Robust decision making for future train maintenance

Illustrating the abilities of decision making under deep uncertainty within the train maintenance context

Kerem Tunca

Delft University of Technology

Robust decision making for future train maintenance

Illustrating the abilities of decision making under deep uncertainty within the train maintenance context

---

Master thesis submitted to Delft University of Technology

in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in **Engineering & Policy Analysis**

Faculty of Technology, Policy and Management

by

Kerem Tunca

Student number: 4601386

To be defended in public on April 13th 2023

Graduation committee

Chair:	Dr.ir. I. (Igor) Nikolic
First supervisor:	Dr. J. (Jazmin) Zatarain Salazar
External supervisor:	T. (Thomas) Milde

Cover: Oskam (2020).

”The best way to predict the future is to invent it.” - Alan Kay

Preface

Before you lies my biggest accomplishment so far: an MSc thesis. It is the final product of six and a half years of studying at the Delft University of Technology. While some moments have been tough, it has been truly inspiring to have been a part of this university as a student. This project in specific contains a hands-on application of how decision making under deep uncertainty can be included in the railway maintenance sector. It was a pleasure being able to introduce a rather novel academic field to Nederlandse Spoorwegen (NS), and I sincerely hope that they will benefit from what is presented in this thesis.

Furthermore, I would like to take this opportunity to thank the people that have supported me throughout this project. I am grateful to my TU Delft supervisors Igor Nikolic and Jazmin Zatarain Salazar for keeping me on the right track. Your feedback has been immensely valuable. I would like to bring out a special thanks to Thomas Milde. Thank you Thomas for the weekly coffee moments and long walks on Tuesdays where I could discuss the obstacles that had to be overcome, while also discussing more demanding topics such as dog psychology and dog behavior.

Kerem Tunca
Delft, April 2023

Summary

A well-maintained train fleet is a top priority for the Nederlandse Spoorwegen (NS), given that they provide a vital service to the Dutch society. This research focuses on improving the robustness of decisions made by decision makers of the NS train maintenance system, considering the uncertainty and risks that NS faces in maintaining their rolling stock during the next 10-15 years. The maintenance system that this research focuses on consists of the four maintenance locations that NS has: Leidschendam, Maastricht, Onnen & Watergraafsmeer. Within these locations, equipment and mechanics form the maintenance capacity. While the current maintenance capacity performs adequately for now, deep uncertainty arises regarding the impact of the outside world on the future performance of NS train maintenance. Current methods applied by NS decision makers do not permit to take deep uncertainty into account.

This research aimed to show the potential benefit decision makers may have by including robust decision making methods. While this research scope regards train maintenance, methods applied in this thesis could be applied in many other complex systems where deep uncertainty is present. Within this research, an agent-based simulation model that simulates the train maintenance until 2035 has been developed. The model allowed for the evaluation of different future scenario's while monitoring the performance of the maintenance capacity. Experts within NS have been consulted to gain knowledge on how train maintenance performance is measured. It became evident that four key performance indicators should be implemented in the model: maintenance throughput time, train withdrawal, equipment occupancy & delivery reliability.

AnyLogic was used to build the simulation model, where entities in the form of trains, maintenance locations, equipment and mechanics are included as separate 'agents'. Two types of maintenance have been distinguished within the simulation model: scheduled and unscheduled maintenance. Trains follow several steps within scheduled maintenance, of which some are obligatory and some are optional. The optional tasks are based on the train condition as it enters scheduled maintenance. Some equipment can only be used to maintain one specific type of train, while others can be used interdisciplinary. Future uncertainty arises concerning train condition, duration of maintenance tasks, mechanic availability and the number of kilometers trains daily drive within operation.

Through exploratory modeling and analysis 2000 scenario's have been tested. The scenario's are constructed by sampling over the uncertainty space of all model uncertainties. Latin Hypercube Sampling (LHS) enabled generating scenario's which allowed for elaborate exploring of the impact of all uncertainties on future train maintenance. The simulation model responded sensitively to three model uncertainties: the number of available mechanics, the daily number of kilometers intercity's drive, and the number of kilometers sprinter train types drive on a daily basis. Scenario discovery showed that these model uncertainties were found to be of significant contribution towards desired model outcomes. If less mechanics will be available compared to the current number of mechanics, issues within train maintenance might arise. In case intercity's and sprinters increase daily driven kilometers by 34% and 17% respectively, issues within train maintenance might occur. For NS it is adviced to monitor these uncertainties carefully.

Few policy interventions were tested, of which the trade-off between those policies has been presented. The advice towards NS is to use this simulation model to evaluate more policies based on what future decisions will be made. This approach is advised to be used as a support tool for deciding where to maintain new train types, as it allows for the quantification of future maintenance performance including those new train types.

Regarding future research, it is recommended to improve current methods by including mechanic qualifications within the simulation model. In addition, it is recommended to compare the effect of current methods with assigning more intelligence to agents within the simulation model. Another recommendation is to evaluate structural uncertainty within the train maintenance system, which could narrow down the scope of what needs to be carefully addressed when making decisions regarding future train maintenance capacity. A final recommendation for further research would be to include the effect of geographic location of the four maintenance locations of NS, given that some trains operate far from Maastricht & Onnen. Maintaining them at Maastricht or Onnen could lead to implications that has been out of scope of this research, but seems valuable for NS to quantify in the near future.

In short, it can be concluded that this research has illustrated the relevance of decision making under deep uncertainty within the train maintenance context. Furthermore, to enhance decision making within NS, the methods used in this research should be applied more extensively. Rather than focusing on optimization issues, the focus should lie on organizing train maintenance to absorb any future perturbations flexibly. This would severely improve future robustness, which is what modern day society requires: reliable train service provided by NS, backed up by a robust future maintenance system.

Contents

Nomenclature

Abbreviations

Abbreviation	Definition
AW	Aardwind
DM	Decision-maker
DMDU	Decision Making under Deep Uncertainty
DDZ	Dubbeldekker Zonering
EMA	Exploratory Modeling and Analysis
FLIRT	Flinker Leichter Innovativer Regionaltriebzug
ICM	Intercity Materieel
ICNG	Intercity Nieuwe Generatie
ICR	Intercity Rijtuig
KPI	Key Performance Indicator
KWB	Kuilwielenbank
LHS	Latin Hypercube Sampling
ML	Maintenance Location
NS	Nederlandse Spoorwegen
PRIM	Patient Rule Induction Method
RDM	Robust decision-making
SLT	Sprinter Lighttrain
SNG	Sprinter Nieuwe Generatie
VIRM	Verlengd Interregiomaterieel

List of Figures

List of Tables

List of Algorithms

1 Introduction

Ensuring a safe, comfortable, and punctual journey for all customers is one of the top priorities of the Dutch railway operator Nederlandse Spoorwegen (NS). To successfully do so, their trains will have to be maintained on a regular basis, which is done at one of the few maintenance workshops located in The Netherlands. NS wishes to make robust decisions regarding their train maintenance to cope with deep uncertainty that they face in their long-term planning. Uncertainties arise at different aspects, such as (future) train failure, future train fleet size, travel demand, and the lack of knowledge concerning future maintenance tasks. Creating insights into the effect of these uncertainties allows decision makers (DMs) to translate operational issues to strategically relevant information that will enable them to make relevant decisions regarding the future of train maintenance. This will permit NS to create policies that will benefit their future robustness. Beside the interest of NS, it is highly relevant for society to have a well-functioning public transport system in the future. Substituting non-sustainable modes of transport for travelling by train is in line with the government's policy that aims for a climate neutral society by 2050 (Ministry of Economic Affairs & Climate Policy, 2020). In addition, NS provides a vital service for the Dutch society (Ministry of Justice & Safety, 2021), which demonstrates the relevance of having a well-maintained train fleet.

1.1. Train Maintenance System

Maintenance is currently performed in one of the 4 workshops of NS: Onnen (O), Watergraafsmeer (W), Leidschendam (L), and Maastricht (M). Figure 1.1 displays each maintenance location on the map of the Netherlands. Currently, train maintenance is performed based on either a time or a distance constraint. Once a train exceeds this constraint, a train is shunted to its maintenance shop. Each maintenance shop handles specific types of trains, pointed out on the right side of figure 1.1.

Once a train arrives at a workshop to perform regular maintenance, it follows a fixed sequence of maintenance tasks: empty the toilet reservoir, replace/fix parts of the carriage, polish the wheels and/or replace the chassis (both tasks are performed only if considered necessary), perform extensive in- and outside cleaning, and finally perform tests to ensure that the train is ready to go into operation. All in all, it takes about 6 to 8 shifts (8 hours per shift) before a train is ready to go back into operation.

Thus, the arrival of a train into a workshop is done based on preventive maintenance, whereas most tasks within the maintenance stop are condition-based. For example, an air conditioner or chassis is replaced if a mechanic determines, after inspection, that the current state of the part is due for replacement. Wear causes the train wheels to vibrate slightly, which is measured by sensors. If these sensors indicate the presence of vibrations, the wheels are polished and calibrated.

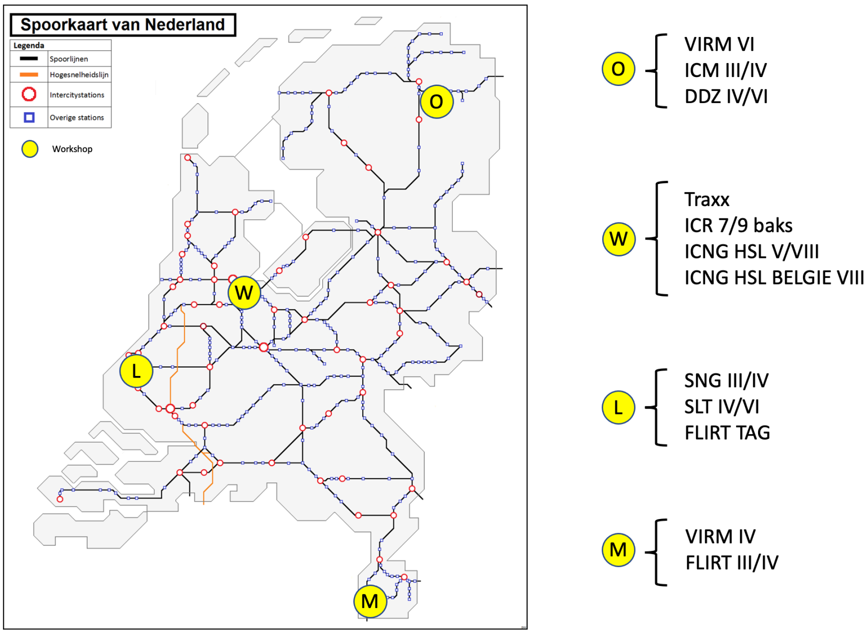


Figure 1.1: Maintenance shop locations of NS together with the specific train types that a workshop maintains

Beside regular scheduled maintenance, a train can also be shunted to a workshop in case of a failure or system error. The mechanics of the workshop then perform so-called pitstops. This regards unexpected, unscheduled maintenance triggered by train failure or a requirement of immediate repair. Mechanics of a workshop aim to fix this train within 24 hours, to minimize the withdrawal time of a train. As soon as a train enters the workshop for a pitstop, a timer begins at 24:00:00 and starts counting down to 0. Tv screens placed throughout the workshop show the time that mechanics have left to repair the train and send it back into operation. This is a way of dealing with sudden train failure to ensure a minimal time of train withdrawal.

1.2. Complexity & Uncertainty

The NS maintenance system, consisting of trains, workshops, and mechanics is a complex adaptive system (Van der Lei, Bekebrede, & Nikolic, 2010). It is an emergent system of dependencies, interaction, and adaptation. Train maintenance frequency depends on its reliability, their reliability depends on investments and maintenance performance, and future investments are done based on for example maintenance capacity and peak hour demand. In addition, the maintenance system responds slowly to system interventions. Activities such as building new facilities, training mechanics, or acquiring new trains are time-consuming. Furthermore, developments regarding self-operating trains and new safety systems such as ERTMS (European Union Agency for Railways, 2023) creates additional uncertainty regarding future train maintenance. The latter system safely shortens the distance that trains have to keep with respect to each other, so that more trains can move on a specific section of the rail track, allowing for a larger frequency of movements.

New trains require different types of maintenance, since equipment of new trains is placed on the roof of the train instead of the bottom of the train. Therefore, uncertainty arises as to what type of maintenance will have to be performed, and how long each task will take. The train roof is less accessible than the bottom, which requires elevated platforms so that mechanics are work efficiently. Not all workshops have unlimited capacity of using these elevated platforms, making the maintenance approach more complex.

1.3. Decision making on the NS train maintenance system

Within NS, a workgroup decides what plans should be made for the mid long-term concerning train maintenance. This plan will be sent to a control group, that will have to agree with these plans. Once formally agreed by the control group, the plan will move to the portfolio board. The portfolio board evaluates plans on all kinds of topics within NS, not just future maintenance plans. They will have to decide if a specific maintenance plan is worth executing, when comparing to other plans that are on the table. In case the plan requires a large investment, the board of directors will evaluate based on a business case whether the plan is worth executing.

Decision makers are tasked with safeguarding the future of NS, while dealing with uncertainty that is inherent with the complexity of the system (Litescu, Viswanathan, Aydt, & Knoll, 2016). Incorrect forecasts can lead to a mismatch between maintenance- demand and capacity, resulting in either excessive costs (overcapacity) or insufficient operational trains (undercapacity). Stranded assets as a consequence of overinvestments, or as a consequence of having to perform less maintenance than anticipated, is costly and should thus be avoided. Currently, the main measure that forms the basis for a DM is the financial benefit that a policy will have for NS. While investing in robustness to deal with deep uncertainty might be more expensive in the short term, it might bring along financial benefits in the long term when situations occur that were less anticipated for. Therefore, an interesting task would be to educate DMs to not only focus on the financial benefit based on the current state of the train maintenance system, but also allow for implementing plans that perform well under a large variety of scenarios.

1.4. Robust policies

Rather than searching for a single optimal solution, policy makers dealing with deep uncertainty wish to aim for creating policies that perform well under different scenarios, also defined as robust policies. Hence, the objective of this research is to support policy-making of NS by analyzing the workflows of workshops, exploring responses of the train maintenance system in case of interventions, and evaluating system robustness under small perturbations (Gribble, 2001). Since incorporating robustness in policy-making is a rather novel approach, especially in the train maintenance field, this research will have an important contribution to scientific knowledge in general.

The robustness perspective is not only relevant for train maintenance systems, but could and should be applied more extensively within any type of asset management where uncertainty is present. Uncertainty could rise in the form of unclear effect of relations within the system, while uncertainty could also be present in the form of more detailed specifics such as future capacity or financial resources. This research aims to show that making an effort to understand a system's uncertainties, and model them adequately, would benefit the needs of any decision-maker, not only within train maintenance (the latter being the scope of this research).

1.5. Relevance

Given the grand challenge to successfully transport travelers during the next years, while dealing with the large proportion of uncertainty, this research is perfectly aligned with the EPA-program. It is highly relevant for society to preserve the well-organized train system in the future, which is a challenge that this research aims to contribute to. NS provides a vital service to the Dutch society, which highlights the importance of adequate business continuity. If NS wishes to incentivize people to make more use of public transportation, it needs to offer a reliable, sustainable transportation system that is backed up by robust maintenance. While the NS train maintenance system may never be fully robust, this research attempts to provide an important contribution into raising awareness on the importance of dealing with uncertainties. Modern day society demands a different approach than what has been done so far.

Renewing the line of reasoning behind large decisions would strongly benefit society in the long run, which is aimed to illustrate by performing this research.

While NS currently transports less passengers than before the pandemic, it is known that it will catch up on (and even exceed) previously realised numbers of train passengers. It is also known that society is globally developing towards being more sustainable. Travelling by train, them being non-pollutant, would perfectly fit within the aim of providing sustainable means of transportation. Given these developments, it would be highly relevant to revise decision-making within NS, steering it towards decisions to be made from a robustness perspective rather than focusing on making decisions conform one key performance indicator. The robustness perspective allows NS to not only perform well under current circumstances, but remain to perform well for whatever lies ahead.

Finally, the presence of highly uncertain factors in the maintenance field of NS combined with the complexity of train maintenance is what makes the problem a suitable one to address from an EPA perspective. Approaches that were taught in the EPA program, such as agent-based simulation, robust decision-making (RDM) and tackling deep uncertainty, will be held in this research.

1.6. Report Structure

In this research, first a literature review on policy-making under (deep) uncertainty with relation to train maintenance and RDM is performed. Next, the research approach is discussed based on the research question that followed from the knowledge gap identified in the literature review. Several sub-questions will be formulated to help answering the main research question. In the subsequent section, the appropriate research method that fits with the research approach is discussed. The following chapters will elaborate on the needs of a DM, followed by the conceptualization of the NS train maintenance system, and formalization of the simulation model. Once these parts have been discussed, analyses will present the quantification and impact of uncertainties regarding train maintenance. Finally, the interpretation of the analyses, conclusions, recommendations & discussion, is presented.

2 Literature Review

This chapter will present a literature review, resulting in the formulation of a research question based on the knowledge gap that is found. This will be done by discussing state-of-the-art literature regarding asset management, maintenance strategies, and (deep) uncertainty regarding railway asset management.

2.1. Search Process

The citation database Scopus and Google Scholar were used to find state of the art, authoritative, and peer-reviewed literature on this topic. Search terms such as 'maintenance AND rolling stock AND train AND asset management' as well as '"robust decision-making" AND maintenance AND transport' provided relevant results that will be presented below. To select relevant literature, abstracts and conclusions have been assessed. The selection was done based on methodology, such as a Robust Decision-Making approach, or on the field of study.

2.2. Uncertainty

Currently, the main complexity arises from many uncertainties that are present in the train maintenance system, such as (future) train failure frequency, the effect of different maintenance strategies, the future size of the train fleet, the introduction of new safety systems such as ERTMS, and maintenance of new trains that are yet to be acquired. Throughout the literature, different methods are used to find optimal maintenance strategies, but little has been studied on an approach to dealing with deep uncertainties in the context of train maintenance. Future train failure frequency is highly uncertain, as the reliability of future trains is unknown. The introduction of new safety systems is one that comes with many uncertainties as well, because it is unknown to whether new software will cause for frequent errors that require immediate updates or repairs. Naturally, failures are inherent with the frequent use and exploitation of trains. Besides train failure, many more uncertainties are present in the train maintenance system overall, making it difficult to make strategic plans for the mid-long term. Studies on uncertainty in complex systems did show that relevant results can be obtained while having information inaccuracy (Litescu et al., 2016; Pan, Demiryurek, & Shahabi, 2012). Within their study, Pan et al. (2012) have tried to obtain better prediction models (traffic related). While information inaccuracy within train maintenance needs to be dealt with, the goal of this thesis is not to develop an accurate prediction model for train maintenance, which is the approach held in the study of Pan et al. (2012). The study of Litescu et al. (2016) provides interesting insights into the contribution of an agent-based simulation model to observe the effect of uncertainty (in the form of information inaccuracy) on a complex system. However, the system they are focusing on is not related to asset maintenance.

To cope with uncertainties within asset maintenance, Shafiee and Sørensen (2019) present different models, methods and strategies regarding asset maintenance. While presenting an insightful analysis on dealing with uncertainty in an asset maintenance field, their study focuses on optimization of future maintenance rather than having a robust maintenance system. Cost effectiveness and time efficiency are the main aspects that form the framework of the study of Shafiee and Sørensen (2019), which are aspects that this thesis avoids. In short, there is ample literature on dealing with uncertainty in such a way that a maintenance system performs adequately for a long period of time.

2.3. Robust Decision Making

Robust Decision Making (RDM) refers to an approach where policies are implemented when they perform well (i.e., robust policies) under a large variety of scenarios. The book by Marchau, Walker, Bloemen, and Popper (2019) presents an extensive overview on how to apply decision-making under deep uncertainty. Instead of having an RDM approach, Van Duin, Bauwens, Enserink, Tavasszy, and Wong (2016) choose to have a strategic roadmapping approach to deal with risk. Beside strategic roadmapping, RDM is suggested by Haasnoot, Kwakkel, Walker, and Ter Maat (2013) as an approach that empowers robust decisions that perform well under uncertain future scenarios. While the study of Haasnoot et al. (2013) provides valuable insights on how RDM can be facilitated, it focuses on a water management case rather than the asset maintenance context. The study of Wurth et al. (2019) demonstrates how an adaptive approach is suitable for infrastructures, that of renewable energy in this case, that are coping with deep uncertainty, allowing for robust ‘no regret’ decisions. Again, the approach of this study is highly promising, but focuses on a different field than train maintenance. This demonstrates the current need for this thesis, since no study has been performed on incorporating RDM within the context of train maintenance.

Lai, Fan, and Huang (2015) as well as Burkhalter and Adey (2020) have used mixed integer linear programming to find optimal maintenance schedules, by which they have been able to achieve utilization improvements as well as cost reductions. While mixed integer linear programming proves to be a suitable method for optimization in train maintenance scheduling, it does not have sufficient means to take (deep) uncertainty into account. Tréfond, Billionnet, Elloumi, Djellab, and Guyon (2017) have studied the combination of maintenance and its uncertainty by taking a robustness perspective. Their approach improves robustness and prevents sub-optimality for some criteria of asset maintenance. Mira, Andrade, and Gomes (2020) have continued on the study by Tréfond et al. (2017), and found that uncertainty regarding maintenance tasks did not affect the maintenance schedules. Yet, in worst-case scenarios (having maintenance tasks taking much longer than planned for), no optimal solution is found, meaning that there are overall delays occurring in the maintenance of trains. Their study however focuses on a small fleet size of one type, which is an oversimplification of reality, making it hard to generalize their results.

2.4. Dynamic Approach

As described in the introduction of this thesis, the current models of NS for the long-term planning are static, while dynamic approaches that consider uncertainty are desired. Marchau, Walker, and Van Wee (2010) have looked into a dynamic approach of handling uncertainties instead, by using dynamic policymaking. Their study showed that a dynamic approach seems ”highly promising in terms of handling the range of uncertainties related to the implementation of long-term transport policies”. While they recommend looking further into the dynamic policymaking approach by using simulations, their study doesn’t focus on train maintenance in specific. To evaluate in what ways simulations, with for example agent-based models (ABMs), have been used to analyze rolling stock maintenance performance of railway infrastructure, a further literature search was done. Pinciroli et al. (2020) have made use of

ABMs to optimize Operations & Maintenance in the energy sector, which is comparable with railway asset management. Again, the intention of this study was to optimize the asset maintenance system instead of setting it up in a robust manner. While they have used ABMs, their study lacks the robustness perspective that this thesis aims to have. Alexandrov, Bannikov, and Sirina (2019) have used ABMs to model rolling stock maintenance but also lack the use of dynamic variation of input variables.

2.5. Knowledge Gap & Research Question

The literature study that has been performed shows the novelty of the field that this thesis places itself in; little has been studied on coping with deep uncertainty through the use of simulation modeling within train maintenance. Studies that have had RDM approaches focused on different fields, while studies that did focus on train maintenance lacked the use of input variation or presented results not suitable to be generalized. A knowledge gap can thus be identified on the effect of having a robust decision-making approach to tackle uncertainty that is present in the NS train maintenance system. This results in the following research question:

How can the robustness of decisions made by decision makers of the NS train maintenance system be enhanced, considering the uncertainty and risks that NS faces in maintaining their rolling stock during the next 10-15 years?

3

Approach & Methodology

This section presents the approach that will be held in this research, together with the formulation of sub-questions. The formulated sub-questions will each contribute to answering the main research question. Next, methods and tools will be selected based on research questions. The selection process will be based on the knowledge on what data is needed to perform adequate research, build a simulation model, and perform relevant scenario analyses to write policy recommendations to a DM of NS.

3.1. Research approach

There are many dependencies and uncertainties to consider when finding a suitable way to approach this problem, given the complex adaptive system of NS. NS wishes to evaluate the impact of interventions on the maintenance approach, while taking into account many of the uncertainties that are present in the strategic planning of maintenance. Managing those uncertainties requires adequate, robust decisions. The further the projection, the more divergence will be observed in plausible future scenarios. Keeping track of changes to the outside world, as well as changes within the train maintenance system, is essential for NS to achieve high efficacy.

This research will have a model-based approach, that is driven by NS data. Once the inner workings of the NS maintenance system have been modeled, different future scenarios will be explored. The performance of policies under these different scenarios allows for the evaluation of their robustness, improving the information quality for NS to decide how to set up their future maintenance system.

3.1.1. Modelling Approach

To support RDM under uncertainty, a modelling approach would be beneficial (Kelly et al., 2013). A modelling approach will enable this research to explore and visualize the impact of interventions on the complex train maintenance system. More specifically, the complexity and adaptivity of the system and its interactions will require a bottom-up approach, an agent-based model, in the contribution to gaining deeper understanding of the functioning of the NS train maintenance system (Van der Lei et al., 2010). In this way, the robustness of decisions regarding future maintenance can be examined.

One could argue that discrete event simulation would, instead of agent-based modelling, be an appropriate method for simulating the NS train maintenance system. Yet, in discrete event simulation, "entities are very passive representations and cannot capture some of the necessary decision-making" (Van Dam, Nikolic, & Lukszo, 2012).

Agent-based modelling thus allows us to create a better understanding of how the maintenance system (that consists of trains, maintenance locations, outillage & mechanics) interacts. It also enables the exploration of consequences of system interventions, the evaluation of system behavior under uncertain future scenarios, and the communication of policy recommendations to NS.

It should be kept in mind that there are limitations to the use of agent-based models. They highly depend on initial conditions, have limited transparency and "results are moderately comparable and reproducible" (Manzo, 2014). To deal with these limitations, interviews will be held with those that are closer to the train maintenance field, together with observance of the maintenance on-site, creating a clearer picture of the train maintenance system and improving the level of detail and accuracy of the model.

3.1.2. Sub Questions

Performing adequate research on the train maintenance system requires several intermediary objectives, which will be translated into sub research questions. These intermediary objectives are: identifying the needs of a DM of the NS train maintenance system; creating a sufficient image of the NS train maintenance system; translating the DM's needs into a model; exploring the effect of system interventions and system behavior under uncertainty to gain useful insights; writing policy recommendations that benefit the robustness of decisions.

Sub questions:

1. What are the needs of a decision maker of the NS train maintenance system to make robust decisions?
2. What does the NS train maintenance system entail?
3. What model(s) can be built to support the decision makers of the NS train maintenance system?
4. How can the model provide valuable insights to enhance the robustness of decisions made by NS decision makers?

3.1.3. Research objective

The main objective of this research is to show that simulation models, together with the exploration of future scenarios, contribute to valuable decisions under the presence of deep uncertainties. This will be done by changing perspectives on decision-making, from basing decisions on a single number to scenario-based decision-making. Instead of prolonging a decision because of the unknown, embrace the known unknowns and use it to construct mid/long term policies. In this way, undesired directions can be avoided in an early stage, which is a major improvement considering the current way of dealing with uncertainty within NS. Rather than accepting that there are many future unknowns, this research will contribute to dealing with uncertainties regarding the future development of the NS train fleet, the evolution of personnel availability, and the uncertainty regarding maintenance task duration.

The aim is to perform exemplary research where data is used to conduct a scenario analysis of a complex adaptive system. This is especially useful for large organizations that are providing a vital service. It is highly relevant for society to preserve the well-organized train system in the future, given that it is a vital service, which is a challenge that this research would aim to contribute to. While this research focuses on the train maintenance system of NS specifically, the approach held in this research could be used in a broader picture. In essence, any system that is dealing with deep uncertainty could use the scenario-based approach to have a picture of what lies ahead. In those cases, exploring many possible future scenarios, supported by data, will be extremely valuable. It not only creates a clearer image of what an organization might expect to happen, it changes the conservative line of reasoning that is behind large decisions. Decision making quality could be improved by adopting the approach that will be held in this research, allowing for well-performing, robust policies.

3.2. Methods

Methods used to perform adequate research will be both qualitative and quantitative. Interviews will be held, combined with desk research and on-site observance, to gain a better understanding of the NS train maintenance system and the uncertainty that NS faces in the next 10-15 years. Next, this knowledge will be translated into a model. This requires quantitative methods, where NS data needs to be acquired to set up the model. By following a research flow the research will be performed in a structured way, where qualitative work precedes the quantitative modelling part. To determine the current decision-making process regarding train maintenance, experts will be interviewed (part of confidential appendices). From their contribution, requirements for the simulation model can be identified, which will form the input of the conceptualization of the NS train maintenance system.

In the next phase of this research, a conceptual model of the NS train maintenance system will be created. Research methods to gather data on the NS train maintenance system are interviews, desk research & on-site observations. Interviews will enable creating a better picture of the relations and inner workings of the NS train maintenance system. Desk research allows gathering relevant data from NS. On-site observations contribute to the understanding of the real world, which will eventually be translated into a simplified version: the simulation model.

3.2.1. Agent-based modelling in AnyLogic

The execution phase of this research consists of building an agent-based model. The agent-based model will be built in AnyLogic, which will be based on knowledge from the conceptualization phase and data that NS possesses. The translation to an agent-based model will be done in the simulation tool AnyLogic, which allows for a friendly way of creating, simulating, and communicating an agent-based model. One of the main benefits of using AnyLogic is the possibility of visualizing the simulation model, increasing the persuasiveness of the model towards a DM. AnyLogic was used in the research of Alexandrov et al. (2019) on maintenance of rolling stock (trains), as well as in the research of Osman (2012) on simulating infrastructure asset management, showing the appropriateness of using AnyLogic as a tool in this context. Based on the agent-based model, there will be a more in-depth view on how, when and where maintenance could be (re-)organized, and what bottlenecks can be identified in the maintenance process. Moreover, the advantage of using AnyLogic is its compatibility with other models that are currently in possession of NS or with simulation models that will be created in the future.

Once the maintenance model is formalized, an iterative process will start where the model is verified and validated. The translation from conceptual model to the agent-based model will be verified by experts. Necessary improvements will be made, after which test runs will be done, generating the first raw output. The agent-based model output will be validated together with historical data and expert knowledge.

3.2.2. Scenario Discovery

During the final phase of this research, the model will be used to perform scenario analysis. Possible future scenarios of the next decade (until 2035) need to be constructed to perform experiments. The output from the scenario analysis will enable us to gain insight into the effect of deep uncertainty. To do so, the model output will be analyzed with the help of the Exploratory Modelling Analysis (EMA) Workbench (Bankes, 1993). The EMA workbench allows for the exploration of uncertainties throughout the range of future scenarios, by evaluating the range of plausible future dynamic developments of the NS train maintenance system (Kwakkel & Pruyt, 2013). The insights gained by EMA will thus allow for the formulation of policy recommendations towards a DM. In this way, the objective of this research, improving the decision-making information quality to support a DM, will be reached.

3.2.3. Data

Once a simulation model has been built, data will be required to configure the model, construct the scenario space, and perform experiments. The XLRM framework poses the external factors that will influence the NS maintenance system in any type of way. Therefore, it is necessary to collect data on each of the external factors. In addition, data will be needed to estimate correct probability distributions for processes have stochastic frequency or duration. The focus of the data collection part should lie on both the external factors (E) and relationships within the system (R) to construct a valid model.

The scenario space will then be constructed once the model formalized, verified, and validated. Data regarding future expectations will be necessary to construct the base-case scenario, which is the scenario that NS is currently expecting to occur during the next 10-15 years. By varying the external factors, the scenario space will be constructed. Tweaking the policy levers will then allow for an exploration of different policies under a large variety of future scenarios. This output is then analysed, by which policies can be compared.

3.2.4. Plausible scenarios

While conducting scenario analysis, this research will have a clear distinction between possible and plausible scenarios. All possible future scenario's could be explored, but this might take extensive computational power. Instead, plausible directions of where the future might head for reduces the number of experiments that will be necessary for EMA. This does not mean that only likely scenarios will be tested, because then it would be impossible to evaluate policy robustness. A DM of the NS maintenance system should be prepared for less likely scenarios as well. In addition, "plausibility does not require the explicit assignment of probabilities" (Wiek, Keeler, Schweizer, & Lang, 2013). Rather than testing only likely scenarios (while at the same time avoiding having to test all scenarios), plausible futures will be constructed by means of the EMA Workbench. Chapter 7 will elaborate more on how scenario's that fit the scope of this research are constructed.

4

Decision-Making Needs

This chapter will be dedicated to answering the first sub question: *What are the needs of a decision maker of the NS train maintenance system to make robust decisions?*. To make robust decisions on any system, a decision maker will need information on system performance, future developments within the system, how to deal with uncertainties, and what strategies can be applied. The XLRM framework presents an overview on these topics in the context of train maintenance.

4.1. XLRM framework for train maintenance

To communicate the understanding of the NS train maintenance system to experts, and to receive their input, the XLRM framework is used (Wong, Srikrishnan, Hadka, & Keller, 2017). The framework creates an uncluttered picture of what needs to be included within the scope of this research. This enables the correct translation to a conceptual model, and eventually a simulation model. The XLRM framework poses a quick overview of what this research will include, and allows for a quick understanding of those that are rather unfamiliar with simulation modelling. The following elements are present in the XLRM framework:

- E**X**ternal factors: the factors that are not influenced by a DM of the NS train maintenance system (considering the research scope), but do affect the system in a certain way
- Policy **L**evers: the fictitious buttons that a DM can press, they form the strategies that a DM can apply
- **R**elationships: the inner workings of the NS train maintenance system, formed by both e**X**ternal factors and policy **L**evers
- Performance **M**etrics: important metrics that a DM uses to base its decisions on, to assess the desirability of future scenarios, or to evaluate the success of an implemented policy

Filling in the XLRM framework for the NS train maintenance system yields the following result, see figure 4.1.

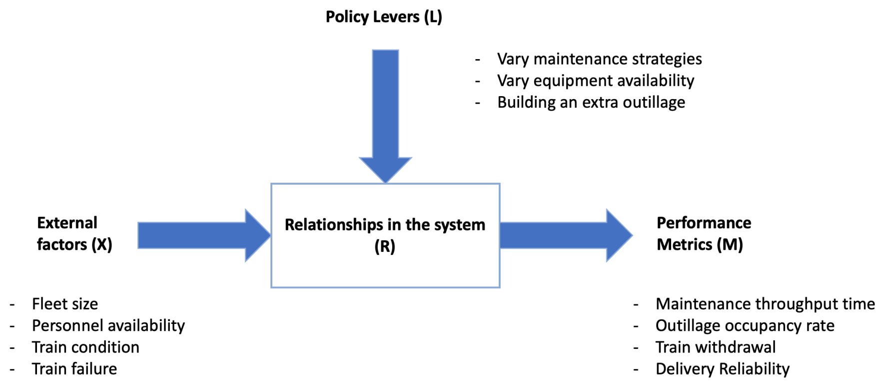


Figure 4.1: Filled out XLRM framework in the context of the NS train maintenance system

Table 4.1 elaborates on the description of each element and how it is related to train maintenance for NS. The XLRM framework poses a clear overview of the scope of this research. For example, fleet size is not an external factor for NS, but it is an external factor when considering the research scope. The fact that there is a whole process happening before the future fleet size is determined, and that people within the organization are deciding about it, is out of scope.

The XLRM framework also poses an overview of the policy levers that are available for NS, together with the metrics that are important for a decision maker on train maintenance. The performance metrics will eventually tell the decision maker how a policy is performing, which creates a concise overview. This improves current methods that are being used to perform policy analysis. Currently, these metrics are separately used in the evaluation of the success different policies. It would be beneficial to compare shifts in the performance metrics while adjusting the policy levers. This would be a major improvement in the context of policy analysis within NS, and is said to be more valuable then optimizing them separately. For instance, when minimizing the withdrawal percentage of the train fleet, the precise effect on other performance metrics is unknown. Let alone the effect of a single policy throughout the range of plausible scenarios that NS may face.

Element	Application to NS train maintenance
X	**Fleet size**: current as well as future fleet size of NS. **Personnel availability**: the availability of mechanics. This number is shifting heavily due to sickness, retirement and overall shortage on the current labor market (Verbeek, n.d.). **Train condition**: the condition of the train determines the amount of work to be put in during maintenance. Their overall condition can vary heavily and is seen as an external factor within the scope of this research. **Train failure**: the reliability of new trains that are yet to be acquired is an unknown external factor. In addition, new safety systems such as ERTMS will be implemented, while their reliability is yet unknown. It could affect the maintenance system substantially in case of unexpected failures. In addition, trains at the end of their lifetime might show higher frequency of train failures.
L	**Maintenance strategies**: various train maintenance strategies can be applied by NS, such as condition-based maintenance, preventive maintenance, predictive maintenance, or corrective maintenance. **Vary equipment availability**: equipment used to perform maintenance on a train is currently only available at a 'home ML', while it can be decided to provide several maintenance locations with different types of equipment and tools. This would enable performing maintenance on multiple train types within 1 ML. **Building an extra outillage**: in case the current capacity is too limited to keep maintaining the train fleet without congestion, an additional outillage could be built.
R	Agent based simulation model of the NS train maintenance system, consisting of Trains, Mechanics, Maintenance Locations & Outillage. Their interaction forms the system relations.
M	**Maintenance throughput time**: the time it takes for a train to complete scheduled maintenance. **Train Withdrawal**: to measure the performance of train maintenance, NS uses withdrawal percentages. These numbers indicate the share of trains that are withdrawn from operation due to maintenance. **Outillage Occupancy Rate**: A rate that keeps track of the share of a ML's capacity that is occupied. **Delivery Reliability**: The share of trains that are delivered on time back to operation after having performed maintenance.

Table 4.1: XLRM elements and their descriptions in the NS train maintenance context

4.1.1. Performance Metrics

To keep track of the system performance is what can be identified as highly important for a decision maker on the train maintenance system of NS. It helps the decision maker to make correct judgments about whether or not to perform a certain policy, such as building a new location facility, building additional equipment in a ML, or performing a different maintenance strategy. There are multiple performance metrics that are key, and their importance will each be discussed separately.

Maintenance Throughput Time

During maintenance, trains start and end their scheduled maintenance at some point in time. This time is collected for each train separately. From the time it takes for a train to complete scheduled maintenance it can be identified whether congestion has occurred. Congestion indicates under-performance of the maintenance system, which is why this has been identified to be an important KPI to monitor.

Train Withdrawal

As mentioned in table 4.1, the train withdrawal indicates the number of trains that are withdrawn from operation due to maintenance. While a train is in maintenance, it cannot be used in the primary service of NS: transporting customers. Ideally, NS wish to operate their trains as much as possible. The current demand for train trips does not permit them to do so, but when they are operating at full capacity, there are still some trains remaining that will have to go into maintenance. This is why NS has to purchase a few extra trains (the precise number is based on the train type), so that when a train is due for maintenance, it can be replaced in operation. If the withdrawal of trains becomes higher, it means that more trains will be in maintenance, which are in that case not available in the process of transporting their customers.

The dependency between the operational capacity and train withdrawal is what makes the latter an important metric to keep track of for a decision maker of NS. If the maintenance process slows down, or if failure rates of trains go up, the fraction of time that trains are undergoing some type of maintenance increases. This strongly affects the operation of NS, and should thus always be monitored. Hence, the needs of a decision maker are partially fulfilled if this number is included in the research outcomes.

Occupancy Rate

Throughout a simulation run, the MLs are occupied to a certain degree, which is represented by the occupancy rate. This information tells a decision maker to what extend the ML capacity is used, which is relevant when exploring the system boundaries. It also shows whether a certain ML has overcapacity: it maintains less trains then it could or should. To evaluate whether additional capacity would be necessary in the near future, the occupancy rate can be an important metric. In case the train fleet size increases (being an external factor), occupancy rates will show whether the maintenance system is able to handle an increased workload.

The occupancy rate is thus an important metric that allows making an informed decision on whether or not to adjust the maintenance capacity. Adjusting the maintenance capacity requires looking years ahead, as increasing capacity has a significant delay. It takes about 2-3 years to train a mechanic for performing maintenance, and it takes about 8-12 years to build a new maintenance facility.

Delivery Reliability

The metric that shows the punctuality of the maintenance system is the delivery reliability. It represents the share of trains that are delivered on time back to operation, after having performed maintenance. This metric is deemed important since it directly affects the train tables. Every time a train is delivered late, the train tables have to be adjusted. It requires additional planning and intense collaboration with schedule developers and operational teams. Therefore, a high delivery reliability is desired. In turn, this pressures the maintenance location to perform their tasks within a given time-frame.

While maintenance locations strive to deliver the train back to operation on time, setbacks might occur. External factors such as personnel unavailability, sickness, or unexpected complex errors might take maintenance longer than anticipated for. In that case, a train is not delivered on time. For a decision maker of NS, it is relevant to know whether a policy intervention affects the delivery reliability or not.

4.2. Model-based insights

Creating useful insights on the effect of different maintenance strategies (or other policy levers) by models is deemed to be valuable according to NS decision makers. Being able to apply different strategies beforehand, and evaluating system behavior accordingly, will benefit the needs of an NS decision maker. The information provided by a simulation model can be used by a decision-maker to set up a plan to potentially change current approaches to train maintenance. Analysing the key performance indicators of different train maintenance policies under a large variety of future scenarios allows decision makers to deal with complexity resulting from deep uncertainty.

Instead of calculating the average number of trains that go into maintenance on a weekly basis, a scenario-thinking will be introduced. Keeping track of the performance metrics under a large variety of plausible future scenarios results in the identification robust policies. It is the task of a decision maker to make correct judgments, yet, EMA will make the life of a decision maker easier by providing tools in making correct judgments that lead to robust policies.

Currently it is unknown what responses of the maintenance system of NS to different strategies could be. Moreover, there is a lack of knowledge what the future will hold precisely: deep uncertainty is accompanied by ill understood system behavior. Therefore, NS is unable to deal with uncertainties, even though they are aware that dealing with them will benefit their maintenance system in the long run. Being more robust towards future scenarios is precisely what was pointed out to be of high value for decision makers. Awareness for possible divergence of future scenarios is what will be highly valuable when making large decisions regarding train maintenance. For example, temporary overcapacity in mechanics might ask for scaling down to avoid costs, while a few years later the capacity might be needed again. If DMs decide to scale down excessive maintenance capacity as soon as possible (reducing maintenance costs), they might create a problem of lacking the maintenance capacity in the future, in case the train fleet increases (while assuming that an increased train fleet requires more overall maintenance). Even though it might be temporarily more costly to keep maintenance capacity on the same level, on the long run it could prevent capacity issues that might turn out to be even more costly. A more robust decision would thus require making judgments that look further ahead, possibly even decades ahead. A simulation model precisely enables those needs.

Naturally, looking years or decades ahead comes with deep uncertainty, which is where EMA comes into play. As has been discussed in chapter 3, EMA allows the decision maker to evaluate, explore and analyse many different scenarios. When robust policies for train maintenance are desired, it is required to perform extensive exploration of many plausible future scenarios. Making robust decisions can then be done by comparing each policy based on their performance regarding withdrawal percentages, occupancy rates of maintenance locations, together with delivery reliability of trains.

4.2.1. Many objective vs single objective decision making

Current strategies on train maintenance concern optimizing one of the performance indicators of NS, without knowing the precise effect on other performance metrics on the long term. While maintenance DMs currently decide to perform a certain policy that fits with the optimization of a single performance metric, they would rather be able to make decisions based on a more comprehensive overview that takes into account multiple objectives. This should then lead to the implementation of robust policies that have been identified by EMA.

5

Conceptualization of the NS Train Maintenance System

To answer the second sub question *What does the NS train maintenance system entail?* this chapter will present a conceptual overview of the scope of this research and system ontologies. To benefit the understanding of relations within NS' train maintenance, narratives are constructed for each agent type in section 5.1. Thereafter, flowchart(s) are presented in section 5.2.

5.1. Narratives

Creating a better understanding of the components that will be the agents in this research, agent narratives are constructed. Within the scope of this research, the agents that will be discussed in terms of narratives are Trains, Mechanics, Maintenance Locations (MLs) & Outillage. Each train requires maintenance at one specific ML, where they have Mechanics and Outillage to their disposal. In why and how agents interact can be learned from this section.

5.1.1. Train

Each 24 hours, a train has a rather similar routine. During the night, trains are stalled, cleaned. In case some minor issues arise, they can be fixed during this time period as well. Depending on the schedule, a train is operating between 5 A.M. till midnight with a single task: transporting passengers from A to B. The main operation process goes on for a predetermined amount of days or kilometers (depending on the train type & whichever comes first). When the maximum allowed days in operation is exceeded, or when the kilometers has been reached, a train has to go to a workshop for scheduled maintenance. Each train type is maintained at one specific ML: their home ML. When they arrive at their home ML, they go through several maintenance phases: emptying the bio-toilet reservoir (BIO), performing maintenance (MTCE), cleaning of in- and outside (CLEAN), grinding of wheels (KWB), replacement of chassis (AW), final tests and checks (CHECK). These phases are currently divided by shifts of 8 hours, making them task blocks of fixed lengths. In reality, the scheduled time is more than the actual time it takes to perform the tasks, just to have some slack in case of setbacks. If the condition of a train allows it, KWB and AW are performed in one shift. While some steps are mandatory, other steps are dependent on the condition of the train as it enters scheduled maintenance (elaborated in section 5.2).

NS possesses many different train types, which require different maintenance. This becomes clear when looking further into the type of maintenance that comes along with performing scheduled maintenance. Important equipment such as air conditioners, electricity transformers, and air filters are revised each time. For new train types, these parts are placed on the roof of the train instead of the bottom of the train. This is a major shift, that requires different type of approaching a train for maintenance, and requires different equipment.

On top of scheduled maintenance, it could occur that a train 'breaks down'. It is then required to immediately repair the train, because most of the time the train schedules expect the train to be operating. The occurrence of failure disrupts the operation, since the train is unable to transport passengers. The train will be shunted to its home ML, where it is being repaired within 24 hours to ensure a minimum time of withdrawal.

5.1.2. Mechanic

A mechanic is part of a team. Team numbers and team size depend on the capacity of the workshop. Teams of mechanics are working around the clock, every day of the week, to make sure that the trains they maintain are safely sent back into operation conform safety standards. The teams operate in shifts of 8 hours in either a night shift, early shift, or late shift. A mechanic works at 1 workshop, on 1 type of maintenance (either pitstop or scheduled maintenance), and works on 1 specific task at a time, such as emptying the toilet reservoir or replacing a chassis. Through qualifications, distinctions are made in assigning mechanics to different types of tasks. However, to prevent over-complication, the qualifications of mechanics is assumed to be out of scope of this research.

During their shift, mechanics are assigned to 1 specific maintenance task. When they take 5 hours to finish a task that was scheduled for 8 hours, they are essentially free for 3 hours. If possible they are assigned to another task, but if there are no minor tasks available they are essentially free of work during the remaining time.

In some cases, a train fails while not being close to its home workshop. When a train fails in another part of The Netherlands and cannot be safely transported to its home workshop, mechanics will travel to the train instead. They have the expertise and tools required to perform a preliminary fix, enabling safe movement to the train's home workshop. This aspect however is not included in the scope of this research. This research solely focuses on the repairing of trains within the MLs. It is therefore assumed that when a train failure occurs, it can always be shunted to its home ML. Hence, mechanics will not leave their ML according to the research scope.

5.1.3. Maintenance Location

NS owns and uses four Maintenance Locations (MLs) in The Netherlands, them being Onnen, Watergraafsmeer, Leidschendam and Maastricht. Each ML has a limited amount of maintenance rail tracks where maintenance can be done, and a limited amount of mechanics that can perform maintenance. The Maintenance Location provides the ability to perform maintenance 24/7. It houses train parts that can be accessed in case parts are broken and are required to be replaced.

Each time a train enters the ML, it fulfills maintenance tasks. The number of tasks to be completed depends on train's condition. Is the job of the ML to enable all tasks to be completed within a certain period of time. Planners within the ML assign train maintenance requests to available outillage and mechanics. This requires a high level of flexibility, since trains might be delayed while being moved to their ML. Delayed trains require adequate adaptation from the planners, since their original schedules need to be revised. It is therefore assumed that train maintenance requests are handled instantly within the availability as the maintenance request is issued. In case the maintenance request cannot be handled instantly, the train will be placed in a waiting cue.

Different MLs have deviating capacities, which can elaborately be observed in table 6.4.

5.1.4. Outillage

Outillage concerns train tracks that are part of a Maintenance Location. Outillage capacity is what partially (together with mechanics) makes up the capacity of MLs. The outillage rail tracks allow for a train to be stalled while it is undergoing maintenance, allows mechanics to perform specific types of maintenance, and in some cases allows mechanics to move freely below the train. Considering the track length varies from 110-200 meters, in some cases two trains can fit at one maintenance rail track, allowing for maintenance to be performed simultaneously on two trains. Some of those tracks are equipped with elevated working platforms. An elevated working platform enables safe and relatively easier maintenance of components that have been placed on the train roof, as it allows mechanics to walk around the train roof and stepping on and off the train roof. This makes the roof of a train more accessible for a mechanic, in contrast to working with portable aerial working platforms. Not all rail tracks of NS workshops have elevated working platforms. In those cases, portable aerial working platforms must be used to perform maintenance on altitude due to the lack of elevated working platform capacity.

5.2. Flowchart

Now that narratives of the NS train maintenance system are understood, they can be displayed visually in a flowchart. A flowchart benefits the understanding of the system's inner workings, what stochastics are present, and where decisions are made within the maintenance process. Flowcharts for a train and the flow within a maintenance process are presented, together with the flowchart of mechanics.

5.2.1. Train flowchart

In this section, the general flowchart of a train entering maintenance is portrayed. Figure 5.1 shows the current process of any train of NS that is shunted to their home workshop.

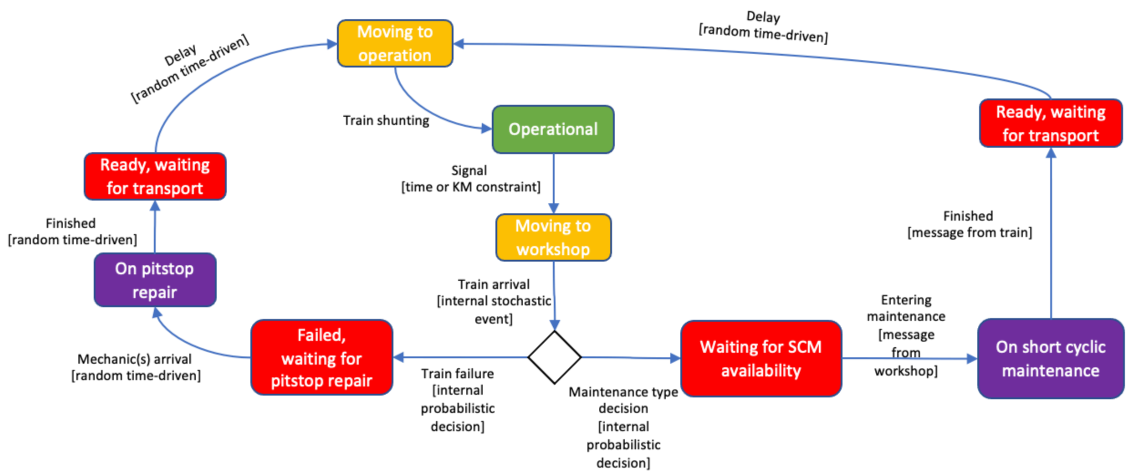


Figure 5.1: General flowchart of current train maintenance at an NS workshop

Figure 5.1 presents the possible maintenance flows that a train can go through. When a train is transporting passengers, it is operational. As soon as it is time for maintenance (based on either the maximum allowed kilometers or the days that have passed since maintenance was done), a train is shunted to its home ML. A train can also be shunted to its home ML in case a spontaneous failure occurs. The yellow boxes represent moving trains towards or from its home workshop. When a train arrives at the ML, it could be for maintenance cyclic maintenance, or to fix the failure which is called a pitstop. The moment the train arrives and the moment mechanics start working on the train are separated, since mechanics might not be available or the maintenance shop has no available outillage yet. The red boxes indicate a waiting, idle train. Once there are mechanics available and there is an empty spot in the workshop's outillage, maintenance/repair is performed. After the maintenance/repair, the train is ready to be shunted back into operation, but has to wait on available train drivers. As soon as the train drivers are ready to move the train, it is shunted back into operation.

The purple 'on short cyclic maintenance' block will now be further specified into separate tasks.

Workflow within maintenance

As soon as a train has entered the maintenance process it sends out a maintenance request, requiring mechanics and outillage to be available. In the theoretical maintenance schedule these tasks have a specific sequence, devised by maintenance planners. In reality, there is room for flexibility in the variation of the order in which tasks are performed. Figure 5.2 depicts the necessary sequence together with the possible variations.

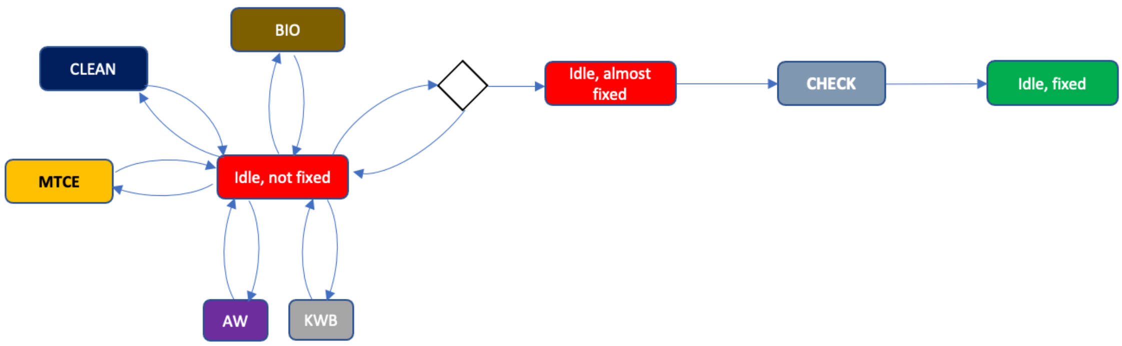


Figure 5.2: Scheduled maintenance flowchart

To gather information on the current state of the train, a pre-check is performed where the train's condition is determined. It then enters the 'Idle, not fixed' state where the train waits to be assigned. Next, regular maintenance (MTCE), emptying the bio-toilet reservoir (BIO), replacing the chassis (AW), cleaning the train exterior and interior (CLEAN), and grinding the wheels (KWB) can be performed independently of each other. Whichever comes first will be decided based on the availability of the workshop and the on-site mechanics. AW & KWB are done conditionally, BIO/MTCE/CLEAN are always performed. After these tasks have been completed, the train reaches the 'Idle, almost fixed' state. It then has to be checked before being returned to operation. The checking (CHECK state) requires outillage and mechanics to be available too.

In reality, cleaning is done at the final part of the maintenance process it to prevent a situation in which mechanics have to work inside a cleaned train. This research assumes that the only sequentially bound task is CHECK. Once final checks are performed the train reaches the 'idle, fixed' state, and becomes ready to be shunted back into operation.

5.2.2. Crew flowchart

The tasks within the sequential maintenance flowchart of figure 5.2 are performed by mechanics that are part of the maintenance crew. Each task takes either 1 or 2 shifts for a maintenance crew. In some occasions, a task might take less time than calculated, e.g. the crew might be finished after 5 hours with performing KWB, while 8 hours were scheduled. As displayed in figure 5.3 crew can thus be working on maintenance, working on pitstop repair, or be on stand-by (in which they are idle). This research assumes that only mechanics qualified to perform pitstop repairs are able to do so. There is a split between mechanics that perform maintenance tasks (regular mechanics), and mechanics that perform both scheduled maintenance tasks and pitstop repairs (pitstop).

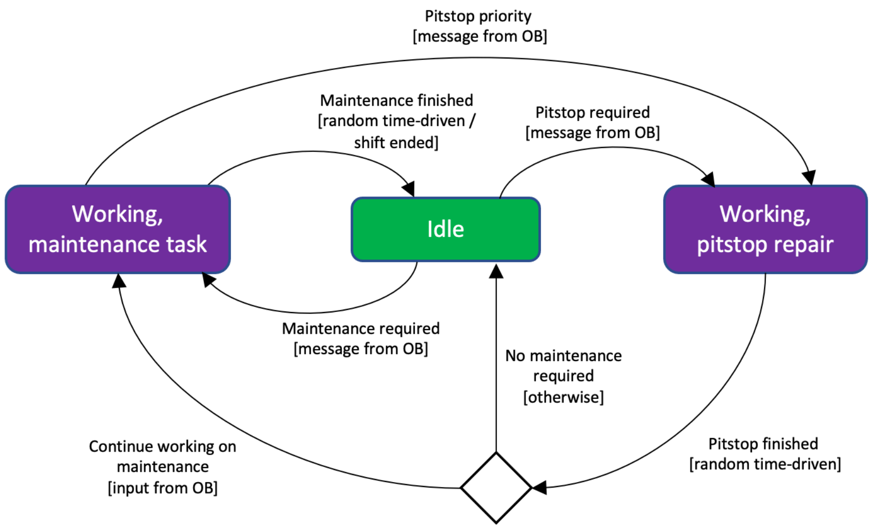


Figure 5.3: Flowchart for the maintenance crew

Purple states indicate a working crew, and the green state is idle. When in idle state, the crew can either receive a message from the ML that there is a train to be maintained, or they can receive a message that pitstop repair is required. While working on a maintenance task, the crew can either go back to an idle state when their task is completed or when their shift ends, or they might be required to perform pitstop repair that has priority. When a pitstop repair is finished (driven by a time drawn from a probability distribution), the crew decides what to do: go back to the maintenance task they were performing initially (in case the ML requires them to do so), or move back to an idle state (in case there are no maintenance tasks left to perform during their shift).

5.2.3. Outillage

Outillage can be either available or in use. Both states are dependent on the arrival of trains and their demand for maintenance. Some types of outillage are used more frequently than others, given that a maintenance cycle has a few optional and a few obligatory components. Trains will always pay a visit to the train track where regular maintenance is done (MTCE), the BIO track where the toilet reservoir is emptied, and to the train washing installation (CLEAN) where the train is washed on the outside. Some MTCE tracks are train-specific, which indicates that only trains of that specific train type can be maintained there. This could be the case when there are elevated platforms needed to access the train roof for example. Optional outillage are the wheel grinding installation called "kuilwielenbank" (KWB) and the chassis removal section called "aardwind" (AW). Current policy is that trains of the type DDZ, VIRM and ICM always go to KWB during scheduled maintenance. After all previously mentioned steps have been completed, the train will has to be checked before it can be returned to operation. This happens at regular maintenance tracks (again of which some are train-specific). Besides being in use or being available, the outillage doesn't have any additional function within the scope of this research.

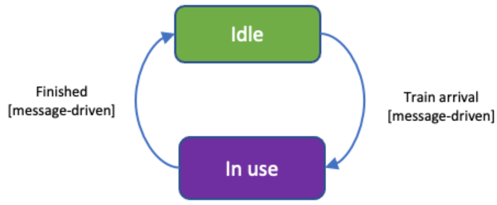


Figure 5.4: Flowchart for Outillage

6

Train Maintenance Model Formalization

This chapter will present the formalization of the simulation model that is based on the conceptual model, aiming to answer the third subquestion ”*What model(s) can be built to support the decision makers of the NS train maintenance system?*”. First, the model setup will be discussed. Next, the different agents will be presented, together with the way they interact. After explaining how agents are distinguished, their behavior is presented in the model inner workings. Algorithms that form the basis of the simulation model are portrayed in pseudo-code. The controls that allow a user to adjust the model settings before and during a run are lastly elaborated on.

6.1. Agents

This section will elaborate on the different types of agents that are included in the simulation model that has been set up in AnyLogic (Appendix A). Figure 6.1 portrays the composition of agent levels in AnyLogic. AnyLogic requires to have one top-level agent, which is essentially a 'moderator' that oversees/manages the environment and interaction of other agents. This agent is referred to as 'Main'. Agents make use of the environment within main to interact. For NS, main could be seen as a control center where maintenance schedules are created, monitored and executed.

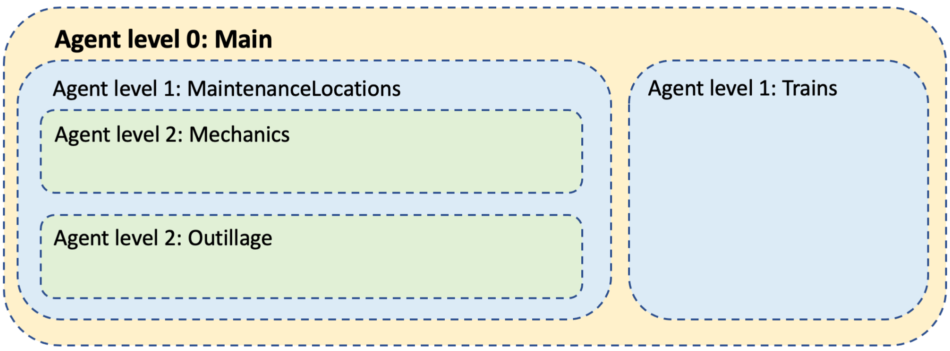


Figure 6.1: Overview of the different agent levels within AnyLogic

Main consists of MaintenanceLocations and Trains. Trains not only require maintenance at a specific MaintenanceLocation, they also require available Outillage & available Mechanics. These belong to the MaintenanceLocation, and are modelled as different agent types.

6.1.1. Main

Within main, two agents are present: Trains and MaintenanceLocations. These are the trains of NS that are operational, together with the maintenance locations where they are maintained periodically. When a simulation is running, the main agent screen offers an overview of statistics that are displayed in histograms, bar charts and time plots. Time plots are essentially line plots that display a certain value during a certain timestep. In the model developed for NS, histograms portray the distribution of waiting times and total throughput times. A time plot is used to display the number of trains that are operational at each timestep. There is always a fraction of trains undergoing maintenance at its home maintenance location at any given time, which can be tracked during a simulation run by checking the timeplot. For those that are less familiar with model specifics, these time plots are convenient to have an instant overview of the system behavior. In case model settings are tweaked, the plots allow for checking the corresponding behavior of the train maintenance system. Plots also allow for an easy and smooth communication of model outcomes towards those that are not familiar with simulation. Therefore, it is main important that the interface of the main agent is well-organized.

Communication between Trains and MaintenanceLocations goes through the Main agent. The Main agent assigns the Train to the corresponding MaintenanceLocation in case a train asks for maintenance or a pitstop repair. The main agent should thus be considered to be a control centre, like the control centers that NS has in reality that are constantly playing into daily changes.

6.1.2. Trains

In the Train agent section, model specifics are determined for the trains of NS. The behavior of 1 train is modelled, and different properties are loaded from a small database in the AnyLogic simulation model. The database enables assigning characteristics of different train series, that each have different home Maintenance Locations, referred to as HomeOB in the model.

When they are due (based on distance travelled between maintenance jobs or on time since last maintenance), trains will have to undergo cyclic maintenance. When a train arrives at a maintenance location, their condition is determined. The duration of maintenance is dependent on the condition of the train. For example, some trains might need to go to the KWB (Kuilwielenbank), while others don't. Each block of maintenance is split into different requests. All trains that undergo maintenance will have to finish steps: 'BIO' 'MTCE' 'CLEAN'. In case their condition is bad, they have to pay a visit to the KWB to polish its wheels. If their condition is even worse, they have to go to the AW (Aardwind), which is the station where the wheels of carriages are replaced. Thus, the condition of the train affects the throughput time of a maintenance visit. The total throughput time of the train is measured, and reported to the Main agent.

When it is time for maintenance, trains will send out maintenance requests. Trains communicate these requests to the top level agent (Main), who forwards this message to the specific maintenance location that corresponds with the Home ML of the train requesting maintenance. Based on the availability of both Mechanics and Outillage within that maintenance location, a train is placed at an available station. This will be further elaborated on in the next subsection. It happens regularly that a train arrives at a maintenance location while the maintenance location is fully occupied. The train will then have to wait for available outillage, before it can start a maintenance task. This waiting time is also measured and reported to the Main agent. These statistics are helpful when analyzing the effect of policies such as upscaling outillage capacity or staff. Suppose one decides to upscale outillage, but there is not enough staff to control the outillage, there will be no chance of using the upscaled availability. Trains would still be waiting, since the MaintenanceLocation needs to assign both outillage as well as available staff to a train before it can start its maintenance task. Tracking the waiting time statistic may thus provide valuable information on the effect of applying policy levers.

6.1.3. Maintenance Locations

There are four maintenance locations of NS: Leidschendam, Maastricht, Onnen and Watergraafsmeer. Each maintenance location handles fixed types of trains. The precise capacity of each ML will be discussed later on in this chapter.

For each outillage type, a time color chart sequentially checks whether they are available or not. Green indicates that the outillage is available, and red indicates unavailable outillage. This allows a user of the simulation model to observe the occupancy of the different outillage types. A lot of red would indicate an overly used outillage type, whereas a lot of green indicates that the specific outillage isn't required much.

6.1.4. Outillage

Different types of outillage are distinguished within the NS maintenance process. Each train has to visit the BIO, CLEAN and MTCE. Based on the condition it also visits KWB and/or AW. These types of outillage can be either available (Idle) or occupied (InUse).

When the "START OUT" message is received, the outillage is set to InUse, becoming occupied. A train is assigned to the specific outillage, and when it is finished the outillage receives the message that it becomes available again. The agent type outillage does not have any other functions other than being available or unavailable. Besides receiving messages from other agents and switching from Idle to InUse it doesn't have any function in the model. This is in line with reality, where outillage simply is a piece of equipment, a train track or tools, that are being used to perform maintenance.

Whenever a train arrives at a type of outillage, a random number is drawn from a outillage-specific triangular distribution. The outcome from drawing the random number then determines the amount of time it takes for mechanics to perform the maintenance task, which can directly be seen as the amount of time that the outillage is unavailable for new maintenance requests. As soon as the train leaves the outillage, the outillage becomes available for new maintenance requests.

BIO

Emptying the bio reservoir requires specific outillage, and is done on BIO-specific tracks. The process is standardized and takes about 3-4 hours to be completed (in case no setbacks occur). It requires a mechanic that has the knowledge and formal certification to perform the tasks that this maintenance step demand. Leidschendam is the only ML that has a capacity of two BIO tracks, the other MLs have only 1 BIO track.

CLEAN

The cleaning of the train is essentially split into two separate jobs: inside and outside. The outside cleaning process is similar to that of a car-wash. The train is shunted to the train washing installation (CLEAN), which takes about 30 minutes to complete. Within this research, it is assumed that during the CLEAN maintenance step, inside cleaning is also done. The inside cleaning is not performed by maintenance mechanics, but it is outsourced. However, this step does require a track at the maintenance location to be available.

MTCE tracks

At all MLs, there are multiple MTCE tracks. These are tracks where solely regular maintenance is performed (not EBKs). This can be either the MTCE step in the simulation, or the Nawerk step in the simulation. The MTCE step is a fixed work-package that takes about two shifts (16hrs). Scheduled maintenance is being performed, together with fixing/replacing based on visual inspections. Some MTCE tracks have elevated platforms, because they are designed to maintain a specific train type where a lot of the equipment is placed on the roof. MTCE tracks also allow mechanics to go under the train to perform maintenance below the train. In reality, the exact duration of this step is highly uncertain and depends on a lot of factors, such as train size, train condition, number of mechanics working on it, and whether track has the presence of elevated platforms. Tracks that don't have elevated platforms, require mechanics to use machines that lifts them in a bucket towards the train roof. Mechanics will then have to perform maintenance from this bucket, essentially hanging above the train. This is more time consuming than working with elevated platforms where mechanics can walk around freely throughout the length of the train roof.

Nawerk is also performed at the MTCE tracks. During this step, mechanics perform final fixes and check-ups to ensure that the train can be sent back to operation safely, until it is due for maintenance again. This requires MTCE tracks because mechanics might need to access the bottom/roof of the train.

KWB

When wheels are not perfectly round, they start bouncing on the train track while the train moving. This can become uncomfortable for customers and might also damage other parts of the train. Sensors placed inside the train tracks measure vibrations of a train as it is passing by, providing the control centre of NS valuable information regarding the condition of the passing train. This data is processed and shared throughout various data platforms. This is why trains are categorised in several conditions by the control centre, as they are constantly being monitored in the software systems of NS. In this research four train conditions are used as they enter maintenance: worse, bad, average & good. Once in a while a train slips on the track, for example due to leaves that lie on the train track, or due to heavy breaking during icy weather conditions. The slipping results in a flat part on the wheel. Thus, occasionally the wheels need to be polished. A small part of the wheel then needs to be peeled off, so that the wheel is perfectly round again. This is done on the kuilwielenbank (KWB). Execution of polishing with the help of the KWB requires (at least) two mechanics that have specific knowledge about handling this outillage; an experienced qualified mechanic and a qualified assistant.

AW

At the AW, a chassis can be detached from the train, for example to replace wheels that have reached the end of their lifetime, or perhaps to replace malfunctioning electricity transformers. In case no replacement is needed of a specific part, but something needs to be repaired, this outillage could also be used. It does not happen often that trains will have to visit this specific outillage, but when they do it is usually a time-consuming task. The chassis a complex, heavy and expensive part of the train, and should thus be handled with great care. Performing maintenance of this part also requires specific knowledge. As goes for the KWB outillage, the condition of the train determines whether it has to pay the AW a visit. Likewise, two mechanics are required to perform maintenance tasks at the AW; an experienced qualified mechanic and a qualified assistant.

6.1.5. Mechanics

Two different types of mechanics are distinguished in the model: regular or pitstop. Regular mechanics are those that are not allowed to perform pitstop repairs, while pitstop mechanics are. In reality, in case there is no pitstop request, all MTCE mechanics are busy, and a new maintenance request arrives, pitstop mechanics will help MTCE mechanics. Because it became to complex and harsh to have pitstop mechanics taking complete responsibility of a maintenance task, the helping of pitstop mechanics is not included in this research.

As goes for Outillage, mechanics too can be in two states: idle or working on a maintenance task. In case a mechanic is of the type 'pitstop', it can only be drawn into the working on pitstop state. Each time a pitstop mechanic agent is finished with a pitstop repair, it returns back to the idle state, where it waits to be assigned to a new train that requires either pitstop repair. 'Regular' mechanics are mechanics that can be drawn into the different train maintenance requests that have been sent out to the control centre. There is a finite set of mechanics available at each maintenance location, which is determined by the amount of work to be done on a weekly basis. For this research, the exact calculation of the required capacity is out of scope, but it results in the following capacities for the MLs displayed in table 6.1.

Maintenance Location	Number of mechanics
Leidschendam	125
Maastricht	100
Onnen	153
Watergraafsmeer	95

Table 6.1: Total number of mechanics available at each Maintenance Location

Because most tasks require multiple mechanics, it is assumed that each task in the simulation is performed by two mechanics at a time. Therefore, the simulation works with mechanic duo's that maintain 1 specific train at a time. Due to its complexity, qualifications of mechanics is not taken into account when assigning mechanics to train maintenance requests. Mechanics will never be assigned to the CLEAN task, because the inside cleaning is being outsourced.

While in reality pitstop mechanics might decide to give priority to a train that undergoes regular maintenance so that it can be sent back to operation as soon as possible (instead of handling an EBK), this does not happen in the simulation. In AnyLogic, there is a clear distinction between mechanics that perform regular maintenance tasks, like BIO/KWB/AW/MTCE/Nawerk, and those that perform EBKs. This is an assumption, because in real life 'pitstop' mechanics might make the decision to first finish a regular maintenance task before starting with the pitstop repair. Several factors could be of influence: remaining maintenance time, but also the requirement of different types of trains in operation. In case the pitstop request comes from a train that is highly required in operation, it would make more sense to immediately start with the pitstop repair, whereas in other cases the pitstop might not require immediate action. It should be kept in mind though that pitstops are aimed to be finished within 24 hours of the request.

Initially the model was set up in a way that pitstop mechanics would pause train maintenance in case an EBK was required. Model runs with this setting have been tested, where pitstop mechanics too help in performing regular maintenance, and would set regular trains on hold as soon as a pitstop request is issued (due to pitstop priority). When the pitstop repair is finished, the pitstop crew would resume the work they were doing on the train that had been set on hold. However, trains that had been set on hold would not always be 'found' by other available mechanics in the simulation model, leading to erroneous output. Trains would get stuck being on hold, or pitstop crew would search for a train that had already been finished by another crew. This implementation raised many errors, and was thus removed from the simulation model. Therefore, the decision has been made to keep the two processes apart from each other; 'regular' mechanics are only assigned to the scheduled maintenance tasked and 'pitstop' mechanics are solely handling EBKs where pitstop repairs are required.

Mechanic employability
The teams of mechanics are scheduled to work around the clock. It is assumed that each ML operates 24/7. Mechanics operate in shifts of 8 hours. For all MLs, there are a total of 5 teams that alternate each other. Suppose a ML has a total of 105 mechanics assigned to them, each team would consist of 21 mechanics. In reality however, there is an employability factor that scales down the team size that is employable compared to a theoretical team size. Experts of the train maintenance system have set this number to 0.667 (2/3 of the team is employable, 1/3 is either on sick leave or is on holiday). This number comes from their the experience with mechanics not being available to work. The type of tasks they perform on a daily basis is highly demanding on their physique, which results in a relatively large share of employees that are not able to work. The following formula is used to determine the number of mechanic agents in the simulation.

$$A_{mechanics} = \frac{T_{mechanics}}{N_{teams}} \cdot f_{employability}$$

Where:

$A_{mechanics}$:	Number of available mechanics in the simulation model
$T_{mechanics}$:	Total number of available mechanics
N_{teams}:	Total number of teams
$f_{employability}$:	Employability factor (0.667)

So in case of the example of a total of 100 assigned mechanics, there is a discrepancy in the theoretical team size (20), and the actual team size (14).

$$M_{agents} = \frac{A_{mechanics}}{2}$$

It should be kept in mind that 1 mechanic agent in the simulation is a duo of mechanics in reality. The number of mechanic agents in the simulation model is calculated by dividing the actual team size by 2, as observed in the formula above.

6.2. Building the model

For each agent type, a small database is used to build the model. For trains, the database contains data on different series, the fleet size (population size) of each series, the home OB of each series, the amount of KM between two maintenance jobs (varies per series), and how much time there should be between two maintenance jobs. AnyLogic loads the data starting up the model, and then automatically assigns the data to the different train series as specified in the database.

Setting up Maintenance Location agents

For Maintenance Locations, the database only contains the name of each location and two booleans: sprinter? & intercity?. Each maintenance location has a 'name', which is based on the string that is in the database, which allows the modeler to make distinctions in the four MLs. In addition, a ML is labeled as a sprinter ML, as a intercity ML, or both (Maastricht). In case the boolean is true, it means that this ML is able to repair those train types.

Type	Sprinter?	Intercity?
Leidschendam	true	false
Maastricht	true	true
Onnen	false	true
Watergraafsmeer	false	true

Table 6.2: Setting up the four Maintenance Location agents

Setting up mechanic agents

For mechanics, the database contains the size of the population and the mechanic type (maintenance crew or pitstop crew), as displayed in table 6.3. As can be observed, team size varies per maintenance location based on the calculation presented in the previous section. The following team sizes are loaded into AnyLogic for each mechanic agent. The database loads the ML that the agent is assigned to, the team size (in duo's), and also makes the distinction in mechanic type: regular or pitstop.

Maintenance Location	Team Size	Type
Leidschendam	8	Regular
Leidschendam	2	Pitstop
Maastricht	6	Regular
Maastricht	1	Pitstop
Onnen	10	Regular
Onnen	2	Pitstop
Watergraafsmeer	5	Regular
Watergraafsmeer	1	Pitstop

Table 6.3: Total number of mechanic agents (duo's) loaded into AnyLogic for each Maintenance Location

Setting up outillage agents

The outillage agents form the maintenance capacity in terms of train tracks. For setting up outillage, the database contains the outillage type, capacity and the maintenance location it is placed in. Some maintenance locations have multiple BIO tracks, others don't. The data specified in the outillage database forms the capacity of the simulation model: the amount of tracks where trains can be placed to perform any kind of maintenance. Some trains require specific outillage during their MTCE/Nawerk maintenance steps, others don't. This is determined by the track column.

Note that Leidschendam does not have 6 tracks for SNG, but it has 3 tracks that can fit 2 trains at the same time. Therefore, the capacity of regular maintenance tracks in Leidschendam is set to 6 for SNG type tracks in table 6.4.

Type	Maintenance Location	Capacity	Track type
BIO	Leidschendam	2	
KWB	Leidschendam	1	
AW	Leidschendam	2	
CLEAN	Leidschendam	2	
MTCE	Leidschendam	2	LDD-SLT
MTCE	Leidschendam	6	LDD-SNG
MTCE	Leidschendam	3	LDD-Alg
BIO	Maastricht	1	
KWB	Maastricht	1	
AW	Maastricht	1	
CLEAN	Maastricht	2	
MTCE	Maastricht	2	MT-Flirt
MTCE	Maastricht	3	MT-Alg
BIO	Onnen	1	
KWB	Onnen	1	
AW	Onnen	1	
CLEAN	Onnen	2	
MTCE	Onnen	3	ON-Alg
BIO	Watergraafsmeer	1	
KWB	Watergraafsmeer	1	
AW	Watergraafsmeer	1	
CLEAN	Watergraafsmeer	2	
MTCE	Watergraafsmeer	2	WGM-Alg
MTCE	Watergraafsmeer	2	WGM-ICNG

Table 6.4: Outillage database to set up outillage agents

6.3. Running the model

During a model run, the four maintenance locations of NS are simultaneously maintaining trains. Data is collected on the occupancy rates of train tracks within the maintenance locations, the throughput time of a train during scheduled maintenance, the delivery reliability of each ML, train withdrawal, and intermediate waiting times between maintenance tasks. Each year the simulation run progresses new trains are added to the train fleet and old trains are removed from the simulation. The maintenance capacity remains the same throughout the run. At the end of a single run, the data that has been collected is processed and saved, so that it becomes available for data analysis.

6.4. Model inner workings

Being able to fully comprehend the model demands that the model inner workings are understood. This section presents the key algorithms that the model uses during a simulation run: the determination of the condition of a train and the processing of a maintenance request.

6.4.1. Determination of train condition

The first algorithm that is implemented in the model is the one where the train condition is determined as it enters a workshop. Generally, the maintenance location 'knows' in what state the train will be before it arrives, but when it has arrived they will perform an inspection to determine the train condition themselves. This is to confirm the state of the train compared to what has been entered in software databases, and to create an image of the work to be done on this specific train. Combining the knowledge on the train's status beforehand, together with the visual inspections, the amount of work that has to be done is inventoried. Algorithm 1 displays the pseudo code of the algorithm as implemented in the AnyLogic simulation model.

Algorithm 1 Train Condition

Let $RandomUniform \leftarrow$ Uniform(0,1) ▷ Draw a random number between 0-1

if $RandomUniform \leq ProbConditionWorse$ **then**
  $condition \leftarrow "Worse"$
else if $RandomUniform \leq ProbConditionBad$ **then**
  $condition \leftarrow "Bad"$
else if $RandomUniform \leq ProbConditionAvg$ **then**
  $condition \leftarrow "Average"$
else
  $condition \leftarrow "Good"$
end if

Add steps to $MTCE\_Steps$ based on train condition

When entering a maintenance location, a train can have four conditions: good, average, bad, worst. Based on this condition, steps left to visit are added to the MTCE_Steps_Left list. The condition is determined by drawing a random number between 0-1. If the random number is lower than the probability that a trains are in the worst condition, the train will get the condition "Worse". If the random number is lower than the probability that a trains are in the bad condition, the train will get the condition "Bad". If the random number is lower than the probability that trains are in the average condition, the train will get the condition "Average". Else, the train gets the condition "Good". These probabilities are assigned by the user that runs the simulation.

6.4.2. Adding MTCE steps

Based on the condition of the train, the steps that a train will have to visit during its maintenance cycle are determined. The pseudo-code shown in algorithm 2 displays the working of the maintenance steps algorithm. A train always has to pay a visit to BIO, MTCE & CLEAN. The optional steps are KWB and AW, and even the duration of KWB is depending on the train's condition. In the simulation model, a train will have to visit both AW and KWB if it arrives at the maintenance location in the "Worse" condition. In case the condition is "Bad", the train only pays a visit to the aardwind (AW), where the chassis is repaired or even replaced. In case the condition of the train is "Average", it will visit the kuilwielenbank (KWB) instead of AW. In case the train is in a good condition, both AW and KWB will not be assigned as an obligatory step for the train to visit.

Algorithm 2 Adding Maintenance Steps

```
Add "BIO" to maintenance steps left
Add "MTCE" to maintenance steps left
Add "CLEAN" to maintenance steps left

if condition = "Worse" then
    Add "AW" and "KWB" to maintenance steps left
else if condition = "Bad" then
    Add "AW" to maintenance steps left
else if condition = "Average" then
    Add "KWB" to maintenance steps left
end if
```

6.4.3. Maintenance request algorithm

As soon as a train enters a maintenance location for regular maintenance, it will send out a maintenance request. The maintenance request concerns all steps that the train has to pass, which is based on the condition of the train (except for BIO, CLEAN and MTCE which are mandatory steps that will have to be passed by all trains). The request will be sent to the top-level agent Main, who then forwards the request to the home ML that is coupled with the train. Algorithm 3 aims to clarify the working of the algorithm that handles the train maintenance requests. All maintenance requests are collected in a list called maintenanceRequests. The algorithm below is called to pick the last request and links the request to available outillage & mechanics (if found).

In this model, the Main agent assigns the correct trains to their Home OB's, and checks whether the outillage and staff of this Home OB are available to fulfill maintenance requests. The algorithm goes as follows: A specific maintenance location checks the oldest maintenance request. It checks for the different types of available outillage if the train (Train T) has yet to visit that specific type of Outillage (O). If there is a match found (that is, a train has to visit O, and O is available), the train T is assigned to outillage O. Next, an available mechanic (M) is searched for. In case there is a mechanic available, train T is sent to the outillage O that is has been assigned to. At the same time (naturally computers don't do things simultaneously but within the same timestep), a message is sent to mechanic M to perform maintenance on Train A at Outillage O. The train will get to know which mechanic is working on it, and the outillage will learn which train will arrive. The latter is done by sending the "START OUT" message to the corresponding outillage. Finally, the maintenance request is removed from the list of maintenanceRequests.

Maintenance Request List

When a match has been found, there might still be other requests in the maintenance request waiting list. The while loop will then continue attempting to match outillage with trains, and send an available

Algorithm 3 Handling Maintenance Requests

```
1:  procedure Maintenance Request(t)                              ▷ Start handling the request for train t

2:      while maintenanceRequests ≠ empty do

3:          Get the first maintenance request: t
4:          Let outillage = null

5:          for o in Available Outillage do
6:              if o ∈ MaintenanceStepsLeft and track category matches train type then
7:                  outillage ← o                   ▷ Match outillage desired with available outillage
8:                  Remove o from available outillage
9:                  break                           ▷ Match found, exit for Available Outillage for loop

10:             else if Size of maintenanceRequests > 1 then
11:                 for length of maintenanceRequests do
12:                     Repeat lines 6-9                      ▷ Try to match other requests in queue
13:                 end for
14:             end if

15:             if outillage ≠ null then
16:                 break                     ▷ Train and outillage match found, exit for loop (line 5)
17:             end if
18:         end for

19:         if outillage = null then
20:             break                                 ▷ No current available outillage, exit while loop
21:         end if

22:         if outillage ≠ CLEAN then                    ▷ Skip in case of CLEAN (outsourced)
23:             Let mechanic = null

24:             if mechanicsAvailable ≠ empty then
25:                 m ← FirstAvailableMechanic                          ▷ Available mechanic found
26:                 Send message to mechanic m to start maintenance
27:                 Inform train t that mechanic m will be working on it
28:                 Inform mechanic m it will be working on train t
29:             end if

30:             if mechanic = null then
31:                 break                                   ▷ Mechanic not found, exit while loop
32:             end if

33:             Send "START OUT" to chosen outillage        ▷ Outillage o will now become occupied
34:             Inform train t about outillage o
35:             Inform outillage o about train t
36:             Remove train t from maintenance requests
37:         end if
38:     end while
39: end procedure
```

mechanic to it. This will go on until the request list becomes empty. Occasionally there is no match between MaintenanceStepsLeft and available outillage, or there is no mechanic available. In those cases, the while loop is stopped, but the maintenance request remains at the top of the request list. Because the request is not removed, the next time the function that handles maintenance requests is called, the first attempt will be handling the request that couldn't be finished.

So as mentioned, the allocation of a maintenance request to available outillage is based on a first in first out method; the oldest request will be the one that the algorithm will tackle. In case no match can be found for the oldest request, the algorithm will continue looking for a match.

Pitstop Requests

The allocation of mechanics to pitstop requests is done similarly to that of a regular maintenance request.

Order of maintenance tasks

Each time a train enters the Idle state, a new maintenance request is sent to Main. When a maintenance task is finished, it is removed from the list of 'MTCE_Steps_Left'. When there are no more maintenance steps left to perform, the train is ready to be shunted back into operation. When there are multiple maintenance steps left to perform, each one of them can be chosen. There is no specific order, as long as all steps will be performed. This is important because it affects the way trains are assigned to outillage. In general, planners of maintenance have preference for a fixed sequence of maintenance tasks, because it is easier to plan ahead if all trains follow the same sequence. In reality, this is not what happens. Trains are simply shunted to the type of outillage that is available at that time, because otherwise they might do unnecessary waiting for a long period of time. If a train has to start with BIO, while MTCE and CLEAN are available, it would not make sense to make the train wait until BIO comes available. In such a case, a train would be shunted to MTCE or CLEAN, so that it doesn't have to wait before undergoing a maintenance step. If all are occupied, of course a train will have to wait until one of them becomes available.

6.4.4. Pitstop crew

The pitstop crew has one responsibility: making sure that pitstop requests are handled as soon as possible, ensuring a minimal time of train withdrawal. It occasionally occurs that the pitstop crew has no pitstop requests to handle. If the maintenance crew is fully occupied, the pitstop crew has nothing to tasks to perform, and a new train wants to receive some kind of regular maintenance, they should help the regular maintenance crew. However, in the simulation model it is assumed that these two processes are kept apart from each other. In reality, when regular mechanics are busy and a new request arrives which could be performed by a pitstop crew (considering their qualifications), they would help the regular mechanics. In the simulation model this is not the case, because this is not how it was initially designed, and because it would add too much complexity to the simulation model within the scope of this research.

6.5. Model controls

To adjust a simulation during a run, AnyLogic provides the ability of using model controls. Model controls can be sliders, buttons, checkboxes, edit boxes for number input, radio buttons and a drop-down menu (referred to as a Combo box in AnyLogic). This section will elaborate on the model controls that are included in the model that has been built for this research.

In fact, model controls are not only used while running a simulation. A modeller might also choose to ask the user to adjust model settings before a run starts. This will be done in the pop-up screen that comes up when running the simulation, but before the simulation is actually started.

6.5.1. Add additional trains

For the new train types, the model user is able to add additional trains before starting a simulation run. This will once trigger the simulation model to add trains based on the user's input. The user specifies the number of trains for a train type it would like to add, and selects the year of introduction for those additional trains. As soon as the simulation is started and arrives at the year of introduction specified by the user, it adds the number of trains as it was told. The simulation will split the number into two types of the train series: the shorter and the longer version. The ICNG has a short version (ICNG V, consisting of 5 carriages), and a longer version (ICNG VIII, consisting of 8 carriages). This model control enables the DM to observe the effect of increased train fleet sizes for the train types that are yet to be introduced, given that the contracts that NS has with the manufacturers allow them to order additional trains if deemed necessary by the long-term planners of the train fleet. The determination of the train fleet size is out of scope, which is why this variable is included as an uncertainty in the simulation model.

6.5.2. Button: shut down OB

In the main agent, model controls are added that allow a user to create a fictitious 'disaster' generated by the button 'Shut Down "OB"'. This button will create the 'disaster'. When the button is pressed, the chosen MaintenanceLocation (OB) will be shut down temporarily. The duration of the disaster can be set by the user when adjusting the slider. During this event, trains that require maintenance at the unavailable disaster OB will be redirected to another OB. The redirection OB can be set in advance, but the user could also select random. In that case, the computer will randomly allocate trains to an alternative OB (e.g. an OB that differs from to the disaster OB. This step is repeated for for each train separately. Until the disaster is finished, trains that were undergoing maintenance at the disaster OB at the start of the disaster will be stuck; their maintenance will not be resumed until the disaster is over.

6.6. Randomness

Within the model, there are several occurrences of stochasticity. Step by step these will be elaborated on. The amount of daily driven kilometers by a train is determined by the Daily KM property, multiplied by a season-dependent factor, and multiplied by a random number uniformly drawn between 0.8-1.2. The determination of the train condition is based on a random number drawn from a uniform distribution between 0-1, as explained in algorithm 1. The duration of maintenance task is drawn from a triangular distribution. Which triangular distribution parameters are being used depends on the maintenance step that the train is in, together with the condition of the train when it entered maintenance.

6.7. Model Verification

The model verification process will be explained in this section. More details can be found in Appendix B. Verifying whether the translation from the conceptual model to a simulation model has been done correctly is crucial, as it forms the foundation for the experiments' results. Verification has been done simultaneously with building the model. After making minor adjustments, the model behavior is examined again and again. Whether the condition of trains results in optional steps is verified and checked. The duration of a complete maintenance job is compared to the allowed maintenance time, which eventually forms the delivery reliability. In case a train exceeds the maximum allowed maintenance time, the delivery reliability of the maintenance location decreases. It was also checked and verified whether different train types use their own characteristics when going into maintenance, for example DDZ/ICM/VIRM train types always have to go to the wheel polishing outillage, regardless of their condition when entering maintenance.

Intensive repetitive checking of whether a mechanic is assigned to the correct train and whether outillage is assigned to the correct train resulted in finding errors in the handling the maintenance requests. For example, in case the first train could not be matched, the whole queue had to wait (while perhaps maintenance request 2 and 4 could be fulfilled). Eventually, these insights were necessary to fine-tune the maintenance request algorithm to its current composition.

6.7.1. Single agent testing

To test the behavior of an agent-based model, a key verification strategy is to start with is to perform single-agent testing. It concerns performing a single run and following the behavior of one single agent, a train undergoing maintenance, without changing any of the input parameters. The train should behave according to the conceptual model, where it requires to fulfill certain steps (some obligatory and some condition-based), communicate with other agents (maintenance locations especially), and collect statistics. By leaving text messages, the pieces of code that are executed can be traced, which allows a smooth verification process.

6.8. Model Validation

To validate the model's behavior, expert judgements have been called upon. By executing the model under the presence of experts on the train maintenance system of NS, the model's behavior has been validated. Various statistics on maintenance were monitored, such as the average maintenance throughput time of a train, the number of trains that are withdrawn at each maintenance location (throughput time), and the delivery reliability of each maintenance location. AnyLogic enables the modeler to translate these statistics to communicable graphs, which can then be analyzed by any train maintenance DM. For example, the number of trains withdrawn by a maintenance location during a run is observed, and compared with realistic numbers.

In addition, to validate whether the model behaves correctly according to the real world, the amount of trains that are maintained weekly by each maintenance location was monitored and compared to empirical data. Under regular circumstances, that is with the base case parameters, the train maintenance system behavior corresponds to reality. Other output generated by the model, such as occupancy have been examined and compared to empirical data or NS standards as well.

7 Analysis of Simulation Model Output

The sub-question that guides the topics discussed in this chapter goes as follows: *How can the model provide valuable insights to enhance the robustness of decisions made by NS decision makers?* To facilitate robust decision making, the model as explained in chapter 6 was utilized. The aim is to enhance decision-making information by means of generating, analyzing and interpreting data that results from the model utilization.

During this chapter, first it is discussed what is considered to be added value for an NS decision maker of the train maintenance capacity. Next, the uncertainties are presented, which will form the input for the experiments. The experimental setup is then discussed, followed by a sensitivity analysis. Subsequently, scenario discovery is performed for the no-policy case and for cases that have been influenced by policy interventions. Lastly, conclusions based on the policy analysis are presented.

7.1. Added value for a Decision Maker

The major obstacle for a DM on the train maintenance system is the presence of deep uncertainty regarding train fleet size, available mechanics, differing duration of maintenance tasks, differing train conditions, and lastly there is uncertainty regarding future daily travelling distances of trains. So far, NS have not been able to quantify the effects of uncertainty on train maintenance. To enhance decisions made on the necessary future maintenance capacity, quantifying the role of uncertainties is highly valuable. It creates insights for DMs that have not been gained before, which is precisely what is aimed to achieve when facilitating decision making under deep uncertainty. The next section will present the main uncertainties present in the system, together with their lower and upper bound (i.e. the uncertainty space). The uncertainty space forms the basis for the experiments that will eventually generate results to be analyzed. Setting up uncertainty space has been done by comparing empirical data, consulting with experts, and by making assumptions regarding lower/upper bounds. For example, experts have shared their knowledge on the maximum number of trains that can be manufactured additionally (according to the contract between NS and the manufacturer), to get an idea of the theoretical upper bound of the future train fleet size. While it is highly unlikely that these bounds will be reached, it provides validated knowledge on how to construct future scenario's. The simulation model allows the DM to evaluate the performance of the current maintenance capacity under a large variety of future scenario's. This adds massive value to the decision-making information quality compared to the previous approach used by NS, where only 1 scenario at a time can be measured in a deterministic way.

7.2. Uncertainties

Table 7.1 presents an overview of all uncertainties that have been included by the simulation model in AnyLogic. Their numbers are used when setting up each input parameter in an AnyLogic experiment. The 'train fleet size' uncertainty and the 'available mechanics' uncertainty are integers, meaning that they can only obtain a rounded number value within the specified boundaries. All other uncertainties are 'real' numbers, meaning that they can obtain any value between the lower and upper bound.

Uncertainty	Data Type	Lower Bound	Upper Bound
Train fleet size	Integer	0	11
Available mechanics	Integer	-3	5
KM's driven per day (Sprinters)	Real	0.75	1.75
KM's driven per day (Intercity's)	Real	0.75	1.75
Worse condition probability	Real	0.01	0.19
Bad condition probability	Real	0.2	0.349
Average condition probability	Real	0.35	0.6
Duration MTCE	Real	10.0	18.0
Duration KWB	Real	3.0	9.0
Duration BIO	Real	3.0	6.0
Duration CLEAN	Real	4.0	9.0
Duration Nawerk	Real	2.0	6.0
Pitstop duration	Real	0.5	1.5

Table 7.1: Model uncertainties and their specifics

7.2.1. Formation of uncertainty space

In this subsection, each uncertainty will be highlighted individually in order to fully comprehend its effect on the simulation model. They define the input parameters of the AnyLogic simulation model, so each composition of input parameters will cause different model behavior. The idea here is to run the AnyLogic model for all years given the composition of input parameters, immediately enabling the comparison between various model settings, of which it is interesting to explore what combination of model settings will yield desired outcomes. The latter forms the basis for the scenario discovery section, presented in 7.7.

Train Fleet Size

The 'train fleet size' uncertainty can take a value between 0-11. Each number represents a scenario pathway, which has been carefully constructed. The amount of trains that leave/enter the train fleet in the near future is known, but when looking ahead for more than 2-3 years, these values could end up different from what is currently expected (base case scenario). Therefore, two factors influence the scenario pathway, the first factor is the year that a certain decrease/increase occurs, and the second factor is the size of the decrease/increase. Combining these two factors results in different scenario pathways, where (e.g.) a train might be introduced later than expected due to unexpected complications, but eventually exceeds the base case scenario due to higher train demand. A visualization of future train fleet scenario's can be found in Appendix C.

Available Mechanics

The available mechanics uncertainty can take a number between -3 and 5. This indicates the number of mechanic agents during a simulation run compared to the current number of available mechanics. For example, if the uncertainty takes a value of -2, all MLs will have 2 regular mechanic agents less to their disposal. Comparing this with table 6.3, it will mean respectively 6, 4, 8, 3 available mechanics for each ML of NS instead of the current number of 8, 4, 10, 5. Keep in mind that each mechanic agent in the simulation represents a duo of mechanics in reality.

KM's driven per day Sprinters/Intercity's
This uncertainty multiplies the daily driven kilometers of a certain train by its value. For example, if this number takes a value of 1.5, all trains of that specific type (either sprinter or intercity's) will drive 50% more on a daily basis.

Worse Condition Probability
Determines quantity of trains that are in worse condition as they enter maintenance, as explained in algorithm 1. This uncertainty represents the value for ProbConditionWorse in the algorithm.

Bad Condition Probability
Determines quantity of trains that are in bad condition as they enter maintenance, as explained in algorithm 1. This uncertainty represents the value for ProbConditionBad in the algorithm.

Average Condition Probability
Determines quantity of trains that are in average condition as they enter maintenance, as explained in algorithm 1. This uncertainty represents the value for ProbConditionAvg in the algorithm.

Duration MTCE
This uncertainty affects the duration of the MTCE step within the scheduled maintenance. The duration of this step is determined by drawing a number from a uniform distribution, with lower bound of 10, and an upper bound of 'Duration MTCE': Uniform(10, Duration MTCE). This uncertainty thus determines the width of the uniform distribution.

Duration KWB/BIO/CLEAN/Nawerk
These uncertainties affect the duration of each a specific step within the scheduled maintenance. The duration of the step is determined by drawing a random number form a triangular distribution, with constant lower/upper bounds. The middle number of the triangular distribution however is determined by the value of this specific uncertainty. So the triangular distribution could be skewed to the left or right, depending on the value of the uncertainty. In this way, there is still stochasticity in the model regarding durations of scheduled maintenance steps, while they are slightly varied throughout the different experiments.

Pitstop Duration
The 'pitstop duration' uncertainty multiplies the number drawn from an empirical distribution, affecting the time it takes for the mechanic to perform and complete the pitstop repair. This specific uncertainty has been included in the model because it is unknown what future train failure will look like. It might become more complicated to detect the cause of train failure, since more and more components will be controlled electronically. Many more underlying causes might be of the issue, parts to be replaced could be more complex, and so on. On the other hand, it could also be the case that minor updates will already fix a failure raised by a train, which would mean a reduced duration of pitstop repair time. Taking into account these divergent possibilities is why this uncertainty is included in the AnyLogic simulation model.

7.2.2. Consensus for uncertainty bounds

While deep uncertainty is present as to what values the uncertainty takes in reality, there needs to be consensus on what the uncertainty space (i.e. lower/upper bound limits) will be.

Uncertainties related to task duration
While precise task duration is unknown, within NS there is consensus on what time a task should theoretically take. The lower bound is what is theoretically possible. E.g., it is known that tasks have at least a duration of x hours. However, delays might occur frequently leading to deeply unknown upper bounds for task duration: how much is x exceeded each time the task is performed?. Usually, more time is scheduled for a task than the time it actually takes to perform it. In this thesis, the upper bound is formed by the shift duration for some of the tasks, including slight delays.

Train fleet size uncertainty bounds
Correspondence with key players within the train acquisition and future fleet planning led to insights in contractual bounds of purchasing additional trains. The lower bound is the number of trains that are currently ordered and being manufactured. While purchasing additional trains does require contracts with manufacturers to be active, there will be no chance of completely utilizing the contract's additional option. Extremes were introduced as upper bounds, for example to create scenario's where 50% additional trains are bought.

Mechanic team size uncertainty bounds
Consultation of experts on future development of mechanic availability led to insights on minimum fte's (full time equivalents) expected to be required at each maintenance location. Again, this formed the basis of the 'delta_mechanics' uncertainty, but extremes were introduced, to illustrate the consequences on maintenance performance of having very little available staff. The upper bound of this uncertainty was set to 5, simply to observe scenario's where mechanic team size would be more than sufficient to compare scenario's where there are still maintenance capacity issues that are not caused by occasional mechanic shortages.

7.3. Generating Experiments

To efficiently generate a large quantity of unique scenario's, the Exploratory Modeling and Analysis (EMA) Workbench has been called upon (Kwakkel, 2022). The EMA Workbench is an open-source data analysis tool, in the form of a Python library. The EMA Workbench is perfectly aligned with the scope of this research, given that it "aims at offering computational decision support for decision making under deep uncertainty and Robust Decision Making" (Kwakkel, 2022). The Workbench comes with tools to setup experiments, run simulation models, and perform extensive data analysis on model outcomes. In this research, the EMA Workbench has been used to perform exploratory research. This is important to notice, because the Workbench also offers tools to perform (multi-objective) robust optimizations. However, given the scope of the research, optimizations are not performed. Rather than searching for an optimal solution, the experiments are generated to evaluate model behavior while varying all uncertain input parameters.

Designing the experiments could be done by hand, but this is highly time consuming, and would thus not be efficient. In this research, designing the experiments is done computationally by generating samples that cover the whole uncertainty space for all uncertainties presented in table 7.1. By default, the EMA Workbench generates samples based on random sampling. However, Latin Hypercube Sampling (LHS) is a more accurate and suitable way of sampling, given that samples are taken from the complete distribution of the uncertainty space. In that way, all interesting combinations are made, including the ones that might be less likely to occur.

Due to the high number of uncertainties that are present, the decision has been made to generate 2000 samples. With 2000 samples, the ranges of the uncertainty spaces are nicely filled with sample points. Each sample represents a combination of the uncertainties from table 7.1, forming the scenario that will be run by AnyLogic. In other words, AnyLogic will perform 2000 runs, where each run is done with different input parameters. Hence, the value of the input parameters for the AnyLogic simulation model has been determined by the EMA Workbench.

Year of Simulation

One of the input parameters that has not been mentioned yet is the year of simulation. For each scenario, the EMA Workbench was asked to provide a year of simulation. This was done because the length of a single run is set to 1 year. So the samples are labeled to a year of simulation, which varies between 2023-2034. The year of simulation determines the composition of the train fleet, given that this composition changes over time (some trains are introduced and others are flowing out, as depicted in Appendix C). An example of 5 samples that each form a scenario can be observed in Appendix D.

7.4. Key Performance Indicators

Within this research, several Key Performance Indicators (KPI's) are used to monitor the performance of the maintenance capacity as it is modeled in AnyLogic. The KPI's allow for communication between the model user and the DM on train maintenance capacity. The model user can tweak the settings, while monitoring model output. Interesting output is defined as output that behaves according to agreed standards. Table 7.2 presents the KPI's, including their thresholds. Thresholds are introduced to enable scenario discovery in a later stage (section 7.7), since outcomes will need to be classified to perform scenario discovery machine learning algorithms (Kwakkel & Jaxa-Rozen, 2016).

Key Performance Indicator	Threshold	Unit
Average throughput time of scheduled maintenance	56	hours
Maximum occupation of outillage	85	%
Maximum average train withdrawal	30	# trains

Table 7.2: Overview of the Key Performance Indicators and their thresholds

Note that the delivery reliability KPI introduced in chapter 4 is not included here, because it is represented by the throughput time KPI. Delivery reliability KPI is determined by calculating the number of trains that have exceeded the 56 hour limit of maintenance throughput time. Therefore, the delivery reliability is directly linked to the first KPI of table 7.2.

1. Average Throughput Time of Scheduled Maintenance

During scheduled maintenance, each train will go through multiple maintenance tasks. Whether a train is able to start each task immediately or not depends on the availability of both mechanics and outillage. In case either of them is occupied, the train will have to wait for them to become available, increasing the eventual throughput time of a train during regular scheduled maintenance. The average throughput time of scheduled maintenance indicates the amount of time (in hours) took for a train to fully complete their scheduled maintenance cycle (including waiting time). Since NS plans maintenance tasks in shifts of 8 hours, and a train is scheduled to complete 6 tasks (of which the MTCE tasks takes two shifts so in total 7 shifts are required to complete scheduled maintenance), 7*8 = 56 hours permitted. The 56 hours that a train is permitted to complete maintenance is considered to be the maximum allowed time that a train should be on maintenance. In case this number is exceeded, it might be caused by two things. Either the train condition was very bad, forcing the train to go through all maintenance steps, including AW and KWB. If then the time it takes to complete those tasks coincidentally takes a long time, the 56 hours limit is exceeded. Another reason might be that a train has to wait a lot between maintenance tasks, due to undercapacity at the ML. Eventually, reporting this KPI permits the model user to analyze the performance of the maintenance capacity under different scenario's. For example, longer average throughput times indicate congestion within the scheduled maintenance processes of one of the MLs.

In case of congestion, the maintenance capacity is insufficient according to the AnyLogic simulation model. To perform robust decision-making on train maintenance, this KPI, among others, should thus be monitored.

2. Maximum Occupation of Outillage
The maximum occupation of outillage is a number composed by two things. First, the monthly occupation of outillage is monitored. Then, the maximum of that number is saved for that year (which equals to a single run). The reason that the maximum number is chosen to be saved, is because at all times the occupation should not exceed a certain limit. If outillage is constantly occupied, it would be an indication of limited capacity compared to maintenance demand. From an efficiency perspective, one would argue that outillage should almost always be fully occupied, otherwise one would waste maintenance capacity. However, setting up maintenance capacity to be fully occupied, leads to almost no flexibility. Yet, being flexible is crucial within the dynamic world of train maintenance, as last minute changes are constantly lurking. Adapting to sudden changes is necessary, meaning that there should always be some room for unexpected changes in the maintenance planning. Therefore, this KPI is monitored, and the maximum number should not exceed 85%. It is assumed that a higher maximum occupation of outillage would be problematic, indicating shortage in maintenance capacity.

3. Maximum Train Withdrawal
Withdrawal can be understood as the number of trains withdrawn by any ML at any point in time. Withdrawal is directly linked to the operation, forming a highly important KPI to track for NS. In case withdrawal numbers exceed a threshold, the operation will feel the effect instantly. More trains withdrawn by MLs means less trains to be used during peak hour demands. Therefore the number of withdrawn trains should be limited, keeping a certain amount of trains available for operation.

Another reason why withdrawal is such an important KPI, is because it determines the amount of new trains to be purchased. The following example illustrates the importance of the withdrawal number for trains to be purchased by NS. If it is deemed necessary to operate 10 similar trains at the same time, but there is always one of them undergoing maintenance, NS will have to purchase a total of 11 trains; 10 trains for operation and 1 train that stands at the ML being at scheduled maintenance. On top of that, trains can be withdrawn by unscheduled maintenance: EBKs. This also adds to the number of trains that have been withdrawn from operation.

7.5. Sensitivity Analysis

This section aims to clarify to what extend variation in the AnyLogic output can be attributed to the variation of the input parameters. This is done by performing regional sensitivity analysis (regional SA). Regional SA applications ”typically consider model parameters as varying inputs, and aim at assessing how their uncertainty impacts model performance” (Pianosi et al., 2016). For a DM it is considered to be useful knowing what model uncertainties impact the simulation model performance, providing insight into possible key risk indicators, or even might raise awareness on the presence of model attractors. Attractors are model settings that drastically change model behavior. For example, if trains will drive more kilometers during operation, there might be a tipping point where the maintenance capacity is no longer able to fulfill all maintenance requests coming in. Regional SA does not point out attractors itself, but it helps identifying parameters that might cause such model behavior.

Figure 7.1 presents an overview of the sensitivity of model performance when tweaking the input parameters. The green line indicates the outcomes of interest that fulfill the KPI conditions, while the orange line indicates the outcomes that are not of interest. Cases of interest are those that have less than 56 hours of average throughput time, less than 85% of maximum outillage occupancy, and finally have less than 30 maximum withdrawal.

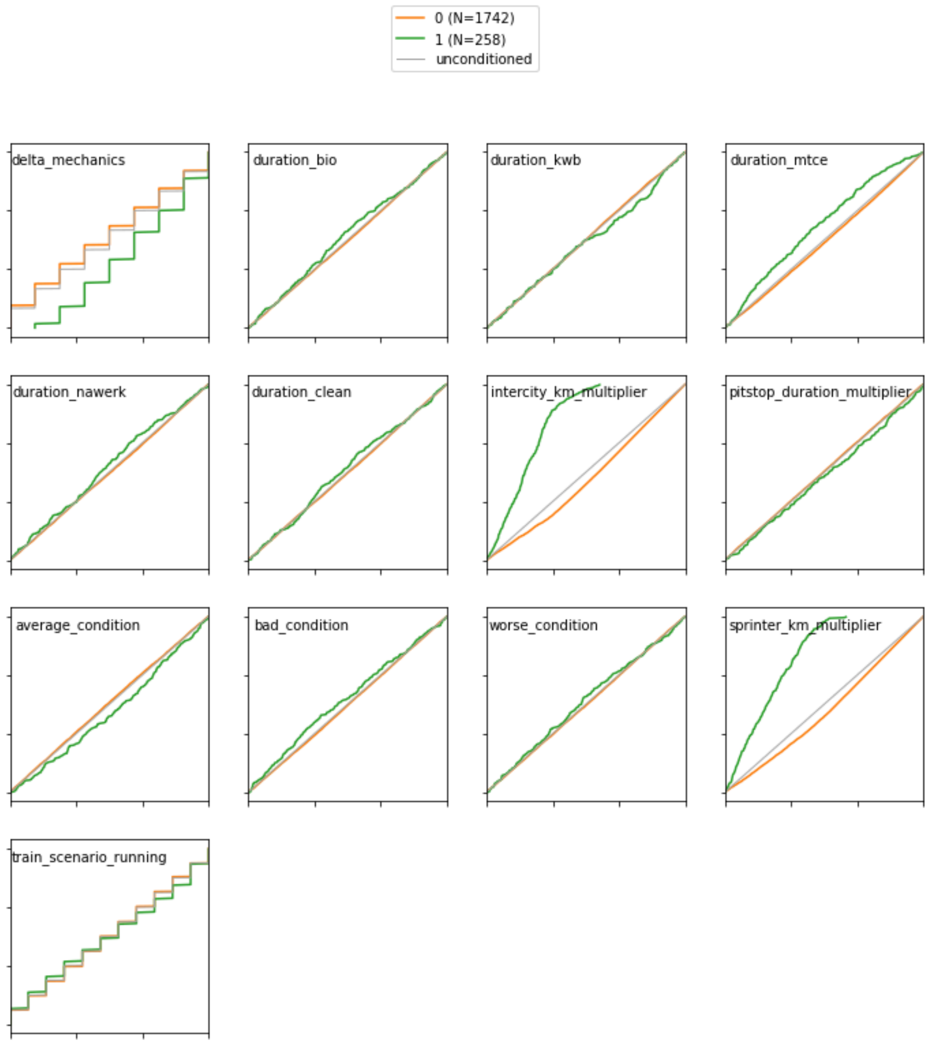


Figure 7.1: Sensitivity of the AnyLogic simulation model towards all model uncertainties

It can be observed that the model output (based on the KPI performance as defined in the previous section) responds sensitively to three uncertainties: the delta number of mechanics available for all MLs, the number of kilometers driven by intercity's, the number of kilometers driven by sprinters. This can be concluded by observing the deviation of the green line, compared to that of the unconditioned (grey) and orange line. Outcomes of interest converge towards unconditioned outcomes for higher values of the delta_mechanics uncertainty; if more mechanics are available at MLs it becomes less likely that the model output is affected.

For the kilometer multiplier uncertainties (both intercity & sprinter), the effect is the other way around. A higher value of the kilometer multiplier, i.e. the more trains drive on a daily basis, causes the maintenance capacity performance to reduce. This makes sense given that when trains have increased daily travel distances, they will have to get back to maintenance sooner. One would expect this effect to be more present for intercity's, since they cover longer distances during operation, meaning that they are already driving more daily kilometers compared to sprinters. Adding the effect of driving more kilometers, intercity's will reach the safety threshold (their maximum allowed of driven kilometers since last maintenance) sooner, making the agents in the model 'ask' for new scheduled maintenance sooner. This effect can be confirmed when comparing the green line of the 'intercity_km_multiplier' uncertainty to that of the the 'sprinter_km_multiplier' uncertainty, where the green 'intercity_km_multiplier' line shows a slightly larger deviation more from the unconditioned grey line.

Conclusions From Regional Sensitivity Analysis
From the Regional SA, it can be concluded that there are three uncertainties to which the AnyLogic simulation model responds sensitively to:

- delta_mechanics
- intercity_km_multiplier
- sprinter_km_multiplier

Above are the model uncertainties that have the largest effect on model output when running the various scenario's. While the 'duration_mtce' uncertainty shows a slight deviation from unconditioned outcomes, the model is not labeled as being sensitive towards this uncertainty.

7.5.1. Factor Prioritization

Factor prioritization (Appendix E) shows what influential uncertainties are of major impact towards model output. While the sensitivity analysis shows to what extend the model responses to tweaking the input, factor prioritization goes a step further by providing insight into KPI-specific effects of tweaking model input. This insight allows the DM to understand which uncertainties of the train maintenance system should be monitored more carefully. At the same time, it shows whether certain model uncertainties are not impacting model outcomes significantly. The relation between input and output is thus shown within one visualization in Appendix E. The color of each tile within the visualization represents the severeness of the relation according to the simulation model, e.g. how much is a certain output KPI influenced by tweaking model input.

From the factor prioritization it can be derived that influential uncertainties differ per Maintenance-Location. The intercity KM multiplier mostly affects Onnen, while the sprinter KM multiplier mostly affects Leidschendam. For Leidschendam and Maastricht maintenance performance the number of available mechanics is highly influential.

To further explore how the KPI's behave according to the simulation model output, the following section will visually present each KPI.

7.6. Visualization of the Model Output

To get a grasp of the model output, this section will discuss the model output in more detail. The aim is to help understanding the model performance according to the different KPI's that have been measured. All KPI's mentioned in table 7.2 will be analyzed individually to indicate the performance of the current maintenance capacity of NS. The outcomes that will be discussed are generated by running the 2000 experiments (that have been generated by the EMA Workbench) in the AnyLogic environment. Subsequently, the AnyLogic output has been exported to Python, enabling data analysis by making use of libraries such as Pandas, Matplotlib & Seaborn.

1. Average Throughput Time of Scheduled Maintenance

Figure 7.2 presents the average duration of scheduled maintenance during that scenario for all trains being maintained at that specific ML. The blue dots represent Leidschendam, which is the largest contributor to outlying outcomes. Next, Onnen & Watergraafsmeer deliver some high values for maintenance throughput time. Data output from Maastricht shows no outlying values for maintenance throughput time, indicating that there is little to no congestion taking place at that ML regarding the current maintenance capacity. Appendix F discusses cases of congestion in more detail.

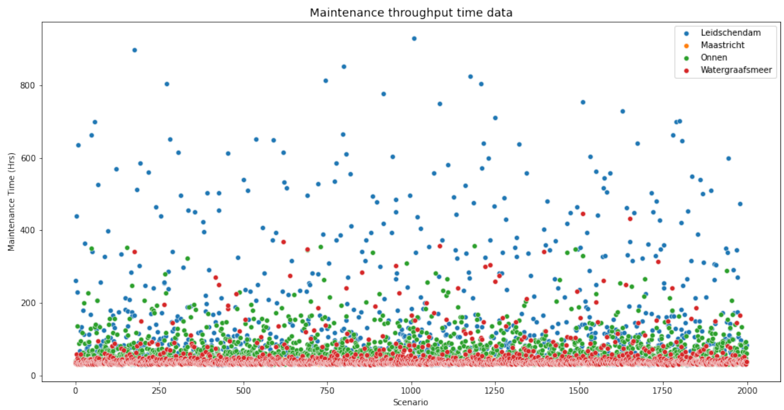
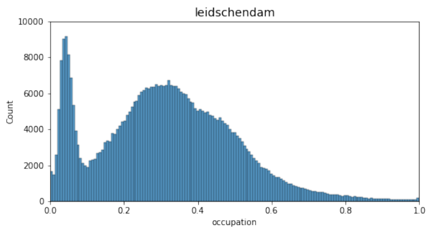


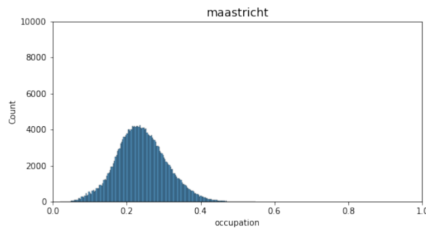
Figure 7.2: Scatterplot that clarifies the data for the maintenance throughput time KPI

2. Occupation of Outillage

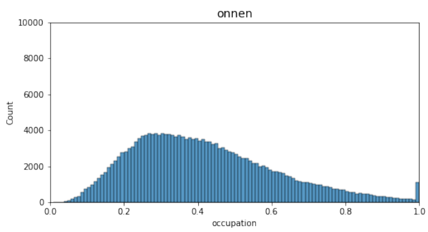
Four subplots are presented in figure 7.3, distinguishing occupation rates in the form of histograms presented for each ML.



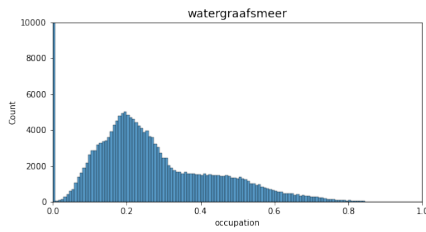
(a) Histogram plot of Leidschendam track occupation



(b) Histogram plot of Maastricht track occupation



(c) Histogram plot of Onnen track occupation



(d) Histogram plot of Watergraafsmeer track occupation

Figure 7.3: Visualization of track occupancy output for all Maintenance Locations of NS under 2000 scenario's

The main takeaway from this figure is that most maintenance locations have occupancy rates well below 0.85, indicating that the threshold of 85% is not reached often. However, for Onnen and to some extend also for Leidschendam, the threshold is exceeded. There is even a slight peak at 1.0 for Onnen, telling us that outillage was fully occupied during a whole month for several occasions.

To further analyse what outillages form a bottleneck for the executed runs, a search has been conducted to find the outillage types that frequently (>100 months within all 2000 scenario's) have had an occupancy rate of 1.0. These were found to be:

- Onnen KWB
- All Onnen regular MTCE tracks

3. Maximum Average Train Withdrawal
The highest average withdrawal for one run is collected. Averages are based on the average train withdrawal for 1 month within the simulation, of which the highest value (i.e. maximum average) is selected. High values for withdrawal can be caused by either congestion in scheduled maintenance, or congestion in unscheduled maintenance (pitstops). Figure 7.4 presents the distribution of the data for each ML. It can immediately be noticed that there is a large discrepancy when comparing the four MLs, where Leidschendam and Onnen have high cases of train withdrawal. Appendix F discusses cases of congestion in more detail.

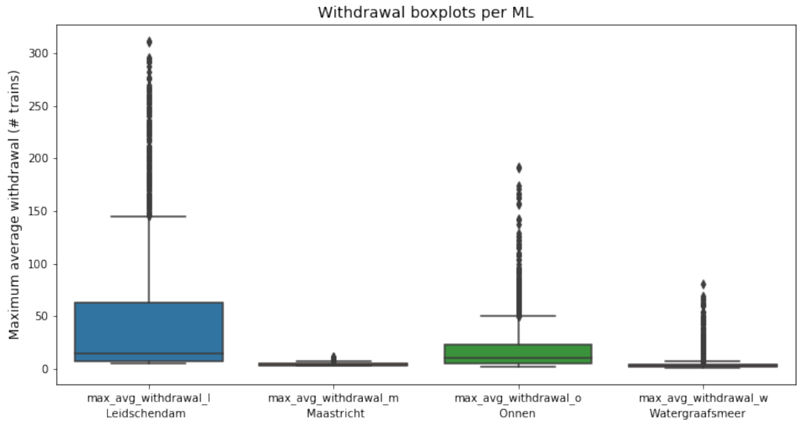


Figure 7.4: Boxplot presenting the spread of the output data regarding train withdrawal

7.6.1. Conclusions From Output Visualization

To create an image of the effect of deep uncertainty on the maintenance capacity, the model output has been visualized for each model KPI. The Average Throughput Time of Scheduled Maintenance showed outlying values for all MLs except Maastricht. From that it can be concluded that the other three maintenance locations, Leidschendam/Onnen/Watergraafsmeer, show behavior of some form of congestion under any of the 2000 scenario's that have been tested. The effect of the 'delta_mechanics' uncertainty can be labeled as the main contributor to congestion at Leidschendam, while having no effect on maintenance throughput time for Onnen. Analyzing the second KPI (Occupation of Outillage) showed that track occupation occasionally reached high levels for Onnen ML, from which it can be concluded that outillage is the main contributor for congestion at Onnen. Maximum Average Train Withdrawal showed outlying values for both Leidschendam & Onnen. The effect of the amount of daily driven kilometers by trains plays an important role in causing such outliers, in combination with low values for the 'delta_mechanics' uncertainty for Leidschendam, from which it can be concluded that the uncertainty regarding daily driven kilometers is an important one to assess by DMs on future train maintenance capacity.

While this section posed a visualization of the KPI's to get a better grasp of what the output data looks like, it would be highly interesting to see what output of interest is caused by which combination uncertainty input. The following section will permit these insights.

7.7. Scenario Discovery

In this section, scenario discovery will be performed to support robust decision making (RDM) on future train maintenance capacity. The EMA Workbench permits a modeler to perform scenario discovery using various techniques. The ones used in this research are the Patient Rule Induction Method (PRIM) as well as dimensional stacking. Both data analysis tools provide visual outcomes that can be conveyed to a DM, which can benefit from the insights gained. The benefit of presenting data analyses in the form of visual outcomes is that it can be explained to those that do not fully comprehend the what way results have been created, but still are able to draw conclusions from it. These conclusions can then be used during the decision-making process when deciding how to set up maintenance capacity, or when to revise current maintenance capacity performance.

Like the approach in the sensitivity analysis, scenario discovery allows the modeler to examine what input is responsible for interesting output, i.e. cases of interest. Interesting output can be be defined as output that fulfills the KPI needs, meaning no congestion for scheduled train maintenance. It has been explained that train maintenance is free of congestion when: average throughput time of all trains during scheduled maintenance does not exceed 56 hours, maximum occupancy rates of outillage stay under 85% for any month during a simulation run, and average withdrawal remains under 30 for all months during a simulation run. Cases of interest are cases that fulfill all of these conditions.

7.7.1. Patient Rule Induction Method

Now that the cases of interest have been defined, it is time to evaluate what uncertainties can be labeled as responsible for generating that specific output. The Patient Rule Induction Method (Kwakkel, 2015) is used to perform exploratory analysis on the AnyLogic output data. The PRIM algorithm performs an iterative process of peeling off data from the model output. It creates smaller boxes of output data, excluding the peeled of data each iteration. The quantity of data that is peeled off each iteration of the algorithm is defined as the peel-alpha. By default, the peel-alpha has been set to 0.1. The lower the peel-alpha value, the less data is removed each iteration.

The algorithm's goal is to find a box that contains as much as possible cases of interest, while minimizing the number of cases within the box that are not of interest. This will always result in a trade-off between coverage of the box and density of the box. Coverage can be understood as: of all cases of interest, how much of them are in the box? Then there is the density of the box, which can be understood as: of all cases in the box, how many are of interest? The minimum coverage threshold of a box is set to 0.8.

PRIM Results

Performing the PRIM algorithm on the AnyLogic output yields the following output, which can be observed in the distribution matrix presented in figure 7.5. The orange dots indicate the cases of interest. These come from a binary classification (True/False), which is the only way PRIM can read the output. 'True' cases are scenarios that haven't exceeded one of the boundary thresholds from table 7.2 (desired), whereas 'False' cases have (undesired). The PRIM algorithm attempts to find a box that covers as much as orange dots as possible, with minimal presence of blue dots inside the box.

The distribution matrix in figure 7.5 shows that three uncertainties were found to be significant in explaining the cases of interest: 'delta_mechanics', 'intercity_km_multiplier' & 'sprinter_km_multiplier'. This is in line with what was expected from the regional SA from the previous section, given that the AnyLogic model output responded sensitively to those uncertainties. What is interesting to observe is that the effect of the model uncertainties can now be quantified due to the large number of scenario's (2000 experiments) that have been tested by AnyLogic. The histogram plot for 'delta_mechanics' shows that most cases of interest lie within the positive range of the uncertainty (orange line).
To select a box, the following rule of thumb has been applied: density should be higher than 0.8. This means that 80% of points in the box are of interest, to convey an interesting part of information without

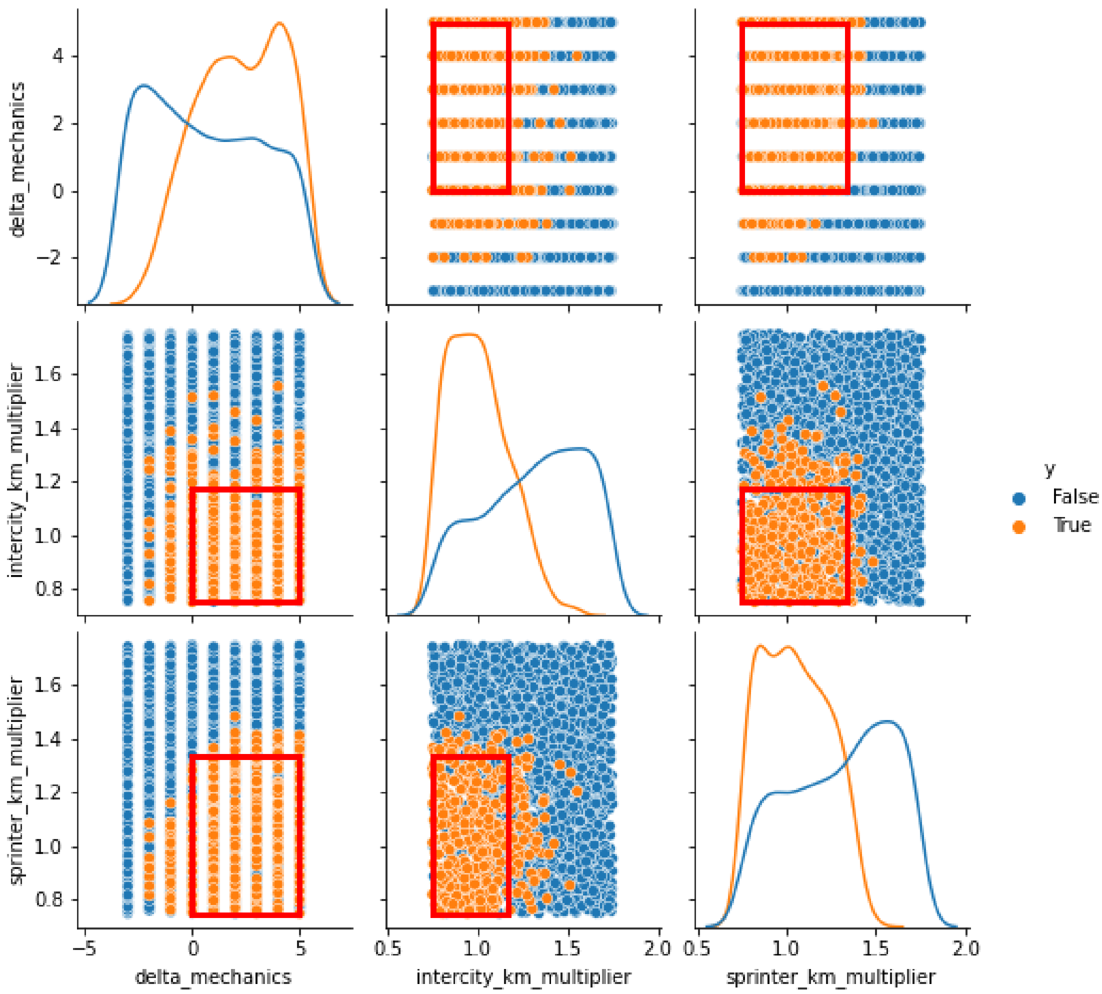


Figure 7.5: Results of performing the PRIM algorithm on the AnyLogic output

losing too much coverage. Plot 7.6 presents the box chosen. Each row is an uncertainty parameter. The blue line indicates the bandwith of the rule on that uncertain factor. The blue lines belong together, as these are restrictions on all factors. Number behind the uncertainty parameter names are P-values, which have to be below 0.05 (shows statistic significance). Non significance could suggest to ignore that specific factor. In this case, all uncertainties have been found to be of significant contribution to the chosen box.

The box found by PRIM will be presented in more detail, so that the allowed lower/upper bound of the three significant uncertainties can be explored. More than 80% of the cases of interest can be explained by the boundaries of the chosen box. From the PRIM algorithm it can be understood that delta_mechanics should not take any value below zero (see figure 7.6), which means that NS should monitor the number of available mechanics for every ML. The number of available mechanics should not decline to ensure robustness towards future scenario's. If the number of available mechanics declines, it might cause operational issues within the scheduled train maintenance. In the same way, the acceptable added number of daily driven kilometers by sprinters and intercity's is 34% and 17% respectively table 7.3, any further and there might arise operational issues.

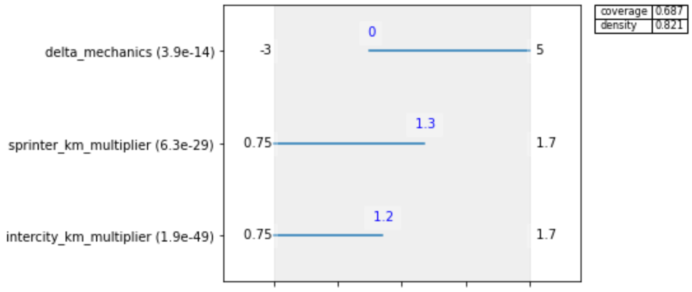


Figure 7.6: Uncertainties within the box found by the PRIM algorithm

Uncertainty	Lower Bound	Upper Bound
Delta Mechanics	0	5
Sprinter KM Multiplier	0.75	1.34
Intercity KM Multiplier	0.75	1.17

Table 7.3: Overview of the accepted boundaries of uncertainties found by the PRIM algorithm

PRIM conclusions

Performing PRIM on the AnyLogic model outcomes permits us to retrieve valuable insights on the effect of uncertainties on the performance of maintenance capacity. It can be concluded that there are three uncertainties which are of significant influence on model outcomes, which are 'delta_mechanics', 'intercity_km_multiplier' & 'sprinter_km_multiplier' (this is consistent with what was found from the sensitivity analysis and factor prioritization). For the 'delta_mechanics' uncertainty, its value should not reach any value below 0. For 'intercity_km_multiplier' & 'sprinter_km_multiplier' their values should not exceed 1.17 and 1.34 respectively.

NS decision-makers can use these insight to monitor the most important uncertainties, which allows them to act well in advance if necessary.

7.7.2. Dimensional Stacking

A more visual approach to scenario discovery is to perform dimensional stacking. Like the PRIM algorithm, dimensional stacking requires the output to be in the form of binary classification. Performing the algorithm in the Python environments for the AnyLogic simulation output returns a pivot table with the most influential uncertainties (figure 7.7). All influential uncertainties are split up into ranges. The performance of the output, in this case the performance of the maintenance capacity, can be observed within each uncertainty range. The color of each block indicates the output performance. The desired direction of the spectrum is towards the yellow end, so the lighter the color of the box, the more the output meets desired performance. For the NS train maintenance system, desired performance is little congestion, low withdrawal and not having overly used outillage.

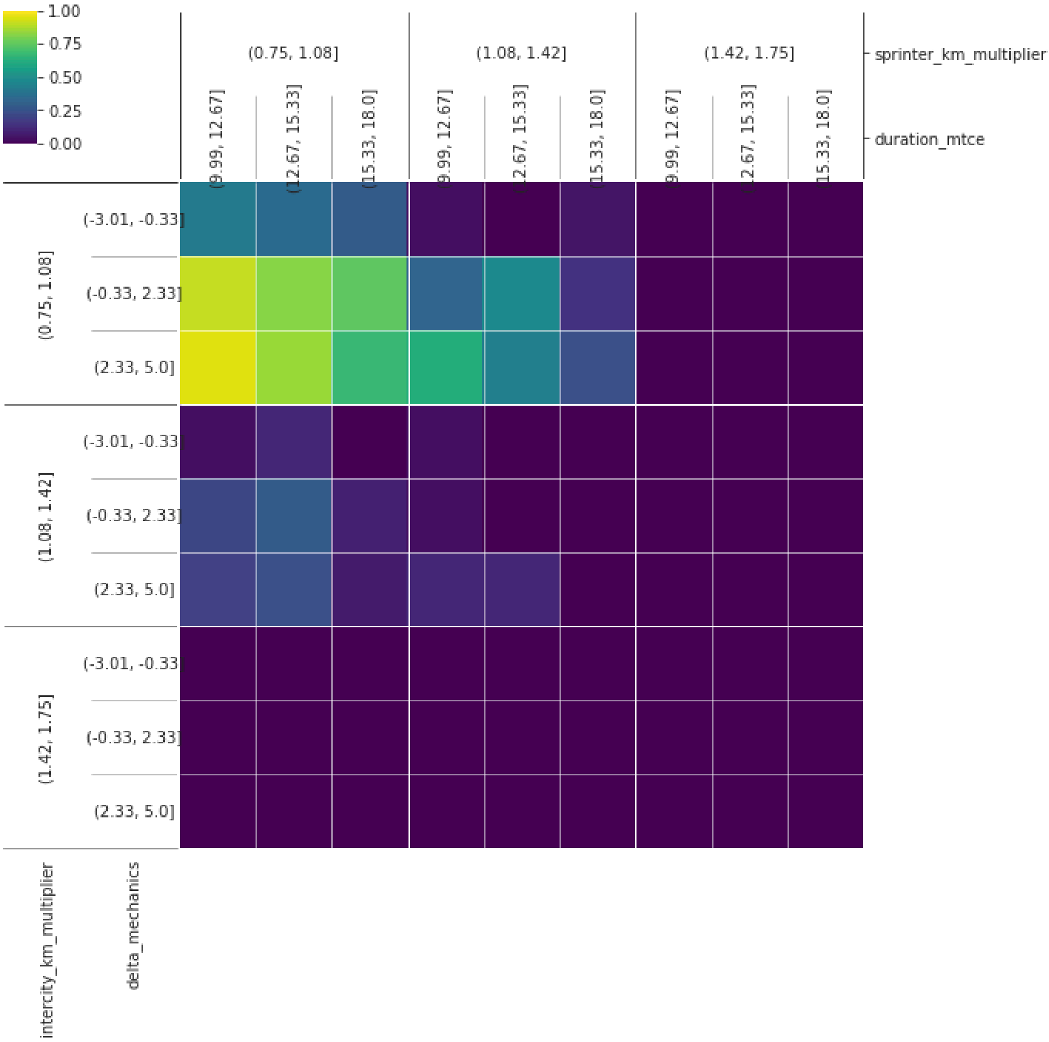


Figure 7.7: Dimensional Stacking Pivot Table presenting the most influential uncertainties

Compared to PRIM in the previous subsection, figure 7.7 conveys that another uncertainty is of important influence: 'duration_mtce'. It can be noticed that a higher value for this uncertainty causes model behavior to become less desired. The same goes for both kilometer multiplier uncertainties, where high values for the uncertainty result in undesired model outcomes. Again, a higher 'delta_mechanics' uncertainty results in more desired model outcomes, meaning that the maintenance capacity performs well under those scenario's.

Dimensional Stacking Conclusions

From the pivot table generated by the dimensional stacking algorithm, it can be concluded that there are four important, influential uncertainties: 'delta_mechanics' 'sprinter_km_multiplier' 'sprinter_km_multiplier' & 'duration_mtce'. Model output that corresponds to desired outcomes, that is outcomes that are classified as True, are caused by low values for 'sprinter_km_multiplier' 'sprinter_km_multiplier' & 'duration_mtce', and non-negative values for the 'delta_mechanics' uncertainty.

7.8. Policy Analysis

So far, all data analyses have been done without taking the implementation of any policy into account. To support RDM, it would be interesting to analyze several policies, and evaluate their robustness towards the 2000 different scenario's. The 2000 same scenario's have been run in the AnyLogic with the implementation of a policy. During the conceptualization phase of this research, few policy levers have been identified to be within scope of the strategic management department of train maintenance. This department has the ability to transform outillage so that it can facilitate maintenance on another train type. In addition, capacity could be increased or decreased (in terms of the number of outillage tracks that a maintenance location possesses). Finally strategic management could decide to adapt the capacity of pitstop handling mechanics, for example by increasing the number of trains that can be repaired at the same time.

The policy levers that have been tested within the AnyLogic environment are summed up below. Their effect is analyzed by comparing the effect of each policy lever model KPI's. Subsequently, each policy lever is explained.

1. None
2. Increased SLT & ICNG Capacity
3. Increased SNG & ICNG Capacity
4. Increased Pitstop Capacity

1. None

Outcomes related to this policy are generated by running the 2000 scenario's in AnyLogic without any implementation of policy levers.

2. Increased SLT & ICNG Capacity

This policy intervention is formed by a combination of two adaptations. Building additional maintenance capacity for SLT train types, which are sprinter trains that are being maintained at Leidschendam. Because the output showed some congestion occurring on SLT-specific outillage, it could be of value to analyse the effect of having such increased capacity. Within this policy intervention, the capacity of Watergraafsmeer too has been transformed. The current Watergraafsmeer capacity is 2 ICNG outillages and 2 regular outillages (where other trains at Watergraafsmeer are being maintained, together with international trains which are out of scope of this research). Since there are currently 2 ICNG (Inter City New Generation) outillage tracks to perform regular maintenance ('MTCE') and 'Nawerk' during scheduled maintenance, future undercapacity might occur. This could be caused by the possibility that additional trains of ICNG might still be bought in the near future. In case operation demands more trains of ICNG, it would mean that the workload for Watergraafsmeer increases as this is the only ML that is able to maintain ICNG train types. As such a scenario might lead to undercapacity in terms of ICNG outillage, a logical step would be to investigate the effect of transforming Watergraafsmeer towards having more ICNG outillage. The 2-2 balance would thus be disregarded and transformed towards a 3-1 balance (3 ICNG tracks, 1 general track).

3. Increased SNG & ICNG Capacity
Like the previous policy, this policy intervention is also formed by a combination of two adaptations. Building additional maintenance capacity for SNG train types, which are new sprinter trains that are being maintained at Leidschendam. Because the output showed some congestion on SNG-specific outillage, it could be of value to analyse the effect of having such increased capacity. At the same time, this policy intervention analyses the effect of transformed capacity within Watergraafsmeer (as explained for the previous policy). It should thus be noted that there has been 'built' both additional capacity for Leidschendam, while having transformed capacity for Watergraafsmeer.

4. Increased Pitstop Capacity
This policy intervention enables analyzing the effect of having increased pitstop capacity. The pitstop capacity is defined as the number of trains that could be repaired simultaneously within the simulation model. This policy lever only affects withdrawal numbers, because that has been set up to be the only KPI that measures both scheduled and unscheduled maintenance in AnyLogic. Withdrawal numbers are formed by the number of trains withdrawn from operation, both scheduled and unscheduled, which is where the effect of this policy lever could be noticed.

7.8.1. Effect of Increased Pitstop Capacity

Policy number 4, Increased Pitstop Capacity, has been analyzed by comparing the maximum average withdrawal output under 2000 scenario's. This is shown in figure 7.8. The reason that this combination of policy lever and KPI is presented is because the only KPI included in the model that can be affected by pitstop capacity is the withdrawal KPI. All other KPI's concern the scheduled maintenance cycle of the simulation model.

It can be observed that the policy has no effect on maximum average withdrawal within the AnyLogic simulation model (figure 7.8), given that both boxplots have an identical shape. What stands out is that both policy outcomes show a large fraction of outlying cases. These can be attributed to cases where there has been some kind of congestion within maintenance. It can be concluded that this congestion is not caused by a lack of pitstop capacity, because having increased pitstop capacity has no effect what so ever on the Maximum Average Withdrawal of trains.

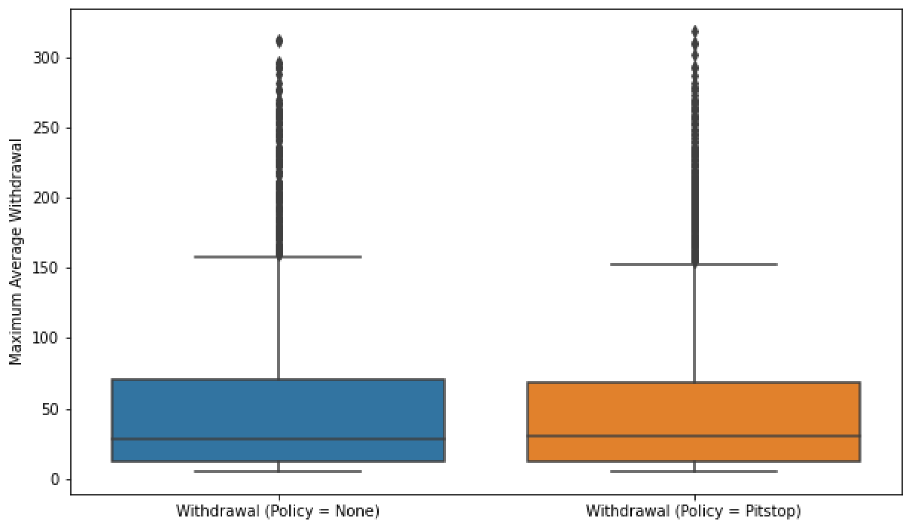


Figure 7.8: Boxplot showing the effect of increased pitstop capacity on train withdrawal

7.8.2. Analyzing All Policies

To compare all policies against each other, this subsection provides a visual representation of their overall effect on model KPI's. The aim of presenting these analyses is to perform exploratory analysis that enhances decision-making information quality. Average values for model KPI's have been calculated for each policy lever. This is a very high level of aggregation which should be kept in mind when interpreting outcomes presented below. Figure 7.9 presents a parallel coordinates plot, generated by the EMA Workbench, which effectively portrays the trade-off within KPI performance of all policies.

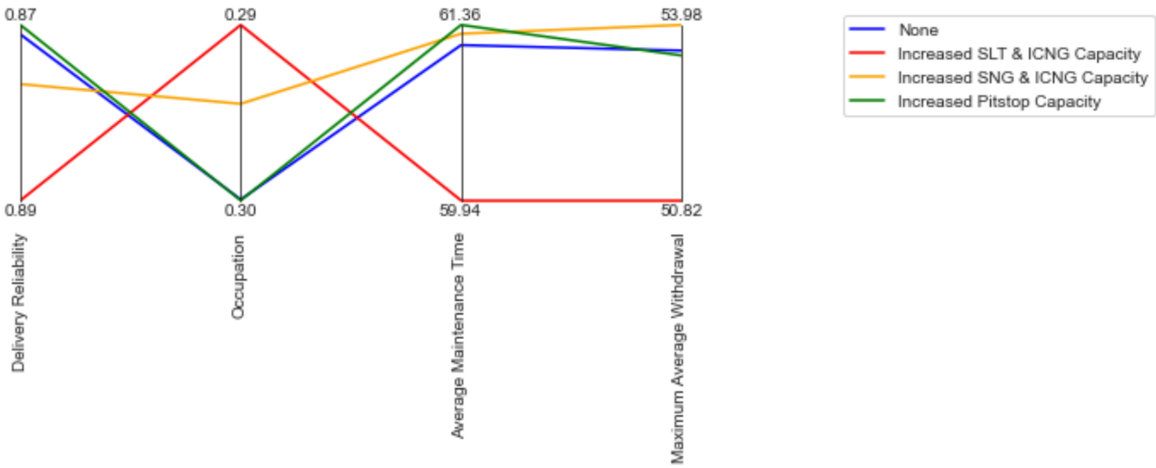


Figure 7.9: Parallel coordinates plot showing the effect of policy interventions on model KPI's

A parallel coordinates plot is formed by laying out the upper and lower boundaries of model outcomes, in this case KPI averages on 2000 scenarios for each policy lever. The lower and upper bound of a KPI forms the spread of one specific axis. Next, the corresponding values of that KPI for a specific policy form the location of the intersection of the coloured line with the KPI axis. The axis of Delivery Reliability and Occupation have been inverted, so that all KPI's have similar desired directions. In this figure the desired direction is towards the bottom of the vertical KPI axes. Each coloured line is represents the model outcomes for a policy. Which colour belongs to which policy can be seen in the legend on the right side of figure 7.9. The axis boundaries are rather close (e.g. only 0.02 difference in overall Delivery Reliability), indicating that the aggregated overall effects of policy interventions are minimal. Nevertheless, it provides one of the many analyses that facilitate robust decision making, which is why it could be of high added value for NS decision-makers. Observing the effects of policies on multiple axes has not been part of the current abilities of NS so far. This indicates the value such analyses could add for DMs on NS train maintenance capacity.

A trade-off can be observed when intersecting colored lines appear. For example, the 'Increased SLT & ICNG Capacity' (red line) policy generates the highest average delivery reliability, but it then intersects with all other policy levers, when observing its value for occupation. The trade-off be made by a DM would thus be: should we invest in high Delivery Reliability with the possibility of having overcapacity or should we not? A more balanced option could be to ensure mediocre Delivery Reliability & Occupation, while under-performing relatively on Average Maintenance Time and Maximum Average Withdrawal (yellow line: Increased SNG & ICNG policy).

Naturally, financial resources could play a large role when making such decisions, but those have not been included within the scope of this research.

7.8.3. Policy Analysis Conclusions

Within this section, four policies have been analyzed, of which one policy was the 'do nothing' policy (1). Two other policies (2 & 3) were related to adapted outillage capacity. The final policy (4) that has been analyzed concerns an increased capacity of simultaneous repairing of pitstops. From the policy analysis it can be concluded that the latter had no effect on the withdrawal KPI, which is the only KPI it could have affected. To decide which adapted outillage policy should be chosen, a trade-off has to be made between the performance of all KPI's. Whichever KPI is deemed more important by a DM leads to the relatively best performing policy from that perspective, as presented in figure 7.9.

7.9. Embeddedness in organization of NS

Now that all analyses have been presented, it would be appropriate to take a few steps back and interpret them within the bigger picture. To what extend can these analyses contribute to enhance decision-making? How can the outcomes of this research be embedded within NS, so that it will be incorporated into their decision-making process? Understanding the main takeaways from the output analysis is one thing, but acting on those outcomes accordingly requires effort as well. This section will provide a more holistic view towards the model outcomes, illustrating the full potential of this research for NS.

First of all, a major step forward is the inclusion of deep uncertainty into decision-making when comparing previously used methods within NS to the methods applied in this thesis. Deterministic models can be replaced by dynamic ones, and deriving results from one scenario is replaced by deriving results from evaluating thousands of scenario's. This could be seen as a major improvement, but it should not go unmentioned that non-experts too should be able to deal with the outcomes presented in this research. In essence, the model that has been developed forms a decision support tool. It supports decisions through evaluating the impact of the presence of deep uncertainty rather than prescribing optimal policies. In case a DM wonders what areas would need to be revised within train maintenance, and where investments should be made, it can design a policy and first run it in the simulation model environment before applying the policy in reality. The model will then show the performance of that policy, given the KPI's that have been set up beforehand. At the same time, the model supports identification of key risk indicators. From the PRIM analysis, it has become clear that few uncertainties need to be monitored in order to preserve desired system performance. Monitoring the amount of available mechanics within the near future, together with monitoring the amount of KMs trains drive when in operation, potential capacity issues can be addressed and prevented accordingly in an early stage. Acting before the situation occurs is what can be realized when monitoring the most important model uncertainties. In this way, decisions can be made to set up a robust train maintenance system.

Decision making under deep uncertainty (DMDU) not only requires a computer model that prescribes what to do and what to avoid, but it goes beyond that. It requires modelers, experts and managers to interact, share information to eventually enhance qualities of decisions to be made. This social part of DMDU should not go unnoticed. DMDU requires a shift in thinking and line of reasoning, and social interactions from that perspective will benefit the organization's view on DMDU. Instead of thinking in terms of efficiency and optimizations, DMDU requires a DM to base decisions on low regret or overall robustness. Accepting that there may never be an optimal solution to highly complex problems reduces the burden on DMs. Nonetheless, their decisions and judgements are still of great value, especially when supported by interpretations of outcomes that have been presented throughout this chapter. The analyses presented in this chapter improve the quality of decision-making information, permitting a DM of NS to incorporate robustness into its decision-making process. Considering that such approaches have not been held before within NS benefits the organization as a whole, which should eventually translate to a smoothly operating, robust, train system. Passengers of NS trains will be able to continue travelling on well-maintained trains for decades, which would be enabled by having a robust future train maintenance system.

8 Conclusion, Discussion & Recommendations

This chapter finalizes what has been presented so far by drawing conclusions, discussing the research, and suggesting recommendations for further research. Section 8.1 presents the conclusion by first answering all sub questions before drawing the final conclusion from this research. Thereafter, section 8.2 discusses the research findings. Finally, section 8.3 provides recommendations for further research possibilities, based on the research that has been conducted so far.

8.1. Conclusion

Main question that guided this research goes as follows:

How can the robustness of decisions made by decision makers of the NS train maintenance system be enhanced, considering the uncertainty and risks that NS faces in maintaining their rolling stock during the next 10-15 years?

From the main research question, several sub questions were set up that were based on intermediary objectives that this research aimed to achieve. Each sub question will be addressed individually before answering the main research question.

Sub question 1: What are the needs of a decision maker of the NS train maintenance system to make robust decisions?
This sub question has been addressed in chapter 4. In order to facilitate the needs of a DM, the most important performance metrics of the train maintenance system were identified. These are: maintenance throughput time, outillage occupation, train withdrawal & delivery reliability. Together, these form the model KPI's that indicate whether policies perform well, which benefits the decision-making information that is needed by a DM to make robust decisions. Robustness of decisions can be increased by leaving behind current deterministic decision-making tools, while introducing decision making under deep uncertainty where a large variety of scenario's are tested. Regarding the strategic management of future train maintenance, it can be concluded that being able to quantitatively compare the performance of the train maintenance system along multiple axes (instead of optimizing maintenance towards a single KPI) would be of great value for a DM.

Sub question 2: What does the NS train maintenance system entail?
To fully comprehend the overall working of the NS train maintenance system, the 2^{nd} sub question has been answered in chapter 5. From this chapter it can be concluded that there are four key players within the NS train maintenance system: Trains, Mechanics, Maintenance Locations & Outillage. Trains undergo maintenance at one of the four specific MLs within this research scope, which is labeled as their home ML. Maintenance tasks are performed by Mechanics that are part of a specific ML. The availability of MLs depends on their capacity, which is formed by the number of train tracks that allow specific types of trains to be maintained: Outillage. Given that Outillage can be train specific, their availability determines the smoothness of scheduled train maintenance. Occasionally, train failure occurs, adding another component to train maintenance on-site. All in all, it can be concluded that the maintenance system can be seen as a complex system, due to the large amount of trains to be maintained and the presence of constant interactions between the key players (the system agents).

Sub question 3: What model(s) can be built to support the decision makers of the NS train maintenance system?
Facilitating the NS decision makers requires the presence of a simulation model that provides insight into the performance of KPI's as identified by sub question 1, within the train maintenance system as identified by sub question 2. The train maintenance model has been built as an agent-based simulation model within the AnyLogic environment. All agents receive their own characteristics, that form the basis of the interaction. During a single run, model KPI's are measured, which are saved and exported after the run has finished to allow for data analyses. From validating the model with experts on train maintenance, it is concluded that the developed model supports DMs even more than current models do. This is mainly due to the opportunity of evaluating and quantifying the effects of deep uncertainty within NS train maintenance, which is something that has not been done so far. It can be further concluded that the developed simulation model supports Robust Decision Making, due to the allowance of testing the current (and even anticipated) maintenance capacity against a large variety of scenario's.

Sub question 4: How can the model provide valuable insights to enhance the robustness of decisions made by NS decision makers?
As mentioned while answering the previous sub question, the usefulness of the developed model shows especially when the effects of deep uncertainty are quantified. Through the analysis of 2000 scenario's this research was able to identify significant uncertainties that influence model KPI's. Desired model outcomes have been defined, which are primarily affected by 3 model uncertainties. These are: the number of available mechanics during simulation for each ML, the number of daily driven kilometers by intercity trains, and the number of daily driven kilometers by sprinter trains. Furthermore, scenario discovery by means of the EMA Workbench allowed for setting up boundaries that these model uncertainties should not exceed considering overall system performance:

- Delta Mechanics uncertainty should not reach a value below 0.
- Sprinter KM Multiplier uncertainty should not exceed a value of 1.34.
- Intercity KM Multiplier uncertainty should not exceed a value of 1.17.

While the duration of the regular maintenance task (MTCE) also played an important role in system performance, it was not identified to be of significant contribution to desired model outcomes. Policy analysis provided an example of how the performance of policy interventions can be measured. Based on the policies that have been set up in chapter 7, it can be concluded that increasing SLT and ICNG maintenance capacity relatively yields the best overall results for model KPI's. The two main contributors to possible congestion can be concluded to be the availability of staff, as well as the number of kilometers that trains will drive on a daily basis. Since Intercity's are already closer to their allowed KMs between maintenance, they have less slack regarding increased daily driving distances.

Final Conclusion

It can be concluded that the developed AnyLogic model adequately simulates train maintenance behavior on an aggregated level. It creates insights into the mid-longterm effect of deep uncertainty present throughout the maintenance system. To enhance the robustness of decisions towards train maintenance of NS, extensive modeling and exploratory analysis have been done. This allowed for the identification of key risk indicators: the boundaries that the uncertainties should stay within to achieve desired train maintenance system performance. Given that there has been no ability of quantifying the effects of the system's uncertainties before this research was conducted, it is concluded that this research eventually enhances future robustness of NS towards the (re-)organization of their train maintenance system. Enhancing future robustness can be realized through monitoring uncertainties in real life by collecting data on the most significant uncertainties that have been presented in this research, and acting upon that data in time if necessary.

To answer the main research question: simulation modelling including Exploratory Modeling and Analysis improves decision-making information quality compared to previous methods. It enhances the robustness of decisions by quantifying the effects of deep uncertainty, providing the ability to prepare for what is anticipated, but more importantly prepare for what is less anticipated as well.

8.2. Discussion

While the approach held in this research severely improves previous methods that NS has applied, it also comes with several limitations. One of those limitations is the perception that the modeler has on the system, the model bias, and the assumptions that have been made accordingly. The model bias can be seen as the consequence of the mental model that has been created after getting a first impression of a system. The first impression is leading in the formation of the mental model, what eventually leads to the conceptualization. Any misinterpretations or wrong understandings could lead to a less accurate simulation model. This effect has been minimized by means of verification and validation of the simulation model with strategic maintenance experts of NS.

Then there are limitations that are not raised by misinterpretation, but caused by decisions of deliberately leaving out relations and interaction: model assumptions. For example, to reduce complexity the decision has been made to separate pitstops and scheduled maintenance within the simulation model. In reality there might be occasions where pitstop mechanics help regular mechanics. For example when regular mechanics become short in staff and train requests start piling up. The help of pitstop mechanics might reduce task duration, allowing the mechanics to finish a train earlier then they might have been able to themselves. This does require pitstop mechanics to have the right qualifications, something that also has not been incorporated in the simulation model that has been built.

One could imagine that when maintenance is performed regularly, trains would raise less failures. Within the simulation model, no link between maintenance performance and train failure has been established. Train condition has been considered to be an external factor within this research, while in reality performing a higher quality of maintenance affects the condition of the trains as they return back into operation. There could always be accidents or random failures, but many of the failures raised by trains could be prevented when performing maintenance well, or when applying different maintenance strategies such as preventative maintenance instead of condition-based maintenance. The current model setup does not permit DMs of train maintenance to evaluate the link between maintenance strategies and maintenance failure.

8.3. Recommendations

Due to the scope of the research, extensive evaluation of different policy levers has not been included. To facilitate RDM, more research could be done to identify policies that perform well under all plausible future scenario's. This would require high capacities of computing power, as many scenario's will have to be run across a various policy levers. Yet, this approach could yield highly valuable results that could form the basis of (re-)organizing NS train maintenance into a highly robust maintenance system. One of the main reasons that such an approach would be highly valuable is because currently it is not known where the newest train types are to be maintained. Outcomes based on extensive policy analysis might severely improve the decision-making information regarding such complex decisions.

In addition, EMA allows the modeler to compare structural uncertainty. The scope of this thesis did not allow comparing the effect of structural uncertainty, but this could be done in further research. It requires two models to be compared, of which one is the complete model and another is a model where some relations that are presumed to be irrelevant are removed. Comparing the output of the two different simulation models allows for evaluating the importance of the relation that has been removed. If the model without the relation yields similar performance results with respect to the original model, it can be concluded that certain relations are not relevant. This could narrow down the scope of a train maintenance DM, as the focus could then lie on the relations that matter the most.

Another recommendation for further research is to include mechanic qualifications within the simulation model. Currently, the model works with mechanic duo's, given that maintenance tasks are never performed by a single mechanic in reality. In reality however, there are fully qualified mechanics which are occasionally assisted by less qualified mechanics. The inclusion of qualifications within the simulation model could contribute to a more accurate way of assigning mechanics to train maintenance requests. Perhaps it might even be achievable to link the current train maintenance model to other models on mechanics, their qualifications and employability. Further research could in that case contribute to more realistic maintenance behavior, as well as providing knowledge on the effect of mechanic qualifications on maintenance performance overall.

Furthermore, a recommendation can be given to enhance the model's intelligence. Creating more autonomous behavior from trains and perhaps even mechanics could raise the possibility of comparison between current maintenance strategies versus those of 'intelligent' trains. Intelligence could be implemented in a form anticipation towards occupation of a train's home ML. Suppose a train itself notices that it will reach the maintenance threshold soon while its home ML is relatively quiet and will remain quiet for the next 56 hours [1], it could indicate that the best time for maintenance is now. Even though the train would not be at the end of the allowed time in operation, such intelligence could reduce congestion when comparing to current train maintenance behavior, which would be interesting and valuable to investigate.

A final recommendation for NS would be to include the geographic locations of trains to evaluate the effect of the geographic location of MLs. Especially Onnen & Maastricht are difficult to reach. Maastricht even has excessive capacity in terms of outillage, so from that perspective it would make sense to maintain a larger fraction of the trainfleet of NS there. However, if trains operate far from Maastricht it is a costly project to maintain them at Maastricht. In those cases it would be time consuming, while requiring financial resources and available train operators that would be able to shunt the train all the way to the ML.

[1] 56 hours is the maximum throughput time of scheduled maintenance without the occurrence of congestion

References

Alexandrov, A., Bannikov, D., & Sirina, N. (2019). Agent-based modeling of service maintenance and repair of rolling stock. In *Iop conference series: Earth and environmental science* (Vol. 403, p. 012193). IOP Publishing.

Bankes, S. (1993). Exploratory modeling for policy analysis. *Operations research*, *41*(3), 435–449.

Burkhalter, M., & Adey, B. T. (2020). Modelling the complex relationship between interventions, interventions costs and the service provided when evaluating intervention programs on railway infrastructure networks. *Infrastructures*, *5*(12), 113.

European Union Agency for Railways. (2023). European rail traffic management system (ertms). Retrieved from https://www.era.europa.eu/activities/european-rail-traffic-management-system-ertms_en

Gribble, S. D. (2001). Robustness in complex systems. In *Proceedings eighth workshop on hot topics in operating systems* (pp. 21–26). IEEE.

Haasnoot, M., Kwakkel, J. H., Walker, W. E., & Ter Maat, J. (2013). Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world. *Global environmental change*, *23*(2), 485–498.

Kelly, R. A., Jakeman, A. J., Barreteau, O., Borsuk, M. E., ElSawah, S., Hamilton, S. H., ... Rizzoli, A. E., et al. (2013). Selecting among five common modelling approaches for integrated environmental assessment and management. *Environmental modelling & software*, *47*, 159–181.

Kwakkel, J. H. (2015). Scenario discovery in python. Retrieved from https://waterprogramming.wordpress.com/2015/08/05/scenario-discovery-in-python/

Kwakkel, J. H. (2022). Exploratory modeling and analysis workbench. Retrieved from https://emaworkbench.readthedocs.io/en/latest/index.html

Kwakkel, J. H., & Jaxa-Rozen, M. (2016). Improving scenario discovery for handling heterogeneous uncertainties and multinomial classified outcomes. *Environmental Modelling & Software*, *79*, 311–321.

Kwakkel, J. H., & Pruyt, E. (2013). Exploratory modeling and analysis, an approach for model-based foresight under deep uncertainty. *Technological Forecasting and Social Change*, *80*(3), 419–431.

Lai, Y.-C., Fan, D.-C., & Huang, K.-L. (2015). Optimizing rolling stock assignment and maintenance plan for passenger railway operations. *Computers & Industrial Engineering*, *85*, 284–295.

Litescu, S. C., Viswanathan, V., Aydt, H., & Knoll, A. (2016). The effect of information uncertainty in road transportation systems. *Journal of computational science*, *16*, 170–176.

Manzo, G. (2014). Potentialities and limitations of agent-based simulations: An introduction. *Revue française de sociologie*, *55*(4), 653–688.

Marchau, V. A., Walker, W. E., Bloemen, P. J., & Popper, S. W. (2019). *Decision making under deep uncertainty: From theory to practice*. Springer Nature.

Marchau, V. A., Walker, W. E., & Van Wee, G. (2010). Dynamic adaptive transport policies for handling deep uncertainty. *Technological forecasting and social change*, *77*(6), 940–950.

Ministry of Economic Affairs & Climate Policy. (2020). Climate policy. Retrieved from https://www.government.nl/topics/climate-change/climate-policy

Ministry of Justice & Safety. (2021). Overzicht vitale processen. Retrieved from https://www.nctv.nl/onderwerpen/vitale-infrastructuur/overzicht-vitale-processen

Mira, L., Andrade, A. R., & Gomes, M. C. (2020). Maintenance scheduling within rolling stock planning in railway operations under uncertain maintenance durations. *Journal of Rail Transport Planning & Management*, *14*, 100177.

Oskam, S. (2020). Icng-5. Retrieved from https://trainspo.com/model/6585/

Osman, H. (2012). Agent-based simulation of urban infrastructure asset management activities. *Automation in Construction*, *28*, 45–57.

Pan, B., Demiryurek, U., & Shahabi, C. (2012). Utilizing real-world transportation data for accurate traffic prediction. In *2012 ieee 12th international conference on data mining* (pp. 595–604). IEEE.

Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B., & Wagener, T. (2016). Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software*, *79*, 214–232. doi:https://doi.org/10.1016/j.envsoft.2016.02.008

Pinciroli, L., Baraldi, P., Compare, M., Esmaeilzadeh, S., Farhan, M., Gohre, B., … Zio, E. (2020). Agent-based modeling and reinforcement learning for optimizing energy systems operation and maintenance: The pathmind solution. In *30th european safety and reliability conference, esrel 2020 and 15th probabilistic safety assessment and management conference, psam15 2020* (pp. 1476–1480). Research Publishing, Singapore.

Shafiee, M., & Sørensen, J. D. (2019). Maintenance optimization and inspection planning of wind energy assets: Models, methods and strategies. *Reliability Engineering & System Safety*, *192*, 105993.

Tréfond, S., Billionnet, A., Elloumi, S., Djellab, H., & Guyon, O. (2017). Optimization and simulation for robust railway rolling-stock planning. *Journal of Rail Transport Planning & Management*, *7*(1-2), 33–49.

Van Dam, K. H., Nikolic, I., & Lukszo, Z. (2012). *Agent-based modelling of socio-technical systems*. Springer Science & Business Media.

Van der Lei, T. E., Bekebrede, G., & Nikolic, I. (2010). Critical infrastructures: A review from a complex adaptive systems perspective. *International Journal of Critical Infrastructures*, *6*(4), 380–401.

Van Duin, R., Bauwens, J., Enserink, B., Tavasszy, L., & Wong, K. J. (2016). Risk-aware roadmapping for city logistics in 2025. In *Proceedings of 6th international conference on information systems, logistics and supply chain*.

Verbeek, J. (n.d.). Personeelstekort, treinuitval, schuld: Voor ns hopen de problemen zich op. Retrieved from https://fd.nl/bedrijfsleven/1443434/personeelstekort-treinuitval-schuld-voor-ns-hopen-de-problemen-zich-op

Wiek, A., Keeler, L. W., Schweizer, V., & Lang, D. J. (2013). Plausibility indications in future scenarios. *International Journal of Foresight and Innovation Policy*, *9*(2), 3.

Wong, T. E., Srikrishnan, V., Hadka, D., & Keller, K. (2017). XLRM framework. doi:10.1371/journal.pone.0178874.g001

Wurth, T. [Ton], Nikolic, I. [Igor], Kwakkel, J. [Jan], Sloot, M. [Maarten], Cuppen, E., Quist, J., … Sloot, M., et al. (2019). Eindrapportage project windmaster: De weg naar een adaptief investeringsbeleid. *Delft University of Technology.(Opdrachtgevers: Gasunie, Ten-neT, Stedin, Port of Rotterdam, Provincie Zuid-Holland) https://doi. org/10.4233/uuid*, *12266*.

A

Simulation Tool: AnyLogic

As mentioned in chapter 3, the simulation model has been set up in the software AnyLogic, a Java based simulation software that allows for agent-based modelling, system dynamics & discrete event simulation (or a combination of them). As mentioned, the agent-based approach is held within the AnyLogic environment. While AnyLogic mainly offers a visual interface where pieces of code can be implemented, the back-end is an extensive java code file. The interface of AnyLogic allows those that are less familiar in programming to easily understand, interact and analyse the simulation model. The communication of the model formalization towards DMs can be done very smoothly through the presentation of state-charts. Even though DMs might have no knowledge about coding, it is not necessary for them to understand java code since the state-charts does most of the explaining itself, enabling smooth verification and validation of the model. The software allows a modeler to create a visual dashboard where statistics are displayed in plots. It enables the modeler to include model controls (input boxes, (radio) buttons, dropdown menu's, sliders etc.) which can all be adjusted by the final user of the model (both before and during the simulation). In that way, a final user is able to pull the strings on different model settings based on the user's preferences, experience, or simply to observe the effect different model settings.

Figure A.1 displays the states that a train goes through in the simulation model. The back-end that forms the inner workings of the transitions to other states is supported by java code which is less readable for a DM, but the states itself are very communicable towards DMs.

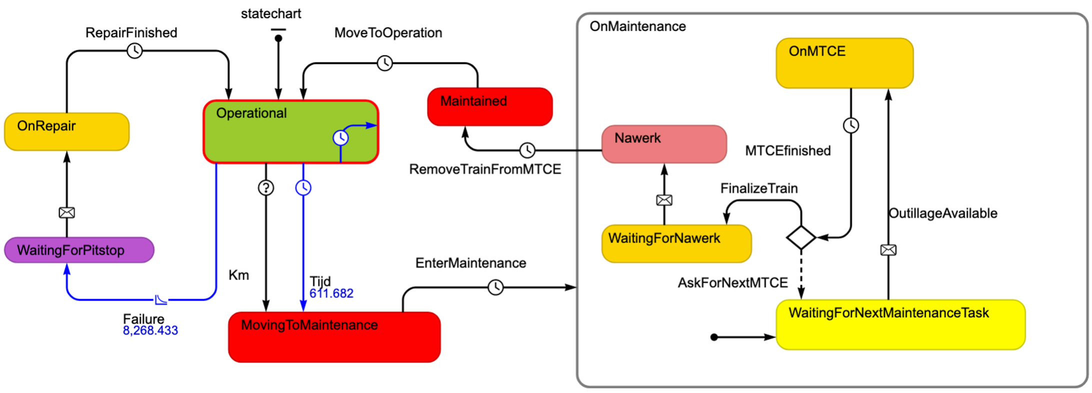


Figure A.1: Train statechart within the AnyLogic simulation environment.

B

Visualisation of the Verification & Validation Process

This appendix provides a more in-depth presentation of the verification an validation process that has been done to ensure that the simulation model has been built correctly and that it produces useful output. First of all, by checking if the model loads the correct number of agents it has been verified whether the model's maintenance capacity equals that of NS, whether the different train types are all included so that the number of trains equals to that of the NS train fleet. Then, together with experts on the management of the train maintenance capacity is has been verified whether the steps that a train undergoes during maintenance correspond to reality. Then by running the simulation, the model intermediary output is observed and validated with what intermediary output would be expected.

B.0.1. Allocation of agents in the simulation model

By checking if the model assigns the correct agents to the train that undergoes maintenance, it has been verified whether the model behaves correctly. For example for train number 51 in the simulation, which happened to be a FLIRT that is being maintained in Maastricht, it can be seen that the mechanic from MaintenanceLocation 1 is assigned to it (B.1), and that it is currently at the outillage of type AW.

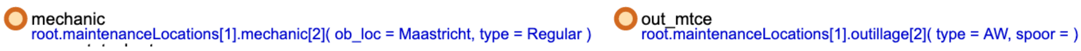


Figure B.1: Agents assigned to the maintenance step AW of train 51, which is a FLIRT-III being maintained at Maastricht.

When subsequently checking whether the AW outillage of Maastricht is indeed occupied by train 51 in the AnyLogic simulation, the outillage's statechart is visited. Figure B.2 shows that the outillage is indeed occupied by train number 51.

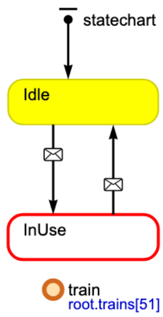


Figure B.2: The outillage AW of Maastricht being occupied by train number 51.

B.1. Train withdrawal validation throughout the simulation

Observing the number of trains being withdrawn without congestion occurring provides an image of what number of trains are undergoing maintenance at any time during the simulation. This will help validating whether the simulation provides useful output. Figure B.3 shows that between approximately 10-30 trains are withdrawn from operation (y-axis), which translates to about 3,5% of trains being withdrawn. This number corresponds to withdrawal numbers that would be expected within NS.

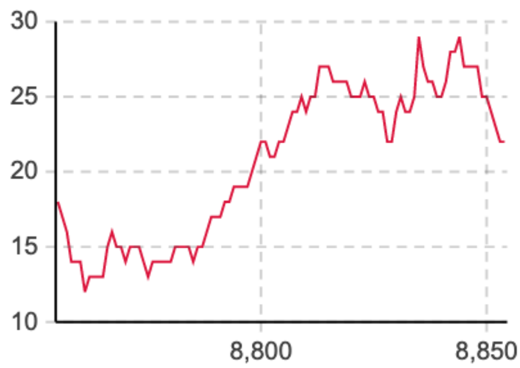


Figure B.3: Real-time withdrawal of trains during the simulation (without congestion).

B.2. Validation of maintenance throughput time

Without major congestion, the maximum allowed throughput time is 56 hours (consisting of 7 shifts of 8 hours). It would be expected that the 56 hour limit is not reached that often, given that many trains have optional steps in maintenance based on their condition. If their condition allows it, the trains skip AW/KWB (except for the DDZ/VIRM/ICM who will always visit KWB). This means that their maintenance takes less shifts, which would result in lower maintenance throughput times. Only when trains have to wait between their maintenance steps, due to occupied outillage or fully occupied mechanics, their maintenance throughput time increases and might exceed the 56 hours of accepted maintenance time. Figure B.4 shows that indeed the 56 hour bound is exceeded with only limited frequency.

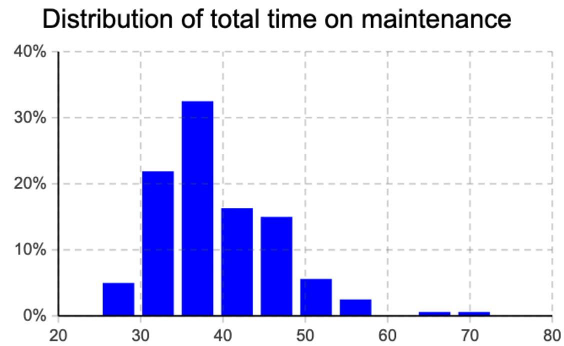


Figure B.4: Distribution of maintenance throughput time (in hours) for all trains within the simulation during one month of maintaining trains.

C

Scenario Pathways Trainfleet

This Appendix presents an example of the scenario pathways for the in- and outflowing trains. Each page shows the basecase scenario on top, which is equal to what's currently expected according to train fleet managers. The bottom figure on each page within this Appendix presents the possible scenario pathways that have been added to the experiments. Each color represents a train type, that has been anonimized due to confidentiality. For inflowing trains, variation in pathways can be observed in terms of final amount of trains of a specific type, while variation can also be observed in terms of timing of the inflow. The same holds for outflowing trains, where variation in terms of outflow timing can be observed. In this way, train fleet uncertainty is included in the scenario analyses.

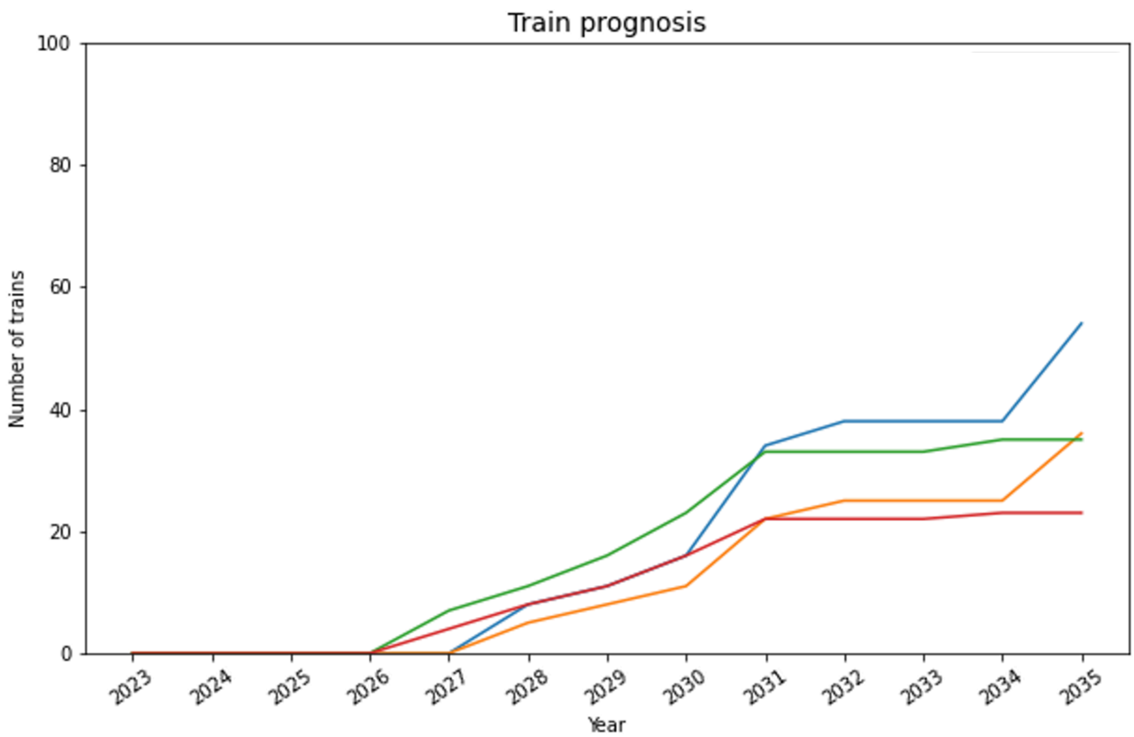


Figure C.1: Basecase scenario pathway inflow

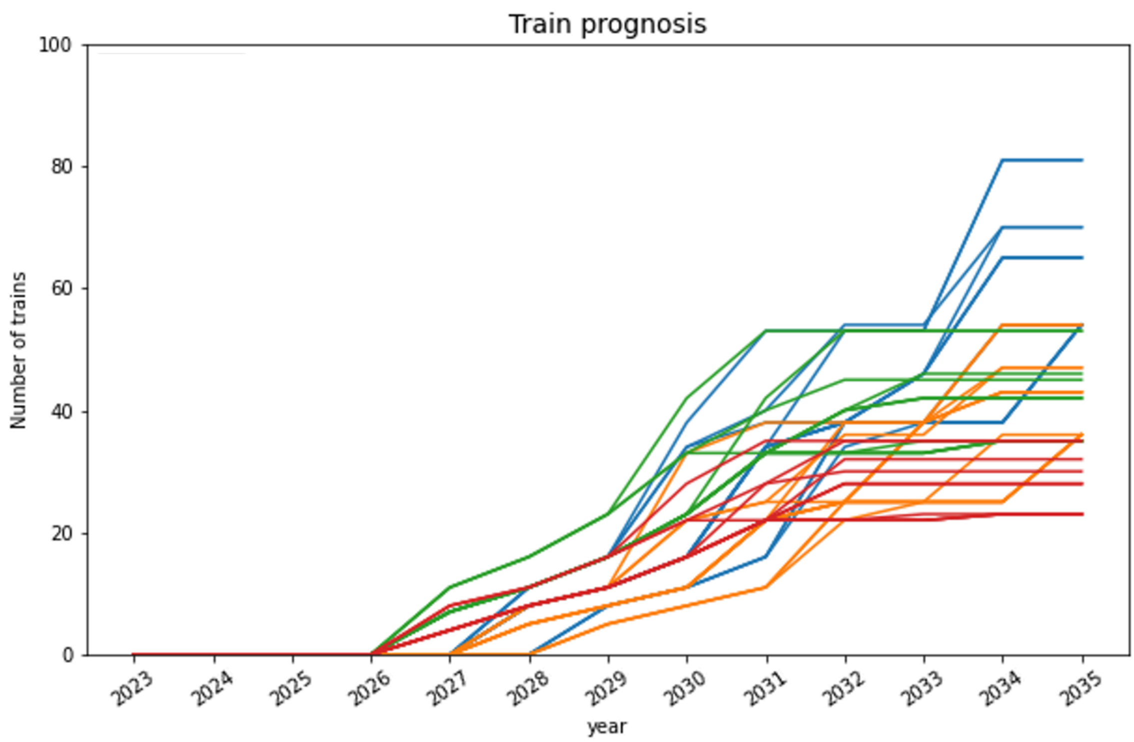


Figure C.2: Scenario pathways inflow

Figure C.3: Basecase outflow

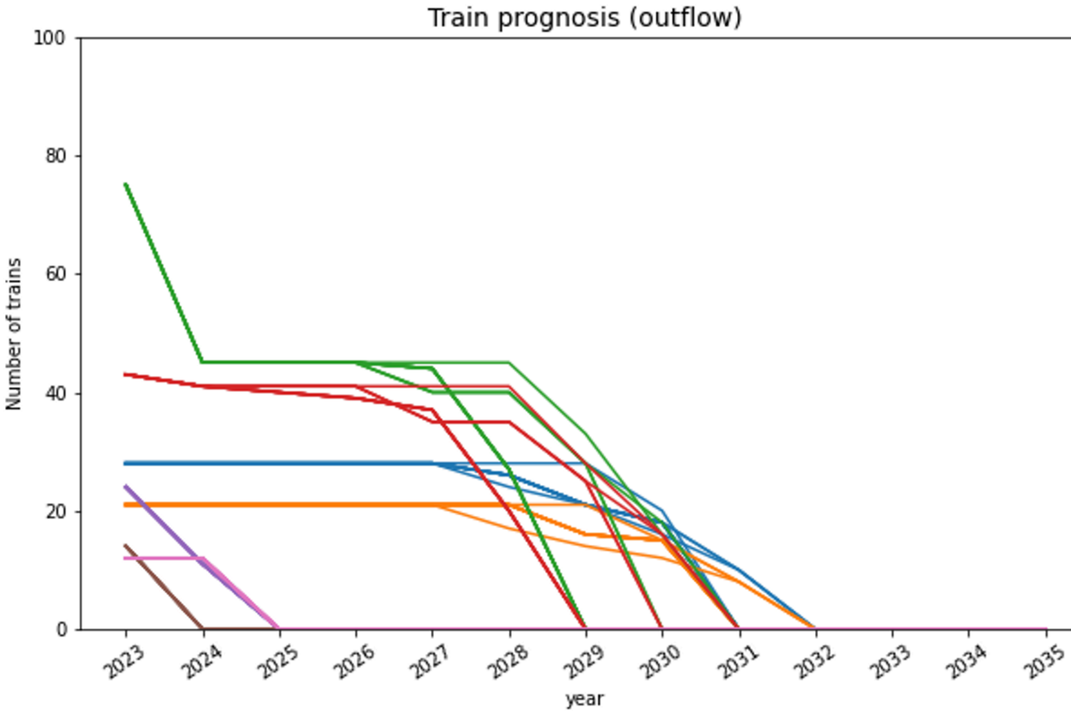


Figure C.4: Scenario pathways outflow

D

Example of LHS Samples

This Appendix shows an example of 5 samples generated by the EMA Workbench, using Latin Hypercube Sampling (LHS). Table D.1 shows the combination of values that are chosen as input parameters for the AnyLogic simulation model. To fit the table in one page, it has been split up into 3 segments, the number on top indicates the scenario number that corresponds to the observed values.

	delta_mechanics	duration_bio	duration_kwb	duration_mtce	duration_nawerk
1	2	4.427855	5.254838	11.369772	3.225928
2	5	4.056786	8.503175	17.745256	3.514747
3	-1	5.402259	7.045784	14.630256	3.440392
4	4	4.585530	5.172699	16.890932	5.111430
5	0	5.791809	3.424110	14.719786	3.495785

	duration_clean	ic_km_multiplier	pitstop_duration_multiplier	prob_condition_avg
1	7.341438	1.490706	0.647825	0.396305
2	7.737634	1.027141	1.138367	0.566786
3	5.575031	1.170828	0.782820	0.506104
4	7.303680	1.316593	0.976617	0.466589
5	5.484670	0.943799	0.766218	0.412851

	prob_condition_bad	prob_condition_worse	sprinter_km_multiplier	train_scenario_running	year_running
1	0.242802	0.179451	1.030893	9	2028
2	0.269071	0.099722	1.522843	10	2033
3	0.215858	0.106928	1.252493	10	2028
4	0.253410	0.010273	1.536484	10	2033
5	0.349084	0.194621	1.053512	10	2028

Table D.1: Example of five scenario's constructed by LHS sampling, designed by the EMA Workbench

E

Factor Prioritization based on the Train Maintenance Simulation Model Output

This Appendix presents the dimensional stacking output from the EMA workbench in figures E.1, E.2 & E.3. For each KPI it has been evaluated to what extend its output is influenced by the model uncertainties. If a cell color is towards the yellow end of the spectrum, it indicates that the relation between the uncertainty and the KPI output is highly important; the uncertainty greatly influences the model output for that KPI.

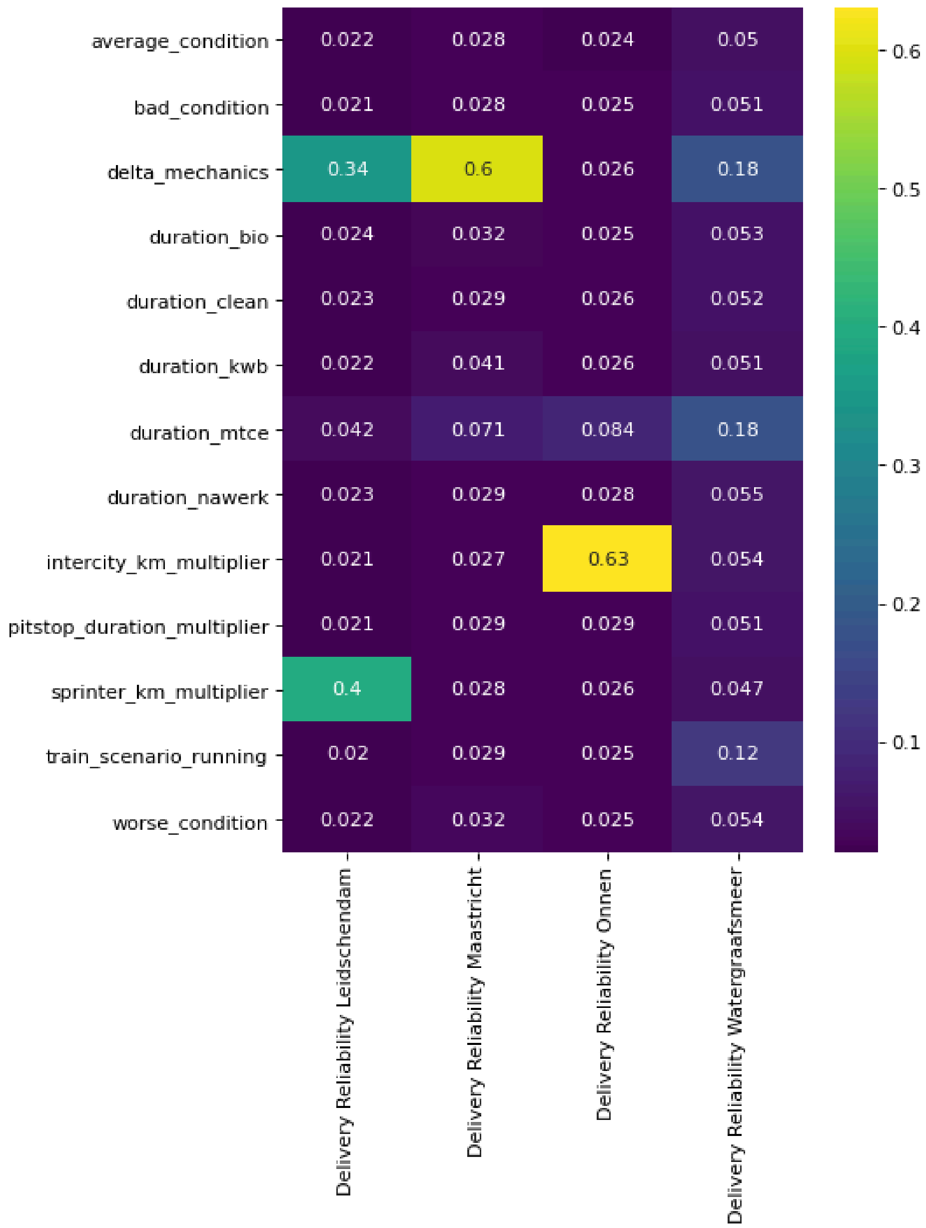


Figure E.1: Effect of model uncertainties on delivery reliability KPI, for each maintenance location separately.

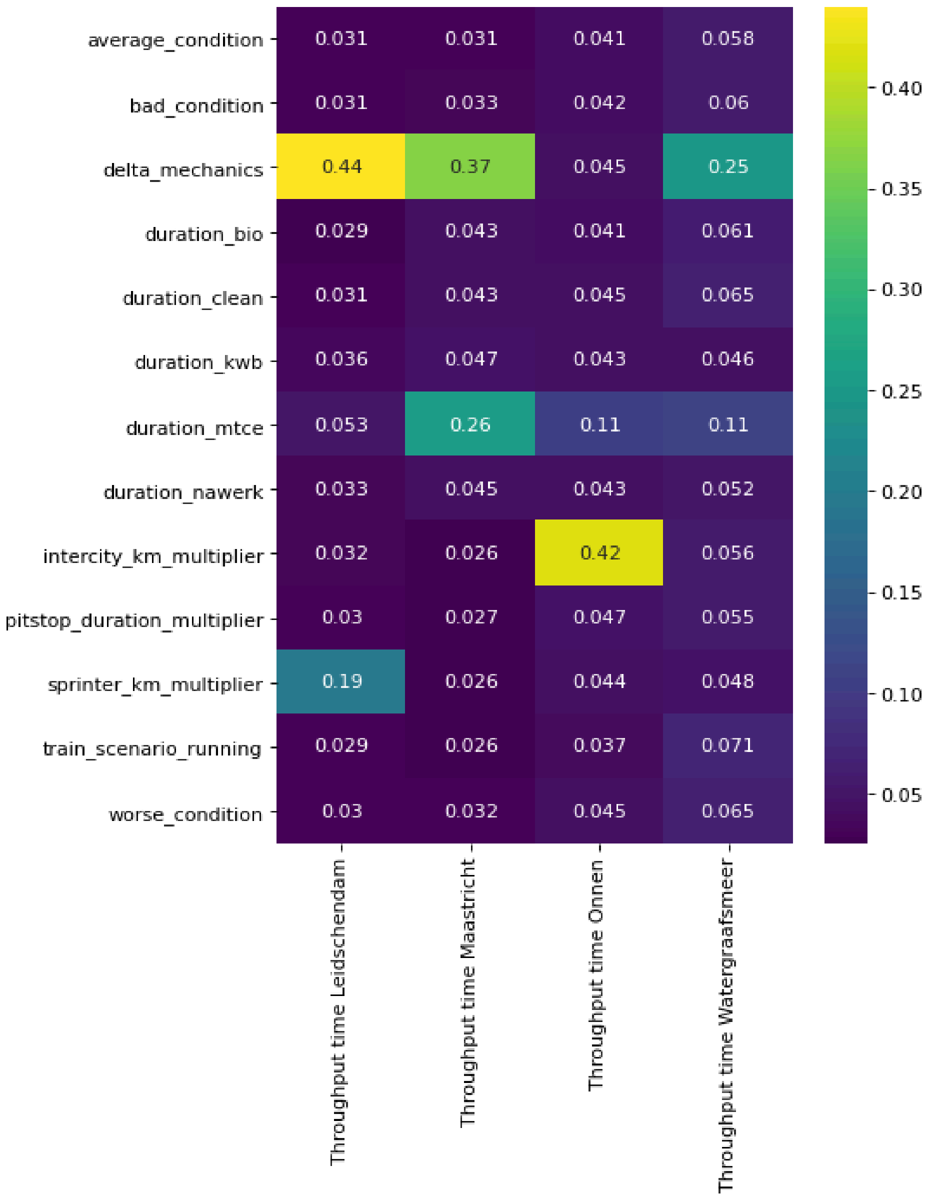


Figure E.2: Effect of model uncertainties on throughput time KPI, for each maintenance location separately.

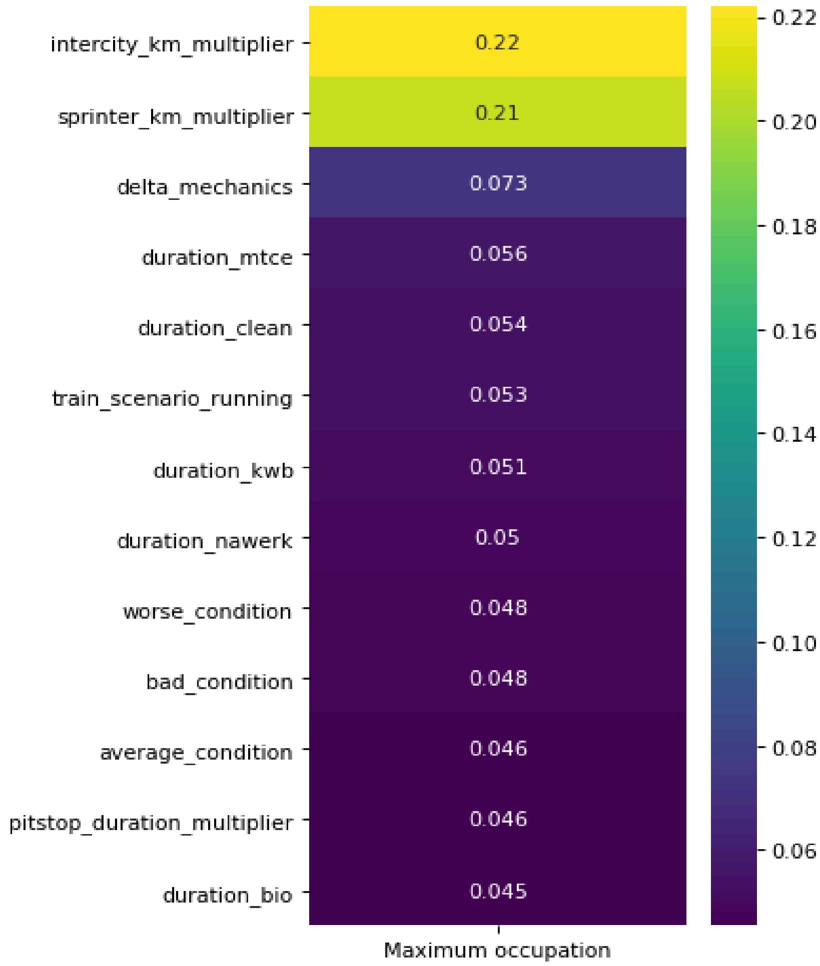


Figure E.3: Effect of model uncertainties on occupation KPI.

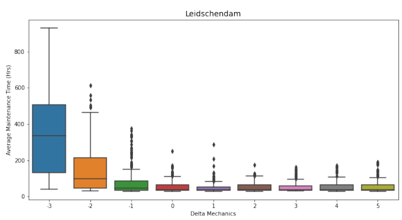
F

Additional output figures for model KPI's

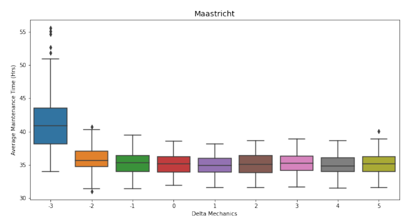
This Appendix presents additional visualizations of model KPI's. The effect of model uncertainties on maintenance throughput time and withdrawal hare discussed in more detail.

F.1. Maintenance Throughput Time

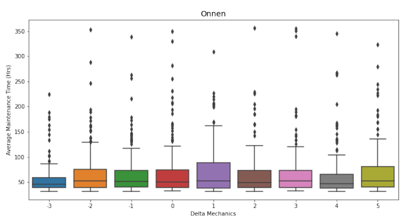
It should be kept in mind that the maximum allowed maintenance time for scheduled maintenance is 56 hours. However, this number is exceeded frequently. To understand where these extremely high values for maintenance throughput time come from, caused by some form of congestion, some further exploring of the effect of mechanic team sizes is done. Python enables comparison of the performance of scheduled maintenance under different mechanic team sizes. The effect of the number of available mechanics on maintenance throughput time for each ML is shown in figure F.1.



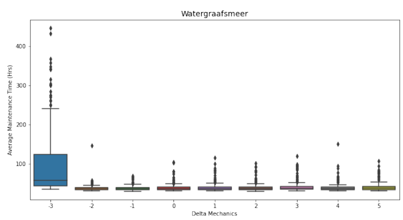
(a) Effect of team size on scheduled maintenance at Leidschendam



(b) Effect of team size on scheduled maintenance at Maastricht



(c) Effect of team size on scheduled maintenance at Onnen



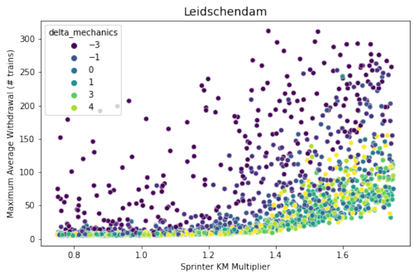
(d) Effect of team size on scheduled maintenance at Watergraafsmeer

Figure F.1: Boxplots visualising the effect of the delta_mechanics uncertainty input parameter

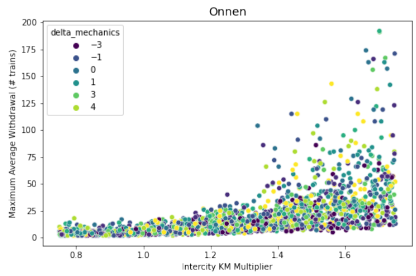
The maintenance location of which it has already been known (within NS) to be short in staff capacity is Leidschendam. This can be confirmed when looking at the data presented in figure F.1a. It can be noticed that as soon as the team size of available regular mechanics takes a value below zero, the average throughput time for train maintenance increases drastically for Leidschendam. Maastricht and Watergraafsmeer only show this behavior when the delta_mechanics input parameter reaches its lowest value of -3. Even though such a scenario is highly unlikely to occur, it gives an idea of the slack that MLs have towards mechanic team size. Average maintenance throughput time of ML Onnen is barely affected by the delta_mechanics uncertainty. From that it can be concluded that for Onnen it is merely the outillage capacity which causes congestion.

F.2. Withdrawal

Higher values for train withdrawal indicate some kind of congestion. Zooming in on the two maintenance locations that show signs of congestion during the execution of 2000 scenario's allows a DM to gain more information on the causes of such model behavior. This is presented in figure F.2.



(a) Scatterplot of Leidschendam train withdrawal



(b) Scatterplot of Onnen train withdrawal

Figure F.2: Scatterplots visualising the effect of input parameters on train withdrawal

The x-axis is based on the train types that are being maintained at the two MLs. Leidschendam is labeled as a 'sprinter' ML, and Onnen is labeled as an 'intercity' ML. Therefore, on the x-axis of figure F.2a the 'Sprinter KM Multiplier' input parameter is shown, together with the corresponding results for train withdrawal. For Onnen, the 'Intercity KM Multiplier' input parameter is shown, together with it's impact on train withdrawal. To show the effect of available mechanics on model output, the output cases are distinguished by their corresponding input for the 'delta_mechanics' input parameter. Dark colored dots correspond to less available mechanics. It can be observed that this is of great influence for train withdrawal levels at Leidschendam, while not being of any influence for train withdrawal at Onnen. Higher levels of train withdrawal at Onnen are in fact triggered when intercity's drive more daily kilometers.