University of Twente

Maintenance Management Optimization by Asset Categorisation

Master Thesis

Ellen Borger

Author

Ellen Borger
S1532049
e.j.r.borger@student.utwente.nl

Study

Master Industrial Engineering and Management
Faculty behavioural management, and social sciences

University of Twente

Drienerlolaan 5
7522NB, Enschede
Nederland
+31 534 899 111

Supply Value

Arnhemse Bovenweg 160
3708 AH, Zeist
The Netherlands
+31 88 0555 999

Supervisor University of Twente

Dr. E. Topan
Dep. Industrial Engineering and Business Information Systems

Dr. I. Seyran Topan
Dep. Industrial Engineering and Business Information Systems

Supervisor Supply Value

Luuk Spanjaards
Consultant

Preface

Dear reader,

In front of you lies my master's thesis, the last phase of my master Industrial Engineering and Management. It represents the end of my time at the University of Twente and the end of my student life. Therefore I want to take the opportunity to thank the people who supported me during the execution of this thesis.

Firstly I would like to thank Engin Topan for being my first supervisor at the University of Twente and guiding me through the last phase of my master's degree. Even though, stress sometimes got the better of me, Engin made sure I stayed focussed and stop stressing. I also would like to thank Ipek Seyran Topan for being my second supervisor at the University of Twente. Ipek helped greatly with structuring my thesis in such a way that not only I understand what I did but everybody hopefully will understand it.

Secondly I would like to thank Supply Value for offering me the opportunity to write my thesis at their company and connecting me with Company A for my data. I would like to thank everybody at Supply Value for making my graduation period so nice during a pandemic. Especially I would like to thank the business unit Supply Chain and operations, my supervisor Luuk Spanjaards and the expert of Company A for offering their help throughout the whole process.

Lastly I would like to thank my parents, boyfriend and friends for supporting me. Your support made my journey as a student filled with many great memories, Thank you!

I hope you enjoy reading this thesis,

Ellen Borger

March 2020

Management summary

Supply Value is a consultancy firm specialized in information management, performance management, project- and change management, procurement and supply chain management. Supply Value grew exponentially the last two years. Even though this growth brought new knowledge Supply Value wants to expand its knowledge even more to better assists its clients. One of the areas where knowledge growth is wanted is maintenance management. Supply Value assists other companies with projects, therefore correctly identifying the challenges with maintenance management within companies and expanding knowledge concerning these challenges is crucial for Supply Value. To understand the challenges of companies seven companies were observed. The current situation, the challenges they experience, and their foreseen future are all identified. As well as the gaps between the literature concerning maintenance management and the execution of maintenance management at the companies.

From the current situation of the companies the gap between the literature and the companies becomes clear. In the literature the maintenance management transition of the last century is divided into four generations. According to the literature we are currently in the fourth generation. This generation is characterised by the integration of safety and maintenance. However, most companies are still situated in earlier generation and experience challenges which obstruct the transition to a next generation. It is also found that the literature concerning maintenance concepts do not fit all companies. The maintenance concepts often focus on identifying the most important systems (MISs) and then focus on optimizing the maintenance for those MISs. This means that assets of the same system type are maintained in the same way. When a company has many assets of the same system type it may be better to categorise these assets of the same system type and specify maintenance per category instead of per important system. Therefore, the central research question of this thesis is: 'How can a maintenance strategy be improved, by implementing asset categorisation?'.

To answer the central research question quantitative data is used. Company A, which is one of the seven companies, supplied this quantitative data. Company A is specialized in infrastructure management and asset management. Company A maintains a high number of assets with a low value per asset for their client. The data for the thesis connects to the asset management part of Company A. The data set, consisting of data of 12911 assets, contains all maintenance orders, called tickets from November 2017 to March 2020, which are 33704 tickets in total. When a failure occurs a maintenance order is created after which corrective maintenance (CM) is applied to repair the asset. A ticket containing information about the failure is then created.

In the analysis we look at what the correct threshold is to divide the assets in to a good and poor condition category and we determine a correct time interval length for this threshold. Furthermore, failure causes are divided into cause groups. 64 failure causes are identified, since multiple causes concern the same component 15 cause categories are identified which consist of all failure concerning the same component. For example, a hinge failure and a lock failure are both part of the cause group door. For each cause group the conditional probability with the next failure is determined. This means that for each cause group *x* is determined that given a failure of cause group *x* occurs what is the probability that the next failure will occur in cause group *y*.

The solution consists of two parts. The first part focusses on achieving the data analysis objective for Company A. The second part focusses on shaping a maintenance concept framework for companies with a high number of the same system type assets with a low value per asset.

The first part of the solution is the solution for Company A. The solution for Company A consists of a threshold and time interval length for the categorisation and of a maintenance type per method. The threshold to determine whether an asset is good or poor condition is two. The interval length belonging to this threshold is three months. Hence if an asset had two or more tickets the last three months it is categorised in the poor condition category, otherwise it belongs to the good condition category. Based on the data set on average 5% of the assets are classified as having a poor condition. If another company than Company A wants to implement asset categorisation based on ticket quantity, they can use the data analysis set up of chapter 4, to determine the correct threshold and time interval length suitable for the company.

The maintenance for the good condition category is equal to the current maintenance at Company A, namely CM. The maintenance for the poor condition category will be a combination of CM and preventive maintenance (PM), more specifically called opportunistic maintenance (OM). Meaning that if an asset fails, a CM action will be executed to repair the failed component and a preventive maintenance (PM) action is executed on another component of the same asset, which is categorised as having a poor condition. The decision of which other component receives PM is based on the conditional probabilities, the component with the highest conditional probability is chosen to receive a PM action. The proposed solution for Company A is validated by setting up a simulation. We simulate the current situation and the proposed solution. Both the current situation as the proposed solution are tested on three KPIs the number of CM actions, the number of maintenance actions and the number of visits to the assets. The simulation results show that the number of CM actions decrease with 2.71%, meaning less failures occur. The number of visits also decrease with 2.71%. However, the total number of maintenance actions increases by 11.76%. Thus, a trade-off must be made whether the decrease in failures and visits to assets are worth the increase of total maintenance actions.

The second part of the solution consists of a framework which can be used when companies want to implement categorisation of assets based on the performance of the assets. The framework is designed after the implementation of the quantitative research at Company A was executed and is based on the on the CRISP-DM steps and the most common steps found in the literature concerning maintenance concept frameworks. The combination of implementation at Company A and the literature led to the proposed framework in this thesis. It must be noted this framework is suitable for companies with many assets of the same system type with a low value per asset. This framework consists of 7 steps described in the table below. The most important systems (MISs) in the framework refer to the assets placed in the poor condition category, thus different to how MISs are described in current literature.

Step	Name	Definition
1	Business understanding and data selection	Identify the overall maintenance objective and determine which data should be collected.
2	Data collection and data understanding	The data selected in step 1 is collected and in this step it must be understood what the collected data means.
3	Data Preparation	Raw data must be transformed into usable data by cleaning the data, construct new attributes and exclude unnecessary attributes.
4	MISs selection and MCCs identification	The MISs are the assets categorised in the poor condition category, these need the focus thus are the most important. The category threshold is determined based on the available data. The MCCs can be determined using FMECA, which indicates the components that affect the reliability of the asset most.
5	Maintenance policy selection	The maintenance policy for both the good and poor condition category are determined.
6	Implementation and evaluation	The strategy is evaluated and implemented in real-life and after implementation evaluated.
7	In service data collection and updating	Since implementing a new maintenance strategy can influence the behaviour of assets and components, e.g. less failures, the parameters should be updated with new data.

The general recommendation is to further research the proposed solution of asset categorisation based on their performance and test it at multiple companies. Besides the general recommendation there are also recommendations for Company A and Supply Value.

The recommendations for Company A are to:

1. Implement the threshold of two tickets over a period of the last three months to categorise in good and poor condition assets.
2. Do a more in-depth research of the causes of failures for all components.
3. Research the best PM action per component, e.g. time-based maintenance or condition-based maintenance, and the optimal parameters related to the PM action.

For the recommendations a global roadmap is made, which shows what actions should be taken now, what action should be taken next and what actions should be taken later on.

Now	Next	Later
• Select a project team to implement the categorisation of assets and determine the correct maintenance actions per group. • Define the business objective and success criteria. • Set up a project plan to start the transition to categorisation of assets using the CRISP-DM method as guideline. • Align the goal of the project with all stakeholders involved.	• Define what data is needed and start the data collection. • Analyse what the correct PM action per component is. • Implement the categorisation of components. • Implement the PM actions which are possible to implement.	• Identify the optimal parameters for all PM actions. • Update category threshold if needed. • Evaluate the results of categorisation and improve if needed.

The main recommendation for Supply Value is to start executing maintenance projects for clients, to further expand their maintenance knowledge. Also, it is recommended that Supply Value specializes in how to overcome the following challenges found connected to maintenance management:

1. The lack of data sharing can impede the shift towards PM;
2. The different cloud-based solution offered cannot be connected with each other;

People within a Company have different ideas about the best maintenance strategy, these ideas are often diametrically opposed.

Table of Contents

List of figures

List of tables

Abbreviations

Abbreviation	Definition
ACPT	Average costs per time
ADS	Advance service
BAS	Basic service
CBM	Condition-based maintenance
CM	Corrective maintenance
CRISP-DM	Cross-industry standard process for data mining
DTM	Delay-time modelling
KPI	Key performance indicator
MCCs	Most critical components
MISs	Most important systems
MRO	Maintenance repair and overhaul
OBC	Outcome-based contracting
OEE	Overall equipment effectiveness
OM	Opportunistic maintenance
PBC	Performance-based contracting
PM	Preventive maintenance
RCM	Reliability centred maintenance
TBM	Time-based maintenance
TPM	Total productivity maintenance

1. Introduction

Maintenance management has gone through a big transition the last decades. Arunraj & Maiti (2007) divides this transition into 4 generations. The first generation ended after world war II. This generation is characterized by basic routine maintenance and reactive maintenance. Maintenance was seen as a necessary evil. The second generation starts after world war II and end circa 1975. From 1950 industries become more complex and more dependent on machines. This increase of dependency leads to a relative increase of maintenance costs compared to other departments within companies. Thus, companies view maintenance management more and more as a core task. Three new maintenance ideas arise in this period, time-based maintenance, planned preventive maintenance and a system for planning and controlling work.

From 1975 the possibilities to use computer programmes to support maintenance management grow, which leads to the third generation. This generation is characterized by an accelerating use of automation, JIT production systems, continued increasing plant complexity, and an increasing demand for standard of product and service quality. The automation of maintenance management leads to the creation of a new maintenance concept, namely reliability centred maintenance (RCM) and new maintenance type condition-based maintenance (CBM). The fourth generation started with the start of the new millennium in 2000. Maintenance management in this generation is characterized by the integration of safety and maintenance, before this generation these where independent activities of a Company. Even though timewise we are currently in the fourth generation, most companies are still implementing aspects of earlier generations. Figure 1 shows the different generations and how they view maintenance.

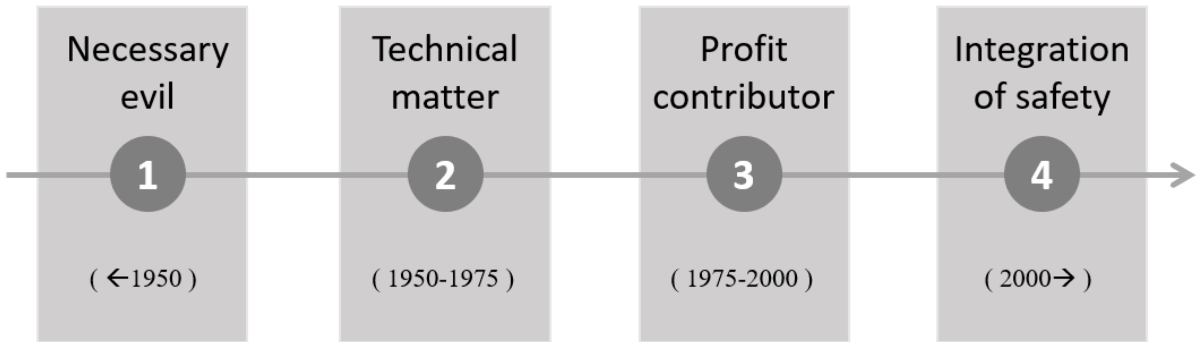


Figure 1 The evolution of maintenance

This thesis will contribute to the concept of CBM, which is an aspect of the third generation. In this study, we focus on CBM of a whole system. The layout of this chapter is as follows, section 1.1 gives a description of the Company, Supply Value, at which this thesis is conducted. In section 1.2 the core problem is determined, and the aim of the research is stated. Sections 1.3-1.6 describe the layout of this thesis. Table 1 shows the specifications per generation.

Table 1 The evolution of maintenance in-depth. Note: from Risk-based maintenance: Techniques and applications, by N.S. Arunraj & J. Maiti, 2007, Journal of Hazardous Materials, 142(3), p. 654 (https://doi.org/10.1016/j.jhazmat.2006.06.069). Copyright 2006 by Elsevier B.V. Reprinted with permission.

* First Generation: * Fix it when it broke * Basic and Routine maintenance * Corrective maintenance	* Second Generation: * Planned Preventive maintenance * Time based maintenance * Systems for planning and controlling work	* Third Generation: * Condition based maintenance * Reliability centred maintenance * Computer aided maintenance management and information system * Workforce multi-skilling and teamworking * Proactive and strategic	* Recent Generation: * Risk based inspection * Risk based Maintenance * Risk based life assessment * Reliability centred maintenance * Condition based monitoring * Computer aided maintenance management and information system
1940 1950	1960 1970	1980 1990	2000 Present

1.1 Company description

Supply Value is a consultancy firm based in Zeist, the Netherlands. It is specialized in performance management, procurement, supply chain management, information management and digital, and project and change management. Their clients are active in multiple industries such as fast-moving consumer goods, industry and high-tech, logistics services, health care, government, and the energy sector. The quantitative analysis will use data received from a client of Supply Value, from now on this client will be called Company A.

1.2 Problem context

Supply Value is a relatively young consultancy firm, it started in 2007. The last two years Supply Value has grown exponentially from 15 to 45 employees. This growth brought new knowledge, but Supply Value wants to expand its knowledge even more. So, it can offer a broader range of solutions to their clients. One of the areas Supply Value has limited knowledge about at the moment is maintenance management. Its wish is to gain more knowledge and insight in this speciality. Supply Value has multiple clients interested in optimizing their maintenance management.

The core problem for Supply Value is: ‘At the moment Supply Value is not able to assist its clients with maintenance management, because it has limited knowledge about maintenance management.’.

Besides the wish of Supply Value to expend its maintenance knowledge, there is also a gap in the literature of maintenance management. In the current literature on maintenance management, various maintenance concepts have been described. The aim of a maintenance concept is to determine per component in a system which maintenance type (corrective maintenance, conditional maintenance, etc.) is the best fit. The maintenance concepts in the literature mainly focus on critical systems and critical components within these critical systems. Each asset of the same system type is maintained in the same way according to the maintenance concepts.

When many assets of the same system type are maintained, these assets may not require maintenance to the same extent. For example, when a Company manages thousands of wind turbines, they will notice differences in the performance between these thousands of wind turbines. Some of these wind turbines will have failures significantly more often than average and some of these wind turbines will almost never experience failures. In this situation it may be useful to categorise wind turbines according to the number of malfunctions and define maintenance plans for each category. From now on we will call this asset categorisation. In this thesis quantitative data of Company A will be used to execute the categorisation of assets. Two categories will be distinguished good condition and poor condition. We determine how assets are categorised and the best fitting maintenance type per category. With maintenance type the choice between corrective maintenance (CM) and preventive maintenance (PM) is meant. The idea of categorisation is that by focussing more on the assets in the poor condition category, these assets will shift toward good condition category. While the assets of the good condition category will stay in the good condition category. Besides, the most impact is made by lowering the number of failures of assets within the poor condition category.

The central research question, belonging to the problem context is: ‘How can a maintenance strategy be improved, by implementing asset categorisation?’. By answering this question Supply Value will have broader knowledge about maintenance management and maintenance strategies. The answer will also support Company A with improving its maintenance strategy.

1.3 Methodology

The methodology followed in this master thesis is the CRISP-DM methodology. In this section a more in-depth explanation of the CRISP-DM is given. In section 1.6 is explained how the thesis chapters relate to the CRISP-DM cycle. CRISP-DM is an abbreviation of Cross Industry Standard Process for Data Mining. It is a methodology created to increase the success rate of datamining projects. The methodology spans the entire lifecycle of a data mining project. Figure 2 shows the CRISP-DM cycle:

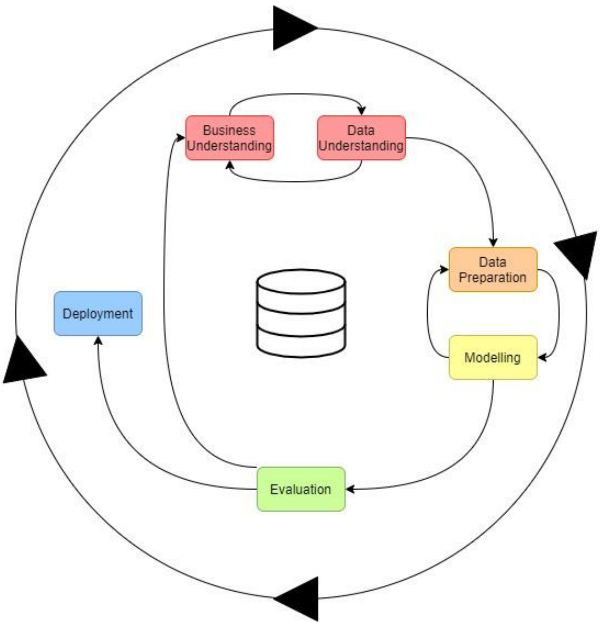


Figure 2 CRISP-DM cycle

CRISP-DM is an iterative cycle which has subcycles within the cycle, the phases will be further elaborated on based on, Wirth & Hipp (2000), Nadali, et al. (2011), Olegas (2015) and Huber, et al. (2019). The initial phase is **business understanding** where the objectives and requirements of the data analysis project are the focus. The goal of the project including the success criteria are also mapped out in this phase. The second phase is **data understanding** the initial data collection will be the first step in this phase. When getting familiar with the available data it can be necessary to rephrase the objectives, requirements, success criteria and the goal of the business understanding phase. The second step in this phase is to determine which data is possibly interesting for future phases.

The third phase is **data preparation** which focusses on preparing the data to create a final dataset. Actions that can be taken in this phase are cleaning the raw data, examples of cleaning are excluding noise and removing duplicate entries. Also new attributes can be constructed in this phase. This phase has a subcycle with the fourth phase **modelling**. In the modelling phase various modelling techniques and / or algorithms are used to construct a solution to reach the goal set in the business understanding phase. When modelling it can occur that one finds out they need to construct new data, thus they will go to the third phase again.

The fifth phase is **evaluation** in this phase one or more high quality models are already built, which will be evaluated in this phase. The model itself as well as the steps carried out to construct the model(s) are compared to the objectives and requirements. The last step of this phase is the decision how the results of the data mining should be used, whether they are applicable in a practical setting.

The last phase is **deployment** the main goal of this phase is to transform the results of the project into a useful deliverable. This deliverable varies based on the set requirements. A simple form of this phase is a report. A higher level is the application of live model(s), based on the created models in phase four, in a Company. In this phase it is important that the results are transformed in such a way that the end user can use the results. This report will be the execution of this phase in the data analysis.

1.4 Research questions

To structure the thesis research questions are composed. This section further elaborates on these questions and with what tools they will be solved.

The central research question is as stated in section 1.2: 'How can a maintenance strategy be improved, by implementing asset categorisation?'. Sub research questions per chapter are composed which aim to together answer the central research question.

At first the current situation of Supply Value and of the client who provides the quantitative data will be determined. The research questions of chapter 2 are:

2. Current Situation

- 2.1. What is the current knowledge of Supply Value concerning maintenance of assets?
- 2.2. How is maintenance currently planned and executed at Company A?
- 2.3. How is maintenance planned and executed at other companies?

The research questions will be answered using the following tools: Observation, meetings, the study of whitepapers, cases and articles published by Supply Values and the companies A to G.

Chapter 3 will answer knowledge questions concerning the main areas of the research, maintenance management and data analysis. The research questions of chapter 3 are:

3. Literature Review

- 3.1. What is known about maintenance strategies in literature? What are the maintenance frameworks and maintenance types?
- 3.2. How does maintenance as a business model influence the way a Company is structured?
- 3.3. How is data mining executed and what type of data mining techniques are described in literature?

The research questions will be answered through a literature study. The sources used will be of a high quality and up-to-date.

In chapter 4 the data of the client of Supply Value will be analysed. The research questions concerning the analysis are:

4. Data Analysis

4.1. What is the data analysis objective? What are the success criteria for achieving the data analysis objective? And what data is available to reach the data analysis objective?

4.2. Which features must be selected for the analysis?

4.3. – 4.4 How can assets be categorised?

4.5. – 4.6 What is the relationship between categories of tickets?

The data analysis will be executed using Excel. Employees of Company A who are specialized in asset management will assist with the analysis process.

In chapter 5 possible solutions for Supply Value's client are designed and tested. Possible solutions for Supply Value are designed as well.

5. Solution Design

5.1. How can the maintenance quality of Company A be improved using asset performance and conditional probability of cause groups?

5.2. How is the data analysis validated? What are the results of the validation?

5.3. How is the framework structured which allows Supply Value to apply the gained knowledge concerning maintenance management?

The answers to the research questions will be obtained from the outcomes of the data analysis and through meetings with experts from Supply Value and Company A.

1.5 Research scope

The scope of the research is limited. Even though all phases of the CRISP-DM cycle will be executed. The evaluation will not be an evaluation of the solution at Company A, but a simulation to validate the solution. This choice is made due to limited time available. To implement the solution at Company A, half a year is too short. If Company A wants to implement the solution, they it is recommended that Company A follows the CRISP-DM cycle again, using the lessons learned from this thesis.

1.6 Report structure

Figure 3 shows the general report structure, where the chapters are connected to the CRISP-DM phases.

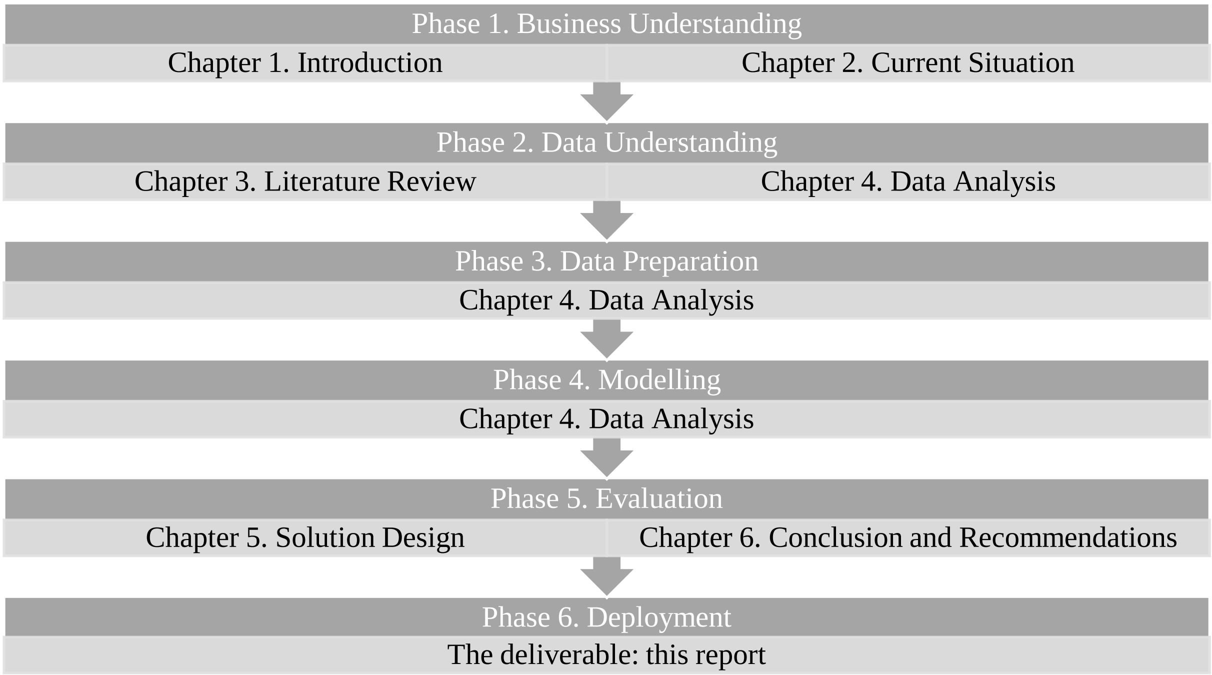


Figure 3 Report structure linked to CRISP-DM cycle

2. Current situation

The current situation concerning maintenance management will be further elaborated on in this chapter. Section 2.1 discusses the current situation at Supply Value, which knowledge about maintenance management they already have. Section 2.2 further elaborates on the current situation of Company A that provides the quantitative data. Section 2.3 describes the current situation concerning maintenance management at other companies and the problems they encounter involving maintenance management.

2.1 Supply Value

As stated in section 1.2, Supply Value has limited knowledge about maintenance management at the moment. In April 2020 Supply Value published a white paper, (Supply Value, 2020), about its knowledge concerning predictive maintenance. A white paper is an in-depth article, which addresses problems and solutions related to a specific subject. Via white papers companies can show their knowledge about a subject and how they can support clients who have similar problems, as the problems described in the paper. In November 2020 Supply Value published an insight on its website. In this insight Supply Value’s knowledge concerning the creation and improvement of a maintenance strategy is shown. Since these two publications are the only knowledge concerning maintenance available at Supply Value these will be explained in this section. Preferably projects executed would be analysed but that is not possible. Thus, the white paper and insight will be explained and analysed to determine the current state of knowledge at Supply Value concerning maintenance management. First the content of the white paper is further elaborated on, at second the content of the insight is further elaborated on. At last, the current situation concerning maintenance of Supply Value is determined.

In this white paper four levels of maturity within predictive maintenance are distinguished. The fourth level represents the highest level of maturity. In level 1 the maintenance intervals are determined based on the expertise of employees. Level 2 combines expertise with measured data. In this level there is so little data measured yet that conclusions solely based on the data are not possible. In level 3 assets are continuously monitored with predetermined critical levels, which trigger alarms. In level 4 the assets are continuously monitored, and critical levels are continuously determined via data analysis, which lead to pure data driven maintenance management. Figure 4 shows an overview of the four levels including how the maintenance strategy is determined. Note that the maturity levels of Supply Value differ from the generations of Table 1, in this thesis we constantly place companies in the generations explained in Table 1.

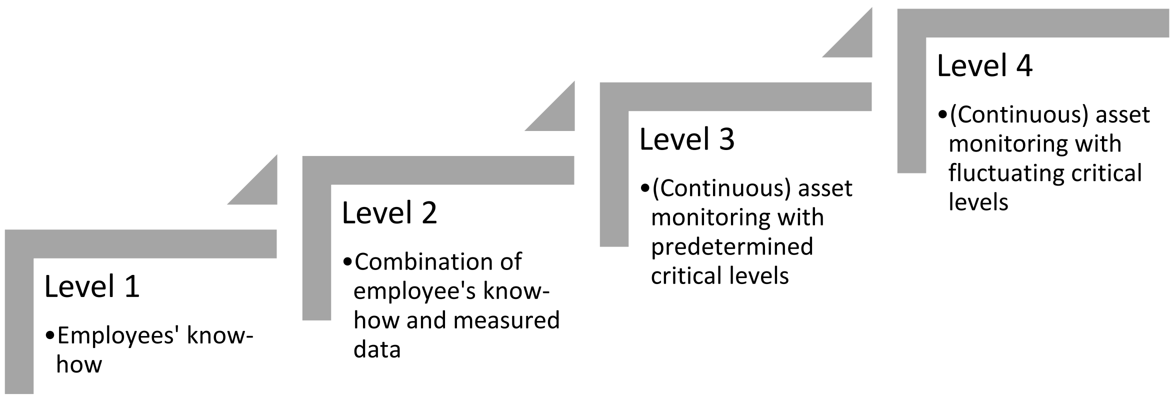


Figure 4 Maturity levels of predictive maintenance

The challenges and possible advantages of implementing predictive maintenance, that Supply Value found are as follows. First the main benefits are the potential cost reduction, the synergy of data and employee expertise, possible prolongation of the asset's lifespan, and an improved resource planning. The main challenge Supply Value found was the initial investment. A Company must invest in a system that supports predictive maintenance, sensors to trustfully measure assets and training of employees. The implementation of predictive maintenance is only worth it if the benefits of predictive maintenance are higher than the initial investment.

Since the implementation of predictive maintenance affects multiple stakeholders within an organisation, Supply Value recommends involving these different stakeholders throughout the whole project. They do not only involve the stakeholders who are directly affected by the change but also those who are indirectly affected.

To successfully integrate preventive maintenance into the existing processes of a Company, Supply Value describes a step-by-step framework. The steps are as follows:

1. Create a strategy.
2. Determine which roles and responsibilities are involved in the integration and assign them to employees.
3. Data collection.
4. Data analysis and modelling predicting models.
5. Asset selection.
6. Determine threshold values and apply them in the predictive maintenance system.
7. Build feedback loops within the predictive maintenance system.
8. Expand the predictive maintenance system to other assets.

Supply Value often uses maturity levels, this to guide clients how well the client is performing compared to similar companies. When looking at the white paper this way, it can be useful for clients to see possible future steps to improve the maintenance strategy. When comparing these levels of maturity to the four generations described in chapter 1, Supply Value's description is more abstract. The four generation link maintenance types such as CBM and reliability bases maintenance to a generation. Where in the maturity levels of Supply Value the levels are not linked to preventive maintenance types. It is called preventive maintenance in the white paper, but the levels are linked to CBM, which is a category within preventive maintenance. Supply Value however also uses perspectives which are less often described in literature. They for example emphasise on including stakeholders throughout the whole project. This is a perspective not looked at in the four generations of chapter 1, but which does have a big impact on the adaption of a new maintenance strategy.

During this research Supply Value published an article about predictive maintenance, Supply Value (2020)[1]. The focus of this article is the overall equipment effectiveness (OEE). The OEE gives insight in the assets of a Company and how they perform. The OEE is divided into three categories and each category has two key performance indicators (KPI's) belonging to the category. These six KPI's are known as the six big losses, Table 2 shows the categories and six big losses. The use of OEE and the six big losses are often described in papers about Total productivity maintenance (TPM), including in Dal, et al. (2000) and Almeanazel (2010). In this section we further elaborate on the explanation of the OEE and six big losses connected to the maintenance strategy of Supply Value.

Table 2 The six big losses

Category	**Six big losses**
Availability	Unplanned stops
	Changeover times, adjustments, and planned stops.
Performance	Idling and small stops.
	Reduced production speed.
Quality	Start-up rejects.
	Defects and reworks during production.

A poor functioning maintenance strategy may lead to a high number of defects and reworks, many unplanned stops, etcetera. This impacts the OEE negatively. The OEE is calculated as follows:

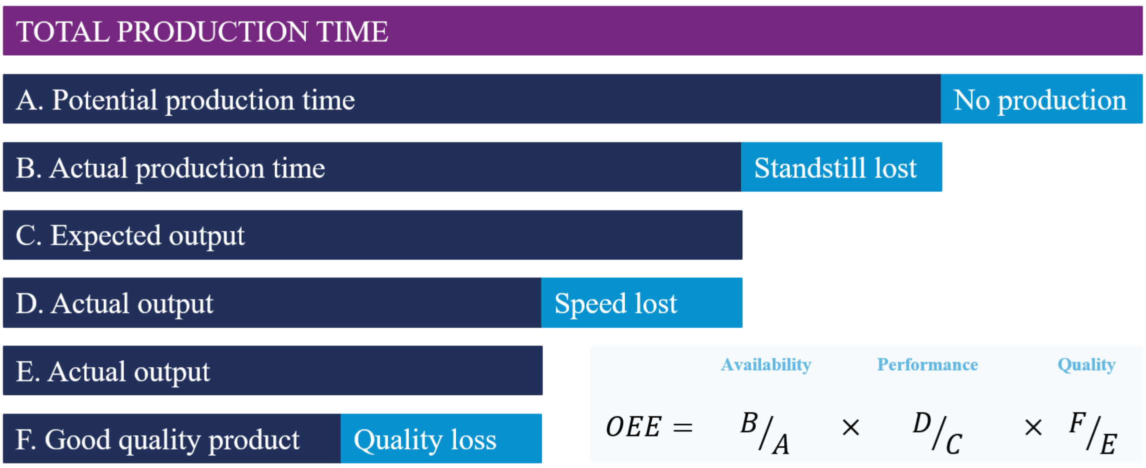


Figure 5 OEE calculation

Supply Value then explains the link between the six big losses and corrective / preventive maintenance that they have found in literature, Table 3. The green colour means it reduces / minimizes the loss, the orange colour means it does not impact the loss and the red colour means is increases the loss.

Table 3 The impact on the six big losses

Six big losses	Corrective maintenance	Preventive maintenance
Unplanned stops	Increases loss	Reduces loss
Changeover times, adjustments, and planned stops	Increases loss	Reduces loss
Idling and small stops	Increases loss	Does not impact loss
Reduced production speed	Does not impact loss	Reduces loss
Start-up rejects	Does not impact loss	Does not impact loss
Defects and reworks during production	Does not impact loss	Reduces loss

Table 3 shows that, according to Supply Value, CM does not reduce the losses, while preventive maintenance has a positive or no impact on the losses. With positive impact is meant that it reduces the loss.

Thus, assuming that Table 3 is correct, it is valuable for a Company to switch from CM to PM when aiming for an increase of OEE. Supply Value created a tool to support the implementation and optimization of preventive maintenance. This tool determines the optimal degradation value M of an asset. When the asset reaches this value M, preventive maintenance should be executed. To determine M, historical data is extrapolated over a significant time horizon to improve the reliability of the analysis. To execute the analysis, Supply Value requires high quality data of at least three months for the following three features:

1. Condition of the asset;
2. Timestamps of the corrective and preventive maintenance executed;
3. Costs of both the corrective as well as the preventive maintenance.

With this historical data the following characteristics of the asset are determined by the Supply Value simulation tool; Asset failures, the effect of preventive maintenance, variation in degradation levels and the ratio between corrective maintenance costs and preventive maintenance costs. These characteristics impact the optimal degradation value M and the average hourly costs of the asset’s lifespan. In the analysis of the simulation tool, the average costs per time (ACPT) unit are determined, for each value M. The M with the lowest ACPT is the optimal value M.

This simulation tool of Supply Value gives some advice for the maintenance strategy. Supply Value does states that it should not be viewed as decisive, since it does not include other Company specific KPI’s, insights and asset characteristics which are not considered. It does give insight into the current maintenance strategy performance. If the current strategy is far from optimal, according to the simulation tool, it could be useful to improve the strategy.

This insight is based on theoretical knowledge found in literature. Thus far Supply Value has not implemented this tool in practise. It is worth noting that the insight tool is generic and simplistic. It may help companies to see the possible advantages of shifting towards preventive maintenance, but it is not decisive.

The white paper and article give a good overview of the current knowledge about maintenance management within Supply Value. Supply Value understands the general concept of maintenance management, the theoretical knowledge of Supply Value can be place in the third generation of Table 1. Practical knowledge is missing. Also are both the insight and the white paper generic. This is logical from a consultancy perspective, since Supply Value wants to appeal to a broad public. The disadvantage of this generic knowledge is the unclearness whether Supply Value is capable to help companies with Company specific maintenance problems. This thesis will contribute to the knowledge of Supply Value concerning maintenance management and how the maintenance within a Company can be improved when dealing with Company specific problems.

2.2 Company A

Company A supplies qualitative and quantitative data for this research, the quantitative data consist of 2,5 years of ticket data, this is explained later in this paragraph. Company A is specialized in infrastructure management and asset management. This thesis connects to the asset management branch within Company A. The data used in this thesis is data about asset malfunctions, which is referred to as *ticket data*, for assets which are operated by a client of Company A. When an asset malfunctions an alarm is triggered, which indicates why the asset malfunctions. This triggered alarm is then converted in the ERP system into a ticket. When converted to a ticket a service employee of Company A will go to the malfunctioned asset and repair it. When the asset is repaired the ticket in the system is updated, it will then include the real time the asset is functioning again. This determines if the service level agreement with the client is met. For each malfunction type, a service level agreement states the maximum allowed downtime. Thus, the current maintenance type is corrective maintenance. The assets maintained by Company A are locate all over the Netherlands. They currently maintain over 13,000 assets.

There are multiple triggers that can cause an alarm set off. These triggers vary from high temperatures to a non-working battery. High temperatures are often caused by the malfunction of mechanical components such as clogged filters, which cannot be monitored with sensors. Thus, these triggers indirectly indicate a malfunction. When a non-working battery triggers an alarm, the trigger directly shows the malfunctioning component.

The malfunction causes can be divided into a passive and active malfunctions. A passive malfunctioning cause is a malfunction of a component that can function without an energy supply. An example is a malfunction of a door. An active malfunction is a malfunction on a component that needs an energy supply to function, such as a battery. Of 8% of all tickets, it cannot be traced back what caused the ticket. Normally the cause can be found in the description of the ticket, in this case the description is unclear. These tickets are therefore labelled unknown. Company A decided that they will be labelled as an active malfunction.

As said Company A currently executes CM. Sometimes an expert from Company A will create a ticket without a triggered alarm, when they think it is necessary to visit an asset. When placing Company A on the four generations explained in chapter 1 they fall between the first and second generation. As in the first generation they mainly execute CM. The creation of tickets by experts can be seen as simple PM, that is why Company A is between the first two generations and not solely in the first generation.

Company A wants to expand the maintenance types it carries out for its client. Besides the CM it offers at the moment, Company A wants to include CBM. This maintenance type expansion will lead to a different business model for Company A. Company A wants to shift towards CBM because of the following reasons:

1. Company A experiences high fluctuation in workload, resulting from when maintenance is needed.
    a. At peak times this leads to staff shortage.
    b. At lows staff who do not have tasks still need to be paid.
2. High number of repeat outages

Company A wants to use CBM to advance its maintenance activities. The slide of Figure 6 from Akkermans (2020) is used to determine how this thesis will contribute to the transition towards CBM for Company A. This thesis will contribute to the failure data analysis. For CBM much knowledge about the cause of the failure is needed and about how failures can influence future failures. At the moment Company A lack this knowledge. This is a bottleneck for the thesis. To overcome this, the relation between consecutive failures will be determined as good as possible. In this case it means that the conditional probability between consecutive failures will be determined. Conditional probability will give insight whether there is a dependency between failures and what the probability between consecutive failures is. It will not prove causality between failures, causality will be out of the scope of this thesis. What conditional probability is will be further explained in chapter 3 the literature review.

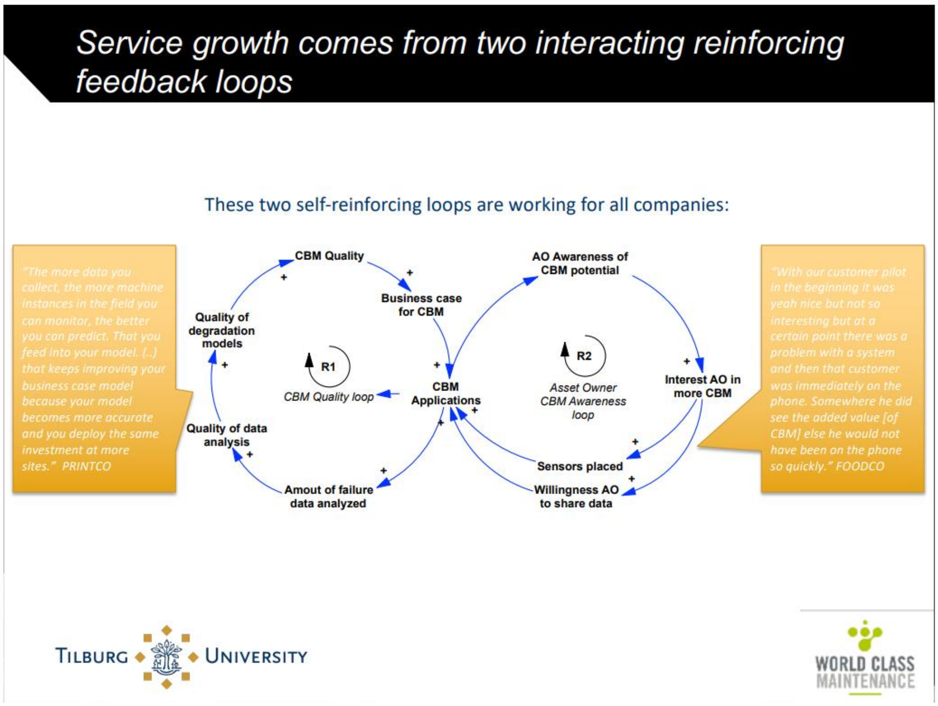


Figure 6 Two self-reinforcing feedback loops, copied from "Business models for CBM-driven smart services" by Akkermans, H, 2020.

To correctly interpret this thesis company A will be placed in the matrix shown in Figure 7. The matrix explains the type of assets that Company A deals with. This matrix shows on the x-axis the asset volume (number of assets) and on the y-axis the asset value (costs per asset). This should be noted, since companies that do not fall into the same quadrant, may not benefit from the same solutions as Company A.

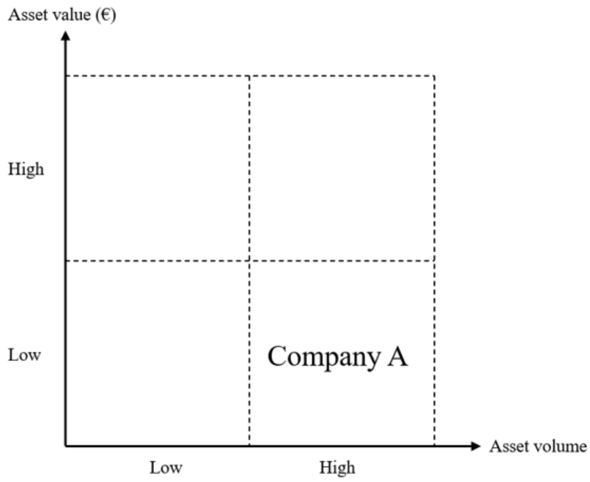


Figure 7 Asset matrix Company A

2.3 Other companies

In section 2.1 it can be concluded that Supply Value is missing insight into the maintenance management of companies. As a consultancy Company it is valuable to know the status quo concerning maintenance management throughout various industries. To get a better idea about the current situation of maintenance management within other companies, the maintenance within six companies in multiple industries are observed. The main goal of these observations is to determine the reality within companies instead of the theoretical reality in literature. In this section the results of these observations are discussed. Table 4 gives an overview of the companies, the definition of the Company size is based on Barahona, et al. (2015).

Table 4 Companies B-G overview

Company	Company size	Number of employees	Sector
B	Medium	251-500	Manufacturing
C	Large	501-1000	Public
D	Large	501-1000	Food
E	Medium	251-500	Manufacturing
F	Enterprise	1001 or more	Transport
G	Enterprise	1001 or more	Transport

2.3.1 Current situation at the six companies

First the current situation is determined, compared to the four generations explained in chapter 1. The six companies are currently not all in the same generation, when placed in the generation model of Figure 1 and Table 1.Where in these generations the companies will be scaled will be explained going from the first generation to the fourth generation. The companies will be addressed as Company B – G, to create clarity what challenges arise with what type of maintenance management strategy. Figure 8 visualizes in which generation every company is, further in this section a more in-depth explanation per company is given.

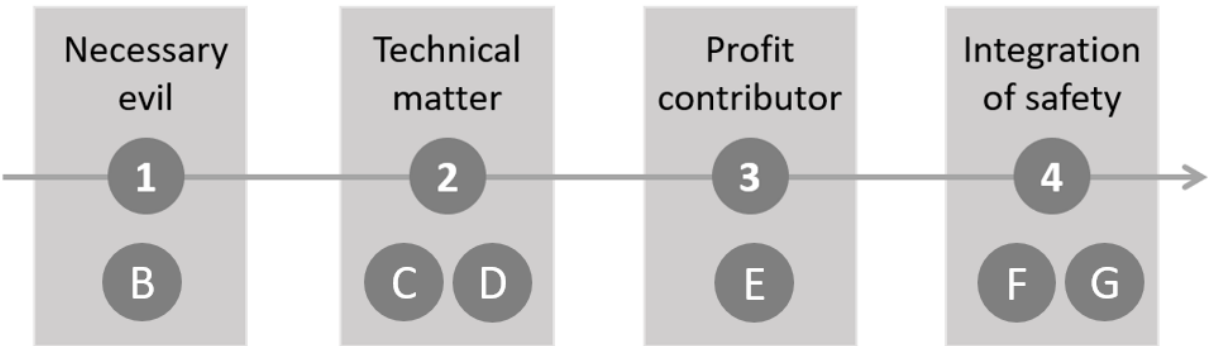


Figure 8 Generation placement for companies B-G

Company B executes only CM when their clients ask for maintenance. Since they maintain assets of their clients and they maintain a large number of asset types, they do not have access to asset data. This makes it more difficult to implement any type of PM. Thus, they are still in the first generation.

Company C currently executes CM and TBM, thus can be placed in the second generation. Currently they are improving their maintenance management within the Company. Their short-term goal is to improve the TBM and to roll out a maintenance plan suitable for the whole Company.

Company D can be placed in the second generation but does apply reliability centred maintenance, which belongs to the third generation and beyond. Compared to Company C, Company D is further developed. Where Company C is focussed on improving aspects of the second generation, Company D is more focussed to move towards the third generation.

Company E executes the maintenance of clients' asset, similar to Company B. The main difference between companies B and E is, Company E is able to collect data where Company B is not able to collect data. The ability to collect data results in, that Company E will execute CBM when enough data is available. When not enough data is available or when clients prefer cheaper but less reliable maintenance CM or TBM is executed. Since Company E does execute CBM they are placed in the third generation.

Company F and Company G apply all aspects of the fourth generation from Table 1. Company F outsources their maintenance, they determine the risk profile of assets and the contractors determine the maintenance type suitable. Company F does execute the data analysis of their assets themselves. Company G maintenance their assets themselves and is constantly researching how to improve their maintenance management. Company G implemented real time monitoring of their assets to executed maintenance on time when needed.

As said in the introduction of this section this section is focussed on the reality instead of the theory, but we also looked at the link between the reality within companies and the literature. The companies are already place in theoretical generations. Now is looked if the companies use the state-of-the-art knowledge in the literature concerning maintenance or make decisions solely on expertise of employees.

Company B and C do not implement the state-of-the-art knowledge available but make decisions solely on expertise. The other companies use a combination of expertise and literature. The difference if and how companies use the literature is big. Where Company B does not use literature since their products are too different from anything described in literature, does Company F carry out their own research and thus as they say create their own literature. What is worth noting is that companies who use little to no literature knowledge said their employees have a more practical mindset and focus more on their own expertise. While Company F stated that many of their employees have an academic degree and have high interest in applying the literature to the real world and creating new literature.

2.3.2 Challenges for the six companies

Within these current situations multiple challenges arise these are partially Company specific and partially general. First the Company specific challenges will be discussed then the general challenges.

Company B does not have data at the moment, since the data is collected and stored at their clients. They want to implement PM but need the data of their clients to do so. The clients do not want to share data at the moment, thus implementing PM is not possible at the moment.

Company F, the Company that outsources their maintenance noticed that the length of contract can be a bottleneck for new innovations on the contractors' side. Since within the contract length it is often not possible to earn back the initial investment on the innovations. Thus, contractors are often focussed on the best maintenance for the duration of the contract. The missed opportunity for Company F within the current structure is that new innovations and other maintenance strategies may extent the lifespan of the asset. To overcome this challenge the expected lifespan of the asset should be included when choosing the contractor.

The first general challenge that most companies experience is the reticent attitude, concerning data sharing, of clients, suppliers, and competitor. This reticent attitude holds back the improvement of maintenance strategies. Since the more high-quality data available the better the predictions concerning this data will be. Companies share little to no data currently since they are afraid of losing their competitive advantage to others. One problem that arises because of this attitude is that many suppliers of components and assets created their own cloud-based solutions to analyse the data of their products. Thus, it can be impossible to connect output data of multiple components within one asset, which complicates applying PM on that asset.

The second general challenge is the shortage of data scientists and technicians. When implementing and improving CBM data scientists are needed who are able to correctly analyse the output data of the assets. To physically maintain the assets technicians are needed. In the Netherlands there is a growing shortage of technicians. Both data professionals and technicians are needed to deliver high level maintenance.

The third general challenge is that people within a Company have different ideas about the best maintenance strategy, these ideas are often diametrically opposed. It can be that the board of directors do not see the advantages of investing in sensors and cloud-based maintenance system to improve the maintenance level, while the maintenance manager sees only disadvantages if they do not invest. Or the maintenance manager wants to invest in sensors and the technicians think the sensors are not as reliable as their expertise about the assets.

The fourth general challenge is related to the Covid-19 pandemic that is currently happening. Maintaining assets of clients in foreign countries is extremely hard at the moment. Besides the normal hours it takes to travel to the client abroad, maintaining the assets and travelling back, the quarantine rules in countries worldwide add at least two weeks extra. Meaning that a technician is not deployable for the time he or she is in quarantine.

In short, the following challenges are found:

- The lack of data sharing can impede the shift towards PM;
- The short length of contracts can impede innovations;
- The different cloud-based solution offered cannot be connected with each other;
- There is a shortage of data scientists and technicians, impeding the shift towards PM.
- People within a Company have different ideas about the best maintenance strategy, these ideas are often diametrically opposed;
- The Covid-19 pandemic hinders the execution of maintenance for clients located in a different country than the country of residence of the Company.

2.3.3 Foreseen Future for the six companies

The future concerning maintenance management that the companies foresee will be further elaborated on. To overcome the shortage of technicians trials already have started in which multiple employees who are not trained technicians are supervised by a technician who tells them how to execute the maintenance. In this trial the technician does not have to travel from one place to another but watches via a video connection. This technique would also be useful to overcome the fourth general challenge concerning traveling abroad during the covid-19 pandemic.

The companies do not agree whether pure predictive maintenance will be the future. With pure predictive maintenance they mean that an algorithm will notify when a component of an asset must be maintained. Meaning that when the notification arises the component must be maintained as soon as possible. The bottleneck that can arise when maintaining assets pure predictive is, that the maintenance schedule with pure predictive maintenance will interfere with current priorities, such as minimizing the disturbance of processes when executing maintenance. The more volatile maintenance schedule also may lead to higher costs due to possible peak maintenance moments and the lack of clustering possibilities. The first trade-off that belongs to this future is between the increase of reliability and lifespan of assets and components on one hand and the more difficult maintenance scheduling and increasing maintenance costs on the other hand. The second trade-off that must be made if the extra knowledge collected with new sensors will be worth the investment. The third trade-off is between retaining insight in how choices are made or giving an algorithm all the knowledge and power to make choice, thus losing insight about why certain choices are made.

Five of the companies do agree that applying more and more CBM is the future. To implement this successfully more data must be collected and applied correctly. One Company stated that how successful CBM will be implemented is dependent whether the whole supply chain is willing to share the data needed for CBM. Another Company transformed this idea to the idea that not only the whole supply chain, but all suppliers of components and assets should work together creating one cloud solution suitable for all components and assets. Company B did not see CBM as the foreseen future, because of the lack of data. Their future is focussed on creating assets with a high reliability which are extremely robust.

2.4 Conclusion

This chapter gave an overview of the current situation of Supply Value, Company A and Companies B-G. This section will give a conclusion, about Supply Value and Company A, what the current situation with its belonging bottlenecks are. Also is explained how the results from section 2.3 will be used.

The current situation of Supply Value is that they have knowledge about maintenance management and maintenance strategies, but this knowledge is very limited and generic. Their main bottleneck is the lack of experience with projects about maintenance at clients.

The current situation of Company A is that they only perform CM at the moment. The bottlenecks that arise with this maintenance type are:

1. High fluctuation in workload, resulting from when maintenance is needed.
    a. At peak times this leads to staff shortage.
    b. At lows staff who do not have tasks still need to be paid.
2. High number of repeat outages.

In section 2.2 a matrix was created to show what type of assets Company A has. The current situation of Companies B-G show that other companies fall into other quadrants of the matrix. Figure 9 shows in which quadrant the other companies fall. It is important to keep this in mind when executing a maintenance project, that the proposed solution in this thesis most likely will not be applicable for Companies B-G. This due to the fact that the value per asset is high, in contrary to Company A.

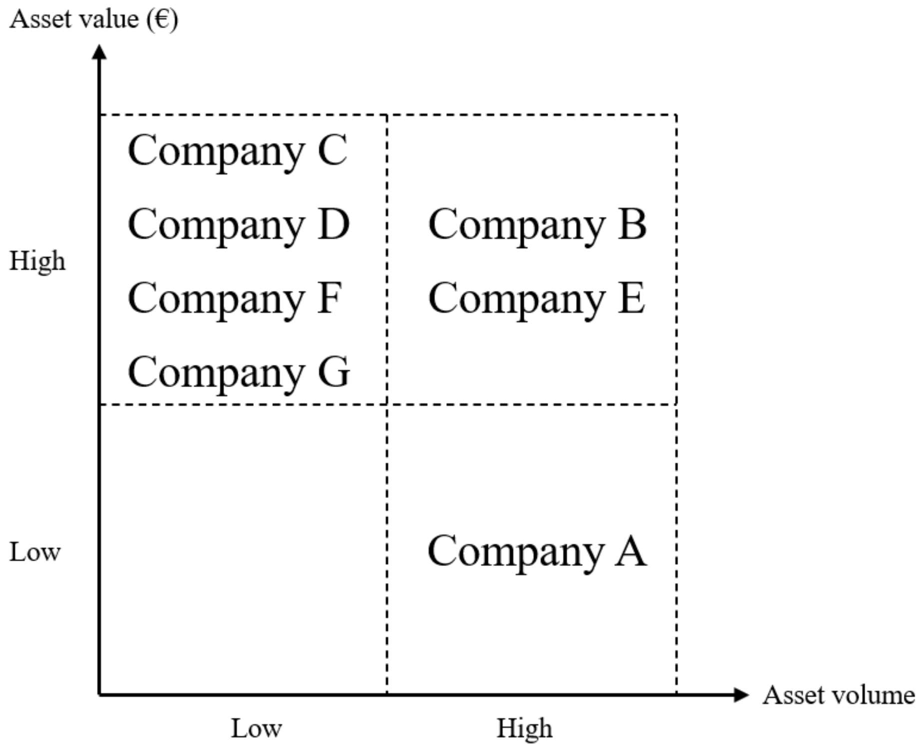


Figure 9 Asset matrix Companies A-G

In chapter 5, the solution design the current situation of Companies B-G will be used to help to shape the solution for Supply Value. Furthermore, in chapter 3, the literature review, there will be a section about maintenance as a business model. What becomes clear is that the companies that execute maintenance for clients (Company A, Company B and E), often experience a lack of data sharing from their clients which limits the maintenance possibilities. With the section of maintenance as a business model is looked at how a business can be shaped in such a way that the communication of data between the client and company can be encouraged.

3. Literature Review

The literature review aims to support the data analysis of Company A as well as providing more insight about maintenance for Supply Value. In section 3.1 the focus is on maintenance strategies; How can a maintenance concept framework support businesses implementing maintenance? (3.1.1); Which types of maintenance exist (3.1.2); And a more in-depth review into condition-based maintenance (3.1.3 – 3.1.4).

In section 3.2 the focus is on maintenance as a business model; What is servitization, a business type which has maintenance as one of the main focusses (3.2.1); And what type of contracts are often used for businesses that have maintenance as one of their main focusses (3.2.2).

In section 3.3 the focus is on supporting the data analysis executed in chapter 4; What types of data mining frameworks are there (3.3.1); Which types of data mining techniques can be used (3.3.2); What is conditional probability (3.3.3); And what is delay-time modelling (3.3.4).

3.1 Maintenance strategies

In this section existing maintenance concept frameworks are discussed in 3.1.1, maintenance types are research in 3.1.2 and a further elaboration on condition-based maintenance is executed in 3.1.3. At last opportunistic maintenance is explained in 3.1.4.

3.1.1 Maintenance concept framework

Garg & Deshmukh (2006) state: “A maintenance concept can be defined as the set of various maintenance interventions (corrective, preventive, condition-based, etc.) and the general structure in which these interventions are foreseen.” (P. 226). Gits (1992) describes the maintenance concept of a technical system as: “The set of rules prescribing what maintenance is required and how demand for it is activated.” (P. 217). There are multiple maintenance concept frameworks described in literature, on the following maintenance concepts will be further elaborated, RCM, Gits, CIBOCOF and the quantitative framework of Faccio, Persona & Sgarbossa.

Moubray (1997) describes RCM: “A process used to determine what must be done to ensure that any physical asset continuous to do what its users want it to do in its present operating context.” (P.7). Multiple studies have built their own framework around the RCM concept. Gupta & Mishra (2016) analysed 19 of these frameworks and divided them in three groups. From these groups group A, focused on continuous improvement, has the most overlap with this research. Thus, a general framework derived from this group is chosen for this literature review.

The framework consists of the following steps:

1. Study preparation: System selection & information collection.
2. Function Failure analysis.
3. Critical item selection.
4. FMECA.
5. Selection of maintenance interventions.
6. Determination of maintenance intervals.
7. Implementation.
8. In-service data collection and updating.

Gits (1992) designed a framework that helps designing the maintenance rules. Step 1 is acquiring technical system data and elementary requirement that should be met. With these inputs the maintenance rules are generated, step 2. The set of maintenance rules generated in step 2 together with composite requirements concerning the set are used for the evaluation of the set of rules in step 3. With the feedback from the evaluation the set of maintenance rules step 2 is repeated, where after step 3 is repeated. When there is no feedback anymore the set of maintenance rules will be used as the maintenance concept.

Waeyenbergh & Pintelon (2009) developed the CIBOCOF framework (The Dutch abbreviation of the Centre for Industrial Management Maintenance Concept Development Framework). They found that a standardized framework is not always the best choice for companies, so they developed the CIBOCOF which supports companies designing a customized maintenance concept. It is an iterative framework which means that after a cycle is completed the next cycle starts. The steps in the cycle are in chronological order: initiation, technical and functional analysis, policy decision and parameter optimization, implementation and evaluation, and feedback.

Faccio, Persona & Sgarbossa (2014), designed a framework based on the CIBOCOF the six steps are divided into three phases, see the overview below.

- Phase I: Equipment analysis
  1. Identification of Most Important Systems (MISs)
  2. Identification of Most Critical Components (MCCs)

  Suggested tools: questionnaires, pareto analysis, AHP, FME(C)A

- Phase II: Survival data collection and analysis
  3. Life-Time data collection for MCCs
  4. Reliability estimation for MCCs

  Suggested tools: For 3. Data collection sheets, producer data sheets, reliability SW. For 4. Empirical function direct to data (EFDD) and theoretical distribution research (TDR)

- Phase III: Decision making process
    5. Maintenance cost estimation for MCCs
    6. Economic evaluation of maintenance policies

    Suggested tools: For 5. Industrial accounting and reporting and for 6. Abacus decision tool and decision graphical curves.

The most common steps mentioned in the literature along with the question that should be answered per step are shown in Table 5.

Table 5 General maintenance framework

Step	Definition	Question
1	Data collection	What data should be collected and why?
2	Identify objectives and resources	What is the overall maintenance objective preference?
3	System selection and definition	What are the most important systems (MISs)?
4	Critical component identification	What are the most critical components (MCCs) in the MISs?
5	Selection of maintenance policy	What maintenance policy should be selected for each possible.
6	Maintenance optimization	What are the optimal maintenance parameters for each failure?
7	Treatment of non-critical items	How should the non-critical items be maintained?
8	Implementation and evaluation	How should the maintenance concept be implemented?
9	In service data collection and updating	How should the estimated parameters be updated?

3.1.2 Maintenance types

Maintenance can be split in two main types, corrective and preventive maintenance, (Waeyenbergh & Pintelon, 2003). Corrective maintenance (CM) means that the system will be repaired when a unit of the system is broken. With preventive maintenance (PM) a system will be maintained regular to prevent system breakdowns as much as possible, while taken costs into account.

CM can be divided into two types, repair and compensating, (Rausand & Hoyland, 2004). With repair the failed unit of the system will be repaired or replaced. With compensating the system can still operate with the failed unit then the unit will go into standby mode. An example of repair is when the batteries of a remote controller are empty, the remote controller as a system cannot function until the batteries are replaced. An example of compensating is when the indication light of the television breaks down, the television still functions. Then the only downside of not repairing is not seeing if the television is switched off or in standby mode.

PM can be divided into multiple types. Different papers distinguish different types, Table 6 shows the types per paper.

Table 6 Preventive maintenance types

Paper	Failure-Based Maintenance	Condition-Based Maintenance	Use-Based Maintenance	Design-Out Maintenance	Detection-Based Maintenance	Time-Based Maintenance
Durocher & Feldmeier (2004)		X				
Gits (1992)	X	X	X			
Muchiri et al. (2011)	X	X	X	X		X
Van Noortwijk (2009)		X				X
Pongpech & Murthy (2006)						X
Waeyenbergh & Pintelon (2003)	X	X	X	X	X	
Yang et al. (2017)		X				X
Yang et al. (2019)		X				X

The two types most mentioned are Condition-Based Maintenance (CBM) and Time-Based Maintenance (TBM). Since CBM is a large part of this research CBM will be further elaborated on in section 4.3.

Kim, Ahn & Yeo (2016), elaborates further on TBM. When, implementing TBM a periodical maintenance schedule is predetermined. To determine the period there are multiple TBM policies described in literature. Van der Heijden (2020) addresses the following TBM policies: age replacement, block replacement, block replacement under minimal repair, block replacement with hidden failure, opportunity-based age replacement.

With the age replacement a unit is repaired / replaced with CM when a failure occurs before time T (predetermined time period). When CM is not needed before time T the unit will be replaced when t (time elapsed since last maintenance) is equal to T. The block replacement policy executes PM at the end of each cycle length T. Independent of possible CM performed in that cycle. With block replacement with minimal repair the CM executed does not lead to an as-good-as new state of the unit. Like block replacement PM is executed at the end of each cycle length T. Block replacement with hidden failure, occurs when failures are only seen when maintenance is executed. Thus, only PM is used, and all hidden failures have fixed costs κ per time unit. At last, there is opportunity-based age replacement. PM of the unit takes place when an opportunity arises, such an opportunity is for example a failure of another unit in the system causes system outage. In that case PM of the current unit will not lead to extra downtime. CM takes place when the unit breaks down, (Dekker & Smeitink, 1991).

3.1.3 Condition-based maintenance

CBM uses real-time data to diagnose impending failures and prognose future equipment health, (Goyal & Pabla, 2015). Therefore, a predictive model is developed that triggers a maintenance alarm when needed. Maintenance schedules often focus on the trade-off between repair costs and inspection cost. Costly failures must be avoided, as well as unneeded maintenance. CBM can balance the cost trade-off better than other maintenance types, (Gupta & Lawsirirat, 2006).

A CBM program consists of three steps, see Figure 10.

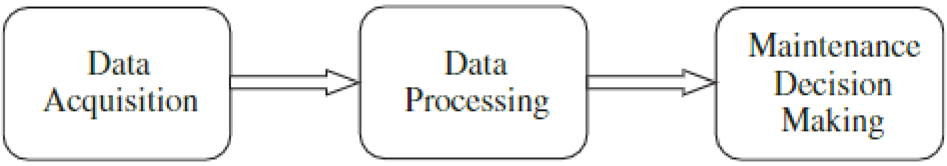


Figure 10 CBM program steps, adapted from "A review on machinery diagnostics and prognostics implementing condition-based maintenance" by Jardine, A.K.S., Lin, D., & Banjevic, D., 2006, Mechanical Systems and Signal Processing, 20(7), 1484. All rights reserved 2005 Elsevier Ltd.

The first step is data acquisition. In this step useful data is collected and stored. The type of data acquired, also known as the observables can be divided into two categories, direct and indirect data.

The second step data processing is dependent on the type of data acquired. Direct data is when the observable, e.g., a sensor, measures the estimator, for example when a sensor measures the crack length, and the crack length is also the estimator. Indirect data is when the observable needs pre-processing to determine the impact on the estimator, (Raheja et al., 2007).

In the last step the information retrieved from the data is combined. The output of this data fusion is the estimated current health of the unit / system. This estimation is then used as input for the decision policy which determines when PM is needed, (Raheja et al., 2007).

An overview of the maintenance types mentioned in 3.1.2 Maintenance types and 3.1.3 Condition-based maintenance is shown in Figure 11.

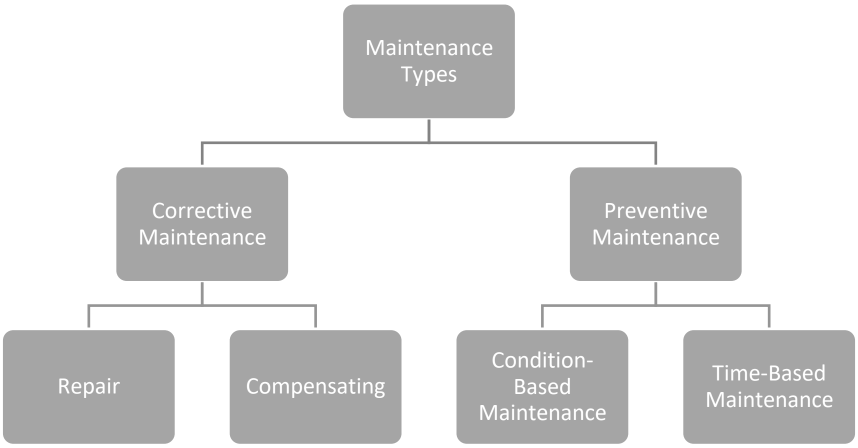


Figure 11 Maintenance types overview

3.1.4 Opportunistic maintenance

CBM as explained in section 3.1.3 can be divided into different CBM types. This section will focus on the CBM type relevant for this thesis, namely opportunistic maintenance (OM) also known as CBM/PM. The idea behind OM is, that when a failure occurs the opportunity is taken to not only perform CM but simultaneously perform PM on the other components of the asset, Ding & Tian (2012). Koochaki et al. (2011) state that where CBM focusses on a single component, OM focusses on a collection of components. Also, the aim is stated as a higher efficiency for the maintenance crew and reducing costs.

Saranga (2004) divides the OM into two categories, age related, and non-age related. In the age related category, PM is only executed on a component if that component has reached a predetermined age threshold. Otherwise, no PM is executed on that component. This category is further divided into three subcategories: hard life, before the component reaches a set age it must be replaced, soft life, a component will be replaced if it has passed a set age, degradation, the component is replaced when it crosses a critical level. In the non-age related category, PM is executed on all components which failures were undetected. Ab-Samat & Kamaruddin (2014) distinguish another category, namely for systems where the failure of a component affects other components. In that case PM is executed for those other components which are affected by the failed component.

The possible advantages of OM stated in the review of Ab-Samat, & Kamaruddin (2014) are as follows: an increase of overall reliability and production rate, and a decrease of set-up costs, total downtime due to failures and overall costs.

3.2 Maintenance as a business model

This section will elaborate further on maintenance as a business model. How this influences the Company and how it influences the contracts with customers. Section 3.4.1will explain servitization, section 3.4.2 discusses different contract types that go hand in hand with servitization.

3.2.1 Servitization

Vandermerwe & Rada (1988) are the first to describe the transformation to servitization. In the past a Company either supplied goods or services, this is described as stage one. Then with the advent of new technologies, such as computers, companies started to move towards business models that offered services and goods, which is stage two. The third stage described combines not only services and goods but also includes knowledge, self-service, and support. Knowledge is the knowhow of the supplier. Self-service is the stimulation of the customer to execute services themselves when possible, with the advantage of lower costs for the customer. Support is the support given to the customer, which ranges from training to use the good to remote maintenance systems. Whether all aspect of the third stage of servitization are offered by the supplier differs per supplier. Some supplying companies view the third stage as a bundle that must be used as a whole, while others view it as a menu from which a customer can chose.

Even though, servitization does not appear in the literature before 1988, there are practical examples of servitization before 1988. Kowalkowski, et al. (2017) compiled a timeline of the servitization and later on deservitization of companies. Which shows types of servitization that arose before 1960 such as IBM offering leasing services for their products.

The possible advantages and disadvantages of servitization will be discussed. First is looked if servitization can lead to a higher overall profit of companies. Wang, Lai & Shou (2018) describes the impact of servitization on the revenue stream of a Company. In 2015 more than half of the total revenue of Rolls Royce arose from the maintenance service of their engine products. Homburg, Fassnacht & Guenther (2003) found the positive impact on overall profitability of servitization in their research. Skaggs and Droege (2004) found positive financial results in manufacturing companies when implementing servitization. In contrary to the positive impact on overall profitability, Neely (2008) found negative financial results when implementing servitization in manufacturing companies. Negative financial results are also found in Visnjic et al. (2012).

A more diverse result is found in Sousa & Silveira (2017), they found that the type of service determines if servitization leads to more or less overall profitability. In the paper a distinction between basis service (BAS) and advance service (ADS) is made. Where BAS focusses on efficiently and effectively installing and maintaining the basic product functionality, BAS is also known as product-oriented services. ADS focusses on co-creating value with the customer, this can be done via e.g., usage training, upgrading of products and consulting. Thus, the focus is more on the relation between the customers actions and the product. The conclusion of Sousa & da Silveira (2017) is, that BAS does not positively influence overall profitability, while ADS does positively influence overall profitability. Thus, it remains unsure if servitization decreases or increases a Company's profit, since studies contradict each other.

When looking if servitization will add value to customers, Baines & Lightfoot (2013) and Rabetino, Kohtamäki & Gebauer (2017) first look at the type of customer. There are three groups of customers distinguished. The first group are customers that do it themselves, only want the goods of the manufacturer. Thus, servitization does not add value for them. The second group does want to do the small maintenance themselves but want the manufacturer to carry out bigger maintenance. For this group servitization can add value if parts of servitization such as small maintenance can be excluded. The last group wants the manufacturer to do everything for them. For this group servitization will add value. Thus, whether servitization adds value to the customer depends on the customer type.

In brief servitization is the business model of not only supplying services or goods but a combination of services, goods, knowledge, self-service, and support. Whether it is beneficial to the overall profit of a Company remains undetermined and whether it adds value to the customer depends on the customer type.

3.2.2 Contract types

There are multiple contracting types associated with servitization, in this section pay-per-use, outcome-based contracting (OBC) and performance-based contracting (PBC) will be further elaborated on.

With a pay-per-use model the customer pays for each time used, instead of paying for the ownership of the good, Bocken, et al. (2018). An example of pay-per-use products is a laundromat, where clients pay to use the washing machine. The maintenance required will not be the responsibility of the customer but of the Company offering the service, Armstrong, et al. (2015). Even though it is a contracting type fitted to servitization, it is less fitting to the maintenance compartment, since pay per use is mainly focussed on offering goods as a service.

With OBC customers only pay when the predetermined outcome, e.g., a certain service level, is delivered. Instead of paying for only tasks and activities, Ng et al. (2009). (Ng et al. (2013) and Visnjic et al. (2017), Argue that OBC is the most advanced servitization contract type. When implementing servitization through OBC a shift to service value co creation is needed, as well as purposeful relationships between the Company and customer, Batista et al. (2017).

PBC focusses on delivering a predetermined level of performance instead of focussing on delivering a certain quantity of goods, Xiang, et al. (2017). Kim et al. (2007) and Guajardo et al. (2012) use the term power by the hour for PBC and it is also seen by some (Ng et al. (2009) and Ng and Nudurupati (2010) as a narrower version of OBC. It is linked to the servitization of businesses and adds to the aspect undervalued in servitization, namely contracting relations and incentive alignment throughout the supply chain, Bastl, et al. (2012). Hypko, Tilebein

& Gleich (2010), sees a shift toward PBC throughout the services industry as well as the manufacturing industry and both in the private and public sector.

Xiang et al. (2017), studied the combination of PBC and CBM, where the reward of the maintenance provider was directly linked to the average availability. Thus, the goal of the maintenance provider was to maximize the performance will minimizing costs. The PBC research is a straightforward profit centred contract and the Company studied was a maintenance repair and overhaul (MRO) service provider. MRO service providers traditionally aim to minimize costs, which can lead to poor performance. When implementing the PBC with profit centred contracts, and rewarding a better performance, the study found that the best results are obtained by focussing most on performance instead of minimizing costs. The PBC with profit centred contracts also lead to higher motivated MRO service providers, due to the changed incentive from minimize costs to maximise performance.

Even though this literature study addresses OBP and PBC two different contracting types, within the literature they are often used interchangeable.

3.3 Data analysis

In this section is described how the data analysis should be performed. At first is looked at data mining frameworks, then data mining techniques, probability theory and delay-time modelling are discussed.

3.3.1 Data mining frameworks

Data mining can be defined as analysing and extracting knowledge or information from enormous data sets by a computer-aid process, Kalyani (2012). To assist this process multiple step-by step frameworks are described. Multiple papers are studied and the frameworks they proposed are combined into a matrix. An X indicates that the paper adressed the step, a blank means the step is not mentioned in the paper.Table 7 shows the matrix overview of which framework addresses which steps.

Kumar & Toshniwal (2015):

1. Data pre-processing
2. Clustering algorithm
3. Association rule mining algorithm

D’Oca & Hong (2015):

1. Data selection
2. Data pre-processing and cleaning
3. Data transformation
4. Data mining

5. Data interpretation and evaluation
6. Knowledge extraction (feedback loop to all previous steps)

Shen, et al. (2012), data framework for hospitals:

1. Data preparation
    a. Input is historical diagnosis database
2. Data pre-processing
3. Mining Module
    a. Output is orders sequence knowledge base
4. Clinical pathway creation system

KDD (knowledge discovery in databases) framework described by Mitra, Pal & Mitra (2002):

1. Data selection
    a. Input is raw data
2. Pre-processing
3. Transformation
4. Data mining
5. Evaluation and interpretation
    a. Output is knowledge

The data mining step uses a data mining algorithm which consists of a combination of the following components:

1. The algorithm model
    a. Focus on its function and its form (e.g. function: classification and form: neural networks).
2. The preference criterion
    a. A set of parameters which are preferred, the trade-off that must be made is that overfitting must be avoided.
3. The search algorithm
    a. An algorithm that finds specific model(s) and parameters, given available data, model(s) and prefferd criteria.

CRISP-DM, described by Wirth & Hipp (2000):

1. Business understanding
2. Data understanding
3. Data preparation
4. Modelling

5. Evaluation
6. Deployment

These steps are within a cycle, thus after step 6 it continuous with step 1. There are also subcycles between, 1 and 2, and, 3 and 4.

Table 7 Data mining frameworks overview

Paper	**Business understanding**	**Data selection**	**Pre-processing**	**Clustering algorithm**	**Data transformation**	**Mining algorithm**	**Data interpretation and evaluation**	**Knowledge extraction**
Kalyani (2012)		X	X		X	X	X	
Kumar & Toshniwal (2015)			X	X		X		
D'Oca & Hong (2015)		X	X		X	X	X	X
Shen, et al. (2012)		X	X			X		X
Mitra, Pal & Mitra (2002)		X	X		X	X	X	X
Wirth & Hipp (2000)	X	X	X			X	X	X

The most common steps in the data mining frameworks found in literature are:

1. Data selection
2. Pre-processing
3. Data transformation
4. Mining algorithm
5. Data interpretation and evaluation
6. Knowledge extraction

3.3.2 Data mining techniques

In this section will be further elaborated on the mining algorithms. The goal of this section is to describe techniques that can assist the category determination of the assets in the data analysis. This section will discuss the following techniques association, classification, clustering and prediction.

Association

With association the aim is to discover relationships between variables in a large data set, by searching for patterns. The most famous example of data association is the market basket analysis. In this analysis it is determined which products in a supermarket are most likely to be bought together, Kalyani (2012).

Huang, Wu & Relue (2002) notated the association techniques as an if-then rule which looks as follows:

X, Y	, Variables in a data set
N	, Number of transactions in a data set
$X => Y$	, X and Y both appear in a transaction
$Support\ (s) = \frac{Frequency\ (X,Y)}{N}$	, Percentage of all transactions that contain both X and Y
$Confidence\ (c) = \frac{Frequency\ (X,Y)}{Frequency(X)}$	, Percentage of transactions that contain X, that also contain Y

A rule between two variables is seen as interesting if the minimum support and confidence thresholds are met.

Classification

Han, Kamber & Pei (2012) describe data classification as a two-step method where in the first step, the learning step, an algorithm is trained to correctly classify preclassified training data. The output of the algorithm consists of a set of classifications rules. The second step, the classification step, classifies test data which is not classified yet. The data of the second set must be independent of the data of the first set to prevent overfit. Decision trees can be used to shape the classification rules of the algorithm.

Aggarwal (2015) states that most classification algorithms consist of two phases, the training phase, and the test phase. These phases are similar to the two steps described in Han, Kamber & Pei (2012). The paper warns, that if the training set is too small, the classification models may perform poor. The classification models often used are neural networks, instance-based classifier, rule-based classifiers, probabilistic models, support vector machines and decision trees.

The two-step method is also mentioned in Rebentrost, Mohseni & Lloyd (2014), Wu, et al. (1996), Dogan & Tanrikulu (2013) & Shah & Jivani (2013). An illustration of the two-step method is shown in Figure 12.

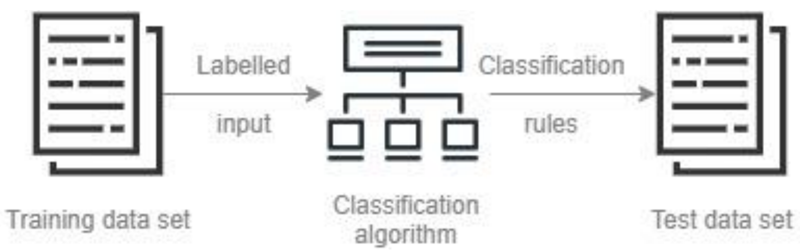


Figure 12 Two-step method

Clustering

With clustering variable are grouped in such a way that variables within a cluster (a group of variables) are more similar than variables of other groups, Kameshwaran, Malarvizhi (2014). Sajana, Rani & Narayana (2016) describe five types of clustering algorithms; partition based, density based, hierarchical based, model based, and grid based.

Partition based algorithms divide variables in a predetermined number of clusters, through iterations the best partitions are found. Types of partition-based algorithms are k-Means, k-Medoids, CLARA and CLARANS, Boomija & Phil (2008).

Density based algorithms differ from other clustering algorithms since it includes outliers. Variables are categorised in core, border, and noise points. Then based on the densities to form clusters, all the core points are connected. Types of density-based algorithms are, DBSCAN, DBCLASD, DENCLU and SUBCLU, Sajana, Rani & Narayana (2016).

Hierarchical based algorithms decompose the data set hierarchical by building a dendrogram where nodes are split into new nodes, the end nodes represent the different clusters, the dendrograms can be build bottom-up or top-down, Kameshwaran, Malarvizhi (2014).

Model based algorithms assume that the data is a result of a finite number of underlying probability distributions. Where each cluster is generated by one underlying probability distribution. The underlying probability distributions are often chosen a priori, Yeung, et al. (2001).

Grid based algorithms divides the data set into a finite number of cells, thus creating a grid. Then the cells that surpass a predetermined threshold are pin pointed and neighbouring cells which surpassed the threshold are grouped into one cluster, Lin, et al. (2008).

Prediction

Prediction techniques aim to determine the relationship between (in)dependent variables, Kalyani, et al. (2012). Hoffman & van der Westhuizen (2019) applies prediction techniques on truck data to determine which input factors can correctly estimate the fuel consumption of a trip. The paper first explores which input factor influence the fuel consumption. Secondly the input factors are divided between solely influenced by the driver, solely influenced by the route, and not clear how influenced. Thirdly the correlation between the route input factors is determined, such that the minimum number of input factors is needed to cover all input factors. Fourthly multiple linear regression analysis is used, in this step the linear model to determine fuel consumption is built.

3.3.3 Probability theory

In the data analysis part, the probability that cause B will be the next ticket cause after the current ticket with cause A is determined. To determine this the probability theory is used. This section will explain the probability theory.

In the foundations of the theory of probability, Kolmogorov describes the five axioms of probability theory. These axioms are nowadays known as the foundation of probability theory, or as Kolmogorov's approach. The approach is as follows, Kolmogorov (1960):

There is a collection of elements (ξ, η, ζ, …) known as *elementary events*. There also is a set of subsets of E called F. The elements in F are known as *random events*.

1. F is a field of sets.
2. Set E is contained in F.
3. In F each set A gets a non-negative real number $P(A)$ assigned, which represents the probability of event A.
4. $P(E)$ is equal to 1.
5. $P(A \cup B) = P(A) + P(B)$, holds only if A and B have no elements in common.

A definite assignment of numbers $P(A)$ together with F, satisfying all axioms above is known as a *field of probability*. Kolmogorov proved that the axioms are consistent.

After the axioms are explained, in Kolmogorov (1960), Kolmogorov continues with independence. Two events are considered independent from each other if one of the following equations is true:

$$P(A \cap B) = P(A) * P(B) \text{ or}$$

$$P(A|B) = P(A)$$

If events are independent the occurrence of event A does not influence the probability that event B occurs.

The probabilities described thus far are unconditional. Next the conditional probability is explained. Conditional probability is the probability that A will occur, if given that B has occurred. Hájek, A. (2011), Defines the conditional probability A→B as the ration of unconditional probabilities:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, where\ P(B) > 0$$

Díaz, Batanero & Contreras (2010), states that conditionality and causation are often confused with one another. It is important to be aware of the difference. Even though the occurrence of event B changes the probability of event A happening, this does not imply that event B causes event A. In short causation between event B and A does mean that $P(A|B) = 1$, but $P(A|B) = 1$ does not necessarily mean that event B causes A. Thus, finding a high $P(A|B)$ does not directly mean that B causes A.

3.3.4 Delay-time modelling

In chapter 5 a simulation will be built to validate the proposed solution. The simulation will be a simplification of the reality. The simulation will be based on the delay-time modelling (DTM). Wang (2012) describes DTM as a mathematical technique used for optimising inspection planning of assets. The technique is often used to optimize inspection intervals and can be seen as a simplified CBM model.

DTM consists of a two-stage process, Berrade et al, (2017). Where the first time a fault in a component can be detected is the first stage called h. The moment the asset fails or when inspection detects the fault is the second stage called u. The delay time is the time between h and u. Figure 13 shows the timeline of the two stages. If an inspection is held during the delay time, failure can be avoided. Wang (2012) distinguishes two types of inspection, imperfect and perfect. With perfect inspection is assumed that a component with a fault will always be noted when inspection, in contrast to imperfect inspection where the fault may be noted.

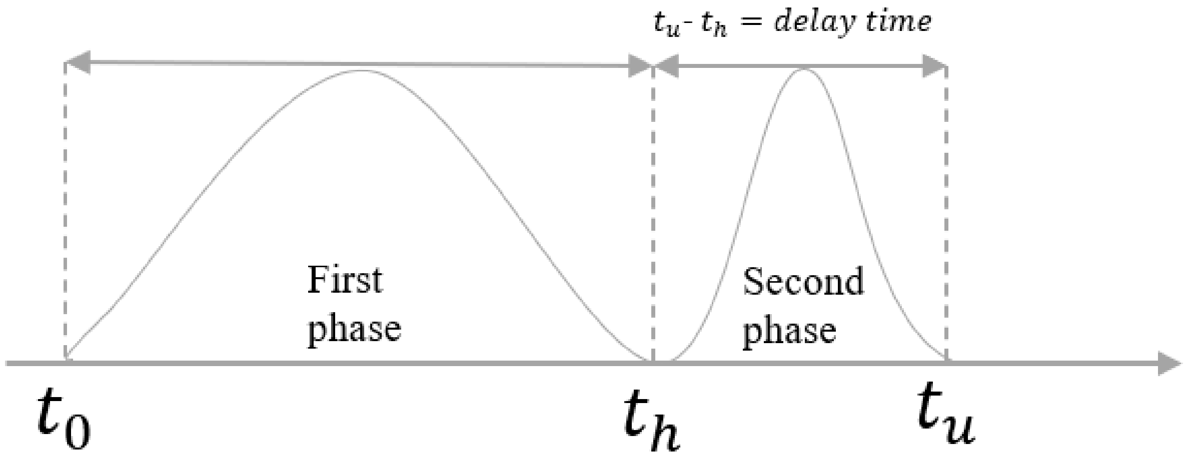


Figure 13 Delay-time model

4. Data analysis

A data analysis is set up which will answer the objective for Company A. First we will give a short overview of the results of the data analysis in this chapter. Since the data analysis consists of multiple steps which are not clearly linked to each other until the conclusion in section 4.7. To clarify the chapter we will show the link between all section in advance.

The objective of this data analysis is to improve the maintenance quality by increasing knowledge on asset categorisation and by increasing knowledge about conditional probability between failure causes of consecutive failures. Section 4.1 further elaborates on the objective. Then the data set is improved in section 4.2.

In sections 4.3 and 4.4 we determine how assets can be categorised. The assets are divided into a good condition category and a poor condition category. In which category an asset belongs is based on the total number of tickets of the last three months (this is the chosen interval length) for that asset. A threshold is set to divide the assets. This threshold is two tickets meaning that if the total number of tickets of an asset is two or higher in the last three months the asset belongs to the poor condition category. It the total number of tickets below the threshold of two then the asset belongs to the good condition category.

In sections 4.5 and 4.6 the, by Company A identified, failure causes are grouped into cause groups and the conditional probabilities between consecutive failures are determined. The grouping of causes will cluster all causes concerning the same component in the same cause group. The conditional probabilities between consecutive failure on cause group level are determined to support the maintenance actions per assets category. If OM is chosen for an asset category, the assets in that category receive CM on the cause group that fails and PM on another cause group. Which cause group should be the one to receive PM is determined by using the conditional probabilities between consecutive failures on cause group level.

Note that section 4.3, 4.4 and 4.6 include a feedback session with experts of Company A. The reason that the feedback sessions are included in the data analysis is because the data analysis is executed for another company. To make sure that the analysis execution and results are in line with the wishes of Company A these feedback sessions are planned. When executing a data analysis for the company you work at it may not be necessary to include these feedback sessions. The feedback sessions ensure that when choices must be made the choice made is preferred by the company.

Figure 14 shows the overview of the data analysis.

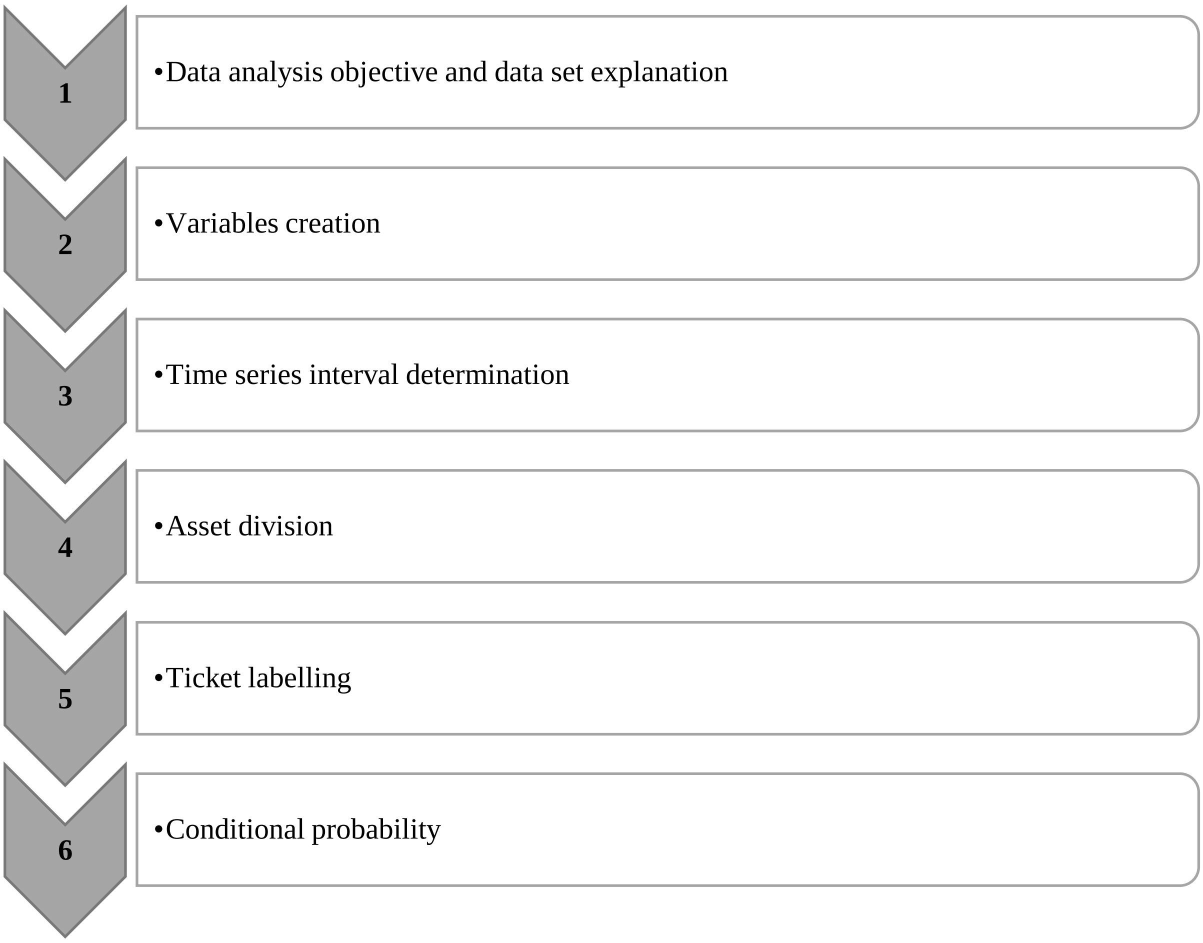


Figure 14 Overview of the chapter sections

The summary of each section is explained below:

1. The defining of the objective of the data analysis and understanding the data set given.
2. The creation of new features in the data set.
3. The determination of the time series length for the analysis.
4. The categorisation of the assets based on ticket quantity.
5. The labelling of ticket causes together with an expert of Company A.
6. The determination of the conditional probability between categories.

4.1 Data analysis objective and data set explanation

The data analysis is executed for Company A. This section of the data analysis elaborates on the business understanding and data understanding. To gain knowledge about Company A and their current challenges two meetings were scheduled. In these meetings Company A and Supply Value together determined the data analysis objective.

Since Company A is the contractor of their clients it is important for them to stay the supplier of maintenance for their clients. To do so, the goal of Company A is to constantly improve their maintenance practices. The data analysis objective belonging to this goal is to improve the maintenance quality by increasing knowledge on asset categorisation and by increasing knowledge about conditional probability between failure causes of consecutive failures.

The success criteria related to the objective are:

1. A set of one or more rules is created which determine(s) the category to which an individual asset belongs.
2. The conditional probability of all follow-up failures is identified, at cause group level.

How the success criteria relate to the objective of the data analysis and the central research question is shown in Figure 15. Figure 15 shows how same system type assets can be categorised and how the maintenance per category can be determined. The two grey squares indicate the success criteria. Note that the category determination is done per asset and the conditional probability is calculated between cause groups. Cause groups are groups that per group cover all failures regarding one component type, for example the cause group door covers all types of door failures.

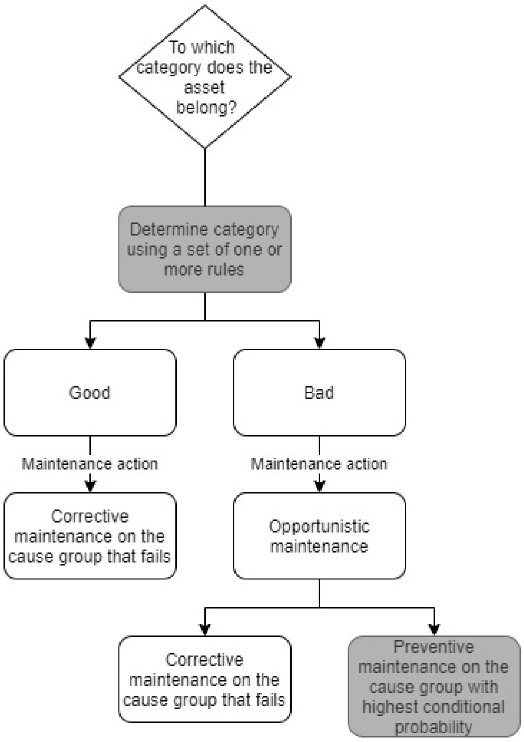


Figure 15 Success criteria related to objective and central research question

Sections 4.2 to 4.4 focus on the first success criterium, sections 4.5 and 4.6 focus on the second success criterium. The data analysis objective is determined, thus this section will continue with the data set explanation.

Company A supplied the data set used for the data analysis. This data set is ticket data from November 2017 until March 2020. The data set consists of 33704 tickets which belong to 12911 same system type assets. Note that the data set only shows the assets that had at least one ticket during the time period covered. The data set is in the form of an excel file, where each column represents one input parameter (= feature) (possible predictor) and each row an individual ticket. A feature is for instance the MCH_code of the assets which is a unique string for each asset. The features will not be mentioned directly in the analysis of this report but are used to execute data analysis. Therefore the data set features are explained to give insight the meaning of the features and how they are formatted they are explained in Table 8. Table 9 shows what features were used for the analysis and which features were not used. Furthermore Table 10 and Table 15 show which features we constructed using the existing features. Again the new constructed features are not explicitly mentioned in the analysis but are shown in the report to give a clearer insight of the analysis.

The features already in the data set are explained in Table 8:

Table 8 Data set features

Number	Feature	Format	Meaning
1	CONTRACT_NAME	String	The division within Company A which solved the ticket
2	SLA_ID	Integer	The Service Level Agreement ID that corresponds to the cause group.
3	Real_f_date	dd-mm-yy hh:mm:ss	The date and time the cause of the ticket was solved.
4	Weekday	String	When the ticket was created during the night, weekend, or week.
5	MCH_CODEtot	String	The asset’s code including the object type.
6	MCH_CODE	String	The asset’s code.
7	MCH	String	The abbreviation of the city where the asset is located.
8	MCH_OBJ_TYPE	String	The object type of the asset.
9	WOAD_ADDRESS_CITY	String	The city where the asset is located.
10	DESCRIPTION	String	The ticket description, indicating the problem.
11	WO_NO	Integer	The work order number.
12	F	Integer	A factor
13	P	Real	A factor
14	Cause	String	The cause of the ticket.

To ensure the data quality, the data is checked for duplicate rows, there were no duplicates found. Then the data was checked for empty cells. There were many empty cells in the column WOAD_ADDRESS_CITY. At this phase the choice was made to fill the cells with not specified. The same decision was made for the column DESCRIPTION. The columns F and P are filled with a 0 since there format is an integer or real.

A selection of features was made, see Table 9 for the selected features. The features that were not selected, will still remain in the original data set sheet, and will only be left out the cleaned data set sheet. This decision is made because at this point the left-out features seem irrelevant to reach the two success criteria. When they are needed after all, they can be added to the cleaned data set. By creating a clean data set the analysis will stay more structured. The following features were chosen:

Table 9 Chosen features from data set

Number	Feature	Format	Meaning
3	Real_f_date	dd-mm-yy hh:mm:ss	The date and time the cause of the ticket was solved.
4	Weekday	String	When the ticket was created during the night, weekend, or week.
6	MCH_CODE	String	The asset's code.
8	MCH_OBJ_TYPE	String	The object type of the asset.
9	WOAD_ADDRESS_CITY	String	The city where the asset is located.
10	DESCRIPTION	String	The ticket description, indicating the problem.
11	WO_NO	Integer	The work order number.
14	Cause	String	The cause of the ticket.

4.2 Feature construction

In this section the features missing in the data are constructed. From the existing features 10 new features are constructed to support the data analysis, these new features are explained in Table 10:

Table 10 New features in data set A-K

Letter	Features	Format	Meaning
A	Asset Code	Integer	The MCH_Code of feature 6 table 9 transformed to an integer instead of a string.
B	Date	dd-mm-yy	The ticket's date derived from feature 3 table 9.
C	Year	Integer (yy)	The ticket's year derived from feature 3 table 9.
D	Month	Integer (mm)	The ticket's month derived from feature 3 table 9.
E	Year&month	String (yy-mm)	Feature B and feature C combined.
F	Quarter	Integer	The ticket's quarter derived from feature 3 table 9.
G	Year&quarter	String(yy-q)	Feature B and feature E combined.
H	Season	String	The ticket's season derived from feature 3 table 9.
I	Week	Integer	The ticket's week number derived from feature 3 table 9.
J	Hour	Integer (hh)	The ticket's hour derived from feature 3 table 9.
K	TotTimeBetween	Real	The time between two consecutive tickets of the same asset. Derived from feature 3 and feature 6 table 9.

Features B to K are constructed for the first success criterium. A set of one or more rules is created which determine(s) the category to which an individual asset belongs. Feature A transforms the unique string per asset to a unique code per asset.

4.3 Time series interval determination

Section 4.3 is linked to the first success criterium. A set of one or more rules is created which determine(s) the category to which an individual asset belongs. The rule will consist of a threshold and a time series interval. If the number of failures of an asset stay below the threshold in the determined time series interval, then the asset is categorised good. Otherwise, the asset is categorised as having a poor condition.

In this section the length of the time series interval used is determined, the threshold is determined in 4.4. Brockwell & Davis (1991) describes a time series as: "A set of observations x_t, each one being recorded at a specified time t." (p. 1). This specified time t can be continuous meaning that all events are recorded as observation. The time interval in which the observations are recorded is specified as $T_i = [x,y]$.

To determine the category to which an asset belongs we look at recent historical data of an asset. This historical data consists of the failures that have happened to the asset. The recent historical data is further specified in this section, by determining the interval length. This consists of the determination of the length of history we look back to and whether we use a fixed or sliding interval length. With fixed is meant that we look at the last complete month / quarter / year etc. prior to today. Hence if it is February 21st and we look at the last month we look at the number of failures of the month January. With sliding is meant that we look at the last two weeks / month / year prior to today. Hence, if it is February 21st and we look at the last month, we look at the number of failures from January 21st to February 21st . In short, in this section the length of the interval is chosen as well as whether the interval will be fixed or sliding.

The maximum time interval possible with the data set of Company A is the whole 2,5 years of data. To determine the preferred length of the time interval, a trade-off must be made between a too short length which does not give enough information and a too long length which consider data points from so long ago that they became irrelevant.

It is determined that a time interval provides useful information, when the average number of tickets per asset are constant as well as the number of tickets per asset of the top 5%, 10%, 15% and 20% worst performing assets. With constant is meant that the threshold is stable throughout multiple periods. If the output is not stable / constant it would mean that for one period the threshold must be 1 ticket and for another period the threshold must be 2 tickets. Then no rule can be created to successfully determine in which category an asset belongs. The top 5% to top 20% worst performing assets are considered based on the pareto principle. The pareto principle states when considering a group of elements 20% of these element account for circa 80% of the effect. Juran (1954) calls this group of 20% of all elements the vital few and the other 80% of the elements trivial many. Juran (1954) states that when improvement is the goal the focus should be the vital few, since improving them is the only way to achieve significant impact. Thus, it is important that the vital few result in a stable output in the eventually chosen time series interval. It must be noted that the pareto principle is normally used to classify a group of elements of which all elements differ from each other. In the case of this analysis the group of elements are all identical, they only differ in how well they perform.

Table 11 and Table 12 show examples of not constant results and constant results over multiple time periods. In the examples below the mean is rounded to its nearest integer.

Table 11 An example of not constant values over time

Time	**Period 1**	**Period 2**	**Period 3**	**Period 4**	**Period 5**	**Period 6**
Mean	0	0	0	1	0	1
Top 5 %	0	1	1	1	0	1
Top 10 %	1	1	1	2	1	2
Top 15 %	1	2	1	2	2	2
Top 20 %	2	2	2	3	2	3

Table 12 An example of constant values over time

Time	**Period 1**	**Period 2**	**Period 3**	**Period 4**	**Period 5**	**Period 6**
Mean	0	0	0	0	0	0
Top 5 %	1	1	1	1	0	1
Top 10 %	1	1	1	1	1	1
Top 15 %	1	1	1	1	1	1
Top 20 %	2	2	2	2	2	2

Company A requested that the fixed interval lengths compared should include interval lengths of:

- A year;
  - 2017, 2018, 2019
- A season;
  - Winter, spring, summer, autumn
- A quarter.
  - 1st quarter (Jan-Mar), 2nd quarter (Apr-Jun), 3rd quarter (Jul-Sep) and 4th quarter (Oct-Dec)
- A month.
  - January, February, March, April, May, June, July, August, September, October, November, December

Furthermore, Company A requested to that the sliding interval lengths compared should include interval lengths of:

- Three months;
  - The last three months prior to today
- Two months.
  - The last two months prior to today

For all the different interval lengths the mean number of tickets was calculated. As well as the thresholds for the 5% to 20% number of tickets. To find the thresholds the assets are first sorted on number of tickets for that quarter from large to small. Then is looked at the number of tickets of the 646th asset, this is the last asset that falls in the group of the top 5% worst performing assets. The number of tickets of that asset (the 646th) would then be the threshold. This was afterwards also done for 10%, 15% and 20%. The thresholds that correspond to each tested interval length are shown in Appendix A.

Beforehand the hypothesis was: the longer the interval length the more constant the outcomes. The results (see appendix A) disproved this hypothesis. When choosing a fixed interval length of a year, the results were not constant over the multiple intervals. The fixed interval length of a month was too short to provide valuable information. The sliding interval length of two months also provided too little valuable information. Since in both the fixed month as the sliding two months no distinction between the top 5% to top 20% worst performing assets could be made.

The fixed interval lengths of a quarter and a season and a sliding interval length of three months provided similar results. Probably because they have the same length. For the quarter, season, and sliding three months the results were stable and did provide enough information.

Feedback session of the time series interval determination

In the feedback session with experts of Company A, the preferred time series time interval is determined. The feedback session consisted of a presentation of the found results and a discussion afterwards. Since a fixed time interval length of a year and a month and a sliding interval length of two months do not give constant and sufficient results these were excluded. Then the choice was between quarters, seasons and sliding three months. Seasons was besides quarters included as a fixed interval length since the idea was that the ticket averages would differ significantly between seasons, due to the weather differences between seasons. This was actually not the case, this became clear when the interval lengths of season, quarter and sliding three months all yielded the same results. Since it was more complicated to work with a fixed time interval length of a season or a sliding interval length of three months instead of quarter, data wise, the interval length of a quarter was chosen to use in section 4.4.

4.4 Categorisation of assets

In 4.4 the assets are categorised based on their performance. This section is linked to the first success criterium. A set of one or more rules is created which determine(s) the category to which an individual asset belongs. An asset either belongs to the good condition category or the poor condition category. A threshold is set to determine in which category an asset belongs. In this section the categorisation is made based on one variable, the total number of tickets of the asset in one period. To determine the threshold, it is first determined how many tickets can be explained by the top worst performing assets. With worst performing is meant that these assets have the highest number of tickets. So, X% of the tickets arise from Y% of the assets. Figure 16 shows the distribution of tickets over all assets. The X-axis shows the top % worst performing assets and the Y-axis shows the ticket % arise from this X value. The threshold will arise from this percentage found, by looking at the last asset that falls within the percentage range. Meaning the asset with the lowest number of tickets which still belongs to the top X% worst performing assets. The number of tickets this asset has will be seen as the threshold.

The dotted line is the trendline best fitted for the plotted graph. To determine what type of trend line was most fitting, the R^2 per trend line type is determined. The rule is the closer the R^2 value is to 1, the more precisely the fit. In Figure 16 the most fitting trendline was the logarithmic function. A logarithmic function has as characteristic that the increase per step either increases or decreases over time. In the case of the assets the increase per step decreases over time. Meaning that improving the reliability of the top 5% worst performing assets has more impact on the overall tickets that improving the reliability of the 5% of the assets after the top 5%.

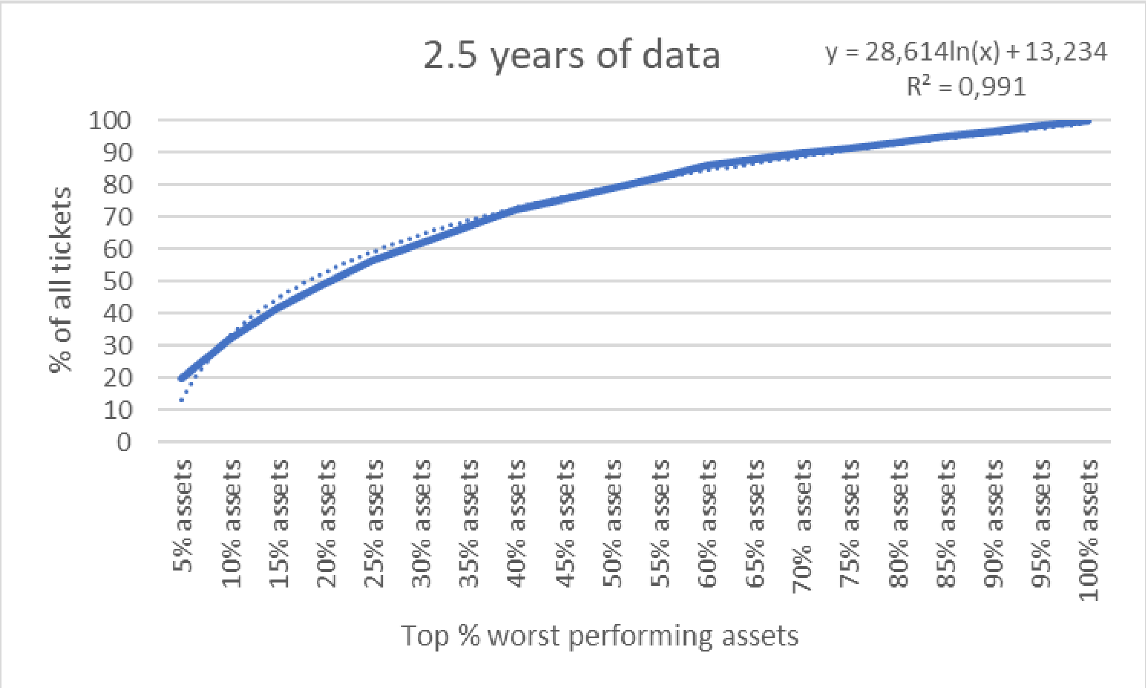


Figure 16 Ticket distribution over the assets (2.5 years of data)

This is also calculated for each quarter, instead of the whole 2.5 years. The results are shown in Figure 17. Here the decrease of the increase per step is even bigger compared to the whole 2.5 years. Where in Figure 16 the top 5% worst performing assets are accountable for 20% of all tickets, when looking per quarter the top 5% worst performing assets are accountable for 40% off all tickets. From this it can be concluded that, the behaviour of individual assets shifts over time. Otherwise the top 5% worst performing assets would account for the same amount of tickets in both the 2.5 years and the quarters.

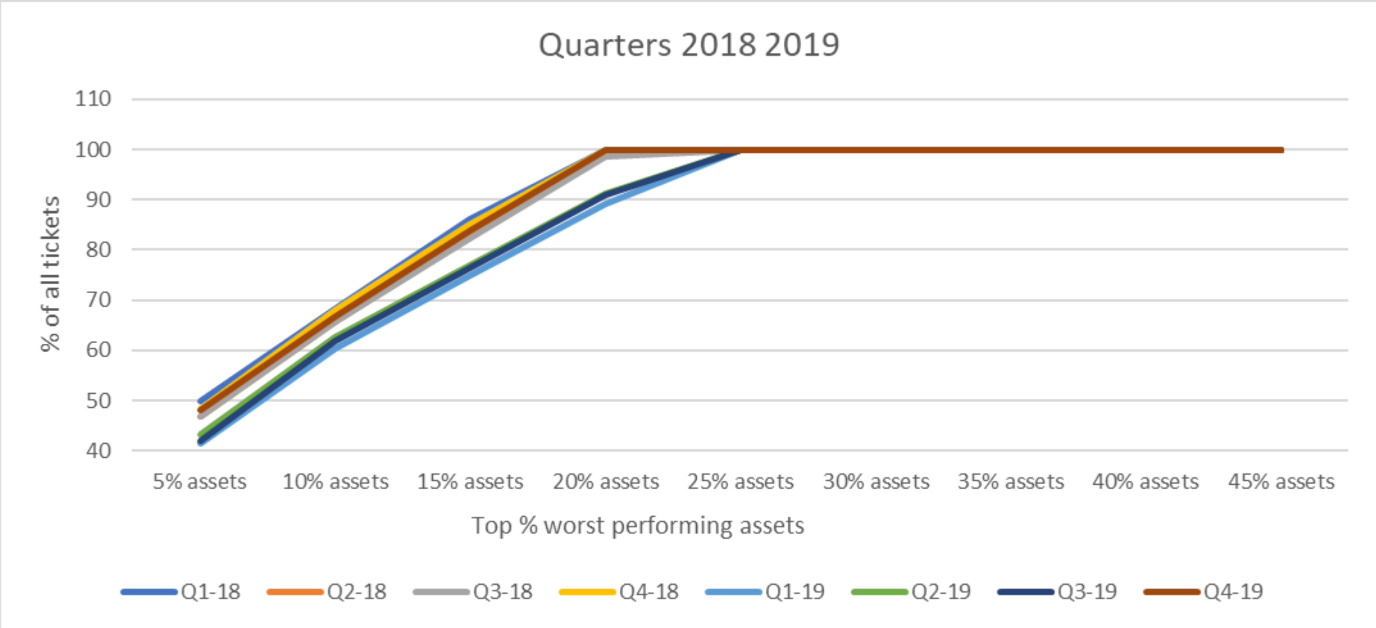


Figure 17 Ticket distribution over the assets (per quarter)

This insight in the ticket distribution is useful for the decision made later on in this section.

The next part of section 4.4 is to determine the actual threshold. There will be one threshold that divides the assets into a good condition category and a poor condition category. For the threshold determination first is looked at the top 5% to 20% worst performing assets, we call this the asset percentage method. Then is looked at the smallest group of assets that together contribute to 5 % to 20% of all tickets, we call this the ticket percentage method. Table 13 shows the names given for all the tested thresholds and their definition.

Table 13 Names and definitions of all the tested threshold groups

Method	Name	Definition
Asset percentage method	5%-worst-assets	This group consists of the top 5% worst performing assets per quarter.
	10%-worst-assets	This group consists of the top 10% worst performing assets per quarter. Thus, the assets which fall in the 5%-worst-assets group also fall within this group.
	15%-worst-assets	This group consists of the top 15% worst performing assets per quarter. Thus, the assets which fall in the 5%-worst-assets group and the 10%-worst-assets group also fall within this group.
	20%-worst-assets	This group consists of the top 20% worst performing assets per quarter. Thus, the assets which fall in the 5%-worst-assets group, the 10%-worst-assets and the 15%-worst-assets also fall within this group.
Ticket percentage method	5%-worst-tickets	This group consists of the smallest group of assets that together contribute to 5% of all tickets per quarter.
	10%-worst-tickets	This group consists of the smallest group of assets that together contribute to 10% of all tickets per quarter. Thus, the assets which fall in the 5%-worst-tickets group also fall within this group.
	15%-worst-tickets	This group consists of the smallest group of assets that together contribute to 15% of all tickets per quarter. Thus, the assets which fall in the 5%-worst-tickets group and the 10%-worst-tickets group also fall within this group.
	20%-worst-tickets	This group consists of the smallest group of assets that together contribute to 20% of all tickets per quarter. Thus, the assets which fall in the 5%-worst-tickets group, the 10%-worst-tickets and the 15%-worst-tickets also fall within this group.

The difference between the asset percentage method and ticket percentage method is shown in Figure 18. With the asset percentage method, the focus is on the 5%-worst-assets, which are accountable for 20% of all tickets. While with the ticket percentage method the focus is on the top 0.27% worst performing assets, which are accountable for 5 % of all tickets.

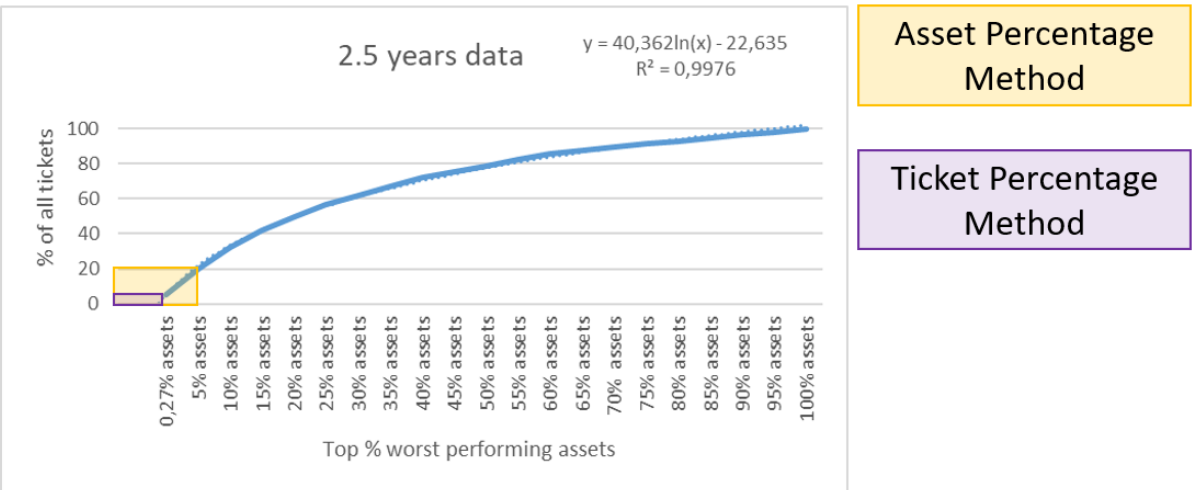


Figure 18 Difference between asset percentage method and ticket percentage method

The thresholds equals the number of tickets of the last asset that falls within the range. To find the thresholds the assets are first sorted on number of tickets for that quarter from large to small. For the asset percentage method is then looked at the number of tickets of the 646th asset (5%-worst-assets), 1291th asset (10%-worst-assets), 1937th asset (15%-worst-assets) and 2582th asset (20%-worst-assets). The number of tickets those assets have result in the threshold for the x%-worst-assets group. For the ticket percentage method it is not always assets on the same ranking place which is the threshold. Thus it is looked what is the smallest group of assets which together make up 5% of all tickets of that quarter. Then this group is sorted based on number of tickets from large to small and the number of tickets of the last assets is the threshold. This is afterwards also done for 10%, 15% and 20% of all tickets. For both the assets percentage method as well as the ticket percentage method the thresholds are determined. For the asset percentage method is found that the 5%-worst-assets, 10%-worst-assets and 15%-worst-assets all yield a constant output, and the 20%-worst-assets does not yield a constant output.

For the ticket percentage method is found that only the 20%-worst-tickets yield a completely constant output. The 5%-worst-tickets has one outlier, the 10%-worst-tickets has three outliers and the 15%-worst-tickets had 2 outliers. Since we earlier determined that a constant / stable output is necessary, it is concluded that only the 20%-worst-tickets may be useful as a threshold. The results per quarter and the threshold that would result from the results are shown in Table 14. All results of this section can be found in Appendix A.

Table 14 Results and thresholds per tested group

	Quarter - year									
Value	**1-18**	**2-18**	**3-18**	**4-18**	**1-19**	**2-19**	**3-19**	**4-19**	**Average**	**Threshold**
Mean	0	0	0	0	0	0	0	0	0	0
20%-worst-assets	0	1	1	0	1	1	1	0	0.625	1
15%-worst-assets	1	1	1	1	1	1	1	1	1	1
10%-worst-assets	1	1	1	1	1	1	1	1	1	1
5%-worst-assets	2	2	2	2	2	2	2	2	2	2
20%-worst-tickets	3	3	3	3	3	3	3	3	3	3
15%-worst-tickets	4	4	4	4	3	4	3	4	3.75	4
10%-worst-tickets	5	4	4	5	4	5	4	4	4.375	4
5%-worst-tickets	6	6	6	6	5	6	6	6	5.875	6

Feedback session of the categorisation of assets

As in the first feedback session this feedback session also consisted of a presentation of the found results thus far and a discussion afterwards. In this feedback session more experts of Company A joined. Therefore, all the sections of this chapter thus far are discussed. First is decided that the asset percentage method was more preferable, for Company A, than the ticket percentage method. Since the experts felt that the thresholds of the ticket percentage method where too high to correctly identify the assets on which should be focussed. Thus that the group of assets to focus on would be too small to create significant impact. The second decision is whether to focus on the 5%-worst-assets, 10%-worst-assets, 15%-worst-assets or 20%-worst-assets. The 20%-worst-assets is excluded since it did not yield a constant output. Thus, the choice is between the 5%-worst-assets, 10%-worst-assets and 15%-worst-assets. The choice made is supported by the findings of the ticket distribution in Figure 16 and Figure 17. From those graphs it is clear that the distribution follows a logarithmic function. To show that indeed the increase per step decreases over time Figure 19 is created. Therefore, is looked at how the worst performing asset account for a percentage of all tickets per quarter. Thus the X axis shows the 5%-worst-asset, the assets that fall between the 5% and 10% worst performing assets, the assets that fall between the 10% and 15% worst performing assets etcetera. The Y axis shows the percentage of tickets explained by the group of the X-axis, this is called the delta. The formula used to calculate the delta is:

$$\Delta(A\ to\ B\ \%) = \%\ of\ the\ tickets\ covered\ by\ A\%worstassets - \%\ of\ the\ tickets\ covered\ by\ B\%worstassets$$

To clarify the delta, the delta of 5-10% is calculated. From the data is calculated that in Q1 2018 the 10%-worst-assets are accountable for 68.36% of all tickets and the 5%-worst-assets are accountable for 49.81% of the tickets. Therefore the delta of 5 to 10% is:

$$\Delta\ (\ 5\ to\ 10\%) = 68.36\ - 49.81 = 18.55$$

The graph in Figure 19 shows that the delta significantly decreases. Therefore, improving the performance of the 5%-worst-assets yields a higher decrease in the total number of tickets, then improving the performance of the assets that fall with in the 5-10% worst performing ticket range. Thus, the choice is made to focus on the 5%-worst-assets. The threshold that corresponds to this choice is a threshold of two tickets. This means that when an asset has two or more tickets in the previous quarter, it will be labelled as a poor condition asset. Otherwise, the asset will be labelled as a good condition asset.

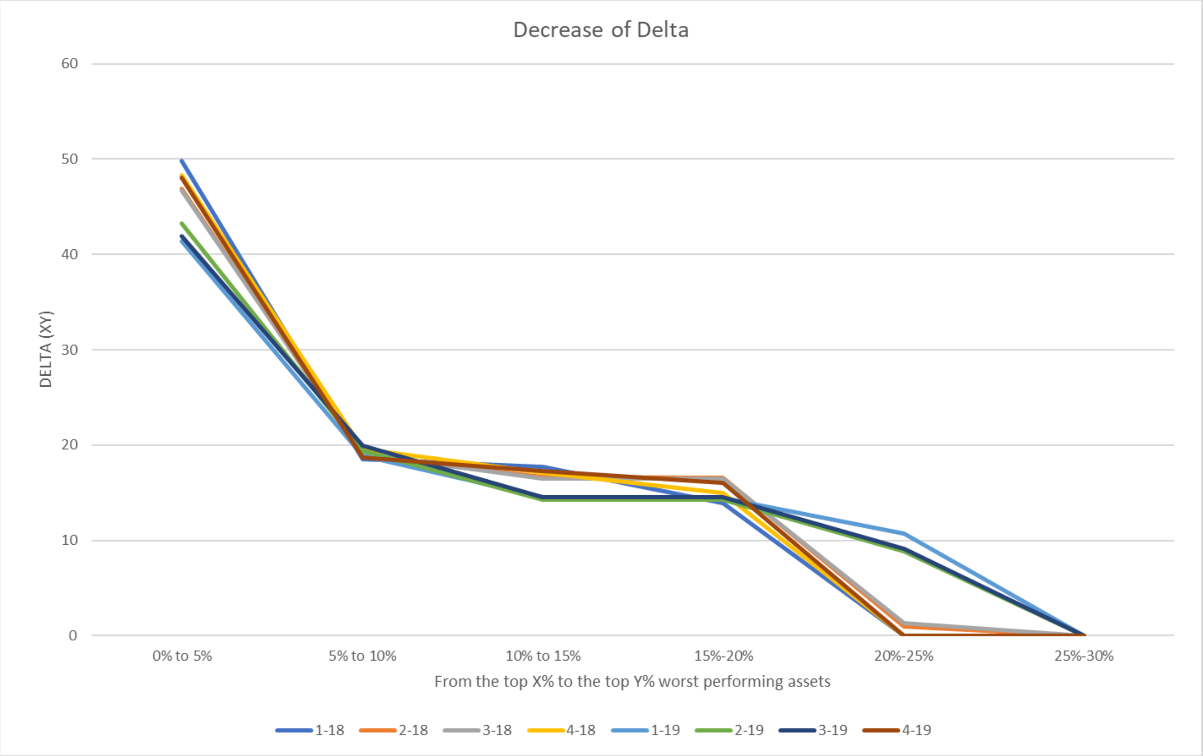


Figure 19 Decrease of Delta

4.5 Ticket labelling

This section is linked to the second success criterium: The conditional probability of all follow-up failures is identified, for at least cause group level. Sections 4.2 to 4.4 focussed on the assets and the ticket quantity attached to these assets. This section will focus on the malfunction that caused the ticket. From the description column in the data set, an expert of Company A already divided the tickets into 64 causes. Considering that there are multiple causes concerning the same component and/ or have the same characteristics, the causes will be aggregated in this section. After the aggragation of causes the tickets of the data set will be labelled with the new aggregated level, meaning that a new feature is constructed which contains the cause group of each ticket. The division shown in Figure 20 starts at the highest level of aggregation all tickets, thus a top-down method is used. The levels are as follows:

0. All tickets
1. Cause group
    a. In this level the tickets are divided based on their cause group. Fifteen cause groups, such as cause group door, are identified.
2. Cause
    a. In this level the tickets are further divided based on their cause. 64 causes, such as hinge failure, are identified.

Levels 1 and 2 will be further elaborated on after the visualisation in Figure 20.

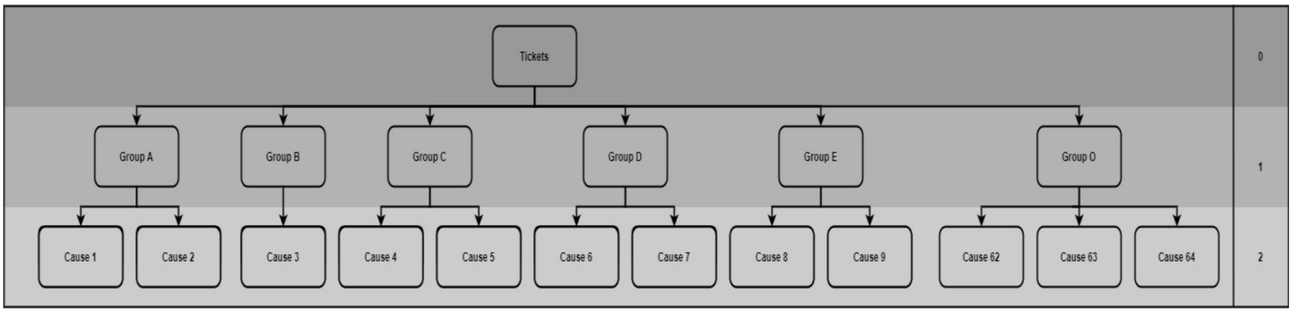


Figure 20 Ticket division through multiple levels

Level 1: Cause group

In this level the causes are divided into fifteen cause groups, Company A has no insight in the best PM action per cause group, therefore these cannot be mentioned here. These 15 cause groups are in Dutch with the English definition is in parentheses:

a. Accu (Battery)
b. Besturingssysteem (Operating system)
c. Deur (Door)
d. Energie (Energy)
e. Ethernet link down (Ethernet link down)
f. Extern (External)
g. Facilitair (Facility)
h. Geen verbinding (No connection)
i. Gelijkrichter (Rectifier)
j. Geluidsklacht (Noise complaints)
k. Onbekend (Unknown)
l. Temperatuur (Temperature)
m. Uitval SC (SC failure)
n. Verkeerd geplaatst (Placed wrong)
o. WOA (WOA)

For illustration, we look at a malfunctioning door. There are multiple causes related to a malfunctioning door. All these causes belong into the cause group door. This grouping of causes supports section 4.6 where the conditional probability per cause group is determined. For this conditional probability it is not important whether the door malfunctions due to a broken lock or broken hinge. It is important what component will malfunction most likely after a malfunctioning door. All cause groups identified are shown in Figure 21, level 1.

Level 2: Cause

Level 2 includes the exact cause, as described in the feature cause:

14	Cause	String	The cause of the ticket.

Examples of such a cause are battery problem and battery fuse, both belong to the cause group battery. Levels 0 and 1 of the division are shown in Figure 21.

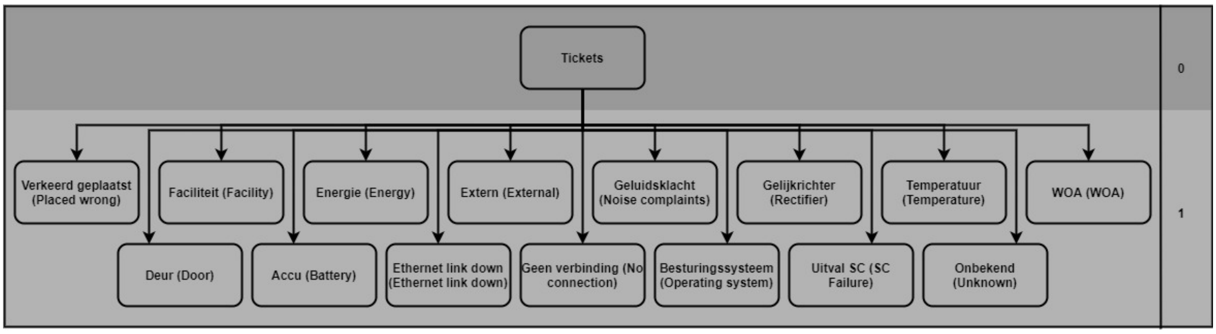


Figure 21 Levels 0 and 1 elaborated

Features construction

Level 1 constructed in the data set as a feature, other features that are also constructed aim to distinguish how many failures an asset has had. The new constructed features are shown in Table 15.

Table 15 New features in data set L-AX

Letter	Feature	Format	Meaning
L	Cause Group	String	The category of the cause belonging to the ticket.
M	CountFailureNumber	Integer	Shows the ticket number belonging to the asset. Thus, the first ticket (based on date) of the asset in the data set is gets the value 1, the second ticket 2, etcetera.
N-AX	Xst failure	String	Shows the ticket cause belonging to feature N (X=P).

4.6 Conditional probability between cause groups

Section 4.6 is linked to the second success criterium: The conditional probability of all follow-up failures is identified, for at least cause group level. As explained in chapter 2 the underlying relationships between are not known by Company A. Since it is not realistic to determine if there is a causality between failures in this thesis, we will determine the conditional probability between failures. This will still provide valuable insights in the data.

Since when the most common next failures are known a PM action can be included concerning these common next failures. This may lead to a lower number of tickets overall.

In this section the conditional probability from consecutive tickets is determined. The theory behind conditional probability is explained in section 3.5.3. In conditional probability the current event is denoted with *x*, in this case the event *x* is occurrence of the current ticket with cause group A. Then the probability that the next ticket will have cause group B is determined. Figure 22 shows the illustration of the conditional probability.

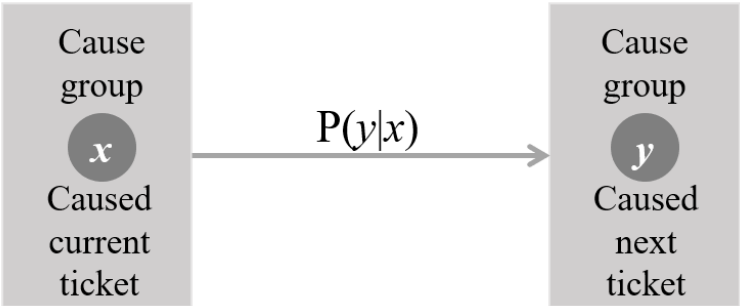


Figure 22 Visualisation of conditional probability of cause group

For each asset, the tickets are arranged on chronological order. This creates a table with each asset being a unique row and each column corresponds to a ticket category.

Table 16 Chronological order of ticket cause groups

1st failure	2nd failure	3rd failure	4th failure	5th failure	…	37th failure
Cause group A	Cause group B	Cause group C	Cause group A	Cause group D	…	Cause group N

To calculate the conditional probabilities, it must be known how often a cause group failure will have another cause group failure as next failure. This must be derived from the data set. As shown in Table 16 a matrix is created, which shows how often one cause group failure was succeeded by another cause group failure in the data set.

Table 17 shows the structure of the matrix the actual matrix is shown in Table 17. The purple cells represent the cause group *x* of the current failure, the blue cells represent the cause group *y* of the next failure and the grey cells represent how often a failure of cause group *x* had a failure of cause group *y* as successor in the data set.

The matrix is constructed using a pivot table. The pivot table is limited, meaning it can only show the consecutive failures per cause group. This means, it only shows how often the failures consecutively occur as 1^{st} and 2^{nd} failure or as 2^{nd} and 3^{rd} failure. Thus, all these pivot table results, $1^{st} \rightarrow 2^{nd} + 2^{nd} \rightarrow 3^{rd} + \ldots + 36^{th} \rightarrow 37^{th}$, need to be added up together, to create a matrix which represents the total number of times that two cause group failures happen consecutively. This matrix is then used to calculate the conditional probability. As said before little is known about the causal relation between cause group failures. For the model is assumed that only the current cause group failure influences the next, thus the cause groups of the earlier failures are left out.

Besides the conditional probability between two consecutive failures also the independency between cause groups are determined. As described in Chapter 3 independency between two cause groups means that the occurrence of the first event (a failure with cause group A) does not influence the next event (a failure with cause group B). Thus, when independency is found between two cause groups the conditional probability of them cannot be used.

The first matrix in Appendix B is the matrix of the transformation step 1. Transformation steps 2 to 7 are shown in the second figure of appendix B. The calculations of the transformation steps are described below.

Table 17 Cause group matrix

	Accu	Besturingssysteem	Deur	Energie	Ethernet link down	Extern	Facilitair	Geen verbinding	Gelijkrichter	Geluidsklacht	Onbekend	Temperatuur	Uitval SC	Verkeerd geplaatst	WOA
Accu	627	158	116	95	3	2	2	45	34	1	69	94	162	0	19
Besturingssysteem	128	1815	629	146	34	7	12	73	109	15	256	240	374	0	35
Deur	110	730	5435	179	73	41	36	157	224	28	430	314	568	0	47
Energie	91	159	158	343	11	8	1	17	63	3	90	79	88	0	8
Ethernet link down	2	35	62	9	98	0	2	5	9	0	15	18	22	0	2
Extern	6	9	49	4	0	36	1	9	6	0	21	12	11	0	1
Facilitair	1	7	19	3	2	1	13	2	2	0	3	6	6	0	2
Geen verbinding	35	86	117	21	8	2	1	192	18	2	60	35	177	0	14
Gelijkrichter	30	110	140	54	5	2	2	15	217	7	41	66	86	1	12
Geluidsklacht	1	6	24	3	1	0	0	0	10	39	5	11	7	0	1
Onbekend	64	284	504	75	23	25	11	67	63	10	552	110	169	0	15
Temperatuur	123	212	296	61	25	7	5	33	71	6	86	603	161	0	16
Uitval SC	148	374	598	78	27	7	10	205	92	10	179	212	1708	0	49
Verkeerd geplaatst	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
WOA	20	48	39	9	2	1	2	8	5	2	13	14	56	0	230

Legend	
	The current event with category *x*.
	The next event with category *y*.
	The sum of all the times *y* is the successor of *x*.

The matrix of Table 17 is transformed into a conditional probability matrix where the grey cells represent $P(y|x)$. The transformation steps (1-4) are as follows:

1. Calculate how often a cause group is the current event as well as how often it is the successor. Table 17 shows the results of this step.
2. Calculate the probability that a cause group is the current event. $P(x) = \frac{for\ x; \sum_{y=A}^{O} x \rightarrow y}{Total\ number\ of\ events}$.
3. Calculate $P(y \cap x) = \frac{\#\ x \rightarrow y}{Total\ number\ of\ events}$.
4. Calculate $P(y|x) = \frac{P(y \cap x)}{P(x)}$ for each x and y.

To confirm that *x* and *y* are dependent on each other the independence is calculated. If

$P(y \cap x) = P(x) * P(y)$ or $P(y|x) = P(y)$ hold, then *x* and *y* are independent, otherwise they are dependent. If an independency is found conditional probabilities between consecutive failures cannot be used. Cause groups are independent if the product of the probability of the occurrence of both cause groups is equal to the probability of the intersection between two cause groups. The independency check steps (5-7) are as follows:

5. Calculate $P(y) = \frac{for\ y; \sum_{x=A}^{O} x \rightarrow y}{Total\ number\ of\ events}$ and $P(x) * P(y)$.
6. Calculate the difference between $P(y \cap x)\ and\ P(x) * P(y)$.
7. Calculate the difference between $P(y|x)\ and\ P(y)$.

The results show that no $x \rightarrow y$ combination yields a result of null in steps 5 and 6. Since an independency would indicate that conditional probabilities could not be used and we found no independencies, the use of conditional probabilities between consecutive failures is justified.

Feedback session after the data analysis

In this feedback session the results of section 4.6 were discussed with an expert of Company A. As well as whether the objective of the data analysis and the success criteria were met. Firstly, the first success criterium is discussed: A set of one or more rules is created which determine(s) the category to which an individual asset belongs. Secondly the second success criterium is discussed: The conditional probability of all follow-up failures is identified, at cause group level.

Currently the threshold in success criterium 1 is solely based on the number of tickets within a quarter. Whether this is a good enough threshold or that more variables should be included, is discussed in the feedback session. In the feedback session is concluded that the experts of Company A prefer a simplistic threshold, thus including other variables is not considered.

In section 4.3 is chosen to continue to section 4.4 with a fixed interval length of a quarter. For the solution for Company A the sliding interval length of three months is chosen instead. The idea behind this is, that when implementing categorisation on asset level it is better to look at the past three months instead of looking at the past season or quarter. Looking at the past quarter would mean that in December the performance of the asset in Q3 (July-September) determines the category of the asset and October plus November are not taken into account. The expert of Company A agrees that a sliding time series interval is better, thus the three months sliding interval length is chosen. In short, a threshold of 2 tickets and a sliding interval length of three months are chosen, thus success criterium 1 is met.

The feedback on section 4.6 was that the cause groups with overall the highest number of tickets also had the highest $P(y|x)$ when they were y. The question from the expert was to redetermine $P(y|x)$ when taking into account the differences between ticket numbers per cause group, thus, to standardise the results. When evaluating this request, we concluded that it would indeed filter out the cause groups which have a high occurrence, but that this would create an unusable conditional probability. The aim of determining the conditional probabilities is to carry out PM for the cause categories y with the highest probability of occurring after x, $P(y|x)$. Thus, to meet the aim the probability should not be standardised. It can be concluded that the second success criterium is met.

4.7 Conclusion

In short, the following choices are made based on the data analysis. The interval length chosen is a sliding interval length of three months. Thus, each day the performance of the three months before that day determine in which category the asset is placed. The threshold for this determination is two tickets. So, an asset is in the poor condition category if it had two or more tickets in the last three months. The conditional probabilities are shown in Figure 23. Per cause group top two probabilities are highlighted. Note that one of the two groups always highlighted is the same cause group as the current cause group. Thus, a repeated failure of the same cause group.

P(A|B)	Accu	Besturingssysteem	Deur	Energie	Ethernet link down	Extern	Facilitair	Geen verbinding	Gelijkrich	Geluidsklacht	Onbekend	Temperatuur	Uitval SC	Verkeerd geplaatst	WOA
Accu	0,439383322	0,110721794	0,081289418	0,066573231	0,002102313	0,001401542	0,001401542	0,031534688	0,023826	0,000700771	0,048353189	0,06587246	0,113524877	0	0,013314646
Besturingssysteem	0,033049316	0,46862897	0,162406403	0,037696876	0,008778725	0,001807384	0,003098373	0,018848438	0,028144	0,003872967	0,066098632	0,061967467	0,09656597	0	0,009036922
Deur	0,013139035	0,087195413	0,649187769	0,021380793	0,008719541	0,004897277	0,004300048	0,018752986	0,026756	0,003344482	0,051361682	0,037505972	0,067845198	0	0,005613951
Energie	0,081322609	0,142091153	0,141197498	0,306523682	0,009830206	0,00714924	0,000893655	0,015192136	0,0563	0,002680965	0,080428954	0,070598749	0,078641644	0	0,00714924
Ethernet link down	0,007168459	0,125448029	0,222222222	0,032258065	0,35125448	0	0,007168459	0,017921147	0,032258	0	0,053763441	0,064516129	0,078853047	0	0,007168459
Extern	0,036363636	0,054545455	0,296969697	0,024242424	0	0,218181818	0,006060606	0,054545455	0,036364	0	0,127272727	0,072727273	0,066666667	0	0,006060606
Facilitair	0,014925373	0,104477612	0,28358209	0,044776119	0,029850746	0,014925373	0,194029851	0,029850746	0,029851	0	0,044776119	0,089552239	0,089552239	0	0,029850746
Geen verbinding	0,045572917	0,111979167	0,15234375	0,02734375	0,010416667	0,002604167	0,001302083	0,25	0,023438	0,002604167	0,078125	0,045572917	0,23046875	0	0,018229167
Gelijkrichter	0,038071066	0,139593909	0,177664975	0,068527919	0,006345178	0,002538071	0,002538071	0,019035533	0,275381	0,008883249	0,052030457	0,083756345	0,109137056	0,001269036	0,015228426
Geluidsklacht	0,009259259	0,055555556	0,222222222	0,027777778	0,009259259	0	0	0	0,092593	0,361111111	0,046296296	0,101851852	0,064814815	0	0,009259259
Onbekend	0,032454361	0,144016227	0,255578093	0,038032454	0,011663286	0,012677485	0,005578093	0,033975659	0,031947	0,005070994	0,279918864	0,055780933	0,085699797	0	0,007606491
Temperatuur	0,072140762	0,124340176	0,173607038	0,035777126	0,014662757	0,004105572	0,002932551	0,019354839	0,041642	0,003519062	0,050439883	0,353665689	0,094428152	0	0,009384164
Uitval SC	0,040032459	0,101163105	0,161752773	0,021098188	0,007303219	0,001893427	0,002704896	0,055450365	0,024885	0,002704896	0,048417636	0,057343792	0,461996213	0	0,01325399
Verkeerd geplaatst	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
WOA	0,04454343	0,106904232	0,086859688	0,020044543	0,004454343	0,002227171	0,004454343	0,017817372	0,011136	0,004454343	0,028953229	0,031180401	0,124721604	0	0,512249443

Figure 23 Conditional probability per cause group

5. Solution design

In this chapter the solution for Company A and Supply Value will be explained. In section 5.1 The Solution for company A will be further elaborated on, the solution will for Company A will be validated in section 5.2. In section 5.3 the solution of Supply Value will be discussed.

5.1 Solution for Company A

In this section the proposed solution for company A is described. The objective stated in chapter 4 is: improve the maintenance quality by using ticket data, increasing the knowledge concerning asset categorisation, malfunction causes and conditional probability between causes. The solution in this chapter aims to reach this objective, by decreasing the number of CM actions and visits to the assets.

The solution consists of two recommendations. The first recommendation for company A is to divide the assets into two categories. The good condition category and the poor condition category. A threshold and an interval length, both based on the data analysis of chapter 4, are used to determine these two categories. The threshold proposed is the threshold of two tickets per the time series interval of the last three months. An asset that had two or more tickets the last three months is a poor condition asset. An asset that had null or one ticket the last three months is an asset in good condition. For the data set this means that on average 5% of all assets are labelled as an asset in poor condition. The next section will explain the validation of this proposed solution.

The second recommendation concerns the maintenance per category. The good condition category will be maintained as all assets are maintained currently. The current maintenance strategy is to only execute CM the assets. Thus, the maintenance strategy for the assets in the good condition category is CM. For the assets in the poor condition category the maintenance strategy is OM. In chapter 3 OM is defined as: When a failure occurs the opportunity is taken to not only perform CM but simultaneously perform PM on the other components of the asset, Ding & Tian (2012). For the proposed solution this means that: If a CM action is executed on an asset belonging to the poor condition category, a PM action on that asset is executed as well. The PM action to execute follows from the conditional probability analysis from chapter 4. Since it is not clear yet how failures of different cause groups interact with each other, it is assumed that by executing PM on the cause group most likely to occur as the next failure, the total number of failures will drop. Since the chance of failure of the component that received PM drops.

5.2 Validation for Company A

In this section we will first explain the KPI's determined for this validation. In 5.2.1 We further elaborate on the simulation and its inputs and outputs. In 5.2.2 the assumptions and simplifications made compared to reality are explained. In 5.2.3 The flow charts belonging to the simulation are explained using two examples to guide you through the flow charts. At last in 5.2.4 the simulation results are discussed.

It is important to research whether the proposed solution will have a positive impact on Company A. Therefore, a simulation, using VBA in Excel, is setup. The simulation simulates the proposed solution as well as the current situation. Due to a limited data set the simulation is a simplification of the reality. Consequently, we also simulate the current situation to have a better comparison. The two simulations are compared on three KPI's these are:

- The number of corrective maintenance executed.
    - The expectation is that the number of CM executed in the proposed solution simulation will be lower than in the simulation of the current situation. Thus, providing insight in the possible improvement when implementing the solution.
- The total number of maintenance executed.
    - The expectation is that the total number of maintenance executed may be higher in simulation of the proposed solution than in the simulation of the current situation. Since it is likely that the number of PM executed is higher than the number of CM that is not needed anymore. This KPI provides insight in the possible downside of the proposed solution.
- The number of visits.
    - The expectation is that the number of visits will be lower in the simulation of the proposed solution compared to the simulation of the current situation. Since in one visit CM and PM both are executed when needed.

Other KPI's such as costs are not included, since we did not have the data concerning that available. The KPI's are all averages per asset over the course of 2,5 years. Since this time span is equal to the time span of the dataset.

5.2.1 Simulation explanation

The simulation is a discrete-event simulation. It simulates the current situation, in which only CM is executed, as well as the proposed solution. In the proposed solution CM and PM are executed for the assets belonging to the poor condition category and only CM is executed for the assets of the good condition category. To simulate both situations a simulation is created which consists of three parts. The first part simulates the assets belonging to the good condition category. The maintenance intervention is equal for both the current situation and the proposed solution. The second part simulates assets of the poor condition category in the current situation, thus only CM is applied on these assets. The third part simulates the assets of the poor condition category in the proposed solution, on them OM is applied. A combination of the results of the first and second part create the current situation. A combination of the results of the first and third part create the proposed solution.

The simulation can be visualised as a black box model. Figure 24 shows the simulation as a black box model with the inputs and outputs belonging to the simulation. Table 18 explains the different in- and outputs mentioned.

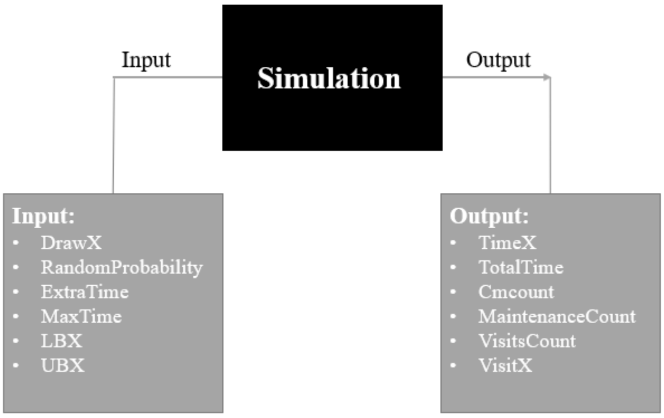


Figure 24 Black box model of simulation

Table 18 Simulation input and output

Input and output	Definition
X	$cause\ group, \quad X: \{A, B, C, D, E, F, G, H, I, J, K, L, M, N\}$
$DrawX$	$Drawn\ time\ from\ the\ empirical\ distribution\ of\ X$
$ExtraTime$	$Extra\ time\ passed\ since\ last\ iteration, drawn\ time\ from\ the\ total\ empirical\ distribution$
$RandomProbability$	$Drawn\ probability\ between\ 0 - 1\ of\ the\ uniform\ distribution$
LBX	$Lower\ bound\ of\ the\ probability\ of\ cause\ group\ X$
UBX	$Upperbound\ of\ the\ probability\ of\ cause\ group\ X$
$TimeX$	$The\ time\ passed\ since\ the\ last\ maintenance\ executed\ on\ cause\ group\ X$
$TotalTime$	$The\ total\ time\ passed\ thus\ far$
$CMcount$	$The\ number\ of\ corrective\ maintenance\ executed\ in\ total$
$MaintenanceCount$	$The\ number\ of\ maintenance\ executed\ in\ total$
$VisitsCount$	$The\ number\ of\ visits\ to\ the\ asset\ in\ total$
$VisitX$	$The\ number\ of\ visits\ to\ the\ asset\ because\ of\ a\ failure\ of\ cause\ group\ X\ in\ total$
$MaxTime$	$The\ maximum\ time\ after\ which\ the\ simulation\ terminates$

For some inputs and outputs the definition is clear enough, others need a more in-depth explanation. In the paragraphs below we further elaborate on those that need more explanation.

The cause groups in the simulation are given a code from A to N to minimize possible typos in the code. The cause group code for each cause group is shown in Table 19.

Table 19 Cause groups and simulation naming

Cause group	Cause group code
Accu	A
Besturingssysteem	B
Deur	C
Energie	D
Ethernet link down	E
Extern	F
Facilitair	G
Geen verbinding	H
Gelijkrichter	I
Geluidsklacht	J
Onbekend	K
Temperatuur	L
Uitval streetcabinet	M
WOA	N

The first input is DrawX, for each cause group a random time from the empirical distribution of that cause group is drawn. Since there was no information available about the interarrival times of failures of the cause groups, we obtained them ourselves. When plotting the found interarrival times in a graph it became clear that the interarrival times of the failures of all cause groups do not follow a generic distribution such as the exponential distribution or the normal distribution. Therefore for each cause group we divided the found interarrival times of failures into bins of 1 day and counted how many interarrival times found belonged to each group. Thus, 7 interarrival times of cause group door are found that have a value between 2 and 3 days, they belong to the 2 day bin of cause group door and that bin then gets the value 7. For each cause group there are 900 bins of which each bin represents 1 day. This is because the total length of the data set available is 900 days. Therefore, no failures are detected which have a interarrival time greater than 900. When for each bin of the cause group door the total number of failures that fall within that bin are counted the probability of each bin can be calculated by : $P(X\ day\ bin) = \frac{Number\ of\ failures\ of\ cause\ group\ door\ counted\ in\ X\ day\ bin}{Total\ number\ of\ failures\ of\ cause\ group\ door\ counted}$. This is done for all cause groups the histograms in appendix C show the empirical distribution of all cause groups.

As said the simulation is a discrete-event simulation which goes from one event to another. To determine the time a next event happens a time is drawn which acts as the time passed since the last event. This time we call ExtraTime. As done with the interarrival times of failures of the cause groups also the interarrival times of all failures are determined by binning the interarrival times found. The ExtraTime is drawn from the empirical distribution each loop.

The input MaxTime is used to determine the run length of the simulation. This is set on 900 days, since we only have information about interarrival times of failures of the cause groups for 900 days.

LBX and UBX define the lower and upper bound per cause group. For each cause group is determine what the probability is that they occur by the formula:

$$P(cause\ group\ X) = \frac{Number\ of\ failures\ of\ cause\ group\ X\ counted}{Total\ number\ of\ failures\ counted}$$

Then the intervals of all cause groups are determine. The interval of accu has a lower bound of 0 and and an upper bound of 0.050659 since P(accu) = 0.050659. The lower bound of besturingssysteem is equal to the upper bound of accu. The upper bound of besturingssysteem is calculated by adding P(besturingssysteem) to the lower bound thus $UBbesturingssysteem = UBaccu + P(besturingssysteem) = 0.050659 + \ 0.1425$ This is done for all cause groups, the lower and upper bound of all cause groups are shown in Table 20.

Table 20 lower bound and upper bound per cause group

Cause group	LB	UB
Accu	0	0,050659
Besturingssysteem	0,050659099	0,193168
Deur	0,193167652	0,568867
Energie	0,568867199	0,608997
Ethernet link down	0,608996632	0,620667
Extern	0,62066679	0,627589
Facilitair	0,627589317	0,631515
Geen verbinding	0,631514734	0,662839
Gelijkrichter	0,662838501	0,6983
Geluidsklacht	0,69829987	0,702809
Onbekend	0,702808795	0,780574
Temperatuur	0,78057449	0,848474
Uitval streetcabinet	0,848473596	0,981646
WOA	0,981646023	1

The output TotalTime is at the start of the simulation 0. Then at the end of the first loop the TotalTime is $TotalTime = \ 0 + ExtraTime$. At the end of each loop that is not the first loop the TotalTime is $Totaltime = TotalTime + ExtraTime$.

5.2.2 Assumptions and simplifications

As said before, the simulation is a simplification of the reality. We consider the following assumptions and simplifications:

- *The maintenance executed is perfect*: This means that after the maintenance the component can be seen as new. This is not necessarily the case when looking at the data set. It is clear from the data set that a malfunction of the same cause group is most likely to be the next malfunction for all cause groups.

- *The failures that are occurred before and detected during random observation is assumed to occur at the time of the observation*: Figure 25 shows how this decision is made. The explanation of the different variables mentioned such as DrawX (drawn time from the emprocal distribution (see also section 5.2.3 for definition of DrawX). When a failure does occur, the assumption is made that the failure occurs at the end of ExtraTime. In practice it should occur at the end of DrawX, but this would create an infeasible simulation, where there are constant jumps back in time. Thus, when a failure occurs it can be seen as a DTM on time DrawX a fault can be noted and the failure happens at TimeX + ExtraTime. In short there are two types of events that can occur either the event that no failure occurs or the event that a failure of component X occurs.

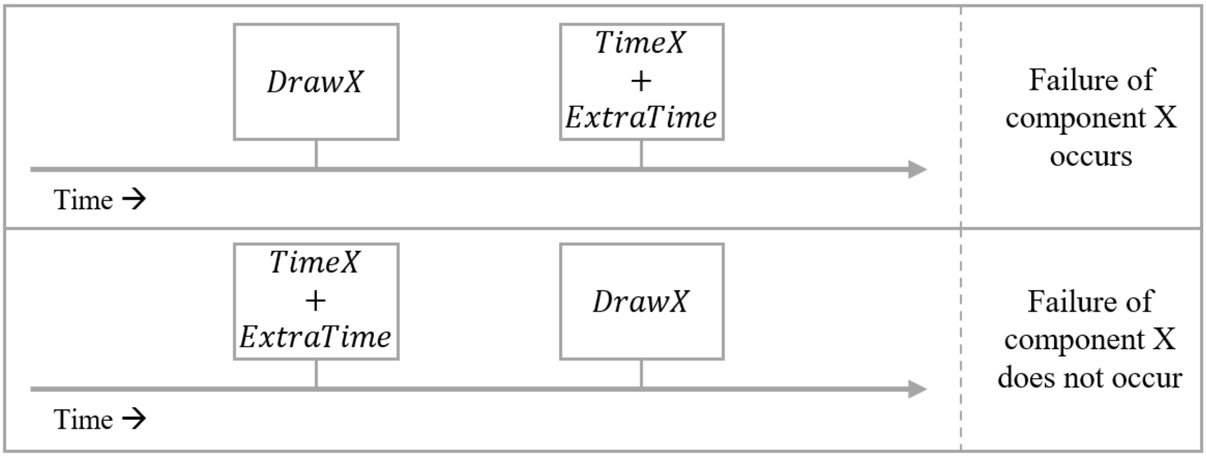


Figure 25 Failure occurrence in the simulation

- Th*e cause group verkeerd geplaatst / wrongly placed is left out*: This only occurred once in the data set which is such a small population that determining the interarrival time of a failure of that cause group cannot be well substantiated. Besides, it is also a failure that according to experts of Company A was highly unusual.
- *The interarrival times of failures ignores the data before the beginning of the data set*: The data set does not include the time until the first failure. Thus, the time between the first day of the data set until the first failure of a cause group is seen as the interarrival time of that failure. This excludes the history of the asset before the dataset.
- *The interarrival times of failures of the poor performing assets differ from the good performing assets*: Based on the worst 5% performing assets of the whole data set the interarrival times of the failures of this group is determined. The same is done for the good performing assets of the data set.
- *The interarrival times per cause group are based on the empirical distribution of that cause group*: Appendix C shows the graphs of the interarrival times of all cause groups. To draw random interarrival times based on an as-good-as true distribution bins per day are made. Thus 900 bins in total. Where in

reality a failure will happen somewhere during the day, the simulation uses steps of one day large. Thus, the interarrival times represent the day not the exact time.

- *A PM action is only made for the cause group that is most likely to occur as next failure*: This cause group is different from the current cause group. Figure 26 shows for each cause group, the other cause group for which PM is executed. Each row shows the current cause group and the yellow highlighted cell per row shows the conditional probability of the cause group that is most likely to occur next. Note that Figure 23 differs from Figure 26. This is because the cause group verkeerd geplaatst / wrongly placed is left out in the simulation.

P(A|B)	Accu	Besturingssysteem	Deur	Energie	Ethernet link down	Extern	Facilitair	Geen verbinding	Gelijkrichter	Geluidsklacht	Onbekend	Temperatuur	Uitval SC	WOA
Accu	0,4394	0,1107	0,0813	0,0666	0,0021	0,0014	0,0014	0,0315	0,0238	0,0007	0,0484	0,0659	0,1135	0,0133
Besturingssysteem	0,0330	0,4686	0,1624	0,0377	0,0088	0,0018	0,0031	0,0188	0,0281	0,0039	0,0661	0,0620	0,0966	0,0090
Deur	0,0131	0,0872	0,6492	0,0214	0,0087	0,0049	0,0043	0,0188	0,0268	0,0033	0,0514	0,0375	0,0678	0,0056
Energie	0,0813	0,1421	0,1412	0,3065	0,0098	0,0071	0,0009	0,0152	0,0563	0,0027	0,0804	0,0706	0,0786	0,0071
Ethernet link down	0,0072	0,1254	0,2222	0,0323	0,3513	0,0000	0,0072	0,0179	0,0323	0,0000	0,0538	0,0645	0,0789	0,0072
Extern	0,0364	0,0545	0,2970	0,0242	0,0000	0,2182	0,0061	0,0545	0,0364	0,0000	0,1273	0,0727	0,0667	0,0061
Facilitair	0,0149	0,1045	0,2836	0,0448	0,0299	0,0149	0,1940	0,0299	0,0299	0,0000	0,0448	0,0896	0,0896	0,0299
Geen verbinding	0,0456	0,1120	0,1523	0,0273	0,0104	0,0026	0,0013	0,2500	0,0234	0,0026	0,0781	0,0456	0,2305	0,0182
Gelijkrichter	0,0381	0,1398	0,1779	0,0686	0,0064	0,0025	0,0025	0,0191	0,2757	0,0089	0,0521	0,0839	0,1093	0,0152
Geluidsklacht	0,0093	0,0556	0,2222	0,0278	0,0093	0,0000	0,0000	0,0000	0,0926	0,3611	0,0463	0,1019	0,0648	0,0093
Onbekend	0,0325	0,1440	0,2556	0,0380	0,0117	0,0127	0,0056	0,0340	0,0319	0,0051	0,2799	0,0558	0,0857	0,0076
Temperatuur	0,0721	0,1243	0,1736	0,0358	0,0147	0,0041	0,0029	0,0194	0,0416	0,0035	0,0504	0,3537	0,0944	0,0094
Uitval SC	0,0400	0,1012	0,1618	0,0211	0,0073	0,0019	0,0027	0,0555	0,0249	0,0027	0,0484	0,0573	0,4620	0,0133
WOA	0,0445	0,1069	0,0869	0,0200	0,0045	0,0022	0,0045	0,0178	0,0111	0,0045	0,0290	0,0312	0,1247	0,5122

Figure 26 Conditional probability per cause group for validation

5.2.3 Flow charts of the simulation model

To clarify how the simulation works two flow charts are made. The flow chart in Figure 28 shows the schematic visualisation of the simulation where OM is executed and Figure 29 shows the schematic visualisation of the simulation where CM is executed. In section 5.2.2 the assumptions and simplification made in the simulation are further elaborated on. The flow chart shows the simulation of a single asset, with a time span of 900 days (denoted by maximum time). In the simulation there will be 10.000 iterations over the flow chart, since all assets are assumed to be identical systems. The flow of cause groups C to M are equal to the flow of cause groups A, B and N.

To clarify how the flow charts work two examples are explained, one where a failure does occur one where a failure does not occur. Figure 27 shows the flows of both examples, both start at the triangle and end at the circle.

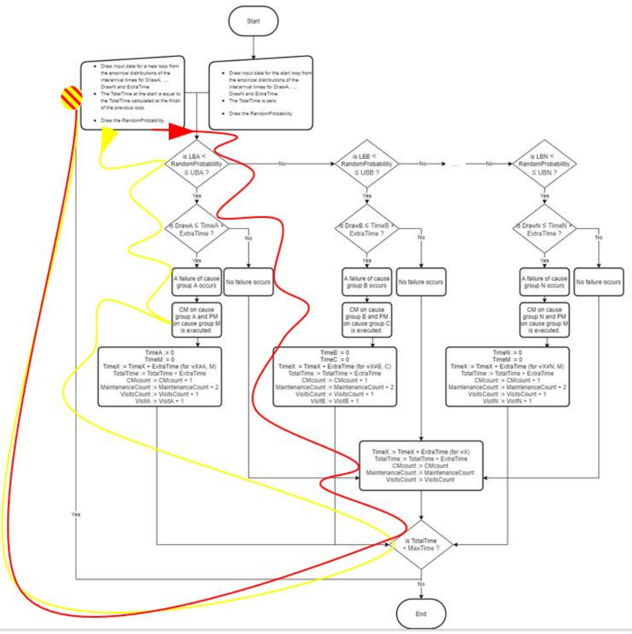


Figure 27 Example flows

In both examples the output of the previous loop which are used as starting position of the current loop are as follows:

TotalTime	= 350
Cmcount	= 6
MaintenanceCount	= 12
VisitsCount	= 6
TimeA	= 50
TimeB	= 0
TimeC	= 0
TimeD	= 200
TimeE	= 350
TimeF	= 350
TimeG	= 350
TimeH	= 350
TimeI	= 350
TimeJ	= 350
TimeK	= 350
TimeL	= 270
TimeM	= 80
TimeN	= 80

First example: failure occurs. In Figure 27 the yellow line follows the flow of the first example. It starts at the yellow triangle and ends at the red/yellow circle. For the first example the following inputs are drawn:

DrawA	= 7
DrawB	= 84
DrawC	= 43

DrawD	= 286
DrawE	= 12
DrawF	= 9
DrawG	= 48
DrawH	= 5
DrawI	= 33
DrawJ	= 435
DrawK	= 31
DrawL	= 129
DrawM	= 7
DrawN	= 87
ExtraTime	= 114
RandomProbability	= 0.0238

It is looked whether the RandomProbability falls within the lower and upper bound of cause group A. This is indeed the case since 0.0238 > 0 an 0.0238 ≤ 0.050659. thus we follow the yes arrow. Then we check if the DrawA ≤ TimeA + ExtraTime. DrawA is 7 which is indeed smaller than 164, so a failure occurs. Therefor CM on cause group A and PM on cause group M are executed. Then everything is updated as follows:

TotalTime	= 464
Cmcount	= 7
MaintenanceCount	= 14
VisitsCount	= 7
TimeA	= 0
TimeB	= 114
TimeC	= 114
TimeD	= 314
TimeE	= 464
TimeF	= 464
TimeG	= 464
TimeH	= 464
TimeI	= 464
TimeJ	= 464
TimeK	= 464
TimeL	= 384
TimeM	= 0
TimeN	= 194

Since TotalTime is smaller than MaxTime (464 < 900) we continue with another loop.

Second example: failure does not occur. In Figure 27 the red line follows the flow of the second example. It starts at the red triangle and ends at the red/yellow circle. For the second example the following inputs are drawn:

DrawA	= 320
DrawB	= 84
DrawC	= 43
DrawD	= 286

DrawE	= 12
DrawF	= 9
DrawG	= 48
DrawH	= 5
DrawI	= 33
DrawJ	= 435
DrawK	= 31
DrawL	= 129
DrawM	= 7
DrawN	= 87
ExtraTime	= 114
RandomProbability	= 0.0238

It is looked whether the RandomProbability falls within the lower and upper bound of cause group A. This is indeed the case since 0.0238 > 0 an 0.0238 ≤ 0.050659. thus we follow the yes arrow. Then we check if the DrawA ≤ TimeA + ExtraTime. DrawA is 320 which is larger than 164, so no failure occurs. Therefore no maintenance actions are executed and everything is updated as follows:

TotalTime	= 464
Cmcount	= 6
MaintenanceCount	= 12
VisitsCount	= 6
TimeA	= 164
TimeB	= 114
TimeC	= 114
TimeD	= 314
TimeE	= 464
TimeF	= 464
TimeG	= 464
TimeH	= 464
TimeI	= 464
TimeJ	= 464
TimeK	= 464
TimeL	= 384
TimeM	= 194
TimeN	= 194

Since TotalTime is smaller than MaxTime (464 < 900) we continue with another loop. Note that if one arc is followed the others are ignored for that loop. In these examples the flow chart of Figure 28 are used. The flow Figure 29 works the same with the only difference that no PM is executed only CM when a failure occurs.

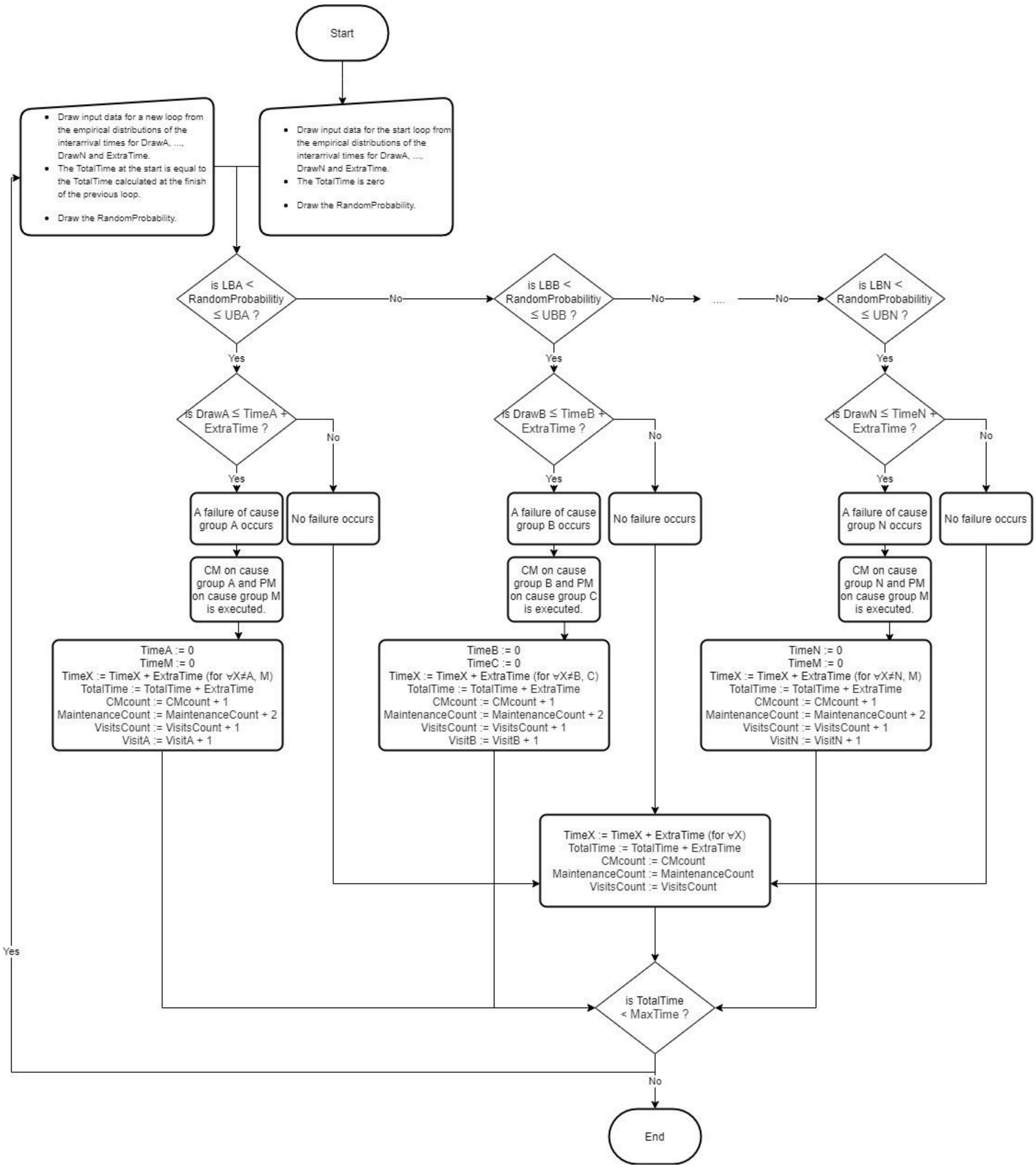


Figure 28 Flow chart of the simulation of the combination of CM and PM

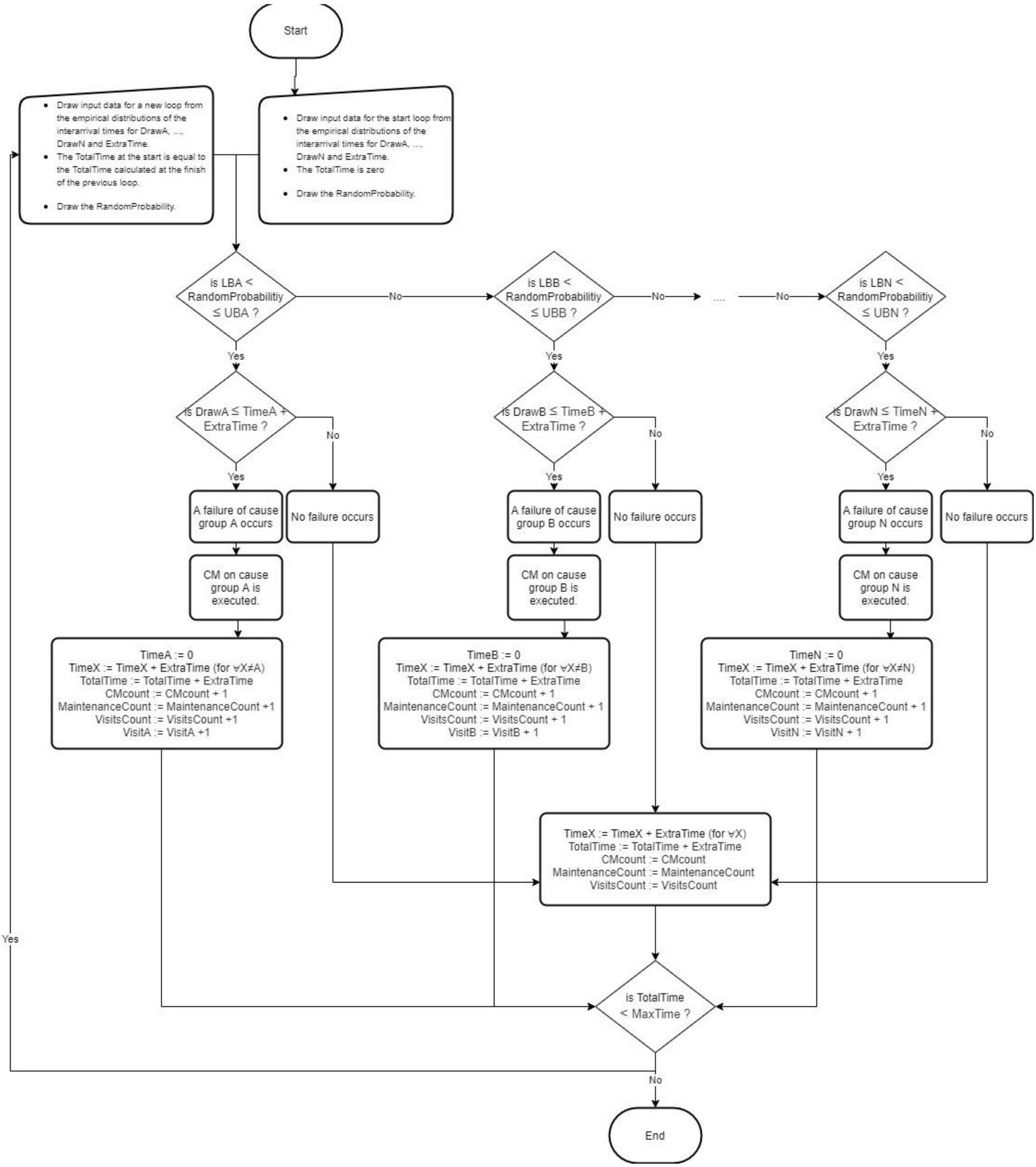


Figure 29 Flow chart of the simulation of only CM

5.2.4 Simulation results

For the simulation, the parts illustrating the poor condition assets are iterated 500 times, thereby creating 500 assets. The part illustrating the good condition assets is iterated 9500 times, which created 9500 assets. Then the current situation and proposed solution results are constructed. By combining the results of the good condition assets with the results of the poor condition assets. Table 21shows the results of the simulation.

Table 21 Validation simulation results

	1. Current situation	**2. Proposed solution**	**Difference between 1 & 2 (%)**
# of CM executed	278331	270775	-2,71475
# of maintenance executed	278331	311071	11,76297
# of visits counted	278331	270775	-2,71475

It is visible that the number of CM executed decreases with 2.71%, while the total number of maintenance executed increases with 11.76%. Thus, the increase in maintenance is $4\frac{1}{3}$ times as large as the decrease in CM.

All in all, the simulation does show that the number of failures and visits to assets decrease when implementing the proposed solution. The downside of the solution is the significant increase in total number of maintenance. The trade-off must be made by Company A whether the increase in the total number of maintenance is compensated by the decrease of failures and visits. This decision is outside the scope of this thesis. Since more information is needed for this decision, such as the impact on costs when implementing the proposed solution.

5.3 Framework for Supply Value

The contribution of this thesis for Supply Value is more knowledge concerning maintenance management. With the main focus on improvement projects for clients similar to Company A. First will be described how Supply Value can be determined whether a client is similar to Company A. Then a framework is described which can be used by Supply Value as a support when executing an improvement project.

The solution of the data analysis is not feasible for all type of companies. In chapter 2 a distinction is made between companies, based on their asset volume and asset value. Figure 30 shows the matrix belonging to the distinction. A company can be placed in one of the four quadrants. The solution is feasible for companies in the fourth quadrant. Further specified, the company must not only have a high volume of assets with a low asset volume, but these assets must also be very similar. Meaning the components of all assets are equal, thus requiring similar maintenance. For companies who fit in quadrant 1 the framework found in the literature review can be used, Table 22. For companies who fit in quadrant 3 or 4 no suggestions are given in this thesis. Since we did not look at those company types.

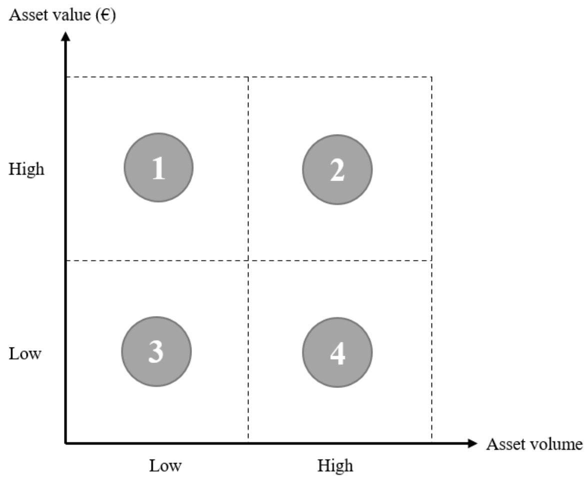


Figure 30 Company type matrix

From the literature review the steps shown in Table 22 are most often used to shape the maintenance concept framework.

Table 22 Maintenance concept framework

Step	**Definition**
1	Data collection.
2	Identify the overall maintenance objective and resources.
3	Selection and definition of the Most Important Systems (MISs).
4	Identification of the Most Critical Components (MCCs) of the MISs.
5	Selection of maintenance policy of the MISs.
6	Determination of the optimal maintenance parameters for each failure type.
7	Maintenance treatment of non-critical items.
8	Implementation and evaluation of the maintenance concept.
9	In service data collection and updating of the estimated parameters.

When applying these steps to a company with a high volume of similar assets a problem arises. Namely, how can the MISs be determined when all are similar. The frameworks studied for the literature review are aimed for companies with multiple types of assets. To determine the MISs in that situation the six big losses explained in chapter 2 can be used. This cannot be used for a company with a high volume of similar assets, such as Company A.

Another obstacle of the steps of Table 22 is the data collection. To successfully implement the framework much data must be collected and available. For instance, for the optimal parameter of step 6 much historic data of the MCCs is needed. For the research in this thesis the data needed for these parameters was not available.

Therefore, instead of the maintenance concept framework the CRISP-DM cycle was used in the thesis. From the CRISP-DM cycle and the steps of the maintenance concept framework of Table 22 a combined step-by-step framework can be created. This framework, Table 23, is aimed for companies which have a high volume of similar assets. The framework is designed after the implementation of the quantitative research at Company A was executed and is based on the on the CRISP-DM steps and the most common steps found in the literature concerning maintenance concept frameworks. The combination of implementation at Company A and the literature led to the proposed framework in this thesis.

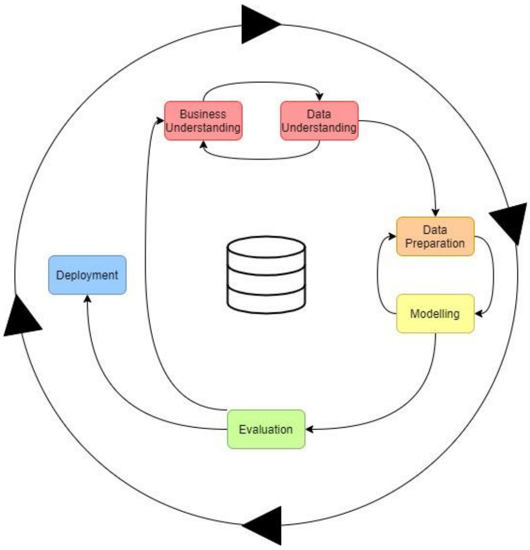


Figure 31 CRISP-DM cycle

Table 23 maintenance concept framework for a high volume of similar assets

Step	Definition
1	Business understanding and data selection. Identify the overall maintenance objective and determine which data should be collected.
2	Data collection and data understanding, collect the data and understand the data collected.
3	Data Preparation.
4	Selection and definition of the Most Important Systems (MISs) and Identification of the Most Critical Components (MCCs) of the MISs.
5	Selection of maintenance policy of both the MISs and non-important systems.
6	Implementation and evaluation of the maintenance concept.
7	In service data collection and updating.

When Supply Value receives a project request from a client who is similar to Company A, the step-by-step framework of Table 23 can be followed. How Supply Value can implement these steps will now be explained. In this explanation with the company, the client of Supply Value is meant.

5.3.1 Step 1: Business understanding and data selection

Step 1 is the business understanding and data selection step. In this step the objective of the project should be specified as well as the success criteria linked to this objective. For this good communication with the company is crucial. Thus, the objective and success criteria should be determined by both the company and Supply Value. It is important that all stakeholders who are directly and indirectly involved in the maintenance strategy change are involved as early as possible. One of the biggest challenges found in chapter 2 are the diametrically opposed ideas concerning the maintenance strategy within a company. By involving all parties early on this can be prevented.

Besides it is also determined which data can be selected and collected of the assets. This depends on the company. We divide the data into two data types, continuous monitoring data, and discrete time events data. Continuous monitoring data describes the continuous status of each asset and its components, the data is derived from e.g., sensors. Discrete time events data describes the past failures of the assets, which are discrete events. The data that can be collected can be limited because of multiple reasons. These reasons are already discussed in chapter 2:

1. The data belongs to the client, instead of the company. A company that maintenance the assets of its clients is often dependent on the client whether they give access to the continuous monitoring data.
2. The continuous monitoring data is available at the company, but not shared with Supply Value.
3. The data available is too little to determine trends in the data. A company that has a low amount of historical continuous monitoring data collected, may not successfully determine trends in their data.
4. The continuous monitoring data of different components / systems cannot be linked to one another due to the use of different software.

Even though the continuous monitoring data is missing in all cases, it is still possible to determine a maintenance strategy. In the first and second case there are two choices. Either base the maintenance strategy purely on the historical discrete time events data available or draw up a contract which does give access to the continuous monitoring data. In the second case it is preferred that continuous monitoring data will be collected as soon as possible. With the limited historical discrete time events data, it is still possible to determine a maintenance strategy but, the strategy should be re-evaluated after enough data is collected. In the fourth case the continuous monitoring data is available, but it will take much effort to correctly link all the data of the different software platforms.

When the continuous monitoring data will not be available for Supply Value, it is still possible to propose an improved maintenance strategy, as shown in the thesis. When the discrete time events data is collected the first step is completed.

5.3.2 Step 2: Data collection and data understanding

The second step revolves around data collection and data understanding. The data selected in step 1 is collected and in this step it must be understood what the collected data means. In step 1 is taken into account which data can be or already is collected. Thus, there should not be a limitation of data collection in this step since the possible limitation are identified in step 1. Besides the collection of the data the meaning of the data is determined in this step, so what does the data represent and in which form do we receive the data.

5.3.3 Step 3: Data preparation

In the third step the data given is prepared. Raw data must be transformed into usable data. To successfully transform the raw data the following steps can be followed:

1. Cleaning the raw data. Exclude noise and remove duplicates. If an observation / input is not complete there are two choice, either delete that observation or fill-in the observation to make it complete.
2. Construct new attributes. If needed new attributes can be constructed to improve the data quality.
3. Exclude attributes. When attributes in the data do not serve a function, they must be excluded.

When these steps are implemented the data is ready to be used.

5.3.4 Step 4: MISs selection and MCCs identification

In step 4 the selection and definition of the Most Important Systems (MISs) and Identification of the Most Critical Components (MCCs) of the MISs are made. First the MISs are determined. When there are various system types the MISs can be selected using the six big losses. In this case there are many similar assets, thus the variety in system types is missing. Therefore, instead of using the six big losses the MISs are selected using the historical discrete time events data and the pareto principle.

The pareto principle states that roughly 20% of the elements (the vital few) cause 80% of the effect, in this case 80% of the failures. With the historical discrete time events data can be determined whether indeed roughly 20% of all assets cause 80% of all failures. If the vital few can be determined these will be the MISs, the idea is that by focussing on these vital few the total number of failures will decrease significantly. In contrary to the MISs determined with the six big losses, the MISs identified with the pareto principle change throughout time. To identify the MISs correctly two variables need to be determined. Firstly, is determined which time span is chosen, thus what is the period looked at when determining if an asset is important or not. The time span can for instance be the last three months. Secondly the failure threshold is determined. If the number of failures of an asset is higher or equal to the threshold during the time span it is important, if the number of failures is lower the asset is not important.

After the MISs are determined the MCCs must be determined. The MCCs are the components that have the biggest impact on the reliability of the system. Often the FMECA to determine the MCCs.

5.3.5 Step 5: Maintenance policy selection

Step 5 Selection of maintenance policy of both the MISs and non-important systems. Figure 32 shows the different maintenance types. In this step the maintenance type per component of the MISs describes the maintenance policy of the MISs and the maintenance type per component of the non-important systems describe the maintenance policy of the non-important systems. When dealing with assets of the fourth quadrant of Figure 30 the best policy is TBM or CM on all components of the non-important systems. For the MISs, the best policy is PM on the MCCs and PM or CM on non-critical components. Besides the choice of maintenance types also the parameters belonging to the maintenance types are determined. An example of such a parameter is the time to replace a component when TBM is applied. Much time is needed to successfully complete this step. Also, enough historical continuous monitoring data is needed to correctly determine the parameters. When time is short it is recommended that Supply Value does not determine the parameters only the maintenance types.

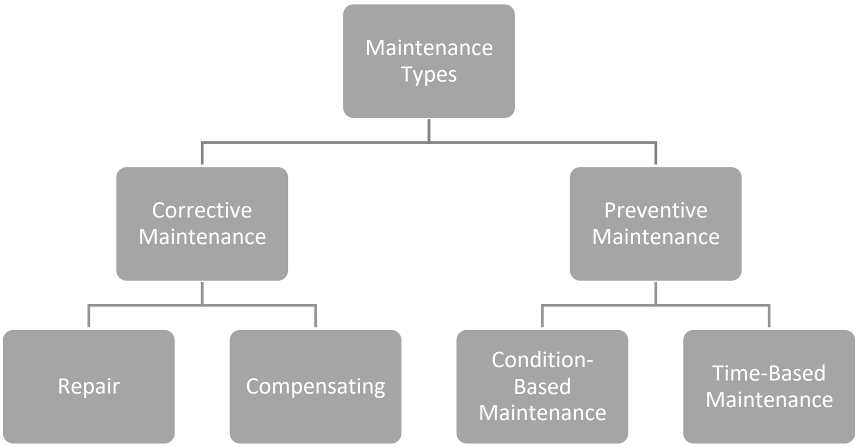


Figure 32 All maintenance types

5.3.6 Step 6: Implementation and evaluation

In step 6 the maintenance concept is implemented and evaluated. In this step the maintenance concept will be implemented. Most likely the implementation and evaluation will be executed by the company, not by Supply Value. To show the company the impact of the proposed maintenance concept, Supply Value can create a simulation to validate the concept.

5.3.7 Step 7: In-service data collection and updating

Step 7 is the in-service data collection and updating step. After the company decides to implement the maintenance concept the parameters must be updated using in service data collection. This step is executed by the company. It is important to not only update the parameters, but also evaluate the quality of the maintenance concept when the concept is in service. It is recommended that this step will not be executed by Supply Value, but by the company.

6. Conclusion and recommendations

In this chapter the conclusion and recommendations of this thesis are discussed. In section 6.1 the conclusion is further elaborated on. In section 6.2 the recommendations for Company A and Supply Value are discussed. At last a discussion is held in section 6.3 concerning how this thesis contributes to the literature and practice.

6.1 Conclusion

This thesis aimed to answer the central research question, with the aid of the CRISP-DM methodology. The central research question was: 'How can a maintenance strategy be improved, by implementing asset categorisation?'. The thesis was executed at Supply Value and the quantitative part of the thesis was executed with data from Company A.

At first the current situation at Supply Value, Company A and Companies B to G were determined. Supply Value has published a white paper and an insight thus far. These contain the current knowledge of Supply Value, which is purely theoretical. Since Supply Value has no practical knowledge yet. In other words, no project concerning maintenance is executed by Supply Value thus far. The theoretical knowledge of Supply Value can be place in the third generation, the practical knowledge is missing.

Company A is specialized in infrastructure management and asset management. This thesis connected to the asset management branch within Company A. Currently Company A executes CM and low-key PM, which places Company A mainly in the first generation. Company A wants to expand the maintenance types it carries out for its client. Besides the CM it offers at the moment, Company A wants to include CBM, because Company A experiences high fluctuation in workload and a high number of repeat outages.

The current situations of companies B-G vary a lot, Figure 33 show for Supply Value (SV) and companies A to G in which generation they currently are. The literature states that companies are currently in the fourth generation, it is clear that this is not the case for most companies, thus there is a gap between the literature and the practice.

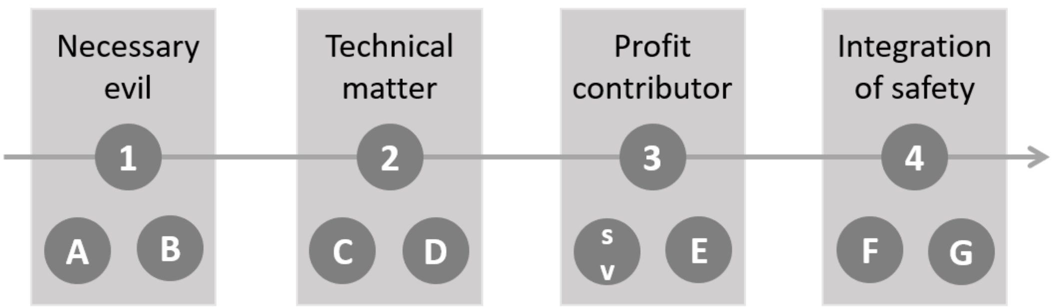


Figure 33 Companies and generations

The data analysis focussed on Company A. For the data analysis a data set including 12911 assets is used. The goal of the data analysis is to improve the maintenance quality by increasing knowledge on asset categorisation and by increasing knowledge about conditional probability between failure causes of consecutive failures.A rule to correctly categorise assets is created, to determine in which category an asset belongs (good or poor condition) it is looked at the number of failures of the last three months. If the ticket number is equal or higher than two the asset belongs to the poor condition category otherwise the asset belongs to the good condition category. Also, the conditional probability between cause groups is determined.

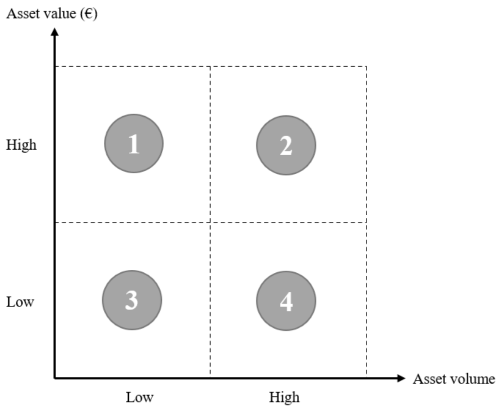
The proposed solution for Company A is to first divide the assets into good and poor condition categories based on the rule above. The good condition category will be maintained as it is maintained currently. Thus, only CM is applied on these assets. For the assets in the poor condition category, it is recommended to implement OM, thus include an PM action when a failure of the asset occurs. The PM actions to execute followed from the conditional probabilities found. Since it is not clear yet how failures of different cause groups interact with each other, it is assumed that by executing PM on the cause group most likely to occur as the next failure, the total number of failures will drop.

A simulation is setup, to determine the impact of the proposed solution. In the simulation the current situation is simulated as well as the proposed solution. Both are tested on three KPI's: the number of corrective maintenance executed, the total number of maintenance executed and the number of visits. The results show a decrease of 2.71% of the number of CM executed and the number of visits, when comparing the current situation to the proposed solution. The total number of maintenance executed increased with 11.76%. Besides this it is visible that the average number of failures per asset decreased with 2,5%.

All in all, the simulation does show that the number of failures and visits to assets decrease when implementing the proposed solution. The downside of the solution is the significant increase in total number of maintenance. The trade-off must be made by Company A whether the increase in the total number of maintenance is compensated by the decrease of failures and visits. This decision is outside the scope of this thesis. Since more information is needed for this decision, such as the impact on costs when implementing the proposed solution.

For Supply Value a framework is created which can guide consultants at Supply Value with future projects concerning similar maintenance management issues. Since the thesis was conducted for a specific company type, Supply Value must first assess whether the company of a future project is similar to Company A, thus if the future company also fits in the fourth quadrant. Then a step-by-step maintenance concept framework is created for Supply Value, that can be followed. Companies that fit in the first quadrant can follow the maintenance concept framework found in the literature review. There are no suggestions for companies that fit in the second or third quadrant.

Step	Definition
1	Data collection.
2	Business and data understanding, identify the overall maintenance objective and understand the given data.
3	Data Preparation.
4	Selection and definition of the Most Important Systems (MISs) and Identification of the Most Critical Components (MCCs) of the MISs.
5	Selection of maintenance policy of both the MISs and non-important systems.
6	Implementation and evaluation of the maintenance concept.
7	In service data collection and updating.



6.2 Recommendations

In this section the recommendations for Company A and Supply Value are given. Furthermore, the general recommendations concerning future research are given. The general recommendation is to further research the proposed solution of dividing assets into categories based on their performance. Since the solution is now only tested at company A. Furthermore, recommendations for Company A and Supply Value are given.

6.2.1 Recommendations for Company A

The first recommendation for Company A is to implement the proposed threshold to categorise the assets and adjust the maintenance for the assets in the poor condition category. The best type of PM for the poor condition category is not determined in this thesis, mainly due to the limited data set available.

The second recommendation is to further research the cause of failure of different components. At the moment of some failure is known what the cause was that led to the failure, but for most failures that is not known yet. To successfully implement PM, it must be known where failures originate from.

The third recommendation is to research the best PM action, e.g. time-based maintenance or condition-based maintenance, for each component. Including the parameters belonging to the PM actions.

The road map in Table 24 gives a guideline which steps should be taken by Company A, if they want to implement the solution and recommendations. The now column indicates what actions should be started as soon as possible. The next column shows the steps after the initial actions and the later indicate the actions in the long term. The identification of the optimal parameters is placed in later since this is a complex process which will take a long time.

Table 24 Roadmap for Company A

Now	Next	Later
• Select a project team to implement the categorisation of assets and determine the correct maintenance actions per group. • Define the business objective and success criteria. • Set up a project plan to start the transition to categorisation of assets using the CRISP-DM method as guideline. • Align the goal of the project with all stakeholders involved.	• Define what data is needed and start the data collection. • Analyse what the correct PM action per component is. • Implement the categorisation of components. • Implement the PM actions which are possible to implement.	• Identify the optimal parameters for all PM actions. • Update category threshold if needed. • Evaluate the results of categorisation and improve if needed.

6.2.2 Recommendations for Supply Value

For Supply Value the main recommendation is to further expand their knowledge about maintenance management, by executing maintenance management projects. Furthermore, it is recommended that Supply Value will specialize in how to overcome one or more of the following challenges:

1. The lack of data sharing can impede the shift towards PM;
2. The different cloud-based solution offered cannot be connected with each other;
3. People within a Company have different ideas about the best maintenance strategy, these ideas are often diametrically opposed.

One of the focus areas of Supply Value is the collaboration within networks. Therefore, it is recommended that they focus on the first challenge. Besides Supply Value already has done projects concerning optimizing IT systems, which have common grounds with the second challenge. The third challenge has to do with not successfully implement chance within a company. Supply Value already has knowledge concerning project and change management, by specifying this knowledge towards maintenance the third challenge can be overcome. Thus, the recommendation for Supply Value is to use their knowledge and implement this on maintenance projects, while at the same time expand their maintenance knowledge.

6.2.3 Recommendations for future research

The data analysis of this thesis is executed at one company, Company A. A research group of only one company is too small to say if the impact of implementing categorisation on assets will be positive in general. Thus a research implementing asset categorisation and executing CM for the good condition category and OM for the poor condition category in multiple companies, that fit in the fourth quadrant of the company type matrix, is needed to determine the general effect of asset categorisation. Researching the categorisation at multiple companies will take a significant amount of time, which will span multiple years.

Another research opportunity is to use the suggested framework of chapter 5 at multiple companies. This framework is based on the data analysis and literature review of this thesis but is not tested in practise yet. Executing this framework would be recommended future research. The execution of this recommendation will span multiple years.

6.3 Discussion

In this section is discussed how this thesis contributes to the literature and practice. First the contribution to the literature is discussed, then the contribution to the practice.

The maintenance concept frameworks found in the literature discussed in chapter 3 all identify the MISs for companies which have several system types. How to determine the MISs for companies with only one system type is not covered in the frameworks found in literature. This thesis comes up with a suggestion how the MISs can be determined if a company has many assets of a single system type. By categorising the assets based on the number of tickets in a predetermined interval. Where the poor condition category will be the MISs. This is a contribution to the literature. The framework guiding the maintenance strategy for companies with a high number of the same system assets, also contributes to the literature concerning maintenance concept frameworks.

The contribution to the practice is divided into the contribution to Company A and the contribution to Supply Value. By using ticket data, increasing the knowledge concerning asset categorisation, malfunction causes and conditional probability between causes, this thesis gained insight into how Company A can improve their maintenance quality. This new insight gained is the contribution to Company A. The contribution to Supply Value is the framework of chapter 5, which Supply Value can use as a guideline when executing maintenance projects at clients which fall in the fourth quadrant of the company type matrix.

References

Ab-Samat, H., & Kamaruddin, S. (2014).Opportunistic maintenance (OM) as a new advancement in maintenance approaches: A review. *Journal of Quality in Maintenance Engineering, 20*(2), 98-121. https://doi.org/10.1108/JQME-04-2013-0018

Aggarwal, C. (2015). Data Classification. *Data mining* (285-344). Cham, Switzerland: Springer

Akkermans, H. (2020, 11, 11). Business models for CBM-driven smart services[pdf]. Retrieved from https://primavera-project.com/wp-content/uploads/2020/11/Lecture-Primavera-Consortium-Nov-11-2020-002.pdf

Almeanazel, O.T. (2010). Total Productive Maintenance Review and Overall Equipment Effectiveness Measurement. *Jordan Journal of Mechanical and Industrial Engineering, 4*(4), 517-522.

Armstrong, C.M., Niinimäki, K., Kujala, S., Karell, E., & Lang, C. (2015). Sustainable product-service systems for clothing: exploring consumer perceptions of consumption alternatives in Finland. *Journal of Cleaner Production, 97*, 30-39. https://doi.org/10.1016/j.jclepro.2014.01.046.

Arunraj, N.S., & Maiti, J. (2007). Risk-based maintenance: Techniques and applications. *Journal of Hazardous Materials, 142*(3), 653-611. https://doi.org/10.1016/j.jhazmat.2006.06.069

Baines, T.S., & Lightfoot, H. (2013). *Made to Serve: Understanding What It Takes for a Manufacturer to Compete Through Servitization and Product-Service Systems.* Hoboken, NJ: Wiley.

Barahona, I., Riba, A., & Yang, J. (2015). Analytical tool adoption level: a case study based on an evidential reasoning approach. *International Journal of Transitions and Innovation Systems*. *4*, 22-44. https://doi.org/10.1504/IJTIS.2015.074644

Bastl, M., Johnson, M., Lightfoot, H., & Evans, S. (2012). Buyer-supplier relationships in a servitized environment: An examination with Cannon and Perreault's framework. *International Journal of Operations & Production Management, 32*(6), 650-675. https://doi.org/10.1108/01443571211230916.

Batista, L., Davis-Poynter, S., Ng, I., & Maull, R. (2017). Servitization through outcome-based contract – A systems perspective from the defence industry. *International Journal of Production Economics, 192*, 133-143. https://doi.org/10.1016/j.ijpe.2016.12.005.

Berrade, M.D., Scarf, P.A., & Cavalcante, C.A.V. (2017). A study of postponed replacement in a delay time model. *Reliability Engineering & System Safety, 168*(2017), 70-79. https://doi.org/10.1016/j.ress.2017.04.006.

Bocken, N.M.P., Mugge, R., Bom, C.A., & Lemstra, H.J. (2018). Pay-per-use business models as a driver for sustainable consumption: Evidence from the case of HOMIE. *Journal of Cleaner Production, 198*, 498-510. https://doi.org/10.1016/j.jclepro.2018.07.043.

Boomija, M.C.A., & Phil, M. (2008). COMPARISON OF PARTITION BASED CLUSTERING ALGORITHMS. *Journal of Computer Applications, 1*(4), 18-21

Brockwell, P.J., & Davis, R.A. (1991). Stationary time series. In *Time series : theory and methods (2nd ed.)* (pp. 1-39). New York, NY: Springer-Verlag.

Dal, B., Tugwell, P., & Greatbanks, R. (2000). Overall equipment effectiveness as a measure of operational improvement – A practical analysis. *International Journal of Operations & Production Management, 20*(12), 1488-1502. https://doi.org/10.1108/01443570010355750

Dekker, R., & Smeitink, E. (1991). Opportunity-based block replacement. *European Journal of Operational Research, 53*(1), 46-63. https://doi.org/10.1016/0377-2217(91)90092-A

Díaz, C., Batanero, C., & Contreras, J. (2010). Teaching independence and conditional probability. *Boletín de Estadística e Investigación Operativa, 26*(2), 149-169.

Ding, F., & Tian, Z. (2012). Opportunistic maintenance for wind farms considering multi-level imperfect maintenance thresholds. *Renewable Energy, 45*(2012), 175-182. https://doi.org/10.1016/j.renene.2012.02.030.

D'Oca, S., & Hong, T. (2015). Occupancy schedules learning process through a data mining framework. *Energy and Buildings, 88*(2015), 395-408. https://doi.org/10.1016/j.enbuild.2014.11.065

Dogan, N., & Tanrikulu, Z.A. (2013). Comparative analysis of classification algorithms in data mining for accuracy, speed and robustness. *Information Technology and Management 14*(2), 105-124. https://doi.org/10.1007/s10799-012-0135-8

Durocher, D.B., & Feldmeier, G.R. (2004). Predictive versus preventive maintenance. *IEEE Industry Applications Magazine, 10*(5), 12-21. https://doi.org/10.1109/MIA.2004.1330766

Faccio, M., Persona, A., & Sgarbossa, F. (2014). Industrial maintenance policy development: A quantitative framework. *International Journal of Production Economics,147*(part A), 85-93. https://doi.org/10.1016/j.ijpe.2012.08.018

Garg, A., & Deshmukh, S.G. (2006). Maintenance management: Literature review and directions. *Journal of Quality in Maintenance Engineering, 12*(3), 205-238. https://doi.org/10.1108/13552510610685075

Gits, C.W. (1992). Design of maintenance concepts. *International Journal of Production Economics, 24*(3), 217-226. https://doi.org/10.1016/0925-5273(92)90133-R

Goyal, D., & Pabla, B.S. (2015). Condition based maintenance of machine tools: A review. *CIRP Journal of Manufacturing Science and Technology, 10*(2015), 24-35. https://doi.org/10.1016/j.cirpj.2015.05.004

Guajardo, J.A., Cohen, M.A., Kim, S.H., & Netessine, S. (2012). Impact of Performance-Based Contracting on Product Reliability: An Empirical Analysis. *Management Science, 58*(5), 961-979. https://doi.org/10.1287/mnsc.1110.1465

Gupta, A., & Lawsirirat, C. (2006). Strategically optimum maintenance of monitoring-enabled multi-component systems using continuous-time jump deterioration models. *Journal of Quality in Maintenance Engineering, 12*(3), 306-329. https://doi.org/10.1108/13552510610685138

Gupta, G., & Mishra, R.P. (2016). A SWOT analysis of reliability centered maintenance framework. *Journal of Quality in Maintenance Engineering, 22*(2), 130-145. https://doi.org/10.1108/JQME-01-2015-0002

Hájek, A. (2011). Conditional Probability. In Philosophy of Statistics, *Philosophy of Science, 7*, 99-135. https://doi.org/10.1016/B978-0-444-51862-0.50003-4.

Han, J., Kamber, M. & Pei, J. (2012). Classification: Basic Concepts. *Data mining concepts and techniques* (3rd ed. 372-391). Waltham, Massachusetts: Morgan Kaufmann Publishers

Van der Heijden, M. (2020). Maintenance optimization I [PowerPoint slides]. Retrieved from https://canvas.utwente.nl/courses/4933/files/1398317?module_item_id=124437

Van der Heijden, M. (2020). Maintenance optimization II [PowerPoint slides]. Retrieved from https://canvas.utwente.nl/courses/4933/files/1405185?module_item_id=124438

Hoffman, A.J., & van der Westhuizen, M. (2019). Empirical model for truck route fuel economy. *Proceedings of the Conference IEEE Intelligent Transportation Systems Conference,* 914-919. https://doi.org/10.1109/ITSC.2019.8917185

Homburg, C., Fassnacht, M. and Guenther, C. (2003). The role of soft factors in implementing a serviceoriented strategy in industrial marketing companies. *Journal of Business-to-Business Marketing, 10(*2), 23-51. https://doi.org/10.1300/J033v10n02_03

Huang, H., Wu, X., & Relue, R. (2002). Association analysis with one scan of databases. *In Proceedings of the IEEE International Conference on Data Mining,* 629-632. https://doi.ort/10.1109/ICDM.2002.1184015.

Huber, S., Wiemer, H., Schneider, D., & Ihlenfeldt, S. (2019). DMME: Data mining methodology for engineering applications – a holistic extension to the CRISP-DM model. *Procedia CIRP*, *79*, 403-408, https://doi.org/10.1016/j.procir.2019.02.106.

Hypko, P., Tilebein, M., & Gleich, R. (2010). Clarifying the concept of performance-based contracting in manufacturing industries: A research synthesis. Journal of Service Management, 21(5), 625-655.

Jardine, A.K.S., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing, 20*(7), 1483-1510. https://doi.org/10.1016/j.ymssp.2005.09.012

Juran, J.M. (1954). Universals in management planning and controlling. *The management review, 43*(November), 748-761.

Kalyani, R.M. (2012). Data Mining Techniques. *International Journal of Advanced Research in Computer Science and Software Engineering*, *2*(10), 439-442.

Kameshwaran, K., & Malarvizhi, K.(2014). Survey on Clustering Techniques in Data Mining. *International Journal of Computer Science and Information Technologies*, *5*(2) , 2272-2276

Kim, S.H., Cohen, M.A., & Netessine, S. (2007). Performance Contracting in After-Sales Service Supply Chains. *Management Science, 53*(12), 1843-1858. https://doi.org/10.1287/mnsc.1070.0741

Kim, J., Ahn, Y., & Yeo, H. (2016). A comparative study of time-based maintenance and condition-based maintenance for optimal choice of maintenance policy. *Structure and infrastructure Engineering, 12*(12), 1525-1536. https://doi.org/10.1080/15732479.2016.1149871

Kolmogorov, A. N. (1960). Foundations of the Theory of Probability. Chelsea Pub Co.

Koochaki, J., Bokhorst, J.A.C., Wortmann, H., & Klingenberg, W. (2011). Condition based maintenance in the context of opportunistic maintenance. *International Journal of Production Research*, *50*(23), 6918-6929. https://doi.org/10.1080/00207543.2011.636924

Kowalkowski, C., Gebauer, H., Kamp, B., & Parry, G. (2017) Servitization and deservitization: Overview, concepts, and definitions. *Industrial Marketing Management, 60*, 4-10. https://doi.org/10.1016/j.indmarman.2016.12.007.

Kumar, S., & Toshniwal, D. (2015). A data mining framework to analyze road accident data. *Journal of Big Data 2*(26). https://doi.org/10.1186/s40537-015-0035-y.

Lin, N.P., Chang, C.I., Chueh, H.E., Chen, H.J., & Hao, W.H. (2008). A Deflected Grid-based Algorithm for Clustering Analysis. *WSEAS Transactions on Computers, 7*(3), 125-132.

Mitra, S., Pal, S.K., & Mitra, P. (2002). Data mining in soft computing framework: a survey. *IEEE Transactions on Neural Networks, 13*(1), 3-14. https://doi.org/10.1109/72.977258.

Moubray, J. (1997). *Reliability-centered Maintenance* (2nd ed.). New York, NY: Industrial press inc.

Muchiri, P., Pintelon, L., Gelders, L., & Martin, H. (2011). Development of maintenance function performance measurement framework and indicators. *International Journal of Production Economics*, *131*(1), 295-302. https://doi.org/10.1016/j.ijpe.2010.04.039

Nadali, A., Kakhky, E., & Nosratabadi, H. (2011). Evaluating the success level of data mining projects based on CRISP-DM methodology by a Fuzzy expert system. *Proceedings of the ICECT 2011 - 2011 3rd International Conference on Electronics Computer Technology, 6*, 161-165. https://doi.org/ 10.1109/ICECTECH.2011.5942073.

Neely, A. (2008). Exploring the financial consequences of the servitization of manufacturing. *Operations Management Research, 1*(2), 103–118. https://doi.org/10.1007/s12063-009-0015-5

Ng, C.L.I., Maull, R., & Yip, N. (2009). Outcome-based contracts as a driver for systems thinking and service-dominant logic in service science: Evidence from the defence industry. *European Management Journal, 27*(6), 377-387. https://doi.org/10.1016/j.emj.2009.05.002.

Ng, I.C.L., & Nudurupati, S.S. (2010). Outcome-based service contracts in the defence industry – mitigating the challenges. *Journal of Service Management*, *21*(5), 656-674. https://doi.org/10.1108/09564231011079084

Ng, I.C.L., Xin Ding, D., & Yip, N. (2013). Outcome-based contracts as new business model: The role of partnership and value-driven relational assets. *Industrial Marketing Management*, *42*(5), 730-743. https://doi.org/10.1016/j.indmarman.2013.05.009.

Van Noortwijk, J.M. (2007). A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety, 94*(1), 2-21. https://doi.org/10.1016/j.ress.2007.03.019

Olegas, N. (2015). CRISP Data Mining Methodology Extension for Medical Domain. *Baltic Journal of Modern Computing, 3*, 92-109.

Pongpech, J., & Murthy, D.N.P. (2006). Optimal periodic preventive maintenance policy for leased equipment. *Reliability Engineering & System Safety, 91*(7), 772-777. https://doi.org/10.1016/j.ress.2005.07.005

Rabetino, R., Kohtamäki, M., & Gebauer, H. (2017). Strategy map of servitization. International Journal of Production Economics, 192, 144-156. https://doi.org/10.1016/j.ijpe.2016.11.004.

Raheja, D., Llinas, J., Nagi, R., & Romanowski, C. (2007). Data fusion / data mining-based architecture for condition-based maintenance. *International Journal of Production Research, 44*(14), 2869-2887. https://doi-org.ezproxy2.utwente.nl/10.1080/00207540600654509

Rausand, M., & Hoyland, A. (2004). *System Reliability Theory: Models, Statistical Methods, and Applications* (2nd ed.), *Reliability of maintained systems* (pp. 361-418). Hoboken, NJ: John Wiley & Sons Inc.

Rebentrost, P., Mohseni, M., & Lloyd, S. (2014). Quantum Support Vector Machine for Big Data Classification. *Physical review letters 113*(13), 130503(1)-130503(4). https://doi.org/10.1103/PhysRevLett.113.130503

Sajana, T., Rani, C.M., & Narayana, V.K. (2016). A Survey on Clustering Techniques for Big Data Mining. *Indian Journal of Science and Technology, 9*(3), 1-12. https://doi.org/10.17485/ijst/2016/v9i3/75971.

Saranga, H. (2004). Opportunistic maintenance using genetic algorithms. Journal of Quality in Maintenance Engineering, 10(1), 66-74. https://doi.org/10.1108/13552510410526884

Shah, C., & Jivani, A.G. (2013). Comparison of data mining classification algorithms for breast cancer prediction. *Proceedings of the 2013 Fourth International Conference on Computing, Communications and Networking Technologies*, 1-4. https://doi.org/10.1109/ICCCNT.2013.6726477.

Shen, C.P., Jigjidsuren, C., Dorjgochoo, S., Chen, C.H., Chen, W.H., Hsu, C.K., … Lai, F. (2012). A Data-Mining Framework for Transnational Healthcare System. *Journal of Medical Systems, 36*(4), 2565-2575. https://doi.org/10.1007/s10916-011-9729-7

Skaggs, B.C., & Droege, S.B. (2004). The performance effects of service diversification by manufacturing firms. Journal of Managerial Issues, 16(3) , 396-407.

Sousa, R. & da Silveira, G.J.C. (2017). Capability antecedents and performance outcomes of servitization: Differences between basic and advanced services. *International Journal of Operations & Production Management*, *37*(4), 444-467. https://doi-org./10.1108/IJOPM-11-2015-0696

Supply Value (2020). *Whitepaper: Predictive maintenance.*

Supply Value (2020)[1]. *Article: Predictive maintenance.*

Vandermerwe, S., & Rada, J. (1988). Servitization of business: Adding value by adding services. *European Management Journal, 6*(4), 314-324. https://doi.org/10.1016/0263-2373(88)90033-3.

Visnjic, I., Neely, A. & Wiengarten, F. (2012). Another performance paradox? A refined view on the performance impact of servitization. (working paper). *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2117043

Visnjic, I., Neely, A., & Jovanovic, M. (2018). The path to outcome delivery: Interplay of service market strategy and open business models. *Technovation, 72–73*, 46-59.https://doi.org/10.1016/j.technovation.2018.02.003.

Waeyenbergh, G., & Pintelon, L. (2003). Maintenance concept development: A case study. *International Journal of Production Economics, 89*(3), 395-405. https://doi.org/10.1016/j.ijpe.2003.09.008

Waeyenbergh, G., & Pintelon, L. (2009). CIBOCOF: A framework for industrial maintenance concept development. *International Journal of Production Economics, 121*(2), 633-640. https://doi.org/10.1016/j.ijpe.2006.10.012

Wang, W., Lai, K., & Shou, Y (2018). The impact of servitization on firm performance: a meta-analysis. *International Journal of Operations & Production Management, 38*(4), 1562-1588. 10.1108/IJOPM-04-2017-0204.

Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining,* 29-3.

Wu, W., Walczak, B., Massart, D.L., Heuerding, S., Erni, F., Last, I.R., & Prebble, K.A. (1996). Artificial neural networks in classification of NIR spectral data: Design of the training set. *Chemometrics and Intelligent Laboratory Systems, 33*(1), 35-46. https://doi.org/10.1016/0169-7439(95)00077-1.

Xiang, Y., Zhu, Z., Coit, D.W., & Feng, Q. (2017).Condition-based maintenance under performance-based contracting. *Computers and Industrial Engineering, 111*, 391-402.https://doi.org/10.1016/j.cie.2017.07.035.

Yang, L., Ma, X., Peng, R., Zhai, Q., & Zhao, Y. (2017). A preventive maintenance policy based on dependent two-stage deterioration and external shocks. *Reliability Engineering & System Safety, 160*(2017), 201-211. https://doi.org/10.1016/j.ress.2016.12.008

Yang, L., Ye, Z.S., Lee, C.G., Yang, S.F., & Peng, R. (2019). A two-phase preventive maintenance policy considering imperfect repair and postponed replacement. *European Journal of Operational Research, 274*(3), 966-977. https://doi.org/10.1016/j.ejor.2018.10.049

Yeung, K.Y., Fraley, C., Murua, A., Raftery, E., & Ruzzo, W.L. (2001). Model-based clustering and data transformations for gene expression data. Bioinformatics, 17(10), 977-987. https://doi.org/10.1093/bioinformatics/17.10.977

Appendix

Appendix A: Interval lengths and ticket distribution

Years

Variables	Value	2017	2018	2019	2020
Mean	.	1	2	2	1
Percentile	0,8	2	3	3	2
	0,85	2	3	3	2
	0,9	2	4	4	2
	0,95	3	6	6	3

Quarters

Variables	Value	1-18	2-18	3-18	4-18	1-19	2-19	3-19	4-19	Average
Mean	.	0	0	0	0	0	0	0	0	0
Percentile	0,8	0	1	1	0	1	1	1	0	1
	0,85	1	1	1	1	1	1	1	1	1
	0,9	1	1	1	1	1	1	1	1	1
	0,95	2	2	2	2	2	2	2	2	2

Seasons

Variables	Value	Winter 17-18	Spring 18	Summer 18	Autumn 18	Winter 18-19	Spring 19	Summer 19	Autumn 19	Winter 19-20	Average
Mean	.	0,240802417	0,303771977	0,332042444	0,28409883	0,304933777	0,361397258	0,376887925	0,295871737	0,25853923	0
Percentile	0,8	0	1	1	0	1	1	1	1	0	1
	0,85	1	1	1	1	1	1	1	1	1	1
	0,9	1	1	1	1	1	1	1	1	1	1
	0,95	1	2	2	2	2	2	2	2	1	2

Months

Column1	2018-1	2018-1	2018-1	2018-2	2018-3	2018-4	2018-5	2018-6	2018-7	2018-8	2018-9	2019-1	2019-1	2019-1	2019-1	2019-2	2019-3	2019-4	2019-5	2019-6	2019-7	2019-8	2019-9
Mean	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
85	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
90	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	0
95	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

2 and 3 Months sliding

Bimonthly	1218	2318	3418	4518	5618	6718	7818	8918	91018	101118	111218	120119	1219	2319	3419	4519	5619	6719	7819	8919	91019	101119	111219
Mean	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
85	0	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	0
90	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
95	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	2	2	1	1	1	1

3months	12318	23418	34518	45618	56718	67818	78918	891018	9101118	10111218	11120119	12010219	12319	23419	34519	45619	56719	67819	78919	891019	9101119	10111219
Mean	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
80	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
85	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
90	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
95	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

Appendix B: Conditional probability

Transformation step 1

TOTAAL(row leads to column)	Accu	Besturing:	Deur	Energie	Ethernet l	Extern	Facilitair	Geen verb	Gelijkricht	Geluidskl	Onbekend	Temperat	Uitval stre	Verkeerd	WOA	Total
Accu	627	158	116	95	3	2	2	45	34	1	69	94	162	0	19	1427
Besturingssysteem	128	1815	629	146	34	7	12	73	109	15	256	240	374	0	35	3873
Deur	110	730	5435	179	73	41	36	157	224	28	430	314	568	0	47	8372
Energie	91	159	158	343	11	8	1	17	63	3	90	79	88	0	8	1119
Ethernet link down	2	35	62	9	98	0	2	5	9	0	15	18	22	0	2	279
Extern	6	9	49	4	0	36	1	9	6	0	21	12	11	0	1	165
Facilitair	1	7	19	3	2	1	13	2	2	0	3	6	6	0	2	67
Geen verbinding	35	86	117	21	8	2	1	192	18	2	60	35	177	0	14	768
Gelijkrichter	30	110	140	54	5	2	2	15	217	7	41	66	86	1	12	788
Geluidsklacht	1	6	24	3	1	0	0	0	10	39	5	11	7	0	1	108
Onbekend	64	284	504	75	23	25	11	67	63	10	552	110	169	0	15	1972
Temperatuur	123	212	296	61	25	7	5	33	71	6	86	603	161	0	16	1705
Uitval streetcabinet	148	374	598	78	27	7	10	205	92	10	179	212	1708	0	49	3697
Verkeerd geplaatst	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
WOA	20	48	39	9	2	1	2	8	5	2	13	14	56	0	230	449
Total	1386	4033	8186	1080	312	139	98	828	923	123	1820	1814	3596	1	451	24790

Transformation steps 2 to 7

P(A∩B)	Accu	Besturingssyste	Deur	Energie	Ethernet link do	Extern	Facilitair	Geen verbindi	Gelijkrichter	Geluidsklach	Onbekend	Temperatuur	Uitval SC	Verkeerd geplaa	WOA
Accu	0,025292457	0,006373538	0,004679306	0,00383219	0,000121017	8,06777E-05	8,06777E-05	0,001815248	0,001371521	4,03388E-05	0,00278338	0,003791852	0,006534893	0	0,000766438
Besturingssyste	0,005163372	0,073215006	0,025373134	0,005889472	0,001371521	0,000282372	0,000484066	0,002944736	0,004396934	0,000605083	0,010326745	0,009681323	0,015086729	0	0,00141186
Deur	0,004437273	0,029447358	0,21924163	0,007220653	0,002944736	0,001653893	0,001452198	0,006333199	0,009035902	0,001129488	0,017345704	0,012666398	0,022912465	0	0,001895926
Energie	0,003670835	0,006413877	0,006373538	0,013836224	0,000443727	0,000322711	4,03388E-05	0,00068576	0,002541347	0,000121017	0,003630496	0,003186769	0,003549818	0	0,000322711
Ethernet link do	8,06777E-05	0,00141186	0,002501008	0,00036305	0,003953207	0	8,06777E-05	0,000201694	0,00036305	0	0,000605083	0,000726099	0,000887455	0	8,06777E-05
Extern	0,000242033	0,00036305	0,001976603	0,000161355	0	0,001452198	4,03388E-05	0,00036305	0,000242033	0	0,000847116	0,000484066	0,000443727	0	4,03388E-05
Facilitair	4,03388E-05	0,000282372	0,000766438	0,000121017	8,06777E-05	4,03388E-05	0,000524405	8,06777E-05	8,06777E-05	0	0,000121017	0,000242033	0,000242033	0	8,06777E-05
Geen verbinding	0,00141186	0,003469141	0,004719645	0,000847116	0,000322711	8,06777E-05	4,03388E-05	0,007745058	0,000726099	8,06777E-05	0,002420331	0,00141186	0,007139976	0	0,000564744
Gelijkrichter	0,001210165	0,004437273	0,005647438	0,002178298	0,000201694	8,06777E-05	8,06777E-05	0,000605083	0,00875353	0,000282372	0,001653893	0,002662364	0,003469141	4,03388E-05	0,000484066
Geluidsklacht	4,03388E-05	0,000242033	0,000968132	0,000121017	4,03388E-05	0	0	0	0,000403388	0,001573215	0,000201694	0,000443727	0,000282372	0	4,03388E-05
Onbekend	0,002581686	0,011456232	0,020330779	0,003025413	0,000927793	0,001008471	0,000443727	0,002702703	0,002541347	0,000403388	0,022267043	0,004437273	0,006817265	0	0,000605083
Temperatuur	0,004961678	0,008551835	0,011940299	0,00246067	0,001008471	0,000282372	0,000201694	0,001331182	0,002864058	0,000242033	0,003469141	0,024324324	0,006494554	0	0,000645422
Uitval SC	0,005970149	0,015086729	0,02412263	0,00314643	0,001089149	0,000282372	0,000403388	0,008269463	0,003711174	0,000403388	0,007220653	0,008551835	0,068898749	0	0,001976603
Verkeerd geplaa	0	0	0	0	0	0	0	0	0	0	0	0	4,03388E-05	0	0
WOA	0,000806777	0,001936265	0,001573215	0,00036305	8,06777E-05	4,03388E-05	8,06777E-05	0,000322711	0,000201694	8,06777E-05	0,000524405	0,000564744	0,002258975	0	0,009277935

P(A)*P(B)	Accu	Besturingssyste	Deur	Energie	Ethernet link do	Extern	Facilitair	Geen verbindi	Gelijkrichter	Geluidsklach	Onbekend	Temperatuur	Uitval SC	Verkeerd geplaa	WOA
Accu	0,003218357	0,009364814	0,019008273	0,00250781	0,000724479	0,000322764	0,000227561	0,001922655	0,002143249	0,000285612	0,004226125	0,004212192	0,008350079	2,32205E-06	0,001047243
Besturingssyste	0,008734895	0,025416905	0,051590078	0,006806411	0,001966297	0,00087601	0,000617619	0,005218249	0,005816961	0,000775175	0,011470064	0,01143225	0,022662829	6,30223E-06	0,002842307
Deur	0,018881626	0,05494199	0,111518753	0,014712956	0,004250409	0,001893612	0,001335064	0,011279933	0,012574128	0,001675642	0,024794055	0,024712316	0,048988693	1,36231E-05	0,006144021
Energie	0,002523715	0,007343536	0,014905576	0,001966531	0,000568109	0,0002531	0,000178444	0,001507674	0,001680656	0,000223966	0,003313969	0,003303044	0,00654782	1,82086E-06	0,000821209
Ethernet link do	0,000629237	0,001830962	0,003716404	0,000490315	0,000141646	6,31053E-05	4,44915E-05	0,000375908	0,000419037	5,58414E-05	0,000826271	0,000823547	0,001632566	4,53995E-07	0,000204752
Extern	0,00037213	0,001082827	0,002197873	0,000289971	8,37694E-05	3,73203E-05	2,63122E-05	0,000222311	0,000247818	3,30245E-05	0,000488655	0,000487044	0,000965496	2,68492E-07	0,00012109
Facilitair	0,000151107	0,000439693	0,00089247	0,000117746	3,40155E-05	1,51543E-05	1,06843E-05	9,02718E-05	0,000100629	1,34099E-05	0,000198424	0,000197769	0,00039205	1,09024E-07	4,91698E-05
Geen verbinding	0,001732094	0,005040068	0,010230101	0,001349683	0,000389909	0,000173709	0,000122471	0,001034757	0,001153479	0,000153714	0,002274467	0,002266968	0,004493946	1,24971E-06	0,000563618
Gelijkrichter	0,0017772	0,00517132	0,01049651	0,001384831	0,000400062	0,000178233	0,000125661	0,001061704	0,001183518	0,000157717	0,002333697	0,002326004	0,004610976	1,28225E-06	0,000578295
Geluidsklacht	0,000243576	0,00070876	0,001438608	0,000189799	5,48309E-05	2,44279E-05	1,72225E-05	0,000145513	0,000162208	2,1616E-05	0,000319847	0,000318792	0,000631961	1,7574E-07	7,92588E-05
Onbekend	0,004447512	0,012941424	0,026267915	0,003465593	0,001001171	0,000446035	0,000314471	0,002656955	0,002961799	0,000394693	0,005840167	0,005820913	0,011539143	3,20888E-06	0,001447206
Temperatuur	0,003845338	0,011189213	0,022711356	0,002996368	0,000865617	0,000385644	0,000271893	0,002297215	0,002560784	0,000341253	0,005049434	0,005032788	0,009976794	2,77441E-06	0,001251261
Uitval SC	0,008337957	0,024261889	0,04924568	0,006497109	0,001876943	0,000836202	0,000589552	0,004981117	0,005552622	0,000739949	0,010948832	0,010912737	0,021632967	6,01584E-06	0,002713145
Verkeerd geplaa	2,25533E-06	6,56259E-06	1,33204E-05	1,7574E-06	5,07693E-07	2,26184E-07	1,59468E-07	1,34734E-06	1,50193E-06	2,00148E-07	2,96154E-06	2,95178E-06	5,85149E-06	1,62722E-09	7,33877E-07
WOA	0,001012643	0,002946602	0,005980879	0,000789073	0,000227954	0,00010157	7,1601E-05	0,000604956	0,000674365	8,98666E-05	0,001329734	0,00132535	0,00262732	7,30623E-07	0,000329511

P(A|B)	Accu	Besturingssyste	Deur	Energie	Ethernet link do	Extern	Facilitair	Geen verbindi	Gelijkrichter	Geluidsklach	Onbekend	Temperatuur	Uitval SC	Verkeerd geplaa	WOA
Accu	0,439383322	0,110721794	0,081289418	0,066573231	0,002102313	0,001401542	0,001401542	0,031534688	0,023826209	0,000700771	0,048353189	0,06587246	0,113524877	0	0,013314646
Besturingssyste	0,033049316	0,46862897	0,162406403	0,037696876	0,008778725	0,001807384	0,003098373	0,018848438	0,028143558	0,003872967	0,066098632	0,061967467	0,09656597	0	0,009036922
Deur	0,013139035	0,087195413	0,649187769	0,021380793	0,008719541	0,004897277	0,004300048	0,018752986	0,026755853	0,003344482	0,051361682	0,037505972	0,067845198	0	0,005613951
Energie	0,081322609	0,142091153	0,141197498	0,306523682	0,009830206	0,00714924	0,000893655	0,015192136	0,056300268	0,002680965	0,080428954	0,070598749	0,078641644	0	0,00714924
Ethernet link do	0,007168459	0,125448029	0,222222222	0,032258065	0,35125448	0	0,007168459	0,017921147	0,032258065	0	0,053763441	0,064516129	0,078853047	0	0,007168459
Extern	0,036363636	0,054545455	0,296969697	0,024242424	0	0,218181818	0,006060606	0,054545455	0,036363636	0	0,127272727	0,072727273	0,066666667	0	0,006060606
Facilitair	0,014925373	0,104477612	0,28358209	0,044776119	0,029850746	0,014925373	0,194029851	0,029850746	0,029850746	0	0,044776119	0,089552239	0,089552239	0	0,029850746
Geen verbinding	0,045572917	0,111979167	0,15234375	0,02734375	0,010416667	0,002604167	0,001302083	0,25	0,0234375	0,002604167	0,078125	0,045572917	0,23046875	0	0,018229167
Gelijkrichter	0,038071066	0,139593909	0,177664975	0,068527919	0,006345178	0,002538071	0,002538071	0,019035533	0,275380711	0,008883249	0,052030457	0,083756345	0,109137056	0,001269036	0,015228426
Geluidsklacht	0,009259259	0,055555556	0,222222222	0,027777778	0,009259259	0	0	0	0,092592593	0,361111111	0,046296296	0,101851852	0,064814815	0	0,009259259
Onbekend	0,032454361	0,144016227	0,255578093	0,038032454	0,011663286	0,012677485	0,005578093	0,033975659	0,031947262	0,005070994	0,279918864	0,055780933	0,085699797	0	0,007606491
Temperatuur	0,072140762	0,124340176	0,173607038	0,035777126	0,014662757	0,004105572	0,002932551	0,019354839	0,041642229	0,003519062	0,050439883	0,353665689	0,094428152	0	0,009384164
Uitval SC	0,040032459	0,101163105	0,161752773	0,021098188	0,007303219	0,001893427	0,002704896	0,055450365	0,024885042	0,002704896	0,048417636	0,057343792	0,461996213	0	0,01325399
Verkeerd geplaa	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
WOA	0,04454343	0,106904232	0,086859688	0,020044543	0,004454343	0,002227171	0,004454343	0,017817372	0,011135857	0,004454343	0,028953229	0,031180401	0,124721604	0	0,512249443

Difference P(A∩B) and P(A	Accu	Besturingssyste	Deur	Energie	Ethernet link do	Extern	Facilitair	Geen verbindi	Gelijkrichter	Geluidsklach	Onbekend	Temperatuur	Uitval SC	Verkeerd geplaa	WOA
Accu	-0,0220741	0,002991276	0,014328967	-0,00132438	0,000603462	0,000242087	0,000146883	0,000107406	0,000771728	0,000245273	0,001442744	0,000420341	0,001815186	2,32205E-06	0,0002808
Besturingssysteem	0,003571522	-0,047798101	0,026216944	0,00091694	0,000594776	0,000593638	0,000133553	0,002273513	0,001420027	0,000170092	0,001143319	0,001750927	0,007576101	6,30223E-06	0,00143045
Deur	0,014444353	0,025494633	-0,107722876	0,007492302	0,001305674	0,000239719	-0,000117134	0,004946734	0,003538226	0,000546154	0,007448351	0,012045918	0,026076228	1,36231E-05	0,0042481
Energie	-0,00114712	0,00092966	0,008532039	-0,011869693	0,000124382	-6,9611E-05	0,000138106	0,000821913	-0,000860692	0,000102949	-0,000316527	0,000116275	0,002998001	1,82086E-06	0,0004985
Ethernet link down	0,000548559	0,000419103	0,001215395	0,000127265	-0,0038156	6,31053E-05	-3,61862E-05	0,000174214	5,59878E-05	5,58414E-05	0,000221188	9,74478E-05	0,000745112	4,53995E-07	0,00012407
Extern	0,000130096	0,000719777	0,00022127	0,000128616	8,37694E-05	-0,001414878	-1,40267E-05	-0,000140738	5,78478E-06	3,30245E-05	-0,000358461	2,97782E-06	0,000521769	2,68492E-07	8,0751E-05
Facilitair	0,000110768	0,000157322	0,000126032	-3,27072E-06	-4,66622E-05	-2,51845E-05	-0,000513721	9,5941E-06	1,99514E-05	1,34099E-05	7,7407E-05	-4,42637E-05	0,000150017	1,09024E-07	-3,1508E-05
Geen verbinding	0,000320234	0,001570927	0,005510456	0,000502568	6,71978E-05	9,30316E-05	8,21324E-05	-0,006710301	0,00042738	7,30363E-05	-0,000145864	0,000855109	-0,00264603	1,24971E-06	-1,126E-06
Gelijkrichter	0,000567035	0,000734047	0,004849071	-0,000793466	0,000198368	9,75552E-05	4,49829E-05	0,000456621	-0,007570012	-0,000124655	0,000679805	-0,00033636	0,001141835	-3,90566E-05	9,4229E-05
Geluidsklacht	0,000203237	0,000466726	0,000470476	6,87827E-05	1,4492E-05	2,44279E-05	1,72225E-05	0,000145513	-0,00024118	-0,001551599	0,000118153	-0,000124935	0,000349589	1,7574E-07	3,892E-05
Onbekend	0,001865825	0,001485192	0,005937136	0,00044018	7,3378E-05	-0,000562436	-0,000129257	-4,57477E-05	0,000420452	-8,69588E-06	-0,016426876	0,00138364	0,004721878	3,20888E-06	0,00084212
Temperatuur	-0,00111634	0,002637378	0,010771058	0,000535698	-0,000142854	0,000103272	7,01984E-05	0,000966033	-0,000303274	9,92199E-05	0,001580293	-0,019291537	0,00348224	2,77441E-06	0,00060584
Uitval SC	0,002367807	0,009175161	0,02512305	0,003350679	0,000787794	0,00055383	0,000186164	-0,003288347	0,001841448	0,00033656	0,003728178	0,002360901	-0,047265783	6,01584E-06	0,00073654
Verkeerd geplaatst	2,25533E-06	6,56259E-06	1,33204E-05	1,7574E-06	5,07693E-07	2,26184E-07	1,59468E-07	1,34734E-06	1,50193E-06	2,00148E-07	2,96154E-06	2,95178E-06	-3,44874E-05	1,62722E-09	7,3388E-07
WOA	0,000205866	0,001010338	0,004407664	0,000426023	0,000147277	6,12177E-05	-9,07665E-06	0,000282245	0,000472671	9,18893E-06	0,000805329	0,000760606	0,000368345	7,30623E-07	-0,00894842

Difference P(A|B) and P(B	Accu	Besturingssyste	Deur	Energie	Ethernet link do	Extern	Facilitair	Geen verbindi	Gelijkrichter	Geluidsklach	Onbekend	Temperatuur	Uitval SC	Verkeerd geplaa	WOA
Accu	0,383473681	-0,051964773	-0,248924378	0,023007277	-0,010483408	-0,004205558	-0,002551665	-0,001865877	-0,013406546	-0,004260907	-0,025063512	-0,007302207	-0,031533614	-4,03388E-05	-0,00487817
Besturingssysteem	-0,022860325	0,305942403	-0,167807393	-0,005869078	-0,003806996	-0,003799715	-0,000854834	-0,014552127	-0,009089197	-0,001088711	-0,007318069	-0,0112072	-0,048492522	-4,03388E-05	-0,0091559
Deur	-0,042770606	-0,075491154	0,318973973	-0,022185161	-0,003866179	-0,000709823	0,000346841	-0,014647579	-0,010476902	-0,001617196	-0,022055018	-0,035668695	-0,077213293	-4,03388E-05	-0,01257887
Energie	0,025412968	-0,020595414	-0,189016298	0,262957728	-0,002755515	0,001542141	-0,003059552	-0,018208429	0,019067513	-0,002280713	0,007012254	-0,002575918	-0,066416847	-4,03388E-05	-0,01104358
Ethernet link down	-0,048741182	-0,037238538	-0,107991574	-0,011307889	0,33866876	-0,0056071	0,003215252	-0,015479418	-0,004974691	-0,004961678	-0,019653259	-0,008658538	-0,066205445	-4,03388E-05	-0,01102436
Extern	-0,019546005	-0,108141113	-0,033244099	-0,01932353	-0,01258572	0,212574719	0,002107399	0,02114489	-0,000869119	-0,004961678	0,053856027	-0,000447394	-0,078391825	-4,03388E-05	-0,01213221
Facilitair	-0,040984268	-0,058208955	-0,046631706	0,001210165	0,017265026	0,009318273	0,190076644	-0,003549818	-0,007382009	-0,004961678	-0,028640581	0,016377572	-0,055506253	-4,03388E-05	0,01165793
Geen verbinding	-0,010336724	-0,0507074	-0,177870046	-0,016222204	-0,002169053	-0,003002933	-0,002651124	0,216599435	-0,013795255	-0,002357511	0,0047083	-0,027601751	0,085410259	-4,03388E-05	3,6347E-05
Gelijkrichter	-0,017838575	-0,023092659	-0,152548821	0,024961965	-0,006240542	-0,003069029	-0,001415136	-0,014365032	0,238147956	0,003921571	-0,021386243	0,010581678	-0,035921435	0,001228697	-0,00296439
Geluidsklacht	-0,046650382	-0,107131012	-0,107991574	-0,015788176	-0,003326461	-0,0056071	-0,003953207	-0,033400565	0,055359837	0,356149433	-0,027120404	0,028677185	-0,080243677	-4,03388E-05	-0,00893356
Onbekend	-0,02345528	-0,01867034	-0,074635703	-0,0055335	-0,000922434	0,007070385	0,001624886	0,000575094	-0,005285493	0,000109316	0,206502164	-0,017393734	-0,059358694	-4,03388E-05	-0,01058633
Temperatuur	0,016231121	-0,038346391	-0,156606758	-0,007788828	0,002077037	-0,001501528	-0,001020656	-0,014045726	0,004409474	-0,001442617	-0,022976818	0,280491022	-0,050630339	-4,03388E-05	-0,00880866
Uitval SC	-0,015877182	-0,061523462	-0,168461023	-0,022467766	-0,005282501	-0,003713673	-0,001248311	0,0220498	-0,012347713	-0,002256782	-0,024999064	-0,015830875	0,316937722	-4,03388E-05	-0,00493883
Verkeerd geplaatst	-0,055909641	-0,162686567	-0,330213796	-0,043565954	-0,01258572	-0,0056071	-0,003953207	-0,033400565	-0,037232755	-0,004961678	-0,0734167	-0,073174667	0,854941509	-4,03388E-05	-0,01819282
WOA	-0,011366211	-0,055782336	-0,243354108	-0,023521411	-0,008131377	-0,003379928	0,000501136	-0,015583193	-0,026096898	-0,000507335	-0,044463471	-0,041994266	-0,020336888	-4,03388E-05	0,49405662

Appendix C: Distribution of interarrival times of the failures per cause group