MASTER

Optimizing maintenance at Tata Steel IJmuiden

Lubbers, W.M.

Award date:
2016

Link to publication
Optimizing maintenance at Tata Steel IJmuiden

by

W.M. Lubbers

BSc. Industrial Engineering
Student identity number 0745492

in partial fulfilment of the requirements for the degree of

Master of Science
in Operations Management and Logistics

Supervisors:

P. Lievers, Tata Steel
ir. Dr. S.D.P. Flapper, Eindhoven University of Technology, OPAC
Dr. G.M. Walter, Eindhoven University of Technology, OPAC
TUE. School of Industrial Engineering

Series Master Theses Operations Management and Logistics

Subject Headings: Maintenance optimization, preventive maintenance, failure prediction
Abstract

In this master’s thesis, conducted at Tata Steel IJmuiden, the optimisation of maintenance has been researched. Production of different products can cause machine components to wear at varying rates. A methodology has been developed to investigate the relation between production and the time to failure of components. This methodology includes the selection of a suitable component for further research, selection of possible influencing factors on failure behaviour and determining the influence of these factors on failure behaviour of the component. Furthermore, a methodology has been developed to determine the optimal preventive maintenance interval of a system with two competing failure modes. Both methodologies have been applied to a production line at Tata Steel IJmuiden.
Preface and acknowledgements

This report is the result of my master’s thesis which has been performed at Tata Steel IJmuiden. It is my final step towards the Master Operation Management and Logistics. Working on this project has been a challenging, but most of all interesting process which has provided me with a lot of new experiences. Therefore, I would like to thank the people that have supported me during this time.

First of all, I would like to thank my supervisor at Tata Steel, Peter Lievers, for giving me the opportunity to perform this thesis at Tata Steel. His enthusiasm for this subject and guidance have been a great help and motivation. I also appreciate the warm welcome I received from my other colleagues at Tata Steel, who really made me feel like a part of the team during my time at Tata Steel.

Furthermore, I would like to thank my supervisors from the university, Simme Douwe Flapper and Gero Walter, for all the hours they spent on this project. Their detailed comments have helped me to see things from a different perspective more than once and have greatly contributed to the quality of the final report.

Finally, I would like to express my gratitude towards my family and friends for their support during the project as well as the rest of my studies. Their encouragement and advice has been invaluable to me. It has helped to make my time at the TU/e unforgettable and I am looking forward to what comes next.

Wouter Lubbers

Eindhoven, July 2016
Executive summary

The research has been performed at Tata Steel Packaging (TSP), a department of Tata Steel IJmuiden. TSP produces packaging steel which can be used for various applications such as food cans and (spray-)paint cans. This steel is produced continuously on large production lines. The first production line of TSP is the pickling line, which removes iron oxides from the steel. The research has focused on this production line.

Research goals

The pickling line consists of different parts with a specific function, such as welding, pickling or oiling the strip. These are called sections. The research aims to find the factors that influence how fast failures occur on these sections and use this to determine the time of failure. These factors will be called predictors. The research has focused on predictors related to the production: how the machine is used (such as processing speed), for what it is used (such as material types) and possibly by whom (e.g. the effect of different crews).

The main goal of the research is to develop a method to predict failure behaviour of machine sections to be able to increase their maintenance performance. Maintenance performance is measured by operating costs (the sum of downtime and maintenance costs), availability of the machine and safety. To reach the main goal, four research questions have been answered. The general approach and results of each research question will be presented below.

Research question 1: Which sections are the most interesting to analyse?

The first criterion used to determine the answer of the first research question are the operating costs of each section. To determine these, a tool has been created in Excel using Visual Basic. With this tool, the yearly downtime and maintenance costs of each section are calculated. The output of this tool is a top 10 of sections that lead to the highest yearly operating costs. For the five sections with the highest operating costs (shown in the figure below), four other criteria have been considered.

Sections that have been subjected to big changes in the last year will be excluded from this research. On top of that, the number of different failure modes of a section (complexity), the completeness of the data of the section and the existence of similar sections elsewhere at Tata Steel have been taken into account.

The 5 sections of the pickling line with the highest yearly operating costs
Each of the criteria has a weight assigned to it with which a score for each of the sections can be determined. For the pickling line of TSP, the side trimming section clearly has the highest score. Therefore, this section has been picked to perform further research on.

The tool for the costs and the general method can also be used for other production lines of TSP. The results are not only useful for the purpose of this research. Identifying the sections with the highest operating costs for each production line can be useful to decide on which sections to focus effort to improve.

Research question 2: Which predictors are interesting and feasible to test?
To determine which predictors can influence failure behaviour, the failure modes of the side trimming section have been analysed first. This has led to three failure modes that can be analysed: failure of the side trimmer blades, failure of the scrap cutter and material getting stuck in the scrap cutter. For each of these failure modes, possible predictors have been determined by interviewing the maintenance engineer and production engineer related to the side trimming section. This information has been supplemented with literature where possible. Predictors that are expected to influence the failure behaviour and can be found in the available data have been included in the research.

During the collection of the data for the predictors as well as the three failure modes, the data of the side trimmer blades has been found to be of insufficient quality. Not all replacements of these blades are recorded and the data often does not mention which out of the four blades has been replaced. Therefore, the remainder of the research has been focused around the scrap cutters.

Research question 3: What relationships exist between the chosen predictors and failure behaviour for the selected section?
For each of the predictors of research question 2, their relation with the failures of the scrap cutter has been tested. This has led to the conclusion that it is not possible to predict upcoming failures of the scrap cutter using production data. In the time until failure of the scrap cutter, it usually processes a similar range of products.

On top of that, the relation between the age of a scrap cutter and the occurrence of scrap getting stuck in the scrap cutter has been investigated. No clear relation between these two failure modes could be found either.

Research question 4: Is it possible to optimise the maintenance schedule of a machine section with the found relations between predictors and failure behaviour?
Since the third research question led to the conclusion that there is no relation between production and failure behaviour, this research question has focused on optimisation of the maintenance strategy based on the failure data. For the failures of the scrap cutter, two separate failure modes have been identified: failures of its blades and other failures. Other failures mostly occur shortly after a new scrap cutter is put to use. These problems are therefore likely to be caused by poor maintenance. The risk of failure due to dull blades increases as the scrap cutter is used. This is shown by the figure below.
The failure distributions of both failure modes have been modelled using a Gamma distribution. Using these distributions, the expected yearly costs of corrective and preventive maintenance have been calculated. The current strategy, corrective maintenance, is optimal in the current situation. The benefits of preventive maintenance for the scrap cutter (less downtime costs) do not outweigh its drawbacks (replacing the scrap cutter more often).

The figure above shows the costs of preventive maintenance for two scenarios. The blue line is the current situation, while the orange line shows the costs in case the failure mode other failures can be eliminated. Elimination of the other failures would lead to a cost saving of €17.000 per year.
<p>| Contents |
|-----------------|-----------------|
| 1  | Introduction .................................................................................................................. 1 |
| 1.1  | Description of Tata Steel Ijmuiden .............................................................................. 1 |
| 1.2  | Problem description ........................................................................................................ 1 |
| 1.3  | Research methodology .................................................................................................... 3 |
| 2  | The pickling line ........................................................................................................... 7 |
| 2.1  | Description of the pickling process .............................................................................. 7 |
| 3  | Data for section selection ............................................................................................. 9 |
| 3.1  | Determining operating costs .......................................................................................... 9 |
| 3.2  | Identifying sections ....................................................................................................... 10 |
| 3.3  | Data sources ................................................................................................................. 11 |
| 4  | Section selection ........................................................................................................... 15 |
| 4.1  | Excel tool ....................................................................................................................... 15 |
| 4.2  | Results for the pickling line ......................................................................................... 17 |
| 4.3  | Further selection methods ............................................................................................ 18 |
| 4.4  | Results for the pickling line ......................................................................................... 19 |
| 4.5  | Conclusion ...................................................................................................................... 21 |
| 5  | Choosing failure predictors .......................................................................................... 22 |
| 5.1  | Analysis of failure modes .............................................................................................. 22 |
| 5.2  | Introduction of predictors ............................................................................................. 27 |
| 5.3  | Available data on predictors ......................................................................................... 30 |
| 6  | Creating the data set ..................................................................................................... 32 |
| 6.1  | Enhancing logbook data ................................................................................................. 32 |
| 6.2  | Merging the logbook data .............................................................................................. 34 |
| 6.3  | Verifying data quality .................................................................................................... 35 |
| 7  | Grouping predictor values ............................................................................................. 39 |
| 7.1  | Creating groups .............................................................................................................. 39 |
| 7.2  | Assigning data to groups ............................................................................................... 44 |
| 8  | Testing predictor influence on failure ........................................................................... 45 |
| 8.1  | Scrap cutter failures ....................................................................................................... 45 |
| 8.2  | Material stuck in scrap cutter ......................................................................................... 53 |
| 8.3  | Conclusion ...................................................................................................................... 54 |
| 9  | Choosing a maintenance strategy .................................................................................. 56 |
| 9.1  | Possible maintenance strategies .................................................................................... 56 |</p>
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.2</td>
<td>Censored data</td>
<td>57</td>
</tr>
<tr>
<td>9.3</td>
<td>Analysis of failure data</td>
<td>59</td>
</tr>
<tr>
<td>9.4</td>
<td>The hazard rate</td>
<td>65</td>
</tr>
<tr>
<td>9.5</td>
<td>Conclusions</td>
<td>66</td>
</tr>
<tr>
<td>10</td>
<td>Optimizing operating costs</td>
<td>67</td>
</tr>
<tr>
<td>10.1</td>
<td>Notation</td>
<td>68</td>
</tr>
<tr>
<td>10.2</td>
<td>Scenario 1: Corrective maintenance costs with two failure modes</td>
<td>68</td>
</tr>
<tr>
<td>10.3</td>
<td>Scenario 2: Preventive maintenance costs with two failure modes</td>
<td>69</td>
</tr>
<tr>
<td>10.4</td>
<td>Scenario 3: Preventive maintenance costs with one failure mode</td>
<td>71</td>
</tr>
<tr>
<td>11</td>
<td>Operating costs of the scrap cutter</td>
<td>72</td>
</tr>
<tr>
<td>11.1</td>
<td>Corrective maintenance, two failure modes</td>
<td>73</td>
</tr>
<tr>
<td>11.2</td>
<td>Preventive maintenance, two failure modes</td>
<td>73</td>
</tr>
<tr>
<td>11.3</td>
<td>Preventive maintenance, failure mode <em>dull blades only</em></td>
<td>74</td>
</tr>
<tr>
<td>11.4</td>
<td>Conclusion</td>
<td>77</td>
</tr>
<tr>
<td>12</td>
<td>Conclusion and recommendations</td>
<td>78</td>
</tr>
<tr>
<td>12.1</td>
<td>Conclusions</td>
<td>78</td>
</tr>
<tr>
<td>12.2</td>
<td>Limitations</td>
<td>78</td>
</tr>
<tr>
<td>12.3</td>
<td>Recommendations for TSP</td>
<td>79</td>
</tr>
<tr>
<td>12.4</td>
<td>Academic relevance</td>
<td>81</td>
</tr>
<tr>
<td>13</td>
<td>References</td>
<td>82</td>
</tr>
<tr>
<td>14</td>
<td>Appendices</td>
<td>83</td>
</tr>
</tbody>
</table>
1 Introduction

1.1 Description of Tata Steel IJmuiden

Tata Steel IJmuiden is part of the Tata Steel Group, a worldwide producer of steel with steelmaking sites in India, UK and the Netherlands. The research has been conducted at the Packaging division of Tata Steel IJmuiden, called TSP. This department receives coils of steel strip which are continuously produced in other parts of the plant. TSP processes them into straight or coiled strip to be used in various packaging applications by customers. The production continues day and night throughout three 8-hour shifts. Most processes are visited by all products and there is a fixed sequence which the products need to follow. Figure 1 shows the detailed product flow through the packaging factory with the simplified names below each general production step. Each little block in the product flow corresponds with a production line and the triangles signify the stocks in between them.

![Diagram of the Product Flow and General Production Steps of TSP](internal_report_TSP)

1.2 Problem description

1.2.1 Main goal

Each of the production lines consists of a lot of sections with very different tasks, such as welding or cleaning the metal strip. This causes a large variety in failure modes and occurrence of failures, which makes it hard to come up with predictions for an entire production line. However, the experience of maintenance engineers is that a large part of the maintenance costs and/or downtime is caused by a few sections. Therefore, research should focus on these sections, since improvements on these parts of the production line will have a large impact on overall line uptime and maintenance costs. This leads to the following main goal for the project:

*A method to predict failure behaviour of machine sections to be able to increase their maintenance performance.*

For maintenance performance, three important indicators are considered: safety, operating costs and availability of the installation.

Safety is considered to make sure that a maintenance policy does not lead to dangerous situations. Critical failures in some sections might lead to dangerous situations for employees, the environment or the production facility. Therefore, every maintenance policy must always satisfy minimum safety requirements.

The operating costs are split up in downtime costs and maintenance costs. Downtime costs are caused by unavailability of the machine and maintenance costs are labour costs for repairs, replacement, inspection or upkeep of a machine part and costs for used materials.

Availability of the installation is defined as the percentage of up time: the time it is capable of producing products conforming to quality standards when desired.
Figure 2 provides a graphical representation of what is meant with failure prediction in the scope of this project. The red and green line show two possible paths through which a piece of equipment goes from a new to a failed state. The red line and green line represent its predicted deterioration according to two different production schedules. The intersection between the deterioration path and the failure line is the estimated moment of failure. Based on this estimate, the equipment can be replaced preventively. This is especially useful in cases where there is no way (available) to observe the deterioration of equipment before a failure occurs. The next section explains how this research will try to predict the failure interval.

1.2.2 Scope

The research aims to find the factors that influence how fast failures occur and use this relation to determine the time of failure. These factors will be called predictors. There are a lot of different predictors that can be considered. At first, this project will focus on predictors related to the production: how the machine is used (such as processing speed), for what it is used (such as material types) and possibly by whom (e.g. the effect of different crews). These predictors will also be referred to as operational parameters in this report. Condition monitoring parameters such as vibrations are not considered initially. If this scope turns out to be too narrow during the execution of the research, more predictors could be added.

The research should focus on one production line to ensure that it can be finished within time constraints. Tata Steel has chosen the pickling line as the research subject. This choice will be reviewed briefly in the next chapter along with a detailed description of the pickling line. It should be noted that the goal of the research is to develop a methodology which is still general enough to be applied to other production lines. Since other production lines can have similar parts, it is possible that data from other lines is used as well.
1.2.3 Deliverables

To reach the main goal of the project, two main deliverables will be created in parallel:

**Deliverable 1:** A combination of methods and tools to enable Tata Steel to predict the failure behaviour of machine sections in order to improve the maintenance policy of production lines.

**Deliverable 2:** An improved maintenance policy for one machine section within the pickling line, based on a model that predicts when the machine section will fail.

The first deliverable provides the way to perform this research and enables Tata Steel to replicate it. The second deliverable will provide Tata Steel with a concrete example of what is possible with the predictive model. It also serves as a way to test and validate the methods and tools produced for the first deliverable.

1.3 Research methodology

1.3.1 General framework

A structured approach is needed in order to reach the two proposed deliverables. Therefore, each deliverable will be created along several intermediate steps with their own deliverables. These steps will follow the general methodology for operations research described by Sagasti and Mitroff (1973). In their paper on operations research they proposed a process consisting of four steps: Conceptualisation, Modelling, Model Solving, and Implementation. It starts and ends in the reality state. Each state corresponds with (intermediate) deliverables and the arrows represent the processes to reach them.

**Conceptualisation**

During the conceptualisation step, choices are made on what subjects and variables (not) to include in the research. On top of that, the researcher develops ideas on how to structure the solution to the problem. This is stored in the conceptual model. Based on this conceptual model, the data that is necessary will be gathered and structured.

**Modelling**

In the modelling phase, the relations from the conceptual model will be quantified to create the scientific model.

**Model solving**

In the model solving phase, the scientific model is used to reach a solution to the original problem. In the context of this research, this means that the model for the relation between predictors and failures is used to determine when to perform maintenance actions.

**Implementation**

After a solution for has been reached in the previous phase, the implementation phase will provide the guidelines to implement this solution and return to reality.
This is a very general approach for a research project. The conceptualisation and modelling phase make up a data mining process. Therefore, the CRISP-DM (Cross Industrial Standard Process for Data Mining) methodology will be applied in this part of the research as it provides more concrete support to reach the scientific model. A thorough explanation of the model is provided by Shearer (2000). Figure 4 provides an overview of the steps of CRISP-DM and shows where it fits in the framework by Sagasti and Mitroff (1973). An important property of the chosen approach is the possibility to take a step back in the process if necessary, making it an iterative process.

---

1.3.2 Research questions

The general framework will be translated to this project using 5 research questions. The answer of each research question will provide the input necessary for the next question.

1. Which sections are the most interesting to analyse?

The goal of this question is to find all interesting sections to analyse for one production line and, consecutively, determine the best subject to focus on in this research. First of all, the sections will be ranked using the sum of downtime and maintenance costs of each section. The sections with the highest costs will then be assessed in a more qualitative way to reach the final answer to this question.

2. Which predictors are interesting and feasible to test?

Once a section has been chosen, the research focus should move to the parameters that influence failure. They should be found in literature and using the experience of involved maintenance engineers. This will result in a list of predictors that are expected to influence failure behaviour and possible to test within this project.

3. What relationships exist between the chosen predictors and failure behaviour for the selected section?

For each of these chosen predictors, its effect on failure behaviour of the chosen section should be investigated. This will result in a model that explains the relation between the predictors and failure behaviour of the section.

4. Is it possible to optimise the maintenance schedule of a machine section with the found relations between predictors and failure behaviour?

In order to allow the use of the model that results from the previous research question, a way to translate this model into a maintenance schedule will be proposed. By combining the model with the future values of the predictors in the production schedule, the optimal moment for maintenance can be determined.

1.3.3 Relation between methodology and report

Each chapter can be related to the methodology described in the past few sections. Table 1 shows the relation between the methodology and the report by linking each chapter to the corresponding phases in the general framework and a research question if applicable.

<table>
<thead>
<tr>
<th>Chapter(s)</th>
<th>Project phase</th>
<th>Corresponding CRISP-DM step</th>
<th>Research question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 2</td>
<td>Conceptualisation</td>
<td>Business understanding</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>Conceptualisation</td>
<td>Data understanding</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Conceptualisation</td>
<td>Data understanding</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Conceptualisation</td>
<td>Data understanding</td>
<td>2</td>
</tr>
<tr>
<td>6, 7</td>
<td>Conceptualisation</td>
<td>Data preparation</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>Modelling</td>
<td>Modelling</td>
<td>3</td>
</tr>
<tr>
<td>9, 10</td>
<td>Model solving</td>
<td>Deployment</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>Implementation</td>
<td>Deployment</td>
<td>-</td>
</tr>
</tbody>
</table>
Part one: Prediction of failures

The first part of this report will investigate the first part of the main goal: *a method to predict failure behaviour of machine sections*. In order to reach this goal, the first three research questions will need to be answered.

First of all, the pickling line will be described in more detail to provide the basics needed for the rest of the report. The next two chapters will focus on the first research question: selecting a section for the research to focus on. First, the necessary data will be collected and subsequently, a method will be presented to determine on which section to focus.

After one particular section has been chosen, its possible predictors of failure will be investigated. Again, the approach is to first identify and gather the necessary data and only then make a choice on which predictors (not) to include in the research. When the definitive set of predictors has been specified, the data set can be created to test the relationship between predictors and failures. This involves gathering, merging and restructuring the data to enable further analysis.

The final step of this part of the research is to test the influence of the predictors on failures. For this purpose, several hypotheses will be developed and tested. The outcome of these tests will lead to a conclusion whether or not it is possible to use the predictors to optimize maintenance performance.
2 The pickling line

As stated in the introduction, the research will focus on the pickling line. This choice has been made together with Tata Steel to reduce the total time that is spent on determining the scope of the project. There are two important reasons for choosing the pickling line:
- First of all, the pickling line is important for the continuity of the entire TSP plant. All products that enter the department have to be pickled first. There is only one pickling line at TSP, so if it breaks down the next steps will no longer receive new coils and the stock of coils before the pickling section will grow rapidly. Of course, the consecutive processes are decoupled by stocks, so a production stop will not lead to problems right away.
- The second reason to focus on the pickling line is that it leads to many corrective and preventive actions each year. The operating costs of the pickling line are among the highest of the production lines at TSP together with the annealing and tinning lines.

While it is hard to judge if the pickling line is the best possible choice without any in-depth research, these two reasons make the pickling line a solid choice as the starting point for this project. Other possible production lines for future application of this project would be the continuous annealing lines and electrolytic tinning lines.

2.1 Description of the pickling process

In the next research steps, the pickling line will need to be discussed in more detail. Therefore, a brief explanation will be provided on why the steel needs to be pickled as well as the process itself.

The material that enters the pickling line has been hot rolled in another plant to a strip with a gauge (thickness) of approximately 3 millimetres. The processing temperature of the hot rolling process varies between 1250 and 850 degrees Celsius. At these temperatures, steel oxidises at a high rate, which causes a layer of iron oxides (called scale) to form on top of the strip. Before the gauge of the strip can be decreased further in the cold rolling process, the scale needs to be removed by the pickling process. Otherwise, the iron oxides will be pressed into the strip, which will result in serious structural defects in the material.

To feed rolls into the pickling line, an automatic crane (which will not be considered in this project) picks up coils in a hall and puts them on a conveyor belt towards the start of the process.
1. When the coil is up for processing it enters the line and is uncoiled with the help of a process operator.
2. Another operator controls the welding machine next to it, where the start of the new coil is welded to the end of the previous one.
3. The strip then enters the entry looping section. When a new roll enters the system, the metal strip needs to be stopped temporarily, while the part of the strip that is being pickled needs to keep moving. To make this possible, the length of the loop can be increased and decreased to act as a
buffer. When a new coil is being uncoiled and welded to the end of the previous one, the loop is decreased in size, allowing the part after the loop to continue. When the welding is completed, the uncoiling section runs at a higher speed than the part after the loop, so the loop size is increased again.

4. After the loop, the strip is led through a set of rolls in order to increase the tension on the strip. They are necessary because the strip receives very little support over a long distance in the next step (the strip is not guided through rolls or resting on a surface for a long time). Without the added tension, the strip would sag too much.

5. Next, the strip enters the actual pickling process. Tata Steel uses hydrochloric acid (HCl) to remove the scale. The strip is guided through five tanks filled with acid and an inhibitor. The inhibitor is a fluid that forms a protective layer on iron. As such, it will prevent the acid from corroding the “good” iron (this is called overpickling). However, it is still important that the strip does not remain in the tanks for too long. After the final tank the strip is cleaned by consecutively squeezing, rinsing and drying the strip. This removes any leftover acid on the strip.

6 + 7. The exit section starts off with another set of tension rolls, followed by a looping section.

8. To protect the strip from corroding, a film of oil is sprayed on the surface. The sides of the strip often contain small cracks, so they are removed by a side trimmer. These side strips are guided away from the line and scrapped and can be reused in the steel making process.

9. Meanwhile, the finished strip is cut off at its original length, coiled and moved to storage by another crane.
3 Data for section selection

The description of the pickling line already showed some of the sections, each with very different tasks and failure modes. More than 100 sections have been specified in SAP for the pickling line. Due to this high variety of functions within the pickling line, it is unlikely that an analysis on production line level will yield reliable predictions of failure behaviour. With so many sections it would also be too time-consuming to analyse all sections separately within this research. The scope needs to be narrowed down further to one section.

The first selection will be made based on operating costs. This chapter will provide a definition of operating costs and subsequently provide an explanation of the data required to calculate these operating costs.

3.1 Determining operating costs

The operating costs of a section $s$ within production line $p$ during a chosen time interval, denoted by $c^O_{s,p}$, consist of two parts: downtime costs and maintenance costs.

**Downtime costs**

Downtime is defined as all time during which a machine is not producing. At Tata Steel Ijmuiden, downtime is split up into three categories: production-related, process-/product-related and technical downtime. Production-related downtime consists of:
- A lack of production resources, for example a lack of crew or coils to process;
- Overcapacity, allowing the production line to stop for a while;
- Human errors, such as a mistake in the welding process, causing the need for a stop to redo the weld.

Process-/Product-related downtime is caused by problems with the production process and/or product, such as the product getting stuck or problems due to poor quality of the received coil. Technical downtime is directly related to the failure or repair of a component within the production line.

Since the goal of the project is to predict failure behaviour of machine sections, only downtime that can be attributed to failure or repairs of specific machine sections is considered. Production-related downtime is not relevant to this research, because none of this downtime is caused by one section in particular or by failures. Process-/Product-related downtime is only relevant to the research when the underlying cause of the downtime can be traced to the failure of a section. All technical downtime is included.
To find the downtime per section over a period of time, the length of the failure and the section where the failure happened are required. The sum of all downtime on a production line \( p \) caused by section \( s \) in the chosen time interval is denoted by \( t_{s,p}^{\text{downtime}} \) in hours.

TSP has estimated how much one hour of downtime costs for each production line. This downtime cost per hour is denoted by \( DC_p \). The total downtime costs over a certain period of time of production line \( p \) caused by section \( s \), \( c_{s,p}^{\text{downtime}} \), can be calculated by multiplying the total downtime caused by this section during this period with the costs of downtime per hour:

\[
c_{s,p}^{\text{downtime}} = t_{s,p}^{\text{downtime}} \cdot DC_p
\]

**Maintenance costs**

Maintenance costs consist of labour costs for repairs, replacement, inspection or upkeep of a machine part and costs for used materials. For each maintenance action, these costs are recorded. To determine the total maintenance costs of a section over a period of time, the costs of all these maintenance actions need to be summed per section. The maintenance costs of a section during a period are defined as \( c_{s,p}^{\text{maintenance}} \).

**Operating costs**

The operating costs for each section are obtained by adding up the maintenance and downtime costs of the same period:

\[
c_{s,p}^{\text{O}} = c_{s,p}^{\text{downtime}} + c_{s,p}^{\text{maintenance}}
\]

For the selection of a section, the preferred period for \( c_{s,p}^{\text{O}} \) is one year. Parts and sections are upgraded over time, for example to increase production or enable production of steel with new specifications. These upgrades happen on a regular basis, especially in sections with a lot of problems, so choosing a longer time interval would increase the risk of including outdated data for some sections. On the other hand, the time period should not be smaller than one year, since every section has one large maintenance stop each year. Picking a period of one year ensures that the planned maintenance of all sections went through one full cycle. It would take too much time to check all sections of a production line individually for changes and adjust the time interval for each section, so this is not an option either.

### 3.2 Identifying sections

To rank the sections based on operating costs, the costs need to be attributed to a particular section. Tata Steel uses functional location codes in SAP to specify locations on different detail levels. The used notation is xxx-xx-xx-..., where each x is a number and each dash signals a new detail level. For example, 248-02-02 is the entry part of the pickling line. The entry consists of the uncoiling, welding and looping section as well as 30 other sections. This can also be seen in Figure 7, which shows the first four detail levels and provides some examples for each level. There are more detail levels below the section level for the components of each section. The amount of detail levels below section level depends on the complexity of the section. As stated in the research questions, this research will focus on the section level.
3.3 Data sources

The maintenance actions and failures at each production line are tracked and saved in SAP. The data is spread over three different sources: (maintenance) orders, notifications and the logbook.

3.3.1 Orders

Orders track all maintenance actions. They can be generated in several ways. Some maintenance actions are generated periodically according to fixed preventive maintenance schedules. They can also be created manually, for example based on a notification. Before a maintenance action is performed, an order is created to specify who should perform which actions on what part of the production line and to estimate the costs. After the maintenance actions have been performed, the actual costs (labour and material costs) are added to the order (Table 2); these are defined as maintenance costs in this project. It is important to note that downtime costs are not included in the orders.

<table>
<thead>
<tr>
<th>Order ID</th>
<th>Functional location ID</th>
<th>Actual costs (€)</th>
<th>Order type</th>
<th>Explanation order</th>
<th>Starting date</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000000</td>
<td>123-12-12-12</td>
<td>1.234,56</td>
<td>PM10</td>
<td>Repair hydraulics oil leakage</td>
<td>01-01-2011</td>
</tr>
</tbody>
</table>

Table 2 shows a part of a maintenance order. A complete order contains over 100 columns, but for the sake of clarity only the most important columns have been included for now.

The different order types in SAP (Table 3) are divided by TSP in maintenance (PM) and not maintenance related (ZZ). The first two PM types are about repair, replacement or upkeep of machine parts and are the only types included in this research. For PM30 orders it can be argued that they do not really belong to maintenance, since their goal is to upgrade a part of the production line. On top of that, they involve large, non-recurring costs, so including this order type could result in misleading output. The PM91 code for orders is used when the technical service department of Tata
Steel IJmuiden has assisted in a job. The costs of these orders are also included in the PM10, PM20 or PM30 order they are associated with.

### Table 3: SAP order types

<table>
<thead>
<tr>
<th>SAP order type</th>
<th>Description</th>
<th>Within scope?</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM10</td>
<td>Corrective maintenance action</td>
<td>Yes</td>
</tr>
<tr>
<td>PM20</td>
<td>Preventive maintenance action</td>
<td>Yes</td>
</tr>
<tr>
<td>PM30</td>
<td>One-time improvement</td>
<td>No</td>
</tr>
<tr>
<td>PM91</td>
<td>Order sub-type</td>
<td>No</td>
</tr>
<tr>
<td>ZZA1/ZZB1/ZZC1</td>
<td>Not maintenance related</td>
<td>No</td>
</tr>
</tbody>
</table>

3.3.2 Notifications and the logbook

*Notifications* indicate that a maintenance action is required. This can be due to machine failure or a problem (for example an upcoming defect) found during inspection. It indicates to which part of the production line it applies (using the functional location ID) and if it is a preventive or corrective notification. In case of a failure, it will also specify if the problem affected production and if so, how long. The notifications only capture technical downtime.

### Table 4: Example of a notification

<table>
<thead>
<tr>
<th>Notification ID</th>
<th>Start of downtime</th>
<th>End of downtime</th>
<th>Functional location ID</th>
<th>Explanation downtime</th>
<th>Effect</th>
<th>Duration (hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>000000000</td>
<td>01-01-2011 00:00</td>
<td>01-01-2011 01:15</td>
<td>123-10-10-10</td>
<td>Hydraulics failure</td>
<td>0</td>
<td>1,25</td>
</tr>
</tbody>
</table>

*The logbook* is directly linked to the production line and is semi-automated. It consists of two parts: stops and production. The production part provides detailed production data during every second of the day, while the stops track whether or not the production line is producing products. When the production line is stopped or a new coil enters the production line, a new entry is created automatically. The operators need to complete these entries. For a stop they need to specify the cause of the stop and if applicable, in which part of the production line the stop was caused.

For now, only the production stop part of the logbook will be considered. Because every production stop is registered, all three types of downtime are included in the logbook. The entries on technical downtime are the only ones for which the functional location name is filled in. Therefore, this is the only type of downtime that will be considered for the section selection.

### Table 5: Example of a logbook entry

<table>
<thead>
<tr>
<th>ID number</th>
<th>Start of downtime</th>
<th>End of downtime</th>
<th>Functional location name</th>
<th>Explanation downtime</th>
<th>Downtime (hr)</th>
<th>Downtime type</th>
</tr>
</thead>
<tbody>
<tr>
<td>000000</td>
<td>01-01-2011 00:00</td>
<td>01-01-2011 01:15</td>
<td>BB12-WELDING</td>
<td>Hydraulics failure</td>
<td>1,25</td>
<td>Technical</td>
</tr>
</tbody>
</table>

As can be seen in Table 4 and Table 5, the logbook and the notifications partially capture the same information: the cause (explanation downtime column), length (downtime) and location (functional location) of all production line stops. Just like the orders, the notification and logbook entries contain many more columns than shown (around 100 and 70 respectively), but these are not used in this part of the research.
However, there are some important differences. While the notifications and logbook both contain some of the same data columns, they have a different purpose and provide a lot of unique information as well.

The notifications are primarily used for the daily control of the production lines. When there is a problem in the line, the notifications are used to draw attention to it, keep track of progress and save details about the problem. When a similar problem occurs in the future, these details can be used to identify and solve it faster. The control function also includes requests for preventive maintenance; when an operator sees that a component is worn and needs to be replaced in the next planned stop he can indicate this in a notification. Therefore, not every notification comes with downtime. The notifications contain a column “effect” which shows if the duration column of the notification is related to downtime.

The logbook tracks every minute of production and is especially useful for reporting functions. Since every stop is automatically added to it, it is less prone to human errors and expected to be more complete than the notifications. There is less focus on details of the stop itself, but more information on the effects of the stop on production.

A flaw of the logbook is that is does not use the functional location codes, but just their names. Since the codes are needed to group the data on section level, these names will need to be converted to the code format explained at the start of this chapter.

<table>
<thead>
<tr>
<th>Notifications</th>
<th>Logbook</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not possible to link with production directly</td>
<td>Same layout as production data (allows linking)</td>
</tr>
<tr>
<td>Uses functional structure codes</td>
<td>Uses functional structure names</td>
</tr>
<tr>
<td>Detailed information about causes and solution for each stop</td>
<td>Information on material characteristics and effects of stop on production</td>
</tr>
<tr>
<td>Entries are created manually</td>
<td>Entries are created automatically, completed manually</td>
</tr>
</tbody>
</table>

Table 6 contains a summary of the differences between the notifications and the logbook. Based on this comparison, the logbook data seems to be more suitable for this project. Its method of recording downtime seems to be more reliable and for future steps, it is a big advantage that it has the same layout as the production data. Before making a final choice on which data source to use, the number of entries and recorded downtime per section over the last year will be compared.
Each dot in Figure 8 represents the total downtime of a section according to both sources between 1 November 2014 and 31 October 2015. For example, the upper right dot shows that for one section, the logbook recorded around 230 hours of downtime, while the notifications recorded only 160 hours approximately. There are 85 sections for which downtime was recorded by either source. The straight line through the graph goes from (0,0) to (240,240). The few sections with over 240 hours of downtime have been left out to make the graph easier to read (increasing the scale further would compress the other values). For all sections above the line, more downtime was recorded by the logbook, while all sections below it have more downtime according to the notifications. The graph shows that overall, the logbook records higher values for downtime, but there are some exceptions.

Further observation of the data shows that the most important reason for the difference is likely to be consistency; there are 623 notifications on the pickling line in the chosen time interval, while the logbook contains 6804 entries over the same period. Even though 4594 of these 6804 entries are shorter than one minute, it does show a large difference between the two sources. Inquiry with staff working with this data revealed that short failures that are occurring multiple times a day are not saved in a notification each time, since this would require a lot of extra effort (both for creating and assessing each entry). This adds up to a large difference over a year.

The few cases where the hours of downtime of a section are higher according to notifications than the logbook can often be contributed to one or two outliers in the notifications for that section. For this reason, the logbook is the preferred source to determine the downtime costs of a section. The notifications will only be used for verification; if the difference between the two types is unusually large or small, this should be checked. After all, the logbook is not perfect either.
4 Section selection

In this chapter, a general methodology will be presented to choose the section within a production line that is most interesting for further research. This is first of all based on the operating costs of each section, using the data presented in the previous chapter. After this initial analysis, the top candidates will be reviewed in more detail to decide on which section the effort should be focused. Hereafter, the proposed methodology will be applied to the pickling line of TSP.

4.1 Excel tool

To use the operating costs per section, a tool has been created in Excel. This makes it faster and easier to perform the analysis and it reduces the probability of making a mistake by reducing the number of steps that need to be performed manually. Figure 9 shows the steps that are needed to reach the operating costs per section and which steps are covered by the tool.

![Flowchart](image)

After determining the desired time period, the data from the logbook, notifications and orders should be imported from SAP. Standard templates for data selection have been created within SAP so that the user only needs to change the period of analysis.

- First of all, this makes sure that the imported data has the right format for the tool; the SAP export will always contain the same data columns in the same order. These columns are comparable to those shown in the example entries of each data source in the previous chapter (Table 2, Table 4 and Table 5).
- The second benefit is that some irrelevant entries are already excluded. For the orders, this means that only the PM10 and PM20 order types are retrieved (see Table 3). From the logbook, only technical interruptions will be retrieved. This leaves out a lot of extraneous
entries that just state the time during which the production line is in operation. Finally, the notifications that contain zero downtime will already be filtered out.

The remaining steps of Figure 9 can be performed using the tool. Figure 10 is a screenshot of the first sheet of the tool. Below is a step-by-step description of all the steps to go through when using the tool. At the start of each step is stated to which part of Figure 10 it applies. This can be either of the four boxes or one of the other sheets. The other sheets are three data sheets (the retrieved data should be copied to these) and two sheets with results generated by the tool.

**Step 1:** BOX 1: Erase old data. There are separate sheets in the Excel file to paste entries of the logbook, notifications and orders. By pressing the button “Erase data” in the starting screen, these sheets are emptied so new data can be copied to these sheets.

**Step 2:** Data sheets. All the collected data from the orders, logbook and notifications can now be copied to the corresponding sheet of this tool.

**Step 3:** BOX 2: Settings. The downtime costs per hour ($D_{CP}$) are filled with the current estimates by TSP by default. However, these values might change in the future. In this case, the user can change this for each production line. On top of that, the production line which will be analysed needs to be chosen.

**Step 4:** BOX 3: Import new data. Now the tool can be used to process the data. As explained at the start of chapter 3, the sections are specified at the fourth level of the functional structure ID. Some maintenance is aimed at areas bigger than one section or just not specified precisely enough. These entries will be removed, since they cannot be attributed to a section. The ones that are specified more precisely than section level will be generalised to section level so they can be analysed together. This step is executed by pressing the process data button for each data type. At the same time, downtime in the notifications and logbook will be converted to downtime costs.
The tool performs one extra step in the logbook entries, since the functional locations are only specified by name and not by notation. The tool searches for and prints the corresponding functional location code for each logbook entry to allow the data to be coupled with the orders.

**Step 5: Data sheets.** In the previous step, each sheet has been sorted by costs in descending order as well. This allows the user to go to the sheet and observe if there are any outliers caused by input errors. When the end date of a notification has been set one day too late for the pickling line for example, this can increase the downtime costs by 50,000 euros, so possible mistakes are easy to identify. Since this judgement can only be made by a person, this action needs to be executed manually in the data sheets.

**Step 6: BOX 4: Process data... . After removal or correction of incorrect values, the button at the bottom of the front page can be pressed. This creates a table which shows the operating costs for all sections as well as a graph showing the top-10 worst performing sections. A slightly modified version of this graph can also be found in the next section (Figure 12). The two different calculations of downtime costs (using either notifications or the logbook) are both used, so they can be compared per section. The user can also choose to sort the top-10 graph by the maintenance costs based on either the logbook or the notifications to check if this causes big changes.

4.2 Results for the pickling line
The tool has been used to find the sections of the pickling line with the highest operating costs. For this research, a period of one year has been chosen, from 1 November 2014 until 31 October 2015. It is best not to use data of less than one month ago, since orders and notifications might still be changed or updated in this period. Under that restriction, this was the most recent period available at the time of the research. The downtime costs per hour, $DC_p$, have been set to €2.109 for the pickling line. The results are shown in Figure 11 and Figure 12. Figure 11 shows the operating costs for all sections in the pickling line. It reveals that there is a clear difference between the sections with the highest costs and the other sections. Figure 12 shows the five sections with the highest operating costs in a higher level of detail.

![Operating costs](image)

*Figure 11: Operating costs of all sections of the pickling line (Nov 2014 – Oct 2015), sorted in descending order*
Figure 12: Operating costs per section (from November 2014 to October 2015)

Each bar consists of two parts: downtime costs and maintenance costs. The maintenance costs are based on the data from the maintenance orders. The downtime costs based on the logbook and notifications are both presented for each section. As decided in Chapter 3, the logbook is considered to be the primary source for downtime costs while the notifications are used for verification. In the case of the pickling section, the difference between the logbook and the notifications is extraordinarily large. Further inspection of the logbook data reveals that one large production stop (of over 30 hours long) which has only been recorded in the logbook mostly caused this difference. This was caused by a planned stop that took longer than planned due to some issues. Apart from that incident, the downtime costs of the pickling line are only half as high as they are presented in Figure 12.

4.3 Further selection methods

A possible approach would be to simply select the section from Figure 12 with the highest costs. However, the differences between the worst performing sections are not that big that either of them should be dismissed already. For the final selection, four more qualitative criteria will be considered:

- Changes to the section. If a section has been revised or improved within the last year, this can make the data unreliable or outdated. The same reasoning applies to ongoing projects within a section. In this case, it is better to wait for at least a year and then assess if the section still has high operating costs. Therefore, this condition needs to be satisfied to continue research on the section.

- Complexity of the section. If the operating costs are spread over a wide variety of failure modes, it will be difficult to assess all of them and find a good maintenance plan for the entire section. The maintenance team of TSP has documented the most important failure modes according to their experience per section. When available, this will be primarily used to assess the complexity. Otherwise, this information will be obtained from the maintenance engineer responsible for the section. This results in a rating of the complexity of low (2 or less failure modes), moderate (3 to 5 failure modes) or high (more than 5 failure modes). The goal is just to assess the complexity, so no detailed information on each particular failure mode needs to be gathered yet.
- The completeness and quality of failure data. When necessary data is not being recorded or unreliable, this will reduce the probability that prediction of failure behaviour of the section is possible. This will result in a low score on completeness. If more data is recorded for a section on top of the standard data (logbook and SAP), this will result in a high score for completeness.

- The presence of the same or similar sections in other production lines throughout Tata Steel IJmuiden. This could provide opportunities for data pooling, which could increase the quality of the used data. At the same time, it makes it more likely that the results are directly applicable in other departments within Tata Steel.

The weights of these four criteria, together with the operating costs of each section, are presented in Table 7. Every decision gets a score assigned to it, which leads to a final ranking for each section. The “changes” criterion does not have a rating linked to it; this criterion must be satisfied to continue. For complexity, completeness and pooling possibility, a high or low score will grant 2 or -2 points respectively. The section with the highest costs will get 5 points, each subsequent section will get one point less. By using this weighting, the criterion costs has the most influence on the final decision, but the other criteria are weighted strongly enough that the section with the lowest costs out of the five options can still be chosen.

<table>
<thead>
<tr>
<th>Changes</th>
<th>Complexity</th>
<th>Completeness</th>
<th>Pooling possibility</th>
<th>Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>No/Go</td>
<td>Low/High</td>
<td>2/2</td>
<td>Yes/No</td>
<td>Highest/5</td>
</tr>
<tr>
<td>Yes/No go</td>
<td>Moderate</td>
<td>0/0</td>
<td>No</td>
<td>.../...</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>-2/-2</td>
<td></td>
<td>Lowest/1</td>
</tr>
</tbody>
</table>

On top of these general criteria, certain special conditions might make a section particularly (un)suitable for this research. Therefore, any special cases not covered by these four criteria should be considered as well to obtain an integral review of all sections.

4.4 Results for the pickling line

The five sections from Figure 12 will be evaluated: the side trimmer section, exit looping section, welding section, pickling section and cleaning section. After the evaluation of each section, the results are presented in Table 10 and a final choice for a section is made.

**Side trimmer section**

As explained in the previous chapter, none of the product-related downtime is included in the analysis for section selection. A lot of product jams are caused by the side trimmer section though due to the side strips getting stuck when they are guided away. This triggered TSP to create specific codes in SAP for the product-related interruptions caused by the side trimmer section. The analysis of the side trimmer section can therefore be extended with product-related downtime, which increases the score of this section on data completeness. The failure modes of the side trimmer section can be found in Table 8. These are taken from the available list of failure modes of the side trimmer section. Their completeness has been confirmed with the maintenance engineer responsible for this section.
Table 8: Failure modes of the side trimmer section

<table>
<thead>
<tr>
<th>Failure mode</th>
<th>Cause</th>
<th>Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dull side trimming blades</td>
<td>Wear due to normal usage</td>
<td>Reduced quality of edge</td>
</tr>
<tr>
<td>Broken side trimming blades</td>
<td>Product deformities / Flaws in sharpening process</td>
<td>Irregular cutting pattern</td>
</tr>
<tr>
<td>Dull scrap cutter blades</td>
<td>Wear due to normal usage</td>
<td>Pieces of scrap get longer</td>
</tr>
<tr>
<td>Material stuck in scrap cutter</td>
<td>Worn side trimmers or scrap cutters / Poor alignment of strip before the section</td>
<td>Automatic emergency stop</td>
</tr>
<tr>
<td>Product cut at wrong width</td>
<td>Side trimmers need to be realigned</td>
<td>Measured band width deviates from required value</td>
</tr>
</tbody>
</table>

Side trimmers can be found in other parts of TSP as well as other plants. There are two more pickling lines at Tata Steel IJmuiden and three galvanizing lines with a similar side trimming section. Side trimming sections of other production lines at Tata Steel often contain a scrap press instead of scrap cutters, so they only match for the first half of the section.

**Exit looping section**

This section is quite large and contains a wide variety of components and corresponding failure modes. There are multiple types of rolls to guide, turn and aim the metal strip as well as a looper car which moves to alter the loop size. For feasibility of the research, it would be necessary to focus on a smaller part of the section, but then the impact in terms of the cost analysis will decrease. On top of that, there is a project underway to solve some of the issues in the looping section. It is better to await the results of the project to see if the exit loop remains a problem section.

**Welding section**

It is very important for the pickling line that the weld between two separate coils is of a consistently high quality. Especially during the pickling process, the metal strip is kept under high tension by the tension rolls. A break in the weld can damage other sections and cause safety hazards. Therefore, the section is manned and controlled continuously by an operator. The quality of the weld is monitored by the machine and samples are regularly taken out and tested with a separate testing bench.

Despite its relatively small size, this section is very complicated. Similarly to the exit looping section, this results in a large variety of failure modes. In December 2015, during the large maintenance stop of that year, the entire machine has been overhauled to solve some recurring issues (problems with its hydraulics amongst others). Some of the larger breakdowns in the past year probably originated from these issues as well. Therefore, the status of the welding section should be reviewed later to see if the operating costs have indeed dropped.

**Pickling section**

There are two things that make the pickling line stand out in Figure 12. It has incurred the highest maintenance costs of all sections and the difference between the notifications and the logbook is particularly high. The first matter is caused by two big orders during the year revision that account for over half of the maintenance costs of the pickling section; a yearly inspection/repair action that incidentally caused higher costs than other years and a leakage. Similarly, a large portion of the downtime costs can be attributed to one entry in the logbook that makes up 40% of the total downtime costs.
All in all, the presence of the pickling section in the top 5 is too largely based on these three entries with three different causes. It is therefore likely that the section will return to much lower costs next year.

Cleaning section
The cleaning section has considerably lower costs than the four above-mentioned sections. However, a large portion of the costs of the section is caused by wringer rolls. These rolls, located after each pickling and cleaning tank, are supposed to remove acid and other contaminations from the metal before it moves towards the next step. The rolls between the pickling tanks ensure that the condition of each tank can be managed separately, while the final roll is used to remove any leftover acid. The two failure modes of the wringer rolls are recorded in Table 9.

<table>
<thead>
<tr>
<th>Failure mode</th>
<th>Cause</th>
<th>Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rolls let through acid and other contaminants</td>
<td>Worn roll surface</td>
<td>Contamination of succeeding sections, marks on product surface</td>
</tr>
<tr>
<td>Rolls stop turning</td>
<td>Bearing failure</td>
<td>Contamination of succeeding sections, marks on product surface</td>
</tr>
</tbody>
</table>

The condition of the rolls is inspected periodically, so there is no continuous condition monitoring data available. When a roll at the cleaning process starts to fail, this can be observed by an operator through a camera which shows the metal surface. If overlooked, it will lead to a degraded surface, which will cause problems in later production steps.

Wringer rolls are present in some other lines such as the cleaning line as well. The presence of acid makes the composition of the rolls of the pickling line specific to this application though. Therefore, when it comes to data pooling, the two other pickling lines of Tata Steel IJmuiden are the only candidates.

4.5 Conclusion
Table 10 provides a summary of the previous section. Based on the outcomes of the cost analysis and the four other criteria, the side trimmer section receives the highest score using the weighting determined in section 4.3. The operating costs associated to this section are the highest and it scores well on the four other criteria. Therefore, the research will continue with this section. The cleaning section seems to be promising as well for further study, but will not be covered in this research.

<table>
<thead>
<tr>
<th>Section name</th>
<th>Recent changes/ Changes in progress</th>
<th>Complexity</th>
<th>Data completeness</th>
<th>Pooling possibility</th>
<th>Costs</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Side trimmer section</td>
<td>No (Go)</td>
<td>Moderate (0)</td>
<td>High (2)</td>
<td>Yes (2)</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Exit looping section</td>
<td>Yes (No-go)</td>
<td>High (-2)</td>
<td>Moderate (0)</td>
<td>-</td>
<td>4</td>
<td>No-go</td>
</tr>
<tr>
<td>Welding section</td>
<td>Yes (No-go)</td>
<td>High (-2)</td>
<td>Moderate (0)</td>
<td>-</td>
<td>3</td>
<td>No-go</td>
</tr>
<tr>
<td>Pickling section</td>
<td>No (Go)</td>
<td>Moderate (0)</td>
<td>Moderate (0)</td>
<td>No (0)</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Cleaning section</td>
<td>No (Go)</td>
<td>Low (2)</td>
<td>Moderate (0)</td>
<td>Yes (2)</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>
5 Choosing failure predictors

To predict the failure behaviour of a section, it is important to understand how and why the section fails. In the previous chapter, a first examination on the main failure modes of the side trimmer has been performed. This uncovered five different failure modes. To predict when each of these failure modes will occur, a set of parameters that are expected to influence that failure mode should be found for each one. These parameters will be referred to as predictors.

First, the failure modes will be studied more in depth. The information gained in this process will then be used to come up with possible predictors. The experience of people who work with the side trimming section, supplemented with literature where possible, will be consulted to determine what parameters might influence each failure mode. This leads to the “Necessary data” part of Figure 13. The available data sources will be checked to see if it is possible to find all the necessary data. The gaps between the necessary data and available data will be assessed to see if more information should be collected. It is also possible that the available data contains possible predictors that had not been considered yet in the previous step. If this is the case, these predictors will be checked with literature and the experts to see if they should be added as well (represented by the arrow from available to necessary data). This will lead to the list of predictors to be used in the next step of the research.

5.1 Analysis of failure modes

In the previous chapter, five failure modes of the side trimming section have been specified (Table 8). First of all, these failure modes will be investigated in more detail to see if they can be included in this research. A detailed explanation of each failure mode will be provided along with this. Furthermore, a final check will be made using the data to ensure that no failure modes have been overlooked in the previous chapter.

To get a better view on the impact of each failure mode, their frequency and costs per year will be analysed. The date range for the used data has been expanded to a period of four years (2012 – 2015). The side trimmer has not been changed in this period and extending the data range will allow for assessment of possible trends over the years. A section with an increasing amount of downtime each year could receive extra priority for this reason. The way in which material getting stuck is being recorded has changed in the middle of 2011, so it would be impractical to include any data from before 2012. In general, it is inadvisable to pick a period larger than 5 years. The data sets become very large which greatly increases the time required to retrieve, link and process them. At the same time, any extra data past this point is not expected to add to the reliability of the research.
5.1.1 Failure mode descriptions

Side trimmer blades

Figure 14 shows a side trimmer. In the left picture, the two blue parts are the blades (in the right picture, the blades are directly right of the red parts). The photo at the right shows a side trimmer of the pickling line. The metal strip is positioned to the left of the blades and will be guided through the red parts during operation. The part of the strip that is cut off is guided away to the right.

The side trimmer blades can fail in two different ways:

1. The blade can get either dull over time
2. A chip can break out of the blade.

The dulling of the blade is a natural wear process. However, a chip breaking out can have multiple causes. The blades are made of hardened steel, which has the benefit of being wear resistant, but also makes them somewhat brittle. Therefore, anything that causes a shock to the blade can cause hairline cracks that can lead to a chip breaking out. This includes careless handling during installation of the blade or processing a weld between two coils (the weld is a lot harder than the rest of the strip).

When either failure mode occurs, the blade is replaced by the production crew. The edge of a new side trimmer blade has a width of 41 mm. When the edge gets dull or a chip breaks out, it can be grinded to sharpen it and use it again. This maintenance action is performed on site by the technical service department of Tata Steel IJmuiden and the costs are charged to TSP afterwards. The blades are transported to their workshop, grinded and transported back to the pickling line. When a blade reaches a width of 22.5 mm, it needs to be discarded. When the stock of trimming blades becomes too low, new ones are ordered by the technical service department.

Once per quarter, the costs of repairs and new blades are charged to TSP. Because each blade can be grinded multiple times before it becomes too thin to use, both failure modes can occur on a blade.
during its lifespan. When a blade is sharpened or discarded, the reason is not specified. This makes it impossible to distinguish between costs related to chipped and dull side trimmer blades.

The frequency and duration of both failure modes per year (and the downtime costs related to it) are filtered out of the total downtime costs of the side trimmer section. This can be done based on the description of each instance of downtime; when the side trimmer blades are switched by the operators, this is specified in the description. In some cases, they specifically mention that it is caused by a chipped side trimmer blade. However, this is not done consistently, so a part of these failures is wrongfully attributed to the default failure mode “dull side trimmer blades”.

In conclusion, it is not possible to separate the two failure modes of the side trimmer blades. They will therefore be considered together in the subsequent steps. The experience of production crew is that chipping is more likely to occur soon after the blade has been deployed, while it takes longer for a blade to get dull.

Failure of the blades can be observed by a reduced quality of the edge of the strip after trimming. The strip is monitored continuously with cameras after the side trimmer. There are no other ways available to judge the condition of the blades, so currently there is no to observe that a failure will happen soon. This rules out the use of condition-based maintenance.

Scrap cutter

![Image of scrap cutter]

**Figure 15: Scrap cutter next to the pickling line**

There are two scrap cutters in use at any given time; at both the right and left side of the metal sheet, a thin strip is cut off by the side trimming blades. These scrap strips need to be cut up into smaller pieces. Figure 15 shows a scrap cutter. At the right side, you can see the green duct through which the strip enters (coming from the side trimmers). It is then guided towards the two turning blades at the left that cut up the strip. Afterwards, the small pieces of scrap (called *fries* due to their shape) come out of the back, where they are transported away by a conveyor belt.
When a failure occurs on a scrap cutter, it is lifted out of the production line and replaced by another one. For each side, there are three scrap cutters available. TSP tries to have one pair of cutters in the production line and at least one pair next to the line at all times to allow fast replacement of cutters whenever one fails.

There are two possible reasons to replace a scrap cutter:

1. When the blades get dull, they will occasionally fail to completely cut through the scrap. This results in longer pieces of scrap coming out of the cutter, which is monitored with cameras.
2. The other indicator of a failing scrap cutter is scrap getting stuck multiple times in short succession, which leads to an automatic emergency stop.

Just like with the side trimmer blades, there is currently no way available to assess deterioration before the failure occurs. Therefore, CBM is ruled out for this failure mode as well.

After a scrap cutter is replaced, it is sent for repairs to the service department, just like the side trimmer blades. The costs are charged back to TSP. A separate list is being used since the end of 2011 to keep track of the dates and cause of all scrap cutter replacements.

**Material stuck in scrap cutter**

This failure mode is both the most occurring and most complicated failure mode of the side trimmer section. After the scrap strip is cut off of the side of the product, it goes through a tube which guides it to the rotating knives of the scrap cutter. This tube is where the material gets stuck. There are many reasons possible for this to happen. The following possible influences have been observed by the production engineer:

- **External influences:** Flaws in the rest of the production line fall in this category. For example, a set of rolls before the side trimming section needs to align the strip to ensure that it is guided through the section in a straight line. If these rolls are not aligning the strip properly, it can change the shape of scrap strips, leading to them getting stuck.
- **Internal influences:** When either the side trimmer or scrap cutter knives need to be replaced, it can influence the quality of the scrap strip. This can in turn lead to the strip getting stuck. Therefore, this failure mode is expected to be related to the previous two.
- Finally, there is some influence of **production.** According to the experience of operators, cutting off a very thin scrap strip leads to more problems. This will be addressed further in the next section.

**Wrong cutting width**

When the side strips are being cut off at the wrong width, this only requires an adjustment of the knives. Therefore, no repair or replacement costs are attributed to this failure mode. However, there is a hidden cost on top of the lost production. When operators find out that the side trimmer is cutting at the wrong width, this can easily cause a complete coil to be rejected and scrapped. This problem usually occurs after switching a pair of side trimmers and is attributed to a flaw in the construction of the side trimmer. A project is underway to specifically address this problem, so this failure mode will not be addressed within this project.
5.1.2 Costs of failure modes

Using the failure data of the logbook during the 2012 – 2015 period, the average yearly frequency, duration and downtime costs for each failure mode have been determined. These have been summarised in Table 11. There is a general failure code for side trimmer blade issues and six failure codes for material stuck in the scrap cutter. Based on these failure codes together with the description of each entry, it is possible to filter out the entries of each failure mode.

For the side trimmer blades and scrap cutters, the repair and replacement costs have been added as well. The repair costs can be found in the maintenance orders, while the purchasing costs of new side trimmer blades can be found in material procurement data in SAP.

<table>
<thead>
<tr>
<th>Failure mode</th>
<th>Frequency (#/year)</th>
<th>Duration (hr/year)</th>
<th>Downtime costs (€/year)</th>
<th>Repair/ replacement costs (€/year)</th>
<th>Operating costs (€/year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dull / Broken side trimmer blades</td>
<td>160</td>
<td>44.23</td>
<td>€ 93.300</td>
<td>€ 100.000</td>
<td>€ 193.300</td>
</tr>
<tr>
<td>Dull scrap cutter blades</td>
<td>23</td>
<td>15</td>
<td>€ 31.635</td>
<td>€ 45.000</td>
<td>€ 76.635</td>
</tr>
<tr>
<td>Material stuck in scrap cutter</td>
<td>2552</td>
<td>192.41</td>
<td>€ 405.800</td>
<td>-</td>
<td>€ 405.800</td>
</tr>
<tr>
<td>Product cut at wrong width</td>
<td>26</td>
<td>7.03</td>
<td>€ 14.800</td>
<td>-</td>
<td>€ 14.800</td>
</tr>
</tbody>
</table>

The success of filtering out the different failure modes depends on the quality of the data. The results are in line with the expectations of Tata Steel though and the average frequency of dull scrap cutter blades is consistent with the replacement list of the scrap cutters. The logbook entries that cannot be attributed to either of these failure modes have been checked for recurring descriptions. No additional failure modes have been identified in these entries, so from this point on these will be discarded. Despite the discarded data, the total downtime costs recorded in Table 11 are significantly higher than the costs recorded during section selection. This difference is caused by adding the product-related failures of the side trimmer section.

The number of entries for material getting stuck in the scrap cutter has dropped from 2726 to 1899 in the last year and the downtime dropped from 207,25 to 175,69 hours. However, production and maintenance staff of TSP cannot explain the reason for this difference. On top of that, the downtime related to this failure mode varies a lot from year to year, so it is too early to consider the change of last year as a trend. The total number of occurrences and downtime from 2012 to 2015 can be found in Figure 16. The number of side trimmer blade issues was far above the average last year, with 193 (compared to the average of 160) entries causing 62,93 (average of 44,23) hours of downtime.
5.1.3 Prioritising failure modes

Based on the previous two sections, an order of priority can be determined for the different failure modes. The product jams in the scrap cutter inflict the highest costs, but they are expected to be related to the condition of the side trimmer and scrap cutter blades amongst others (see material stuck in scrap cutter, section 5.1.1). Therefore, the focus will be on these two underlying problems first. The side trimmer blades have shown a sharp increase in operating costs last year and are clearly ranked second among the operating costs related to each failure mode. They will be prioritised first for this reason.

If a relation between product jams and the condition of the side trimmers and scrap cutters can be quantified, this would offer new opportunities to determine an optimal point for their preventive maintenance. First, it is necessary to know the time at which the side trimmers and scrap cutters fail.

The failure mode “product cut at wrong width” will not be regarded since it scored low on both costs and feasibility regarding this project.

5.2 Introduction of predictors

Now that each failure mode has been specified, the possible predictors for each failure mode can be determined.

5.2.1 Predictors by experience

The following predictors are grounded on the experience of employees of TSP. Using the failure modes from the previous section, semi-structured interviews have been conducted with the maintenance engineer and the production engineer related to the side trimmer section.

*Processed length since last failure*

Each coil is of a different length, corresponding to the specifications of the customer. The length of strip produced before a piece of equipment fails can be measured and indicates clearly how much it has been used. It is considered to be an important influence on the degradation of both the side trimmer blades and the scrap cutters.
Metal strip gauge
The gauge (or thickness) of the metal strip varies for each coil. In general, the strip that is processed by the pickling line has a gauge of 1 to 3 millimetres. This is a large relative difference, so this could make influence the effort needed to cut the strip. On top of that, it can influence how easily the scrap strips of a coil will bend. This could have an effect on how easily it gets stuck in the scrap cutter.

Material resistance
Even though almost every roll going through the pickling line will end up as tin plate, there are a lot of differences in material characteristics. A strip with a higher resistance will inflict more strain on the machinery and could therefore reduce its lifetime. This is also an important reason for the manufacturer of side trimmer blades not to give an estimate of their time until failure. The effects of differing material resistances have not been observed by the production staff, since they do not know the resistance of the coils they process.

Location within the coil
There are two areas in a coil that have specific conditions. First of all, the weld that connects the strip from two coils is always harder than the rest. This briefly causes extra stress on the blades. On top of that, the steel received by the pickling line is quite crude. The material characteristics are often not completely stable within one coil. The start and end of a strip can suffer from a problem called necking. In the warm rolling process, the strip is constantly pulled in the length direction of the strip while it is being pressed. This makes sure that the width of the strip remains relatively stable throughout. When the tension gets a bit too high, the strip will deform; it gets thinner and can form waves on the surface. It can also locally change material characteristics, which could influence the side trimming process.

Strip width
The difference between the width of the roll when it is received by the pickling line and the required end width determines how much strip is cut off from the sides. Thin strips can be harder to shear, causing the strip to rather bend instead of breaking. On top of that, thinner strips bend more easily, which can cause them to bend when they hit the side of the tube leading to the scrap cutters. Thinner strips have been observed to lead to a higher occurrence of scrap getting stuck.

Scrap cutter set used
This failure is specific to the scrap cutter failures and products getting stuck. There are three scrap cutter sets for each side. At any given time, the aim is to have at least one scrap cutter set in reserve next to the line. This allows for quick changes in case of failures. There are slight differences between each set, which could influence the amount of failures. For example, the tube guiding the scrap to the cutters, which is part of a cutter set, is slightly different in one set.

Side trimmer blade age
As explained in the first section of this chapter, dull or broken side trimmer blades can be grinded and reused multiple times before they are discarded. The grinding process is supposed to return the condition of the blades to as good as new, but operators believe that grinded blades are likely to fail faster than new ones. This could be attributed to multiple influences, such as the quality of the grinding process and the transport of grinded blades between the service department and TSP.
5.2.2 Predictors from data

The data from the logbook has been checked for possible predictors that did not come up in the initial interviews. These four additional predictors have been verified with literature and the experts to decide whether or not they should be considered.

Time since last failure

Since the time at which a failure happened is known, it is possible to calculate the time between two failures. This measure is commonly used in reliability analysis. However, in this case, time is not expected to be a good predictor. Not all strip needs to be processed by the side trimming section. The production data table contains a column that indicates if the side trimmer is used. On average, 40% of all coils need to be trimmed. All other coils already have the right specifications to continue to the next production line. To avoid too many switches between the two types, they are each produced in batches. Therefore, the production times would need to be converted to only reflect the time that the side trimmer was in operation in order to have any predictive value. On top of that, the production time does not take into account when there was a production stop in the middle of a roll. It only records when the roll entered and left the production line. Based on these reasons, time since last failure will not be considered as a predictor, as processed length is considered to be a much better suited predictor and will be used instead.

Processed tonnes since last failure

Just like with time or produced length since last failure, it is possible to use the weight of the produced coils to calculate the processed weight between two failures. In some production lines within Tata Steel IJmuiden, this measure is used for preventive maintenance actions. However, the weight of the coil mostly depends on its length, width and gauge. These variables are already being considered as predictors themselves though. Therefore, it is assumed that if there is a relationship between processed weight and failures, it should also show up through the predictors strip length, width and gauge. This makes the processed tonnes since last failures redundant and it will not be used within this research.

Human error

There are certain procedures that production employees need to follow, but in practice, each operator might perform the same task in a slightly different way. This has been verified in the interviews. These small differences could affect the failure behaviour of the studied section. However, the production engineer does not believe that there will be a relation between degradation of the side trimmer section and the different crews. What could be checked for are differences in “replacement behaviour” of the crews. For example, some crews might decide to replace side trimmer blades or scrap cutters sooner than others.

Time of the day

Since production continues for 24 hours a day over three shifts, there might be some difference between night shifts and daytime shifts. On top of that, it is possible that certain mistakes occur more frequently at a particular time in a shift (for example towards the end of the shift). Checking the times at which the failures occur might show a relation between this predictor and failures. This assumption is supported by research of Folkard and Lombardi (2006). Their research combined the results of multiple studies relating different aspects of shiftwork to accidents and injuries. They found an increased risk in case of successive night shifts as well as varying risk levels for each hour of the
day. In general, the fifth and eighth hour of a shift lead to the most injuries and/or accidents. Therefore, it would be interesting to see if these risks are present at TSP as well.

Presenting these predictors to the production engineer led to another possible influence. During the summer holidays, TSP needs to hire external personnel to fill in for the people. Each year, a peak in failures is observed during this period, so there should be a seasonal influence here.

5.3 Available data on predictors

5.3.1 Primary data source

Until this point, three data sources have been considered: the maintenance orders, notifications and the logbook. Only a very limited amount of failures leads to a maintenance order, so orders do not provide a good overview of the total amount of failures. On top of that, the orders are not linked in any way to production. Therefore, the orders are not useful for the prediction of failures. The notifications mostly share these problems. Chapter 3 has shown that the notifications miss a lot of failures that are recorded in the logbook and since they are entered manually, the times of the failures are less reliable as well. This would make it very hard to link the notifications to production data in a reliable way.

The logbook contains more information than what has been used in the previous parts of the research. As explained in chapter 3, the logbook is split into a production and a production stop part. The production stop entries do contain information about production at the time of the failure. Each entry contains approximately 70 columns with data, including the identification number (ID) of the coil in production at the time of the production stop and the customer order to which it belongs. Some of these columns provide the same information or are not filled consistently. A list of all useful columns can be found in Appendix B.

The production part contains approximately 70 columns of data as well. There is a separate entry for each produced coil, filled with product and production information. A few of the recorded characteristics are the coil its length, width, gauge and identification number (ID). All useful columns can be found in Appendix B.

5.3.2 Linking necessary and available data

The data recorded in the logbook has been compared with the predictors proposed in the previous section to see if it provides the necessary data to test all these predictors. Product specifications such as gauge and width are recorded, so this data can be used directly. The influence of human errors can be tested using the different crew colours: each of the five crews has a colour which is recorded in each logbook entry. The location within the coil when a failure occurs is only recorded for one failure mode: scrap getting stuck. Because of the suspicion that there is a relation between scrap getting stuck and the location within the coil, six codes have been created in the column “TIB code”. These specify the side where the cut off strip of scrap got stuck (left or right side) and if it was close to the weld between two coils, in the middle of a coil or just after a notch. A notch is a cutout in the side of the strip. When this “hole” is positioned over the side trimmer, the cutting width of the blades can be changed. Figure 17

Figure 17: A notch between metal strips of two coils
shows a transition between two coils with a different width where a notch has been made.

Product hardness cannot be measured with the data in the logbook. A general description of the material type is included, but the hardness can vary within one material type. The process and product development department has a list of all customer orders and the specifications of the metal used for this order. Linking this list with the logbook data will allow for analysis of the relationship between product hardness and failures.

The scrap cutters in use are tracked separately in a file, together with the time and date at which they are changed. Based on the times at which they are replaced, this data can be merged with the logbook data as well.

The last predictor to consider from the previous section is the age of the side trimmer blades. Even though it is easy to judge this from the width of the blade, this is not recorded in the production and failure data. This makes it currently impossible to keep track of this predictor. A way to measure the difference between new and grinded blades would be to add a variable to the failure data. For example, a column could be added in which either a one or a zero should be recorded to signal whether the failed blade was new or grinded. With this variable, separate analyses could be made for new or grinded blades.

The available and necessary data for each predictor has been summarized in Table 12. If a predictor has been rated as necessary and can be assessed with the available data, it will be used.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Necessary</th>
<th>Data availability</th>
<th>To be used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processed length</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Gauge</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Resistance</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Location</td>
<td>Yes</td>
<td>Only for material getting stuck</td>
<td>Only for material getting stuck</td>
</tr>
<tr>
<td>Scrap strip width</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Used scrap cutter</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Side trimmer blade age</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Human error</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time of the day/year</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Processed weight</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Time since last failure</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>
6 Creating the data set

Now that the predictors that will be investigated have been chosen, it is necessary to link the various data sources. Four sources have been named throughout the previous chapter: the stops and production database of the logbook, the table with resistance codes of different metal types and a list with replacement dates of the scrap cutter. Figure 18 gives a first overview of the different data sets and in which order they can be linked. Further explanation of this process will follow throughout this chapter. The *failure data* part shows multiple boxes. This indicates that multiple sets of failure data (one for each failure mode) can be extracted from the production stops data file. When creating the final data set out of the production and failure data, a choice can be made on which failure data to include. Certain failure modes can be left out or analysed separately if needed.

![Figure 18: Linking of the different data sets](image)

### 6.1 Enhancing logbook data

#### 6.1.1 Production data

The product development department of TSP has calculated the material resistance of the different types of steel that are being produced. This calculation is based on two factors: the steel recipe code (indicating the specific composition of the steel) and the temperature at which the steel strip has been coiled at the hot rolling plant. The latter is influenced by how fast the strip has been cooled after hot rolling, which affects the hardness of the steel. The combinations of recipe codes and coiling temperatures that are currently in use at TSP are saved in a data table with their corresponding resistance number (“Resistance numbers” in Figure 18).

A second list contains all the customer orders of TSP. For each customer order, an order number, the recipe code, coiling temperature, customer and a date are recorded. The resistance number can be added to this list by checking the first table for a corresponding combination of a recipe and coiling temperature. This operation can be performed easily using a vertical searching function in Excel, which is explained in Appendix C.
The resulting list needs to be linked to the production data from the logbook. To relate each produced coil to a customer order, the customer order number can be used, since this number is included in both data sets. The customer names are available in both data sets as well, but these are not unique enough to link a produced coil to a specific customer order. The customer order numbers have three issues that need to be addressed to link the two data sets:

1. In the past, the pickling line of TSP has produced coils for other departments. The orders for these coils are not available and cannot be linked because of that. As of now, the capacity of other pickling lines at Tata Steel IJmuiden is high enough, so this problem will not be present in new data. On the 2012-2015 data set, this makes up about 1.6% of the produced coils. The resistance number for these coils will be set to “unknown”. Removing the coil is not an option, since the other parameters of the coil can and should still be used.

2. Customer order numbers are not unique. They consist of 5 digits followed by one or two letters. The digits identify the order, while the letters only distinguish between different batches within that order. However, the order numbers are not unique over a period longer than one year. Therefore, some other attribute needs to be used to prevent linking the wrong order numbers. There are two options: customer orders and dates. The customer names are only included in 40% of the coils in the production data. The dates of the orders and the production data are the order creation date and the production date respectively, so these are never the same. The difference between these two dates can vary from weeks to multiple months. By just using the year in combination with the order, the right matches can be made for double order numbers.

3. Not all order numbers of the production data can be found in the customer order list. If the 5-digit part of the number in the production list does not exist in the customer order list, there is no way to create a link. When just the letter is missing, it still can be linked. For example, if order 12345A and 12345C are in the customer order list but 12345B is missing, the resistance code of batch A and C can be used for batch B as well.

After dealing with these issues, 96% of the 81,400 coils in the production data could be linked to a material resistance number. The used SQL code has been added to Appendix C.

6.1.2 Production stops data

The scrap cutter replacement list contains the number (1, 2 or 3) of the scrap cutter being used at either side of the strip (left/right) and the replacement dates. This is not tracked anywhere in the logbook, but this information is of vital importance to analyse the scrap cutter failures. Without it, it is impossible to see which cutter has been replaced (left, right or both) and to assess the performance of each individual scrap cutter (1, 2 or 3). Table 13 shows the layout of the scrap cutter replacement list. The SAP notification related to the replacement can be looked up for more information. The cell containing the SAP notification number is replaced by the text “preventive” if a cutter has been replaced preventively, since no notification is made in case of preventive maintenance. The description provides extra information on the replacement.

<table>
<thead>
<tr>
<th>Cutter left side</th>
<th>Cutter right side</th>
<th>Replacement date</th>
<th>SAP notification nr.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>R1</td>
<td>01-01-2015</td>
<td>12345678</td>
<td>...</td>
</tr>
</tbody>
</table>
There is no way to automatically link the replacements to any entries in the logbook. However, it is possible to perform this step manually. Using the information provided by the replacement list (replacement date, related SAP notification and description), it is possible to find the corresponding production stop in the logbook. When using the SAP notification to find the time of the replacement, one minor detail to keep in mind is that there is a one hour difference between the times of the notifications and the logbook during daylight savings time (the logbook time does not take daylight savings time into account). The preventive replacements are always carried out during a planned stop. During these planned stops, multiple maintenance actions are performed across the pickling line. These maintenance actions are not specified individually in the logbook, so the preventive scraper replacements will not be recorded explicitly either. However, they can be linked to the planned stop that occurred on the replacement date.

6.2 Merging the logbook data

Since both the production and production stop part of the logbook are recorded automatically by the same system, all of their entries should be consistent. However, they are saved in separate modules within the system and cannot be retrieved together from the system as one homogeneous set of data. There are some columns in both data sets that provide the same information, making it possible to establish a link using these columns.

A new entry is created automatically in the production data for each produced coil. Each coil receives a number to identify and keep track of the coil. In case of a production stop, the number of the coil that was being processed at that time is added to the entry. Therefore, this would be an ideal way to link the production stops to the production data. However, there are two issues that complicate this process.

1. First of all, there are two IDs for one coil: the TSP coil number and the number it received at the hot rolling plant, called the mother coil number. This term is used since one hot rolled coil can be occasionally split up in a few smaller coils when parts of the coil are not of sufficient quality to be processed into packaging steel. In this case, the TSP number of each smaller coil will be appended with a letter (A, B, ...), while the same mother coil number is used each time. The production stops use the mother coil number to indicate at which coil the failure occurred. In case the mother coil has been split up into a few smaller coils, this causes one stop to be linked to multiple coils. To solve this issue, every production stop will be linked to the first coil created with a mother coil.

2. One mother coil number can be used for different coils over the course of years. Therefore, a new identifier has been created, combining the mother coil number with the week and year in which the coil has been produced. Based on this unique identifier, the production and production stop data sets can be coupled without creating incorrect combinations between produced coils and production stops.

The SQL code for all steps to create this link can be found in Appendix C. To test the code, all 11,083 production stops related to the side trimming section from 2012 to 2015 have been linked to the production data. The resulting table has been checked to see if the link created any double entries of production stops. On top of that, the completeness has been checked: 10,831 stops could be linked. The others contain a coil number that does not exist in the production data. There is no explanation available for this problem and since it only affects less than 2.3% of the production stops, these will be excluded from the research.
6.3 Verifying data quality
Now, the data quality of the final data set should be assessed to determine if it satisfies all requirements for this project.

6.3.1 Determining quality
Olson (2003) defined data quality as the extent to which it satisfies the requirements of its intended use. Within this research, intended use should receive extra attention; the used data has been recorded for other purposes (such as production control) and not for this research. For this purpose, Olson (2003) determined six measures for data quality: accuracy, timeliness, relevance, completeness, understood and trusted.

Accuracy indicates the amount of wrong, missing and double values within a data set. Olson (2003) states that this is the most important measure of quality; when a large portion of the values in a data set is missing or wrong, their analysis will never be accurate either.

Timeliness considers the time it takes before a data set is ready to be used. When data on certain events needs to be available within a week, but it takes a month to become available, it is not timely.

Relevance measures if the data set provides the information that is necessary. In the context of this research, relevance is used to assess which of the proposed predictors can be tested using the available data.

Completeness is the extent to which all data points are recorded in the data set. The production data is considered to be complete when all produced coils are recorded, while the failure data should contain every failure or replacement of components related to the research to be complete.

The data is understood if the researcher knows how to interpret the data. The most important implication of this for the TSP data is a good understanding of the used descriptions of different failures. The failures are not split up into different failure modes within the data set, so a good understanding is necessary to ensure that failures are attributed to the right failure mode.

Trust of data determines how willing users of the data are to believe its results and use them. This can be affected by earlier (negative) experiences with the data. This measure is purely subjective and there is no reason to assume that trust will be an issue. The data quality will therefore only be assessed based on the first five criteria.

6.3.2 Production data
First of all, the quality of the production data is judged. Apart from the age of the side trimmer blades, all necessary predictors are available. On top of that, the accuracy of the data is very high. Wrong values can be found by comparing the recorded values to the known installation limits. For example, a scrap strip should be between 6 and 50 mm wide. The biggest accuracy issue is present in the used scrap cutters: during a period of almost 2 months, the replacements were not tracked. This period needs to be excluded from analysis, but is not a big issue on the total period of four years.

Overall, the production data is easy to understand and interpret. However, it is important to note that this set of production data only includes coils that pass the side trimmer section. The production dates/times and coil numbers are therefore not always consecutive. Secondly, a change of the used scrap cutter does not necessarily mean that it has failed. The failure data need to be checked to see
the reason for replacement; it could be a preventive replacement. The quality of the production data has been summarized in Table 14.

Table 14: Assessment of production data quality

<table>
<thead>
<tr>
<th>Quality criterion</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>- Material resistance: 4% missing values</td>
</tr>
<tr>
<td></td>
<td>- Scrap width: 0,88% wrong/missing values</td>
</tr>
<tr>
<td></td>
<td>- Material gauge: &lt;0,01% wrong/missing values</td>
</tr>
<tr>
<td></td>
<td>- Length: &lt;0,01% wrong values</td>
</tr>
<tr>
<td></td>
<td>- Used scrap cutter: Unknown for 2-month period</td>
</tr>
<tr>
<td><strong>Timeliness</strong></td>
<td>Production data is complete and available within a day</td>
</tr>
<tr>
<td><strong>Relevance</strong></td>
<td>7 out of 8 predictors available (5 out of 8 without extra added data)</td>
</tr>
<tr>
<td><strong>Completeness</strong></td>
<td>All processed coils included</td>
</tr>
<tr>
<td><strong>Understood</strong></td>
<td>- Only side trimmed coils included</td>
</tr>
<tr>
<td></td>
<td>- Scrap cutter change is not necessarily failure</td>
</tr>
</tbody>
</table>

6.3.3 Failure data

Because the failure data needs much more manual input than the production data, there is a higher probability of input errors, so the overall quality is expected to be lower than the production data quality. Currently, the intended use of the failure data is mainly general downtime analyses. Examples include:
- Identifying the amount of up- and downtime of a production line;
- Being able to measure the proportion of planned and unplanned downtime;
- Measuring the amount of technical, production-related and product-related downtime;
- In the case of the side trimmer section, the production stops are also used for some performance indicators, such as the percentage of coils that are side trimmed and how often material gets stuck per week.

The data is accurate enough for these purposes, but needs to be assessed to judge if it is also accurate enough to predict failure behaviour. The accuracy of coil numbers is quite high: only 2.3% of the coils has a coil number that does not exist in the production data. The amount of missing failures and wrong entries is difficult to assess. It is not always possible to identify that a failure is attributed to the wrong section. On top of that, a failure that has not been recorded at all will be impossible to trace back.

Based on the failures for the side trimmer section, the accuracy issues that have been identified are:

- Failures that are attributed to the wrong section of the production line. These include failures are attributed to other sections, while their description or failure code points to a failure that is specific to the side trimmer section. It is very likely that some of these wrongly attributed failures have not been discovered and are missing from the final data set.
- The same issue applies to the failure codes. When a side trimmer blade is replaced because the scrap was getting stuck multiple times in short succession, the failure code for scrap getting stuck is sometimes used. For that specific instance of downtime it is not the right failure code though, which possibly leads to a failure being attributed to the wrong failure mode.
The relevance of the failure data shows a gap between the usage at TSP and the intended use within this research. A side trimmer consists of two blades (upper and lower) and there is a side trimmer at both the left and the right side of the strip. Whenever a failure occurs, the operator can decide to only switch one blade, one side or all blades at the same time. This is not specified consistently in the description of the failure, which makes it impossible to track the time until failure for each specific side trimmer blade.

The biggest quality issue is related to the completeness of the side trimmer blade data. This becomes apparent when comparing the amount of side trimmer blade failures in the logbook to the costs of grinding and new blades. The grinding costs are about €55,000 per year on average, while there are 160 recorded failures each year. The exact costs of grinding one blade are not known, but are estimated at €100. The average amount of blades replaced per failure can be estimated with the formula below:

\[
€55,000 = 160 \times \text{Average amount of blades replaced per failure} \times €100
\]

This would lead to an estimate of 3.44 blades replaced per failure, which is unrealistically high according to TSP employees. It is possible to change the side trimmer blades when the side trimming section is not in use. This could be either due to production stops elsewhere in the production line or coils that require no side trimming. Since these replacements do not lead to a production stop, they are not recorded. Because of these data quality issues, it will be impossible to give a reliable estimate (or prediction) for the time until failure for each side trimmer blade.

The quality of the scrap cutter data is a lot higher, especially due to the separate replacement list. An important part of its data understanding is that it can experience some other, incidental failures besides the blades getting dull over time. The replacements of the scrap cutters also include some preventive replacements. Therefore, when assessing the failure mode dull scrap cutter blades, not all of the scrap cutter replacements can be used as failures.

Table 15 on the next page summarizes all issues with the failure data quality. Based on this assessment, the scrap cutter failure data is of the highest quality and should be analysed first. The side trimmer blades need more detailed data collection to fit the goal of this research, but some insights on the predictors of their failure might be developed by analysing the current data set.
<table>
<thead>
<tr>
<th>Quality criterion</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Coil number: 2,3% wrong values</td>
</tr>
<tr>
<td></td>
<td>Failure attributed to the wrong section</td>
</tr>
<tr>
<td></td>
<td>Failures are attributed to the wrong category</td>
</tr>
<tr>
<td><strong>Timeliness</strong></td>
<td>Data is complete and available within a day</td>
</tr>
<tr>
<td><strong>Relevance</strong></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>For side trimmer blades, the exact blades that have been replaced are not always specified</td>
</tr>
<tr>
<td></td>
<td>There is some ambiguity whether or not the blades have been replaced</td>
</tr>
<tr>
<td><strong>Completeness</strong></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Side trimmer blades that have been replaced during other stops or production of coils that require no side trimming are not recorded</td>
</tr>
<tr>
<td></td>
<td>Scrap cutter failures: Unknown for 2-month period</td>
</tr>
<tr>
<td><strong>Understood</strong></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>A replacement is not necessarily a failure, this difference should be checked in the description</td>
</tr>
<tr>
<td></td>
<td>Failures that happened on the change of two shifts are split up and should not be seen as two failures</td>
</tr>
</tbody>
</table>
7 Grouping predictor values

The predictors material resistance, gauge and scrap strip width can contain many different values. For example, the gauge of the material is being tracked with an accuracy of a hundredth of a millimetre. To make these values easier to interpret and analyse, they will be grouped before further analysis.

7.1 Creating groups

7.1.1 Theoretical background

An important use of data grouping is the construction of histograms for continuous variables. By clustering multiple data points into groups, patterns can be identified. However, if too many values are clustered into one group, it is possible that too much information is lost. Figure 19 shows the risk of choosing groups which are too large. The left graph shows the distribution of an imaginary data set which has been divided into three groups by the dashed lines. The right figure shows the same observations plotted in a histogram in these three groups. The groups are all the same and the underlying pattern is no longer observable, which could lead to misinterpretation of the data. Based on the histogram, one would think that the line in the left graph would be straight (horizontal) as well.

![Example of wrong grouping](image)

Sturges (1926) developed a formula to determine an appropriate group size based on the binomial distribution, called Sturges’ rule:

\[
\text{Number of groups} = 1 + \log_2(n),
\]

(7.1)

where \( n \) represents the number of data points. The rule assumes normally distributed data and is therefore not always usable. Other rules have been proposed by Scott (1979):

\[
\text{Width of each group} = 3.5 \times \text{standard deviation} \times n^{-1/3}
\]

(7.2)

and Freedman and Diaconis (1981):

\[
\text{Width of each group} = 2 \times \text{interquartile range} \times n^{-1/3}
\]

(7.3)

The latter two rules also take into account how much the data is spread out, making them more robust than Sturges’ rule. The data set for this project contains 81,400 coils, leading to a high number of groups for any of these rules. The predictor material gauge from the data set can be used to illustrate this. The standard deviation of the material gauge over all coils is 0.37 mm. The interquartile range can be determined by sorting the data and calculating the difference between the largest and smallest value in the middle 50 percent of all values. For the material gauge data this is
equal to 0.7 mm. Using the rules by Scott (1979) and Freedman and Diaconis (1981) we get the following proposed group widths:

For Scott’s rule: \( \text{Width of each group} = 3.5 \times 0.37 \times 81.400^{-\frac{1}{3}} \approx 0.0299 \text{ mm}. \)

For Freedman and Diaconis’ rule: \( 2 \times 0.7 \times 81.400^{-\frac{1}{3}} \approx 0.0323 \text{ mm}. \)

The material gauge ranges from 1.36 to 3.05 mm, a range of 1.69 mm. With the two proposed group widths (0.0299 and 0.0323 mm), this would lead to 57 and 52 groups respectively. There are two problems with having this many groups:

- There are a lot of groups that will be almost empty each period or have a lot of variance between periods, which makes it hard to notice the influence of each specific group. This effect is reinforced by the fact that material is produced in batches;
- With very small group sizes (such as the 0.03 mm for gauge), there are no meaningful differences between the groups. A 0.03 mm difference in gauges of the processed material cannot even be measured reliably by the pickling line.

Bhardwaj (2008) mentions Sturges’ rule, but also proposes a more practical approach for determining the amount of groups. In the end, the number of groups should be determined in such a way that it results in easy interpretable data. For irregular data, it can also be appropriate to use varying group sizes.

For this research, each predictor will be assessed individually to determine a practical number of groups. An important consideration here is that the amount of groups should be suitable for visual comparison in graphs. The specific ways in which the groups for each predictor are determined are explained in the subsequent sections.
7.1.2 Material resistance groups

Material resistance is measured with resistance numbers ranging from 5.5 to 28.9. The exact relation between different resistance codes is not clear: material with a resistance number of 27.5 is not five times harder to cut than material with a resistance of 5.5. On top of that, some numbers have been rounded to whole numbers, while others contain a decimal. Therefore, the resistance numbers should be treated as ordinal variables. This means that their order from low to high is clear, but the exact difference between two numbers cannot be interpreted.

Because of this, the resistance numbers will be ordered into groups with closely related resistance numbers. In Figure 20, the resistance numbers that are present in the data set are represented by blue dots. Each set of resistance numbers that are close to each other has been clustered into one group. The range of each group is represented in the figure by a line. Some groups, such as the group around resistance number 10, contain a slightly bigger spread of resistance numbers. The reason for this is that coils with these resistance numbers are rarely produced. Creating separate groups would therefore only reduce the interpretability in these cases.

Next to the six groups shown in Figure 20, one extra group will be created for unknown resistance numbers. Table 16 shows how often each of these resistance group occurs.

<table>
<thead>
<tr>
<th>Resistance group</th>
<th>Number of coils in the group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>20827</td>
</tr>
<tr>
<td>Group 2</td>
<td>3359</td>
</tr>
<tr>
<td>Group 3</td>
<td>14844</td>
</tr>
<tr>
<td>Group 4</td>
<td>31612</td>
</tr>
<tr>
<td>Group 5</td>
<td>1911</td>
</tr>
<tr>
<td>Group 6</td>
<td>5558</td>
</tr>
<tr>
<td>Unknown resistance</td>
<td>3280</td>
</tr>
</tbody>
</table>
7.1.3 Material gauge groups

The gauges of material that can be processed by the pickling line vary from 1.35 to 3.05 millimetres. To determine how many groups should be created, the gauges of all coils have been plotted in a graph. Figure 21 shows all 81,400 coils on the x-axis and their corresponding gauges (sorted in ascending order). Most of the time, the graph shows an almost horizontal line which suddenly increases now and then. This means that there are a few values for the gauge that are very common (for example 2 mm), while all other values of the gauge are very uncommon. Based on the number of horizontal parts, six groups should be chosen. The group numbers are shown in the graph. The coils in the left- and rightmost part of the graph will be included in groups 1 and 6 respectively. Because these values are within the limits of 1.35 and 3.05 millimetres they are valid and should not be removed. They will not be placed in a separate group either, since these groups would contain very little coils.

![Figure 21: Gauges of all produced coils](image)

Figure 22 shows the size and bounds of each of the 6 groups that have been identified above. Any values outside of the possible range are identified as outliers. These values have been either erased or corrected if possible (in case of a wrongly placed comma for example).

![Figure 22: Grouping of gauges](image)
7.1.4 Scrap width groups

A similar process has been used for the different values for scrap strip width. Figure 23 shows the distribution of different scrap widths over all produced coils. The possible values range from 6 to 50 millimetres. Because every extra millimetre that is cut off is extra waste, the width of the scrap strip should be kept as low as possible. Therefore, scrap strips rarely exceed a width of 16 millimetres. All the exceptionally wide scrap strips (more than 16 millimetres) are grouped together to prevent the creation of a large number of groups that are hardly used.

The scrap widths between 6 and 16 millimetres all occur regularly and there are no gaps between possible values of the width. Therefore, there is no reason to use different group sizes for these values. The chosen group size is 2 millimetre. Wider scrap is expected to cause more stress on the scrap cutter blades due to its bigger surface. By choosing a group size of 2 millimetres, there is enough difference between the groups to observe a difference if there is any. Figure 24 shows the resulting six groups.

![Figure 23: Occurrence of difference scrap widths, grouped per 0.5 millimetre](image)

![Figure 24: Grouping of scrap strip widths](image)
7.2 Assigning data to groups

To use the created groups of predictors in practice, the data set needs to be adjusted. To test for the influence of the different groups of resistance, width or gauge on failure, the proportion in which each group has been processed until each failure needs to be calculated. Table 17 provides some example data to explain this process. The left three columns are available in the initial data set. The length, resistance number and the scrap cutter ID are known for each coil (the resistance number is represented by a group letter for this example).

Based on the resistance number, the group in which the coil belongs can be determined. The length of the coil is then added to that group. This process is repeated for each coil until the failure which is investigated occurs (in this case replacement of the scrap cutter). When this happens, all groups will be reset to 0. The row of the last coil that was produced before the scrap cutter replacement (the bold row in Table 17), contains the total production of that scrap cutter until failure. The production of that scrap cutter until its failure has now been split up into the resistance groups.

Table 17: Example data set for grouping

<table>
<thead>
<tr>
<th>Coil length</th>
<th>Resistance number</th>
<th>Scrap cutter ID</th>
<th>Resistance group A</th>
<th>Resistance group B</th>
<th>Resistance group C</th>
<th>Resistance group D</th>
</tr>
</thead>
<tbody>
<tr>
<td>800</td>
<td>A</td>
<td>1</td>
<td>800</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>900</td>
<td>B</td>
<td>1</td>
<td>800</td>
<td>900</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>700</td>
<td>B</td>
<td>1</td>
<td>800</td>
<td>1600</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>1100</strong></td>
<td>C</td>
<td>1</td>
<td>800</td>
<td><strong>1600</strong></td>
<td><strong>1100</strong></td>
<td>0</td>
</tr>
<tr>
<td>600</td>
<td>A</td>
<td>2</td>
<td>600</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

This process needs to be repeated multiple times on a large set of data (the 81,400 coils). First of all, the scrap cutters on the left and right side need to be considered separately since they are replaced independently. On top of that, the grouping needs to be performed for the predictors resistance, gauge and width. Therefore, it is necessary to automate this process as much as possible. For this reason, a script has been written in the programming language Visual Basic for Excel.

The Visual Basic script creates a large array similar to Table 17 and prints the last rows before failure (such as the bold line in Table 17) in an Excel sheet. This row contains the total amount of strip processed by the scrap cutter until failure, broken down into the different resistance groups. The proportion of each group is also written as a percentage of the total processed strip until failure.

There is no way to observe any patterns in such a data set. Therefore, the failures should be sorted by total amount of strip processed. This leads to a data table such as Table 18. In the next chapter, the data sets such as this one will be analysed to find a relation between the predictors and failures.

Table 18: Example of grouped resistance data

<table>
<thead>
<tr>
<th>Total processed amount until failure (in m)</th>
<th>Group 1 (in %)</th>
<th>Group 2 (in %)</th>
<th>Group 3 (in %)</th>
<th>Group 4 (in %)</th>
<th>Group 5 (in %)</th>
<th>Group 6 (in %)</th>
<th>Unknown (in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100.000</td>
<td>30</td>
<td>10</td>
<td>15</td>
<td>25</td>
<td>7</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>150.000</td>
<td>25</td>
<td>10</td>
<td>20</td>
<td>25</td>
<td>10</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>210.000</td>
<td>35</td>
<td>5</td>
<td>20</td>
<td>20</td>
<td>6</td>
<td>12</td>
<td>2</td>
</tr>
</tbody>
</table>
8 Testing predictor influence on failure

In this chapter, the influence of the separate predictors on failure behaviour will be tested. For each failure mode, one or more hypotheses will be generated based on the research up to this point (especially section 5.2). By testing these hypotheses, a better understanding of the failure modes can be achieved before attempting to create a general model. The emphasis will be on the scrap cutter failures, since their data has been found to be of the highest quality for this research in Section 6.3 of the report.

8.1 Scrap cutter failures

Over the 2012-2015 period, 161 scrap cutter replacements have taken place. Out of these, 83 were on the left side (called the drive side) and 78 were on the right side (called the control side). Not all of these replacements can be used for this research. All the predictors that have been gathered are based on the gradual wear of equipment. In case of the scrap cutters, this corresponds with the failure mode dull scrap cutter blades.

Other replacements can be divided into two groups:

- **Preventive replacements**: Some scrap cutters have been replaced before they failed in any way. Including these replacements in the analysis would only contaminate the data set. There is no way to judge how much longer the scrap cutter would have lasted if it had not been replaced preventively.
- **Other failures**: The scrap cutter also suffers from other failures that occur independently of the condition of the blades. These can be roughly divided into two causes: damage (for example due to a scrap strip that got stuck and damaged the scrap cutter) and misalignment. The latter problem can be created during revision of the scrap cutter. A lot of these failures lead to very early replacements. Again, there is no way to judge how much longer the blades would have lasted otherwise.

Removal of the preventive replacements and other failures leaves 37 replacements on the left side and 36 replacements on the right side.

8.1.1 Scrap cutter used

Hypothesis: *The mean time to failure depends on which scrap cutter is used*

As stated in Chapter 5, there are three scrap cutters for each side of the strip (left, L, and right, R). They are named L1, L2, L3 and R1, R2, R3. While each scrap cutter is essentially the same and has the exact same blades, they are believed to have different mean times to failure. If the hypothesis is true, the different scrap cutters might need to be tested separately in subsequent hypothesis tests. Otherwise, the variance caused by different scrap cutters might obscure other relations.

In the previous chapter (Section 7.2), the amount of strip that was processed until failure of the scrap cutter has been calculated. Each of these observations is linked to either of the six scrap cutters. To test the influence of different scrap cutters, the observations of each scrap cutter have been grouped. Figure 25 shows all the observed metres to failure for each scrap cutter using boxplots. First of all, there is one extremely high value for scrap cutter R1. Since this value is more than two times higher than all other values of R1, it will be treated as an outlier and removed. It is not certain if this scrap cutter really survived for this long or if some replacements were not recorded in between.
There are some differences between the six different scrap cutters. However, the differences are not very large. The difference between the left and right side for each set (for example L1 and R1) seems to be especially small. An explanation for this is that if one scrap cutter fails, the other side will sometimes be replaced under the assumption that it is probably close to failure as well. While it is not necessary to have scrap cutters with the same number on both sides (i.e. a combination of L1 and R2 is also possible), the sets with the same numbers have been kept together for about 60% of the produced coils (48400 out of 81400).

To test the hypothesis in an objective way, a t-test will be used. A t-test compares two average values and determines whether they are significantly different from each other. This judgment is based on the difference between the average values, the sample size and the standard deviation within the samples. Because of the small sample sizes (there are only 73 observations left spread out over six scrap cutters), it is hard to prove that there is a difference. The less observations are available, the larger the risk that differences are based on randomness. This is taken into account by the t-test. However, the test can still provide an indication of differences between scrap cutters.

Each scrap cutter has been compared with the two other cutters of the same side as well as the cutter with the same number of the other side (i.e. L1 is compared with L2, L3 and R1). The results are summarized in Table 19, based on a significance level of 0.05. If the p value is below 0.05, one can say with at least 95% confidence that two scrap cutters have different mean times to failure. The comparison between scrap cutters R2 and R3 is the only significant one. What should be kept in mind
is that due to the number of tests (9), the risk of a false positive increases as well. More information on this problem and the t-test can be found in Appendix D. Based on these results, the hypothesis will be rejected. There is no hard evidence that the differences between the times to failure of the scrap cutters are based on anything else than randomness.

Table 19: Cross table of t-test results. Values below 0,05 indicate significant difference of means

<table>
<thead>
<tr>
<th></th>
<th>L1</th>
<th>L2</th>
<th>L3</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td></td>
<td>0,96</td>
<td>0,485</td>
<td>0,332</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>L2</td>
<td>0,96</td>
<td></td>
<td>0,525</td>
<td>X</td>
<td>0,073</td>
<td>X</td>
</tr>
<tr>
<td>L3</td>
<td>0,485</td>
<td>0,525</td>
<td></td>
<td>X</td>
<td>X</td>
<td>0,366</td>
</tr>
<tr>
<td>R1</td>
<td>0,332</td>
<td>X</td>
<td></td>
<td>0,754</td>
<td>0,241</td>
<td></td>
</tr>
<tr>
<td>R2</td>
<td>X</td>
<td>0,073</td>
<td>X</td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>R3</td>
<td>X</td>
<td>X</td>
<td>0,366</td>
<td>0,241</td>
<td></td>
<td>0,032</td>
</tr>
</tbody>
</table>

X = not tested

Because there is no observed difference between the scrap cutters on the left and right side, no distinction will be made between the left and right scrap cutters for the remainder of this research. All subsequent hypothesis tests in this Chapter will consider the failure data of the right and left scrap cutter together. However, any duplicate observations (these occur when the right and left scrap cutter were both in use for the exact same time) will be removed. In these cases, only one of the two scrap cutters has really failed, while the other has been replaced under the assumption that it should be close to failure as well.

8.1.2 Material resistance

Hypothesis: *Material resistance influences the degradation of the scrap cutter blades: material of a higher resistance causes a blade to wear more quickly.*

In Section 7.1.2, seven groups have been created for the predictor material resistance. If the hypothesis is right, there should be a relation between the amount of strip processed by a scrap cutter until failure and the distribution of this amount over the seven groups. In general, a scrap cutter that failed fast is expected to have processed a high percentage of high-resistance material, while a scrap cutter that processed a lot of material before failure should have processed more low-resistance material.

The data, as depicted in Table 18, can be used to create the graph of Figure 26. The x-axis contains the total processed material at the time of failure of the scrap cutter, while the y-axis shows the percentage of this total which has been produced in each material resistance group. To reduce the variation slightly and make the figure easier to interpret, the values have been turned into a moving average of the last 3 failures.

The figure should show any trends in the proportion of each resistance group relative to the total processed material. The farther to the right you go on the x-axis, the longer the scrap cutter has survived. Therefore, according to the hypothesis that material gauge influences failure, the farther to the right you move in the graph, the higher the proportion of the low-resistance groups should be.
However, the differences in proportion of the resistance groups between the different occurrences of failure are very small and seemingly random. Group 6 for example, which contains the material with the highest resistance codes, makes up between 5 and 10 percent of the total processed strip on average and remains within this range for almost all observations. The scrap cutters that survived very long (the right part of the line) processed approximately the same percentage of hard material as the ones that survived less long.

What should be taken into account is that all lines show percentages and not absolute values. If you compare the right-most failure (over 3400 km of processed material) to failures on the left side (less than 500 km) in Figure 26, it should be noted that much more material has been produced in all resistance groups. This further refutes the idea that material resistance influences failure behaviour of the scrap cutter.

Based on these observations, the hypothesis should be rejected. The collected data does not support it in any way.
8.1.3 Material gauge

Hypothesis 1: Material gauge influences the degradation of the scrap cutter blades: material of a higher gauge causes a blade to wear more quickly.

Hypothesis 2: Material gauge influences the degradation of the scrap cutter blades: material that deviates a lot from the average gauge (in either direction) causes a blade to wear more quickly.

The same approach has been used for the predictor material gauge. In this case, two hypotheses have been generated. The first hypothesis is based on the assumption that it takes more force to cut a thicker strip. The two turning blades that cut the scrap into smaller pieces have a small gap in between them. This gap is set in a way that is ideal for the average material gauge and is never changed. Therefore, it is possible that scrap with a very low gauge is actually harder to cut than scrap with an average gauge.

![Proportion of gauge groups processed until failure](image)

**Figure 27: Proportion of production of each gauge group relative to time to machine failure (2012-2015)**

Based on the two hypotheses, attention should focus on the groups with especially low (1 and 2) and high gauges (5 and 6). None of the lines seem to be increasing or decreasing as they move towards the right, so both hypotheses should be rejected.
8.1.4 Scrap width

Hypothesis: Scrap width influences the degradation of the scrap cutter blades: material of a higher width causes a blade to wear more quickly.

The hypothesis for scrap width is based on the assumption that it takes more force to cut a wider strip, since it has a larger surface area. To confirm this hypothesis, the groups of relatively narrow strips (groups 1 and 2) should be more prevalent in scrap cutters that survived for a longer period of time (the right side of the graph) and wide strips (groups 5 and 6) should be more prevalent on the left side of the graph. Figure 28 does not support this hypothesis in any way. The percentages of different width groups are very similar and do not show any trend (constantly increasing/decreasing) throughout the different observations on the x-axis. The hypothesis will be rejected.

![Proportion of width groups processed until failure](image-url)
8.1.5 Scrap cross section

Hypothesis: The cross section of the scrap strip influences the degradation of the scrap cutter blades: scrap with a higher cross section causes a blade to wear more quickly.

By multiplying the width and gauge of the scrap, the cross section of the scrap strip can be calculated. The reason to combine these two predictors is that they are very similar. They both rely on the assumption that a larger scrap surface will lead to faster wear of the blades. When high-width strips have a very small gauge or vice versa, the effect of each predictor might be obscured when looking at these variables separately. Because the range of values for cross sections is larger than those of width and gauge individually, the amount of groups has been increased to 8. The range of each of the groups is equally large.

Figure 29: Proportion of production of each cross section group relative to time to machine failure (2012-2015)

Based on the resulting figure, Figure 29, the hypothesis on the cross section of scrap needs to be rejected as well. Again, there does not seem to be any trend across the different observations on the x-axis.
8.1.6 Crew

Finally, the differences between crews will be investigated. This analysis differs from the other ones, as it will not focus on production until the moment of replacement, but on the replacement itself. Section 5.2.2 of this report has explained that differences between crews are not expected to lead to faster wear of the scrap cutters.

The crews follow each other up according to a fixed schedule. In the time until replacement of a scrap cutter, each crew will have operated that scrap cutter for an equal amount of hours. What can be checked with the available data is whether some crews replace the scrap cutter more often than others. The replacements of the scrap cutters on both the left and right side have been used together. If both scrap cutters were replaced, this will still be counted as one replacement.

Figure 30 shows that over the time period 2012-2015, the green and white crew have performed twice as many replacements as the three other crews. Due to the small number of replacements, this cannot be seen as hard statistical evidence, but it might be interesting to look into why these two crews replace the scrap cutter more often. This is not within the scope of this research and will therefore not be investigated in depth. It should be noted though that this difference cannot be attributed to human error.

Figure 30: Number of scrap cutter replacements per crew colour (2012-2015)
8.2 Material stuck in scrap cutter

In Section 5.1 of this report has been shown that the highest costs of the side trimming section are caused by the failure mode of scrap getting stuck in the scrap cutter. On top of that, this failure mode is expected to be related to the condition of the scrap cutter amongst others. Because the failure mode has been logged including the side of the failure (left or right scrap cutter), it is possible to relate this failure mode to replacements (and therefore the age) of scrap cutters.

**Hypothesis:** The condition of the scrap cutter influences how often material will get stuck in the scrap cutter; the older the scrap cutter, the more often material will get stuck.

The hypothesis has been represented graphically in Figure 31. At the start of the graph, the scrap cutter is new and there are very little incidents of scrap getting stuck. As the age of the scrap cutter increases, the amount of incidents increases more and more. To test this hypothesis, the failure mode “scrap getting stuck” has been plotted against replacements of the scrap cutters. If the result looks like Figure 31, the hypothesis will be tested further, otherwise it will be rejected.

![Figure 31: Expected relation between scrap cutter age and material getting stuck](image)

![Figure 32: Scrap getting stuck versus strip processed by a scrap cutter](image)

Figure 32 shows the first part of the results of this test. The full results can be found in Appendix E. The x-axis shows how much strip has been processed (in thousands of kilometres), while the y-axis shows how many times the scrap got stuck in the scrap cutter since the last replacement. Every time the scrap cutter is replaced, the number on the y-axis is reset to zero. Therefore, the parts between two replacements should have the form of Figure 31. However, there is no increase of scrap getting stuck just before the replacements of the scrap cutter, so this hypothesis will be rejected as well.
8.3 Conclusion

Based on the results of this chapter, it is not possible to predict any failures related to the scrap cutter using the proposed predictors with the available data. Therefore, there seems to be no reason to adjust maintenance intervals to the levels of any of the considered predictors.

An observation across the different tests is that there are no large differences between the values of the predictors over time. Material of different resistance, gauge or width is produced in similar proportions over longer periods of time. This can be explained by the fact that customer demand is based on long-term contracts. Therefore, the production does not change too much from day to day. On top of that, scrap cutters do not fail very often, so by the time one fails, it usually processed most different types of material.

One drawback of the method used in this chapter is that all replacements due to other reasons than dull scrap cutter blades have been left out. For the scrap cutter, these make up approximately half of the observations. Appendix F provides a method that takes into account all replacements. While this does not lead to a different conclusion for the scrap cutters, it can be useful in situations where most replacements happen preventively and the method described in this chapter is not usable.
Part two: Maintenance optimization

In the introduction of this report, the main goal of the project has been defined as: A model/method to predict failure behaviour of machine sections to be able to increase their maintenance performance. Up to this point, research has focused on the first part of this goal (prediction of failure behaviour). Even though no relation could be found between any of the predictors and the failure behaviour, it may still be possible to increase the maintenance performance of the scrap cutter. The predictors will no longer be considered in this step, since they did not seem to influence failure behaviour.

First of all, the current maintenance strategy is discussed. The scrap cutters of the pickling line will run until failure and are replaced at this point. This is a corrective maintenance policy. In the past (mostly from November 2011 until June 2012), the scrap cutters have also been replaced preventively before a failure occurred. During this period, preventive replacements were executed every 1 or 2 weeks.

In parallel to this research, a plan is being developed by TSP for a new preventive maintenance policy, based on the amount of processed kilometres of strip by the scrap cutter. The amount of processed strip until replacement in this plan is 2000 kilometres and is based on the preventive maintenance policy of the scrap cutters at another pickling line at Tata Steel.

In the next chapters, different maintenance strategies will be considered. Based on the available data, a selection of strategies which can be used for the scrap cutter will be made. The optimal parameter values for these strategies will be determined and then compared to the current situation (corrective maintenance). This will lead to recommendations for the maintenance of the scrap cutter. On top of that, the plan by TSP to replace the scrap cutter every 2000 kilometres will be evaluated.
9 Choosing a maintenance strategy

In this chapter, a choice will be made on which maintenance strategies are to be considered for the scrap cutter of the pickling line. For this purpose, different maintenance strategies will be discussed briefly along with the prerequisites for their usage. After this, the failure data of the scrap cutter will be analysed.

9.1 Possible maintenance strategies

Van Dijkhuizen (2000) provides a general perspective on different maintenance strategies. Three main strategies are defined: corrective, preventive and predictive maintenance.

Corrective maintenance consists of maintenance performed upon or after failure. Van Dijkhuizen (2000) breaks it down further into opportunity- and failure-based maintenance. Opportunity-based maintenance is a special case, where failed components do not have to be replaced immediately (for example due to redundancy). In this case, the failed component will be ignored until an opportunity for replacement arises. This is not applicable for the scrap cutter and will not be considered.

Preventive maintenance is performed before failure according to a planning. This can be based on either time or usage. In the analysis of predictors (Section 5.2) it has been concluded that time is not a good option for the side trimming section since it is not constantly in use (not all coils need to be side trimmed). Therefore, only usage-based maintenance will be considered in this chapter. Usage will be measured by the amount of processed kilometres, just like in the first part of the report.

Finally, predictive maintenance is introduced. The goal of prediction is to predict the time of failure to perform maintenance just before the failure occurs. It can therefore be seen as a special case of preventive maintenance. Condition-based maintenance tries to accomplish this by monitoring the condition of the components. If a relation between production and failures would have been found in the first part of this report, this could have led to a predictive maintenance plan. There is no way to assess the condition of the side trimmer blades during production at the moment. Therefore, condition-based maintenance will not be considered either.

This leaves two options for the scrap cutter: usage-based maintenance and failure-based maintenance (the latter will be referred to as corrective maintenance from this point on). When choosing a preventive maintenance strategy over a corrective strategy, the number of replacements will increase to prevent that the equipment will break down. Intuitively, this increases costs, so it is only worth it if there is some benefit to perform preventive replacements. This leads to two prerequisites to choose preventive maintenance over corrective maintenance:
1. The costs of a preventive maintenance action should be lower than the costs of a corrective maintenance action. An example is when failure of one component can cause damage to other parts of the machine. If this prerequisite is not satisfied, preventive maintenance increases the number of replacements (and therefore costs) while offering no benefits.

2. The failure rate of the component should be increasing. The failure rate is the probability that a component will fail at a moment in time, given that the component has survived until then. When the probability that a component will fail decreases over time (or remains constant), it is not beneficial to perform preventive maintenance. This can be illustrated with the following example: When the probability that a component fails is the highest just after replacement (which is true in case of a decreasing failure rate), you would never want to replace it preventively, because by replacing it, you have only increased the probability of failure.

The first prerequisite is satisfied in the case of the scrap cutter. A corrective scrap cutter replacement requires the pickling line to be stopped, which leads to downtime costs (this has been discussed in Section 3.1 of this report). When it is replaced preventively however, it can be performed during a planned stop of the production line. These stops occur weekly for regular maintenance on the pickling line. A scrap cutter replacement would not extend the duration of the stop. Therefore, a preventive replacement would not cause any downtime costs. This makes it cheaper than a corrective replacement. The second prerequisite will be tested in the next sections.

9.2 Censored data

Before continuing with the analysis of the scrap cutter failures, the concept of censored data should be covered. Censoring is a common problem in data analysis caused by incomplete data points. In case of failure data of a system, this means that either the starting time (when the system was started to be used) or the time of failure is unknown. This can be caused by wrong or missing values in the data set. In this case, some data points might have to be removed completely. However, it is also possible that a system was replaced before it actually failed (for example due to preventive maintenance). In these cases, it is not possible to determine when it would have failed. This data should not be removed though, since the time that it did survive is known and is useful to know. It did not fail either, so these data points should be treated differently.

![Figure 34: Example of multiply censored data](image-url)
There are a few different types of censored data. For this research, only multiply censored data will be considered. In multiply censored data, the censored data points occur at different points in time. An example of this can be found in Figure 34. Other types of censoring can be found in Ebeling (2010).

In the case of the scrap cutters, data is censored when the scrap cutter is replaced preventively before it failed. The times after which the preventive replacements happened are different for each observation, so the replacement times are multiply censored.

The failure data of the scrap cutter has already been collected and structured for the first part of the research. Therefore, replacements due to dull blades have already been identified, just like replacements for preventive replacements or other failures (see Section 8.1). This distinction is important to make. The deterioration of the blades of a scrap cutter over time is a logical result of their usage. Other failures of the scrap cutter are not linked to the blades, but to other parts of the scrap cutter (for example the tube that leads towards the blades). On top of that, these failures tend to happen especially during the early life of the scrap cutter. Due to the very different nature of these two failure modes, they should be analysed separately.

Figure 35 shows 6 imaginary data points where the scrap cutter has been replaced for various reasons. For each of the two separate failure modes, these data points need to be used differently:

1. Failure mode **dull blades**: All points labelled F in Figure 35 will be treated as failures, while the data points labelled O and P (other failures or preventive replacements) will be treated as censored data for this failure mode;
2. Failure mode **other failures**: All points labelled O will be treated as failures, while the data points labelled F and P will be treated as censored data for this failure mode.

![Figure 35: Example of censored data specific to the scrap cutter](image)
9.3 Analysis of failure data

The data of the scrap cutter replacements can be used to create an empirical distribution of the failure times. An empirical distribution can provide some insight in the failure distribution of the sample data (in this case the amount of kilometres each scrap cutter processed until replacement). However, the amount of observations is limited. If a theoretical distribution can be found that fits the observations well, it is possible to model future failure behaviour with this distribution.

Therefore, the preferred option is to “fit” an existing theoretical distribution (e.g. a Normal distribution) to the sample data. This means that the parameter values of such a distribution are chosen in a way such that the distribution closely follows the pattern of the observed values. The method of Ebeling (2010) for identifying failure distributions will be used for this purpose. It consists of the following steps:

1. Identify candidate theoretical distributions;
2. For each candidate theoretical distribution, determine the parameter values that result in the best possible fit for that distribution;
3. Determine which of the candidate distributions of step 2 achieves the best fit with the data.

The idea of this method is to first come up with theoretical distributions that could possibly fit the data well. Step two results in the best possible fit for each separate distribution. Depending on the data set, some distributions will be able to achieve a better fit then others. By comparing the distributions and picking the best one (step 3), the best possible fit can be achieved.

9.3.1 Identifying candidate distributions

To get a first impression of the failure data, they are presented graphically in a histogram. To split up the data in groups to be used in the histogram, Sturges’ rule will be applied (see formula 7.1). In Chapter 7, each group needed to be analysed separately, which made Sturges’ rule very impractical due to the large number of groups it proposed. In this chapter, the goal is to create a histogram which provides a good representation of the overall distribution of the failure data. For this purpose, Sturges’ rule is better suitable. It also provides an objective and easily usable way to determine the amount of groups for histograms of failure data.

The data set contains 82 scrap cutter replacements. However, only 37 of these replacements are caused by failures due to dull blades and 25 are due to other failures. The remaining 20 replacements were performed preventively. Two histograms will be made: one for the replacements due to dull blades and one for the other failures. The amount of groups determined by Sturges’ rule is:

\[
Number \ of \ groups \ (dull \ blades) = 1 + \log_2(37) \approx 6
\]

\[
Number \ of \ groups \ (other \ failures) = 1 + \log_2(25) \approx 6
\]

The groups are each of equal size, so the width of each group is equal to the highest value divided by 6. The highest observed amount of kilometres processed until failure is 3411 for the failures due to dull blades. Based on Sturges’ rule, these observations will be divided in 6 groups of 570 km (3411 / 6 \approx 570). This leads to the histograms of Figure 36 (replacements due to dull blades) and Figure 37 (replacements due to other failures). The y-axis shows the probability of failure per kilometre.
Based on the visual inspection of these two histograms, some general remarks can already be made on the possible theoretical distributions for the failure data. First of all, both histograms are not symmetrical. Because of this, it would make no sense to try to fit a symmetrical distribution such as the normal distribution.

Three distributions that are used frequently in failure analysis and will be considered to model the failures are the lognormal distribution, Weibull distribution and gamma distribution. All three distributions are commonly used for modelling of failure distributions (Ebeling, 2010). Based on the values of their parameters, they can model a decreasing, constant and increasing failure rate, making them very flexible. Their probability density functions, \( f(t) \), are given in Table 20.

### Table 20: Probability density functions of Lognormal, Weibull and Gamma distribution

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Probability density function ( f(t) )</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull</td>
<td>( f(t) = \frac{\beta}{\theta} \left( \frac{t}{\theta} \right)^{\beta-1} e^{-\left( \frac{t}{\theta} \right)^\beta} ) for ( t \geq 0 )</td>
<td>( \theta = \text{scale parameter} )&lt;br&gt;( \beta = \text{shape parameter} )</td>
</tr>
<tr>
<td>Lognormal</td>
<td>( f(t) = \frac{1}{t \sigma \sqrt{2\pi}} e^{-\frac{\ln(t-\mu)^2}{2\sigma^2}} ) for ( t &gt; 0 )</td>
<td>( \mu = \text{location parameter} )&lt;br&gt;( \sigma = \text{shape parameter} )</td>
</tr>
<tr>
<td>Gamma</td>
<td>( f(t) = \frac{t^{\gamma-1} e^{-\frac{t}{\alpha}}}{\alpha^\gamma \Gamma(\gamma)} ) for ( t \geq 0 )</td>
<td>( \alpha = \text{scale parameter} )&lt;br&gt;( \gamma = \text{shape parameter} )</td>
</tr>
</tbody>
</table>

#### 9.3.2 Estimating the optimal parameter values

To determine the best fitting parameter values for a distribution, the method of Maximum Likelihood Estimation (MLE) is presented by Ebeling (2010). This method can also be used for multiply censored data. Likelihood is a score which indicates how likely it is that a set of observed values follows a certain theoretical distribution. The higher the likelihood, the better the distribution fits. The parameter values of a distribution that produce the highest likelihood are considered best according to this method. Ebeling (2010) proposes the following formula for the likelihood of a theoretical distribution in the case of multiply censored data:

\[
L(\theta) = \prod_{i \in A} f(t_i^-; \theta) \prod_{i \in B} R(t_i^+; \theta).
\] (9.1)

In this formula, \( L \) is likelihood, \( t_i^- \) is an observed failure time, \( A \) is the set of observed failure times, \( t_i^+ \) is a censored failure time, \( B \) is the set of censored failure times and \( \theta \) symbolizes the unknown value(s) of the parameter(s) of the distribution. \( f \) is the probability density function of the
distribution, while \( R \) (the reliability function) is equal to 1 minus the cumulative density function of the distribution. The specific likelihood functions for the Weibull, lognormal and gamma distribution can be found by filling in their probability density function and reliability function in formula (9.1). As an example, the likelihood function for the Weibull distribution is presented below:

\[
L(\theta, \beta) = \prod_{i \in A} \beta \left( \frac{t_i}{\theta} \right)^{\beta-1} e^{-\left( \frac{t_i}{\theta} \right)^{\beta}} \prod_{i \in B} e^{-\left( \frac{t_i}{\theta} \right)^{\beta}}.
\]

(9.2)

To determine the maximum of this function, the natural logarithm (ln) of this function is taken. By setting the derivative of this resulting function equal to 0, the optimal values of the parameters can be determined. In case of formula 9.2, this would mean that the derivative with respect to \( \theta \) should be taken and set equal to 0 to determine the optimal value of \( \theta \) and the same procedure should be repeated for \( \beta \). For each of the candidate distributions, this leads to equations which can only be solved in a numerical way. Therefore, this procedure will be executed with a Matlab script, which can be found in Appendix G.

The resulting optimal parameter values for each distribution are presented in Table 21.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Failure mode dull blades</th>
<th>Failure mode other failures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull</td>
<td>( \theta = 2100,4 )</td>
<td>( \theta = 7141,4 )</td>
</tr>
<tr>
<td></td>
<td>( \beta = 1,83 )</td>
<td>( \beta = 0,59 )</td>
</tr>
<tr>
<td>Lognormal</td>
<td>( \mu = 7,39 )</td>
<td>( \mu = 8,71 )</td>
</tr>
<tr>
<td></td>
<td>( \sigma = 0,77 )</td>
<td>( \sigma = 3,04 )</td>
</tr>
<tr>
<td>Gamma</td>
<td>( \alpha = 761,46 )</td>
<td>( \alpha = 13329 )</td>
</tr>
<tr>
<td></td>
<td>( \gamma = 2,52 )</td>
<td>( \gamma = 0,55 )</td>
</tr>
</tbody>
</table>

9.3.3 Determining the best fitting distribution

To compare the three candidate distributions with the observed data, the empirical reliability function will be determined and compared to those of the three candidate distributions. The reliability function, also called survival function, shows the probability that the scrap cutter will survive after having processed a certain amount of strip.

The empirical reliability is estimated with the following formula (Ebeling, 2010):

\[
\hat{R}(t_i) = \left( \frac{n+1-i}{n+i+1} \right) \delta_i \hat{R}(t_{i-1}).
\]

(9.3)

\( \hat{R}(t_i) \) is the estimate of the probability that the scrap cutter will survive until kilometre \( t_i \)

\( n \) is the total amount of replacements (including censored data)

\( i \) is the number of the replacement

\( \delta_i \) indicates if observation \( i \) is censored or not: censored = 0, not censored = 1

In order to use this formula, all replacement times (including censored data) should be sorted in ascending order.

By estimating the empirical reliability for all available data points and plotting them, the empirical reliability curve can be created. For both failure modes, the empirical reliability curve will be created. The three candidate distributions with the parameter values determined in the previous section will be compared to this reliability curve to test their fit.
Failure mode I: dull blades

First, the failure mode dull blades will be considered.

Figure 38 shows the empirical reliability curve for dull blades resulting from formula 9.3, together with the fitted theoretical distributions using formula 9.1. The three smooth lines are the theoretical curves (Red = Weibull, Green = Lognormal, Black = Gamma), while the blue rough line (going down in steps) represents the empirical reliability curve. The closer the fitted line is to the empirical reliability curve, the better its fit. All three theoretical distributions follow the empirical curve closely and seem to provide a good fit to the empirical distribution. However, the lognormal distribution seems to underestimate the reliability in the range between the 1000 and 1500 processed kilometres in Figure 38, since its line is quite far from the empirical reliability function. To determine which distribution provides the best fit, an objective statistical test should be used.

Multiple statistical methods to test the hypothesis that a theoretical distribution fits the empirical distribution well are available. The Chi-square test can be used for multiple distributions, but is weak for small sample sizes and cannot be used for multiply censored data. Other tests include the Kolmogorov-Smirnov test, which can be used for normal and lognormal distributions, and Mann’s test, which is specific to Weibull distributions. However, these tests do not account for multiply censored data either.

Without these tests, it is not possible to prove statistically that the distribution of failures can be described by any of the candidate functions. However, based on how close the candidate distributions follow the empirical reliability function (Figure 38), there is no reason either to reject any of the candidate distributions.
However, this leaves the issue which of the three candidate distributions is closest to the empirical function. This can be tested using the Akaike Information Criterion (AIC), first proposed by Akaike (1973). It is also based on the likelihood function and suitable for any type of censored data:

$$AIC \text{ score } = 2k - 2 \ln (L),$$

(9.4)

where $k$ is the number of parameters (which is 2 for all candidate distributions) and $L$ is the likelihood function. The AIC score cannot judge if a model fits the sample data. It is merely a comparison of the set of models and shows which one is closest to the original data. This is indicated by the lowest AIC score. For small data sets, the AIC score should be corrected slightly:

$$AIC_c = AIC + \frac{2k(k+1)}{n-k-1},$$

(9.5)

where $n$ is the total number of replacements (82 in the case of the scrap cutters). More information on usage of the AIC can be found in a review by Burnham, Anderson and Huyvaert (2011). Both the AIC and corrected AIC ($AIC_c$) are provided in Table 22. Since all three candidate distributions have the same number of parameters (2), the correction of the AIC does not influence the difference between the distributions in this case.

<table>
<thead>
<tr>
<th>Candidate distribution</th>
<th>AIC score</th>
<th>AICc score</th>
<th>$l_i$</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull distribution</td>
<td>635,74</td>
<td>636,09</td>
<td>0,878</td>
<td>44%</td>
</tr>
<tr>
<td>Lognormal distribution</td>
<td>637,92</td>
<td>638,28</td>
<td>0,098</td>
<td>5%</td>
</tr>
<tr>
<td>Gamma distribution</td>
<td>635,61</td>
<td>635,96</td>
<td>1</td>
<td>51%</td>
</tr>
</tbody>
</table>

Table 22: Results of the AIC for dull blades

Based on these results, the Gamma distribution provides the best fit. It should be noted that the difference with the Weibull function is very small. Out of the three tested distributions, a relative probability that a distribution is the best fitting one can be calculated with the following two formulas (Burnham et al., 2011):

$$l_i = e^{-\frac{1}{2} \Delta_i}$$

(9.6)

$$\text{Probability}_i = \frac{l_i}{\sum_{j=1}^{M} l_j}$$

(9.7)

In these formulas, $l_i$ is the likelihood of distribution $i$ compared to the distribution with the lowest AIC score. $\Delta_i$ is the difference between the lowest AICc score (635,96 in this case) and the score of distribution $i$. $M$ is the total number of distributions that were tested (4 in this case).

Formula 9.6 is closely related to the AIC formula. In this formula, the logarithm of the likelihood is multiplied by minus two: $-2 \ln (L)$. In formula 9.6, the difference in the AIC score of two models is transformed back to a likelihood, $L$, since $e^{-\frac{1}{2} \Delta}$ is the inverse of $-2 \ln (\Delta)$.

Formula 9.7 turns the likelihoods into probabilities. This is the probability that the distribution is the best fitting distribution out of the tested options. These can be found in Table 22 as well.

Based on these results, the gamma distribution with parameter values $\alpha = 2,5174$ and $\gamma = 761,4635$ will be used to model the failures mode dull blades.
Failure mode II: other failures

For the other failures, the empirical and theoretical reliability functions have been determined in the same way as for the failure mode *dull blades*. This leads to the reliability curves in Figure 39. The Weibull and gamma distributions produce very similar reliability curves. The values of their shape parameters (β and γ respectively), are below 1, which indicates a decreasing failure rate. This will be discussed in the next section. The lognormal distribution (the green line) starts to deviate from the empirical curve in the right part of the graph.

![Empirical and theoretical reliability curves of failure mode other failures](image)

The three theoretical distributions follow the empirical curve closely, so none of the distributions will be rejected based on the graph. Therefore, the AIC scores will be used again to determine the best fitting distribution out of the three candidates.

Based on the results in Table 23, the gamma distribution provides the best fit for this failure mode. It slightly outperforms the Weibull distribution again. Therefore, the failure mode *other failures* will be modelled by a gamma distribution with parameter values $\alpha = 13329$ and $\gamma = 0.5457$.

<table>
<thead>
<tr>
<th>Candidate distribution</th>
<th>AIC score</th>
<th>AIC$_c$ score</th>
<th>$l_i$</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull distribution</td>
<td>451.27</td>
<td>451.82</td>
<td>0.79</td>
<td>40%</td>
</tr>
<tr>
<td>Lognormal distribution</td>
<td>452.81</td>
<td>453.36</td>
<td>0.17</td>
<td>9%</td>
</tr>
<tr>
<td><strong>Gamma distribution</strong></td>
<td><strong>451.04</strong></td>
<td><strong>451.58</strong></td>
<td><strong>1</strong></td>
<td><strong>51%</strong></td>
</tr>
</tbody>
</table>
9.4 The hazard rate

As stated in Section 9.1, one of the two prerequisites for preventive maintenance is an increasing failure rate (also known as hazard rate). Using the failure functions for the two failure modes that have been determined in the previous sections, this hazard rate can be determined. The hazard rate function can be calculated using the probability density function and reliability function:

\[
h(t) = \frac{f(t)}{R(t)}.
\]  

The two resulting hazard rate functions are presented in Figure 40. The red and black line represent the hazard rates of dull blades and other failures respectively.

![Figure 40: Hazard rate of the two failure rates](image)

The hazard rate of failure mode **dull blades** is continuously increasing. That means that the longer they have survived, the higher the probability becomes that they will fail. This satisfies the second prerequisite for preventive maintenance, stated in Section 9.1. However, the hazard rate of the other failures is strongly decreasing during the first 200 kilometres after which it continues to decrease less strongly. This is caused by a high amount of early failures. Ebeling (2010) defined this period of early failures as the **burn-in period**. He describes the main cause of this type of failures as manufacturing defects. Since the scrap cutters are repaired after replacement, the defects can also be caused by the maintenance action.
Because the two failure modes result in the same maintenance action (replacement of the scrap cutter), it makes no sense to look at one of the two failure modes without considering the other. The failure mode other failures creates a high risk of failure just after replacement, so there is always a risk that the scrap cutter will break down again shortly after replacement. However, based on the constantly increasing failure rate of the failure mode dull blades, preventive maintenance can still be considered.

At the same time, it is recommended to try to reduce the amount of failures that happen just after replacement of the scrap cutter. A possible approach for this would be to first inspect the quality of a scrap cutter after it has been repaired. On top of that, it is important to make sure that the scrap cutters are not damaged during transport back to TSP, storage or installation.

9.5 Conclusions
Based on the results of this chapter, a preventive maintenance strategy can still be considered for the scrap cutters. Two prerequisites for preventive maintenance have been considered: a preventive maintenance action should be cheaper than a corrective maintenance action and the failure rate should be increasing. Both of these conditions have been satisfied. On top of that, a theoretical distribution has been fitted to both failure modes of the scrap cutter. These distributions will be used again later in the report to determine the costs of different maintenance strategies.
10 Optimizing operating costs

In this chapter, the formulas to calculate the expected operating costs (as defined in the first part of this report) of three different maintenance scenarios for the scrap cutter will be provided:

1. Corrective maintenance (current situation)
2. Preventive maintenance (current situation, so considering both failure modes of the scrap cutter)
3. Preventive maintenance (only considering the failure mode dull blades)

The expected costs of the corrective maintenance policy will be calculated to have a baseline scenario. The other scenarios can be compared to this scenario to test if preventive maintenance provides a cost benefit and, if so, quantify the expected benefit.

Scenario 2 will provide the optimal preventive maintenance interval for the current situation. Next to the optimal replacement interval, the expected costs of replacing the scrap cutter after processing 2000 kilometres will be calculated (the preventive plan which is currently under consideration).

The third scenario, preventive maintenance only considering the failure mode dull blades, has two uses. First of all, it has a theoretical use, by providing the method to calculate preventive maintenance costs in a situation with only one failure mode. On top of that, the scenario can generate some practical insight in how much the second failure mode influences the operating costs. If the maintenance of the scrap cutters could be improved in such a way that the other failures are no longer occurring (or at least less frequently), the optimal preventive maintenance interval would move towards the outcome of this scenario. Therefore, the difference between scenarios 2 and 3 will show the possible benefits of eliminating burn-in failures.

The operating costs will be approximated using renewal theory. In the context of this research, a renewal is equal to each time a system is repaired or replaced. By estimating the expected time between two renewals and the average costs for each renewal, the expected costs per time unit can be calculated. More information on renewal processes and their application in maintenance modelling can be found in Chapter 5 in Kulkarni (2010).

It should be noted that all formulas proposed in this chapter are based on the assumption that the two failure modes are independent. In the case of the scrap cutter, the causes of the two failure modes are very different. First of all, dull blades and other failures are based on different components of the scrap cutter (blades versus other components). Chapter 9 has also shown that the two failure modes show completely different patterns. The other failures usually occur early, while the blades slowly seem to wear down, resulting in a steadily increasing failure rate.

Furthermore, the notation and formulas in this chapter will use kilometres to measure the usage of the system (in this case the scrap cutter). However, it is possible that time would be a suitable way to measure usage for other systems (this corresponds to the time-based maintenance strategy of Figure 33). In this case, the kilometres can be swapped out by any measure of time (e.g. minutes, hours) and the formulas will still hold, as long as the same unit of measurement is used throughout all formulas.
10.1 Notation
First of all, the notation that will be used in this chapter will be covered:

\( C_{cm} = \text{Cost of a corrective maintenance action} \)

\( C_{pm} = \text{Cost of a preventive maintenance action} \)

\( ECC = \text{Expected Cycle Costs (one cycle is the period between two renewals)} \)

\( ECL = \text{Expected Cycle Length in kilometres} \)

\( f_i(.) = \text{Probability density function of failure mode } i \ (i = 1, 2) \)

\( F_i(.) = \text{Cumulative density function of failure mode } i \ (i = 1, 2) \)

\( g = \text{Expected costs per kilometre} \)

\( R_i(.) = \text{Reliability function (also called survival function) of failure mode } i \)

\( T = \text{Amount of kilometres processed until failure of a component} \)

\( \tau = \text{Preventive maintenance interval } (\tau > 0) \text{ in kilometres} \)

The relation between \( f_i(.) \), \( F_i(.) \) and \( R_i(.) \) is defined as follows:

\( F_i(t) \) gives the probability that failure mode \( i \) has occurred before or during kilometre \( t \) and is defined as:

\[
F_i(x) = \int_0^x f_i(t) \, dt \tag{10.1}
\]

\( R_i(t) \) gives the probability that failure mode \( i \) has not occurred up until kilometre \( t \) and is defined as:

\[
R_i(x) = \int_x^\infty f_i(t) \, dt \tag{10.2}
\]

From this, it follows that \( F_i(x) + R_i(x) = 1 \).

10.2 Scenario 1: Corrective maintenance costs with two failure modes
The expected costs of a maintenance strategy can be defined in the following way:

\[
g = \frac{ECC}{ECL} \tag{10.3}
\]

The expected costs of one cycle divided by the expected length of one cycle results in the expected costs per kilometre. This formula will form the basis for the cost calculations of all three scenarios.

In case of a corrective strategy, the ECC is equal to the corrective replacement costs of the scrap cutter. Each cycle both starts and ends with the replacement of the scrap cutter, which will always be a corrective replacement in this scenario. This formula only holds under the assumption that both failure modes lead to the same costs. This assumption holds for the scrap cutter.

\[
ECC = C_{cm} \tag{10.4}
\]
However, it is possible that the two corrective maintenance actions have different costs. Appendix J contains the ECC for this case (formula AJ.1). These formulas will not be used for the scrap cutter.

The ECL is equal to the expected time to failure, $E(T)$. This is the combined mean of the two failure modes: the average amount of kilometres after which one of the failures occurs. This can be calculated with the following formula:

$$E(T) = \left( \int_{t=0}^{t=\infty} \left( \int_{u=t}^{u=\infty} u \cdot f_1(u) \, du \right) f_2(t) \, dt \right) + \left( \int_{t=0}^{t=\infty} \left( \int_{u=0}^{u=t} f_1(u) \, du \right) t \cdot f_2(t) \, dt \right)$$  (10.5)

The left part of the formula gives the average time to failure for cases where failure 1 occurs, the right part the average time for the cases where failure 2 occurs. The derivation and more detailed explanation of this formula is provided in Appendix H.

10.3 Scenario 2: Preventive maintenance costs with two failure modes

For preventive maintenance, a decision variable $\tau$, the amount of kilometres after which a preventive maintenance action will be performed, is introduced. The smaller the value of $\tau$, the more often preventive maintenance actions will be performed. By choosing too low values of $\tau$, you risk replacing equipment that was still in good condition, which may lead to unnecessarily high maintenance costs. If the value of $\tau$ is chosen too high, the equipment may fail often before the planned moment for preventive maintenance, leading to a high number of corrective replacements. Therefore, the optimal value of $\tau$ is found by taking both risks into account.

The reliability of a component with $n$ independent failure modes is equal to the reliability of each failure mode:

$$R(t) = R_1(t) \cdot R_2(t) \ldots \cdot R_n(t) = \prod_{i=1}^{n} R_i(t)$$  (10.6)

Based on this property of independent reliability functions, the ECC can be determined:

$$ECC = \left( 1 - (R_1(\tau) \cdot R_2(\tau)) \right) \cdot C_{cm} + (R_1(\tau) \cdot R_2(\tau)) \cdot C_{pm}$$  (10.7)

$(R_1(\tau) \cdot R_2(\tau))$ is the probability that none of the failure modes will occur before the preventive maintenance action, while $(1 - (R_1(\tau) \cdot R_2(\tau)))$ is the probability that either failure mode will occur before the preventive maintenance moment. Again, these two probabilities add up to 1. Similarly to the first scenario, this formula only holds under the assumption that corrective and preventive maintenance costs are the same for both failure modes. The formula for different corrective maintenance costs can be found in Appendix J (formula AJ.2), but will not be used for the scrap cutter.

The ECL gets more complicated compared to the corrective strategy. There are five cases that need to be considered:

1. The probability that failure mode 2 happens before $\tau$, failure mode 1 happens before failure mode 2 (failure mode 1 occurs);
2. The probability that failure mode 2 happens before $\tau$, failure mode 1 happens after failure mode 2 but before $\tau$ (failure mode 2 occurs);
3. The probability that failure mode 2 happens before $\tau$, failure mode 1 happens after failure mode 2 and after $\tau$ (failure mode 2 occurs);
4. The probability that failure mode 2 happens after \( \tau \), failure mode 1 happens before \( \tau \) (failure mode 1 occurs);

5. The probability that failure mode 2 happens after \( \tau \), failure mode 1 happens after \( \tau \) (so preventive maintenance is performed).

These 5 probabilities are visualised in Figure 41. In case 5, it does not matter which failure mode occurs first, since the preventive maintenance moment \( \tau \) occurs before either failure mode. This is the only case out of the 5 in which preventive maintenance action will be performed, the other four cases are corrective replacements. By calculating the probability that each case occurs, the following formula for the ECL will be reached:

\[
ECL = \left( \int_{t=0}^{\tau} \left( \int_{u=0}^{\tau} u * f_1(u) \, du \right) f_2(t) \, dt \right) + \left( \int_{t=0}^{\tau} \left( \int_{u=0}^{\tau} f_1(u) \, du \right) t * f_2(t) \, dt \right) + \left( \int_{t=\tau}^{\infty} \left( \int_{u=0}^{\tau} u * f_1(u) \, du \right) f_2(t) \, dt \right) + \tau * R_1(\tau) * R_2(\tau)
\]

(10.8)

This formula gives the expected amount of kilometres until replacement for each case, multiplied by its probability of occurring. Appendix I provides the proof that the probabilities of the five cases add up to 1, as well as the derivation of this formula.

With the formulas for the ECC and ECL, the cost per kilometre, \( g(\tau) \), can be calculated.

\[
g(\tau) = \frac{ECC(\tau)}{ECL(\tau)}
\]

(10.9)

To minimize \( g(\tau) \), its derivative with respect to \( \tau \) should be set equal to zero. By inserting the resulting optimal value of \( \tau \) in \( g(\tau) \), the costs per kilometre for the optimal preventive maintenance interval are obtained.
10.4 Scenario 3: Preventive maintenance costs with one failure mode

The procedure to calculate the optimal preventive maintenance interval for this scenario can be seen as a simplified version of the second scenario. The expected cycle costs can be calculated in the following way:

\[ ECC(\tau) = F(\tau) * C_{cm} + R(\tau) * C_{pm} \]  \hspace{1cm} (10.10)

This formula uses the probability that the component will fail before the preventive maintenance action is performed, \( F(\tau) \). If it fails before the preventive maintenance, the corrective maintenance costs will be incurred. If it survives until the scheduled maintenance, \( R(\tau) \), the preventive maintenance costs are incurred. Since \( F(\tau) + R(\tau) = 1 \), the \( ECC(\tau) \) can only take on values between \( C_{cm} \) and \( C_{pm} \).

Each cycle is ended by a failure (after \( T \) kilometres) or a preventive replacement (at kilometre \( \tau \)), depending on which of the two possibilities occurs first. This leads to the following expected cycle length:

\[ ECL(\tau) = \mathbb{E}[\min(T, \tau)] \]  \hspace{1cm} (10.11)

This expected value of the minimum of \( T \) and \( \tau \) can be calculated with the following formula:

\[ \tau * R(\tau) + \int_0^\tau (x * f(x)) \, dx \]  \hspace{1cm} (10.12)

The first part, \( \tau * R(\tau) \), is the probability that the component will be replaced preventively (so it did not fail before kilometre \( \tau \)), multiplied by the amount of kilometres until replacement \( \tau \). \( x * f(x) \) is the probability that a failure will happen at kilometre \( x \), multiplied by the amount of kilometres until that failure, \( x \). The integral of this function from 0 to \( \tau \) will result in the expected cycle length for all cases where the component failed before the preventive maintenance. Adding up the two parts results in the ECL.

The value of \( \tau \) that leads to the lowest cost per kilometre, \( g(\tau) \), can be calculated in the same way as for the second scenario, formula 10.9. The derivative of \( g(\tau) \) with respect to \( \tau \) should be set equal to zero. By inserting the resulting optimal value of \( \tau \) in \( g(\tau) \), the costs per time unit for the optimal preventive maintenance interval are obtained.
11 Operating costs of the scrap cutter

The formulas presented in the previous chapter, combined with the fitted probability density functions for the two failure modes found in Chapter 9, can be used to determine the expected costs for each of the three scenarios.

First, all the necessary parameter values for the scrap cutter need to be determined. Using these values, the expected operating costs for each scenario will be calculated. On top of that, the input values will be varied to determine how much effect this has on the results. Based on these results, an assessment can be made on how robust the results are.

\( C_{cm} \) : Costs of a corrective maintenance action

These costs consist of two parts for the scrap cutter. Downtime costs and maintenance costs. Its downtime costs are determined by the amount of time that the pickling line needs to be stopped. TSP counts €2.390 for each hour of downtime for the pickling line. Each corrective scrap cutter replacement requires the pickling line to stop for about 20 minutes. This is equal to €800 of downtime costs.

On top of that, the scrap cutter will need to be repaired after replacement. These repairs are performed at the technical service department and charged to TSP per quarter. Due to this way of working and the lack of a detailed specification, it is hard to determine the exact costs of revision of one scrap cutter. The cost estimate that is being used by TSP is €1.000 per scrap cutter. This is based on the total repair costs divided by the number of scrap cutters. These costs also include the labour costs for the repairs. Together with the downtime costs of €800 this makes for €1.800 per corrective maintenance action.

\( C_{pm} \) : Costs of a preventive maintenance action

By performing the same maintenance action preventively, it can be fit within an existing stop. In this way, the downtime costs of €800 are avoided. The costs of revision of a preventively replaced scrap cutter are not different from those of a correctly replaced (failed) scrap cutter. Therefore, each preventive maintenance action costs €1.000.

The two failure modes of the scrap cutter are both modelled by a Gamma distribution, as covered in Chapter 9. This leads to the following specification of \( f_1 \) (failure mode dull blades) and \( f_2 \) (failure mode other failures):

\[
\begin{align*}
  f_1(u) &\sim \text{Gamma}(761.46; 2.5174) \\
  f_2(t) &\sim \text{Gamma}(13329; 0.5457).
\end{align*}
\]

Now, the costs of the three scenarios can be calculated using the formulas of Chapter 10. In case of the preventive scenarios (2 and 3), the optimal preventive maintenance interval (\( r \)) will be determined as well. The calculations will be performed with Matlab using the script in Appendix K. This script is only an implementation of the formulas introduced in the previous chapter. By adjusting the values of the parameters in the script, it is possible to use this script for other components as well. Appendix L provides a basic verification of the results of the script to ensure that the results are correct.
11.1 Corrective maintenance, two failure modes

\[ ECC = C_{cm} = €1800 \]

\[ ECL = \left( \int_{t=0}^{t=\infty} \left( \int_{u=0}^{u=t} u \cdot f_1(u) \, du \right) f_2(t) \, dt \right) + \left( \int_{t=0}^{t=\infty} \left( \int_{u=t}^{u=\infty} f_1(u) \, du \right) t \cdot f_2(t) \, dt \right) = 1381.8 \text{ km} \]

\[ g = \frac{1800}{1381.8} \approx 1.30 \text{ €/km}. \]

Using the formulas of Chapter 9, the expected costs of a corrective maintenance strategy for the scrap cutter can be calculated. Over the last four years, the scrap cutter has processed approximately 23,250 kilometres of scrap on average per year. At any moment, there are two scrap cutters in use, so the costs need to be multiplied by two. This leads to the following expected costs per year under a corrective strategy:

\[ 2 \times 1.30 \times 23,250 \approx €60,574 \text{ per year}. \]

11.2 Preventive maintenance, two failure modes

The influence of the preventive maintenance interval \( \tau \) on the costs is found by filling in increasing values of \( \tau \) in the formulas from the previous chapter. The results are presented in Figure 42 for values of \( \tau \) of up to 5000 km by the blue line. As \( \tau \) moves towards infinity, the costs per kilometre descend down to approximately 1.30 €/km. These costs are equal to those of scenario 1, so these calculations are consistent with each other. From this graph, it can be concluded that the preferred maintenance strategy is corrective.

![Figure 42: Relationship between \( \tau \) and expected costs per km for scenario 2](image)
Next to the current situation, two other situations have been looked at. The first situation which has been considered is if the repair costs for all maintenance could be reduced down to €750. This is represented by the orange line in Figure 42. In this situation, the optimal maintenance strategy would still be corrective. Even if only preventive maintenance would lead to cheaper repairs (€750 per scrap cutter instead of €1,000), the optimal maintenance strategy would be corrective, with the same costs as in the current situation. This is shown by the grey line.

In Chapter 9, it has already been mentioned that the failure mode *other failures* will decrease the benefits of a preventive maintenance strategy due to its decreasing failure rate. On top of that, the difference in costs of a preventive and a corrective maintenance action is fairly small. Therefore, it is not surprising that corrective maintenance is preferred in the current situation.

Table 24: Results for scenario 2 for different values of $\tau$

<table>
<thead>
<tr>
<th></th>
<th>$\tau = 1000$</th>
<th>$\tau = 2000$</th>
<th>$\tau = 3000$</th>
<th>$\tau = \infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECC</td>
<td>€ 1353</td>
<td>€ 1606</td>
<td>€ 1729</td>
<td>€ 1800</td>
</tr>
<tr>
<td>ECL</td>
<td>761 km</td>
<td>1147 km</td>
<td>1301 km</td>
<td>1382 km</td>
</tr>
<tr>
<td>$g(\tau)$</td>
<td>1,778 €/km</td>
<td>1,401 €/km</td>
<td>1,328 €/km</td>
<td>1,303 €/km</td>
</tr>
<tr>
<td>$R(\tau)$</td>
<td>55,8 %</td>
<td>24,2 %</td>
<td>8,9%</td>
<td>0 %</td>
</tr>
</tbody>
</table>

Table 24 shows that both the expected cycle costs and expected cycle length increase as $\tau$ becomes higher. Under the current costs of preventive and corrective maintenance, the benefits of preventive maintenance (lower cycle costs) do not outweigh its drawback (more replacements). The expected costs per year for the current situation are:

$$2 \times 1,303 \times 23.250 \approx € 60,574 \text{ per year}.$$  

It should be noted that the plan to change scrap cutters after every 2000 kilometres would lead to higher costs: $2 \times 1,401 \times 23.250 = € 65,146 \text{ per year}.$

11.3 Preventive maintenance, failure mode *dull blades only*

The costs of the third scenario, preventive maintenance with only the failure mode *dull blades*, have been analysed similarly to the second scenario. The costs for different values of $\tau$ have been plotted and are presented in Figure 43. Besides the current situation, the situations with lower maintenance costs and lower preventive maintenance costs have been plotted again (similarly to the previous section). Even though the failure mode with the decreasing failure rate (*other failures*) has been removed, the lowest costs for this scenario are still achieved with a corrective maintenance strategy.

If the cost of both preventive and corrective maintenance is reduced, the optimal replacement moment moves to approximately 4200 kilometres (the location of the marker on the orange line). In the case where only the costs of a preventive maintenance action become cheaper, the optimal moment for preventive maintenance is after 2700 kilometres of production. However, in both situations, the costs do not rise a lot after the optimal replacement moment, so the cost benefit of preventive maintenance would still be limited.
Table 25 provides some more information on the scrap cutter for different values of \( \tau \). This also helps to explain why preventive maintenance is more expensive than corrective maintenance in the current situation. The difference in the expected cycle costs is relatively small for varying values of \( \tau \). If the difference between the costs of a preventive maintenance action and a corrective maintenance action would have been bigger, preventive maintenance would be optimal. The grey line in Figure 43 showed this as well.

This leads to the following estimate of the yearly costs for scenario 3:

\[
2 \times 0.939 \times 23.250 \approx \text{€} \, 43.664 \text{ per year}.
\]

Even though the recommended maintenance strategy would still be corrective in this scenario, the yearly costs are considerably lower than those of scenarios 1 and 2. If it were possible to completely eliminate the second failure mode (other failures), this could save \( 60.574 - 43.664 = \text{€} \, 16.910 \) each year.

However, it might be difficult to eliminate this failure mode completely. A more realistic goal in practice would therefore be to reduce the probability of occurrence of the second failure mode. It is possible that the biggest cost savings are already reached after a small reduction of this probability, in which case it would not make sense to put a lot of effort into completely eliminating the failure.
mode. Similarly, it is possible that the biggest cost savings are only reached when the probability of the failure mode is almost reduced to zero.

To gain some insight in the relation between the probability of occurrence of the second failure mode and the cost savings, its probability distribution determined in Chapter 9 needs to be changed. The scale parameter of the Gamma distribution can be increased to stretch it out over the x-axis. Figure 44 shows two Gamma distributions with the same value for the shape parameter, but different values for the scale parameters. The distribution of the blue line has a much larger scale, which leads to a much smaller probability of failure per kilometre. The other parameter of the Gamma function (the shape parameter) should be kept constant, as modifying its values would alter the shape.

![Figure 44: Effect of increasing the scale parameter of a probability distribution](image)

By manipulating the values of the scale parameter of the second failure mode and using these with the formulas of scenario 2, the relation between the probability of occurrence of the second failure mode and the cost savings can be found. The values for the scale parameter are chosen such that the failure rate will be either 20, 40, 60 or 80 percent of the original failure rate. The result is shown in Figure 45.
Figure 45: Cost savings achieved by reducing failure mode other failures

This graph shows that the relation between the reduction of failure mode other failures and the cost savings is close to linear. However, the first 20 percent of reduction of failure mode other failures leads to a slightly higher cost reduction relative to the last 20 percent, with approximately €4,000 and €3,000 cost savings respectively. Based on this result, it is beneficial to reduce the probability of occurrence of other failures. Even a small reduction would lead to lower operating costs for the scrap cutter.

11.4 Conclusion

Based on the results of this chapter, it can be concluded that corrective maintenance is the preferred maintenance strategy for the scrap cutter. However, if the failure mode other failures can be reduced, this could lead to costs savings of around € 17,000 per year.
12 Conclusion and recommendations

12.1 Conclusions

This report has been divided into two parts: prediction of failures and maintenance optimization. The main conclusions of each part will be presented.

Part one: prediction of failures

The pickling line of TSP has over 100 sections, but only five of these sections cause the majority of the operating costs. This shows that it is important to determine where to focus the research on. Using the tool presented in Chapter 4, it has been shown that the side trimmer section has the highest operating costs of the pickling line.

The data collection part (Chapter 6) has shown the importance of a high data quality. Although there is a lot of data available on production and production stops, it is not suitable for failure analysis in many cases. The side trimmer blades are a prime example of this problem, as their data was of insufficient quality to analyse their failures. A second issue is that the necessary data is spread out across different systems, which makes it much harder to find and analyse the right data. These issues will be addressed in more detail in the recommendations.

Finally, the research has shown that there is no relation between the failure predictors and the failure behaviour of the scrap cutter.

On top of that, the relation between the age of a scrap cutter and the occurrence of scrap getting stuck in the scrap cutter has been investigated. No clear relation between these two failure modes could be found either.

Part two: maintenance optimization

The analysis of the failures of the scrap cutters has led to the identification of two separate failure modes: failures of its blades and other failures. Other failures mostly occur shortly after a new scrap cutter is put to use. These problems are therefore likely to be caused by poor maintenance or installation of the scrap cutter. Reducing the other failures will decrease the yearly operating costs of the scrap cutter. Elimination of this failure mode would reduce the operating costs of the scrap cutter by €17,000 each year.

The current strategy, corrective maintenance, is optimal in the current situation. The benefits of preventive maintenance for the scrap cutter (no downtime costs) do not outweigh its drawbacks (replacing the scrap cutter more often). Implementing a preventive maintenance strategy would lead to higher costs. The plan to replace the scrap cutter every 2000 kilometres, which is currently under consideration, would lead to higher operating costs of the scrap cutter of €4,500 per year.

12.2 Limitations

The replacements of the side trimmer blades are not documented consistently and precisely. This has made it impossible to analyse the failure behaviour of the side trimmer blades. The relation between their failure behaviour and production has not been investigated. The maintenance of the side trimmer blades could not be optimized either. When data is available on the exact failure times of each blade, it will be possible to perform these analyses using the methods used in the report. This would also allow for testing the relation between the condition of a blade and scrap getting stuck in the scrap cutter.
For prediction of failure behaviour, research has focused only on finding possible predictors and testing their influence. For the scrap cutter, there is no relation between the predictors and failure behaviour, so the research has not continued past this point. If a component can be found for which there is a relation between production (predictors) and failure behaviour, the exact effect of each predictor still needs to be determined. A possible way to determine these effects would be to use proportional hazards modelling, first introduced by Cox (1972).

Finally, condition monitoring has not been covered in this research. No condition data was available for the side trimmer or the scrap cutter blades, so this was not possible. Since no relation could be found between the predictors and failure behaviour of the scrap cutter, it could be interesting to focus on condition monitoring in future research. This could also provide opportunities for condition based maintenance.

12.3 Recommendations for TSP

12.3.1 Data quality

Data uniformity
In SAP, the functional location structure is used (see Section 3.2). The logbook uses this structure as well, but only uses the names of the locations instead of the location code (for example “welding machine” instead of “150-02-02-02”). The section names in the logbook are not identical to the names in SAP, for example due to extra spaces. This creates extra difficulties when comparing logbook data to SAP data. The codes are shorter, unambiguous and known for all production lines. Using the location codes instead of the section names in the logbook data would allow for easier use of its data.

This is an example of two databases that use the same structure, but different naming standards. Data on maintenance and production is spread out over several systems at TSP. In order to link or compare these systems, it is important that such different standards are avoided as much as possible. This would make it a lot easier to perform a data analysis of data coming from multiple systems.

Data uniqueness
Material identifiers (the coil numbers) and customer order numbers are not unique, even though they should be. For example, a customer order number consists of 5 digits. The first digit needs to be a 6 or an 8 for TSP orders. Therefore, there are only 20,000 order numbers that can be used by TSP. Due to the high number of orders each year, it is inevitable that one number must be used for different orders. Every two years, the same order numbers are being used for new orders.

For this research, the existence of double identifiers (IDs) in data made it a lot harder and more time-consuming to link data in a reliable way (see Sections 6.1.1 and 6.2). Therefore, it is recommended to ensure that any identifying numbers are always unique. First of all, the amount of digits in each number should be increased. For the customer order numbers, adding two digits would already solve the problem. On top of that, the databases should not allow users to link a coil or customer order to an ID which has been used before.
Data completeness

To perform any analysis on the failure behaviour of components, the available data should provide the answers to two vital questions for all replacements: *when* the component has been replaced and *which* component has been replaced. If any of these questions cannot be answered, it is unlikely that any analysis will be possible. This leads to the following two recommendations for TSP.

There is very little reliable data on what happens exactly during planned maintenance. When components are replaced during a planned stop, this is very difficult to track down in the data. When investigating a large period of time, this makes it impossible to determine *when* components have failed or have been replaced. The scrap cutters are an exception since each replacement is tracked separately in a list. Without this list, it would not have been possible to trace back all replacements. This would have caused an underestimation of the number of failures during the investigated period, leading to an unreliable analysis. Therefore, it is recommended to keep better track of all replacements, both planned and unplanned. This would increase the overall quality of the data. A possible way to do this is by keeping track of the replacements separately, similarly to the scrap cutter replacement list. However, since this requires a lot of time and effort, the preferred way would be to include this information in an existing database such as the notifications in SAP or the production stop part of the logbook.

In case of the side trimmer blades, it is often not clear *which* of the four blades has or have been changed. Without specifying this, it is impossible to determine the time to failure of each individual blade. By creating a standard way of documenting these replacements, the data can be made much easier to process and much more valuable for analyses. This could be achieved by creating a new data column in the logbook to store this information.

12.3.2 Scrap cutter repairs

The costs of maintenance of the scrap cutter are hard to determine. These costs are charged by the technical service department to TSP once every month or quarter, without a clear breakdown into individual repairs. This makes it hard to control what has been done and how much scrap cutters have been repaired. Charging each repair separately would make the maintenance documentation a lot more transparent. In case of any problems with the quality of a scrap cutter, it could be traced back to one specific repair, making it easier to detect what went wrong. This also applies to the maintenance costs of the side trimmer blades, which are charged quarterly without a proper specification.

This recommendation is also connected to the analysis of the failure behaviour of the scrap cutters. Some scrap cutters fail after a very short amount of time. These failures should be investigated and eliminated as much as possible, as indicated in the report. However, without a clear specification of what has been done during the repair of each individual scrap cutter, it will be hard to find the cause of these issues.

12.3.3 General recommendations

With the data that is currently available, corrective maintenance is the best option for the scrap cutters. Since its failure behaviour cannot be predicted using production data, it is recommended to change the focus to condition monitoring. Possible suggestions to determine the condition of a blade would be to keep track of the cutting sounds, vibrations or the quality of the scrap output. Using any of these methods, it might be possible to determine when a blade is close to failure and schedule
maintenance accordingly to decrease the operating costs of the scrap cutter. However, more research will be necessary to confirm if either of these condition monitoring methods is applicable to the scrap cutters.

For this research, the pickling line has been investigated. Although this production line is important for the continuity of coil production of TSP, it is not the bottleneck in the production process. It has more capacity than what is needed, so when it breaks down, it will not lead to production losses right away. For future research within TSP, it can be interesting to apply the method for maintenance optimization on a bottleneck production line. This would increase the costs of downtime due to lost production. In such a situation, a preventive maintenance policy is more likely to outperform a corrective maintenance policy in terms of costs.

12.4 Academic relevance
Most literature on predictive maintenance is closely related to condition-based maintenance. The researchers try to predict the moment of failure using measurements of the condition of the system. However, condition monitoring data is not always available. By using production-based predictors such as the characteristics of the produced material (e.g. resistance, gauge), this research offers a different method to predict the moment of failure of a system. Furthermore, this research does not start with a predefined set of possible predictors. Instead, it offers a methodology to find and select the possible predictors.
13 References


14 Appendices

Appendix A: Definitions and abbreviations
Appendix B: Data columns of the logbook
Appendix C: Linking different data sources
Appendix D: T-test for comparison of means
Appendix E: Scrap getting stuck in the scrap cutter
Appendix F: Alternative method for testing predictor influence on failure
Appendix G: Matlab code for distribution fitting
Appendix H: Derivation and proof of correctness of formula 10.5
Appendix I: Derivation and proof of correctness of formula 10.8
Appendix J: ECCs for different corrective maintenance costs of failure modes
Appendix K: Matlab code for optimization of operating costs
Appendix L: Verification of Matlab code for optimization of operating costs
Appendix A: Definitions and abbreviations
When a definition refers to another term, it is printed in bold.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burn-in period</td>
<td>Period in the early life of new components during which a lot of failures occur due to manufacturing defects</td>
</tr>
<tr>
<td>Censored data</td>
<td>Incomplete data. In case of failure data, this can happen when a component is replaced before the failure occurred.</td>
</tr>
<tr>
<td>Coil</td>
<td>Rolled up sheet metal, this is done to transport the metal between production lines or to customers.</td>
</tr>
<tr>
<td>Condition-based maintenance</td>
<td>A maintenance strategy where the condition of a component is monitored (through either sensors or inspections) to be able to notice an upcoming failure and act accordingly (through repair or replacement of the component).</td>
</tr>
<tr>
<td>Corrective maintenance</td>
<td>A maintenance strategy where a component will only be repaired or replaced upon or after failure.</td>
</tr>
<tr>
<td>Downtime</td>
<td>The time during which a production line is not in operation. TSP distinguishes three types of downtime: Technical downtime, production-related downtime and process-/product-related downtime.</td>
</tr>
<tr>
<td>Failure</td>
<td>When a production line stops working or needs to be stopped because it is no longer capable of producing products that comply to set standards.</td>
</tr>
<tr>
<td>Failure mode</td>
<td>Way in which a machine can fail</td>
</tr>
<tr>
<td>Failure rate (also called hazard rate)</td>
<td>The probability that a component will fail at a moment in time, given that the component has survived until that point</td>
</tr>
<tr>
<td>Hazard rate</td>
<td>See Failure rate</td>
</tr>
<tr>
<td>Maintenance costs</td>
<td>Labour costs for repairs, replacement, inspection or upkeep of a machine part and costs for used materials (e.g. replacement parts).</td>
</tr>
<tr>
<td>Material gauge</td>
<td>The thickness of the metal strip</td>
</tr>
<tr>
<td>Notch</td>
<td>A part cut out of the side of the metal strip. For side trimmers, a notch is used to create a small part in the strip that does not go through the side trimmers. This can be used to change the width at which the side trimmers cut.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Operating costs</td>
<td>The sum of <strong>downtime</strong> costs and <strong>maintenance costs</strong>.</td>
</tr>
<tr>
<td>Operational parameter</td>
<td>Parameter that is specifically linked to operational choices: how the machine is used, for what it is used and by whom.</td>
</tr>
<tr>
<td>Pickling</td>
<td>The process of removing the layer of oxides from steel that form during the steel making process (detailed description in chapter 2).</td>
</tr>
<tr>
<td>Predictor</td>
<td>Parameter that can be used to predict how long a machine can run until <strong>failure</strong> occurs. For example, the hardness of a metal strip could be a predictor for the shear that needs to cut it.</td>
</tr>
<tr>
<td>Preventive maintenance</td>
<td>A maintenance strategy where maintenance on a component is performed according to a planning with the aim of replacing or repairing it before failure occurs. This can be split up in <strong>time-</strong> or <strong>usage-based maintenance</strong>.</td>
</tr>
<tr>
<td>Process-/Product-related downtime</td>
<td><strong>Downtime</strong> caused by problems with the process and/or product, such as the product getting stuck.</td>
</tr>
<tr>
<td>Production line</td>
<td>A line of connected machinery with one common goal. In general, if one of these machines stops, the entire production line needs to be stopped. The <strong>pickling</strong> line is an example of this.</td>
</tr>
<tr>
<td>Production-related downtime</td>
<td><strong>Downtime</strong> caused by a lack of production resources (lack of available crew or <strong>coils</strong> to process)</td>
</tr>
<tr>
<td>Response time</td>
<td>The time it takes until people and materials are available to start repairs in case of failure of a <strong>production line</strong>.</td>
</tr>
<tr>
<td>Section</td>
<td>Part of a <strong>production line</strong> with one specific task, e.g. welding.</td>
</tr>
<tr>
<td>Technical downtime</td>
<td><strong>Downtime</strong> due to a <strong>failure</strong> or repair on the <strong>production line</strong>. This downtime is recorded from the moment the production stops until it is resumed, so this includes possible time lost due to cooling down/setup times of a machine or <strong>response time</strong>.</td>
</tr>
<tr>
<td>Technical failure</td>
<td>A production stop caused by machine failure.</td>
</tr>
<tr>
<td>Time-based maintenance</td>
<td>A type of <strong>preventive maintenance</strong> where maintenance actions are scheduled according to the time since the last maintenance action or failure (whichever occurred last).</td>
</tr>
<tr>
<td>Usage-based maintenance</td>
<td>A type of <strong>preventive maintenance</strong> where maintenance actions are scheduled according to the amount it has been used since the last maintenance action or failure (whichever occurred last). Usage can be expressed in operating hours or production for example.</td>
</tr>
<tr>
<td><strong>Abbreviation</strong></td>
<td><strong>Definition</strong></td>
</tr>
<tr>
<td>------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>AIC</td>
<td>Akaike Information Criterion</td>
</tr>
<tr>
<td>CBM</td>
<td>Condition-based maintenance</td>
</tr>
<tr>
<td>MLE</td>
<td>Maximum Likelihood Estimator</td>
</tr>
<tr>
<td>RUL</td>
<td>Remaining Useful Life</td>
</tr>
<tr>
<td>TSP</td>
<td>Tata Steel Packaging</td>
</tr>
</tbody>
</table>
## Appendix B: Data columns of the logbook

### Production part of the logbook

<table>
<thead>
<tr>
<th>Column name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Materiaal ID Moederrol</td>
<td>Coil number of the entering coil, originating from the warm rolling plant</td>
</tr>
<tr>
<td>Materiaal ID</td>
<td>Coil number used by TSP (every coil produced by the pickling line receives a new number, when a received coil is cut up into smaller pieces, each of these gets a separate number)</td>
</tr>
<tr>
<td>HO Jaar Kwartaal</td>
<td>Year + quarter</td>
</tr>
<tr>
<td>HO Jaar Week</td>
<td>Year + week number</td>
</tr>
<tr>
<td>HO Jaar Maand Dag</td>
<td>Date of production according to a shift day (starts at 06:00)</td>
</tr>
<tr>
<td>Wachtdatum En Nummer</td>
<td>Date including the number of the shift (1, 2 or 3)</td>
</tr>
<tr>
<td>Ploegkleur</td>
<td>Colour identifying the crew of the shift (Blue, red, yellow, white or green)</td>
</tr>
<tr>
<td>Kantschaar Indicator</td>
<td>Indicates whether the coil is processed by the side trimmer</td>
</tr>
<tr>
<td>KIM waarde In</td>
<td>Weight of the coil divided by its width at entrance</td>
</tr>
<tr>
<td>KIM waarde Uit</td>
<td>Weight of the coil divided by its width after exit</td>
</tr>
<tr>
<td>KKR Eind Datetime</td>
<td>End time of coil processing</td>
</tr>
<tr>
<td>Klant code</td>
<td>Customer name</td>
</tr>
<tr>
<td>Badhev Code</td>
<td>Code describing the desired end product</td>
</tr>
<tr>
<td>Omschrijving Badhev</td>
<td>Explanation of the code</td>
</tr>
<tr>
<td>Omschrijving Bekleding</td>
<td>Desired coating of the final product (usually tinned in case of TSP)</td>
</tr>
<tr>
<td>Order + Postletter</td>
<td>Customer order number + letter to indicate different batches</td>
</tr>
<tr>
<td>Ordemnummer</td>
<td>Customer order number</td>
</tr>
<tr>
<td>Actuele dikte Uit in mm</td>
<td>Gauge of the strip at exit</td>
</tr>
<tr>
<td>Actuele breedte In mm detail</td>
<td>Coil width at line entry</td>
</tr>
<tr>
<td>Actuele breedte Uit mm detail</td>
<td>Coil width after exit</td>
</tr>
<tr>
<td>Inzet Bruto</td>
<td>Weight of the coil in kilos at entry</td>
</tr>
<tr>
<td>Inzet Bruto lengte</td>
<td>Length of the coil in meters at entry</td>
</tr>
<tr>
<td>Opbrengst Directe Lijn</td>
<td>Weight of the coil in kilos after exit</td>
</tr>
<tr>
<td>Opbrengst Directe Lijn lengte</td>
<td>Length of the coil in meters after exit</td>
</tr>
</tbody>
</table>

### Production stops part of the logbook

<table>
<thead>
<tr>
<th>Column name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wachtdatum</td>
<td>Date of production according to a shift day (starts at 06:00)</td>
</tr>
<tr>
<td>Wachtdatum En Nummer</td>
<td>Date including the number of the shift (1, 2 or 3)</td>
</tr>
<tr>
<td>HO Jaar Kwartaal</td>
<td>Year + Quarter of stop</td>
</tr>
<tr>
<td>HO Jaar Week</td>
<td>Year + Week of stop</td>
</tr>
<tr>
<td>Ploegkleur</td>
<td>Colour identifying the crew of the shift (Blue, red, yellow, white or green)</td>
</tr>
<tr>
<td>Installatie Sectie Naam</td>
<td>Section name</td>
</tr>
<tr>
<td>Toelichting Stilstand</td>
<td>Description of the reason of the production stop</td>
</tr>
<tr>
<td>Duur (in Uren)</td>
<td>Duration of the stop in hours</td>
</tr>
<tr>
<td>Begintijd Stilstand</td>
<td>Starting time of the stop</td>
</tr>
<tr>
<td>Eindtijd Stilstand</td>
<td>Ending time of the stop</td>
</tr>
<tr>
<td>TIB Verlies Gepland</td>
<td>Indicates whether losses were planned or unplanned</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>TIB Categorie Omschrijving</td>
<td>Category of the failure (&quot;technical&quot;, &quot;side trimmer&quot;, scrap problems&quot;)</td>
</tr>
<tr>
<td>TIB Code</td>
<td>Failure code indicating what type of failure it is</td>
</tr>
<tr>
<td>TIB Code Omschrijving</td>
<td>Description of the TIB code (examples are &quot;electrical&quot;, &quot;mechanical&quot; or &quot;side trimmer blade&quot;)</td>
</tr>
<tr>
<td>TIB Code Type Omschrijving</td>
<td>Distinguishes between Technical failures, product-related failures and production-related failures</td>
</tr>
<tr>
<td>Stilstand Volgnummer</td>
<td>Stop number</td>
</tr>
<tr>
<td>Materiaal ID</td>
<td>Coil number (the “old” number of the warm rolled coil is still being used in this table)</td>
</tr>
<tr>
<td>Klant code</td>
<td>Customer name</td>
</tr>
<tr>
<td>Badhev Code</td>
<td>Code describing the desired end product</td>
</tr>
<tr>
<td>Badhev omschrijving</td>
<td>Explanation of the code</td>
</tr>
<tr>
<td>Order + Postletter</td>
<td>Customer order number + letter to indicate different batches</td>
</tr>
<tr>
<td>Ordernummer</td>
<td>Customer order number</td>
</tr>
<tr>
<td>Duur</td>
<td>Duration of stop in seconds</td>
</tr>
<tr>
<td>Ml Thickness</td>
<td>Gauge of the coil</td>
</tr>
<tr>
<td>Ml Width</td>
<td>Width of the coil at entry</td>
</tr>
<tr>
<td>Ml Cal Weight</td>
<td>Weight of the coil at entry</td>
</tr>
<tr>
<td>Stilstand uniek Id</td>
<td>Unique stop number</td>
</tr>
</tbody>
</table>
Appendix C: Linking different data sources

Linking resistance numbers to customer orders

Below is an example entry of the resistance number table, showing the important rows. First the recipe code and coiling temperature should be merged in a new column. For the example entry below, this would be “123A600” (Considering these cells would be located at cells A2 and B2, the Excel formula will be “=A2&B2”)

<table>
<thead>
<tr>
<th>Recipe code</th>
<th>Coiling temperature</th>
<th>Resistance number</th>
</tr>
</thead>
<tbody>
<tr>
<td>123A</td>
<td>600</td>
<td>5,5</td>
</tr>
</tbody>
</table>

The same thing should be done in the customer order table (this table contains the same two columns).

=VLOOKUP(“Cell of the customer order table which contains the Recipe code + coiling temperature”; “Table

Linking production (logbook) to production stops (logbook)

All the SQL queries used to link the production data with production stops data. They are presented in the order in which they need to be executed:

1. From the production data, remove all the cases where the side trimmer section was not used and empty rows:
   SELECT * INTO Opbrengsten_Mrol
   FROM Opbrengsten
   WHERE Opbrengsten.[Materiaal ID Moederrol] IS NOT NULL
   AND Opbrengsten.[Kantschaar Indicator] = "J";

2. Create a unique ID for the failure data by creating a value which is a combination of the material ID and the production week:
   UPDATE Stilstanden
   SET ID = [Materiaal ID] & "-" & [HO Jaar Week] & "-1";

3. Create an ID for the production data in the same way:
   UPDATE Opbrengsten_Mrol
   SET ID = [Materiaal ID Moederrol] & "-" & [HO Jaar Week];

4. Of all rows in the production data with an ID that is not unique, select the first row:
   SELECT First(A.NR), A.ID, First(A.[Materiaal ID]), ... all other columns of the table ... First(A.[Opbrengst ongepland werk lengte]) INTO Opbr_dups
   FROM Opbrengsten_Mrol AS A
   INNER JOIN (SELECT ID, COUNT(ID)
   FROM Opbrengsten_Mrol
   GROUP BY ID
   HAVING (COUNT(ID)>1)) AS B
   ON A.ID = B.ID
   GROUP BY A.ID;
5. Remove these duplicate IDs from the production data:
SELECT Opbrengsten_Mrol.* INTO Opbrengsten_2 
FROM Opbrengsten_Mrol 
LEFT JOIN Opbr_dups 
ON Opbrengsten_Mrol.NR=Opbr_dups.NR 
WHERE Opbr_dups.NR IS NULL;

6. Select all other rows of the duplicates in the production data:
SELECT A.* INTO Opbr_dups2 
FROM Opbrengsten_2 AS A 
INNER JOIN Opbr_dups 
ON A.ID = Opbr_dups.ID;

7. Remove these rows from the production data as well:
SELECT Opbrengsten_2.* INTO Opbrengsten_3 
FROM Opbrengsten_2 
LEFT JOIN Opbr_dups2 
ON Opbrengsten_2.NR=Opbr_dups2.NR 
WHERE Opbr_dups2.NR IS NULL;

8. Add the selection of step 3 back to the production data (the result of this is the production data set with only the first occurrence of each ID):
INSERT INTO Opbrengsten_3 
SELECT * 
FROM Opbr_dups;

9. Add "-1" to the end of all IDs in this list
UPDATE Opbrengsten_3 
SET ID = ID & "-1";

10. Add all the remaining duplicate values back to the production data:
INSERT INTO Opbrengsten_3 
SELECT * 
FROM Opbr_dups2;

11 Merge the production and failure data based on the IDs
SELECT * INTO Opbr_Stilst 
FROM Opbrengsten_3 AS A 
LEFT JOIN Stilstanden AS B 
ON A.ID=B.ID;
Appendix D: T-test for comparison of means

The goal of this appendix is to provide some information on how to interpret the results of the t-test. More information on the t-test, including the formulas, can be found in Montgomery and Runger (2011). Below is the full output of the statistical program SPSS of the t-test for scrap cutters L1 and L2. The most important cells have been coloured grey. The results can be interpreted in two steps:

1. Check if the variances of the two samples are equal using Levene’s test (the left part of the table). This is necessary because the formula for the t-test is slightly different depending on the outcome. The null hypothesis of the test is that the variances of the two samples are equal. The null hypothesis is rejected when the test statistic is lower than 0,05. In Table 26, the value of the test statistic can be found under “Sig.” (significance). The value is 0,135. Therefore, the null hypothesis is not rejected and equal variances will be assumed.

2. Since Levene’s test showed that the variances can be assumed to be equal, the row “Equal variances assumed” should be used for the t-test. The means of the two samples are considered to be significantly different if the value is lower than 0,05. In this case, the test value is 0,960, so there is no significant difference between the mean time to failure of scrap cutters L1 and L2.

Some care should be taken when using the t-test. A significance value of 0,05 of the t-test in this context means that if you would take 100 samples of the two scrap cutters, you would find that the means of the two samples are different 5 times, even though they are not. Therefore, when performing a lot of t-tests, the risk of finding a “wrong” value increases: a 95% confidence interval over 10 tests results in a $1 - 0.95^{10} \approx 40\%$ probability that the test will imply a difference between two samples, even though there is none.

Table 26: Example output of a t-test for comparison of means of scrap cutters L1 and L2

<table>
<thead>
<tr>
<th></th>
<th>Levene’s Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td>Processed length until failure</td>
<td>Equal variances assumed</td>
<td>2,394</td>
</tr>
<tr>
<td></td>
<td>Equal variances not assumed</td>
<td>-054</td>
</tr>
</tbody>
</table>
Appendix E: Scrap getting stuck in the scrap cutter

This figure provides the rest of the data of Figure 32 (Figure 32 only showed the region from 0 to 10 in this graph in more detail). More information on this graph can be found in section 8.2 of the report.
Appendix F: Alternative method for testing predictor influence on failure

This section will provide an alternative method to test if there is a relation between predictors and the failure behaviour of a system. Although it is more complicated than the method described in Chapter 8, it has the benefit of being able to take all replacements into account. The method of Chapter 8 will only consider failures caused by the failure mode which is being tested, as mentioned at the start of Section 8.1. Preventive replacements and replacements due to other failure modes are left out of that analysis. Therefore, when a component is only replaced preventively, the method of Chapter 8 cannot be used. In such cases, the method of this section can be used instead.

The method in this section still uses the groups for the predictors which have been created in Chapter 7. The method will be demonstrated using the predictor material resistance for the scrap cutter. Table 27 shows three imaginary replacements, in the same layout as used in Table 18. For each replacement, the percentage of how much was processed of each resistance group is shown. For example, replacement 1 shows that the system was replaced after 100,000 meters of production. 31% of this produced strip (31,000 meters) belongs in resistance group 1.

<table>
<thead>
<tr>
<th>Replacement</th>
<th>Total processed amount until replacement (in m)</th>
<th>Group 1 (in %)</th>
<th>Group 2 (in %)</th>
<th>Group 3 (in %)</th>
<th>Group 4 (in %)</th>
<th>Group 5 (in %)</th>
<th>Group 6 (in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100,000</td>
<td>31</td>
<td>10</td>
<td>16</td>
<td>25</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>150,000</td>
<td>40</td>
<td>5</td>
<td>30</td>
<td>15</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>210,000</td>
<td>29</td>
<td>10</td>
<td>18</td>
<td>25</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

The first step of the method is to find replacements that have similar percentages of each (resistance) group. In Table 27, the replacements 1 and 3 have very similar percentages in each group, while the second observation has very different percentages. Based on this, there seem to be two different clusters.

This information can be used as input for the second step: a clustering algorithm. A clustering algorithm divides a set of observations into clusters with similar characteristics. The number of clusters should be defined by the user. In the example above, this would mean that replacement 1 and 3 are placed in one cluster and the second replacement is placed in a second cluster.

The final step is to compare the times to failure of the different clusters. The full method will be explained below in more detail.

1. Parallel plotting

For each scrap cutter, the total length it produced until replacement is known. On top of that, the material it has processed is divided into 6 resistance groups and a group for which the resistance is unknown. This data has been ordered like in Table 27. Each observation can be plotted as a line in a graph with the different groups on the x-axis and the percentage of each group on the y-axis. This creates the parallel plot of Figure 47. Each line is drawn transparently, so in places where a lot of lines come together, the blue colour will be darker.

The purpose of the graph is to identify certain clusters of lines that are very similar. In Figure 47, it is hard to clearly define certain clusters. For resistance groups 2, 5, 6 and the unknown group almost all lines are close to each other. For resistance group 1, the observations are scattered without any clearly definable groups. In resistance group 3, there seem to be two darker areas near 10 and 20 percent. Group 4 is more scattered, but lines that have a lower percentage of group 3 seem to have a
larger percentage of group 4 often. Based on this difference, two clusters will be considered in the next step. In general, the number of clusters should not be chosen too high. Picking a high number of clusters on a small data set will lead to small clusters, which leads to problems in the final step of this method. This will be discussed in that section of the method (Kaplan-Meier curves).

Furthermore, there are a few clear outliers observable. For example, three observations consist for 100% of group 3 or 6. These outliers are caused by scrap cutters that were replaced after only a few hours of production and can be left out in this case. Otherwise, there is a possibility that they will distort the clustering step.

2. K-means clustering

For the clustering step, the k-means clustering algorithm will be used. This algorithm is implemented in most statistical packages such as SPSS and Minitab. For this research, SPSS will be used to perform this analysis. As input, the k-means algorithm needs the number of desired clusters, \( k \), and a set of data points, which are the observations of the scrap cutter in this case. The data should be structured similarly to Table 27.

The output of the algorithm consists of \( k \) points, called centroids, that are the central point of each cluster. Each observation is assigned to the centroid which it is closest to. This leads to the parallel plot of Figure 48. This plot is the same as Figure 47, but the outliers are taken out and the two lines are coloured green and red, depending on the cluster they belong to. The green lines represent cluster 1 and the red lines represent cluster 2.

The centroids of the two clusters are presented in Table 28. All replacements have been placed in the cluster of the centroid they are closest to. This leads to 31 replacements in cluster 1 and 45 replacements in cluster 2.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Resistance group 1</th>
<th>Resistance group 2</th>
<th>Resistance group 3</th>
<th>Resistance group 4</th>
<th>Resistance group 5</th>
<th>Resistance group 6</th>
<th>Resistance group unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>28 %</td>
<td>2 %</td>
<td>14 %</td>
<td>39 %</td>
<td>2 %</td>
<td>6 %</td>
<td>9 %</td>
</tr>
<tr>
<td>2</td>
<td>15 %</td>
<td>4 %</td>
<td>22 %</td>
<td>46 %</td>
<td>2 %</td>
<td>8 %</td>
<td>2 %</td>
</tr>
</tbody>
</table>
3. Kaplan-Meier curves

To compare the clusters that have been determined in the previous step, Kaplan-Meier curves will be used. These curves are an estimate of the probability of survival of a component over time (or in this research, processed meters of strip). In Section 9.3, a similar procedure will be discussed in more detail. The plot of the two clusters has been created using SPSS and is shown in Figure 49. The vertical dashes on the green and red lines are the replacements of the scrap cutter due to other reasons than blade failures (these are called *censored data points*). These values were not considered in the analysis of Chapter 8, but are included in these plots.
Figure 49 shows that cluster 1 (the green line) is more likely to survive during the first 1,500,000 metres than cluster 2. However, after 1,500,000 metres, the probability of survival of both clusters is approximately 60%. After this, cluster 2 has a higher probability of survival. From this result, we can conclude that neither cluster has a higher chance of survival for the entire lifetime of a scrap cutter. Only if the line of one cluster would be consistently higher than other clusters, it would make sense to consider different maintenance schedules for the different clusters. Therefore, the results of this method are in line with the conclusion of the method used in Chapter 8.

Usage of the method
Some care should be taken when using the methods described above. Even though they are fairly easy to use, their results can be easily misinterpreted. For the clustering steps, it is important to check that the resulting clusters make sense. For example, performing the clustering step in the example above for 3 clusters would have led to one cluster consisting of less than 5 observations. This would be too little to draw any conclusions from the Kaplan-Meier plot. On top of that, one should check that a cluster does not only contain censored data. It is possible that the censored data points are similar and that the clustering step will put a lot of the censored data in the same cluster.
Appendix G: Matlab code for distribution fitting

1. Create an empirical failure distribution, fit theoretical distributions and test the fit

```matlab
% load the amount of km run until failure and whether or not it is censored
km = xlsread('E:/TSP/Data/Deelvraag 4/Testdata.xlsx','Incl_Censored_L','D2:D83');
censor = xlsread('E:/TSP/Data/Deelvraag 4/Testdata.xlsx','Incl_Censored_L','E2:E83');
% fit a Weibull distribution to the data according to the MLE
pd = wblfit(km,{},censor);
pd2 = lognfit(km,{},censor);
pd3 = gamfit(km,{},censor);
% Now some code to create a histogram which shows the distribution of
% failures over time
% n = the total number of observations (in this case n = 82)
% binWidth is determined by Sturges' rule:
% - Number of columns (called bins here) for the histogram = 1+log2(n)
%   (which evaluates to 7 for n = 82)
% - The maximum of variable "km" is approximately 3500, so seven groups
%   leaves groups of 500 each, which is the binWidth
n = length(km);
binWidth = 500;
% binCenter gives the middle value of each bin, which will be printed on the
% x-axis of the histogram
binCenter = 250:binWidth:3250;
% Hist_bars counts how many values fall in each bin of the histogram
% This value is then divided by n*binWidth to get the probability that a
% failure will happen for each single kilometre
% This probability is necessary to compare the histogram to the fitted
% distributions
Hist_bars = hist(km,binCenter);
Hist_bars = Hist_bars / (n*binWidth);
% Finally, the histogram itself is created and the color of the bars is
% adjusted to grey to make it a bit more clear
% Finally, the names of the axes and the range of the y-axis are defined
bar(binCenter,Hist_bars,'hist');
h_gca = gca;
h = h_gca.Children;
h.FaceColor = [.8 .8 .8];
xlabel('Amount of kilometres processed');
ylabel('Probability Density');
ylim([0 0.001]);

% Now, the Weibull- and Lognormal distribution are both plotted on the
% histogram to give a first impression of the quality of the fit.
xgrid = linspace(0,3420,1000);
pdfEst = wblpdf(xgrid,pd(1),pd(2));
line(xgrid,pdfEst,'Color','r')
pd2Est = lognpdf(xgrid,pd2(1),pd2(2));
line (xgrid,pd2Est,'Color','g')
pd3Est = gampdf(xgrid,pd3(1),pd3(2));
line (xgrid,pd3Est,'Color','k');

% Finally, a measure is printed to compare the quality of the two fitted
% distributions. This measure is the Akaike Information Criterion (AIC)
%AIC = 2*(nr of estimated parameters in the model)-2*ln (Log-likelihood)
% nr of estimated parameters = 1 for both weibull and lognormal function
% wbllike returns the negative log-likelihood, so the "-" turns to "+" in
```
% the formula
AIC_weib = 2+(2*(wbllike(pd,km,censor)));
AIC_logn = 2+(2*(lognlike(pd2,km,censor)));
AIC_gam = 2+(2*(gamlike(pd3,km,censor)));

2. Create a reliability graph with the empirical and theoretical distributions

%First, read the data on "Kilometres processed until failure" (km) and
%reliability (rel) from the Excel file
km = xlsread('E:/TSP/Data/Deelvraag
4/Testdata.xlsx','Incl_Censored_L','H2:H38');
rel = xlsread('E:/TSP/Data/Deelvraag
4/Testdata.xlsx','Incl_Censored_L','I2:I38');
% Create a figure with the 'real-life reliability data' on the y-axis and
% the kilometres processed on the x-axis
stairs(km,rel)
xlabel('Kilometres processed');
ylabel('Reliability');
ylim([0 1]);

% Now, plot the expected reliability of the two fitted distributions on the
% same graph to check how well they fit the real-life data
% Reliability = 1 - (Cumulative distribution function (cdf) of the fitted
% distribution)
% The parameters for the distributions are taken from the results of the
% Censored_data_fit code
% Weibull is drawn in red ('r'), Lognormal in blue ('b')
xgrid = linspace(0,3420,1000);
pdfEst = 1-wblcdf(xgrid,2100.4,1.831);
line(xgrid,pdfEst,'Color','r')

pd2Est = 1-logncdf(xgrid,7.3864,0.7701);
line(xgrid,pd2Est,'Color','b')

pd3Est = 1-gamcdf(xgrid,2.5174,761.4635);
line(xgrid,pd3Est,'Color','k')

for i = 1:length(km)
    Resid_gam(i) = rel(i) - (1-gamcdf(km(i),2.5174,761.4635));
end
for i = 1:length(km)
    Resid_wbl(i) = rel(i) - (1-wblcdf(km(i),2100.4,1.831));
end
for i = 1:length(km)
    Resid_logn(i) = rel(i) - (1-logncdf(km(i),7.3864,0.7701));
End

3. Create the hazard graph with the best fitting theoretical distribution

G = graph
xlabel('Kilometres processed');
ylabel('Hazard');
ylim([0 0.0025]);
xgrid = linspace(0,3420,1000);
Fail = wblpdf(xgrid,1530.6,1.8751);
Rel = 1-wblcdf(xgrid,1530.6,1.8751);
HazFun = wblpdf(xgrid,1530.6,1.8751)./(1-wblcdf(xgrid,1530.6,1.8751));
line(xgrid,HazFun,'Color','r')
line(xgrid,Fail,'Color','k')
line(xgrid,Rel,'Color','g')
Appendix H: Derivation and proof of correctness of formula 10.5

To get the expected time to failure (which is equal to the expected cycle length for corrective maintenance) in case of one failure mode, the formula

\[ \int_{t=0}^{t=\infty} tf(t) \, dt \]  

(AH.1)

should be used. For each possible value of \( t \) between 0 and \( \infty \), this formula calculates the probability that the failure mode will occur, multiplied by the time after which it occurs. If a second failure mode is added, this integral needs to be split up into two parts: a case where failure mode 1 occurs first and a case where failure mode 2 occurs first.

First, the probability of each of these two cases will be considered. The probability that failure mode 1 occurs before failure mode 2, can be explained in the following way: for every probability that failure mode 2 happens at time \( t \), we need the probability that failure mode 1 happens before \( t \).

Figure 50 represents this in a graph for one arbitrary value of \( t \).

The two lines in Figure 50 represent the probability density functions of two failure modes. The red dot provides the probability that failure mode 2 happens exactly at time \( t \), which equals \( f_2(t) \). The blue area in the figure is the probability that failure mode 1 happens before this time \( t \) and can be calculated with the integral of \( f_1 \) over all points between 0 and this point \( t \):

\[ \int_{u=0}^{u=t} f_1(u) \, du. \]  

(AH.2)

However, we need to know this probability for every possible value of \( t \) between 0 and \( \infty \), since failure mode 2 can happen at any of these moments. This results in the following double integral:

\[ \int_{t=0}^{t=\infty} \left( \int_{u=0}^{u=t} f_1(u) \, du \right) f_2(t) \, dt. \]  

(AH.3)
Opposite to formula A5.2, the probability that failure mode 1 occurs after \( f_2(t) \) can be described by the integral of \( f_1 \) over all points between \( t \) and \( \infty \):

\[
\int_{u=t}^{u=\infty} f_1(u) \, du. \tag{AH.4}
\]

Applying the same reasoning as for formula A5.3 leads to this formula:

\[
\int_{t=0}^{t=\infty} \left( \int_{u=t}^{u=\infty} f_1(u) \, du \right) f_2(t) \, dt. \tag{AH.5}
\]

To prove that all probabilities have been covered, formulas A5.3 and A5.5 should add up to 1. For any failure mode \( i \) with a probability density function that does not allow negative values of \( t \), it holds that:

\[
\int_{t=0}^{t=\infty} f_i(t) \, dt = 1. \tag{AH.5}
\]

Combining formulas A5.2 up to A5.5 leads to the following:

\[
\int_{u=0}^{u=t} f_1(u) \, du + \int_{u=t}^{u=\infty} f_1(u) \, du = \int_{u=0}^{u=\infty} f_1(u) \, du = 1
\]

\[
\int_{u=0}^{u=t} f_1(u) \, du = a
\]

\[
\int_{u=t}^{u=\infty} f_1(u) \, du = b
\]

\[
\int_{t=0}^{t=\infty} a \ast f_2(t) \, dt + \int_{t=0}^{t=\infty} b \ast f_2(t) \, dt = \int_{t=0}^{t=\infty} 1 \ast f_2(t) \, dt = 1. \tag{AH.6}
\]

Now we know that the probabilities add up to 1. To get the expected cycle length, formula A5.1 must be considered. If failure mode 1 occurs first, the probability that this failure mode occurs should be multiplied by the time at which it fails. The same applies to failure mode 2. This yields the formula used in Section 10.2 of the report (formula 10.5):

\[
\mathbb{E}(T) = \left( \int_{t=0}^{t=\infty} \left( \int_{u=0}^{u=t} u \ast f_1(u) \, du \right) f_2(t) \, dt \right) + \left( \int_{t=0}^{t=\infty} \left( \int_{u=t}^{u=\infty} f_1(u) \, du \right) t \ast f_2(t) \, dt \right). \tag{AH.7}
\]
Appendix I: Derivation and proof of correctness of formula 10.8

This formula is similar to the formula of Appendix G. However, the addition of a preventive replacement time $\tau$ increases the number of possible cases that needs to be considered. The different cases that need to be considered are:

1. Probability that failure mode 2 happens before $\tau$, failure mode 1 happens before failure mode 2 (failure mode 1 occurs);
2. Probability that failure mode 2 happens before $\tau$, failure mode 1 happens after failure mode 2 but before $\tau$ (failure mode 2 occurs);
3. Probability that failure mode 2 happens before $\tau$, failure mode 1 happens after failure mode 2 and after $\tau$ (failure mode 2 occurs);
4. Probability that failure mode 2 happens after $\tau$, failure mode 1 happens before $\tau$ (failure mode 1 occurs);
5. Probability that failure mode 2 happens after $\tau$, failure mode 1 happens after $\tau$ (so preventive maintenance is performed).

An example of cases 1, 2 and 3 is shown in Figure 51 for two failure modes. The red dot represents the probability that failure mode 2 happens at that particular moment. The three differently coloured areas are numbered corresponding to the case which they represent.

![Diagram showing three cases](image)

Figure 51: The three cases for a situation where failure mode 2 occurs before $\tau$

These 5 probabilities should add up to one. Using the same logic as in Appendix 5, this leads to a formula consisting of five double integrals:

$$
\int_{t=0}^{t=\tau} \left( \int_{u=0}^{u=t} f_1(u) \, du \right) f_2(t) \, dt + \int_{t=0}^{t=\tau} \left( \int_{u=t}^{u=\tau} f_1(u) \, du \right) f_2(t) \, dt + \int_{t=0}^{t=\tau} \left( \int_{u=\tau}^{u=\infty} f_1(u) \, du \right) f_2(t) \, dt + \\
\int_{t=\tau}^{t=\infty} \left( \int_{u=0}^{u=t} f_1(u) \, du \right) f_2(t) \, dt + \int_{t=\tau}^{t=\infty} \left( \int_{u=t}^{u=\infty} f_1(u) \, du \right) f_2(t) \, dt.
$$

(AI.1)

By replacing each of the double integrals above by a letter, we get $a + b + c + d + e$. 

101
The first three parts of A6.1 \((a, b, \text{and} c)\) have the same outer integral. Their inner integrals add up to 1:

\[
\int_{u=0}^{U=t} f_1(u) \, du + \int_{u=t}^{U=e} f_1(u) \, du + \int_{u=0}^{U=e} f_1(u) \, du = \int_{u=0}^{U=e} f_1(u) \, du = 1. \tag{A1.2}
\]

The two last parts of A6.1 \((d)\) and \((e)\) have the same outer integral as well. Their inner integrals add up to 1:

\[
\int_{u=0}^{U=t} f_1(u) \, du + \int_{u=t}^{U=\infty} f_1(u) \, du = \int_{u=0}^{U=\infty} f_1(u) \, du = 1. \tag{A1.3}
\]

Based on formulas A6.2 and A6.3, we can combine \(a, b, \text{and} c\) into one integral, as well as \(d\) and \((e)\). This leads to a shortened version of formula A6.1:

\[
\int_{t=0}^{t=\tau} 1 \cdot f_2(t) \, dt + \int_{t=\tau}^{t=\infty} 1 \cdot f_2(t) \, dt. \tag{A1.4}
\]

These two integrals also add up to 1:

\[
\int_{t=0}^{t=\tau} f_2(t) \, dt + \int_{t=\tau}^{t=\infty} f_2(t) \, dt = \int_{t=0}^{t=\infty} f_2(t) \, dt = 1. \tag{A1.5}
\]

Therefore, it holds that formula A6.1 adds up to one.

Now, the five parts of formula A6.1 need to be transformed to reflect the expected time to replacement. In the cases where failure mode 1 occurs first, \(f_1(u)\) needs to be multiplied by \(u\), where failure mode 2 occurs first, \(f_2(t)\) needs to be multiplied by \(t\), and in case of preventive maintenance, the integral needs to be multiplied by \(\tau\). This gives the following formula:

\[
ECL = \left( \int_{t=0}^{t=\tau} \left( \int_{u=0}^{U=t} u \cdot f_1(u) \, du \right) f_2(t) \, dt \right) + \left( \int_{t=0}^{t=\tau} \left( \int_{u=t}^{U=e} f_1(u) \, du \right) t \cdot f_2(t) \, dt \right) + \\
\left( \int_{t=0}^{t=\tau} \left( \int_{u=t}^{U=\infty} f_1(u) \, du \right) t \cdot f_2(t) \, dt \right) + \left( \int_{t=\tau}^{t=\infty} \left( \int_{u=0}^{U=e} u \cdot f_1(u) \, du \right) f_2(t) \, dt \right) + \\
\tau \left( \int_{t=\tau}^{t=\infty} \left( \int_{u=t}^{U=\infty} f_1(u) \, du \right) f_2(t) \, dt \right) = a + b + c + d + e. \tag{A1.6}
\]

Parts \(b\) and \(c\) of this formula can be combined:

\[
\left( \int_{t=0}^{t=\tau} \left( \int_{u=t}^{U=t} f_1(u) \, du \right) t \cdot f_2(t) \, dt \right) + \left( \int_{t=0}^{t=\tau} \left( \int_{u=t}^{U=\infty} f_1(u) \, du \right) t \cdot f_2(t) \, dt \right) = \\
\int_{t=0}^{t=\tau} \left( \int_{u=t}^{U=\infty} f_1(u) \, du \right) t \cdot f_2(t) \, dt. \tag{A1.7}
\]

The last part of formula A6.6 \((e)\) can be rewritten:

\[
\tau \int_{t=\tau}^{t=\infty} \left( \int_{u=t}^{U=\infty} f_1(u) \, du \right) f_2(t) \, dt = \tau \cdot R_1(\tau) \cdot R_2(\tau). \tag{A1.8}
\]

Making this substitution in formula A6.6 leads to the formula presented in Section 10.3 of the report (formula 10.8):

\[
ECL = \left( \int_{t=0}^{t=\tau} \left( \int_{u=0}^{U=t} u \cdot f_1(u) \, du \right) f_2(t) \, dt \right) + \left( \int_{t=0}^{t=\tau} \left( \int_{u=t}^{U=\infty} f_1(u) \, du \right) t \cdot f_2(t) \, dt \right) + \\
\left( \int_{t=\tau}^{t=\tau} \left( \int_{u=t}^{U=\infty} u \cdot f_1(u) \, du \right) f_2(t) \, dt \right) + \tau \cdot R_1(\tau) \cdot R_2(\tau). \tag{A1.9}
\]
Appendix J: ECCs for different corrective maintenance costs of failure modes

Scenario 1
From formulas AH.3 and AH.5, we already have the probability that either failure mode will occur first. If failure mode 1 occurs, the corresponding costs $C_{cm1}$ will be incurred. If failure mode 2 occurs, the costs will be equal to $C_{cm2}$. Combining this leads to the following formula:

$$ECC = \int_{t=0}^{\infty} \left( \int_{u=0}^{u=t} f_1(u) \, du \right) f_2(t) \, dt \cdot C_{cm1} + \int_{t=0}^{\infty} \left( \int_{u=t}^{u=\infty} f_1(u) \, du \right) f_2(t) \, dt \cdot C_{cm2} \quad (AJ.1)$$

Scenario 2
For the second scenario, the calculation of the ECC is similar. However, there is a third possibility, namely that preventive maintenance will occur. Therefore, the formula needs to be adjusted. The upper bound of the failure mode that occurs first needs to be changed to $\tau$. If the failure happens after $\tau$, there will be a preventive replacement and the costs will be equal to $C_{pm}$.

$$ECC = \int_{t=0}^{\infty} \left( \int_{u=0}^{u=t} f_1(u) \, du \right) f_2(t) \, dt \cdot C_{cm1} + \int_{t=0}^{\infty} \left( \int_{u=t}^{u=\infty} f_1(u) \, du \right) f_2(t) \, dt \cdot C_{cm2} +$$

$$(R_1(\tau) \cdot R_2(\tau)) \cdot C_{pm} \quad (AJ.2)$$

Note: both formulas AJ.1 and AJ.2 are not used in the Matlab script (Appendix K).
Appendix K: Matlab code for optimization of operating costs

% Determine the costs per time unit for corrective maintenance, preventive maintenance for a single failure mode and prev. for two failure modes
% In case of preventive maintenance, the optimal maintenance interval will be determined with regard to costs.
clear
% First, fill in the input parameters for costs below

% Costs of downtime for replacing a component correctly
C_dt_cm=800;
% Costs of downtime for replacing a component preventively
C_dt_pm=0;
% Costs of maintenance (repair/new component costs) for corrective action
C_m_cm=1000;
% Costs of maintenance (repair/new component costs) for preventive action
C_m_pm=1000;

% Calculate the total costs for a protective and corrective maintenance action
C_cm = C_dt_cm + C_m_cm;
C_pm = C_dt_pm + C_m_pm;

% Specify the distribution of each failure mode
Fm_1_distr = 'GAMMA';
Fm_2_distr = 'GAMMA';

% Now, fill in the distribution and parameter values for both failure modes
Value_1_f1=2.5174; % First parameter value for failure mode 1
Value_2_f1=761.4635; % Second parameter value for failure mode 1
Value_1_f2=0.5457; % First parameter value for failure mode 2
Value_2_f2=13329; % Second parameter value for failure mode 2

% Define variables t and u for f_1(u) and f_2(t)
syms t u;

% Formulas to be used in the various integrals (coming from table 19) in the report: These can be replaced by formulas of other failure distributions

% Probability density function (pdf) of failure mode 1
f1 = ((u.^(Value_1_f1-1)).*exp(-u./Value_2_f1))./(Value_2_f1.^Value_1_f1).*gamma(Value_1_f1);
% Probability density function of failure mode 1 * u
uf1 = (u.*((u.^(Value_1_f1-1)).*exp(-u./Value_2_f1))./(Value_2_f1.^Value_1_f1).*gamma(Value_1_f1));

% Probability density function of failure mode 2
f2 = ((t.^(Value_1_f2-1)).*exp(-t./Value_2_f2))./(Value_2_f2.^Value_1_f2).*gamma(Value_1_f2);
% Probability density function of failure mode 2 * t
uf2 = (t.*((t.^(Value_1_f2-1)).*exp(-t./Value_2_f2))./(Value_2_f2.^Value_1_f2).*gamma(Value_1_f2));

% CASE 1: Corrective, two failure modes

%------------------------- -----------------------------------------------
%------------------------- CASE 1: Corrective, two failure modes ----------------
%------------------------- -----------------------------------------------
%Inner integrals:
int_u_f1_0_t = int(u_f1,u,0,t);
int_f1_t_inf = int(f1,u,t,Inf);

%Multiply the outcomes of the inner integral by either f2(t) or t*f2(t) and
%turn the result into a matlab function for the next integral
int_2x1 = matlabFunction(int_u_f1_0_t * f2);
int_2x2 = matlabFunction(int_f1_t_inf * t_f2);

%Calculate the outer integrals and add them up to get the ECL
part1 = integral(int_2x1,0,Inf)
part2 = integral(int_2x2,0,Inf)
ECL_corr = part1 + part2

%ECC is equal to the corrective costs in this case, so the cost per km is:
g_corr = C_cm/ECL_corr;

%Turn the formula for u*f(u) into a matlab function
uf1u = matlabFunction(u_f1);

% The optimal value of tau for case 2 is calculated with a loop. Increasing
% values of tau are tried and stored in an array. These values can be
% plotted to find the value of tau which leads to the lowest costs per km
% The range and precision of the loop can be adjusted below:
% Start_value determines the lowest value of tau at which will be looked
% This value must be set to a whole number, equal to or bigger than 1.
% Precision determines how much will be added to tau each loop (so the size
% of each step)
% Nr_of_cycles defines the number of loops (don't set this bigger than 100
% to keep calculation times fairly low)
Start_value1 = 100;
Nr_of_cycles1 = 100;
Precision1 = 100;
Start_value1 = Start_value1 / Precision1;

% Create the array for the results of the cost function under different tau
% values
g_2 = zeros(Nr_of_cycles1,4);
for i = 1:1:Nr_of_cycles1
    tau_2 = (Start_value1+i-1)*Precision1;
    if strcmpi(Fm_1_distr,'GAMMA')==1
        CDF_1 = gamcdf(tau_2,Value_1_f1,Value_2_f1);
    end
    ECC_2 = CDF_1*C_cm+(1-CDF_1)*C_pm;
    ECL_2 = (tau_2*(1-CDF_1)) + integral(uf1u,0,tau_2);
end

%This ECC is equal to formula (10.10) of the report
ECC_2 = CDF_1*C_cm+(1-CDF_1)*C_pm;
%This ECL is equal to formula (10.12) of the report
ECL_2 = (tau_2*(1-CDF_1)) + integral(uf1u,0,tau_2);

%Print the results in this array. The first two columns provide the
%values of tau and the corresponding costs per time unit g(tau)
\begin{verbatim}
g_2(i,1) = tau_2;
g_2(i,2) = ECC_2 / ECL_2;
g_2(i,3) = ECC_2;
g_2(i,4) = ECL_2;
end

% Function to optimize g(tau) automatically for this strategy (first line is ECC, second
% is ECL)
if strcmpi(Fm_1_distr, 'GAMMA') == 1
    fun_g_2 = @(tau) ((gamcdf(tau,Value_1_f1,Value_2_f1))*C_cm+(1-gamcdf(tau,Value_1_f1,Value_2_f1))*C_pm)...
    / (tau*(1-gamcdf(tau,Value_1_f1,Value_2_f1))+integral(uf1u,0,tau));
end

% Plot the results of the optimization algorithm in a graph
options = optimset('Display','iter','PlotFcns',@optimplotfval);
min_g_2 = fminbnd(fun_g_2,0,10000,options)

% The optimal value of tau for case 3 is calculated with a loop. Increasing
% values of tau are tried and stored in an array. These values can be
% plotted to find the value of tau which leads to the lowest costs per km

% The range and precision of the loop can be adjusted below:
% Start_value determines the lowest value of tau at which will be looked
% This value must be set to a whole number, equal to or bigger than 1.
% Precision determines how much will be added to tau each loop (so the size
% of each step)
% Nr_of_cycles defines the number of loops (don't set this bigger than 100
% to keep calculation times fairly low)
Start_value = 100;
Nr_of_cycles = 50;
Precision = 100;

Start_value = Start_value / Precision;

for i = 1:1:Nr_of_cycles
    tau_3 = (Start_value+i-1)*Precision;
    %integral of u*f_1(u) from 0 to tau
    int_u_f1_0_tau = integral(u_f1,u,0,tau_3);
    %Define F(u) for failure mode 1
    if strcmpi(Fm_1_distr, 'GAMMA') == 1
        CDF_1 = gamcdf(tau_3,Value_1_f1,Value_2_f1);
    end
    %Define F(t) for failure mode 2
    if strcmpi(Fm_2_distr, 'GAMMA') == 1
        CDF_2 = gamcdf(tau_3,Value_1_f2,Value_2_f2);
    end

    %This ECC is equal to formula (10.7) of the report
    ECC_3 = (1-CDF_1)*(1-CDF_2)*C_pm + (1-(1-CDF_1)*(1-CDF_2))*C_cm;
    %This ECL is equal to formula (10.8) of the report
\end{verbatim}
ECL_3 = (tau_3*(1-CDF_1)*(1-CDF_2)) + integral(matlabFunction(int_u_f1_0_t * f2),0,tau_3) + integral(matlabFunction(int_f1_t_inf * t_f2),0,tau_3) + integral(matlabFunction(int_u_f1_0_tau * f2),tau_3,Inf);

%Print the results in this array. The first two columns provide the values of tau and the corresponding costs per time unit g(tau)
g_3(i,1) = tau_3;
g_3(i,2) = ECC_3 / ECL_3;
g_3(i,3) = ECC_3;
g_3(i,4) = ECL_3;
g_3(i,5) = (1-CDF_1)*(1-CDF_2); %perc. of survival until prev. maintenance

end
Appendix L: Verification of Matlab code for optimization of operating costs

The formulas for optimization of operating costs have been processed in the Matlab script of Appendix K. To verify the results, some parameter values will be chosen for which the result can also be calculated without using the script. The results of the script should match the expected results.

**Scenario 1: Corrective maintenance with both failure modes:**

The corrective maintenance costs per kilometre, $g$, consist of two parts: the Expected Cycle Costs (ECC) and the Expected Cycle Length (ECL). The Expected cycle costs should always be equal to the corrective maintenance costs (€1.800 for the scrap cutter). The ECL should be similar to the average time to failure of a scrap cutter. The failure data, which has also been used to determine the probability distribution in Chapter 9, can be compared to the ECL calculated by the Matlab script.

The average moment of failure for the scrap cutter is after 1358 kilometres of material processed according to the failure data. The ECL produced by the Matlab script is 1382 kilometre. This is only a very small difference, so the calculated ECL of the Matlab script can be expected to be correct.

**Scenario 2: Preventive maintenance with both failure modes:**

If the value of the preventive maintenance interval $\tau$ is set to 0, the costs should be infinite. This makes sense from an intuitive point of view, but can also be deducted from the formulas. If the time until preventive replacement is equal to zero, the Expected Cycle Length will also be 0. Since the costs are calculated by the formula ECC/ECL, the resulting cost per km is infinite or simply non-existing as you cannot divide by 0. The Matlab script for scenario 2 crashes with the error message "Infinite or Not-a-Number value encountered" for $\tau = 0$.

When $\tau$ is set to an extremely large number (or infinity), the ECC should be equal to the corrective maintenance costs (€1.800 for the scrap cutter) and the ECL should be equal to the result of scenario 1. Therefore, the costs per kilometre should be equal to scenario 1 as well. This is confirmed by the results in Sections 11.1 and 11.2 of the report. Both scenarios lead to the same cost per km of €1,30. Therefore, they are consistent with each other.

**Scenario 3: Preventive maintenance with only the failure mode dull blades:**

Similarly to scenario 2, when $\tau$ is set to 0, the ECC should be €1.000, the ECL should be 0 and $g(\tau)$ should be infinity. The result of the Matlab script in this scenario is different from scenario 2. Entering the value $\tau = 0$ into the script results in the result $g(\tau) = \infty$. As explained above, this answer is acceptable, so this check is passed. On top of that, the ECC should be equal to the preventive maintenance costs (€1.000 for the scrap cutter). This is the output of the script as well.

If $\tau$ is set to infinity, the ECC should be equal to the corrective maintenance costs. Again, the results of the script are in line with the expected results.

Finally, since one failure mode is removed, the ECL of this scenario should be longer than the ECL of the previous scenario. This is confirmed by the results and can also be seen in Section 11.1 and 11.2 of the report.
Varying the probability distribution:

On top of the tests of the 3 scenarios which mostly deal with varying values of $\tau$, the probability distribution can be changed. If the value of the scale parameter of the Gamma distribution(s) is increased, the time to failure should be longer. Therefore, the ECL should increase. If the value of the scale is reduced, the ECL should be lower as well. This has also been confirmed with the calculations at the end of Section 11.3 (especially Figure 45).