Minimal utilization rates for railway maintenance windows: a cost-benefit approach

VTI Working Paper 2021:8

Abderrahman Ait-Ali and Tomas Lidén

Transport Economics, VTI, Swedish National Road and Transport Research Institute

Abstract

Methods for economic assessment, e.g., cost-benefit analysis, are often used in the rail sector to evaluate large infrastructure investments such as building new high-speed railway lines. With larger railway networks and aging infrastructure, these methods can also be used for maintenance planning decisions. In this paper, we focus on basic maintenance and the newly introduced concept of maintenance windows in Sweden. These are pre-allocated slots in the annual train timetable dedicated to performing, among others, periodic/frequent maintenance activities such as inspections, maintenance and repairs. To justify the pre-allocation of such windows, this study presents a method to find minimal utilization rates depending on window designs and traffic situations. Using a cost-benefit approach, the maintenance windows are assessed using a total social cost including maintenance work costs, loss in traffic production and reliability gains in future traffic. Based on a case study from the Southern main line in Sweden, we study the minimal utilization rate in different test scenarios, i.e., night or day shifts, asset degradation functions and designs of maintenance windows. The results show that lower utilization rates (5-50%) can be accepted during low-volume traffic or for partial closures, while higher utilization rates (50-90%) are required for full closures during high-volume traffic. Whether the rates are measured as share of used window time or share of utilized windows is less important, especially when higher utilization is required.

Keywords

Maintenance windows; rail infrastructure; cost-benefit analysis

JEL Codes

R42

Minimal utilization rates for railway maintenance windows: a cost-benefit approach

Working Paper

Abderrahman Ait-Ali and Tomas Lidén
The Swedish National Road and Transport Research Institute (VTI), Box 55685, SE-102 15, Stockholm, Sweden

Abstract

Methods for economic assessment, e.g., cost-benefit analysis, are often used in the rail sector to evaluate large infrastructure investments such as building new high-speed railway lines. With larger railway networks and aging infrastructure, these methods can also be used for maintenance planning decisions. In this paper, we focus on basic maintenance and the newly introduced concept of maintenance windows in Sweden. These are pre-allocated slots in the annual train timetable dedicated to performing, among others, periodic/frequent maintenance activities such as inspections, maintenance and repairs. To justify the pre-allocation of such windows, this study presents a method to find minimal utilization rates depending on window designs and traffic situations. Using a cost-benefit approach, the maintenance windows are assessed using a total social cost including maintenance work costs, loss in traffic production and reliability gains in future traffic. Based on a case study from the Southern main line in Sweden, we study the minimal utilization rate in different test scenarios, i.e., night or day shifts, asset degradation functions and designs of maintenance windows. The results show that lower utilization rates (5-50%) can be accepted during low-volume traffic or for partial closures, while higher utilization rates (50-90%) are required for full closures during high-volume traffic. Whether the rates are measured as share of used window time or share of utilized windows is less important, especially when higher utilization is required.

Keywords: maintenance windows; rail infrastructure; cost-benefit analysis

1. Introduction

Economic assessment approaches in the rail sector, such as cost-benefit analysis (CBA), are often used to evaluate large infrastructure investments, e.g., building new high-speed railway lines. Such analysis is lacking when it comes to assessing and planning maintenance of the infrastructure. With larger railway networks, aging infrastructure requires more maintenance investments. In addition to the direct work costs such as for material and workers, maintenance activities affect traffic production leading to additional substantial social costs.

The infrastructure manager is responsible for maintaining and renewing the different assets or infrastructure components, and hence the overall rail infrastructure system (Zoeteman, 2001). In particular, the manager needs to decide when assets should be inspected, maintained or renewed. Such decisions are often a trade-off between costs (e.g., risk of disruptions, safety) and benefits (e.g., reliability).

Major expenses are at stake due to the large costs of rail investment and asset management of the rail infrastructure system. **Figure 1** shows that Swedish expenditures on railway infrastructure have been steadily increasing since 2000. The share of maintenance is also increasing partly due to the maintenance debt accumulating over several years. Moreover, Trafikverket, the Swedish rail infrastructure manager, has indicated in its maintenance plan for 2020-2023 that it will allocate a yearly budget of around 10 billion SEK, mainly split almost evenly between renewals and basic maintenance (Honauer and Ödeen, 2020).

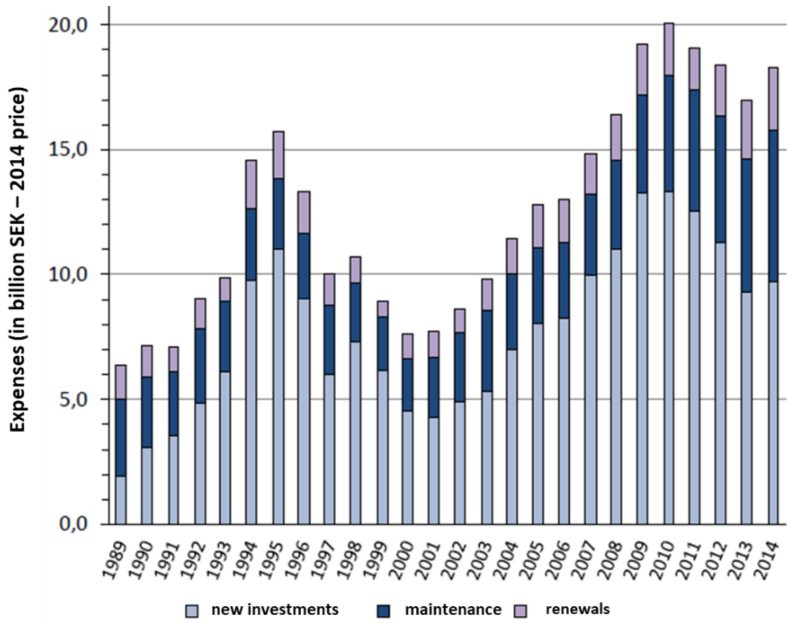

Figure 1. Expenditures on railway infrastructure in Sweden between 1989 and 2014 (Alexandersson, 2015).

Given the expenses at stake and the limits in public funds, the infrastructure managers are faced with several prioritizations to make. In Europe, EC (2018) states that social value should be used as an important prioritization criterion. Thus, the importance of using cost-benefit analysis (CBA) as a tool to study the social value of the investments. Moreover, such planning decisions are also important for the punctuality of the whole rail system. However, knowledge about the planning of maintenance is still weak (Kristoffersson, 2019).

As part of basic maintenance planning, this paper studies maintenance windows (MWs) which were recently introduced in Sweden (Göransdotter and Dyrssen, 2017). These MWs are pre-allocated slots or reserved capacity in the annual train timetable. Such capacity guarantees access to the track for the contractors in order to perform regular maintenance and inspection activities. Using a cost-benefit approach, the study presents a method to find minimal utilization rates that can economically justify the reservation of such MWs in the annual timetable.

Since their implementation in 2016 by Trafikverket, several studies have found that MWs were sometimes not efficiently utilized by maintenance contractors (Göransdotter and Dyrssen, 2017). Other studies have attempted to improve the planning of such windows, e.g., in parallel with traffic planning (Lidén, 2018). Although important, we do not aim to find the optimal planning of MWs. Instead, the aim is, given a fixed schedule of the MWs, to find the minimal utilization rate of the schedule so that the social value is positive. Such minimal utilization would justify the negative effect of the windows (for the train traffic that could otherwise have taken place). Establishing such minimal utilization rates also forms the basis for setting contractual requirements and when evaluating the performance of maintenance contractors.

Hence, the study excludes long-term maintenance activities such as infrastructure investments and large renewal works, also called PSB (*planerade större banarbeten* in Swedish). It also does not include short-term activities such as corrective maintenance after accidents or sudden failures. Moreover, the maintenance of the rolling stock or train vehicles is also excluded, although important for fixed stock and hence the rail infrastructure system. Another important limitation aspect is the exclusion of funding and tax considerations, e.g., track access charges.

Using a case study from Sweden, both costs and benefits are studied. In particular, costs include short- and long-term components, e.g., work and material costs, reduced available capacity and disturbances whereas benefits include the increased reliability of the future traffic production. The model is applied to several test scenarios with different characteristics such as time period, asset degradation rate and design of MWs. Based on the different tested scenarios, the numerical results indicate which utilization rates should be required for different designs of MWs and traffic situations. These results show that MWs during low-volume traffic may only require 5-10% utilization rates. During high-volume traffic, the minimal utilization rates increase to 25-50% for partial closure of the tracks (single-track) and to 50-90% for full closure. In the latter case, similar rates are obtained whether the rates are measured as share of used window time or share of utilized windows.

The paper starts with this introductory section. Section 2 reviews the existing related literature. The model is described in section 3. The case study is presented in section 4 including test scenarios and results. Section 5 ends the paper with conclusions.

2. Literature overview

In this section, we introduce some background information and terminology concerning infrastructure maintenance in railways. Many presented concepts are specific to the Swedish railways, but references are sometimes given to several other international resources.

2.1. Infrastructure assets

Train operations in railway systems require the existence of a solid infrastructure. The latter consists of various assets that need continuous inspections, maintenance, corrections/repairs and renewal. Several important assets are illustrated in **Figure 2** by Lander and Petersson (2012). Other vital assets include switches, signaling, overhead lines/catenary wires, power transformers, and civil engineering structures such as bridges, viaducts and tunnels.

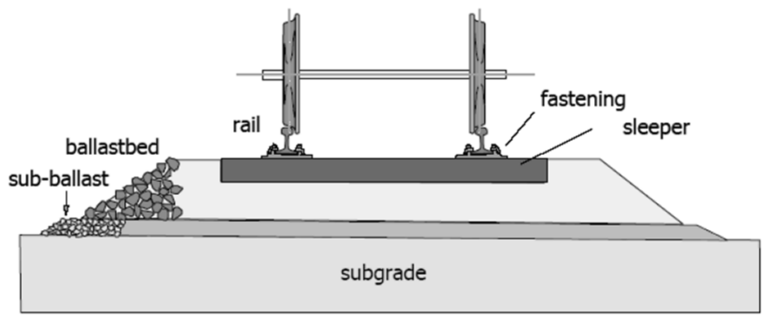

Figure 2. Illustration of some infrastructure assets in railway systems (Lander and Petersson, 2012).

Using a system perspective, railways can be seen as a system of infrastructure assets. The well-functioning of the asset system is conditioned by the efficiency of the maintenance activities of such assets/components. A study by WSP (2014) made an inventory of different assets and corresponding maintenance activities in Sweden, see **Appendix 1: Examples of infrastructure assets and maintenance activities.** Other assets and maintenance activities can also be included in this inventory such as civil structures (tunnels and bridges/viaducts), terminal stations and marshaling yards.

2.2. Assets management

Assets are generally maintained by different activities such as cleaning, lubrication, straightening, calibration, reparation, renovation and replacement. Some of these activities are more frequent and/or cheaper than others. Instead of schedule-based maintenance (e.g., sleepers, ballast), some assets are treated based on their status or conditions such as rail and overhead wires.

There are generally two categories of maintenance activities that are performed on assets in railway systems. One is the renewal (*reinvestering* in Swedish) after the maximum lifetime is reached whereas the second category is the basic maintenance (*basunderhåll* in Swedish) which is often frequently performed before renewal.

Basic maintenance can take one of the following two forms:

- Corrective maintenance (*avhjälpande underhåll* in Swedish) is happening after the occurrence of one or more failures in the asset.
- Preventive maintenance (*förebyggande underhåll* in Swedish) is aimed at preventing the potential occurrence of failures. It is also known in the literature as schedule or plan-based maintenance (Rausand and Vatn, 2008).

A taxonomy of the different types of maintenance activities is illustrated in the timeline in **Figure 3**. It also indicates when these are performed.

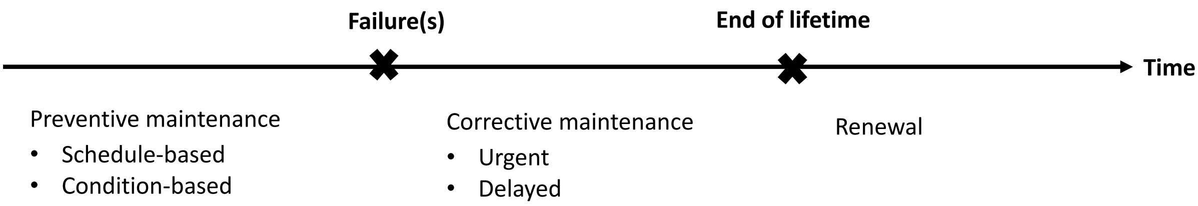

Figure 3. Maintenance timeline and the corresponding types of activities.

To choose the correct time and place to perform activities on the infrastructure, some assets are periodically inspected. These inspections provide information about the status or condition of the asset so that condition/status-based preventive maintenance can be planned and potentially take place.

Inspections are traditionally done manually but recently more automatic and digital measurement tools are increasingly adopted. For instance, some trains include cars that are equipped with tools to inspect track and rail structure by measuring the track gauge and rail profile. These new technologies have the advantage of increasing safety and reducing costs for maintenance.

The inspection frequency does not only depend on the asset but also other parameters such as traffic intensity and train speeds. For instance, Trafikverket classifies its infrastructure assets for inspection in different categories, called inspection classes (*Besiktningsklass* in Swedish, noted Bx). **Figure 4** shows these classes for various speeds and yearly traffic intensity. Other considerations such as weather are also considered.

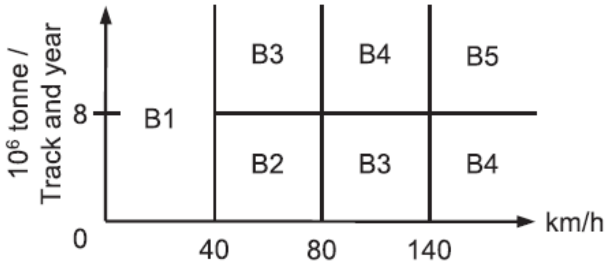

Figure 4. Infrastructure classes for inspection in Sweden (Stenström et al., 2016).

Preventive maintenance can alternatively take place according to a fixed schedule, i.e., schedule/periodic-based maintenance. Such period is the result of knowledge information about the assets, e.g., from past experience or instructions by the manufacturer. Other factors such as the traffic volume and/or weather conditions may also have substantial effects on the choice of the schedule. Some examples of schedule-based activities include maintenance of assets in marshaling yards, electric power, signaling and telecommunication systems. In some cases and based on inspection results, such activities can also be performed ahead of schedule and are thus part of status-based maintenance (Göransdotter and Dyrssen, 2017).

In case of an accident or infrastructure failure(s), corrective maintenance is performed either immediately or scheduled for later, usually combined with some operative restrictions, e.g., lowered speeds, until the repair(s) have been completed. Such activities are performed on assets that are inspected with poor-quality condition. Other examples include urgent interventions due to accidents, crime or failures as well as winter maintenance, e.g., snow removal (Göransdotter and Dyrssen, 2017). In Sweden, failures are marked in different categories. For instance, A-marking means acute (*Akut* in Swedish) condition requiring immediate repairs whereas V/M/År-markings require scheduling corrective maintenance within a week/month/year (Alexandersson, 2015).

Besides maintaining assets, renewal activities are also performed when judged to be better than maintenance or when the lifetime of the asset is reached. For instance in Sweden, the recommended lifetime for major railway assets, e.g., rails, tunnels, bridges, is around 60 years but further inspections are generally performed (Trafikverket, 2020a). Renewals are generally more expensive than maintenance but can contribute to enhanced service quality and asset lifetime and thus reduce the need for inspections and preventive and/or corrective maintenance.

Although not discussed in this project, the choice of whether to renew or otherwise maintain the infrastructure is important and is often based on the balance between several elements, e.g., safety, costs/economy, and type of the infrastructure/asset(s). Interested readers are referred to a more detailed recent analysis by Nilsson and Odolinski (2020).

2.3. Planning for maintenance

National infrastructure managers, such as Trafikverket in Sweden, are generally responsible for the different assets and their maintenance. However, differences exist between countries depending on the level of vertical separation as well as competition (horizontal separation) in the market for maintenance. Moreover, many countries in Europe with similar market organizations have different external/internal actors for planning and performing maintenance activities. **Table 1** provides a comparison between various possible organizations in some European railways.

Table 1. Comparison between the maintenance organization in some European railways (Alexandersson, 2015).

	Procurement (outsourced)	Internal (in-house)	Both
Inspection	**SE**, FI, NL	DK, BE, FR, CH, NO	GB
Maintenance	**SE**, FI, NL, NO	DK, BE, FR, GB	CH
Renewal	**SE**, FI, NL, GB	-	CH, NO, DK, BE, FR
Investments	**SE**, FI, NL, GB, FR	-	CH, NO, DK, BE

Unlike many European railways where maintenance and/or inspection is planned and performed in-house by internal units of the infrastructure manager, Trafikverket outsources them to procured external stakeholders, also called contractors or entrepreneurs. Some countries choose to combine internal works with external procurements, e.g., Switzerland/CH (for maintenance) and Britain/GB (for inspections). In addition to the infrastructure manager, often the main owner of the infrastructure assets, there are several stakeholders which are involved in the different levels of planning for maintenance such as government regulator(s), representatives for passengers and freight customers, train operators (Kobbacy and Murthy, 2008).

Different strategies for planning maintenance exist, but the literature distinguishes between two main variants, i.e., condition- and reliability-based maintenance (Ling, 2005). The former involves the continuous inspection of the assets and maintenance is performed based on the measured condition of the assets. The latter is, however, focused on evaluating the failures risks, frequencies and lifetime characteristics, maintenance activities are then planned accordingly. See **Figure 5** for an example of tamping and ballast inspection data by Gaudry et al. (2016).

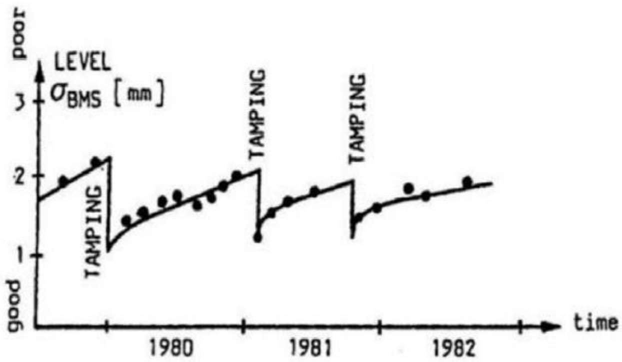

Figure 5. Degradation curve of the ballast based on several inspection data points (Gaudry et al., 2016).

Large investments, e.g., renewals and substantial reparations, are often planned strategically. However, minor activities, e.g., inspections and small reparations, are considered in the short-term planning process, also called tactical. Urgent activities, e.g., corrective maintenance, are planned during operations. With a focus on asset management and maintenance, **Figure 6** illustrates the different planning levels for railway maintenance. The figure is inspired by Kobbacy and Murthy (2008) and Lidén (2014).

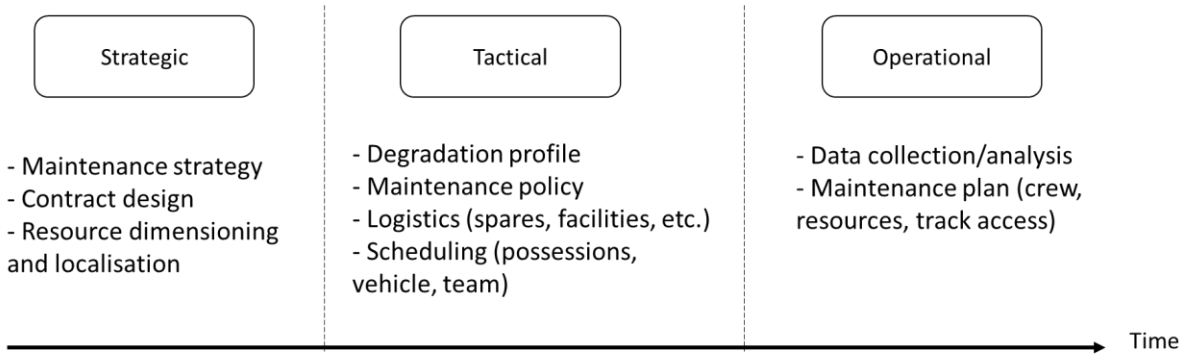

Figure 6. Maintenance planning levels and problems, inspired by Kobbacy and Murthy (2008) and Lidén (2014).

Different planning levels are concerned with various maintenance planning problems. Strategic questions are treated a long time beforehand, often several years in advance, to set overall maintenance strategic goals, performance-based/traditional contracts and resource locations. More specific plans are made during the tactical planning level, e.g., traffic and maintenance schedules based on maintenance policies and other information such as asset degradation profiles. Data is collected/analyzed during the operational level where detailed maintenance plans are performed including possessions guaranteeing access to the track for maintenance and the corresponding crew/resource allocation.

In Sweden, strategic planning, of large maintenance projects (*Planerade Större Banarbeten – PSB*) and traffic-affecting actions (*Trafik Påverkande Åtgärder - TPÅ*), is done first as a result of strategic directions by the government together with Trafikverket. The latter publishes a periodic maintenance plan (*underhållplan*), see (Honauer and Ödeen, 2018) and (Honauer and Ödeen, 2020). The plan describes prioritized major investment and maintenance plans for both road and rail infrastructure. The plan covers 4 years and is continuously updated every 2 years.

At the tactical level, the allocation of railway capacity for track access is generally part of an annual process described in the national network statement (Trafikverket, 2020d). The latter, also called *Järnvägsnätsbeskrivningen* or JNB, lists at the beginning of the annual process the main maintenance activities that are planned to be included in the annual timetable. Based on the draft timetable and additional details of the PSBs, a maintenance plan (*Banarbetsplan - BAP*) is specified. Further applications for maintenance slots and/or train paths can be done as part of the short-term ad hoc application process (Trafikverket, 2020d). Based on a similar figure by Hedström (2020), **Figure 7** gives a schematic summary of the different scenarios to plan and allocate capacity for maintenance activities in Sweden. The figure illustrates when the different maintenance activities are planned. Activities responding to immediate/acute needs for maintenance are planned at the operational level, i.e., real-time traffic management. For earlier planning of maintenance, the figure distinguishes between activities that can/cannot take place as part of MWs (Nilsson et al., 2015).

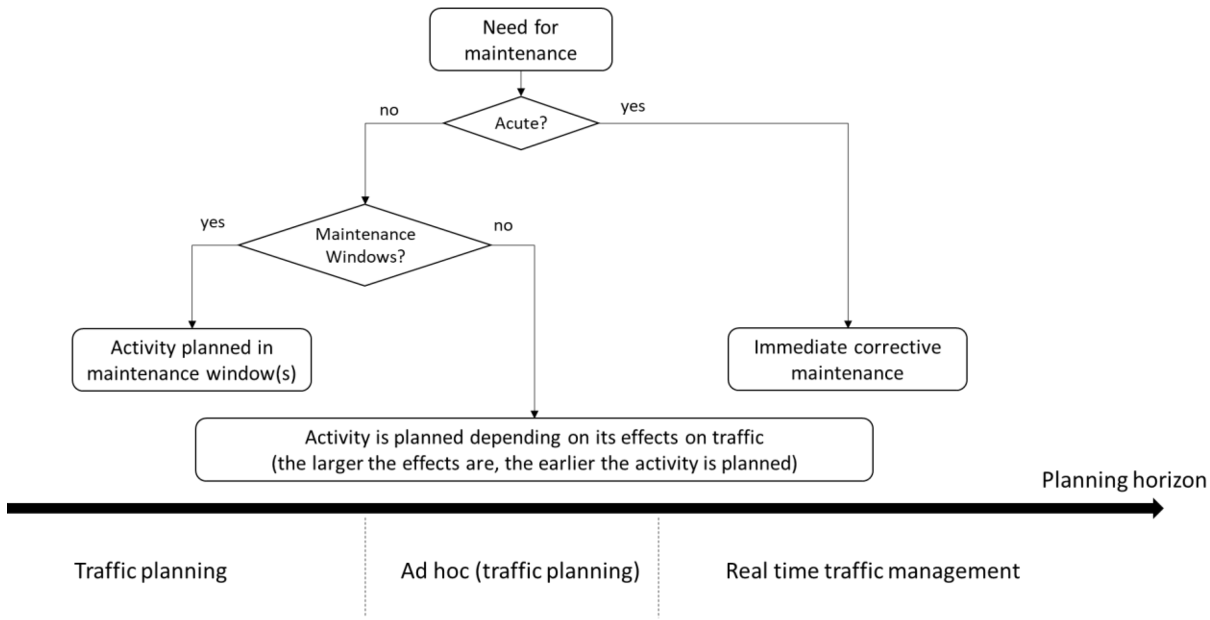

Figure 7. Different scenarios to allocate capacity for maintenance activities in Sweden, inspired by Hedström (2020).

Many interesting problems relate to planning for maintenance activities. For instance, the need for maintenance can either be based on inspection, i.e., condition/status-based maintenance, or repeated periodically, i.e., schedule-based maintenance. For different assets, each approach has costs and benefits hence the problem of optimizing maintenance by finding the best approach for the different assets. For more details on this problem, interested readers are referred to the book chapter on maintenance optimization by Mazzuchi et al. (2007). Other relevant problems that relate to maintenance planning include (but are not limited to) concurrent planning of maintenance and train traffic (Lidén, 2018), safety during maintenance works (Fokkert et al., 2007), eMaintenance (Al-Douri et al., 2016) and predictive maintenance (Ran et al., 2019).

2.4. Maintenance windows (MWs)

Inspired by other countries such as France and the Netherlands, Sweden has adopted, since 2016, a new form of capacity slots during the allocation process. Such new forms are called MWs (*servicefönster* in Swedish) and are pre-allocated in the annual timetable for recurrent maintenance activities (Lidén, 2018). Scheduling slots for maintenance activities within such windows have therefore no effects on train traffic. Another advantage is that they guarantee access to the track for the contractors and thus allow them to estimate their costs more accurately (Honauer and Ödeen, 2018).

”Maintenance access windows” is an alternative and more precise term that is also found in the literature (Kalinowski et al., 2020). The term indicates that the windows may not always be used but only provide capacity to access the tracks for maintenance if needed. For brevity, we adopt instead the term ”maintenance windows” or MWs for short throughout this paper.

Since its introduction in Sweden, the new concept of maintenance windows has not been efficiently used in maintenance planning as expected by Trafikverket (2020b). The latter is continuously working on improving the utilization rates of these windows. Olauson (2020) indicates that these inefficiencies are because MWs are often not well-suited to perform some types of maintenance activities.

Several inspection and maintenance activities are planned within MWs. In practice, the contractors apply for possessions to access the tracks within the pre-defined MWs. Such capacity applications should be sent at least 20 days beforehand. Otherwise, the remaining capacity is released for the short-term ad hoc application, i.e., typically between 5 to 12 days beforehand (Alexandersson, 2015).

MWs can have different designs. It is important to mention that there is generally a trade-off in the size of the slots in MWs. Although preferred by entrepreneurs, longer slots have more effects on traffic production and risk having lower overall utilization rates. However, shorter slots have a lower effect on traffic but are often not long enough for entrepreneurs to perform certain activities and thus may also tend to not be fully utilized (Olauson, 2020).

Moreover, activities that are planned within MWs also have different configurations. If the required time to perform the maintenance activity is longer than the possessions that are possible within MWs, the activity may be divided and performed over multiple possessions. However, this will require additional overhead costs, e.g., setup and transport expenses. **Figure 8**, by Lidén and Joborn (2016), illustrates the distinction between three different configurations.

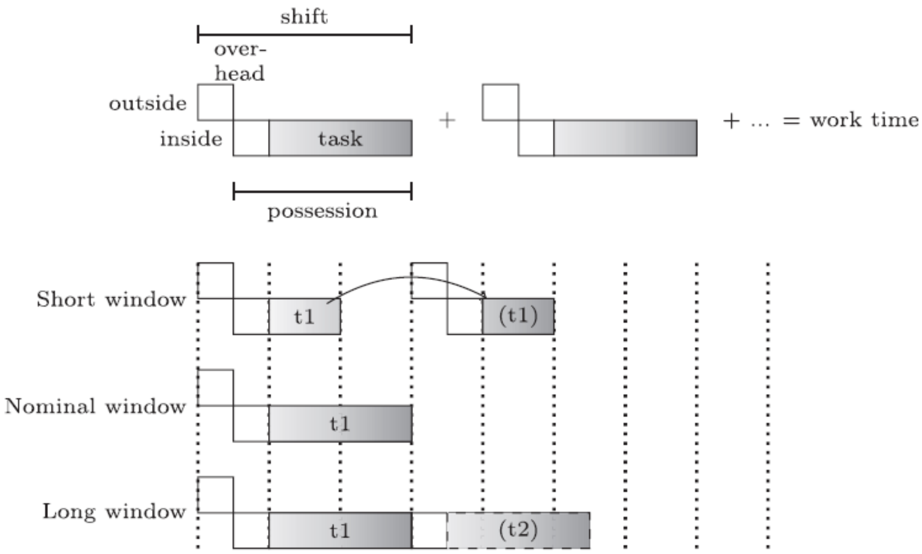

Figure 8. Illustration of different designs for MWs and planning configurations (Lidén and Joborn, 2016).

Finding the most efficient configuration depends on the following elements (Lidén et al., 2020):

1. Efficient utilization level of the service windows which depends on the accessibility of maintenance resources, e.g., cost of transport and logistics.
2. Required frequency for track access which depends on the nature of the maintenance activity.
3. Tolerance for negative effects on traffic which depends on the type of traffic.

2.5. Assessing maintenance

The basic railway maintenance often accounts for a large share of the total expenses in railways. For instance, more than 50 % of the total maintenance expenses in Sweden between 2017 and 2021 are spent on basic maintenance, including preventive and corrective maintenance activities (Honauer and Ödeen, 2018).

Cost-benefit analysis (CBA)

Several studies suggest that the maintenance costs in railways are driven by various factors. A study by UIC (2015) short-listed aspects such as asset density, electrification, tonnage, speed, maintenance strategy and service quality. Traffic density, i.e., higher tonnage and/or capacity usage, is one aspect that attracted substantial attention in the literature (Andersson, 2006). An early literature review by Hedström (1996) has also shown that the effects of traffic volumes are significant. Subject to different traffic volumes, similar assets, i.e., having the same lifetime characteristics, may therefore need to be maintained with different frequencies.

Many studies attempted to quantify the marginal costs for railway maintenance. Odolinski and Boysen (2019) have attempted to estimate the marginal costs from capacity utilization which is useful for planning maintenance activities for parts of the infrastructure with different traffic volumes. Such estimates are also useful for track access charging (Odolinski and Wheat, 2018) and/or planning the renewal of assets (Nilsson and Odolinski, 2020). In the newly deregulated market in Europe, train operators pay track access charges to the infrastructure manager as compensation for the effects of the train traffic on the infrastructure. Charging the track access is hence based on marginal cost pricing. In addition to the cost-recovery, such charges are also used for the efficient use of capacity (Ait Ali, 2020).

Moreover, additional costs can be incurred when scheduling maintenance activities on infrastructure sections with high traffic as a result of the loss in traffic production. MWs are therefore in competition with potential train services. A thesis by Lidén (2018) investigated how such windows can be designed in, among others, cost-efficient ways.

In addition to traffic disruptions due to maintenance activities (FR8RAIL-II, 2020), failures in the infrastructure, due to poor maintenance, can have costly effects. For instance, failures in certain assets such as rails, switches, overhead lines and signaling systems are responsible for a large percentage of the delays in the system. In Sweden, Kristoffersson (2019) finds that around 16% of the total hours of delays in 2017 are due to infrastructure-related problems, of which 21 % are due to problems with switches.

Most activities in MWs have pre-planned/expected and continuous effects on train traffic. Infrastructure-related delays are due to repair activities that have unexpected effects on traffic. In this study, the former is included as a social cost where the latter is used as a proxy/benefit for increased punctuality or reduced risk for delays when performing maintenance activities.

The literature includes few other CBA studies of maintenance. For instance, Stenström et al. (2016) consider that maintenance activities have both direct and indirect costs. In their CBA model, the authors include the following:

- Material and labor costs, i.e., direct costs
- Maintenance times which include logistic time and active repair time
- Production or service losses, e.g., delays
- Inspection costs for preventive maintenance, e.g., using tack geometry cars
- Failure costs for corrective maintenance

Moreover, the same authors state that for planning maintenance activities, the following characteristics should also be considered (Stenström et al., 2016): cost of downtime, redundancy of the infrastructure, i.e., network connectivity, and reliability characteristics.

Andersson et al. (2011) studied the relation between the costs of maintenance activities and their socio-economic effects including benefits. Train operations together with maintenance or renewal activities define several characteristics in the system, e.g., speed, reliability, level of comfort, tonnage/load, safety and emissions. These in turn lead to different effects that include, among others, travel/transport time or cost, delay, crowding and accidents. **Figure 9** gives an illustration of this framework.

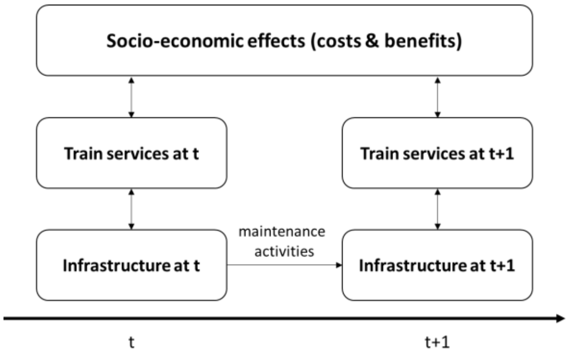

Figure 9. Illustration of the relationship between maintenance activities and socio-economic effects.

Based on this framework, Eliasson and Börjesson (2014) highlight the importance of the assumptions that are used to construct train timetables for train services in the resulting total socio-economic effects. Using a similar methodology, Lidén and Joborn (2016) compared the costs of infrastructure maintenance activities with the socio-economic effects of train traffic by optimizing the train timetable using mathematical programming. Another CBA study focusing on planning MWs in Sweden used simulation-based optimization instead (Sweco and WSP, 2015). Stenström et al. (2016) have developed analytical formulations of the total costs for planning preventive and corrective maintenance. Together with a CBA, the authors formulated the benefits and costs of performing preventive and maintenance before describing a case study based on historical maintenance data. The authors presented the resulting benefit/cost ratios indicating the most economically efficient maintenance plans.

According to the Swedish railway legislation, the infrastructure manager or Trafikverket must prioritize maintenance activities according to their socio-economic effects. These are defined by strategic directions of the government, e.g., safety, punctuality, environment (Ekström, 2015). Thus, important sections of the infrastructures connecting large cities and railway hubs are often given priority. To be able to quantify the different effects, Trafikverket publishes regular updates to its guideline for cost-benefit analysis, also known as ASEK, to be used in economic assessment projects of the Swedish infrastructure (Trafikverket, 2020a). To rank the different possible projects or policies in comparison to the do-nothing scenario (also called comparison alternative), the net present value ratio (*nettonuvärdeskvoter - NNK)* is often used to present the CBA results based on profitability (Bångman, 2012).

Life cycle cost (LCC)

To assess infrastructure maintenance and lower the operations costs, it is also important to know the costs throughout the life cycle of the assets, also called life cycle cost/costing or LCC for short (Zoeteman, 2001). Quantitative LCC studies aim at assessing the total cost of acquiring, owning and disposing of assets. This can serve in decision-making tools, e.g., to improve maintenance strategies for these different infrastructure assets such as finding the optimal trade-off between investment and maintenance. In Sweden for instance, there have been several attempts to use LCC-analysis for different railway assets, e.g., rail tracks (Patra, 2007) and switches & crossings (Nissen, 2009).

LCC analysis can be used in combination with other approaches such as CBA to find which asset and when to replace it. It can also be used at an earlier stage to choose between different types and/or combinations of assets (Nissen, 2009). LCC has even been recently used in a KPI model to study the potential of different innovations of Shift2Rail projects, e.g., in railway infrastructure and rolling stock (Perreal et al., 2019).

There are 6 phases in an LCC analysis, i.e., concept and definition, design and development, manufacturing, installation, operation and maintenance, disposal. The simpler alternative analysis includes only 3 main phases, i.e., development, operation, phase-out. The latter is the approach that was commonly used in Sweden (Nissen, 2009). We focus in this study on the performances of the assets during operations, i.e., asset degradation. In this context, **Figure 10** illustrates two typical examples of degradation curves/functions that are studied in this paper. Similar examples exist in survival analysis and are often called survival functions (Rayhusthwaite, 2009).

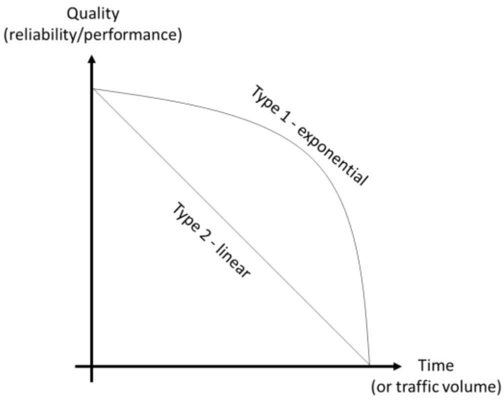

Figure 10. Examples of two studied degradation functions.

In LCC analysis, the total costs of the phases (and components) are evaluated at the time of the analysis. It is therefore important to use the so-called total present value (TPV) where all the costs and benefits, including the ones in the future, are all converted/discounted to an equivalent present value using a discounting factor. For more details on the importance of such discounting in maintenance planning, see (van der Weide et al., 2010). Other metrics are used to present the costs in LCC studies such as internal rate of return and/or annuity. The resulting values are often validated using sensitivity or uncertainty analysis (Zoeteman, 2001).

For rail infrastructure assets, there is a trade-off between the maintenance period and the risk for failure or disruptions. It is ideally preferable to schedule maintenance just before the failure of the asset. Thus, the loss from maintaining an already functional asset is reduced/absent. However, e.g., due to varying conditions (increased traffic volume and/or harsher weather), some assets can fail before the usual period and such uncertainties risk provoking costly disruptions, e.g., delays and discomfort. **Figure 11** illustrates, among others, that the main advantage of a perfectly scheduled maintenance, guaranteeing a minimal accepted quality, is the reduction in the loss (in red color) due to early maintenance activities.

Decision support systems (DSS) use LCC to analyze the long-term impacts of design and maintenance activities on the total cost but also on other aspects such as reliability, traffic performance or punctuality, availability and safety (Zoeteman, 2001). However, the degradation function is often not known or uncertain even if traffic conditions, e.g., weather, trainloads, are not highly variable.

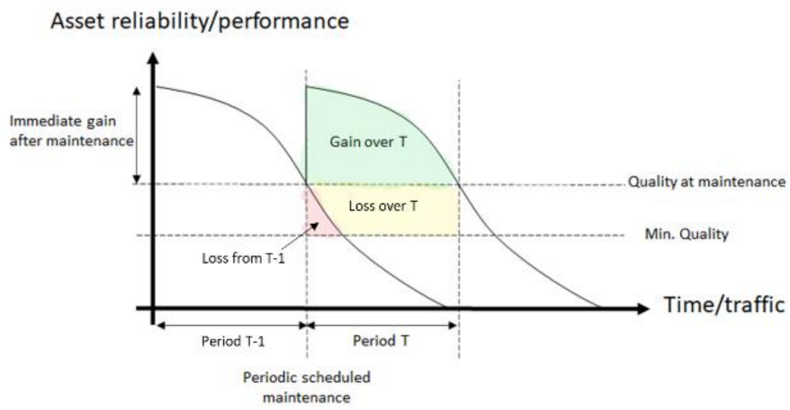

Figure 11. Hypothetical degradation curve and periodic maintenance of an asset.

Reliability (RAMS)

Maintenance assessment is not restricted to analyzing the costs and benefits over the life cycle of assets. It is therefore important to also consider other aspects, one more general approach is the so-called RAMS analysis which stands for Reliability, Availability, Maintainability and Safety. These four are all used as indicators of the quality and performance of the infrastructure assets (Patra, 2007). RAMS aims to predict the specific functionalities of a product over its complete life cycle (Ghodrati et al., 2017).

In RAMS analysis, each indicator is defined and specifically characterized with different RAMS parameters. For instance, reliability can be defined as the probability that an item can perform a required function under given conditions for a given time interval. It is characterized by the failure rate or the potential number of failures over a given period. As for availability, it is defined and characterized by the ability of an asset to be functional over a certain time (Ghodrati et al., 2017).

RAMS has been used, in combination with LCC, as the basis for infrastructure-related decision-making in several European research projects, e.g., InnoTrack (Nissen, 2009). Using Monte Carlo simulations, Patra (2007) identified RAMS parameters that have the highest effect on the LCC costs for railway track maintenance in Sweden. Such parameters are either affected by the system, operating and/or maintenance conditions. The resulting relations can be used to optimize planning condition-based maintenance, see the discounted cost model by van der Weide et al. (2010) using reliability evaluation.

3. Model

Once we have reviewed the existing literature as well as the main components of a possible CBA model, i.e., effects of railway maintenance, this section describes the developed CBA model. It combines different assessment methods which are presented in the literature overview.

3.1. Utilization rates for MWs

Designs of MWs

As described in the literature, MWs are pre-allocated slots in the annual timetable that guarantee access to the tracks for entrepreneurs to perform various periodic or frequent maintenance activities.

The slots are generally reserved for few hours (typically 2-6 hours) per day over several weeks. Let $W = (n_{weeks}, n_{days}, n_{hours})$ denote the main characteristics of the maintenance window that is performed on a section of the rail infrastructure, i.e., the number of weeks, days and hours. The total access time (in hours) is then $T(W) = n_{weeks} \times n_{days} \times n_{hours}$.

On the Southern main line, Trafikverket (2019) in its annual network statement has for instance reserved 6 hours in six days over several weeks as capacity windows for performing maintenance activities. Certain windows allow for single-track traffic often in day-times whereas others correspond to total traffic shut-down typically during night-times.

Note that different MWs can be utilized at the same time in two or more sites on the same line. In this case, we assume that the total access time $T(W)$ corresponds to the longest window.

To study the utilization rates for such needs, we separately look at two different cases, namely weekdays ($n_{days} = 5$ with work performed during night-time) and weekends ($n_{days} = 2$ with work performed during day-time). In particular, we will look at the following two alternative designs for MWs over a period of $n_{weeks} = 4$ or a month:

- **MW-day**: $n_{hours} = 2$ during day-time and weekdays, in total $2 \times 5 = 10$ h/week
- **MW-night**: $n_{hours} = 5$ during night-time and weekends, in total $5 \times 2 = 10$ h/week

We assume that MWs, noted W, are already designed and fixed, and we study the subsequent total costs and benefits for different utilization rates.

Dynamics of maintenance activities

MWs and the activities that are performed within these windows have various costs and benefits. **Figure 12** shows a graph of the different possible activities in the life cycle of a rail infrastructure asset (studied activities are in red color). In this study, we consider several costs and benefits that relate to the different maintenance activities that are planned as part of MWs. The benefits include information about the asset status and the gain in future train production which mainly relates to train service improvement thanks to increased reliability (Ling, 2005). The costs consist of the work-related expenses and the loss in the potential current production which is also known as the cost of downtime or opportunity cost of train traffic (Kobbacy and Murthy, 2008). Some other costs/benefits are not considered in this study, e.g., taxes, track access charges and external effects such as noise reduction and environmental effects.

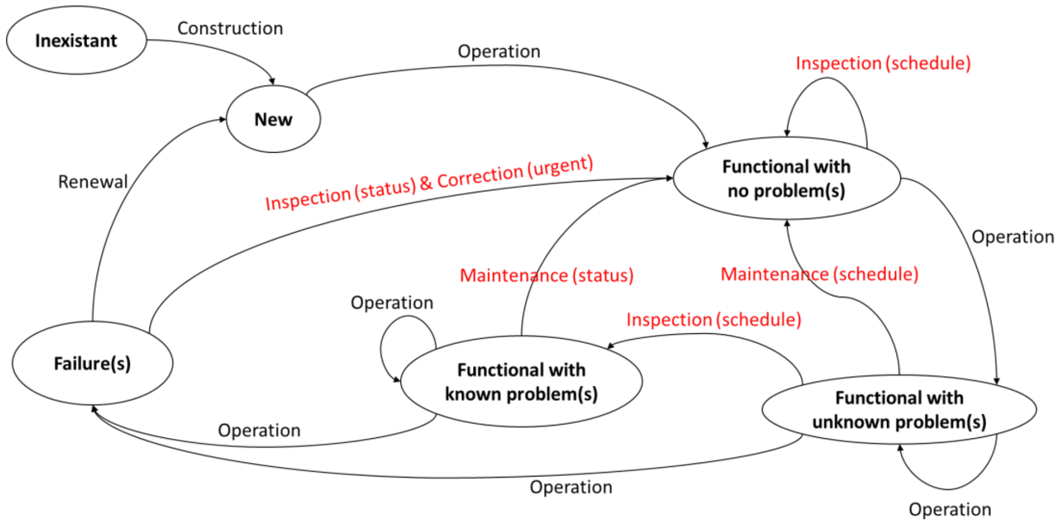

Figure 12. Evolution of the status of an asset illustrated as a directed graph.

Given W, let $u = \frac{T_{eff}}{T(W)}$ denote the utilization rate of these windows where T_{eff} is the effective average time spent on track and $T(W)$ is the total access time of the pre-allocated MWs W. There are several alternative ways to characterize the rate of utilization u. For instance, Granberg and Rehn (2020) state that some possible characterizations include:

- Share of utilized slots, which can be useful to assess the costs in traffic production
- Share of utilized access time, which can indicate if MWs are long
- Share of canceled or performed activities, which can assess short-term planning

Note that the definition $u = \frac{T_{eff}}{T(W)}$ corresponds to the second characterization. The same authors also conclude that such characterizations are only useful to study and follow-up for few important parts of the infrastructure, referred to as "hot" MWs. Moreover, large complex parts of the infrastructure should use adapted characterization to study the utilization rates of MWs in these regions, referred to as "islands".

In this study, we mainly adopt and focus on the two first definitions. Both are compared in the case study but the second one is analytically advantageous in that it leads to a differentiable total cost function, hence marginal costs can be expressed analytically.

Given W, the total social costs TC can be formulated as the sum of the work cost CC and the loss in (current traffic) production LP minus the gain in future traffic production GP, i.e., $TC(u) = CC(u) + LP - GP(u)$.

MWs are allocated during capacity planning. Thus, regardless of the utilization rate u of such windows, some costs (e.g., loss in production LP) persist and are independent of u, while the work costs and future benefits depend on the utilization rate.

Note that unused MWs may allow for capacity that can be used, instead of maintenance activities, for increasing the robustness of the timetable in case of disruptions. Such effects are not included in this study.

Minimal utilization rates

This study aims to find the minimal utilization rate $u_{\min}$ that can justify the pre-allocation/reservation of the MWs W in the annual timetable. The minimal utilization rate is illustrated in **Figure 13**. Although the curve can look different in reality, the illustration assumes that the benefits increase more than the costs when the utilization rate increases.

The comparison alternative is the do-nothing scenario, i.e., no utilization of MWs ($u = 0\%$). In this case as shown in **Figure 13**, the total net social value is negative and includes both the loss in production (costs of cancelled traffic) as well as the total corrective maintenance costs. The latter, as will be explained later, includes both the work and delays costs.

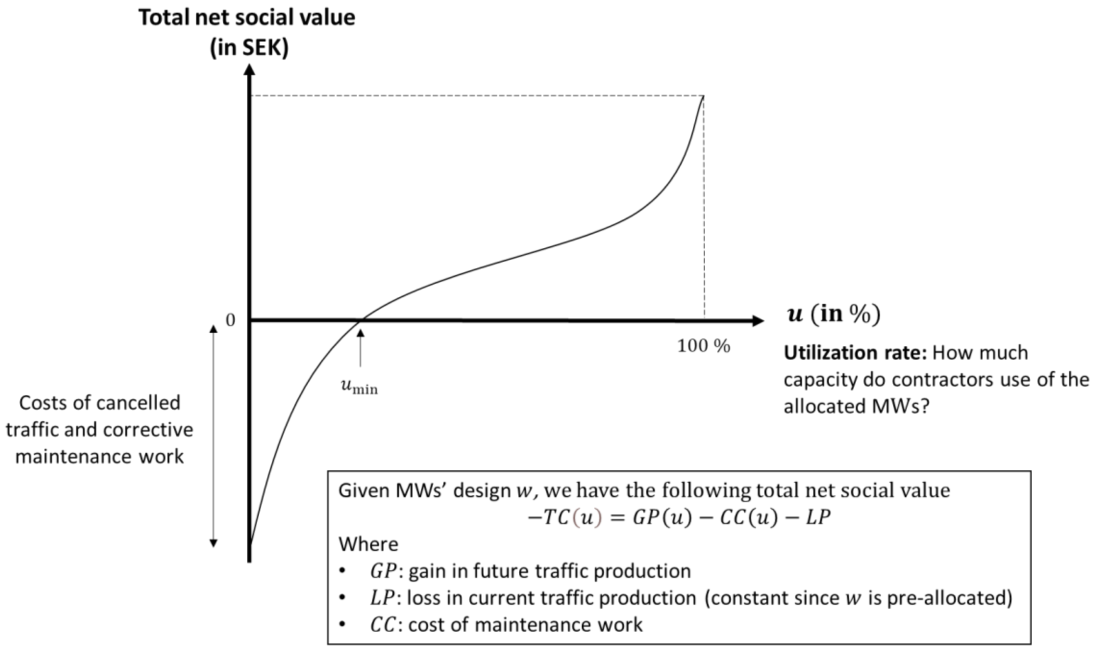

Figure 13. Illustration of the minimal utilization rate for MWs.

As mentioned in the literature, the net present value ratio is often used to present the resulting profitability of the different utilization rates. However, we use the total net social value since MWs' profitability always increases with higher utilization rates, i.e., full utilization (or $u = 100\%$) is always the most profitable policy. We are therefore interested in finding the minimal utilization rates after which the total net social value becomes positive.

3.2. The cost of maintenance work

The cost of construction work CC depends, among others, on the time spent on the track before finishing the maintenance activity. We consider the direct time needed for the steps to perform the different maintenance activities (Hedström, 2020). Although not included in this study, there are additional indirect costs for maintenance work such as license assessment, analysis requirement, project management, and costs for procurement.

Given utilization rate u, the cost of work $CC(u)$ can be formulated as a function of the effective work time $T_{eff} = u\,T(W)$ as in equation (1).

$$CC(u) = (1 + \rho)\, k_{time} \left(t_{transport} + T_{eff}\right) + uK_{fixed}\,, \tag{1}$$

where the additional overhead time includes the time for transport (noted $t_{transport}$). K_{fixed} is the average fixed cost of the material needed for work, k_{time} is the cost of work per time unit, and ρ is the compensation factor for night shifts. In general, activities may include additional delay and/or waiting times. However, we do not consider these additional overhead times in this study. Hedström (2020) has attempted to estimate the time needed for different types of maintenance works such as the replacement of track, rail or overhead lines, and Lidén (2014) gives an indicative list of different maintenance activities (including inspection) and their corresponding typical possession times.

The transport time $t_{transport}$ depends on the number of shifts n_{shift} that are performed. If share of utilized access time is used, such number can be calculated as $n_{shift} = \left\lceil \frac{T_{eff}}{n_{hours}} \right\rceil = \left\lceil u\, n_{weeks}\, n_{days} \right\rceil$.

A unitary transport time can be used, i.e., for traveling between the site and depot. In the case study, we assume that the average transport or setup time is $t_{transport} = 1$ hour for day-time and 3 hours during night shifts. In reality, this average time depends on, e.g., contract (response time requirement for corrective maintenance), site (remote or dense areas) and labor law (work conditions). Moreover, the average cost K_{fixed} for material can be estimated based on statistics of the total yearly costs, e.g., 1 200 MSEK/year for switches and crossings in Sweden. Different activities have different cost parameters, we consider the same parameter values that are used by Lidén and Joborn (2016) and Lidén et al. (2020), see **Table 2**.

Table 2. Adopted values for maintenance costs, also used by Lidén and Joborn (2016) and Lidén et al. (2020).

	Value		
Parameter (notation)	**Day**	**Night**	**Unit**
$t_{transport}$	1	3	Hour
k_{time}	1 250	10 000	SEK per hour
ρ	0	60	%

3.3. The loss in traffic production due to maintenance work

Allocating capacity for MWs can either lead to the reduction or removal of potential traffic. This will lead to a certain loss $LP = LP(W)$ of traffic production that depends only on the characteristics w of the maintenance window, i.e., independent from the utilization rate u of the window.

Trafikverket (2020c) uses priority criteria to estimate the costs of train path cancellation. Given a type of traffic k (e.g., freight, commuter or highspeed), there are $N_k = N_k(W)$ trains paths which are canceled to pre-allocate capacity for MWs W. Based on such criteria, the loss in traffic production, due to the slots for MWs W, is given in equation (2).

$$LP = \sum_{\mathrm{k}} N_k(\mathrm{Time_k} \times (100\% + \mathrm{K_k}) \times (100\% + \mathrm{J_k}) \times \mathrm{B_k} + \mathrm{Distance_k} \times \mathrm{C_k}) \ , \quad (2)$$

where $\mathrm{B_k}$ and $\mathrm{C_k}$ are, respectively, time and distance cost parameters for excluding a path of train type k. The percentage parameters $\mathrm{K_k}$ and $\mathrm{J_k}$ are used to account for the exclusion of train paths of type k. They refer to the correction factor for the base time and the utility threshold of the train path, respectively. For more details on these cost parameters, see (Trafikverket, 2020c).

N_k can be assumed to be dependent on the frequency of the train services as well as W. Assuming a frequency F (in number of departures per hour) on a single-track/direction, we have $N(W) = F \times T(W)$ canceled trains in that direction. The number N of canceled train paths can also depend on the number of tracks that are closed/unavailable for maintenance. In case both directions/tracks are unavailable, the number of canceled trains is doubled assuming similar traffic in both directions.

If trip distributions, e.g., origin-destination matrix often for commuter train services using smart cards, are known, there are alternative and more accurate methods to calculate the social costs of certain traffic services. For instance, Ait-Ali et al. (2020) developed a model to compute such loss in traffic production when canceling/modifying commuter train services.

The cost parameters for the loss in traffic production depend on the type of traffic that corresponds to the excluded train path. Based on recommended values from Trafikverket (2020c), we adopt the parameters that are provided in **Table 3** for the two studied train path categories, namely commuter (SP) and intercity passenger trains (FX) and freight services (GS).

Table 3. Parameter values for train path exclusion, also used by Trafikverket (2020c).

Traffic	Category	Cost parameters			
	A	B (SEK/min)	C (SEK/km)	J (%)	K (%)
Commuters in large cities	SP	1238	104	15	20
Intercity (higher speed)	FX	816	71	20	6
Freight (higher speed)	GS	269	61	15	2

Nelldal et al. (2019) report that around 78% of train traffic (in train-km) in Sweden is for passengers whereas the remaining 22% is for freight. In the absence of better statistics, we use similar shares in this case study. We also assume that affected train paths are from the following traffic categories: 20% of traffic is for freight and the other 80% is for passengers, of which half is intercity (40%) and the other half for regional/local commuting trips (40%). For MWs that are scheduled during the nights, we assume that the train traffic is predominantly for freight (100%).

3.4. The gain in future production

The gain in future production includes both the improvements in service reliability (for the customers) as well as the reduction in corrective maintenance and/or inspections (for the infrastructure manager). We first assume that maintenance contractors have enough knowledge about the studied assets so that they can perfectly schedule these activities. This means that the gains in terms of reduced need for future maintenance/inspection activities are negligible.

The total gain in future traffic production $GP(u)$ can be captured by the benefits BR in terms of traffic reliability (thanks to performed activities with utilized windows) minus the costs CR from unreliability risks (due to unperformed activities in non-utilized windows). One possible formulation is $GP(u) = u\,BR - (1-u)\,p\,CR$, where parameter p is the likelihood of a failure that requires immediate corrective maintenance after inspection. This failure risk often depends on the type of assets and can be affected by other factors such as inspection duration/frequency, traffic volume and weather.

Note that the benefits from increased reliability BR are projected in the future whereas the previously presented costs, i.e., loss in production and maintenance costs, are in present value. We will therefore use, in what follows, a discounting factor r when calculating the future gains for conversion into a present value.

Benefits of increased traffic reliability

Several maintenance activities are scheduled to take place periodically, e.g., as recommended by the manufacturer or based on knowledge of the assets, i.e., life-cycle analysis. We assume that maintenance contractors have enough knowledge of the asset to be able to efficiently schedule or select the right time and location for maintenance activities. Thus, the losses (in red color in **Figure 11**) which are due to early maintenance are neglected, see **Figure 14**.

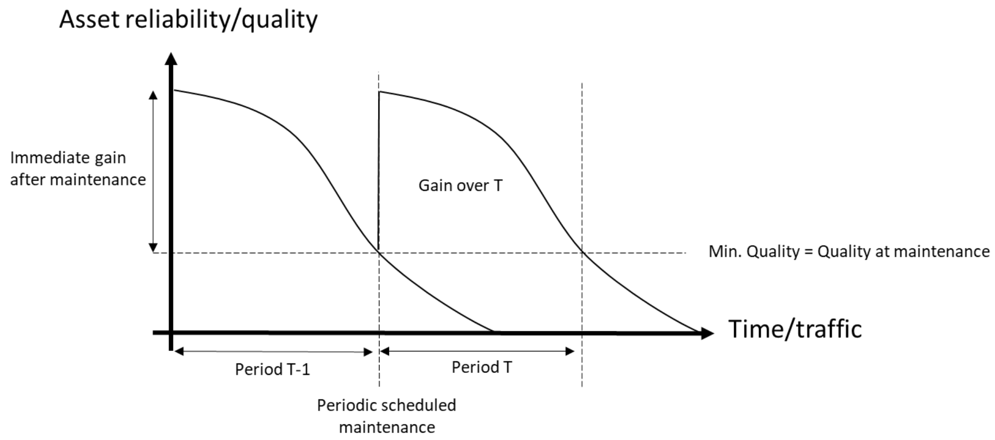

Figure 14. Gain in performance when maintenance is perfectly scheduled.

Let T_{asset} be the period in years between two consecutive scheduled maintenance activities of the asset, e.g., 1 year in average for switches and crossings. To account for the discounted gains of maintenance over the period T_{asset}, we set the minimal asset quality/reliability to zero and use the formulation in equation (3).

$$BR(u) = \int_0^{T_{asset}} \frac{Q(t)}{(1+r)^t} dt \ , \tag{3}$$

such that $r = 0.35$ is the discount factor and $Q(t)$ is the production quality at time $t \in [0, T_{asset}]$.

Since it is difficult to estimate the monetary value of an asset given a certain level of production quality, we use the proxy benefits from avoiding disruption costs such as corrective maintenance costs, delays and/or discomfort. Statistics exist on average delays for different asset failures (Lidén, 2019).

In this study, we estimate the benefits just after the maintenance $Q_0 = Q(t = 0^+) = C^{corr} + C^{del}$, where C^{corr} is the cost for corrective work assumed to be similar to maintenance work with an overhead cost and C^{del} is the cost of delays for both passenger and freight services. How these are calculated is presented in the following subsection about the failure risks due to non-utilized windows, i.e., **Costs of the failure risk**.

How the future gain function $Q(t)$ varies over time, from $t = 0^+$ to $t = T_{asset}$, depends on many factors, e.g., weather conditions, type of asset, the volume of train traffic (Andersson et al., 2016) and axle load (Odolinski, 2019). Let i be the curvature parameter (we will study two different variants), we formulate the future gain function as $Q(t) = Q_0 \sqrt[i]{1 - \left(\frac{t}{T_{asset}}\right)^i}$. We will focus on the cases, as in **Figure 10**, where $i = 1$ and $i = 3$, i.e., linear/type 2 and exponential/type 1, respectively. Both are typical survival/degradation processes which are relevant for rail infrastructure assets.

Costs of the failure risk

To account for the costs from failure risk, we calculate the total costs due to infrastructure failure requiring immediate corrective maintenance. This total cost is noted Q_0 and consists of the costs for corrective work C^{corr} and disruption delays C^{del}.

The costs C^{corr} for corrective maintenance work is calculated in a similar way to that of the schedule-based maintenance work CC with a higher overhead cost parameter $\rho^{corr} > \rho$. Moreover, the effective access time T_{eff} is assumed to be the average required time to repair the asset, e.g., the average total delay.

To calculate the costs C^{del} due to disruption delays, we use existing statistics about delays due to asset failure, e.g., reported by Lidén (2019) for assets such as switches and crossings (SC), overhead or contact wires (CW) and track circuits (TC). SC are reported to be the cause of considerable delays which accounts for more than 13% of the maintenance costs in the Swedish railways (Ghodrati et al., 2017). Based on statistics from *Ofelia* database (see **Appendix 2: Data and information systems**), the authors found that failures in switches lead to an average delay of 20 minutes. Given such average delays, Trafikverket (2020a) provides the cost parameters in **Table 4** to calculate the total social costs for delays.

Table 4. Valuation of delays (e.g., after disturbances) in SEK per hour and person or ton, data by Trafikverket (2020a).

Type of traffic	Average market share (in %)	Valuation of delays
Passenger long distance	40	298 SEK/hour
Passenger local/regional	40	282 SEK/hour
Freight	20	3.85 SEK/ton

Note that the valuations that are presented in **Table 4** correspond to the private commuting trips for passenger traffic and the average over all goods for freight services. Moreover, the total cost of delay over all types of traffic is calculated using the same average market shares that were mentioned in the end of **Section 3.3**.

3.5. Limitations

Several assumptions and simplifications are used in the model. Although not used for maintenance, many pre-allocated slots part of MWs are potentially used for traffic production or to increase its robustness in case of disruptions. These benefits, which could reduce the total social cost, are not included in the CBA model.

In addition to the limitations of the CBA methodology (Van Wee, 2007), the developed model does not include several costs, benefits and externalities which could affect the resulting total costs. For instance, track access charges (main revenues for the infrastructure manager) and environmental effects (emissions and noise) are not included. Moreover, the costs of associations or missing connections are not considered when calculating the loss in production using the priority criteria by Trafikverket (2020c).

An important limitation of the model is the assumption that maintenance activities are perfectly scheduled. This in turn assumes that the maintenance contractors have perfect knowledge of its assets and can therefore perfectly plan the time and place to perform the maintenance activities. This is however not the case in reality as there is often absence of reliable relevant information of the assets. This uncertainty is further increased by the reforms in railways such as deregulation (Andersson and Hultén, 2016). This means that that maintenance activities may be performed earlier or later which leads to additional costs due to excessive maintenance (if earlier) or increased failure risks (if later).

4. Case study

In this section, we present the input data and test scenarios. Results are presented and discussed later in the section.

4.1. Input data

We choose in this case study to focus on the intersection between the Southern Main Line (*Södra Stambanan*) and the line between *Hallsberg* and *Malmö*. Both lines connect important centers for passenger and freight traffic in Sweden. Another reason for the choice is the data availability.

Figure 15. The southern main line, the case study focus is between *Mjölby* and *Malmö (Kavelgrisen, 2017).*

In **Table 1**, we present the main input data that is used to generate the results in this case study. Such data is collected from various sources, e.g., infrastructure manager, operators and academic reports such as (Nelldal et al., 2019).

Table 5. Summary of the main characteristics of the case study.

	Commuter	Highspeed	Freight	Unit
Line	Norrköping - Mjölby	Stockholm - Malmö	Hallsberg – Malmö (vial Mjölby)	-
Distance (speed)	79 (140)	614 (200)	450 (135)	km (km/h)
Travel time	0:49	4:25	3:20	h:min
Passengers	66	138	-	pax/train
Work (business)	50 (20)	0 (69)	-	pax/train
Goods	-	-	800	ton/train

4.2. Test scenarios

In this case study, we look at different test scenarios. Some of these are mainly based on use cases from related projects, e.g., FR8RAIL-II (2020). Scenarios are constructed so that the results can be compared to draw meaningful conclusions.

To study the effect of MWs on the minimal utilization rate, we consider various designs for these windows. **Table 6** presents the three different studied designs, it also provides some of the main characteristics of each design. For instance, night traffic is assumed to be predominantly for freight services.

Table 6. Overview of the studied designs of MWs and their main characteristics.

Notation	Design of maintenance window (MW)	Potential traffic	Production loss	Work cost
day-all	MW-day & all track closure	mostly passenger	High	Low
night-all	MW-night & all track closure	mostly freight	Low	Medium
day-single	MW-day & single-track (speed reduction)	mostly passenger	Medium	High

We have previously described different variants of the model, i.e., degradation function of the studied asset (noted k), definition of the utilization rate (noted u). These are also studied by considering different test scenarios. **Table 7** presents the different scenarios that are used to generate results. These are compared and discussed in the following section.

Table 7. Overview of the scenarios that are tested and discussed in the result section.

MW design	Degradation function k	Utilization rate u as a share of utilized	Test scenario (notation)
D-all	Exponential (k=3)	access time	day-all-exp-time
		slots	day-all-exp-slot
	Linear (k=1)	access time	day-all-lin-time
		slots	day-all-lin-slot
N-all	Exponential (k=3)	access time	night-all-exp-time
		slots	night-all-exp- slot
	Linear (k=1)	access time	night-all-lin-time
		slots	nigh-all-lin- slot
D-single	Exponential (k=3)	access time	day-single-exp-time
		slots	day-single-exp-slot
	Linear (k=1)	access time	day-single-lin-time
		slots	day-single-lin-slot

4.3. Results and discussions

Using previously described input data, we generate and discuss some of the results corresponding to the different studied test scenarios.

First, we look at the variation of the total gross social costs, i.e., without considering the benefits or gains in future production thanks to the increased reliability of the infrastructure assets. **Figure 16** presents three subplots where each is showing the variation of total gross social costs as a function of the utilization rate of either the access time (X=time) or slots in the MWs. Each subplot corresponds to a particular design of the studied maintenance window. Note that all the results in the figure are based on an exponential degradation function, the linear variant will be studied later in this section.

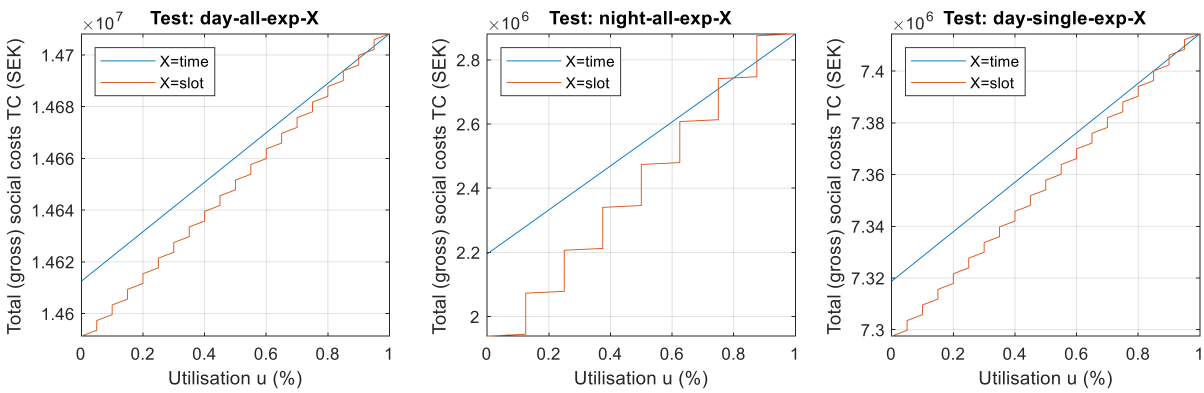

Figure 16. Total (gross) social costs as a function of the utilization rate for different test scenarios.

With no considerations of the benefits/gains, the total gross social costs increase with higher utilization rates of MWs, mainly due to increased work costs. The difference between the two definitions (time or slot) is lower especially for higher utilization rates which is due to lower differences in terms of setup expenses, e.g., transport costs.

Note that the total costs when no slots are used (i.e., $u = 0$) correspond to the cost for train path cancellation which is pre-allocated for MWs instead of train traffic.

Second, we study the variation of the total net social costs including benefits and the corresponding minimal utilization rate of MWs. **Figure 17** shows such variation for both definitions of utilization rates. For each definition, the figure compares the variations for different test scenarios, i.e., designs for MWs.

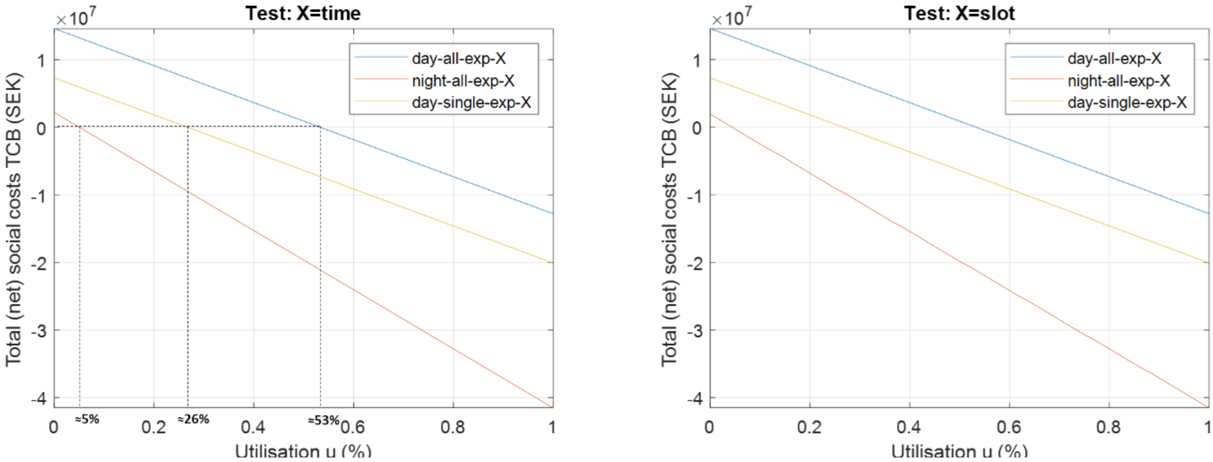

Figure 17. Variation of the total net social costs for different utilization rates.

Similar to **Figure 16**, **Figure 17** shows that there are only slight differences between the resulting social cost variation when using the two definitions. The minimal utilization rates are 5%, 26% and 53% depending on the test scenarios. As expected, night-time windows require lower minimal utilization whereas day-time windows have higher minimal utilization requirements, especially in case of full closure of the tracks. These are mainly driven by the higher loss in traffic production in peak times.

The calculated minimal utilization rate of MWs may also depend on the variant of the model that we use. For instance, we have so far used an exponential type 1 degradation function. In **Table 8**, we present the resulting minimal utilization using different variants of the model, i.e., shifts, track closure, definition of the utilization rate and degradation functions.

Table 8. Minimal utilization rates for other variants or test scenarios of the model.

			Minimal utilization rate (in %)	
Shifts	**Track closure**	**Test scenario**	**X=exp**	**X=lin**
Night	**Full**	night-all-X-time	5	9
		night-all-X- slot	4	7
Day	**Full**	day-all-X-time	53	91
		day-all-X-slot	53	91
	Partial (single-track)	day-single-X-time	26	45
		day-single-X-slot	26	46

Numerical results in **Table 8** show that different minimal utilization rates are required in the various studied situations. Low rates (5-10%) are obtained during night shifts where train traffic volumes are lower. During day shifts with high volume of traffic, the obtained rates are higher (25-50%) in case of partial closure of the tracks (single-track) whereas the utilization rates are the highest (50-90%) when all tracks are closed.

Moreover, the results in the table indicate that there are no significant differences between slot and time-based definitions of the utilization rates. However, minimal rates are shown to depend on the degradation function of the assets (exp or lin), the time (day or night) and the space (single track or full closure). The minimal utilization rate of MWs is the highest (up to 91%) when the asset quality degrades the quickest (lin) and when slots are scheduled during high-volume traffic (day). However, such results assume, among others, perfect knowledge of the asset degradation function as well as a minimum functional quality set to zero. It is possible to study, for instance, the effect of the latter using a sensitivity analysis combined with an analytic expression of the minimal rates given certain fixed parameters such as MW design and degradation function.

In a recent follow-up of the utilization rates of MWs on the Southern main line, Granberg and Rehn (2020) report a level of 51-52% compared to a level of 78-93% on another single-track line in Sweden (*Värmlandsbanan*). When the MWs were introduced, Trafikverket (2015) aimed at, at least, 80 % of all windows being effectively used for maintenance, and that the same level of allocated windows should be fixed in the annual timetable. Recent goals have an even higher goal of 85 % utilization rate (Trafikverket, 2019). The results we have obtained indicate that having higher goals is more appropriate for highly intrusive MWs while lower goals could be used during low-volume traffic or when the windows are less intrusive. To use one goal for all MWs is not justifiable from a social cost-benefit viewpoint.

5. Conclusions

MWs were initially introduced to secure enough capacity or track access time in the annual train timetable that can be used to perform, among others, essential recurrent maintenance activities including inspections, schedule/status-based maintenance, and repairs. One of the purposes for these pre-allocated slots is to make these maintenance activities have as few disturbances to train traffic as possible. Follow-up studies have however shown that such windows are not efficiently used and their utilization rates are low compared to the goals that were initially set, see for instance the recent study by Granberg and Rehn (2020). Alexandersson (2015) mentioned that there are many reasons, e.g., not including all stakeholders in the planning and not enough information about the maintenance needs, in addition to the infrastructure manager lacking enough previous experience.

In this study, we investigated the utilization rates of MWs using a cost-benefit approach. Based on a case study from the Swedish Southern main line, we studied the variation of the total social costs as a function of the utilization rates to estimate the minimal rates for justifying the capacity allocation of the MWs. Such minimal rates are estimated for different test scenarios including designs of MWs and variants of asset life models.

Trafikverket (2020b) has indicated that it is working with projects (e.g., within *Precision banarbete)* and several actions to improve the utilization rates of MWs. Some of these actions are:

1. Possibility to cancel slots for maintenance activities for alternative use, e.g., train traffic
2. Earlier schedule of maintenance plans at key operation sites where all tracks are used
3. Improved coordination between stakeholders and activity areas
4. Segmentation of sites into so-called islands for efficient maintenance plans

In relation to action 4, our results indicate that Trafikverket can improve their MWs-related policies by specifying segment-based minimal utilization rates. The case study provides some numerical examples regarding which utilization rates should be required for different segments, e.g., designs of MWs and traffic situations. These results indicate that MWs during low-volume traffic may only require a 5-10% utilization rate, while partial closures and full closures during high-volume traffic increases the minimal utilization rates to 25-50% and 50-90%, respectively. Moreover, we show that the adopted definition of the utilization rate does not play an important role especially in case of MWs requiring higher utilization rates, e.g., "hot" MWs and important "islands" as mentioned before.

In addition to the previously described improvement actions, a minimal rate of utilization could be used in setting incentive levels for performance-based maintenance contracts. Clearer agreements between the stakeholders, e.g., maintenance contractors and infrastructure manager, including the required minimal utilization rates could give incentives to increase the capacity utilization and therefore reduce the overall maintenance and traffic costs.

Future studies include several improvement and application aspects of the model. In addition to the improvement possibilities that are given by the limitations mentioned in Section 3.5, the model can be extended and applied for more efficient planning of maintenance, e.g., design more economically efficient MWs. Given the importance of prediction compared to reaction in maintenance planning (Ran et al., 2019), similar models can be developed to assess the efficiency of preventive maintenance (Stenström et al., 2016) and renewal strategies (Nilsson and Odolinski, 2020). Other extensions could also include economically smarter planning of maintenance activities using artificial intelligence (AI) and asset data (Kobbacy and Murthy, 2008). For instance, AI methods guided by the CBA model can use GUS data to learn economically efficient maintenance plans (Olauson, 2020), e.g., when and where to schedule maintenance activities.

Acknowledgments
This work was conducted as part of the 2nd working package of the FR8RAIL III research project within the Shift2Rail program of the European Union. The authors are grateful to Matts Andersson for reviewing the paper and for improvement recommendations. We are also grateful to Kristofer Odolinski for earlier references and to Jan-Eric Nilsson for earlier comments and suggestions, and to Magnus Wahlborg for project administration.

References
AIT-ALI, A., WARG, J. & ELIASSON, J. 2020. Pricing commercial train path requests based on societal costs. *Transportation Research Part A: Policy and Practice,* 132**,** 452-464.

AIT ALI, A. 2020. *Methods for Capacity Allocation in Deregulated Railway Markets.* Doctoral thesis, comprehensive summary, Linköping University Electronic Press.

AL-DOURI, Y. K., TRETTEN, P. & KARIM, R. 2016. Improvement of railway performance: a study of Swedish railway infrastructure. *Journal of Modern Transportation,* 24**,** 22-37.

ALEXANDERSSON, G. 2015. Koll på anläggningen. Stockholm: Statens offentliga utredningar.

ANDERSSON, M. 2006. Marginal cost pricing of railway infrastructure operation, maintenance, and renewal in Sweden: From policy to practice through existing data. *Railways***,** 1-11.

ANDERSSON, M., BJÖRKLUND, G. & HARALDSSON, M. 2016. Marginal railway track renewal costs: A survival data approach. *Transportation Research Part A: Policy and Practice,* 87**,** 68-77.

ANDERSSON, M. & HULTÉN, S. 2016. Transaction and transition costs during the deregulation of the Swedish Railway market. *Research in Transportation Economics,* 59**,** 349-357.

ANDERSSON, M., NYSTRÖM, J., ODOLINSKI, K., WIEWEG, L. & WIKBERG, Å. 2011. A strategy for developing a framwork for economic analysis of operation, maintenance and renewal of road and rail infrastructures. *VTI rapport.* Linköping: Statens väg- och transportforskningsinstitut.

BÅNGMAN, G. 2012. *Introduktion till samhällsekonomisk analys*, Trafikverket.

EC 2018. Assessment of unit costs (standard prices) of rail projects (CAPital EXpenditure).

EKSTRÖM, C. 2015. Underhållbehov på järnväg - Jämförelse mellan två regionala banor. *CODEN: LUTVDG/(TVTT-5248)/1-77/2015*.

ELIASSON, J. & BÖRJESSON, M. 2014. On timetable assumptions in railway investment appraisal. *Transport Policy,* 36**,** 118-126.

FOKKERT, J. V. Z. D., DEN HERTOG, D., BERG, F. V. D. & VERHOEVEN, J. 2007. The Netherlands schedules track maintenance to improve track workers' safety. *Interfaces,* 37**,** 133-142.

FR8RAIL-II 2020. D 3.3 - Requirements for a Decision Support Tool. *Shift2Rail.*

GAUDRY, M., LAPEYRE, B. & QUINET, É. 2016. Infrastructure maintenance, regeneration and service quality economics: A rail example. *Transportation Research Part B: Methodological,* 86**,** 181-210.

GHODRATI, B., AHMADI, A. & GALAR, D. 2017. Reliability Analysis of Switches and Crossings: A Case Study in Swedish Railway. *International Journal of Railway Research,* 4**,** 1-12.

GRANBERG, A.-S. & REHN, M. 2020. Öka nyttjandegrad av servicefönster. *In:* TRAFIKVERKET (ed.). Resultatrapporter: Trafikverket.

GÖRANSDOTTER, H. & DYRSSEN, E. 2017. *Servicefönster i järnvägsanläggningen - En studie om varför nyttjandet ser ut som det gör.* LTH.

HEDSTRÖM, R. 1996. *Samband mellan trafikbelastning-banstandard-underhållskostnader: en litteraturstudie*, Statens väg-och transportforskningsinstitut., VTI meddelande 806.

HEDSTRÖM, R. 2020. Tider i spår för underhållsarbeten. *VTI rapport.* Linköping: Statens väg- och transportforskningsinstitut.

HONAUER, U. & ÖDEEN, S. 2018. Underhållsplan 2018–2021. *In:* TRAFIKVERKET (ed.).

HONAUER, U. & ÖDEEN, S. 2020. Underhållsplan 2020–2023. *In:* TRAFIKVERKET (ed.).

JUNTTI, U. & JÄGARE, V. Enhanced collaboration models in eMaintenance projects. International Heavy Haul Association (IHHA) STS 2019 Conference, 2019. 925-931.

KALINOWSKI, T., MATTHEWS, J. & WATERER, H. 2020. Scheduling of maintenance windows in a mining supply chain rail network. *Computers & Operations Research,* 115**,** 104670.

KAVELGRISEN 2017. Karta över södra stambanan enl Trafiksäkerhetsverket 15 feb 2017.

KOBBACY, K. A. H. & MURTHY, D. P. 2008. *Complex system maintenance handbook*, Springer Science & Business Media.

KRISTOFFERSSON, I. 2019. Indikatorer för ökad punktlighet på järnväg : slutrapport inom projektet Nypunkt. *VTI rapport.* Linköping: Statens väg- och transportforskningsinstitut.

LANDER, S. & PETERSSON, J. 2012. *Tamping planning in railway maintenance: improvement potential for optram as decision support.*

LIDÉN, T. 2014. *Survey of railway maintenance activities from a planning perspective and literature review concerning the use of mathematical algorithms for solving such planning and scheduling problems*.

LIDÉN, T. 2018. *Concurrent planning of railway maintenance windows and train services.* Doctoral thesis, comprehensive summary, Linköping University Electronic Press.

LIDÉN, T. 2019. STAPLA-F research report. *In:* PRESS, L. U. E. (ed.) Version 3 ed. Linköping: Department of Science and Technology.

LIDÉN, T., BRUNSSON, L. & LUNDSTRÖM, F. 2020. Utformning av servicefönster för varierande trafik- och underhållssituationer. Linköping.

LIDÉN, T. & JOBORN, M. 2016. Dimensioning windows for railway infrastructure maintenance: Cost efficiency versus traffic impact. *Journal of Rail Transport Planning & Management,* 6**,** 32-47.

LING, D. 2005. *Railway renewal and maintenance cost estimating.* Doctoral degree PhD thesis, Cranfield University.

MAZZUCHI, T. A., VAN NOORTWIJK, J. M. & KALLEN, M.-J. 2007. Maintenance Optimization. *Encyclopedia of Statistics in Quality and Reliability.*

NELLDAL, B.-L., ANDERSSON, J. & FRÖIDH, O. 2019. *Utveckling av utbud och priser på järnvägslinjer i Sverige 1990-2019: Avreglering och konkurrens mellan tåg, flyg och buss samt jämförelse mellan tåg-och resenärspunktlighet,* Rapport, Kungliga Tekniska högskolan (KTH).
NILSSON, J.-E. & ODOLINSKI, K. 2020. When should infrastructure assets be renewed? : the economic impact of cumulative tonnes on railway infrastructure. *Working papers in transport economics.* Stockholm.
NILSSON, J.-E., PYDDOKE, R., KARLSSON, R. & JOHANSSON, A. 2015. Tidtabelläggning : principer, tumregler och utfall. *VTI rapport.* Linköping: Statens väg- och transportforskningsinstitut.
NISSEN, A. 2009. LCC analysis for switches and crossings : a case study from the Swedish Railway Network. *International Journal of COMADEM,* 12**,** 10-19.
ODOLINSKI, K. 2019. Estimating the Impact of Traffic on Rail Infrastructure Maintenance Costs : The Importance of Axle Loads. *Journal of Transport Economics and Policy,* 53**,** 334-350.
ODOLINSKI, K. & BOYSEN, H. E. 2019. Railway line capacity utilisation and its impact on maintenance costs. *Journal of Rail Transport Planning & Management,* 9**,** 22-33.
ODOLINSKI, K. & WHEAT, P. 2018. Dynamics in rail infrastructure provision: Maintenance and renewal costs in Sweden. *Economics of Transportation,* 14**,** 21-30.
OLAUSON, E. 2020. Framtidens Järnvägsunderhåll. *SOU.*
PATRA, A. P. 2007. *RAMS and LCC in rail track maintenance.* Luleå tekniska universitet.
PERREAL, Y., HAINZ, S., VANNIER, E., KRISTOFFERSSON, I. & MEYER ZU HÖRSTE, M. 2019. A methodology to assess the impact of end-user centric innovations on railway transportation attractiveness.
RAN, Y., ZHOU, X., LIN, P., WEN, Y. & DENG, R. 2019. A Survey of Predictive Maintenance: Systems, Purposes and Approaches. *ArXiv,* abs/1912.07383.
RAUSAND, M. & VATN, J. 2008. Reliability Centred Maintenance. *Complex System Maintenance Handbook.* London: Springer London.
RAYHUSTHWAITE 2009. Survivorship curves. *In:* CURVES.JPG, S. (ed.). English Wikipedia, CC BY 3.0, .
STENSTRÖM, C., NORRBIN, P., PARIDA, A. & KUMAR, U. 2016. Preventive and corrective maintenance – cost comparison and cost–benefit analysis. *Structure and Infrastructure Engineering,* 12**,** 603-617.
SWECO & WSP 2015. Kapacitetstilldelning och prioriteringar i tågplaneprocessen : problem, möjligheter och förslag till förbättringar. *Trafikverkets publikationer.* Borlänge: Trafikverket.
TRAFIKVERKET 2015. Långsiktigt hållbara affärer för Trafikverket på den svenska underhållsmarknaden.
TRAFIKVERKET 2019. Järnvägsnätsbeskrivning 2019.
TRAFIKVERKET 2020a. Analysmetod och samhällsekonomiska kalkylvärden för transportsektorn: ASEK 7.0.
TRAFIKVERKET 2020b. En punktligare tågtrafik: sammanställning av Trafikverkets åtgärder 2017–2019. *In:* TRAFIKVERKET (ed.).
TRAFIKVERKET 2020c. Järnvägsnätsbeskrivning 2020 - Prioriteringskriterier. Swedish Transport Administration.
TRAFIKVERKET 2020d. Network Statement 2020. Swedish Transport Administration.
UIC 2015. Key cost drivers in railway asset management. *In:* RAILWAYS, I. U. O. (ed.).
VAN DER WEIDE, J. A. M., PANDEY, M. D. & VAN NOORTWIJK, J. M. 2010. Discounted cost model for condition-based maintenance optimization. *Reliability Engineering & System Safety,* 95**,** 236-246.
VAN WEE, B. 2007. Rail Infrastructure: Challenges for Cost–Benefit Analysis and Other ex ante Evaluations. *Transportation Planning and Technology,* 30**,** 31-48.
WSP 2014. Inventering, uppdatering och sammanställning av effektsamband inom investering, trafikering och drift och underhåll på järnväg.
ZOETEMAN, A. 2001. Life cycle cost analysis for managing rail infrastructure. *European Journal of Transport and Infrastructure Research,* 1.

Appendix 1: Examples of infrastructure assets and maintenance activities

Asset	**Maintenance activities**	***Swedish translation***
Tracks	Track replacement	*Spårbyte*
	Track grinding	*Spårslipning*
	Sleeper replacement	*Slipersbyte*
	Track fastener replacement	*Befästningsutbyte*
	Track straightening	*Spårriktning*
Ballast	Ballast maintenance	*Underhåll av ballast (exkl. ballastrening)*
	Ditch cleaning, water removal	*Dikesrensning, Avvattning*
	Revision of under-ballast structure	*Revision av underballast*
	Ballasting	*Ballastrening*
Switches	Change of switches	*Växelbyte*
	Warming up switches	*Växelvärme*
Signaling	Signaling system	*Signalsystemsåtgärder*
	ATC	*ATC*
	Upgrade to ERTMS	*ATC till ERTMS*
	Detectors	*Detektorer*
Electricity	Overhead line replacement	*Kontaktledningsbyte med/utan stolpbyte*
	Converter	Omformare
	Energy supply system	*kraftförsörjningssystemet*
	Electrification	*Elektrifiering*
trains and capacity	Speed increase measures (tracks and signaling)	*Hastighetshöjande åtgärder (spår- och signalåtgärder)*
	Extension of sidings	*Förlängning av mötesstationer*
	Extension of platforms	*Plattformsåtgärder för ytterligare tåg*
Others	Maintenance of fences	*Underhåll av stängsel*
	Level crossing	*Plankorsningar för effektivare trafikstyrning*
	Winter measures (snow or ice removal)	*Vinteråtgärder (snöröjning, isrivning, beredskap)*
	Protection from trees	*Trädsäkring*

Appendix 2: Data and information systems in Swedish railways

Maintenance activities can generate large amounts of data that can be used to follow-up the efficiency of the maintenance plans and to potentially improve them. In Sweden, a number of databases have already been used to store and access valuable data on different aspects of the traffic and/or infrastructure, e.g., *BANSTAT* (about train traffic) and *TFÖR* (about traffic delays).

Many of the early databases are being revised and combined for better data structuring (Andersson et al., 2011). There are however a number of currently used computer systems including databases:

- *Ofelia* is where infrastructure failures and repairs are reported. It is used by both contractors (to report) and analyst (to study statistics).
- *LUPP* is useful for statistics on train punctuality and traffic disturbances. It has replaced both *BANSTAT* and *TFÖR*.
- *BIS* is a digital system for information about current railway infrastructure assets. An interactive geographical visualization is also available.
- *BESSY* is a digital tool for documenting and planning inspections of railway infrastructure assets.
- *OPTRAM* is where periodic measures are reported from track recording trains. Measures include data on assets such as tracks, overhead lines, ballast and rail profile.

New systems and projects have also been introduced to improve the usability of existing ones, e.g., *RufusOnline, GUS* (Ekström, 2015), *ePilot* project (Juntti and Jägare, 2019). For instance, *GUS* is a new system that aims to help supervise, plan and follow-up road and railway (preventive) maintenance.