i = g(\in)$.

Neglecting the uncertainty $\in$ in (1) leads to underestimation of the failure probability.

Then, Let $\varphi_i$ be the failure event of the rail stretch $i$ and $\bar{R}_i$ be the maximum tolerable value for the failure risk $R_i = \Pr\{\boldsymbol{\varphi}_1\} D_i(\boldsymbol{\varphi_i})$. Assuming that the loss $D_i(\boldsymbol{\varphi_i})$ is known, the maximum failure probability is $\bar{P}r\{\boldsymbol{\varphi}_i\} = \bar{R}_i / \ D_i(\boldsymbol{\varphi_i})$.

Therefore, the model allows to evaluate that hard deadline $T_H^i$ for the maintenance activity on rail stretch $i$ consisting of the instant at which $F_{T_{\bar{\delta}}^i}(t) \geq \bar{P}r\{\varphi_i\}$. Analogously the soft deadline is defined as the time instant $T_s^i$ at which $F_{T_{\bar{\delta}}^i}(t) \geq \gamma s \bar{P}r\{\varphi_i\}$, $\gamma s < 1$. The soft deadlines should be respected although not in a mandatory way, to minimize the failure probability even if the threshold $\bar{P}r\{\varphi_i\}$ is always guaranteed. Therefore, the soft threshold allows to consider the failure probability not only as a constraint but also to consider the failure probability not only as a constraint but also as a term to be further minimized.

The same approach can be applied to compute the release time, which is defined as the time instant $T_R^i$ at which, given the degradation process, $F_{T_{\bar{\delta}}^i}(t) \geq \gamma s \bar{P}r\{\varphi_i\}$,$\gamma R < \gamma s$, ie., when the failure probability of generic rail stretch $i$ becomes non negligible. By iteratively updating these release times, the subset $A \subseteq R$ of rail stretches, whose maintenance has to be scheduled, can be identified. More details about the iterative definition of the set $A$ will be provided in Section IV. In the case-study section, the customization of the above mentioned model will be discussed with reference to the European Standards [43]–[45] and the Swedish National Railway Regulation [46], which specify safety-related limits for each track geometry parameter.

# 7.Optimization Model for Maintenance Planning

This section describe the mixed integer linear programming (MILP) model. The relevant notation is reported in table II.
In this section, to simplify the problem notation, the reference to RH interval $[t_h, t_h + \delta t]$ is dropped. In other words, all the considered sub-problems are written imposing $t_h = 0, \forall h$. This corresponds to a leftward translation of the deadlines by the quantity $t_h$ on the time axis, which does not affect the characteristics of the solution: the new deadlines are $\hat{T}H = TH - t_h$, and $\hat{T}s = Ts - t_h$,which can be interpreted as the remaining

| Symbol | Meaning |
|---|---|
| $\mathcal{M}$ | set of $\lvert\mathcal{M}\rvert$ maintenance teams |
| $\mathcal{H}$ | set of $\lvert\mathcal{H}\rvert$ frames in the rolling-horizon framework |
| $\omega_i$ | priority of the maintenance activity on the rail stretch $i \in \mathcal{A}$ |
| $\hat{\tau}_H^i$ | translated hard deadline for maintenance activity of rail stretch $i \in \mathcal{A}$ |
| $\hat{\tau}_S^i$ | translated soft deadline for maintenance activity of rail stretch $i \in \mathcal{A}$ |
| $\pi_i^m$ | processing time of maintenance activity on rail stretch $i \in \mathcal{A}$ by team $m \in \mathcal{M}$ |
| $\lambda_c$, $\lambda_q$ | weights of the different terms of the cost function |
| $\mathcal{T}$ | set of $\lvert\mathcal{T}\rvert$ train-free time intervals |
| $r$ | generic train-free interval $\in \mathcal{T}$ |
| $\ell_r$ | length of the train-free interval $r \in \mathcal{T}$ |
| $I_r$ | start time of the train-free interval $r \in \mathcal{T}$ |
| $i^\star$ | tamping machine parking slot nearest to the rail stretch $i$ |
| $\theta_{i,i^\star}$ | tamping machine travel time from the rail stretch $i$ to its nearest parking slot $i^\star$ |
| $s_{i,j}^m$ | setup time required by team $m$ to prepare the maintenance activity of $j$ after that of $i$, if they are executed in the same train-free interval $t$ |
| $s_{i^\star j}^m$ | setup time required by team $m$ to prepare the maintenance activity of $j$ if it is the first of the $(r+1)^{th}$ train-free interval and $i$ was the last of the $r^{th}$ interval |
| $\eta_{i,j}^m$ | actual setup time required by team $m$ to prepare the maintenance activity of $j$ after concluding the maintenance of $i$ |
| $q_i$ | tardiness of the maintenance activity of $i \in \mathcal{A}$ |
| $M$ | large integer number (big M) |
| $x_{i,j}^m$ | binary assignment variable equal to 1 if the activity of $j$ is performed soon after the activity of $i$ by team $m$ and 0 otherwise |
| $w_i^r$ | binary assignment variable equal to 1 if the activity of $i$ is performed during the period $r \in \mathcal{T}$ |
| $t_i^m$, $c_i^m$ | beginning and completion times of the maintenance activity of $i \in \mathcal{A}$ |
| $(\cdot)^\star$ | optimal value of the variable $(\cdot)$ |
| $\mathcal{Y}$ | set gathering the optimization problem variables $c_i, t_i, w_i, x_{i,j}^m, \forall i, j \in \mathcal{A}, \forall m \in \mathcal{M}$ |

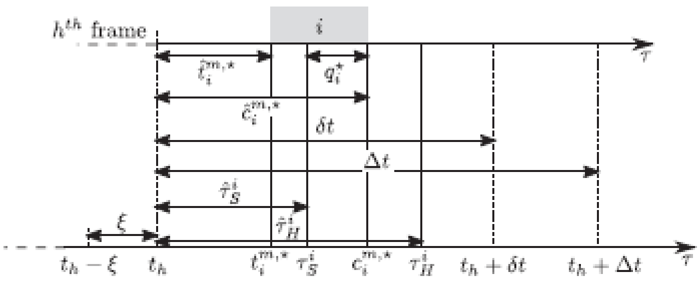


Figure 22: Correspondence between the solution of each sub- problem and the complete maintence problem.

Time until the deadlines at the beginning of the hth frame. Similarly, once the optimization problem is solved, the optimal maintenance starting times $\hat{t}_i^{m,*}$ and completion times $\hat{c}_i^{m,*}$ are translated to the right of quantity $t_h$, that is $t_i^{m,*} = t_h + \hat{t}_i^{m,*}$ and $c_i^{m,*} = t_h + \hat{c}_i^{m,*}$ , respectively. A sketch shows the relations between the relative and absolute times is depicted in Fig.22

**A. Assumptions**

In the formulation of MILP problem, the following assumptions have been considered to model real-world operational aspects.

1) Any interval $[t_h, t_h + \Delta t], h = 0,1 \ldots$ can be divided into (discontinuous) train-free subintervals, gathered in the set $T$, during which train circulation is forbidden and maintenance activities can be performed.

2) Preemption of maintenance activities is not allowed.

3) All maintenance teams are available at the initial time.

4) All maintenance teams are unrelated, and each maintenance activity can be processed by any free team.

5) To allow for normal circulation outside train- free intervals, the tamping machines must be able to reach the nearest parking rail at the end of the last maintenance activity in the each train-free interval Fig.  shows the assignment of the parking slots to the different segments of the rail line. Therefore, the setup time is different depending on whether the consecutive activities on the rail stretches $i$ and $j$ are executed in the same interval □ or not. In particular, the setup time is composed of the travel time from the location of the rail stretch $i$ to the location of rail stretch $j$ if they are maintained in the same interval; on the other hand, if the activity of $j$ is performed in the interval $r + 1$, the setup time consists of the travel time from the tamping machine parking slot $i^*$ nearest to $i$ (usually consisting of a secondary rail in a station).

These assumptions are representative of the considered maintenance process consisting of tamping activities along a rail line. Nevertheless, Assumptions 3 and 4 can be easily removed by modifying the related problem constraints.

**B. MILP Formulation**

This section introduces and describes the optimization problem for the maintenance of a set of assets distributed along a railway line. Thanks to the previously described translation scheme, the proposed formulation is valid for all frames $h \in H$ making up the RH framework. However, the problem in [1} can be used in the place of the one described on this section to easily switch to the network problem. In addition, a new characteristics of the present problem consists of the introduction of a particular constraint aimed at taking into account the time needed by tamping machine to reach a parking slot. Before the is described, two dummy activities $i$ =0 and $j = |A| + 1$ are artificially introduced to a correctly identify the first and last real activities.
Any instance of considered scheduling problem turns out to be

$$y^* = \arg min_y \ J\,(Y) \qquad (6)$$

$$J(Y) = \lambda_C \sum_{m=1}^{|M|} \sum_{i=1}^{|A|} w_i \hat{c}_i^m + \lambda_q \sum_{i=1}^{|A|} q_i \qquad (7)$$

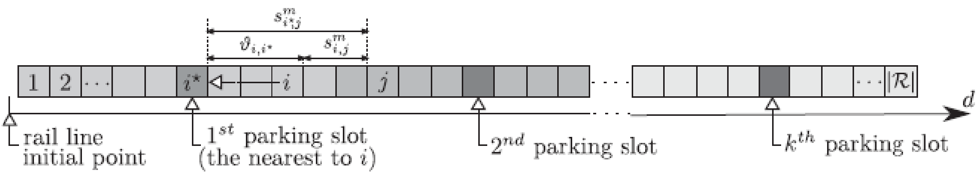


Fig.6. Scheme of the |R| rail stretches and of the k tamping machine parking slots of the considered line. All rail stretches with the same color have the same nearest parking slot. An indication of the setup times is also reported, highlighting the difference between the cases in which j follows i in the same time window and in which i and j are performed in different time frames.

$$\hat{c}_i^m = \hat{t}_i^m + \sum_{j=1}^{|A|+1} \Pi_i^m \, x_{i,j}^m \ \forall m \in M \qquad (8)$$

$$\hat{c}_i^m \leq \widehat{T_H^i} \ \ \forall i \in A, \forall m \in M \qquad (9)$$

$$q_i = \max\{0, \hat{c}_i^m - \hat{T}s^i\} \ \forall i \in A, \forall m \in M \qquad (10)$$

$$\eta_{i,j}^m = \sum_{r=1}^{|T|} s_{i,j}^m w_i^r w_j^r + S_{i*,j}^m (1\text{-}w_i^r) w_j^r$$

$$\forall m \in M, \begin{cases} i = 0, \dots\dots, |A| \\ j = 1, \dots., |A| + 1 \end{cases} \qquad (11)$$

$$\hat{t}_j^m \geq max\{\hat{c}_i^m, I_r w_{j,r}\} + \boldsymbol{\eta}_{i,j}^m - M(1 - x_{i,j}^m)$$

$$\forall r \in T, \forall m \in M, \begin{cases} i = 0, \dots\dots, |A| \\ j = 1, \dots., |A| + 1 \end{cases} \qquad (12)$$

$$\hat{c}_i^m \leq I_r + l_r - \vartheta_{i,i*} + M\,(1 - w_i^r)$$
$$\forall i \in A, \forall r \in T\ , \forall m \in M \qquad (13)$$

$$\sum_{j=1}^{|A|+1} x_{0,j}^m \leq 1, \forall m \in M \qquad (14)$$

$\sum_{m=1}^{|M|} \sum_{i=0,i\neq j}^{|A|+1} x_{i,j}^{m} = 1, \quad j \in A$ (15)

$\sum_{m=1}^{|M|} \sum_{j=1,j\neq 1}^{|A|+1} x_{i,j}^{m} = 1, \ i \in A$ (16)

$\sum_{k=1,h\neq j}^{|A|+1} x_{j,k}^{m} - \sum_{k=0,h\neq j}^{|A|} x_{k,j}^{m} = 0, \ \forall j \in A, \forall m \in M$ (17)

$\sum_{r=1}^{|T|} w_i^r = 1, \quad \forall i = 0,1 \dots, |A| + 1$ (18)

$\hat{t}_i^m, \hat{c}_i^m, q_i \in \mathbb{R} \geq 0, \quad \forall i \in A, \forall m \in M$ (19)

$\eta_{i,j}^m \in \mathbb{R} \ \geq 0, \ \forall m \in M,$
$i = 0, \dots, |A|, j = 1, \dots, |A|+1$ (20)

$x_{i,j}^m \in \{0,1\}, \forall m \in M,$

$\forall i = 0, \dots, |A|, \forall j = 1, \dots, |A| + 1, i \neq j$ (21)

Where:

- The cost function in (7) consists of the weighted sum of the tardiness with respect to the soft deadlines and the total completion time;
- The constraints in (8) define the completion times of the maintenance activities. If asset i is assigned to team m, these constraints, together with the cost function, set all completion times
- the constraints in (9) guarantee that the maintenance activity on each asset $i$ is completed before the relevant hard deadline;
- The constraints in (10) define the tardiness of the maintenance activities with respect to the soft deadline
- the constraints in (11) define the setup time3, which is different if the maintenance activities are executed by maintenance team $m$ in two different train-free intervals as described in assumption 5;
- the constraints in (12) guarantee that if the activity on asset $j$ is performed soon after the activity on asset $i$, it starts after the completion of $i$; at the same time, they guarantee that if the activity on $j$ is the first of the train-free interval $h$, it starts after the beginning of that interval;
- the constraints in (13) guarantee that all maintenance activities finish within the train-free subinterval, leaving enough time to reach the nearest parking slot; while these constraints apply to all activities in a train-free interval, the one associated with the last activity dominates all the others;
- the constraints in (14) guarantee that at most one activity is scheduled as the first work of each maintenance team;
- the constraints in (15) and (16) guarantee that every maintenance activity has exactly one predecessor and one successor (also considering the dummy activities), respectively;
- the constraints in (17) state that a predecessor/successor pair of activities has to be assigned to the same maintenance team m;

**Algorithm 1:** MATHeuristic Solution Algorithm.

$u \leftarrow 0$
$T \leftarrow 0$
**while** $T \leq T_{\max}$ **do**
  randomly choose $\tilde{\mathcal{Y}}^u \subset \mathcal{Y}^u$ and let $\mathcal{Y}_{\text{fix}} = \mathcal{Y}^u \setminus \tilde{\mathcal{Y}}^u$
  $\tilde{\mathcal{Y}}^{u,\star} \leftarrow \arg\min_{\tilde{\mathcal{Y}}^u} J(\tilde{\mathcal{Y}}^u \cup \mathcal{Y}_{\text{fix}})$
  $T \leftarrow T + T_u$
  **if** $J(\tilde{\mathcal{Y}}^{u,\star} \cup \mathcal{Y}_{\text{fix}}) < J(\mathcal{Y}^u)$ **then**
    $\mathcal{Y}^{u+1} \leftarrow \tilde{\mathcal{Y}}^{u,\star} \cup \mathcal{Y}_{\text{fix}}$
  **else**
    $\mathcal{Y}^{u+1} \leftarrow \mathcal{Y}^u$
  **end if**
  $u \leftarrow u + 1$
**end while**

- The constraints in (18) guarantee that each activity is only performed in one and only one train-free subinterval;
- The constraints in (19)–(21) define the problem variables.

The weights $\lambda_c$ and $\lambda_q$ in (7) are chosen by the maintenance service provider in line with the strategic goals of its organization (more importance attributed to minimization of the completion time or to minimization of tardiness with respect to the soft deadline). Regardless, this choice cannot affect the safety level of the system performance since the constraints in (9) guarantee that the hard deadlines are always met. If the maintenance service provider gives a low value to the weight $\lambda_c$ , making fulfillment of the soft deadlines negligible in the cost function, the assets remain in an acceptably degraded condition. Moreover, the obtained experimental results are quite insensitive to the weights $\lambda_c$ and$\lambda_q$, as a variation of the cost function of 10%was obtained by varying the weights by 50%.

**C. Problem solution**

The problem defined by (6) – (21) is Np- hard: therefore, it is characterized by a very high computational effort and requires the definition of effective heuristic strategies to be solved. In this article, a MATHeuristic approach is considered.

The development of the solution algorithm is outside the scope of this article; nevertheless, some information is provided. The algorithm is based on an iterative approach, which provides new solutions by performing successive optimizations of a reduced subset of randomly chosen variables. The considered algorithm is reported in Algorithm 1, while the relevant notation is defined in Table III. The advantage of the hybrid MATHeuristic approach is the possibility of combining the strength of exact methods with the flexibility of an approximated metaheuristic This methodology has been proven to be a very competitive alternative to solving large-scale instances of complex optimization problems [47]. The algorithm starts from an initial admissible top-to-end solution, consisting of executing the activities on the assets as sorted along the line while assigning the first $|A|/2$ activities to the first team and the remaining ones to the second team. The choice of the top-to-end initial solution is due not only to its simplicity but also to the fact that such

TABLE III
NOTATION OF ALGORITHM 1

| Symbol | Meaning |
|---|---|
| $u, T$ | iteration and total solution time counters |
| $T_u$ | solution time of the $u^{th}$ iteration |
| $T_{\max}$ | maximum solution time |
| $\mathcal{Y}^u$ | best solution at the beginning of the $u^{th}$ iteration |
| $\tilde{\mathcal{Y}}^u$ | subset of $\mathcal{Y}^u$ |
| $\mathcal{Y}_{\text{fix}}$ | fixed variables in the $u^{th}$ iteration |
| $\tilde{\mathcal{Y}}^{u,\star}$ | optimal value of $\tilde{\mathcal{Y}}^u$ at the end of the $u^{th}$ iteration |
| $J(\mathcal{Y}^u)$ | best cost function at the beginning of the $u^{th}$ iteration |

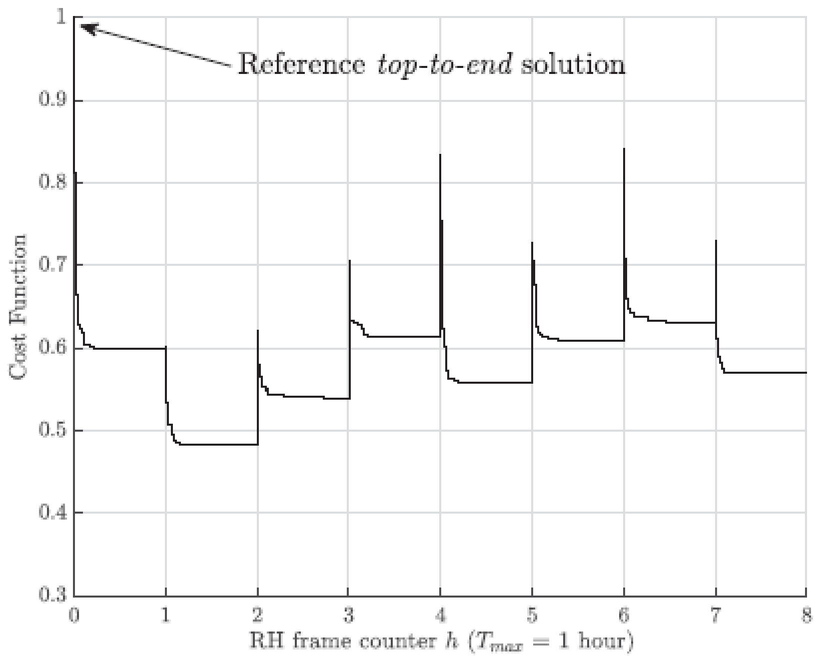


Figure 23: Cost function values over 8 RH frames for the case study described in Section VI

a strategy is often applied in the real-world maintenance of geographically distributed assets. Regarding the performance of the *MAT Heuristic* technique, the applications of the proposed approach to the considered case study yield average reductions of approximately 45% of the cost function with respect to the reference top-to-end solution

As reported in Algorithm 1, at the end of each iteration, if the whole solution improves, the new values of the optimized variables are accepted; otherwise, they are dropped. Actually, the solution never worsens, as$\tilde{y}^u \cup \tilde{y}_{fix}$. Is always a feasible solution; nevertheless, if $\tilde{y}^{u,*} = \tilde{y}^u$, the whole solution does not improve, as shown in fig. 7, where the flat shapes indicate that there are no improvements. The iterative approach is applied for $T_{max}$ hours. As the solution never worsens within an RH window, the peaks in Fig. 7 represent the transition between RH frames: since new maintenance activities are considered, the cost function may be much greater than the optimal value of the previous period. Finally, since the final solution of one RH frame represents the initial solution for the next one, this initial solution may be infeasible. This infeasibility is simply overcome by the *MATHeuristic* approach by varying the subset of considered variables. In particular, in the new RH window, if the previous solution is infeasible, the first iterations of the*MATHeuristic* will find a feasible initial solution for that RH window, while the successive iterations will optimize the feasible initial solution by randomly choosing a subset of variables. Therefore, the Fig. 8. B&B and MATHeuristic results comparison. The reference point (100%) corresponds to the optimal cost function found via the B&B for an instance with 15 maintenance activities. Note that the best B&B solution is certainly not the optimal one, as Algorithm 1 finds better ones.

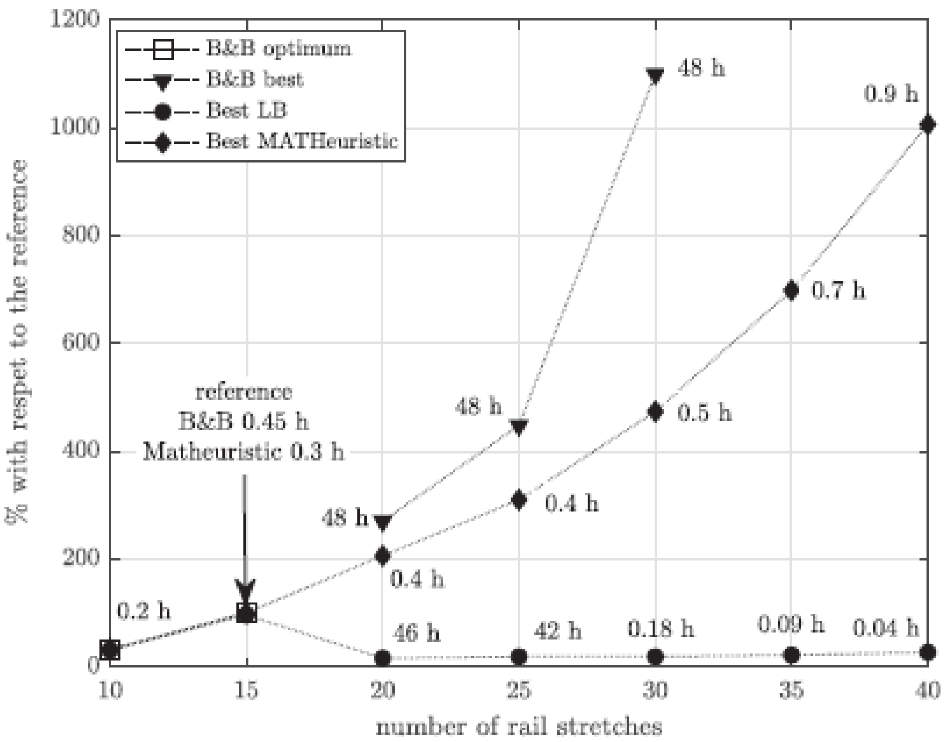


Figure 24: B&B and MATHeuristic results comparison. The reference point (100%) corresponds to the optimal cost function found via the B&B for an instance with 15 maintenance activities. Note that the best B&B solution is certainly not the optimal one, as Algorithm 1 finds better ones

process for finding the initial feasible solution and the process for improving it are the same.

More details regarding the approach are available in [47] and [1] where a general framework of the *MATHeuristic* approach and an example of its application to railway maintenance are provided.

Regarding the comparison of the *MATHeuristic* performance with respect to that provided by the generic branch and bound (B&B) approach implemented by the IBM-Ilog Cplex solver, the outputs are reported in Fig. 8 for different instance dimensions of the problem. In particular, in this figure, the relative values of the cost functions for each solution are depicted, with the relevant labels being the time required to find the solutions.

The generic B&B approach does not find the optimal solution for instances with $|A| \geq 20$ or even a feasible solution for instances with $|A| \geq 35$, considering a maximum running time of 48 h. Nevertheless, in less than one hour, the considered MATHeuristic finds better solutions.

Moreover, even the lower bound (LB), which is the solution of a relaxed problem providing an evaluation of the goodness of the solution, is very hard to find; in fact, for instances with $|A| \geq 20$ the best LBs are far from the best solutions. In addition, for instances with $|A| \geq 30$, the LB found in the initial instant of the solution search is not improved

**D. Solution Example**

In this section, an example is reported with the aim of explaining the RH approach in detail. In this regard, for the sake of clarity, it is assumed that $t_1 = 0, \delta t = 1$ days and $\Delta t = 4$ days. Then, Let $|A| = \{1,3,8,10,12,19,23,26,29,30\}$ be the subset of rail stretches that need to be considered in the maintenance plan

Example 1. Delay of an activity: In this example, during the first time window, the problem in section V-B is able to fins the optimal solution that assigns the activities $M_1 = \{3; 10; 19; 12; 8\}$ to the first team and $M_2 = \{30; 23,26; 29; 1\}$

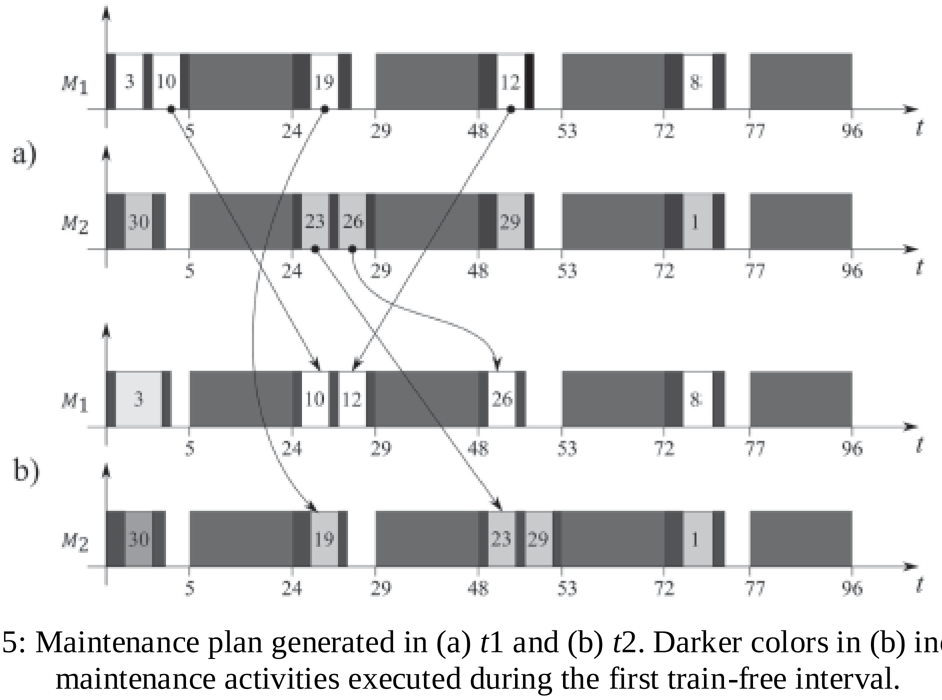


Figure 25: Maintenance plan generated in (a) $t1$ and (b) $t2$. Darker colors in (b) indicate the maintenance activities executed during the first train-free interval.

to the second team. In such sets, the maintenance activities assigned to different train-free intervals are separated by semicolons, and the relevant representation is depicted in Fig. 25, where the dark gray boxes represent the intervals in which trains circulate and maintenance activities cannot be performed, while the black boxes represent the maintenance teams' travel time. Finally, the white and light gray boxes represent the maintenance activities assigned to the first team and the second team, respectively. With reference to the RH framework, the dynamic evolves as follows:

1) at $t_1$, the maintenance teams start their maintenance activities and at the end of the first train-free nterval, maintenance activity 30 is finished on schedule, while a delay in activity 3 makes activity 10 impossible to finish within the first train-free interval. Therefore, that maintenance activity has to be reconsidered in the problem stated for the interval $t_1, t_2 + \Delta t)$;

2) the new schedule to be applied in $t_2$ is $M_1 = \{3; 10, 12; 26; 8\}$ for the first team and $M_2 = \{30; 19; 23, 29; 1\}$ for the second team, where the bold entries indicate the already- executed maintenance activities

The solutions depicted in Fig.9 show that due to the rescheduling of the activity on rail stretch 10, in the new plan, some activities previously assigned to one team are then assigned to the other one, some activities are brought forward, and others are delayed. Nevertheless, despite some modifications, the activities scheduler is able to keep all the maintenance activities within the first four train-free intervals

# 8.Case Study

The considered case study is about testing a maintenance plan provided by an optimization algorithm; this optimization algorithm assigns the set of jobs and assets to be maintained by certain number of teams. We have two teams optimizing the sequence. The sequence and assignment to teams is performed by the ExtendSim and from our perspective as an input.

## 8.1 MODEL DESCRIPTION

ExtendSim is a powerful, leading edge simulation tool. Using ExtendSim, you can develop dynamic models of real-life processes in a wide variety of fields. Use ExtendSim to create models from building blocks, explore the processes involved, and see how they relate. Then change assumptions to arrive at an optimum solution. ExtendSim and your imagination are all you need to create professional models that meet your business, industrial, and academic needs.

ExtendSim is an easy-to-use, yet extremely powerful, tool for simulating processes. It helps you Understand complex systems and produce better results faster.

With ExtendSim you can:

- Predict the course and results of certain actions
- Gain insight and stimulate creative thinking
- Visualize your processes logically or in a virtual environment
- Identify problem areas before implementation
- Explore the potential effects of modifications
- Confirm that all variables are known
- Optimize your operations
- Evaluate ideas and identify inefficiencies
- Understand why observed events occur
- Communicate the integrity and feasibility of your plans

In the simplest terms, ExtendSim models are made up of blocks and connection. you can see in the model window. As the model runs, information goes into a block, is processed and/or modified, and is then sent on to the next block via a connection.

**Blocks**

Each block in ExtendSim represents a portion of the process or system that is being modeled. Blocks have names, such as Math or Queue, that signify the function they perform. A Queue block, for example, will have the same functional behavior in every model you build. You can also add your own label to a block to indicate what it represents in your specific model, such as a Queue block labeled Waiting Line.

**Create**

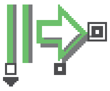

Figure 26: Create Block

Generates items or values, either randomly or on schedule. If used to generate items, it pushes them into the simulation and should be followed by a queue-type block.

**Shift**

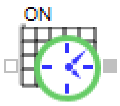


Figure 27: Shift Block

This block provides a schedule; we can work from “00:00am

**Queue**

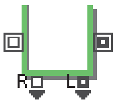

Figure 28: Queue Block

Acts as a sorted queue or as a resource pool queue. As a sorted queue, holds items in FIFO or LIFO order, or sorts items based on their attribute or priority.

**Activity**

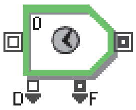

Figure 29: Activity Block

Processes one or more items simultaneously. Processing time is a constant or is based on a distribution or an item’s attribute.

**History**

Figure 30: History Block

Reads the information of an item as soon as an item passes through these blocks and has all the information

**Exit**

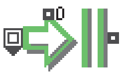

Figure 31: Exit Block

Removes items from the simulation and counts them as they leave.

**Icons**

A block's icon is usually a pictorial representation of its function. Its icon symbolizes an actual tank that can have quantities added or removed from it. The small squares attached to the sides of the icon are connectors, which are discussed in more detail in the following section.

**Connectors**

Most blocks in ExtendSim have input and output connectors (the small squares attached to the Block). As you might expect, information flows into a block at input connectors and out of the Block at output connectors. A block can have many input and/or output connectors; some blocks have none. For instance, an input connector on the left for values to enter. The output connector on the right reports the results of the block's computations; in the tank it reports the contents at each time step. Additional inputs on the bottom are for controlling specific item behavior.

The main phases are:

- Open the relevant libraries, if necessary. (Most ExtendSim libraries are automatically opened When you launch ExtendSim)
- Add the blocks to the model.
- Move them to the desired positions.
- Add connections between blocks.

Now that you have placed blocks on the model, connected them, and configured their dialogs with data, it is time to run the model.

Select Run > Run Simulation or click the Run Simulation button in the toolbar.

In the following the model will be described in detail.

The model would be created using the following,

- **Creation of maintenance activities**

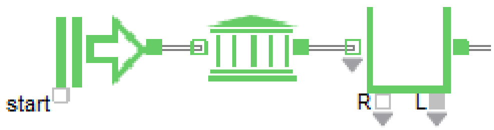


Figure 32: Creation of maintenance activities

.

The model starts with a create block which creates items (maintenance activities to be performed) and assign to each of them a position attribute.

- **Degradation attribute**

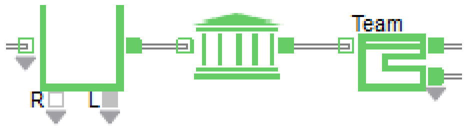


Figure 33: Degradation attribute

The Hard deadline of intervention, that is the time deadline within which the maintenance should be executed and

The Soft deadline of intervention, that is the time deadline within which the maintenance should be executed in order to avoid delay, is generated according to the degradation model described in Chapter 5.

- **Maintenance activities**

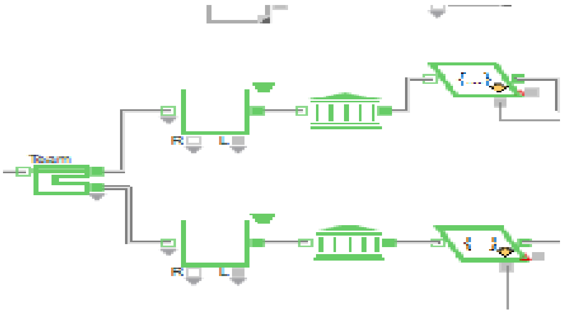

Figure 34: Activities are divided into two teams

The assets enter into an activity block named preventive maintenance. For uncertainties we attach a random number block, representing the execution time of the maintenance activity.

- **Resources availability**

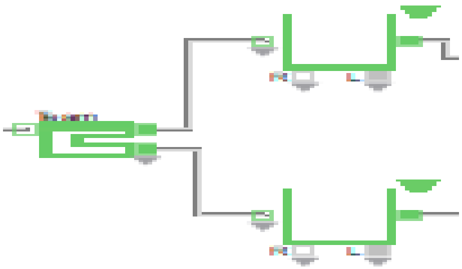

Figure 35: Resources Availability

Resource pool release block which releases the available maintenance team. The total number of resources is defined in the Resources pool block.

The set block guarantees that when a maintenance team is performing a maintenance activity, the available resources are reduced of one unit.

- **History Block**

Figure 36: History Block

The history blocks reports items statistics about the starting time and ending time of each activity.

- **Exit Block**

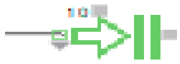

Figure 37: Exit Block

Finally, the exit block passes items out of the simulation.

We would be considering scenarios for a comparison as follows:

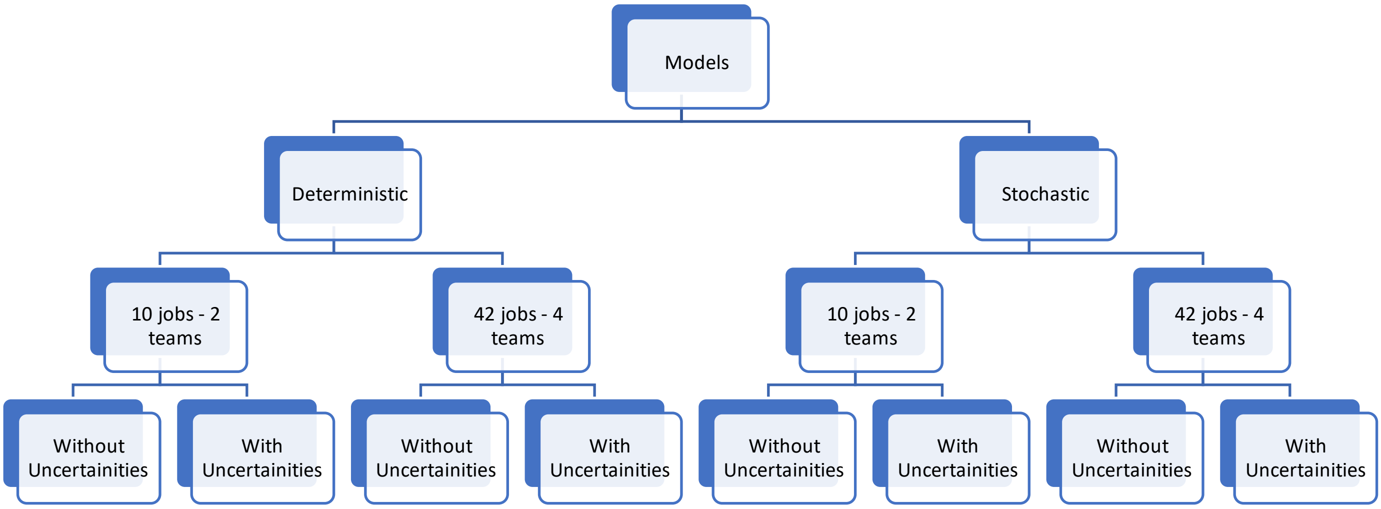


Figure 38: Pictorial representation of the scenarios that are considered in Deterministic and Stochastic models

Uncertainties is a word defined to the state of limited knowledge of activities and reasons that could cause a delay in the processing time.

## Deterministic Model

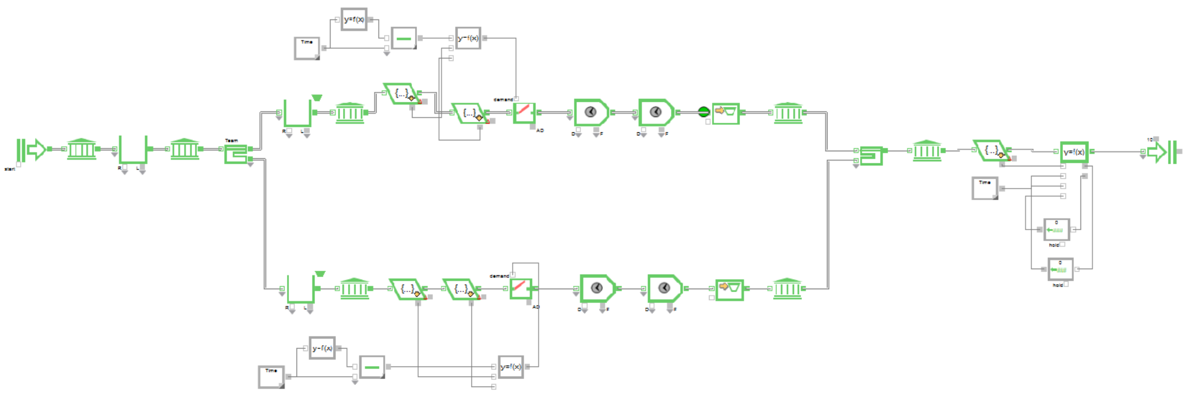

Figure 39: Scenario D1 - Deterministic Model- 10 jobs without uncertainties

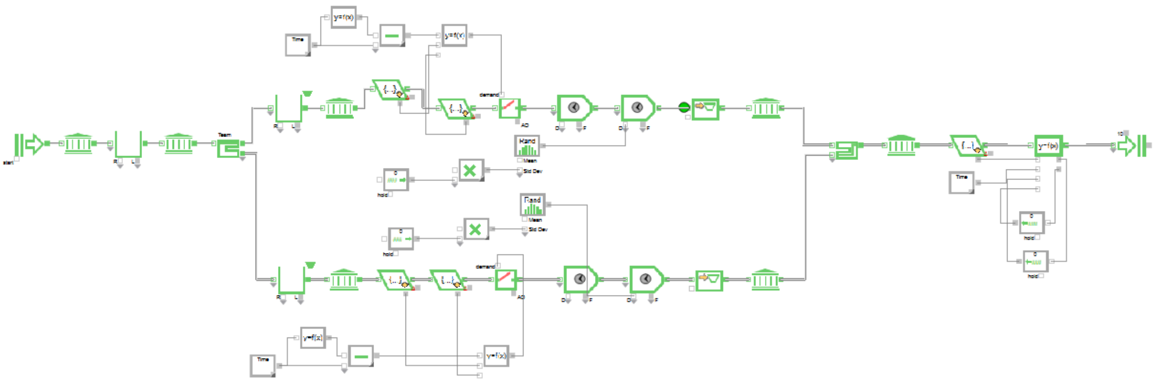

Figure 40: Scenario D2 - Deterministic Model- 10 jobs with uncertainties

The jobs are assigned as per the table 3, positioning them, assigning them to respective teams and mentioning the deadlines. This data is entered duly in create block of Scenario D1 - Deterministic Model- 10 jobs without uncertainties and for Scenario D2 - Deterministic Model- 10 jobs with uncertainties.

Table 3: Input data for 10 jobs and 2 teams

| | create time | item quantity | item priority | Position | Teams | Sequence | Processing Time | Hard Deadline | Soft Deadline | travel time |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 1 | 1 | 2 | 5 | 120 | 6200 | 3500 | 104 |
| 2 | 0 | 1 | 1 | 2 | 1 | 1 | 110 | 7400 | 2800 | 16 |
| 3 | 0 | 1 | 1 | 3 | 1 | 5 | 130 | 6700 | 4500 | 32 |
| 4 | 0 | 1 | 1 | 4 | 1 | 2 | 115 | 7400 | 5300 | 38 |
| 5 | 0 | 1 | 1 | 5 | 1 | 4 | 125 | 8500 | 7400 | 41 |
| 6 | 0 | 1 | 1 | 6 | 1 | 3 | 138 | 6300 | 6100 | 44 |
| 7 | 0 | 1 | 1 | 7 | 2 | 2 | 112 | 6500 | 4200 | 41 |
| 8 | 0 | 1 | 1 | 8 | 2 | 3 | 118 | 7000 | 5800 | 26 |
| 9 | 0 | 1 | 1 | 9 | 2 | 4 | 122 | 8000 | 6900 | 26 |
| 10 | 0 | 1 | 1 | 10 | 2 | 1 | 128 | 6300 | 6000 | 97 |

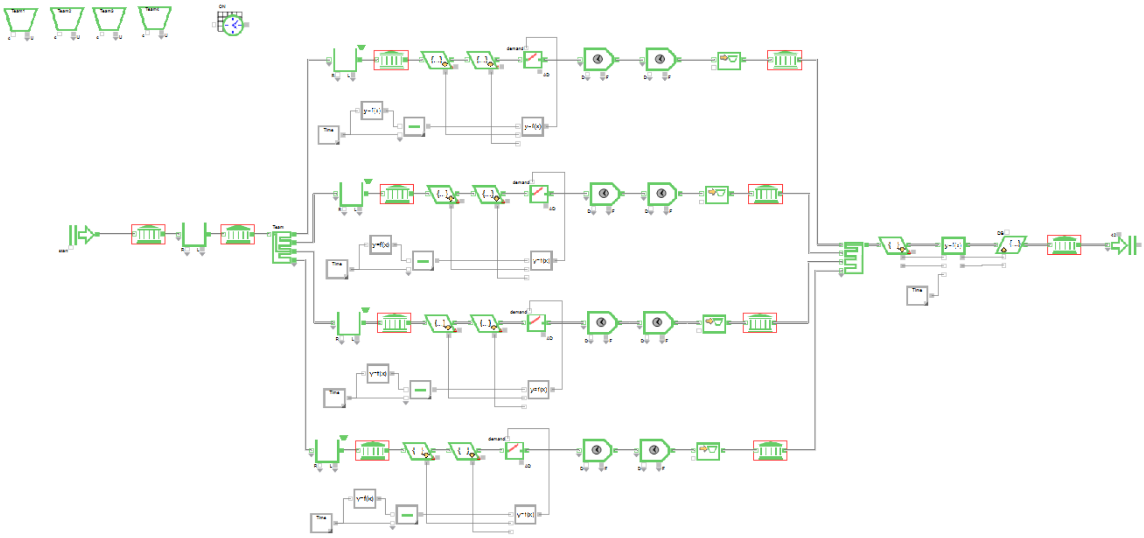

Figure 41: Scenario D3 - Deterministic Model- 42 jobs without uncertainties

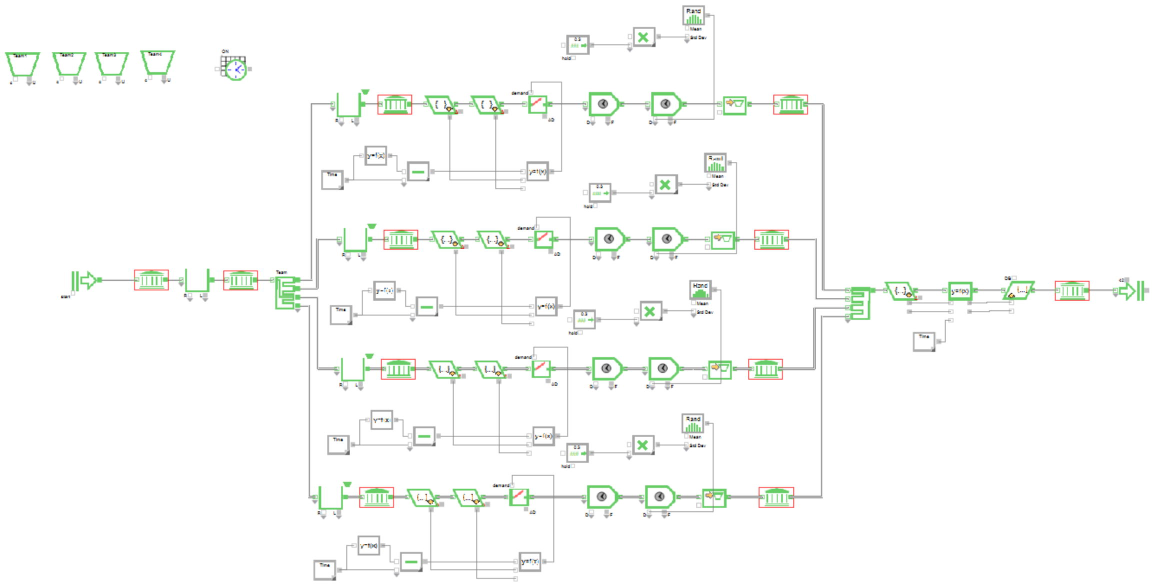

Figure 42: Scenario D4 - Deterministic Model- 42 jobs with uncertainties

The jobs are assigned as per the table 4, positioning them, assigning them to respective teams and mentioning the deadlines. This data is entered duly in create block of Scenario D3 - Deterministic Model- 42 jobs without uncertainties and for Scenario D4 - Deterministic Model- 42 jobs with uncertainties.

Table 4: Input data for 42 jobs and 4 teams

| Create time | Item quantity | Item priority | Position | Teams | Sequence | Processing Time | Hard Deadline | Soft Deadline | travel time |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 1 | 1 | 1 | 7 | 113 | | | 35 |
| 0 | 1 | 1 | 2 | 1 | 1 | 85 | | | 85 |
| 0 | 1 | 1 | 3 | 3 | 9 | 113 | | | 44 |
| 0 | 1 | 1 | 4 | 4 | 9 | 92 | | | 62 |
| 0 | 1 | 1 | 5 | 4 | 8 | 115 | | | 23 |
| 0 | 1 | 1 | 6 | 4 | 1 | 81 | | | 11 |
| 0 | 1 | 1 | 7 | 3 | 5 | 92 | | | 10 |
| 0 | 1 | 1 | 8 | 4 | 2 | 73 | | | 23 |
| 0 | 1 | 1 | 9 | 1 | 4 | 101 | | | 14 |
| 0 | 1 | 1 | 10 | 3 | 4 | 95 | | | 116 |
| 0 | 1 | 1 | 11 | 2 | 7 | 101 | | | 30 |
| 0 | 1 | 1 | 12 | 1 | 8 | 80 | | | 40 |
| 0 | 1 | 1 | 13 | 3 | 1 | 122 | | | 47 |
| 0 | 1 | 1 | 14 | 2 | 1 | 129 | | | 27 |
| 0 | 1 | 1 | 15 | 4 | 3 | 129 | | | 32 |
| 0 | 1 | 1 | 16 | 1 | 9 | 129 | | | 99 |
| 0 | 1 | 1 | 17 | 4 | 4 | 80 | | | 42 |
| 0 | 1 | 1 | 18 | 3 | 2 | 82 | | | 13 |
| 0 | 1 | 1 | 19 | 4 | 6 | 106 | | | 22 |
| 0 | 1 | 1 | 20 | 1 | 11 | 86 | | | 136 |
| 0 | 1 | 1 | 21 | 2 | 6 | 112 | | | 37 |
| 0 | 1 | 1 | 22 | 1 | 10 | 139 | | | 95 |
| 0 | 1 | 1 | 23 | 1 | 6 | 112 | | | 128 |
| 0 | 1 | 1 | 24 | 3 | 10 | 83 | | | 8 |
| 0 | 1 | 1 | 25 | 2 | 11 | 110 | | | 38 |
| 0 | 1 | 1 | 26 | 1 | 2 | 94 | | | 95 |
| 0 | 1 | 1 | 27 | 2 | 4 | 132 | | | 55 |
| 0 | 1 | 1 | 28 | 4 | 7 | 85 | | | 53 |
| 0 | 1 | 1 | 29 | 3 | 3 | 74 | | | 74 |
| 0 | 1 | 1 | 30 | 3 | 6 | 71 | | | 13 |
| 0 | 1 | 1 | 31 | 1 | 3 | 82 | | | 20 |
| 0 | 1 | 1 | 32 | 2 | 10 | 124 | | | 55 |
| 0 | 1 | 1 | 33 | 3 | 7 | 116 | | | 43 |
| 0 | 1 | 1 | 34 | 3 | 8 | 139 | | | 109 |
| 0 | 1 | 1 | 35 | 3 | 11 | 112 | | | 108 |
| 0 | 1 | 1 | 36 | 2 | 9 | 85 | | | 53 |
| 0 | 1 | 1 | 37 | 4 | 5 | 99 | | | 42 |
| 0 | 1 | 1 | 38 | 2 | 5 | 92 | | | 65 |
| 0 | 1 | 1 | 39 | 1 | 5 | 83 | | | 63 |
| 0 | 1 | 1 | 40 | 2 | 2 | 108 | | | 33 |
| 0 | 1 | 1 | 41 | 2 | 8 | 122 | | | 30 |
| 0 | 1 | 1 | 42 | 2 | 3 | 95 | | | 30 |

## Stochastic Models

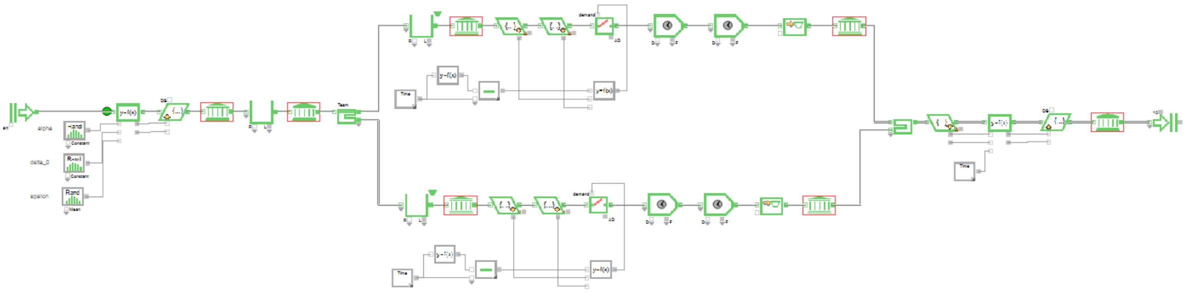

Figure 43: Scenario S1 - Stochastic Model- 10 jobs without uncertainties

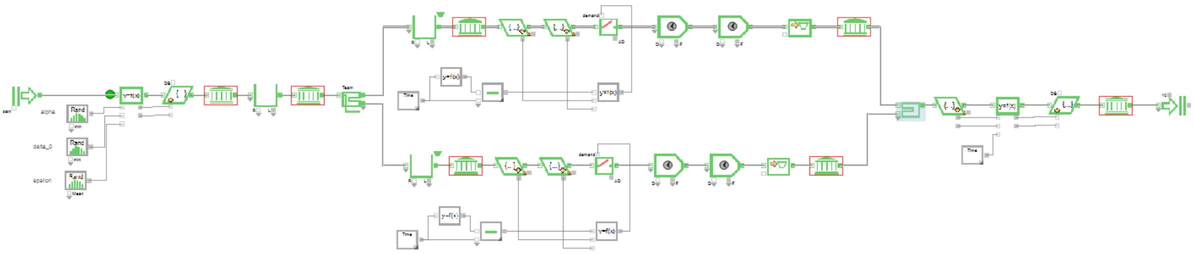

Figure 44: Scenario S2 - Stochastic Model- 10 jobs with uncertainties

The jobs are assigned as per the table 3, positioning them, assigning them to respective teams and mentioning the deadlines. This data is entered duly in create block of Scenario S1 - Stochastic Model- 10 jobs without uncertainties and for Scenario S2 - Stochastic Model- 10 jobs with uncertainties.

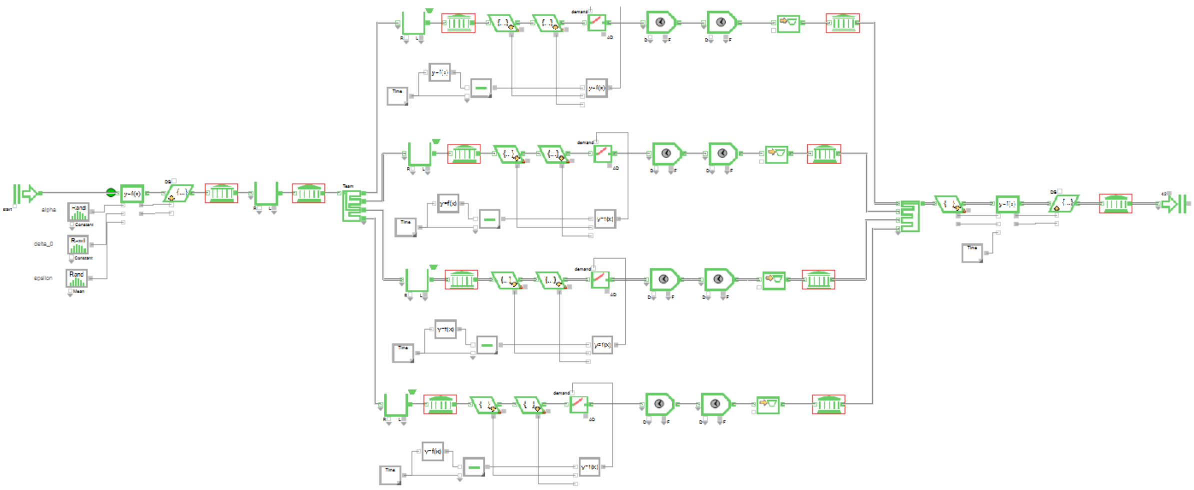

Figure 45: Scenario S3 - Stochastic Model- 42 jobs without uncertainties

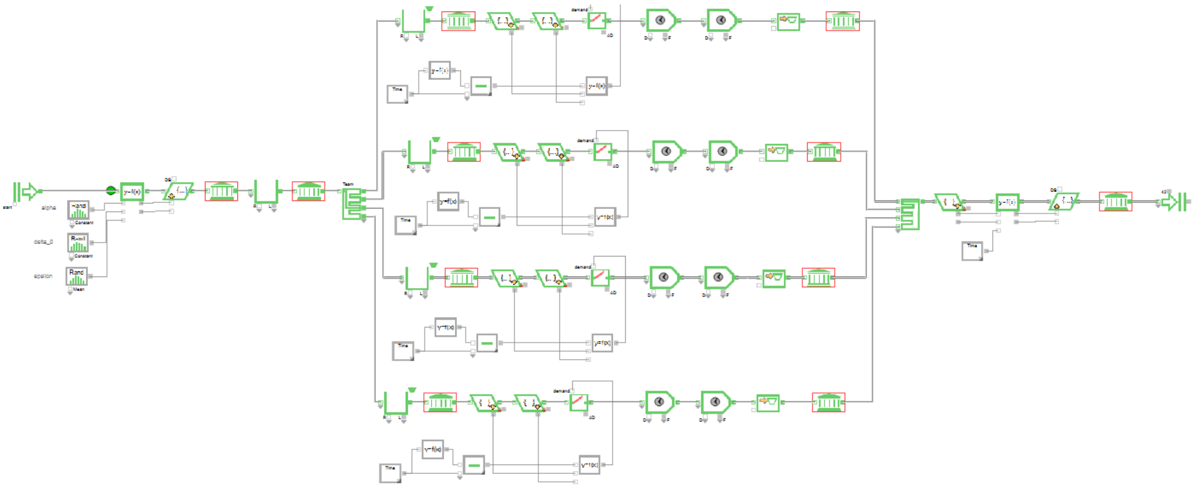

Figure 46: Scenario S4 - Stochastic Model- 42 jobs with uncertainties

The jobs are assigned as per the table 4, positioning them, assigning them to respective teams and mentioning the deadlines. This data is entered duly in create block of Scenario S3 - Stochastic Model- 42 jobs without uncertainties and for Scenario S4 - Stochastic Model- 42 jobs with uncertainties.

In the above model's addition of uncertainties differs between deterministic and stochastic models.

Uncertainties, are considered in terms of processing time for both deterministic and stochastic models.

In deterministic model, uncertainties are introduced using the following, and then calculating the standard deviation of processing time.

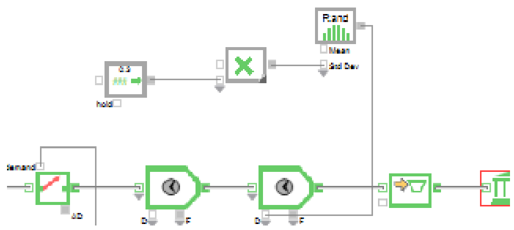


Figure 47: Uncertainties in Deterministic Model

And, we compare the delay based on the total time taken.

Whereas, in Stochastic models the variation of alpha, delta and epsilon numbers from constant to Uniform real is our approach to introduce uncertainties. We modify the model by considering the equation

$$T_{\delta}^{i} = g(\epsilon) = \frac{1}{\alpha_i} 1n \frac{\bar{\delta} - \epsilon}{\delta_i(T_k^i)}$$

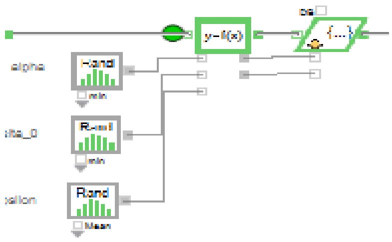

Figure 48: Uncertainties in Stochastic Model

And we compare delay in terms of hard and soft deadlines.

# 9.Results

## Comparison of Delay/Earliness time between Deterministic and Stochastic model for two teams and 10 jobs

Comparison of Delay/Earliness in Hard Deadlines

| Position | D1 | D2-0% | D2-10% | D2-20% | D2-30% | S1 | S2-Sample1 | S2-Sample1 |
|---|---|---|---|---|---|---|---|---|
| 1 | 1432 | -18 | -18.0426 | -17.9871 | -18.275 | 0 | 0 | 0 |
| 2 | 7148 | 7128 | 7128.163 | 7127.949 | 7128.358 | 0 | 0 | 0 |
| 3 | 2056 | 2056 | 2055.968 | 2055.788 | 2056.195 | 0 | 0 | 0 |
| 4 | 6968 | 5793 | 5793.133 | 5792.9 | 5792.88 | 0 | 0 | 0 |
| 5 | 5288 | 5283 | 5283.072 | 5283.173 | 5282.934 | 0 | 0 | 0 |
| 6 | 4496 | 4490 | 4490.155 | 4489.897 | 4489.477 | 0 | 0 | 0 |
| 7 | 4754 | 4736 | 4735.898 | 4735.831 | 4736.136 | 0 | 0 | 0 |
| 8 | 5119 | 3820 | 3819.846 | 3820.246 | 3819.884 | 0 | 0 | 0 |
| 9 | 4824 | 3376 | 3376.076 | 3375.876 | 3375.488 | 0 | 0 | 0 |
| 10 | 5850 | 5848 | 5847.975 | 5847.756 | 5848.178 | 0 | 0 | 0 |

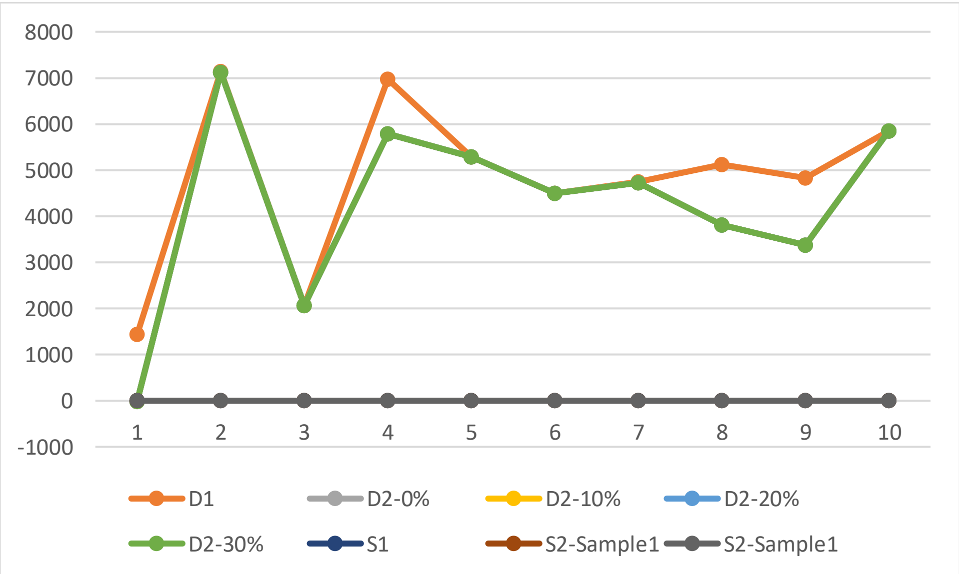


Figure 49: Comparison of Delay/Earliness in Hard Deadlines obtained in Deterministic and Stochastic models

Comparison of Delay/Earliness in Soft Deadlines

| Position | D1 | D2-0% | D2-10% | D2-20% | D2-30% | S1 | S2-Sample1 | S2-Sample1 |
|---|---|---|---|---|---|---|---|---|
| 1 | -1268 | -2718 | -2718.04 | -2717.99 | -2718.27 | 0 | 0 | 0 |
| 2 | 2548 | 2528 | 2528.163 | 2527.949 | 2528.358 | 0 | 0 | 0 |
| 3 | -144 | -144 | -144.032 | -144.212 | -143.805 | 0 | 0 | 0 |
| 4 | 4868 | 3693 | 3693.133 | 3692.9 | 3692.88 | 0 | 0 | 0 |
| 5 | 4188 | 4183 | 4183.072 | 4183.173 | 4182.934 | 0 | 0 | 0 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 6 | 4296 | 4290 | 4290.155 | 4289.897 | 4289.477 | 0 | 0 | 0 |
| 7 | 2454 | 2436 | 2435.898 | 2435.831 | 2436.136 | 0 | 0 | 0 |
| 8 | 3919 | 2620 | 2619.846 | 2620.246 | 2619.884 | 0 | 0 | 0 |
| 9 | 3724 | 2276 | 2276.076 | 2275.876 | 2275.488 | 0 | 0 | 0 |
| 10 | 5550 | 5548 | 5547.975 | 5547.756 | 5548.178 | 0 | 0 | 0 |

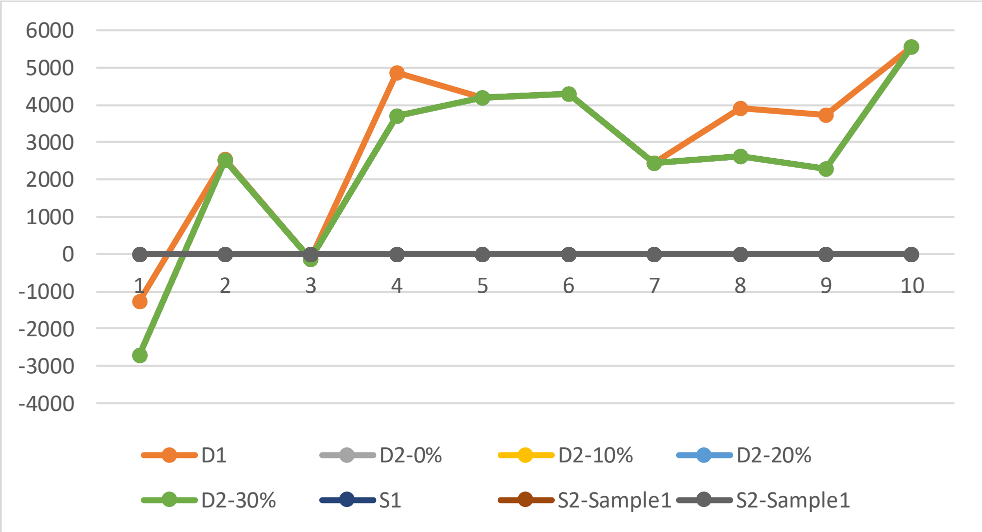


Figure 50: Comparison of Delay/Earliness in Soft Deadlines obtained in Deterministic and Stochastic models

## Comparison of Delay/Earliness time between Deterministic and Stochastic model for four teams and 42 jobs

Comparison of Delay/Earliness in Hard Deadlines

| Position | D3 | D4-0% | D4-10% | D4-20% | D4-30% | S3-delayHard | S4 - Sample 1-delayHard | S4 - Sample2 - delayHard | S4 - Sample3 - delayHard | S4 - Sample4 - delayHard |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | -149 | -149 | -149 | -149 | -149 | 0 | 979.8968 | 0 | 4.460279 | 0 |
| 2 | -171 | -171 | -171 | -171 | -171 | 0 | 0 | 0 | 0 | 0 |
| 3 | -158 | -158 | -158 | -158 | -158 | 1143.576 | 0 | 294.9675 | 1325.298 | 935.0132 |
| 4 | -155 | -155 | -155 | -155 | -155 | 0 | 0 | 702.1965 | 0 | 861.7695 |
| 5 | -139 | -139 | -139 | -139 | -139 | 0 | 0 | 0 | 0 | 0 |
| 6 | -93 | -93 | -93 | -93 | -93 | 0 | 0 | 0 | 0 | 0 |
| 7 | -103 | -103 | -103 | -103 | -103 | 0 | 0 | 0 | 0 | 0 |
| 8 | -97 | -97 | -97 | -97 | -97 | 0 | 0 | 0 | 0 | 0 |
| 9 | -116 | -116 | -116 | -116 | -116 | 0 | 0 | 0 | 0 | 0 |
| 10 | -212 | -212 | -212 | -212 | -212 | 0 | 0 | 0 | 0 | 0 |
| 11 | -132 | -132 | -132 | -132 | -132 | 0 | 0 | 0 | 0 | 0 |
| 12 | -121 | -121 | -121 | -121 | -121 | 0 | 321.6798 | 1820.716 | 744.4565 | 0 |
| 13 | -170 | -170 | -170 | -170 | -170 | 0 | 0 | 0 | 0 | 0 |
| 14 | -157 | -157 | -157 | -157 | -157 | 0 | 0 | 0 | 0 | 0 |
| 15 | -162 | -162 | -162 | -162 | -162 | 0 | 0 | 0 | 0 | 0 |
| 16 | -229 | -229 | -229 | -229 | -229 | 1214.576 | 2944.465 | 995.1147 | 3506.225 | 705.2955 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 17 | -123 | -123 | -123 | -123 | -123 | 0 | 0 | 0 | 0 | 0 |
| 18 | -96 | -96 | -96 | -96 | -96 | 0 | 0 | 0 | 0 | 0 |
| 19 | -129 | -129 | -129 | -129 | -129 | 0 | 0 | 0 | 0 | 0 |
| 20 | -223 | -223 | -223 | -223 | -223 | 4088.576 | 4558.129 | 5499.853 | 3338.621 | 4934.877 |
| 21 | -150 | -150 | -150 | -150 | -150 | 0 | 0 | 0 | 0 | 0 |
| 22 | -235 | -235 | -235 | -235 | -235 | 2660.576 | 1925.71 | 1135.591 | 4899.252 | 2841.443 |
| 23 | -241 | -241 | -241 | -241 | -241 | 0 | 0 | 0 | 0 | 0 |
| 24 | -92 | -92 | -92 | -92 | -92 | 0 | 0 | 0 | 0 | 0 |
| 25 | -149 | -149 | -149 | -149 | -149 | 2574.576 | 2242.523 | 1854.109 | 2306.564 | 3001.06 |
| 26 | -190 | -190 | -190 | -190 | -190 | 0 | 0 | 0 | 0 | 0 |
| 27 | -188 | -188 | -188 | -188 | -188 | 0 | 0 | 0 | 0 | 0 |
| 28 | -139 | -139 | -139 | -139 | -139 | 0 | 0 | 0 | 0 | 0 |
| 29 | -149 | -149 | -149 | -149 | -149 | 0 | 0 | 0 | 0 | 0 |
| 30 | -85 | -85 | -85 | -85 | -85 | 0 | 0 | 0 | 0 | 0 |
| 31 | -103 | -103 | -103 | -103 | -103 | 0 | 0 | 0 | 0 | 0 |
| 32 | -180 | -180 | -180 | -180 | -180 | 1165.576 | 1936.642 | 1632.624 | 0 | 1233.031 |
| 33 | -160 | -160 | -160 | -160 | -160 | 0 | 0 | 0 | 0 | 0 |
| 34 | -249 | -249 | -249 | -249 | -249 | 0 | 0 | 0 | 1553.583 | 0 |
| 35 | -221 | -221 | -221 | -221 | -221 | 2646.576 | 1690.095 | 3443.482 | 3172.186 | 1419.88 |
| 36 | -139 | -139 | -139 | -139 | -139 | 0 | 797.3266 | 0 | 1086.003 | 207.7981 |
| 37 | -142 | -142 | -142 | -142 | -142 | 0 | 0 | 0 | 0 | 0 |
| 38 | -158 | -158 | -158 | -158 | -158 | 0 | 0 | 0 | 0 | 0 |
| 39 | -147 | -147 | -147 | -147 | -147 | 0 | 0 | 0 | 0 | 0 |
| 40 | -142 | -142 | -142 | -142 | -142 | 0 | 0 | 0 | 0 | 0 |
| 41 | -153 | -153 | -153 | -153 | -153 | 0 | 619.7052 | 0 | 0 | 1208.277 |
| 42 | -126 | -126 | -126 | -126 | -126 | 0 | 0 | 0 | 0 | 0 |
| Average | -154.119 | -154.119 | -154.119 | -154.119 | -154.119 | 368.9056 | 428.9565 | 413.7775 | 522.3011 | 413.0582 |

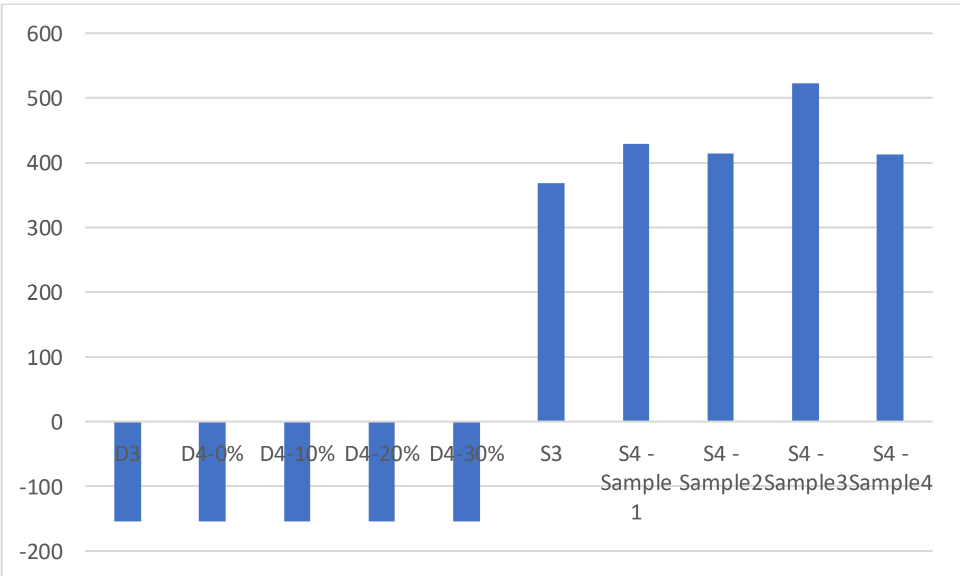


Figure 51: Comparison of Average Delay/Earliness in Hard Deadlines obtained in Deterministic and Stochastic models

Comparison of Delay/Earliness in Soft Deadlines

| Position | D3 | D4-0% | D4-10% | D4-20% | D4-30% | S3 | S4 - Sample 1 | S4 - Sample2 | S4 - Sample3 | S4 - Sample4 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | -149 | -149 | -149 | -149 | -149 | 0 | 0 | 0 | 0 | 0 |
| 2 | -171 | -171 | -171 | -171 | -171 | 0 | 0 | 0 | 0 | 0 |
| 3 | -158 | -158 | -158 | -158 | -158 | 0 | 0 | 0 | 0 | 0 |
| 4 | -155 | -155 | -155 | -155 | -155 | 0 | 0 | 0 | 0 | 0 |
| 5 | -139 | -139 | -139 | -139 | -139 | 0 | 0 | 0 | 0 | 0 |
| 6 | -93 | -93 | -93 | -93 | -93 | 0 | 0 | 0 | 0 | 0 |
| 7 | -103 | -103 | -103 | -103 | -103 | 0 | 0 | 0 | 0 | 0 |
| 8 | -97 | -97 | -97 | -97 | -97 | 0 | 0 | 0 | 0 | 0 |
| 9 | -116 | -116 | -116 | -116 | -116 | 0 | 0 | 0 | 0 | 0 |
| 10 | -212 | -212 | -212 | -212 | -212 | 0 | 0 | 0 | 0 | 0 |
| 11 | -132 | -132 | -132 | -132 | -132 | 0 | 0 | 0 | 0 | 0 |
| 12 | -121 | -121 | -121 | -121 | -121 | 0 | 0 | 0 | 0 | 0 |
| 13 | -170 | -170 | -170 | -170 | -170 | 0 | 0 | 0 | 0 | 0 |
| 14 | -157 | -157 | -157 | -157 | -157 | 0 | 0 | 0 | 0 | 0 |
| 15 | -162 | -162 | -162 | -162 | -162 | 0 | 0 | 0 | 0 | 0 |
| 16 | -229 | -229 | -229 | -229 | -229 | 0 | 0 | 0 | 0 | 0 |
| 17 | -123 | -123 | -123 | -123 | -123 | 0 | 0 | 0 | 0 | 0 |
| 18 | -96 | -96 | -96 | -96 | -96 | 0 | 0 | 0 | 0 | 0 |
| 19 | -129 | -129 | -129 | -129 | -129 | 0 | 0 | 0 | 0 | 0 |
| 20 | -223 | -223 | -223 | -223 | -223 | 0 | 0 | 0 | 0 | 0 |
| 21 | -150 | -150 | -150 | -150 | -150 | 0 | 0 | 0 | 0 | 0 |
| 22 | -235 | -235 | -235 | -235 | -235 | 0 | 0 | 0 | 0 | 0 |
| 23 | -241 | -241 | -241 | -241 | -241 | 0 | 0 | 0 | 0 | 0 |
| 24 | -92 | -92 | -92 | -92 | -92 | 0 | 0 | 0 | 0 | 0 |
| 25 | -149 | -149 | -149 | -149 | -149 | 0 | 0 | 0 | 0 | 0 |
| 26 | -190 | -190 | -190 | -190 | -190 | 0 | 0 | 0 | 0 | 0 |
| 27 | -188 | -188 | -188 | -188 | -188 | 0 | 0 | 0 | 0 | 0 |
| 28 | -139 | -139 | -139 | -139 | -139 | 0 | 0 | 0 | 0 | 0 |
| 29 | -149 | -149 | -149 | -149 | -149 | 0 | 0 | 0 | 0 | 0 |
| 30 | -85 | -85 | -85 | -85 | -85 | 0 | 0 | 0 | 0 | 0 |
| 31 | -103 | -103 | -103 | -103 | -103 | 0 | 0 | 0 | 0 | 0 |
| 32 | -180 | -180 | -180 | -180 | -180 | 0 | 0 | 0 | 0 | 0 |
| 33 | -160 | -160 | -160 | -160 | -160 | 0 | 0 | 0 | 0 | 0 |
| 34 | -249 | -249 | -249 | -249 | -249 | 0 | 0 | 0 | 0 | 0 |
| 35 | -221 | -221 | -221 | -221 | -221 | 0 | 0 | 0 | 0 | 0 |
| 36 | -139 | -139 | -139 | -139 | -139 | 0 | 0 | 0 | 0 | 0 |
| 37 | -142 | -142 | -142 | -142 | -142 | 0 | 0 | 0 | 0 | 0 |
| 38 | -158 | -158 | -158 | -158 | -158 | 0 | 0 | 0 | 0 | 0 |
| 39 | -147 | -147 | -147 | -147 | -147 | 0 | 0 | 0 | 0 | 0 |
| 40 | -142 | -142 | -142 | -142 | -142 | 0 | 0 | 0 | 0 | 0 |
| 41 | -153 | -153 | -153 | -153 | -153 | 0 | 0 | 0 | 0 | 0 |
| 42 | -126 | -126 | -126 | -126 | -126 | 0 | 0 | 0 | 0 | 0 |
| Average | -154.119 | -154.119 | -154.119 | -154.119 | -154.119 | 0 | 0 | 0 | 0 | 0 |

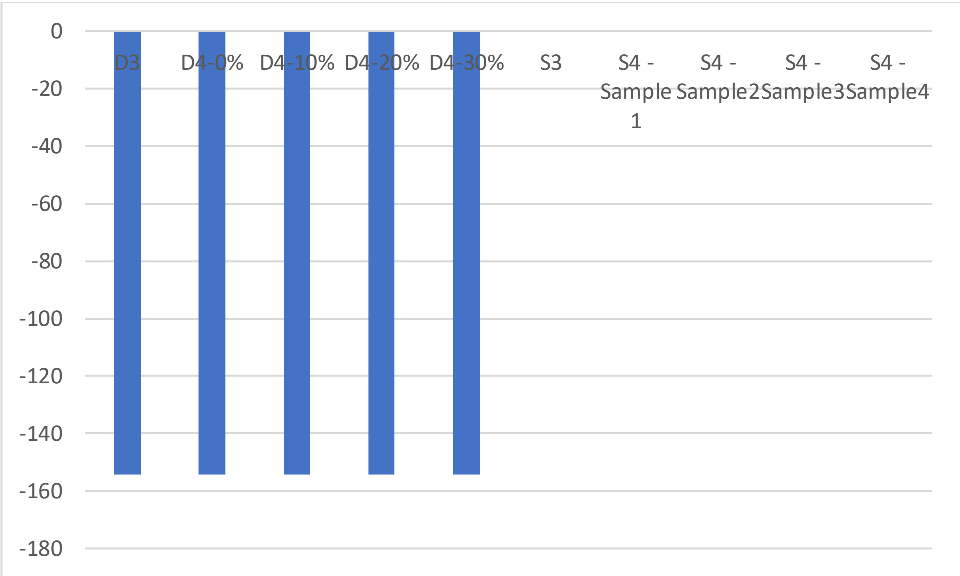


Figure 52: Comparison of Average Delay/Earliness in Soft Deadlines obtained in Deterministic and Stochastic models

# 10. Conclusion

Railways is the lifeline of any Country and contributes greatly in its economic growth. Rail Transport, functions round the clock which increases the importance of adopting a cost-effective and efficient Rail Maintenance Strategy.

In this thesis, deterministic and stochastic approaches are adopted to plan maintenance activities within a maintenance strategy.

Various scenarios are taken into consideration for better understanding. Scenario where 2 teams are considered for 10 jobs and scenarios for 4 teams are considered for 42 jobs. Each are again divided based on whether uncertainties are considered or not. Then, these are compared graphically for better understanding.

The results show that the delay obtained using stochastic approach is lesser or negligible as when compared to delay obtained using deterministic approach. Which indicates that the efficiency shall be higher for stochastic models than deterministic models.

# References


[1] A. Consilvio, A. Di Febbraro, R. Meo, and N. Sacco, “Risk-based optimal scheduling of maintenance activities in a railway network,” EURO JTransp. Logistics, vol. 8, no. 5, pp. 435–465, 2019.

[2] Asset management overview, principles and terminology, ISO 55000:2014,International Organization for Standardization, Standard, Geneva,Switzerland, 2014.

[3] A. Zoeteman and C. Esveld, “State of the art in railway maintenancemanagement: Planning systems and their application in Europe,” Proc.IEEE Int. Conf. Syst., Man Cybern., 2004, vol. 5, pp. 4165–4170.

[4] N. Jimenez–Redondo, N. Bosso, L. Zeni, A. Minardo, F. Schubert, F.Heinicke, and A. Simroth, “Automated and cost effective maintenancefor railway (ACEM–Rail),” Procedia-Social Behavioral Sci., vol. 48,pp. 1058–1067, 2012.

[5] S. S. Soh, N. H.M. Radzi, and H. Haron, “Review on scheduling techniquesof preventive maintenance activities of railway,” in Proc. 4th Int. Conf.Comput. Intell., Model. Simul., 2012, pp. 310–315.

[6] R. Macedo, R. Benmansour, A. Artiba, N. Mladenovi´c, and D. Uroševi´c,“Scheduling preventive railway maintenance activities with resourceconstraints,” Electron. Notes Discrete Math., vol. 58, pp. 215–222,2017.

[7] S. Khalouli, R. Benmansour, and S. Hanafi, “An Ant Colony algorithmbased on opportunities for scheduling the preventive railway maintenance,”in Proc. Int. Conf. Control, Decis. In. Technol., Apr. 2016, pp. 594–599.

[8] A. Higgins, “Scheduling of railway track maintenance activities andcrews,” J. Oper. Res. Soc., vol. 49, no. 10, pp. 1026–1033, Oct. 1998.

[9] G. Budai, D. Huisman, and R. Dekker, “Scheduling preventive rrailwamaintenance activities,” in Proc. IEEE Int. Conf. Syst., Man Cybern., Oct.2004, vol. 5, pp. 4171–4176.

[10] F. Peng, S. Kang, X. Li, Y. Ouyang, K. Somani, and D. Acharya, “Aheuristic approach to the railroad track maintenance scheduling problem,”Comput.-Aided Civil Infrastructure Eng., vol. 26, no. 2, pp. 129–145, 2011.

[11] C. Borraz-S ánchez and D. Klabjan, “Strategic Gang scheduling forrailroad maintenance,” CCITT, Center for the Commercialization of Innovative Transportation Technology Northwestern University, Evanston,IL, USA, Tech. Rep., 2012. [Online]. Available: https://rosap.ntl.bts.gov/view/dot/25019

[12] M. Dell’Orco, M. Ottomanelli, L. Caggiani, and D. Sassanelli, “NewDecision support system for optimization of rail track maintenance planning based on adaptive neurofuzzy inference system,” Transp. Res. Rec.,Vol. 2043, no. 1, pp. 49–54, 2008.

[13] P. Umiliacchi, D. Lane, F. Romano, and A. SpA, “Predictive maintenanceOf railway subsystems using an ontology based modelling approach,” inProc. 9th World Conf. Railway Res., 2011, pp. 22–26.

[14] R. Shingler, G. Fadin, and P. Umiliacchi, “From RCM to predictiveMaintenance: The InteGRail approach,” in Proc. 4th IET Int. Conf. RailwayCondition Monit., 2008, pp. 1–5.

[15] N. Jimenez-Redondo, S. Escriba, F. Benitez, F. Cores, and N. Caceres,“Towards automated and cost-efficient track maintenance. Final developments of the ACEM-Rail project,” in Proc. Transport Res. Arena, 2014,pp. 1–11.

[16] A. Consilvio, M. Iorani, V. Iovane, M. Sciutto, and G. Sciutto, “Real-timeMonitoring of the longitudinal strain of continuous welded rail for safetyImprovement,” in Proc. IMechE F, J. Rail Rapid Transit, 2019, pp. 1–10.

[17] S. Sharma, Y. Cui, Q. He, R. Mohammadi, and Z. Li, “Data-drivenOptimization of railway maintenance for track geometry,” Transp. Res.C, Emerg. Technol., vol. 90, pp. 34–58, 2018.

[18] A. Jamshidi et al., “A decision support approach for condition-basedMaintenance of rails based on big data analysis,” Transp. Res. C, Emerg.Technol., vol. 95, pp. 185–206, 2018.

[19] P. C. Lopes Gerum, A. Altay, and M. Baykal-G ürsoy, “Data-drivenPredictive maintenance scheduling policies for railways,” Transp. Res. C,Emerg. Technol., vol. 107, pp. 137–154, 2019.

[20] S. Simson, L. Ferreira, and M. Murray, “Rail track maintenance planning:An assessment model,” Transp. Res. Rec., J. Transp. Res. Board, vol. 1713,pp. 29–35, 2000.

[21] M. Burrow, S. Naito, and H. Evdorides, “Network-level railway trackMaintenance management model,” Transp. Res. Rec., J. Transp. Res.Board, vol. 2117, pp. 66–76, 2009.

[22] T. Zhang, J. Andrews, and R. Wang, “Optimal scheduling of track maintenance on a railway network,” Quality Rel. Eng. Int., vol. 29, no. 2,pp. 285–297, 2013.

[23] I. Soleimanmeigouni, A. Ahmadi, and U. Kumar, “Track geometry degradation and maintenance modelling: A review,” Proc. Institution Mech.Eng., F, J. Rail Rapid Transit, vol. 232, no. 1, pp. 73–102, 2018.

[24] M. C. Jeong, S.-J. Lee, K. Cha, G. Zi, and J. S. Kong, “Probabilistic model Forecasting for rail wear in Seoul metro based on Bayesian theory,” Eng. Failure Anal., vol. 96, pp. 202–210, 2019.

[25] J. Carretero et al., “Applying RCM in large scale systems: A case study With railway networks,” Rel. Eng. Syst. Safety, vol. 82, no. 3, pp. 257–273, 2003.

[26] U. Bharadwaj, “A risk based approach to maintenance optimisation of Business critical railway structures/equipment,” in Proc. Institution Eng. Technol. Semin. Appl. New Technologies Railways, 2007, pp. 53–60.

[27] E. Baglietto, A. Consilvio, A. Di Febbraro, F. Papa, and N. Sacco, "A Bayesian Network approach for the reliability analysis of complex railway Systems," in Proc. IEEE Int. Conf. Intell. Rail Transp., 2018, pp. 1–6.

[28] A. Consilvio, A. Di Febbraro, and N. Sacco, "A modular model to schedule Predictive railway maintenance operations," in Proc. Int. Conf. Models Technol. Intell. Transp. Syst., 2015, pp. 426–433.

[29] Y. F. Wang, Y. F. Zhang, J. Y. H. Fuh, Z. D. Zhou, P. Lou, and L. G. Xue, "An integrated approach to reactive scheduling subject to machine Breakdown," in Proc. IEEE Int. Conf. Autom. Logistics, 2008, pp. 542–547.

[30] M. Wang, K. Ramamohanarao, and J. Chen, "Dependency-based risk evaluation for robust workflow scheduling," in Proc. IEEE 26th Int. Parallel Distrib. Process. Symp. Workshops Ph.D. Forum, 2012, pp. 2328–2335.

[31] Y. Gao, Y. Ding, and H. Zhang, "Job-shop scheduling consideringRescheduling in uncertain dynamic environment," in Proc. Int. Conf.Manag. Sci. Eng., 2009, pp. 380–384.

[32] R. M. Pathan, "Recovery of fault-tolerant real-time scheduling algorithmFor tolerating multiple transient faults," in Proc. 10th Int. Conf. Comput.Inf. Technol., Dec. 2007, pp. 1–6.

[33] L. Ma, J. Kang, C. Zhao, and S. Liu, "Modeling the impact of prognosticErrors on CBM effectiveness using discrete-event simulation," in Proc. Int.Conf. Quality, Rel., Risk, Maintenance, Safety Eng., 2012, pp. 520–525.

[34] S. Narayanaswami and N. Rangaraj, "A MAS architecture for dynamic,Real-time rescheduling and learning applied to railway transportation,"Expert Syst. Appl., vol. 42, no. 5, pp. 2638–2656, 2015.

[35] L. Meng and X. Zhou, "Simultaneous train rerouting and rescheduling onAn N-track network: A model reformulation with network-based cumulative flow variables," Transp. Res. B, Methodol., vol. 67, pp. 208–234,2014.

[36] A. D'Ariano, L. Meng, G. Centulio, and F. Corman, "Integrated stochasticOptimization approaches for tactical scheduling of trains and railwayInfrastructure maintenance," Comput. Ind. Eng., vol. 127, pp. 1315–1335,2019.

[37] S. V. Aken, N. Bešinovi´c, and R. M. Goverde, "Designing alternativeRailway timetables under infrastructure maintenance possessions," Transp.Res. B, Methodol., vol. 98, pp. 224–238, 2017.

[38] M. H. Bin Osman, S. Kaewunruen, and S. Dindar, "Disruption: A newComponent in the track inspection schedule," in Proc. IEEE Int. Conf.Intell. Rail Transp., Aug. 2016, pp. 249–253.

[39] Z. Su, A. Jamshidi, A. Núñez, S. Baldi, and B. D. Schutter, "Multi-levelCondition-based maintenance planning for railway infrastructures – aScenario-based chance-constrained approach," Transp. Res. C, Emerg.Technol., vol. 84, pp. 92–123, 2017.

[40] A. Consilvio, A. Di Febbraro, and N. Sacco, “Stochastic schedulingApproach for predictive risk-based railway maintenance,” in Proc. IEEEInt. Conf. Intell. Rail Transp., 2016, pp. 197–203.

[41] S. M. Famurewa, T. Xin, M. Rantatalo, and U. Kumar, “OptimisationOf maintenance track possession time: A tamping case study,” Proc.Institution Mech. Eng., F, J. Rail Rapid Transit, vol. 229, no. 1, pp. 12–22, 2013.

[42] S. M. Famurewa, U. Juntti, A. Nissen, and U. Kumar, “Augmented utilisation of possession time: Analysis for track geometry maintenance,”Proc. Institution Mech. Eng., F, J. Rail Rapid Transit, vol. 230, no. 4pp. 1118–1130, 2015.

[43] Railway applications – track – track geometry quality – part 5: GeometricQuality levels – plain line, switches and crossings, EN 13848–5:2017,European Committee for Standardization, Standard, Brussels, Belgium,2017.

[44] Railway Applications. Track. Track Geometry Quality. Characterization ofTrack Geometry, EN 13848–1:2019, European Committee for Standardization, Standard, Brussels, Belgium, 2019.

[45] Railway Applications – Track – Track Geometry Quality – Part 6:Characterisation of Track Geometry Quality, EN 13848–6:2014, European Committee for Standardization, Standard, Brussels, Belgium,2014.

[46] Trafikverket (2014), “Track structure – track quality – requirement during Construction and maintenance,” Trafikverket, Standard, 2014.

[47] F. Della Croce, A. Grosso, and F. Salassa, “A matheuristic approach for The total completion time two-machines permutation flow shop problem,” In Evolutionary Computation in Combinatorial Optimization, P. Merz and J.-K. Hao, Eds. Berlin, Germany: Springer, 2011, pp. 38–47.