Industrial Engineering & Management
Master thesis

Transitioning to a counter-based maintenance strategy at the packaging filling lines of Heineken

A thesis submitted in fulfillment of the requirements for the degree of Master of Science in the field of:

Production & Logistic Management

Manufacturing Logistics

Author:

N. Meulenbroek

Company supervisor:

J. Campfens

University of Twente supervisors:

Dr. E. Topan

Dr. I. Seyran Topan

August 30, 2024

UNIVERSITY OF TWENTE.

TRADE MARK Heineken®

Management Summary

This thesis explores transitioning from a time-based to a counter-based maintenance strategy for the packaging filling lines at Heineken's Zoeterwoude brewery. Time-based maintenance follows a fixed schedule (e.g. every two weeks), while counter-based maintenance is planned based on actual usage of equipment. Currently, the time-based approach for packaging filling lines leads to inconsistent maintenance intervals, causing over- and under-maintenance due to variable production hours. With expected decreases in total production, the company demands an adaptive strategy, which schedules maintenance based on actual machine usage; counter-based maintenance.

To arrive at a solution to this problem, the main research question is this:

"How can the maintenance planning activities of the packaging filling lines in the brewery at Zoeterwoude of Heineken be adjusted to change from a time-based plan to a counter-based plan?"

To address this, the research starts by looking at the current situation, studying what others have done, creating new solutions, and testing them. An analysis at the packaging filling lines found that production hours between maintenance activities vary a lot. For example, one year had 2,000 hours between two activities, and another year had 3,000 hours between the same activities. This shows the company sometimes does too much or too little maintenance.

The literature review highlights the importance of transitioning from time-based to counter-based maintenance strategies for improved efficiency and cost savings. For example, Tinga (2013, p. 204) states that, considering the results of a case-study, transitioning from time- to counter-based systems notably reduces replacements with about 30%, especially during low operating hours. Although current theories support the benefits of counter-based maintenance, they lack practical guidelines for transitioning from time-based systems.

The solution outlines a clear approach of the transition. It introduces methods to calculate usage counters, incorporating the current time-based intervals. Table 1 reveals that the progression of the usage counters are lower or higher than the current time-based progressions, meaning maintenance should not be planned as early or late as with the time-based strategy. The percentage represents the amount of time passed since the last execution divided by the set interval, either in weeks for time-based or in production hours for counter-based maintenance.

TABLE 1: Five results of applying counter-based maintenance

Code	**Description**	**Last execution**	**Time-based progress**	**Interval** (hours)	**Counter** (hours)	**Counter-based progress**
PM1	SIXO812.1, 1J	2-11-2023	64%	3,375	1,942	58%
PM1	Losdok81, 6M	26-1-2024	82%	2,447	2,165	88%
PM1	TRANS-VL-DS81, 6M	30-1-2024	80%	1,688	1,266	75%
PM2	DOZO81, 1J	24-8-2023	84%	3,375	2,331	69%
PM4	ETIMA, 3J	31-3-2022	75%	15,098	12,176	81%

Take for example the fourth row in Table 1. Regarding its time-based interval, it will be scheduled in the near future because it is already at 84%, while the counter-based

progress is only at 69%. Executing this activity too early is an example of over-maintenance. These examples in the table indicate that the company could perform less over- and under-maintenance if they consider the counter-based progressions.

It is difficult to move from a time-based to a counter-based maintenance strategy. This means predicting future production is necessary if the company wishes to prepare activities. The research tested short-term (13 weeks aggregated data) and long-term (1 week non aggregated data) prediction models to find the best way. Aggregating data is the process of summing information from multiple data-points to provide a comprehensive view, reducing noise, but it fails to show the exact information of a specific week. The research uses production data (in minutes) from 2016-2021 to train models like Linear Regression, Decision Tree, Random Forest, Gradient Boosting Machine, K-Nearest Neighbor, Holt-Winters, and a simple baseline model. These models predict 2022-2023 values. If the baseline model works best, it means using simple averages is better than complex models. The values of RMSE, MAD, sMAPE, and Bias show the success of the models. The first three should be low, and Bias shows if the model guesses too high or too low. Table 2 shows which model performed best (see grey cells).

TABLE 2: Average Performance Metrics of Model Types for both Aggregated and Non Aggregated Data

	Aggregated Data				Non Aggregated Data			
Model	**RMSE**	**MAD**	**sMAPE**	**Bias**	**RMSE**	**MAD**	**sMAPE**	**Bias**
Baseline	13908.33	11873.16	19.65%	-9351.49	2048.40	1610.90	38.65%	-615.61
LR	12011.24	9929.66	16.25%	-3829.79	1784.44	1406.62	36.66%	-431.28
DT	13500.02	11229.74	18.99%	-7084.88	1900.13	1464.80	37.62%	-402.50
RF	10875.16	9466.41	15.43%	-5224.77	1808.92	1437.25	36.18%	-391.02
GBM	11401.90	10029.58	16.40%	-6301.94	1825.33	1453.03	36.87%	-478.37
KNN	13139.49	10926.30	18.68%	-6478.31	1976.68	1539.52	39.44%	-430.09
HW	12414.19	9970.86	15.14%	7103.20	2176.81	1693.78	40.71%	-276.97

Table 2 indicates that for 13-week aggregated production data, the Random Forest model outperforms others, while for the non aggregated production data the Linear Regression model is the best. In both cases they are better than the baseline, because these models are more sophisticated. Their expertise comes from the fact that they use the input variables such as the moment in the year or workforce hours from the tactical planning.

This research also develops a mathematical model to schedule preventive maintenance, reducing over- and under-maintenance using a rolling horizon approach. This means the model sets a plan for the next year but adjusts it every quarter to stay flexible with production changes. It was tested how different production levels impact the model for both time-based and counter-based maintenance. If one maintenance activity costs €200 and using real production data from 2023, it is possible to see how well the model works. The NPM shows the average hours deviating from the counter-based intervals per activity. A perfect NPM is 0, meaning no over- or under-maintenance, but some deviation is expected. Table 3 shows the NPM, maintenance costs (MC), and the split between over-maintenance (OM) and under-maintenance (UM) for different scenarios with decreasing production.

The first column in Table 3 shows the weekly average production hours randomly subtracted from the data in year 2023. Comparing counter-based (CBM) to time-based strategies (TBM) in the year 2023 with no reduced production, the former

TABLE 3: Sensitivity Analysis of planning model applied to 2023

	NPM (hours)		**MC** (€)		**OM/UM** (%)	
Avg. reduction per week	**TBM**	**CBM**	**TBM**	**CBM**	**TBM**	**CBM**
0 hours	295.08	200.60	€71,800.-	€74,200.-	22%/78%	25%/75%
5 hours	256.07	208.91	€71,800.-	€71,016.-	31%/69%	31%/69%
12.5 hours	218.44	202.11	€71,600.-	€63,316.-	50%/50%	29%/71%
25 hours	278.71	178.58	€71,400.-	€55,784.-	78%/22%	35%/65%
37.5 hours	390.48	196.82	€71,200.-	€46,016.-	90%/10%	49%/51%
50 hours	505.03	188.12	€70,400.-	€38,920.-	93%/7%	52%/48%

reduces over- and under-maintenance by about 32%, while it costs €3,400.- more. These extra costs come from additional activities to reduce under-maintenance. Table 3 reveals the counter-based schedules better maintain a low level of total over- and under-maintenance, while it also better divides the proportions. For example, if the company expects an average reduction of production of 37.5 hours in the upcoming years, observing the maintenance costs in this situation, the newly introduced counter-based strategy with the planning model could save about €25,000.- in a year compared to the time-based strategy, in addition to a reduction of almost 50% of over- and under-maintenance, which are also more evenly distributed. All these values show the possible benefits of a counter-based maintenance strategy!

The biggest limitation of this research was in the planning model, which assumes a fixed number of activities at a single moment for maintenance, while in reality, this depends on many factors like activity duration, skill sets, and availability of spare parts and workforce. The model needs to account for these constraints for extra validity.

A specific example of the contribution to theory is the challenge of predicting production hours for counter-based maintenance. While time-based systems allow easy scheduling due to constant intervals, counter-based systems struggle with this. This research introduces a method to predict production using machine learning.

The practical contribution mainly is cost reduction. Companies with decreasing production can save costs by using counter-based maintenance (when coming from time-based systems). Table 3 shows how the model is especially cost-effective when production decreases, helping the companies maintain competitiveness.

Future research could explore more advanced machine learning techniques, like neural networks, to improve production prediction accuracy. Incorporating diverse datasets with additional features could refine these models further.

The most important recommendations are:

1. Implement the Random Forest model to forecast production and plan maintenance 13 weeks in advance, refining the model with more data and features.

2. Extend the mathematical model from this research for scheduling, with further investigation into a holistic model considering more constraints.

3. Create a data model integrating different data sources, to automatically update usage counters.

Preface

This master thesis is the grand finale of my student adventure at the University of Twente. My time in Enschede has been an absolute blast, a whirlwind of academic challenges and personal growth that I will treasure forever. Nobody achieves anything all alone, so I would like to take the opportunity to give my thanks to the people who were part of this wonderful and challenging period.

To begin, I wish to thank Heineken for giving me the chance to finish my graduation assignment within the Maintenance Engineering Team of the brewery in Zoeterwoude. Everyone has been decent and kind. There is a welcoming atmosphere and people are very dedicated, enthusiastic, and open for everyone. A special thanks to my supervisor, John Campfens, and his team. I know that it has been a very busy period for the manager and the other maintenance engineers. The fact that all of them still made the time and effort to help me with at best of their abilities says a lot about the people that work here. Not only was everyone very professional, I have made some great personal relationships with my colleagues, and I wish them all the best in the future. John specifically deserves a lot of credit. He manages to keep everyone motivated, and has pushed me in the correct direction whenever I needed it. Most importantly, from the very beginning up to the end, he has been very enthusiastic about all the work that the team and myself have been performing. Positive energy is contagious, and he understands this very well.

Secondly, a special thanks to my first supervisor at the University of Twente, Engin Topan. I have known him for quite some time, and I am very grateful for all his help during this assignment. His feedback and guidance have helped me in being motivated, staying positive, and getting the best out of myself as a student. In all honesty, I can say that I am very proud of the final result, and this was not possible without Engin's help. On top of that, the support of my second supervisor, Ipek Seyran Topan, has also been vital. Whatever Engin and myself missed, Ipek noticed. She delivered the critical feedback that was essential to finish this thesis. Therefore, she also deserves my thanks.

And last but definitely not least, a huge thanks to my amazing friends, girlfriend, and family. Your support during the past months and throughout my student life has given me memories to proudly look back on.

I hope you enjoy reading this thesis as much as I enjoyed creating it!

Nils Meulenbroek

Utrecht

August 2024

Contents

List of Figures

List of Tables

List of Abbreviations

AvgL Average Length

BOW Bottle One Way

CBM Condition-Based Maintenance

CTBM Calendar Time-Based Maintenance

DT Decision Tree

GBM Gradient Boosting Machine

HNS Heineken Netherlands Supply

HW Holt-Winters

KNN K-Nearest Neighbor

LBM Load-Based Maintenance

LR Linear Regression

MAD Mean Absolute Deviation

MC Maintenance Costs

MILP Mixed Integer Linear Programming

NPM Numeric Performance Metric

OM Over-Maintenance

PD Planned Downtime

PM Preventive Maintenance

PMSP Preventive Maintenance Scheduling Problem

RF Random Forest

RMSE Root Mean Squared Error

sMAPE Symmetric Mean Absolute Percentage Error

UBM Usage-Based Maintenance

UM Under-Maintenance

UPD Unplanned Downtime

Chapter 1

Introduction

The production industry is a very large sector with many different companies creating various types of products. Many of them play a large role in society and some firms even send their products all over the world. The Heineken Company, specifically Heineken Netherlands Supply (HNS), is one of these companies and they were kind enough to introduce a new problem in one of their departments. This thesis will explore strategical concepts considering the planning of maintenance activities at the packaging filling lines of the brewery in Zoeterwoude. Section 1.1 introduces the company, and Section 1.2 gives the problem description. Section 1.3 provides the methodology of this research to arrive at a useful solution.

1.1 The Company

Established in 1864, Heineken N.V. is a prominent Dutch multinational brewing company. With a presence in over 70 countries, the company operates more than 168 breweries and boasts a diverse portfolio of more than 350 beers and ciders, including international, regional, local, and specialty variants. Employing around 90,000 individuals, Heineken has solidified its position as a major player in the global brewing industry. (N.V., 2023)

In 2023, Heineken N.V. achieved significant milestones, producing a staggering 24.26 billion liters of beer and generating global revenues to a total of 30.308 billion euros. This performance secured its standing as the leading brewer in Europe and positioned it among the world's largest brewers by volume. (N.V., 2023)

The biggest Heineken Dutch breweries are strategically situated in Zoeterwoude and 's-Hertogenbosch. This thesis focuses on Zoeterwoude, considering the activities of its Maintenance Engineering Team. This team has the mission to improve the reliability and performance of technical installations by optimizing maintenance strategies at the lowest possible costs. With their technical expertise the team advises the production departments and use their specialisms for the correct technical documentation, policies and plans across the departments. A brewery consists of many sections, such as spaces for the brewing process, packaging facilities, storage warehouses, et cetera. Packaging facilities contain the filling lines of many types of products, such as metal cans, glass bottles, kegs. The packaging filling lines use multiple machines for each step of the packaging process, and these machines sometimes fail or need maintenance. Therefore, maintenance planning is an important aspect of the packaging process, considering when to plan maintenance, how to plan it, and how maintenance can be integrated in the total brewing process and with all the other departments/teams.

1.2 Problem Description

This section delivers a clear description of the problem. Section 1.2.1 gives the perspective from the management of the company, and Section 1.2.2 delivers an analysis of all observed problems. Then Section 1.2.3 provides a statement that describes the core problem.

1.2.1 Management Problem

When the management of the company considers the planning of maintenance at the packaging filling lines, many concerns arise. Production is important since downtime at the filling lines, which is the time where work is being halted for any reason, leads to downtime costs. Downtime costs are the profits that the company loses when its equipment or network stops functioning, e.g. the filling lines being stopped. It was found that in 2023, if a packaging filling line's total amount of downtime increases with 1%, it would cost the company tens of thousands of euros. One of the causes of unexpected downtime is an insufficient maintenance strategy.

Currently there is a maintenance strategy in use, meaning they define rules for the sequence of maintenance activities. The management of the company observes that many maintenance activities at the packaging filling lines are based on time. Time-based maintenance is a maintenance plan based on a fixed time interval, e.g. once every two weeks. A counter-based maintenance plan is of another type used for planned maintenance or reactive maintenance based on asset counter registrations. A time-based maintenance plan for the packaging filling lines would be adequate if the interval of the maintenance activities contains the same amount of production hours. However, for the packaging filling lines, the production between two consecutive maintenance activities currently differs in the thousands of operating hours.

Calendar time-based maintenance follows a set schedule, which may lead to over-servicing some equipment and under-servicing others, also called over- and under-maintenance, resulting in inefficiencies and increased risks of breakdowns. Literature gives the indication that a counter-based maintenance strategy could improve these inefficiencies. Tinga (2013, p. 204) states that, considering the results of a case-study, going from a time-based maintenance strategy (CTBM: calendar time-based method) to a counter-based maintenance strategy (UBM: Usage-based method) reduces the number of replacements considerably, especially in the years with low numbers of operating hours. Cichelli (1977, p. 69) found that going from time-based maintenance to counter-based maintenance can increase productivity with 0.6%, which could be even more in reality. Christer and Doherty (1970, p. 924) shows a case where their maintenance activities were recommended to be controlled on a tonnage throughput basis as opposed to a calendar time basis, because of the cost reductions it would provide. Finally, Liu, Wang, and Tan (2024) discusses that counter-based preventive maintenance policy has the advantage of lower operational cost rate, but that time-based maintenance systems can achieve similar results. Further explanations on the benefits of counter-based maintenance are in Section 3.2.

Based on the arguments from the management of the company and the literature, it is simply understood that there are inefficiencies in the maintenance strategy, and there are unnecessary costs due to over- and under-maintenance. Moreover, the management confirms that the transition from time-based to counter-based maintenance is a complex process that needs to consider stakeholders, data analysis, planning optimization, and integration of departments. Not only is counter-based maintenance an intricate strategy on itself, transitioning from a time-based system for such a large operation as the brewery in Zoeterwoude becomes a very delicate project. Therefore, the problem perceived by the management of the company is the lack of knowledge on how to change the maintenance planning process from time-based strategies to counter-based strategies.

1.2.2 Problem Cluster

Due to the broad scope of the management problem, this section gives a logistical analysis to identify a suitable core problem. After further analysing the perceived problems, a cause-and-effect relationship between them is observed, allowing the possibility to visualise a problem cluster. A problem cluster is a model used to map different problems and their mutual relationships. A problem cluster serves as a means of structuring the problem context and is used to identify the core problem (Heerkens and Winden, 2017, p. 51). An arrow visualises the causes and effects, where the cause points to the effect. The action problem observed by the company is visualised with the grey colour. An action problem is a discrepancy between the norm and reality, as perceived by the problem owner (Heerkens and Winden, 2017, p. 22). Figure 1.1 presents the problem cluster.

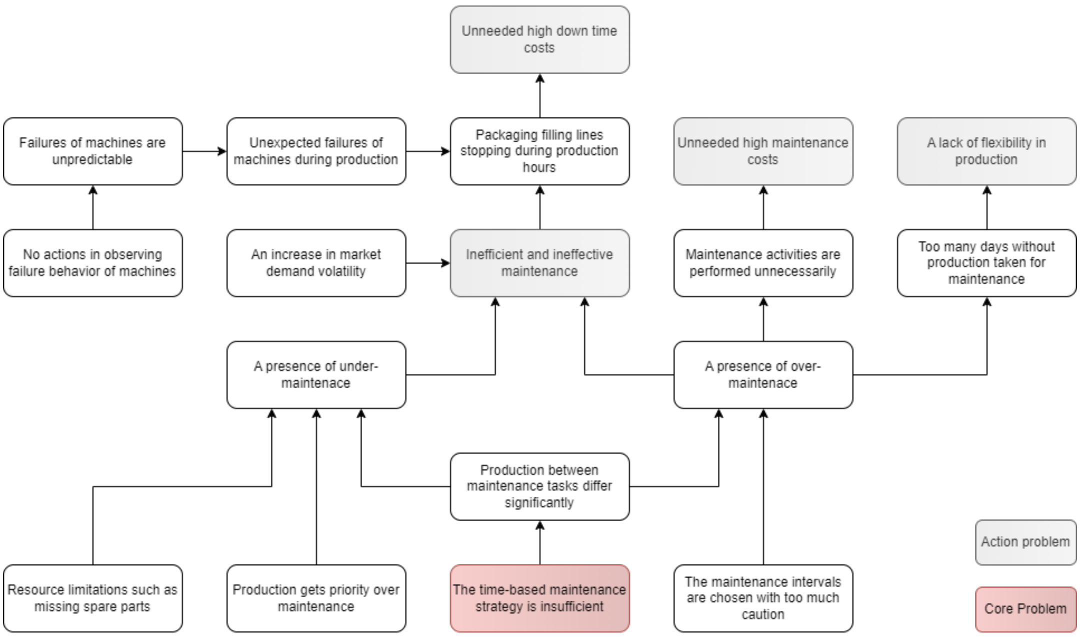


FIGURE 1.1: Problem cluster

Figure 1.1 shows the relationship between causes and effects of observable problems at the packaging filling lines. The management problems of Section 1.2.1 are shown in the grey boxes. The company observes high down time costs which are unnecessary. This is caused by unexpected failures of the machines during production hours, and due to the maintenance being done inefficiently and ineffectively. The unexpected failures of the machines is a result of it being very unpredictable. The fact that the company does not take any action into observing the failure behavior of the machines is the main cause of this problem chain.

Effective maintenance means applying a maintenance strategy that ensures the prevention of failures, while efficient maintenance suggests that it achieves this outcome with with minimal waste of resources. More information on these terms in Section 3.1. The maintenance being inefficient and ineffective has multiple causes, one of them being the increase in volatility of the market demand. Market demand volatility refers to the degree of fluctuation in the demand for goods or services within a market over a certain period of time, where high volatility means that demand levels change rapidly and unpredictably, and low volatility indicates more stable and consistent demand patterns. Market demand volatility can disrupt traditional maintenance planning, but counter-based maintenance offers a more adaptive approach by basing maintenance decisions on actual equipment usage. This enables organizations to better manage maintenance activities in response to fluctuations in market demand.

Another cause of inefficient and ineffective maintenance is the presence of under- and over-maintenance. Over-maintenance causes the maintenance activities to be unnecessary at times, resulting in unneeded high maintenance costs. Additionally, over-maintenance leads to many days during the year spent on maintenance, making the production schedules less flexible, which is another action problem of the company. All maintenance engineers commonly understand that too much maintenance may also lead to extra unplanned stops due to the excessive work on the machines that are dis- and reassembled frequently. Under- and over-maintenance are results of the amount of production varying between the maintenance activities. This is a direct result of the maintenance intervals being based on time (e.g. every two weeks). Other causes for under-maintenance are resource limitations, meaning that spare parts sometimes are missing, and the fact that production is considered as too important to perform maintenance. Finally, whenever the interval of maintenance is chosen with too much caution it can result in over-maintenance.

1.2.3 Core Problem

Core problems are those whose solutions will make a real difference (Heerkens and Winden, 2017, p. 41). The potential core problems are always at the beginning of a causation chain in a problem cluster. The choice of the current time-based maintenance plan is causing the hypothesis of inefficient maintenance planning because production activities differ each week and change dynamically during the year. This problem (shown in red in Figure 1.1) is chosen as the core problem. This is the better core problem to tackle because its solution is feasible, and the solution of the problem will have the biggest impact on the management problem. This problem also gives a clear suggestion on what would be a better method, namely a counter based maintenance strategy. The selected core problem is defined as:

In the current situation, the maintenance planning activities at the packaging filling lines in the brewery at Zoeterwoude of Heineken are time-based, while the production activities differ throughout the year and change dynamically, which indicates that a counter-based maintenance strategy is more suitable.

1.3 Problem Solving Approach and Research Design

This section explains the problem-solving approach and reports on the research design. Section 1.3.1 gives the scope of the research, and Section 1.3.2 delivers the methodology and the research questions. Finally, Section 1.3.3 follows with the deliverables.

1.3.1 Research Scope

Heineken has different breweries across many countries, a few of them present inside of the Netherlands. One of them, the brewery at Zoeterwoude has various departments, all of them needing maintenance. The whole production process consists of brewing, fermentation & lagering, filtration, intermediate storage (tanks), bottling & packaging, and again storage (pallets) before the product is sent away via multiple transportation methods. Figure 1.2 visualizes these steps of the production process in a flow chart.

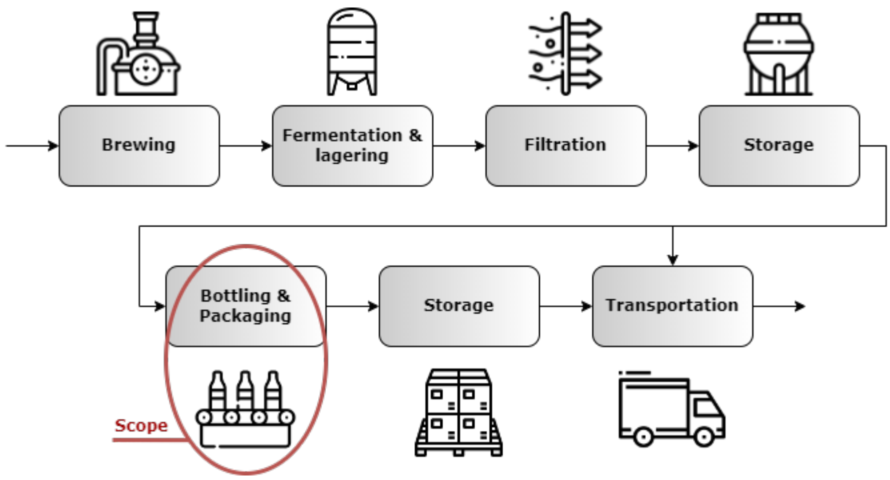


FIGURE 1.2: Production process flow chart with the identified scope of the research

For this thesis to solve the core problem for different breweries and at all departments would be too broad and time extensive. The Maintenance Engineering Team has observed the presence of the problem at the packaging filling lines. Putting the focus of the research here will have the biggest impact, because this section of the process considers the most assets and maintenance tasks. More than 50% of the total number of maintenance activities are dedicated to the bottling and packaging process. There is also the fact that there are five maintenance engineers present in the team all dedicated to the packaging filling lines, and only two for all the other processes, again showing the importance of the packaging filling lines. Therefore, the scope of the research lies at the brewery in Zoeterwoude and focuses on the packaging filling lines. The gathering of knowledge and information can be done at all lines, but calculations and tests should be done at only one. Note that different lines are often used for various product types. This thesis will focus on lines 81 and 82, which are part of the Bottle One Way (BOW) packaging lines, meaning these products go to external markets outside of the Netherlands. Both lines are exactly the same, except for the fact that line 81 has an additional set of machines to package products in boxes. The Maintenance Engineer dedicated to these lines is responsible for the entire process from the moment empty bottles arrive, until the moment filled bottles are put on pallets and continue to the warehouse or transport services. The research will only consider this part of the process. According to the team manager, the fluctuations of production amounts are the largest at lines 81 and 82, considering the whole brewery. Setting the scope of the research to these lines will have the largest effect and lead to the most interesting findings.

1.3.2 Methodology and Research Questions

The main research question of this thesis is:

"How can the maintenance planning activities of the packaging filling lines in the brewery at Zoeterwoude of Heineken be adjusted to change from a time-based to a counter-based strategy?"

The goal of the research is to answer this main research question, solving the selected core problem. This is done by dividing the research in different steps, each acquiring more knowledge on how to arrive to an answer. Figure 1.3 shows the structure of the research and the thesis, where the consecutive stages show the subjects in focus and in what order the research considers them.

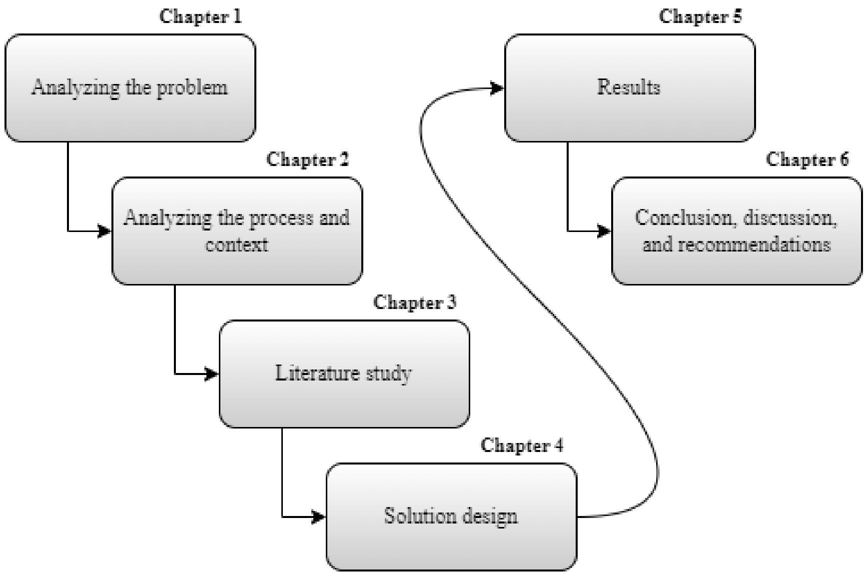


FIGURE 1.3: Research & Thesis Structure

Dividing the main research question into more specific smaller research questions and connecting them to different chapters, creates a simple pathway to arrive at the final answer to the problem. Since the first chapter discusses the analysis of the problem and the creation of the main research question, only the chapters after the first get their own research questions. The following are the specified smaller research questions:

1. "How is Heineken currently planning the maintenance activities at the packaging filling lines based in the brewery in Zoeterwoude?"

The answer of this research question leads to the analysis of the current situation, the current process, and the context of the problem. Chapter 2 delivers this answer. The following sub research questions are useful in this section of the research:

- *Who are the stakeholders of the maintenance activities performed at the packaging filling lines?*
- *What are the differences between the lines, line parts, machines, and which should be considered for the counters?*
- *How does the Maintenance Engineering Team apply maintenance on the packaging filling lines?*
- *How are the current time intervals found of the maintenance activities?*
- *How do the current time intervals perform based on an analysis of a single activity?*
- *What would be an adequate counter for the packaging filling lines?*
- *What quantitative evidence is there that production differs between maintenance activities?*

To answer this research question, and its applicable sub research questions, the research makes use of interviews, organizational documents, and gathered knowledge by working alongside the maintenance engineers of the packaging filling lines.

2. "What is the theoretical background on counter-based maintenance, (dynamic) maintenance planning activities, and prediction models?"

The answer of this research question in Chapter 3 gives the needed theoretical background with a literature study. The following sub research questions are useful in this section of the research:

- *When does theory state that a counter-based maintenance strategy is adequate and in what problem context is it commonly used?*
- *How is the optimal maintenance interval determined in the literature?*
- *How is maintenance planning described in literature, and what models are present in the theory?*
- *How is preventive maintenance of production lines described in literature?*
- *How are prediction models described in literature?*

To answer this research question, and its applicable sub research questions, the research uses a systematic literature review, relevant articles and theory books, and other relevant papers/theses with a similar problem context.

3. "How should this research collect, select, and process the data to create the design of the solutions?"

The answer of this research question in Chapter 4 provides the calculation of the newly introduced maintenance counters of the relevant machines, in addition to the design of both the prediction and planning models. The following sub research questions are useful in this section of the research:

- *What is the approach to arrive at a solution of the problem?*
- *How should the Maintenance Engineering Team calculate the counters, and how often should they perform the calculations?*
- *What data should be collected for this thesis, what data is currently available, and how should the situation of missing data be handled?*

- *How should the Maintenance Engineering Team predict the production hours of the machines on the production filling lines?*
- *How should the Maintenance Engineering Team model the planning of preventive maintenance activities?*

To answer this research question, and its applicable sub research questions, the research collects, processes and selects relevant data from different outputs. Furthermore, it describes the solution models that uses the data.

4. "What are the results of the solution design and how do they affect the performance of the maintenance activities?"

The answer of this research question in Chapter 5 discusses the results of the solution design and how the models have an effect on the performance of the maintenance activities. It serves as a validation of the model and therefore a validation of this research. The following sub research questions are useful in this section of the research:

- *How do the new usage counters improve the current situation?*
- *What are the results of the predictive models?*
- *What are the results of the planning model?*
- *How do the solutions improve the current situation?*
- *How sensitive is the model to new situations?*

To answer this last research question, and its applicable sub research questions, the research creates the models and solve them in an adequate software program (Excel or Python). Furthermore, the research evaluates different performance measures, compares the new with the current situation, and performs a sensitivity analysis.

1.3.3 Deliverables

The deliverables of this bachelor thesis are the following:

- *Literature review that delivers knowledge on a counter-based maintenance strategy*
- *Analysis on usage counters for maintenance activities*
- *Prediction model for upcoming production amounts*
- *Planning tool/model for maintenance activities*
- *Comparison with current situation plus numerical study as validation*
- *Step-by-step method document for the transition of a packaging filling line to a counter-based maintenance strategy*

Chapter 2

Problem Context

This chapter addresses the first research question:

"How is Heineken currently planning the maintenance activities at the packaging filling lines based in the brewery in Zoeterwoude?"

It frames the current situation in the context of the problem. The research question unfolds into several sub-questions, namely:

- *Who are the stakeholders of the maintenance activities performed at the packaging filling lines?*
- *What are the differences between the lines, line parts, machines, and which should be considered for the counters?*
- *How does the Maintenance Engineering Team apply maintenance on the packaging filling lines?*
- *How are the current time intervals found of the maintenance activities?*
- *How do the current time intervals perform based on an analysis of a single activity?*
- *What would be an adequate counter for the packaging filling lines?*
- *What quantitative evidence is there that production differs between maintenance activities?*

Section 2.1 discusses the stakeholders, and Section 2.2 introduces the specifications on the filling lines and its parts/machines. Section 2.3 addresses the maintenance activities and how the company handles them, while Section 2.4 elaborates on the decisions on the current time intervals for maintenance activities. Then, Section 2.5 discusses the performance of the current time-based counters. Section 2.6 considers an adequate counter, and Section 2.7 provides quantitative evidence on the difference in production between maintenace activities. Section 2.8 concludes the chapter.

2.1 Stakeholders

Many departments come in contact with the maintenance activities of the packaging filling lines. Initially, the Tactical Planning Team creates an annual operational schedule indicating when lines will be operational and when they will undergo maintenance activities. The team constructs this planning in collaboration with the maintenance engineers. To execute these activities, spare parts are necessary. The Spare Parts Management Team controls the inventory of spare parts, making sure the right resources are present whenever maintenance activities should be performed. The Spare Parts Warehouse is closely related to the latter, receiving, collecting, and combining these spare parts accordingly. Line Operators and Mechanics are the stakeholders directly involved in carrying out the activities on the line. Maintenance Coordinators and Planners have the specific task to assign mechanics to maintenance tasks and coordinate the activities, while the maintenance engineers are responsible for the application of the right maintenance strategy. Figure 2.1 visualises these stakeholders and which are closely correlated via the dotted circles.

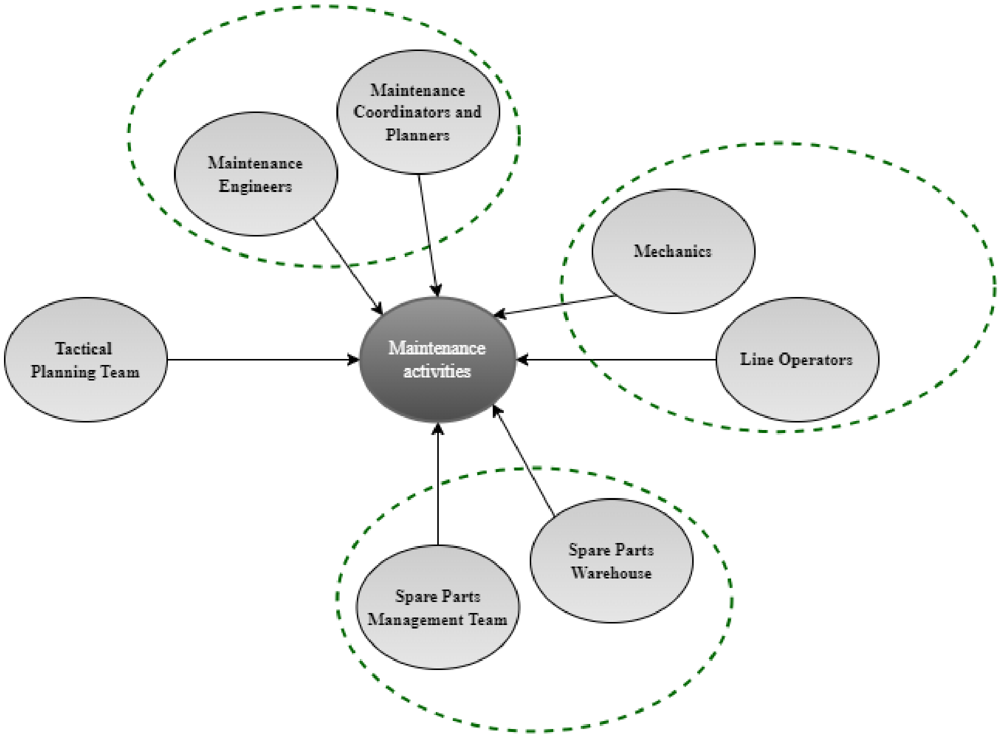


FIGURE 2.1: Stakeholders considering maintenance activities

When this thesis considers decisions on the maintenance activities, it could have an influence on all of the stakeholders. Each party has value in the collection of information in the following sections and chapters. This research aims to collect knowledge, considering all perspectives. Therefore, the previous figure gives the definition of who the stakeholders are, and how they are related.

2.2 Packaging Filling Line

To find the differences between the lines, line parts, machines, and how this affects the maintenance activities, it would be useful to find information on the production system tree of the packaging filling line. Figure 2.2 shows an example, giving the four typical hierarchical levels. Section 2.2.1 discusses the sections of the line, while Section 2.2.2 explains on which levels the counters would be adequate. Section 2.2.3 considers how the company measures the performance of a packaging filling line.

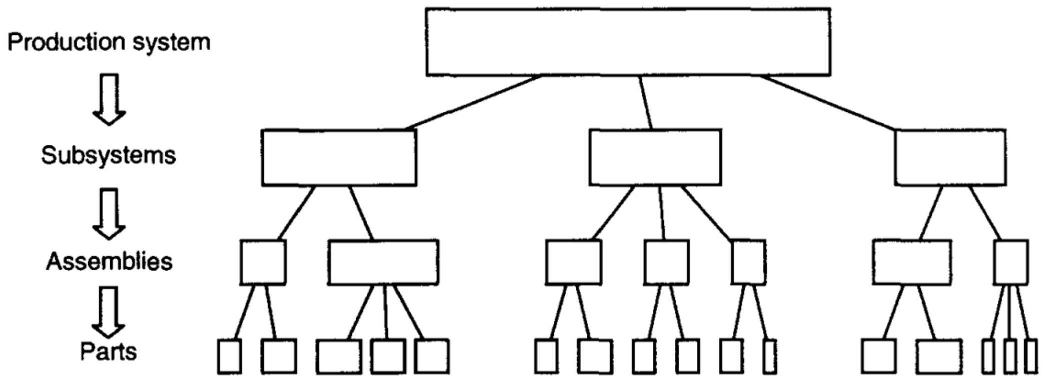


FIGURE 2.2: Example of a production system tree consisting of four hierarchical levels (Dijkhuizen and Harten, 1997)

2.2.1 Line Sections

In the context of this research, the creation of a production system tree would be too extensive since there are hundreds of parts to consider. However, there is information on the general structure of the packaging filling line, and what parts make up these sections. The company calls these parts *assets*. To give some insight into the relevant systems, the line contains the following sections (in between these sections the product is moved on conveyor belts, also called Conveying):

- Unpacking
- Bottle washing
- Bottle filling
- Inspection
- Pasteurization
- Labelling
- Packing

These sections all contain assemblies and subsequent assets. The company uses the software IBM Maximo for the process of monitoring assets. IBM Maximo Application Suite is a set of applications for asset monitoring, management, predictive maintenance and reliability planning. In this software, the Maintenance Engineering Team can find each item of a specific packaging filling line, accompanied by the specific maintenance tasks that correlate to these items. Section 2.3 discusses these tasks further.

2.2.2 Maintenance Structure

A maintenance counter indicates the age of some part of the packaging filling line that is in need of maintenance. Such a counter is chosen for a specific level in the system. For example, the counter could be used for a machine entirely, or for the assets in the machine specifically. To answer which sections and what asset get their own counter, information on the structure of maintenance is required. Not only does the company work with an asset management system, there are more subjects to consider. Subsequently, Figure 2.3 shows the structure of the maintenance tasks.

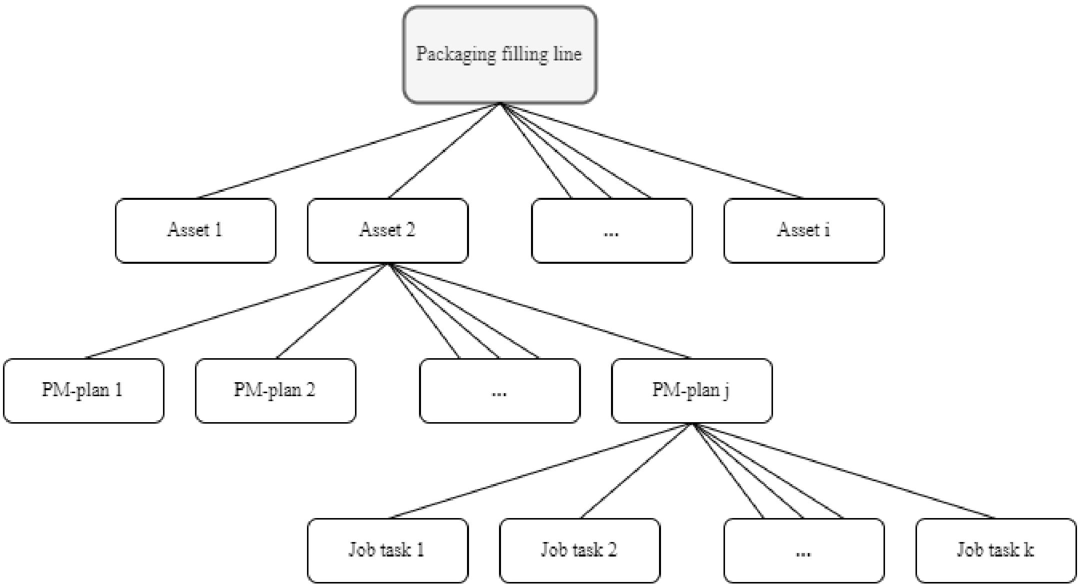


FIGURE 2.3: Structure of Maintenance Tasks

Section 2.2.1 has previously discussed the various sections of the packaging filling line and their respective functions. These sections consist of machinery organized into assets, as depicted in Figure 2.3. The company employs a specific organizational framework whereby each asset possesses its own dedicated Preventive Maintenance (PM) plan. Such a PM-plan contains a detailed outline of job tasks to be executed collectively when the plan is scheduled. These job tasks encompass information regarding spare parts, labor hours, costs, and other relevant details. A PM-plan may encompass one or multiple tasks.

Maintenance engineers utilize these PM-plans to schedule preventive maintenance, ensuring that the prescribed tasks within a PM-plan are consistently performed together. Presently, each PM-plan indicates the time interval for its implementation (refer to Section 2.4). Note that while an asset may have its own PM-plan, it can also be comprised of subordinate assets. For instance, within the packaging filling line, the asset "Filler Rinser" has its designated PM-plans, yet it also incorporates smaller assets like "Fill Valves," each with its own set of PM-plans. These plans may differ in tasks, intervals, and maintenance strategies.

This research chooses to adopt usage-based counters for the PM-plans instead of assets or job tasks, due to the following reasons. Firstly, assigning a counter at the asset level would imply uniformity across all PM-plans associated with that asset, despite potential variations in their purposes. For example, an inspection PM-plan and a replacement PM-plan for the same asset should not have similar usage counter-based intervals. Similarly, assigning counters to specific job tasks would prove complications, as these tasks are grouped together for efficiency, often performed concurrently due to shared preparatory setup activities. Individual counters for each task would disrupt this synchronized approach, thereby undermining the efficacy of the current maintenance structure.

2.2.3 Line Performance

One of the most important ways how the company evaluates the performance of a packaging filling line is by measuring the amount of downtime. In manufacturing, downtime can be categorized into two types: planned and unplanned downtime. Planned downtime refers to periods when production is deliberately halted for reasons such as (routine) maintenance activities, changeovers, setups, employee breaks, or shift changes. Planned downtime is essential for long-term operational health. Unplanned downtime is unexpected and can have dangerous consequences for production schedules. It can occur due to machine failures, material shortages, or unschedules stops (e.g. human error). At the packaging filling line there are two types of unplanned downtime: short and long stops. The company wishes to accomplish two things: optimizing planned downtime to keep it at a minimum, and reducing unplanned downtime to keep operationability at a maximum. Currently, planned downtime takes about 10 to 15 percent of the available production time, and on general workdays there is about 20 to 30 percent of unplanned downtime.

2.3 Maintenance Activities

The team leader of the Maintenance Engineering Team states they apply two types of maintenance activities. The first are *reactive* activities, where the maintenance is a result of a disruption. This type considers aspects, such as corrective maintenance, Break Down Analyses, and Root Cause Failure Analyses. The other type consists of *proactive* activities, where the team tries to control the situation before disruptions occur. This considers other aspects of maintenance, such as revisions, preventive maintenance, condition monitoring. The company also performs some FMECA's but this has not been done for the filling lines in the scope of this research. Section 2.3.1 discusses types of maintenance tasks, and Section 2.3.2 explains the whole process of performing these maintenance activities. Finally, Section 2.3.3 provides the scheduling method of maintenance activities.

2.3.1 Maintenance Types

Depending on the asset (see Section 2.2), some preventive maintenance plans should be performed accordingly. The company uses a code system to distinguish different preventive maintenance plans. Table 2.1 shows these codes, their general task, and how forthcoming they were at line 81 in the past years.

TABLE 2.1: Preventive Maintenance Codes

Code	**Task**	**Division of tasks**			**Average downtime**
		2021	2022	2023	Duration in hours
PM1	Inspection	30%	30%	33%	2:36
PM2	Lubricating	49%	52%	51%	1:31
PM3	Calibration	1%	1%	1%	2:04
PM4	Replacement	7%	6%	7%	11:48
PM5	Revision	12%	1%	13%	15:11
PM6	Condition Monitoring	1%	1%	1%	2:03
PM7	Cleaning	4%	3%	3%	4:14
PM8	Safety & Compliance	7%	5%	6%	6:09

Inspection involves visually examining equipment or systems to detect any signs of wear, damage, or other issues that may require attention. *Lubricating* means applying grease, oil, or other lubricants to machinery components to reduce friction and wear and extend the lifespan of equipment. *Calibration* involves adjusting equipment to ensure it meets specified performance standards or accuracy requirements, which is crucial for measuring devices to maintain precision in their readings. *Replacement* implies replacing worn-out or damaged parts with new ones to maintain the functionality and safety of equipment. *Revision* entails a more thorough inspection that takes up more time due to the extensive check on all parts of an assembly. *Condition Monitoring* involves using various techniques such as vibration analysis, or temperature readings to assess the health and performance of equipment. *Cleaning* means removing dirt, debris, or other contaminants from equipment or systems to prevent interference with their operation. *Safety and Compliance* involves ensuring that equipment and maintenance procedures comply with safety regulations and standards. Safety inspections and compliance checks help prevent accidents and ensure a safe working environment. All the PM tasks are considered scheduled preventive maintenance tasks, except for PM6, which is (predictive) condition-based maintenance.

Next to these PM tasks, there is a revision each two years. These are the maintenance activities performed during an extensive revision. Such a revision is planned in cooperation with the Tactical Planning Team, often takes op to 5 to 10 workdays, and considers the whole packaging filling line. There is an inspection on all assemblies and their subsequent assets, and whenever necessary, the mechanics replace the asset. Line 81 had a revision in 2021 and 2023, and this explains the increase in the proportion of PM5 tasks in these years (see Table 2.1).

Since the Heineken brewery in 's-Hertogenbosch already performs a partially counter-based maintenance strategy, plentiful communication with their maintenance engineers is appropriate. In one of the interviews, the maintenance engineers advise to only consider the first five PM tasks (PM1 to PM5), because of the following reasons: (i) Condition Monitoring means that the proactive decision if a task should be performed, depends on the state of what the sensors measure, not on the counter of the machine. (ii) The necessity of Cleaning tasks does not depend on the amount of production time, but depends on the total time that has passed. (iii) Safety and Compliance tasks are necessary whenever the regulations state that they are necessary, and this should not depend on the amount of production of a line. To conclude, it would not be beneficial for the company to adjust the maintenance strategy for the tasks PM6 to PM8 to a counter-based maintenance strategy. Both considering the previous arguments and the divisions of PM tasks in the previous years, this research decides to focus on the tasks with the code PM1 to PM5. This still leaves more than 80% of the PM tasks to consider, creating enough positive impacts for the company. Note that the PM4 and PM5 tasks also consider the largest amount of average planned downtime.

2.3.2 Maintenance Process

To understand how the Maintenance Engineering Team applies the maintenance activities, a workflow diagram shows the details of the process. It shows actions, activities, responsibilities, and relations. Figure 2.4 shows this workflow diagram.

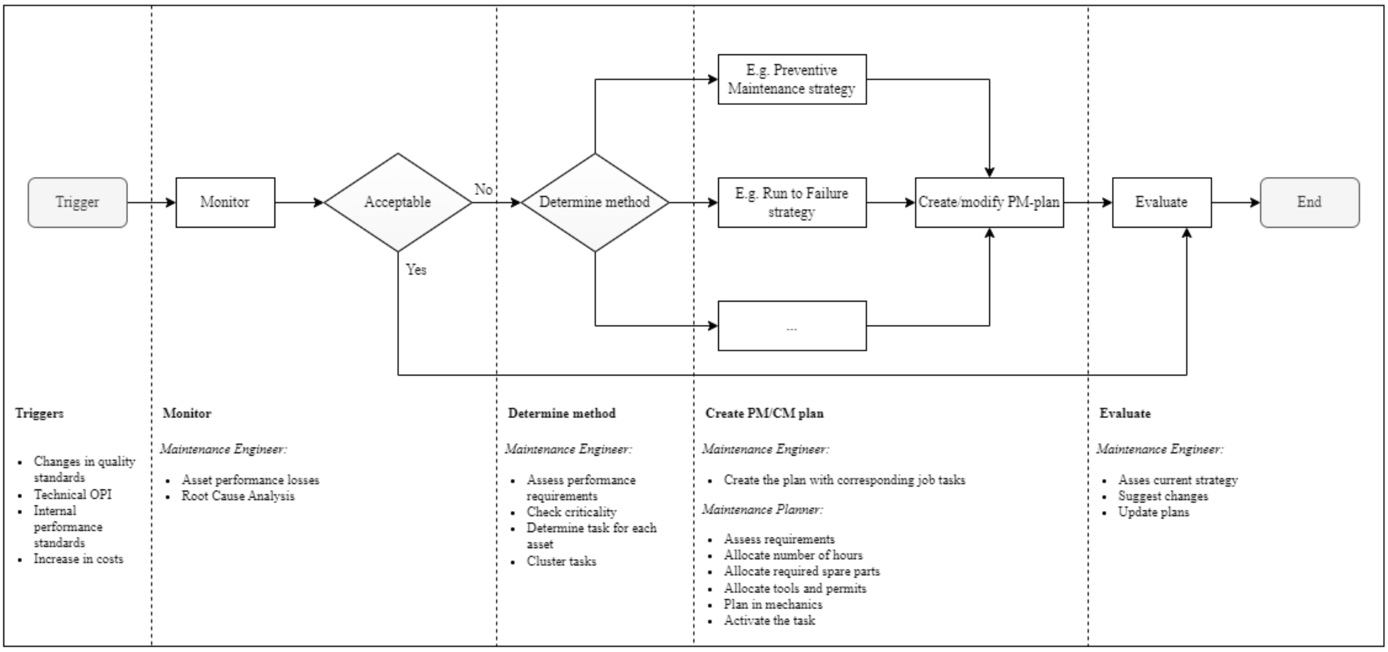


FIGURE 2.4: Workflow of Maintenance Process

Triggers initiate the activities of the Maintenance Engineering Team. These triggers may include alterations in quality or internal performance standards, escalations in costs, or shifts in the technical Operational Performance Indicator (OPI). In response to these triggers, the maintenance engineer starts monitoring activities, which involve analysing asset performance or initiating a Root Cause Analysis. This monitoring includes the assessment of the time-based intervals of the PM-plans. Based on this assessment, the maintenance engineer determines whether the performance meets acceptable criteria. If deemed acceptable, they proceed directly to evaluation. If not, the maintenance engineer selects an appropriate improvement method. Depending on the circumstances, a maintenance strategy is chosen, taking into account the general prerequisites and criticality of the tasks involved. Subsequently, the maintenance engineer digitally creates the plan, while the Maintenance Planner evaluates specific requirements such as spare parts procurement, task assignment to mechanics, and task scheduling. Ultimately, the evaluation phase starts, during which the maintenance engineer evaluates the current strategy, proposes alterations, and updates the Maintenance Plan as necessary. This workflow shows how the PM-plans are used, and when the maintenance engineer monitors the time-based interval.

2.3.3 Maintenance Planning

The company uses a simple method to schedule the preventive maintenance activities. Figure 2.4 shows that the Maintenance Planner considers all requirements of the tasks and then schedules them whenever possible. Currently, three months prior to the deadline of the time-based interval, the maintenance engineer activates the PM-plan (if it is not activated automatically), resulting in the Maintenance Coordinator and Planner checking the requirements and scheduling the mechanics. These preventive maintenance activities can only be scheduled during revision or on what the company calls *stop-days*. A stop-day is a specific day on which production is (partially) stopped to perform maintenance activities. There are short and long stop-days, which alternate every week to provide some possibilities for maintenance activities. Each packaging filling line has a yearly schedule showing when the line is in revision, and when the stop-day are planned. The Tactical Planning Team and Maintenance Engineering Team work together prior to the start of the year to finalise this schedule. Figure 2.5 shows a fabricated example of such a stop-day planning. Note that there is no planning model in use for the planning of these preventive maintenance activities. It is all done by hand by the Maintenance Planners.

Date	Col. 81	Col. 82
1-01-25		
2-01-25	81	
3-01-25	81	
4-01-25	81	82
5-01-25	81	
6-01-25		
7-01-25		
8-01-25		
9-01-25		
10-01-25		
11-01-25	81	82
12-01-25		
13-01-25		
14-01-25		
15-01-25		
16-01-25		
17-01-25	81	
18-01-25	82	82
19-01-25		
20-01-25		
21-01-25		

Legend	
	Stop-day
	Short stop-day (e.g. for cleaning)
	Extra stop
	Revision
	Vacation
	Public holidays
	IT / Brewery stop

FIGURE 2.5: Fabricated example of stop-day planning of packaging filling lines 81 and 82

As is common in the industry, the company uses three management strategy levels to address the planning of production; the operational level, the tactical level, and the strategic level. The operational level deals with day-to-day activities and short-term decisions, with a time horizon of two weeks. The tactical level involves medium-term planning to optimize efficiency and resource use, with a time horizon of 12 to 13 weeks. The strategic level focuses on long-term direction and growth opportunities, looking at multiple years. Regarding the planning of preventive maintenance, only the operational level is in contact with the Maintenance Planner, weekly sending the operational schedule for the upcoming two weeks every week. This gives the Maintenance Planner opportunities to reschedule both corrective and preventive maintenance activities. Rescheduling happens a lot due to clustering opportunities, or unexpected changes in the availability of mechanics or spare parts. The tactical production planning is not considered in the planning of preventive maintenance.

2.4 Maintenance Time Intervals

Inside IBM Maximo, the application software for maintenance, each maintenance activity contains a specified time-based interval. For example, this could be eight weeks (code: 8W), three months (code: 3M), or two years (code: 2Y). This time-based interval indicates how much time is in between the execution of a specific PM-plan. These intervals are chosen according to the combination of the experience of the stakeholders, and the advice given by the machine suppliers. For example, somewhere a few years ago a new machine was added to the packaging filling line, and the advise given was to replace a certain part every 8 weeks. This was followed by the stakeholders and they found out that they performed more corrective maintenance then they replaced the part proactively. Therefore, they changed the interval to 6 weeks and they kept this interval ever since. In conclusion, the present time intervals are determined based on recommendations from machinery suppliers, augmented by the company's own experiential insights.

The current time-based maintenance intervals for the preventive maintenance tasks vary from one week to eight years. This simply depends on the type of task of the preventive maintenance activity. For example, a common activity would be an inspection plan (PM1), taking place every one or two weeks. A less common activity would be the replacement of valves (PM4), taking place only once every two years. For the Maintenance Engineering Team to monitor the counters of all PM-plans that are very common would take too much of their time. Moreover, maintenance tasks are always planned on a stop-day (see Section 2.3.3), meaning the shortest amount of time between maintenance activities that is possible would be one week for short tasks and two weeks for longer tasks. Therefore it is not necessary to investigate if a PM-plan should be planned on a shorter notice than one or two weeks. Therefore, this research continues to focus on the preventive maintenance tasks with their current time-based interval larger than two weeks. In other words, the company should not alter the maintenance strategy of preventive maintenance activities with a current time-based interval of one or two weeks.

2.5 Performance Time Intervals

The importance of this section transpires from the assumption that the current time-based intervals are adequate enough to translate them to counter-based intervals, made by the manager of the Maintenance Engineering Team. To simply take this statement as true is unscientific. Therefore, a thorough assessment on the current time-based intervals should take place. To do this for all PM-plans is time extensive labor and not the focus of this research. Consequently, this research takes one PM-plan and performs an analysis on its interval to attain a thorough evaluation of the adequacy of the assumption.

The chosen PM-plan has the following description: PM4 VULM81 Kaarsfilter 4W. This is a replacement activity of specific filters in the filler/rinser machine. Currently the time-based maintenance interval is set at four weeks, making this PM-plan a good candidate for this analysis because it has been executed many times in the past years. This section collects data from SAP to find the numbers of executions in the past four years of the PM-plan, the corresponding corrective maintenance actions of this part specifically, and the total corrective maintenance actions on the entire asset. Figure 2.6 shows these numbers.

Figure 2.6 displays an increase of executions of the PM-plan over the years while the amount of corrective activities remains relatively stable. The Maintenance Engineering Team decreased the time-based maintenance interval of this PM-plan which causes this rise. The argument for this modification of the interval is based on the

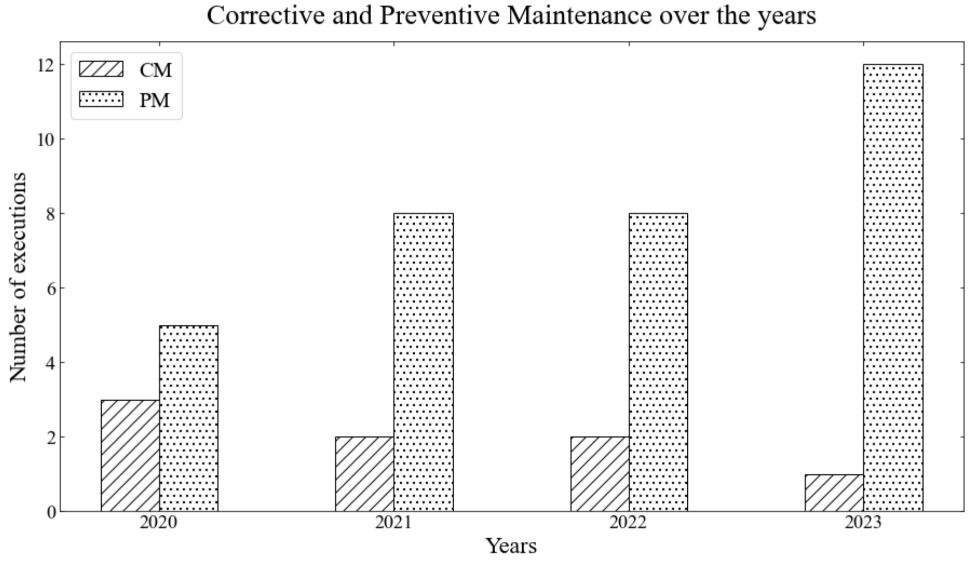


FIGURE 2.6: Corrective and Preventive Maintenance over the years

experience of too much corrective activities. Costs also show that the modifications were justified. Both the packaging manager and the software systems deliver information on corrective and preventive maintenance costs for this PM-plan specifically. Table 2.2 shows how the maintenance costs and unplanned downtime (UPD) change over the years.

TABLE 2.2: Corrective and preventive maintenance costs

Year	**2020**	**2021**	**2022**	**2023**
CM Costs	€ 3,014.25	€ 2,009.50	€ 2,009.50	€ 1,004.75
PM Costs	€ 983.40	€ 1,573.44	€ 1,573.44	€ 2,360.16
Total Costs	€ 3,997.65	€ 3,582.94	€ 3,582.94	€ 3,364.91
UPD	41%	37%	35%	31%

Table 2.2 displays how the Maintenance Engineering Team responds to an excess of corrective maintenance actions. After 2020 with a high amount of corrective maintenance, the team responds with an increase of preventive actions by lowering its time-based maintenance interval, ultimately lowering the total costs of maintenance in the year. Although the maintenance engineers did not evaluate the PM-plan with using a fitted failure distribution and evaluating the corrective/preventive costs, they change the strategy of a PM-plan according to their experience and the advice from the OEM of the machine. Note that the unplanned downtime (UPD) of the whole machine decreases over the years, indicating a positive performance that further justifies the adjustment of the time-based interval. This one example establishes why the knowledge from employees of the company and OEM should not be ignored, and why the current time-based intervals show to be adequate enough to be translated to usage-based counters.

2.6 Counter Measure

The company makes use of Total Productive Management (TPM) as their approach, also known as Total Productive Maintenance. TPM activities focus on eliminating the following six major losses (Rausand, Barros, and Høyland, 2021, p. 396):

- Equipment failure (breakdown) losses
- Setup and adjustment losses
- Idling and minor stoppages
- Reduced speed losses
- Defects in process and reworking losses
- Yield losses

Figure 2.7 shows how these six major losses affect the time concepts used to evaluate the performance of a packaging filling line. This extends the concept of downtime from Section 2.2.3. These time concepts are relevant because the company gathers this data to evaluate the production hours of the machines on the packaging filling line. From these time concepts a decision on the measuring unit can be made for the usage-based counters.

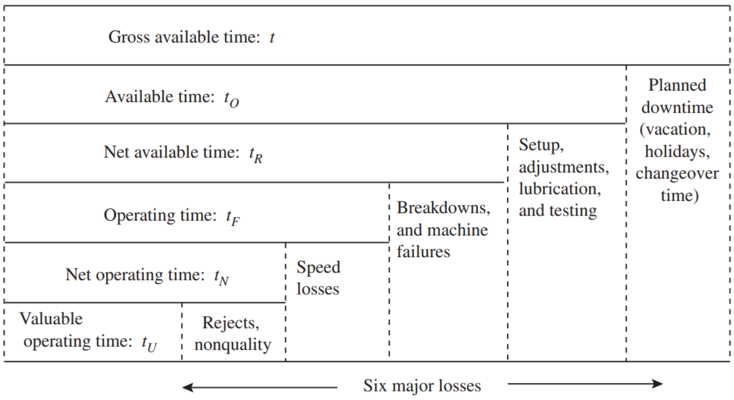


FIGURE 2.7: Time concepts of the six major losses in TPM (Rausand, Barros, and Høyland, 2021, p. 397)

Some engineers would argument that the Net operating time is an adequate unit to measure operating hours per machine, but in the context of this research, this does not consider all states of the machine. Therefore, in addition to these time concepts, the company gathers information per machine on its state. The following states are known:

- Emptying idle time
- Filling idle time
- Production time
- Production stops
- Malfunction time

The *Emptying idle time* entails the period when the machine is not actively engaged in production due to the absence of products from preceding stages in the production process. This happens, for example, at the end of a shift when the last products are moving through the line. The *Filling idle time* is the opposite, being the period when the machine is actively working but cannot get rid of its discharge due to malfunction or start ups downstream the production line. The *Production time* denotes the time during which the machine is actively engaged in its primary task, with no malfunctions. The *Production stops* are the instances when the machine halts its production activities temporarily, where it is deliberately stopped by the operator. The *Malfunction time* signifies periods when the machine is unable to perform its intended functions due to technical issues or malfunctions. The difference with Production stops is the Malfunction time considering unplanned interruptions.

The Manufacturing Execution System collects per day the total time of all machines in each state. In a thorough investigation by the Maintenance Engineering Team, a final decision was made on the best representation of real life operating hours, meaning the machines are actively moving around. This being the Production time, the Emptying idle time, and the Filling idle time. In these states the machines are up and running and the assets are in use, meaning they are operating and moving. The other states represent the time where the machine is standing still. Considering this decision of the team, the most adequate counter for each machine would be the sum of the Operating time, the Emptying idle time, and the Filling idle time. This research defines this sum as Real Operating Time.

2.7 Quantitative Evidence

The company has historical data, giving information on all executed preventive maintenance activities in the past, their PM-code, a description, and its execution date. The goal is to combine this knowledge with the historical output on the Real Operating Time from the Manufacturing Execution System, to finally evaluate the differences in production times between maintenance activities for every PM-plan. Section 2.7.1 provides evidence showing that production varies throughout the years, while Section 2.7.2 gives an analysis on all PM-plans of line 81, indicating the differences in production times between maintenance activities. Lastly, Section 2.7.3 analyzes the multipacker machines of both line 81 and 82 to show the historical differences between counter-based maintenance and time-based maintenance.

2.7.1 Production Variation

For one machine on line 81, Figure 2.8 shows the production times in minutes per week of years 2022 and 2023. It perfectly represents how production activities differ throughout the years and change dynamically. This fact motivates this research to further investigate how the amount of production between the execution of PM activities varies.

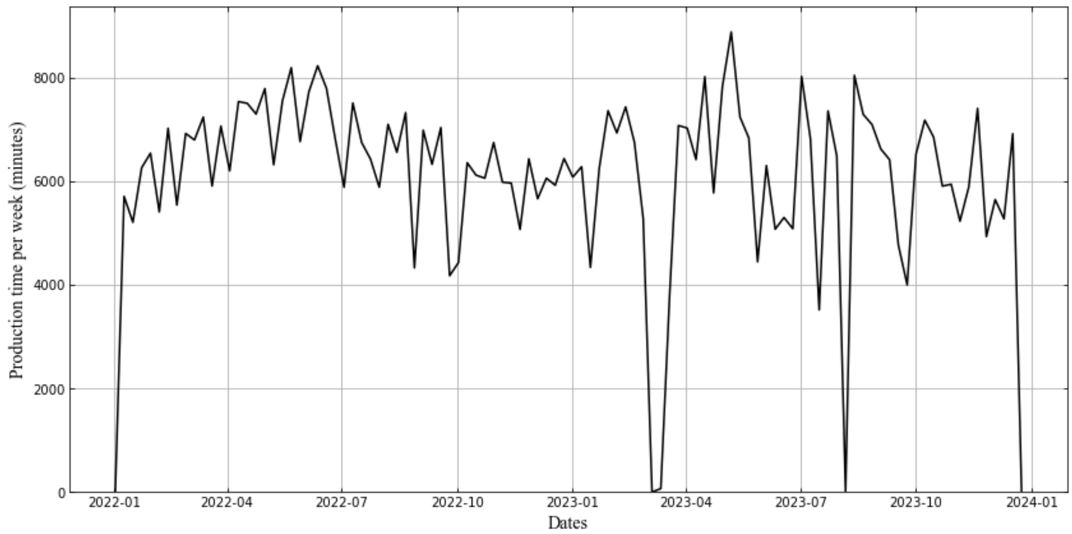


FIGURE 2.8: Production per week for one machine on filling line 81

2.7.2 PM-plans Analysis

Subsequently, a box plot visualises the differences in production times between maintenance activities, giving an indication on the range of differences in production hours, and showing outliers when present. To finalise this quantitative evidence, only the box plots of the PM-plans that are still currently in use are valuable. This results in a total of 101 box plots. Appendix A shows all box plots, and Figure 2.9 shows the first ten PM-plans found.

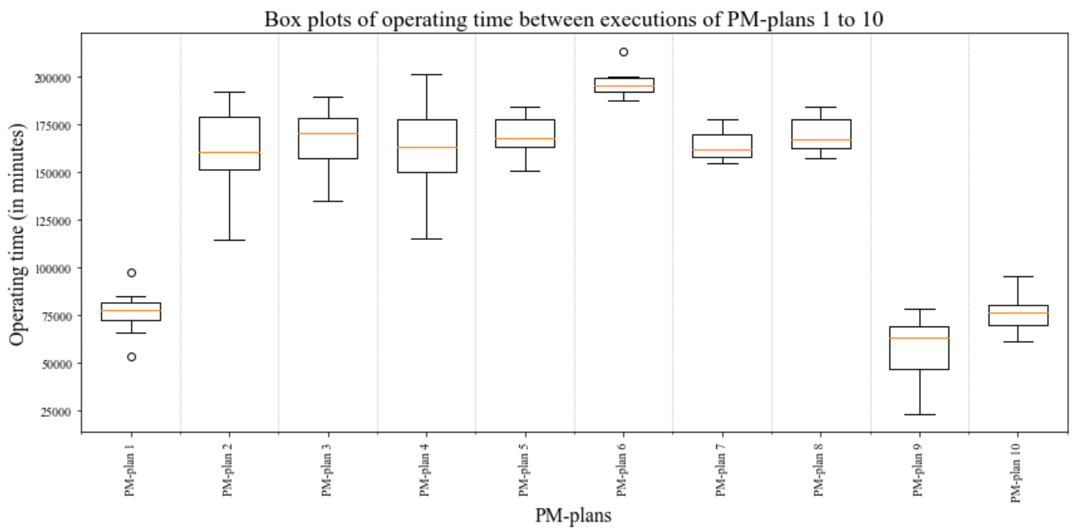


FIGURE 2.9: Box plots showing differences in production time between executions of PM-plans

Figure 2.9 reveals that there is a significant difference in Real Production Time between the past executions of the maintenance activities. The differences in the time-based intervals chosen for the PM-plans shows itself by having different ranges across all box plots. What is consistent for most PM-plans is that there is large range, meaning the amount of Real Operating Time between maintenance activities of these PM-plans has not been constant in the past. The ranges for PM-plan 4 differs from 125,000 to 200,000 minutes. The difference, 75,000 minutes, means that at one point, the machine corresponding to this PM-plan had been operating for about 2,000 hours since its last preventive maintenance activity, and at another point, it has been operating more than 3,000 hours since that same preventive maintenance activity. The difference is a result of fluctuations in production demands, differences in operational practices and scheduling, variability in maintenance planning, and potential inaccuracies in the data. It is clear quantitative evidence that the production on the packaging filling line differs between maintenance activities. With the current time-based maintenance strategy, the operating hours of the machines diverge with significant amounts between specific preventive maintenance executions.

2.7.3 Multipackers Analysis

One of the maintenance engineers started a project to apply a counter-based strategy more than eight years ago. This counter-based system was created for only the multipacker machine of packaging filling line 81. As mentioned earlier, line 81 and 82 are close to identical. This past project creates a unique situation where this research can analyse the differences between the multipacker machines of both lines, considering their preventive maintenance. The analysis includes four components of the multipacker, each having their own preventive maintenance inspections. Table 2.3 gives some information on the number of executions (in 2019 to 2023) of the PM activity per component of each line, while also giving the set interval of these PM-plans.

TABLE 2.3: Multipacker information per line

	Line 81 (counter-based)		**Line 82 (time-based)**	
Component	**Executions**	**Interval (hours)**	**Executions**	**Interval (months)**
Inserter	8	750	34	2
Cluster entry	4	1500	19	3
Bottle entry	3	2250	10	6
Emitter	4	1500	19	3

Immediately, one can observe that the amount of executed PM activities is larger for line 82. In 2019 to 2023, the multipacker of line 82 has been operating in total for about 21.000 hours, while this machine line 81 operated only for about 8500 hours. Because line 82 was busier, the maintenance engineer took a more safe and conservative approach for line 82 and set the interval lower, which lead to the increase of PM activities. To further analyze the differences, Figure 2.10 shows the averages, minima, and maxima of the values indicating the amounts of production between the executions of PM. The figure shows these values for the four components of the multipacker machine of both lines.

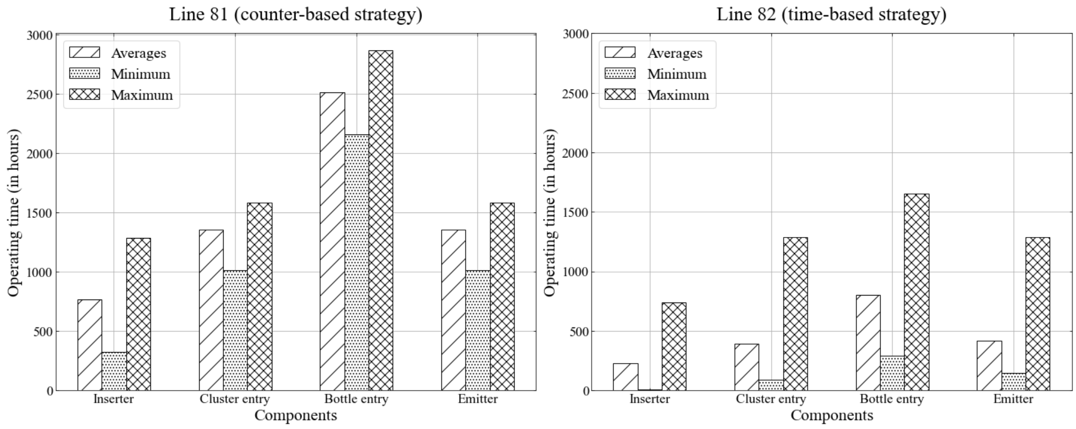


FIGURE 2.10: Bar charts showing differences between the multipackers of both lines

Again, the safe and conservative approach of the maintenance engineer is present in Figure 2.10, because the averages of the line 82 multipacker are lower in comparison to line 81. Still, considering that the proportion of corrective maintenance through these years between both multipackers has been stable, these graphs show that introducing a counter-based strategy could reduce the total amount of PM activities with no impact on corrective activities, thereby reducing the amount of over-maintenance. Additionally, the spread between the minimum and maximum value of the amount of production between maintenance activities is lower for the counter-based strategy, indicating that there is less variety of wear and tear for the components when the company applies preventive maintenance. Also the averages of line 81 in Figure 2.10 are close to the set intervals from Table 2.3, indicating that the maintenance engineers successfully applied the counter-based strategy.

Still, a link is missing to the amount of unplanned downtime (UPD) between these two machines. Table 2.4 shows the percentage of UPD of the total available production time over the years 2019 to 2023.

TABLE 2.4: Unplanned Downtime over the years

Line	2019	2020	2021	2022	2023
Line 81 (counter-based)	42%	38%	38%	36%	33%
Line 82 (time-based)	38%	41%	34%	33%	34%

The UPD values over the years are very similar for both the multipackers. Considering that the number of time-based maintenance activities on line 82 were more prevalent, while the UPD was similar to the line with a counter-based strategy, it suggests that there is a presence of over-maintenance. The introduction of a counter-based strategy leads to less preventive maintenance costs, an increase of efficient and effective maintenance, all while the percentage of unplanned downtime remains at its current value. To conclude, in spite of the fact that line 82 was to produce more products during these years, the past project of applying a counter-based strategy has shown some improvements in reducing over-maintenance and decreasing the variety in the amount of production between maintenance activities.

2.8 Conclusion

The examination into Heineken's current approach to maintenance activities at the packaging filling lines in Zoeterwoude has revealed a multifaceted landscape involving various stakeholders, maintenance types, time intervals, and performance evaluation metrics. Through an in-depth exploration of these elements, an answer to the first research question emerges; Heineken currently plans maintenance activities through a combination of reactive and proactive measures, guided by preventive maintenance codes and time-based intervals. Collaboration between stakeholders, such as the Tactical Planning Team with the Maintenance Engineers and Planners, ensure the smooth execution of these activities. However, the existing time-based maintenance strategy, while incorporating insights from machinery suppliers and historical experience, exhibits significant variability in production hours between maintenance executions.

Quantitative evidence received from historical data highlights this variability, showcasing substantial differences in Real Production Time between successive maintenance activities. These discrepancies show the limitations of the current time-based maintenance approach, indicating a need for a more dynamic and context-aware strategy. Additionally, a past project of applying a counter-based strategy has shown some improvements in reducing over-maintenance and decreasing the variety in the amount of production between maintenance activities, with no negative effect on the amount of unplanned downtime.

In conclusion, Heineken's current maintenance planning methodology, while comprehensive, lacks the adaptability to account for the nuanced operational dynamics of the packaging filling lines. Moving forward, there's a clear imperative for the integration of usage-based counters, specifically focusing on Real Operating Time, to optimize maintenance scheduling and enhance operational efficiency. By embracing a more nuanced approach that considers machine states, production losses, and historical performance data, Heineken can elevate its maintenance planning to better align with the demands of its production environment.

Chapter 3

Literature Review

This chapter delves into addressing the second research question:

"What is the theoretical background on counter-based maintenance, (dynamic) maintenance planning activities, and prediction models?"

It explores current knowledge on failure distributions, maintenance strategies, and planning models across various timeframes. The research question unfolds into several sub-questions, including:

- *When does theory state that a counter-based maintenance strategy is adequate and in what problem context is it commonly used?*
- *How is the optimal maintenance interval determined in the literature?*
- *How is maintenance planning described in literature, and what models are present in the theory?*
- *How is preventive maintenance of production lines described in literature?*
- *How are prediction models described in literature?*

Section 3.1 introduces counter-based maintenance, and Section 3.2 tackles the subject of determining the optimal maintenance intervals. Section 3.3 addresses the planning models present in current theory, and Section 3.4 considers the maintenance of production lines specifically. Then, Section 3.5 discusses prediction models, and Section 3.6 concludes the chapter.

3.1 Counter-based Maintenance

Rausand, Barros, and Høyland (2021) outlines different types of preventive maintenance (PM) tasks: age-based, clock-based, condition-based, opportunity-based, and overhaul tasks. Companies are increasingly adopting degradation-based PM tasks due to enhanced data collection. Each PM task type is explained with examples such as age replacement policies, clock-based maintenance, and condition-based maintenance. Clock-based tasks occur at set calendar times (time-based maintenance), while age-based tasks occur at a specific item age (counter-based maintenance). A counter-based strategy is advisable when failure costs exceed planned replacement costs and when item failure rates rise. In general, a counter-based strategy is more complex due to extra data collection and monitoring activities, a necessity for more intricate prediction methods, and a higher dependence on technology in the context of its application.

The article of Tinga (2010) contains some statements on the effectiveness and efficiency of maintenance strategies. Effective maintenance entails the prevention of system failures or breakdowns. In calendar time-based maintenance, this means setting adequate intervals to prevent failures, which becomes challenging when system usage varies over time. If intervals are not correctly set to address the system's needs, maintenance becomes ineffective. Efficient maintenance occurs when desired outcomes are achieved with minimal waste. Time-based maintenance aims to minimize unnecessary actions, but in uncertain usage scenarios, it can lead to inefficient resource allocation. Usage-based maintenance eliminates uncertainty, allowing more precise intervals tailored to the system's needs, reducing both failure risks and unnecessary actions, subsequently improving the effectiveness and efficiency.

To address the fatigue of items, Tinga (2013) states that the following sources cause the uncertainty in the state (considering its damage) of a given item at a certain point in time:

- variations actual usage
- variations in effect of usage on (internal) loads
- variations in the life consumption for a given internal load

By evaluating the actual usage over time with a counter-based strategy, it is possible to reduce this uncertainty in the amount of damage of a given item. Tinga (2013) states that going from a calendar time-based system to a usage counter-based system results in an increase of a component's service life. It shows that with this new strategy, replaced components are more damaged at replacement, reducing remaining lifetime spillage (over-maintenance). Incorporating usage data changes replacement timing from calendar time to equivalent hours, with operating hours often more relevant than calendar time for industrial machinery.

Furthermore, Tinga (2010) delves into maintenance methods: calendar time-based (CTBM), usage-based (UBM), usage severity-based (USBM), load-based (LBM), and condition-based (CBM). CBM is most efficient, but UBM and LBM are useful without sensors or system accessibility. Tinga (2013, p. 204) states that, considering the results of a case-study, transitioning from CTBM to UBM notably reduces replacements with about 30%, especially during low operating hours, without changing failure probability. Traditional calendar-based maintenance may not suit systems with variable usage or failure patterns, favoring UBM and LBM for accuracy and efficiency. Figure 3.1 shows the results of this case study, indicating the differences between the maintenance strategies.

Christer and Doherty (1970) suggests shifting from calendar-based to overhauls ba— sed on the tonnes of steel throughput. The suggested solution, an example of maintenance in the steel production sector, is robust enough to accommodate interruptions in the overhaul scheduling without losing effectiveness. The study suggests that the transition leads to cost savings in thousands of pounds per year. Studies like Cichelli (1977) advocate for usage-based over calendar time-based systems, showing productivity gains via simulation models. Liu, Wang, and Tan (2024) compare calendar-time-based and age-based maintenance, considering repair times and factors like replacement cost and repair effectiveness.

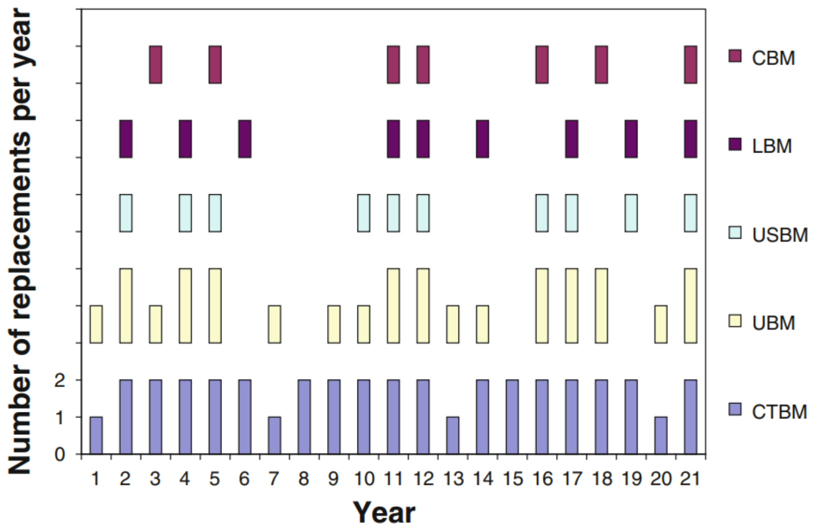


FIGURE 3.1: Overview of item replacements for five different maintenance strategies (Tinga, 2013, p. 204)

Studies by Kim, Ahn, and Yeo (2016) and Ahmad and Kamaruddin (2012) indirectly discuss usage counter-based maintenance while focusing on condition-based methods. By highlighting the condition-based strategy's advantages in maintaining condition under consistent inspection, it suggests the importance of considering usage patterns in maintenance strategies.

The study of Deloux, Fouladirad, and Bérenguer (2016) extends the concept of cou—nter-based maintenance by exploring maintenance policies that consider both deterioration level and usage profile, suggesting a nuanced approach to maintenance decision-making. By incorporating usage data into maintenance policies, it aligns with the principles of usage counter-based maintenance, which emphasizes the importance of usage patterns in determining maintenance needs.

Wang (2002) categorizes maintenance policies for deteriorating systems, while Tiddens, Braaksma, and Tinga (2023) offer a decision framework for predictive maintenance method selection. Practitioners still often follow a costly trial-and-error process in selecting the most suitable predictive maintenance method, and the article delivers a framework to support asset owners in selecting the optimal predictive maintenance method for their situation.

Overall, the theory suggests that a counter-based maintenance strategy is adequate when failure costs outweigh planned replacement costs, when traditional time-based approaches are ineffective due to variable usage or failure patterns, and when usage data is critical for accurate maintenance decision-making. It is commonly used in industries where equipment up-time is critical, and the consequences of failure are significant, such as manufacturing, transportation, and energy sectors.

3.2 Optimal Maintenance Interval

The book of Rausand, Barros, and Høyland (2021) gives the method of calculating the optimal replacement interval for any measuring unit of time. This may be measured by many different time concepts, such as calendar time, time in operation, number of work cycles, and so on. The strategies that the book considers are Age Replacement, Block Replacement, and P–F Intervals. The methods of the book can only be applied when a failure distribution can be fitted to the failure behaviour of an item. Determining the maintenance interval requires failure data, and the book delivers many other sources for this if it is missing.

A maintenance optimization problem that involves only one component with an increasing failure rate in time has a common formula for the optimal interval of preventive maintenance activities. Note that this formula finds the optimal interval over an infinite time horizon using the Weibull distribution and the costs for corrective and preventive maintenance. Tan and Kramer (1997) delivers this general formula:

$$C(T^*) = \frac{C_{cm}F(T^*) + C_{pm}(1 - F(T^*))}{\int_0^{T^*}(1 - F(t))\,dt} \tag{3.1}$$

where, $C(.)$ is the cost rate function, C_{cm} are the corrective maintenance costs, C_{pm} are the preventive maintenance costs, $F(T)$ is the cumulative failure distribution, and T^* is the optimal interval.

The book of Tinga (2013) gives further categorization of the several approaches that can be followed to determine the preventive maintenance intervals. Figure 3.2 shows this categorization, where the moment in life cycle represents the specific stage of the component's life where the maintenance intervals are determined. The condition assessment is the method used to determine the system condition during the service life. The final criterion considers the prognostic approaches used to predict the future behavior or condition of a system or component.

Percy (2008) reviews basic models for complex repairable systems, explaining their use for determining optimal PM intervals. The book gives a summary of models with corresponding sources in the literature. These models contain the Renewal process, Nohomogeneous Poisson process, Delayed renewal process, and many more. Again most of these models use the distribution of failures for components.

Holland and McLean (1975), Basker and Husband (1978), Siswanto and Kurniati (2018), and Sharma and Rai (2021) all perform a case study or practical application on determining the optimal preventive replacement policy. All show significant changes in performances such as improvement in availability, reduction of total cost per unit time, and annual savings.

To sum up, the theory shows that through data collection, modeling, and optimization, there are many methods and examples of arriving at the optimal interval for preventive maintenance activities. This interval may be measured by different time concepts, such as calendar time, or time in operation. Literature seems to lack knowledge on the process of transitioning specifically, going from a calendar time-based to a usage counter-based strategy. Nothing was found on general rules, regulations, or methods of addressing the calculation of intervals with an absence of failure data. Jonge et al. (2015) do deliver an analysis on the influence of uncertainty in failure distribution parameters on the optimal maintenance interval.

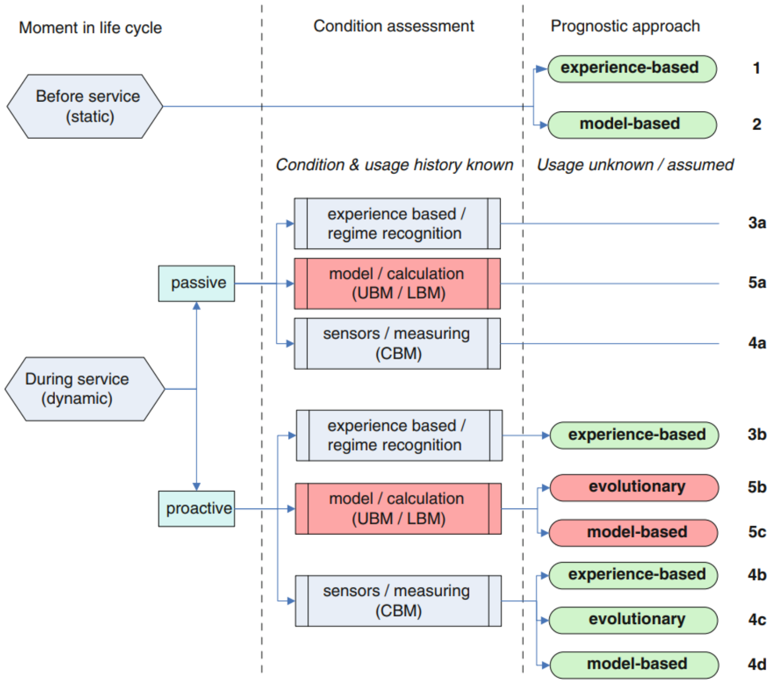


FIGURE 3.2: Classification of preventive maintenance policies (Tinga, 2013, p. 170)

3.3 Maintenance Planning Optimization Models

Many review articles have been written on maintenance optimization models. For example, Dekker (1996) gives an overview of applications of maintenance optimization models, analyzing their role and discussing the factors which may have hampered applications. The review of Ben-Daya and Rahim (2001) contains more specifics on the aspects of production, quality, and maintenance, and which mathematical models in literature aim to integrate these issues. Finally the thesis of Budai-Balke (2009) discusses a review of maintenance planning models in different business sectors.

Several studies address optimal maintenance policies and their operational schedules for various systems. Dedopoulos and Shah (1995) analyze preventive maintenance parameters for equipment in multipurpose plants, balancing maintenance benefits with costs. They integrate production and maintenance planning to optimize maintenance policies. Vatn, Hokstad, and Bodsberg (1996) present a flexible approach for determining maintenance schedules considering safety, cost, and production objectives, adaptable to resource availability and management priorities. Frost and Dechter (1998) frame preventive maintenance scheduling for power generating units as constraint satisfaction problems, aiming to minimize operating and maintenance costs over a planning period. Vaurio (1999) develops cost functions for components subject to random failures, considering periodic inspections and preventive maintenance to minimize total cost rate, accounting for repair, maintenance, and production losses. Dijkhuizen (2000) introduces a hierarchical model for clustering preventive maintenance jobs in multi-component production systems, aiming to find maintenance frequencies that minimize average cost per unit of time. Haghani and Shafahi (2002) propose a mathematical programming approach to scheduling bus maintenance, optimizing daily inspection schedules to minimize interruptions in bus operations and maximize system reliability. Fokkert et al. (2007) creates with the help of operations research techniques a Mixed-Integer-Programming Model to arrive at a maintenance schedule for the rail-track activities of ProRail, showing the workload distribution across several weeks. Their research includes the consideration of safety factors, and the company has accepted the resulting schedule, which has been in operation since the year 2000.

Furthermore, some studies present mathematical models focusing on maintenance schedules with a larger time horizon, representing time tables up to one year in the future. Alardhi and Labib (2008) presents a method for solving a maintenance scheduling problem and the method has been illustrated for a co-generation plant in Kuwait. The basic idea of the method is to model the problem as zero-one integer problem. Mixed integer programming has been shown to be a useful model, and an illustrative example shows the applicability, delivering an equipment maintenance schedule of seven units in two plants across 52 weeks.

Budai-Balke (2009) introduced the Preventive Maintenance Scheduling Problem (PMSP) in the context of railway infrastructure, aiming to minimize possession and maintenance costs by clustering maintenance activities. The models of this thesis relate to the machine scheduling problem as defined by Graham et al. (1979). Two versions were presented: one with fixed intervals (RPMSP) and one with only a maximum interval (PMSP). The PMSP has the following mathematical formulation (Budai-Balke, 2009, p. 71 - 72).

Indices

T	Set of discrete time periods (e.g. months, weeks) in which the maintenance activities need to be scheduled, i.e. $\mid T \mid$ is the planning horizon.
PA	Set of projects
RA	Set of routine maintenance works
A	$PA \cup RA$ Set of all activities
C	{(m,n) $\mid$ work m is combinable with n, $\forall m, n \in A$ }

Parameters

L_a	Cycle length of the routine work $a \in RA$
F_a	Frequency of the routine work $a \in RA$
G_a	Number of periods elapsed since routine work $a \in RA$ was last carried out before the planning horizon starts
LC_a	$\{t \in T \mid 1 + \lvert T\rvert - L_a \leq t \leq \lvert T\rvert\} \subseteq T$, set of time periods from the last planning cycle for routine work $a \in RA$
b_{at}	$\frac{\lvert T\rvert - t}{L_a}$, length of the remaining interval until the end of the planning horizon divided by the length of the planning cycle for routine work $a \in RA$ and for time period $t \in LCa$
T_p	$\subseteq T$ set of possible start points of project $p \in PA$
D_p	Duration of project $p \in PA$
pc_t	possession cost in period $t \in T$
mc_a	maintenance cost per time period for carrying out work $a \in A$

Decision variable

- x_{at} Binary variable that denotes whether activity $a \in A$ is assigned to period $t \in T$ ($x_{at} = 1$), or not ($x_{at} = 0$)
- z_{at} Binary variable that denotes whether activity $a \in RA$ is carried out for the last time in the planning horizon at time $t \in LCa$ ($z_{at} = 1$), or not ($z_{at} = 0$)
- m_t Binary variable that denotes whether the track is used for preventive maintenance work at time $t \in T$ ($m_t = 1$), or not ($m_t = 0$)
- y_{pt} Binary variable that denotes whether the execution of project $p \in PA$ starts at time $t \in T$ ($y_{pt} = 1$), or not ($y_{pt} = 0$)

Objective function

Minimize:

$$\sum_{t \in T} pc_t m_t + \sum_{a \in A} \sum_{t \in T} mc_a x_{at} + \sum_{a \in RA} \sum_{t \in LCa} mc_a b_{at} z_{at} \tag{3.2}$$

Constraints

$$\sum_{t=1}^{L_a - G_a} x_{at} \geq 1 \quad \forall a \in RA \tag{3.3}$$

$$\sum_{s=0}^{L_a - 1} x_{a,t+s} \geq 1 \quad \forall a \in RA,\ 1 \leq t \leq |T| - L_a + 1 \tag{3.4}$$

$$\sum_{t \in LCa} z_{at} \geq 1 \quad \forall a \in RA \tag{3.5}$$

$$z_{at} \leq x_{at} \quad \forall a \in RA,\ t \in LC_a \tag{3.6}$$

$$x_{mt} + x_{nt} \leq 1 \quad \forall t \in T,\ (m,n) \notin C \tag{3.7}$$

$$\sum_{t \in Tp} y_{pt} = 1 \quad \forall p \in PA \tag{3.8}$$

$$x_{ps} \geq y_{pt} \quad \forall p \in PA,\ t \in Tp,\ s = t, \ldots, t + Dp - 1 \tag{3.9}$$

$$m_t \geq x_{at} \quad \forall a \in A,\ t \in T \tag{3.10}$$

$$x_{at},\ z_{at},\ y_{pt},\ m_t \in \{0,1\} \quad \forall a \in A,\ p \in PA,\ t \in T \tag{3.11}$$

Constraints 3.3 ensure each work is carried out at least once; 3.4 schedule works no more than L_a periods apart; 3.5 - 3.6 define the last interval length; 3.7 ensure only combinable activities are carried out simultaneously; 3.8 guarantee each project is executed once; 3.9 assign projects to the correct number of periods within the specified start time range; 3.10 occupy periods for preventive maintenance if work is planned; and 3.11 ensure decision variables are binary. The PMSP was proven to be NP-hard. Four heuristics were developed to approximate solutions efficiently, with PMSP showing lower costs but longer computation times. The thesis also uses Meta-heuristics for their robust search capabilities, implementing Genetic and memetic algorithms, along with iterative and opportunity-based heuristics. Computational results show significant improvements over CPlex solver solutions, particularly for instances with low possession costs. Memetic algorithms, especially using simulated annealing, outperformed genetic algorithms in most cases. The solutions devised for railway maintenance scheduling can be applied to maintenance scheduling in other sectors as well.

Henriksson (2019) uses Mixed Integer Linear Programming (MILP) to optimize preventive maintenance planning and operation for a generic fleet of equipment. The maintenance requirements are mixed conditions of calendar based and operation based constraints. Also, maintenance of hierarchy types are handled. The objective is to minimize the number of maintenance events, considering even operation and spread in planned maintenance. The results show that it is possible to optimize preventive maintenance and operation scheduling with MILP. There are some limitations in size depending of included constraints and parameter values. Limitations in size can be handled with step-wise calculations, manual manipulations of the results, or if satisfying end conditions can be achieved, splitting of the planning period.

To conclude, many articles and theses in literature explore various maintenance planning methodologies and models. Many operations research techniques are present, such as mathematical programming, Mixed-Integer-Programming Models, and Meta-heuristics. Examples of industries present in the theory are the railway industry, the power industry, co-generation plants, and the water fleet industry.

3.4 Maintenance of Production Lines

Maintenance of production lines is a crucial aspect of manufacturing operations, focusing on preventing untimely breakdowns, improving reliability, and ensuring efficient production. Literature provides a variety of approaches and methodologies for maintenance management, including the integration of advanced technologies and strategies.

In the context of production lines, integrating production planning and preventive maintenance is crucial for enhancing operational efficiency and reducing costs. Aghezzaf and Najid (2008) addresses this by proposing mathematical models to optimize production and maintenance scheduling in parallel failure-prone production lines, employing cyclic and non-cyclical preventive maintenance policies. Cadi et al. (2015) extends this model to series-parallel production lines, emphasizing cost minimization and incorporating buffer stocks between subsystems. Purnomo, Wahab, and Singh (2023) focuses on optimizing planned preventive maintenance across multiple production lines using an optimized Weibull distribution and Bayesian optimization within a simulation framework, successfully minimizing total maintenance costs while ensuring production continuity. These studies underscore the importance of integrated production-maintenance strategies and ongoing efforts to refine optimization techniques for improved operational efficiency and system reliability in manufacturing environments.

Furthermore, in the context of preventive maintenance scheduling at production lines, Ebrahimipour, Najjarbashi, and Sheikhalishahi (2015) addresses a multi-objective preventive maintenance scheduling problem across multiple production lines, considering reliability, maintenance costs, and system downtime. Qing et al. (2010) extends this discussion to imperfect preventive maintenance on parallel production lines, proposing integrated models to minimize completion costs. Xu et al. (2017) explores maintenance planning in an unreliable production line with a branch buffer, developing an optimization model considering system throughput, buffer inventory, and device reliability. Lastly, Wang, Li, and Zhang (2016) focuses on parallel production lines, introducing a scheduling model with urgency-based objective functions and constraints on multi-product production and maintenance windows. These studies collectively highlight the complexity of preventive maintenance scheduling and offer diverse approaches to address it effectively.

Some studies focus on delay-time analyses for maintenance on production lines. Christer and Waller (1984) explores the use of delay-time analysis and snapshot modeling to predict downtime consequences and enhance maintenance practices. The study underscores the iterative nature of maintenance improvements, demonstrating how insights from modeling facilitated efficient inspection practices and plant modifications. Meanwhile, Li et al. (2018) presents a preventive maintenance strategy for automatic production lines based on a delay-time analysis. Through criticality assessment and maintenance modeling, the study establishes tailored maintenance plans for different machine groups, aiming to optimize reliability, operating rates, and maintenance costs.

The use of machine learning is also present in the literature considering maintenance on production lines. Kang, Catal, and Tekinerdogan (2021) presents a novel machine learning-based approach for predicting the Remaining Useful Life (RUL) of equipment in production lines, focusing on turbo engines using NASA datasets. The study demonstrates the efficacy of employing interpolation and multi-layer perceptron neural network (MLP) algorithms, alongside pre-processing techniques such as normalization and principle component analysis, to enhance predictive maintenance in production environments. The results underscore the potential of utilizing machine learning for RUL prediction and its significance in proactive maintenance management, while acknowledging challenges such as data variability and the need for representative training data in real-world production settings.

To conclude, in literature, preventive maintenance of production lines is depicted as crucial for preventing breakdowns, improving reliability, and ensuring efficient production. Studies highlight the integration of production planning and maintenance scheduling, optimization techniques, delay-time analyses, and the utilization of machine learning for predictive maintenance as key strategies to enhance operational efficiency and system reliability.

3.5 Prediction Models

Literature discusses many types of prediction models. Efron (2020) compares modern prediction algorithms with standard regression models, centered on the differences between prediction and estimation or prediction and attribution. Shmueli (2010) aims to clarify the distinction between explanatory and predictive modeling, and discusses its sources. Hand (2006) argues that simple methods typically yield performances almost as good as more sophisticated methods. Lastly, Donoho (2017) discusses the difference between data science and statistics. Section 3.5.1 continues to discuss the different types of prediction models, and Section 3.5.2 considers the evaluation metrics. Section 3.5.3 shortly discusses one-hot encoding. Then, Section 3.5.4 tackles the subject of parameter tuning and cross-validation, while Section 3.5.5 delves into the current literature of prediction modelling concerning maintenance or production.

3.5.1 Model Types

The book of James et al. (2013) discusses several models, including the following:

- Linear Regression
- Decision Trees
- Random Forest
- Gradient Boosting Machines
- K-Nearest Neighbor

Linear Regression is a statistical method used to model the relationship between a dependent variable and one or more independent variables by fitting a linear equation to the observed data. *Decision Trees* are versatile and easy-to-interpret models that partition the feature space into segments and make predictions based on the majority class or the average target value within each segment. *Random Forests* are ensemble learning methods that construct multiple decision trees during training and output the average prediction of the individual trees. *Gradient Boosting Machines* are a powerful ensemble technique that builds models sequentially, with each new model attempting to correct the errors of the previous ones. Lastly, *K-Nearest Neighbor* is a non-parametric method used for regression tasks, where the predicted value is the average of the values of its k nearest neighbors.

Another type is the *Holt-Winters model*; a forecasting technique used in time series analysis. The book of Axsäter (2006) discusses it as an extension of exponential smoothing where it takes into account trends and seasonality in the data. The model includes three components: level (the average value of the series), trend (the direction of the series), and seasonality (patterns that repeat at regular intervals). By incorporating these components, the Holt-Winters model can provide more accurate forecasts for time series data.

3.5.2 Evaluation Metrics

Again, literature discusses many subjects on methods of assessing models. Steyerberg et al. (2001) assesses different variants of split-sample, cross-validation, and bootstrapping methods to validate a logistic regression model. The book of James et al. (2013) also considers model accuracy assessment methods. This research chooses to assess the models based on the Root Mean Squared Error (RMSE), the Mean Absolute Deviation (MAD), and the Symmetric Mean Absolute Percentage Error (sMAPE). The following equations calculate the RMSE, the MAD, and the sMAPE:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{f}(x_i))^2} \tag{3.12}$$

where $\hat{f}(x_i)$ is the prediction that $\hat{f}$ gives for the ith observation, y_i is the actual value, and n is the number of predicted observations. The RMSE will be small if the predicted responses are very close to the true responses, and will be large if for some of the observations, the predicted and true responses differ substantially.

$$\text{MAD} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{f}(x_i)| \tag{3.13}$$

where $\hat{f}(x_i)$ is the prediction that $\hat{f}$ gives for the ith observation, y_i is the actual value, and n is the number of predicted observations. The MAD measures the average absolute difference between the predicted and actual values. Similar to RMSE, a lower MAD value signifies a better model.

$$\text{sMAPE} = 100\% * \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - \hat{f}(x_i)|}{(|y_i| + |\hat{f}(x_i)|)/2} \tag{3.14}$$

where $\hat{f}(x_i)$ is the prediction that $\hat{f}$ gives for the ith observation, y_i is the actual value, and n is the number of predicted observations. The sMAPE measures the relative accuracy of predictions. It scales the absolute errors by the average of the absolute values of the actual and predicted values. sMAPE ranges from 0 to 100%. Lower values indicate better accuracy.

$$\text{Bias} = \frac{1}{n} \sum_{i=1}^{n} (\hat{f}(x_i) - y_i) \tag{3.15}$$

where $\hat{f}(x_i)$ is the prediction that $\hat{f}$ gives for the ith observation, y_i is the actual value, and n is the number of predicted observations. A bias close to zero indicates that the model's predictions are, on average, close to the actual values, whereas a positive bias indicates a tendency to overestimate, and a negative bias indicates a tendency to underestimate.

3.5.3 One-Hot Encoding

The previously discussed models use input data for the creation of accurate predictions. This data consists of so called features. Sometimes, it is necessary to process these features to adequately prepare them for the prediction models. One-hot Encoding is a method in machine learning to represent categorical variables as binary vectors. It converts each category into a binary feature. In this feature, all the elements are zero except for the position corresponding to the specific category, which is set to one.

The book of Bishop (2006) discusses various encoding methods for categorical data, including one-hot encoding, and their impact on machine learning algorithms. It shows that if the months or weeks in the year represent categorical data rather than a sequential trend (e.g. for seasonality), you can use one-hot encoding. This is useful when you want to treat each month or week as an independent category.

3.5.4 Parameter Tuning and Cross-Validation

The book of James et al. (2013) discusses parameter tuning as the method that adjusts the parameters of a model to optimize its performance. Parameter tuning involves experimenting with different values for the model's parameters to find the combination that yields the best results in terms of model accuracy, generalization, or other performance metrics. This iterative process typically involves techniques such as grid search, random search, or more sophisticated optimization algorithms to find the optimal parameter values for a given model and dataset.

One method of experimenting with different parameter values is Cross-Validation. This is a technique used to assess the performance of a machine learning model. It involves dividing the dataset into multiple subsets and then performing the training and validation process multiple times. Different types of data series have unique characteristics that necessitate specialized splitting techniques for effective model evaluation. The paper of Schnaubelt et al. (2019) explains these differences. It considers time series data, which are observations collected at successive points in time, making them inherently dependent on previous values. This dependency requires preserving the temporal order during model evaluation. To maintain the temporal structure, ensuring that future values are not used to predict past values, the paper suggests to split the data into training and testing sets such that the training set always precedes the testing set chronologically. For example, train on the first few time points and test on the next few, then move the window forward. Figure 3.3 shows how to split Time Series Data in Cross-Validation methods.

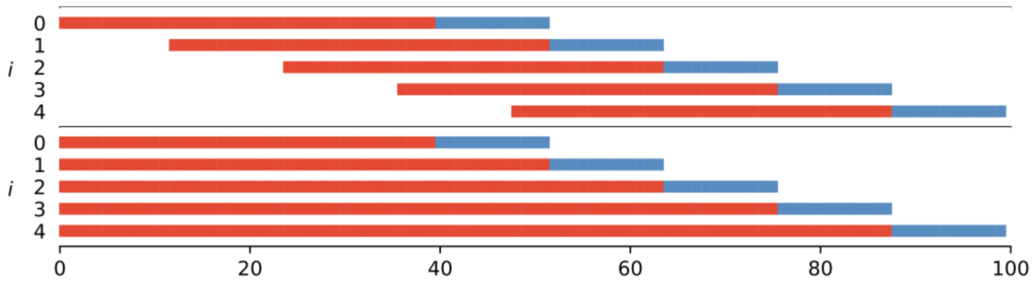


FIGURE 3.3: Illustration of validation data splitting schemes for Time Series Data (Schnaubelt et al., 2019)

Cross-Validation involves dividing the dataset into multiple subsets/folds and performing the training and validation process multiple times. In each iteration, there is one validation subset while the remaining subsets are for training. This process repeats multiple times, with different parameter combinations in each iteration. The goal is to find the set of parameters that results in the best average performance across all iterations. By doing so, cross-validation ensures that the selected parameters generalize well to unseen data, helping to avoid over-fitting and improving the model's overall performance.

3.5.5 Recent Literature

This section delivers the most recent articles on prediction modelling in the context of maintenance or production. Ayvaz and Alpay (2021) developed a data-driven predictive maintenance system using IoT sensor data and machine learning algorithms like RF and XGB to prevent production stops. Hu et al. (2023) introduced Knowledge Enhanced Reinforcement Learning (KERL) to optimize production and maintenance scheduling, enhancing performance with prior knowledge. Kang, Catal, and Tekinerdogan (2021) proposed using multilayer perceptron neural networks (MPNN) to predict the remaining useful life (RUL) of equipment in continuous production lines. In mining, Koomson, Temeng, and Ziggah (2024) used metaheuristic algorithms (PSO, GA, WOA) to predict dump truck tire life, with PSO-MLPNN as the best model. Koulinas, Paraschos, and Koulouriotis (2024) applied decision tree algorithms informed by reinforcement learning for optimization in degrading manufacturing and remanufacturing systems. Li et al. (2024) combined STL, DBN-ELM, and SVR for accurate prediction of rock drilling operation times in underground mining. Tucci, Piazzi, and Thomopulos (2024) predicted electricity production from solar photovoltaic installations in Italy using KNN and RF models, highlighting feature selection and retraining frequency. Pan et al. (2023) developed a conditional health status prediction structure for EVA copolymer reactors, finding Extremely randomized trees as optimal for predicting reactor bearing's RUL. Finally, Skachkova, Alenin, and Mokshin (2022) used correlation analysis and Bayesian regularization neural networks (BRANN) for feature selection and optimization in oil drilling, demonstrating superior prediction accuracy and training efficiency. These studies highlight the potential of advanced machine learning and optimization techniques in improving maintenance, production, and operational efficiencies across various industries. Table 3.1 provides the conceptual matrix of these articles, showing the models, metrics and other details.

TABLE 3.1: Conceptual matrix for prediction modelling

Source	Models	Metric	Tuning	Predicting
Ayvaz et al. 2021	RF, XGB, GBM, SVR	R2, MAE, MAPE, RMSE	no	Prod. stops
Hu et al. 2023	KERL	overall business reward	no	Prod. workload
Kang et al. 2021	MPNN	MSE	yes	RUL
Koomson et al. 2024	PSO, GA, WOA	VAF, NASH, R2, MAPE	yes	Component life
Koulinas et al. 2024	DT	TPR	yes	Prod. policies
Meulenbroek 2024	**LR, DT, RF, GBM, KNN, HW**	**RMSE, MAD, sMAPE**	**yes**	**Prod. time**
Li et al. 2024	STL, DBN-ELM, SVR	MAPE	no	Prod. time
Tucci et al. 2024	KNN, DT, KRR, LR, SVR, RF, GBM	NRMSE	yes	Electricity Prod.
Pan et al. 2023	RF, XGB, SVM, LR	RAE	yes	RUL
Skachkova et al. 2022	BRANN	Learning error	yes	Oil prod.

This comparison highlights the effectiveness of diverse machine learning models (e.g., RF, GBM, neural networks) across industries, emphasizing the importance of model tuning and evaluation metrics like the RMSE. Recent literature often focusses on individual cases where machine learning is applicable. To extend these findings, this research aims to provide a more holistic analysis inside a relatively large organization. Whenever there is a need for many separate predictions due to a complex system of multiple aspects, the current literature fails to address how to assess multiple prediction methods across different prediction data sets. This research aims to fill this gap.

3.6 Conclusion

Theoretical background on counter-based maintenance emphasizes scheduling tasks based on equipment usage rather than fixed calendar intervals, especially when failure costs outweigh replacement costs and traditional time-based approaches are ineffective due to variable usage patterns. Case studies have demonstrated the superiority of counter-based maintenance over time-based methods, showing that the number of replacements during a given time period can decrease with about 30%. Transitioning from a calendar time-based system to a usage counter-based system leads to an extension of a component's service life, by reducing the uncertainty of its state. This transition reveals that under the new strategy, replaced components exhibit greater damage at the time of replacement, thereby reducing the occurrence of over-maintenance and minimizing remaining lifetime spillage. Studies show that the transition can result in cost savings of thousands of pounds. Nevertheless, theory lacks to discuss the specific calculations necessary to transition from a time-based to a counter-based strategy.

One large gap is the absence of theory on the practicalities of transitioning from a time-based to a counter-based maintenance strategy. There are no written methods of using the knowledge and data of time-based systems to adjust to usage counters. Additionally, in the context of such a transition, there are no methods of arriving at adequate maintenance intervals while failure data is lacking in an organization. Furthermore, while there is a lot of theory on prediction models, there are no sources that apply these methods in the context of counter-based maintenance. Having usage counters for the application of maintenance results in the problem of having to predict future usage amounts. There is no theory that applies prediction models in this context. Finally, there are many examples of planning models in current literature. However, there are no specific planning models that consider planning maintenance activities based on their usage counters. This thesis aims to fill these gaps in the theory.

The purpose of this chapter was to find the present theory on the comparisons/transitions between time-based maintenance and counter-based maintenance, the determination of the interval between maintenance activities, the mathematical models for maintenance planning, the maintenance of production lines, and prediction models. Table 3.2 shows the conceptual matrix of this literature review, stating which concepts the found sources tackle. This table indicates the gaps in the theory, thereby showing the contribution of this research by incorporating all concepts.

To conclude, this research chooses to create a method to transition from the time-based to a counter-based strategy. It finds the best method for the company to predict the amount of production on the filling lines, by evaluating the DT, RF, GBM, KNN, and the HW models, with additional tuning and cross-validation methods to improve parameter settings. Finally, it creates a new planning model similar to the model from Budai-Balke (2009), while adjusting the goal to comply with a counter-based maintenance strategy, with a focus on minimizing under- and over-maintenance.

TABLE 3.2: Conceptual matrix

Sources	**Comparison TBM and UBM**	**Interval Determination**	**Planning Models**	**Production Lines**	**Prediction Models**
Alardhi et al. 2008			x		
Aghezzaf et al. 2008			x	x	
Ahmad et al. 2012	x				
Axsäter 2006					x
Ayvaz et al. 2021				x	x
Basker et al. 1978		x		x	
Ben-Daya et al. 2001			x	x	
Bishop 2006					x
Budai-Balke 2009			x		
Cadi et al. 2015			x	x	
Christer et al. 1970	x		x	x	
Christer et al. 1978				x	
Cichelli 1977	x			x	
Dekker 1996		x	x	x	
Deloux et al. 2016	x				
Dedopoulos et al. 1995			x	x	
Dijkhuizen 2000		x	x	x	
Dohono 2017					x
Ebrahimipour et al. 2015			x	x	
Efron 2020					x
Fokkert et al. 2007			x		
Frost et al. 1998			x		
Graham et al. 1979			x		
Haghani et al. 2002			x		
Hand 2006					x
Henriksen 2019			x		
Holland et al. 1975		x			
Hu et al. 2023			x	x	x
James 2013					x
Jonge et al. 2015		x			
Kang et al. 2021				x	x
Kim et al. 2016	x				
Koomson et al. 2024					x
Koulinas et al. 2024			x		x
Li et al. 2018				x	
Li et al. 2024				x	x
Liu et al. 2024	x				
Meulenbroek 2024	**x**	**x**	**x**	**x**	**x**
Pan et al. 2023					x
Percy 2008		x			
Purnomo et al. 2023			x	x	
Qing et al. 2010			x	x	
Rausand et al. 2021	x	x		x	
Schnaubelt et al. 2019					x
Sharma et al. 2020		x			
Shmueli 2010					x
Skachkova et al. 2022				x	x
Steyerberg et al. 2001					x
Tinga 2010	x	x			
Tinga 2013	x	x			
Tinga 2023	x				
Tucci et al. 2024				x	x
Vatn et al. 1996			x	x	
Vaurio 1998		x	x	x	
Wang 2002	x				
Wang et al. 2016			x	x	
Xu et al. 2017			x	x	

Chapter 4

Solution Design

This chapter addresses the third research question:

"How should this research collect, select, and process the data to create the design of the solutions?"

It explores data collection and analysis methods. The research question unfolds into several sub-questions, including:

- *What is the approach to arrive at a solution of the problem?*
- *How should the Maintenance Engineering Team calculate the counters, and how often should they perform the calculations?*
- *What data should be collected for this thesis, what data is currently available, and how should the situation of missing data be handled?*
- *How should the Maintenance Engineering Team predict the production hours of the machines on the production filling lines?*
- *How should the Maintenance Engineering Team model the planning of preventive maintenance activities?*

Section 4.1 starts with a short description of the approach to arrive at a solution to the problem. Then, Section 4.2 discusses the calculations of the counters, and Section 4.3 addresses the prediction models for production in the future. Section 4.4 considers the planning model, with its corresponding modelling decisions and mathematical formulation. Then, Section 4.5 concludes the chapter.

4.1 Solution Approach

To arrive at a solution to the core problem, this chapter divides and combines three separate solutions. Firstly, the company and present theory lack a method of translating the current time-based approach to a sufficient counter-based maintenance strategy. The first solution are definitions of formulas to arrive at counter-based intervals between maintenance activities. Secondly, introducing usage counters for the application of maintenance results in the problem of having to predict future usage amounts. The second solution entails finding a sufficient prediction method of future production amounts on the packaging filling line. Finally, the company requires a solution to the planning activities of the preventive maintenance actions. The third solution is the planning model that generates a schedule which minimizes over- and under-maintenance. Figure 4.1 visualizes this solution approach.

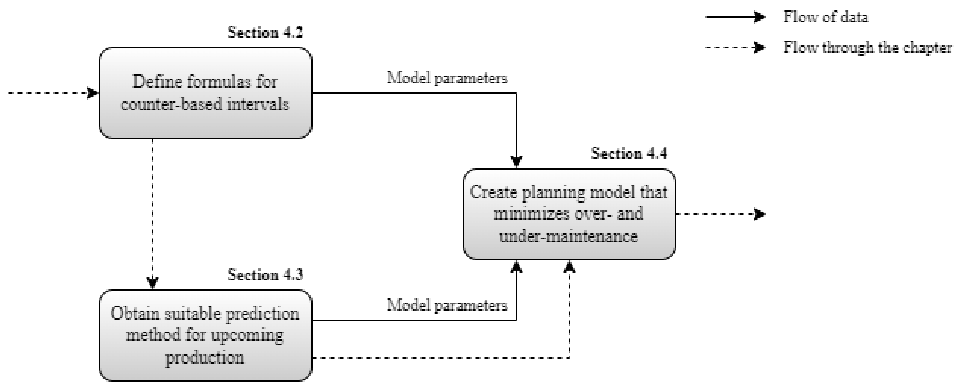


FIGURE 4.1: Solution Approach

Figure 4.1 displays the chronological order of this chapter, starting with the definition of the formulas for the counter-based intervals, following with the method of obtaining the most suitable prediction method, and ending up with a planning model for the maintenance activities. Note how Figure 4.1 shows that the first two solutions generate model parameters of the final planning model, meaning that the thee sections are integrated.

4.2 Determination of Usage Counters

With the knowledge from Chapters 2 and 3, this section discusses the calculations of the counters for the maintenance of a packaging filling line. Section 3.2 already states how literature uses failure distributions of parts to find the optimal maintenance interval with the use of equation 3.1. The company however lacks data to assign failure distributions to its assets, resulting in a demand for a simple transitional formula going from the time-based intervals to the counter-based intervals. Section 2.5 has shown the performance and value of the current intervals, justifying the following equations for the calculations of the usage counters.

Consequently, this research defines a usage counter as the total amount of executed production hours in proportion to the set translated counter interval of a particular PM-plan. The following equation determines this usage counter:

$$C = \frac{\sum_{t=l}^{T} P_t}{CI} \tag{4.1}$$

where C is the usage counter, CI is the counter-based interval, P_t is the amount of production on day t, l is the day on which the activity was last executed, and T is the present date. The formula sums up the total amount of production hours since its last execution and divides it by the counter-based interval. The following equation determines the counter-based interval:

$$CI = \frac{TI}{365} * A \tag{4.2}$$

where CI is the counter-based interval, TI is the time-based interval in days, and A is the average annual amount of production hours for the past five years. Note that this is a simple linear function, making it relatively easy for the company to perform these calculations. An example would be a specific PM-plan that currently uses a time-based interval of six months. This PM-plan has an average of 6,000 hours annually. This would result in a translated Counter-Based Interval of about 3,000 hours (182/365 * 6,000). Do note that these Counter-Based Intervals are useful parameters for the planning model in Section 4.4

The usage counter C from equation 4.1 indicates how close the PM-plan is to its interval. For example, a value of 0.5 indicates that half of the amount of production has passed until it reaches the interval. If the value is 1.5, it indicates that there has been too much production and the execution of the PM-plan should be considered as too late. Ideally, the value of the usage counter should be as close as possible to the value 1.0 at the execution of the PM-plan. Knowing when to schedule the PM-plan based on the value of the usage counter (at $C = 0.8$ or $C = 0.9$) depends on knowing what amount of production the company can expect in the future, and this is what the next section discusses.

Because the PM-plans are always planned 13 weeks beforehand for preparations, monthly monitoring of the counters should be adequate, if an automatic check is impossible. After the calculations a maintenance engineer evaluates the preventive maintenance plans that will reach the counter interval in the upcoming 13 weeks, and send these activities to the maintenance planner to start preparing its execution. Obviously, it is necessary to further investigate how a maintenance engineer can know that the amount of production in the upcoming 13 weeks exceeds the counter-based interval. The next section tackles this issue.

4.3 Prediction Models

This section focuses on the prediction models and the different steps to generate useful output. The goal of the prediction models is to predict the amount of production of the machines from the packaging filling line 81. Section 4.3.1 starts with the objective of the prediction models. Then, Section 4.3.2 discusses the time horizon, while Section 4.3.3 considers the collection of raw data. Section 4.3.4 tackles the processing of this data, and Section 4.3.5 explains the application of One-Hot Encoding. Finally, Section 4.3.6 elaborates on the training of the model.

4.3.1 Objective

Whenever there is the requirement to plan maintenance activities with a substantial amount of time before its execution due to preparations, introducing usage counters results in the problem of having to predict future usage amounts. For example, if one maintenance activity should be scheduled whenever its corresponding machine reaches 3,000 hours, and currently the machine has been active for only 2,000 hours. How would one know when to plan the maintenance activity before it reaches this interval? Additionally, if one wishes to plan multiple activities across a large planning horizon, it becomes increasingly difficult to know when production reaches the counter-based intervals.

These difficulties require the use of two prediction models. The first provides knowledge on when to plan any PM-plan with enough time (13 weeks) for preparations. The second improves the planning activities of multiple PM-plans in the upcoming period by providing the expected amount of production in between stop-days, which are opportunities for preventive maintenance. Note that the latter provides the information for parameters of the planning model in Section 4.4

4.3.2 Forecasting Aspects

In the context of forecasting, the terms *time horizon*, *time buckets*, and *aggregation level* are crucial for understanding the different dimensions and granularity of the forecasts. The time horizon refers to the length of time into the future for which a forecast is made. The time horizon for the first prediction model is set at one week because the goal is to acquire knowledge on when to start planning a given PM-plan. Therefore it is only necessary to know the predicted value of one week at any given moment during the year. The time horizon for the second prediction model is set at one year because the goal is to acquire knowledge before the start of the following year. These predictions are less accurate but provide useful insights for the maintenance planner.

Time buckets refer to the intervals that divide the forecast period. For both prediction models the time buckets are set to weeks. Daily predictions are not possible since the second source of data only provides weekly information, and monthly predictions lack knowledge on the precise moment when to plan maintenance in a given week.

The aggregation level refers to the level of detail at which the model forecasts the data. To aggregate means to group/sum multiple data values to simplify the predictions. For the first prediction model, the goal is to find in what amount of production will follow in the following 13 weeks. Therefore, the aggregation level is set at 13 weeks. The Maintenance Engineering Team should be able to predict during any given week in a year. Therefore, each data point now shows the sum of the following 13 weeks, for both the input and output variables. The second prediction model aims to find the details of every week in the following year and has therefore an aggregation level of one week.

4.3.3 Data Collection

This research collects data from several software systems. The Manufacturing Execution System (MES) is the software that collects and stores all kinds of data from each production filling line. MES also stores the production data per filling line, giving the amount of minutes that each machine is running per day. Section 2.6 discusses the possible states of a machine, and MES collects on each calendar day and for each machine on the line the amount of minutes that the machine is in each state. Section 2.6 argues why only three states show the most reliable representation of real production amounts. Therefore, this research requires data on the states that are Production time, Emptying idle time, and Filling idle time.

For line 81, MES currently is able to export data of the years 2016 to 2024. In total it contains data for 31 machines. Together with one of the maintenance engineers a thorough analysis has been done on the MES output of these 31 machines and five of them are unusable due to either censored data or unrealistic outcomes. For each remaining machine the only feature from MES is the calendar day.

The second source of data originates from the Tactical Planning Team. This department makes use of a central database with historical figures on working hours and breakdown requests. In close collaboration with the Workforce Planner of the packaging filling lines and the Tactical Planning Team, this research gathers data that states the number of working hours on line 81 per week from years 2013 to 2023. The years 2013 to 2015 are invaluable because it is dated and misrepresents reality. Therefore, this research collects data from all sources from years 2016 up until and including 2023.

4.3.4 Data Processing

Before a model can train with the data, the data requires processing. This section discusses the feature modifications, outliers, and missing data. The book of Géron (2019) covers processing techniques extensively, demonstrating their application in machine learning pipelines. Some of these techniques are applicable to the data of this research. MES contains data on each calendar day of the year. From the argumentation of Section 4.3.2, this research processes the daily data from MES such that it merges into weekly data values, where the calendar day transforms into three features: the year, the week, and the week number inside this month. The data from all sources do not contain any outliers and it has no missing data, meaning it does not require further processing.

One of the techniques from the book of Géron (2019) is handling temporal data. This suggests that if a feature represents temporal data and trends over time are important, it is essential to transform it to new features that capture the passage of time. For instance, if the year represents the time of an event, you could create a feature that measures the number of years since a baseline year (e.g. years_since_2000 = year - 2000) to account for long-term trends. For the data in this section, this research applies this method on the feature that represents the year.

One last modification is aggregating the data for the first model. Therefore, each data point now shows the sum of the following 13 weeks, for both the working hours and the production amounts of the machines. Table 4.1 shows the final data for the first prediction model with the aggregated values, and Table 4.2 shows the final data for non aggregated values.

TABLE 4.1: Data for the first prediction model (aggregated values)

Input variables (features)				Output variables		
Year	**Week**	**Week of Month**	**Tactical Planning**	**Machine 1**	**...**	**Machine 26**
0	1	1	1,693	81,717.85	...	63,723.49
0	2	2	1,719	83,733.84	...	67,513.81
0	3	3	1,735	83,895.49	...	68,560.35
...	...	...	...	...	...	...
7	40	1	1,588	90,834.31	...	73,710.34

Table 4.1 contains four input variables, 26 output variables, and a total of 404 data-points. The first three columns in Table 4.1 define the numerical value of the year, week, and week number of the month of a data-point, e.g. the first week of January in 2016 (0 meaning 2016, 1 meaning the first month, 1 meaning the first week). The fourth column contains the aggregated numerical amount of production hours planned for the workforce in the upcoming 13 weeks. The first data-point shows that the company planned the workforce to be working a total of 1,693 hours in the upcoming 13 weeks. There are 26 output variables, all containing information of one machine on the line. The value represents the aggregated amount of minutes that

the machine is running in the upcoming 13 weeks. For example, the first data-point shows that machine 1 was running almost 82,000 minutes in the upcoming 13 weeks.

TABLE 4.2: Data for the second prediction model (non aggregated values)

Input variables (features)				Output variables		
Year	**Week**	**Week of Month**	**Tactical Planning**	**Machine 1**	**...**	**Machine 26**
0	1	1	126	2,880	...	0
0	2	2	136	8,032.81	...	5,196.48
0	3	3	130	6,917.86	...	6,317.85
...	...	...	...	...	...	...
7	52	4	120	0	...	0

Table 4.2 is similar to Table 4.1. Again, it has 4 input variables, 26 output variables, with a slightly higher total of 416 data-points because the values are not aggregated. Only the fourth feature and all output variables adjust to be non aggregated values. For example, now the first data-point shows that the company planned the workforce to be working 126 hours in the that specific week. Also machine 1 shows to have run 2,880 minutes in that week.

4.3.5 Application of One-Hot Encoding

For completeness, this research aims to find the best possible method of predicting the production. Another method to hypothetically increase the accuracy of the predictions is to use One-Hot Encoding, as Section 3.5.3 discusses. One-Hot Encoding makes binary features, often used in statistical modeling to represent categorical data. Each new feature, typically with values 0 or 1, indicate the absence or presence of a particular category. They allow categorical features to be included in regression models by converting qualitative data into a quantitative format.

For example, in both the datasets the weeks of the year are typically categorical input features:

- Week_1: 1 if the data-point is the first week, and 0 if otherwise
- Week_2: 1 if the data-point is the second week, and 0 if otherwise
- Week_3: 1 if the data-point is the third week, and 0 if otherwise

These features in prediction models help analyze the effect of each specific week on the outcome variables. This research chooses to create One-Hot Encoded features for the week, and the week of the month. These features might have specific seasonal effects (e.g., holiday weeks, vacation periods, or monthly targets). One-Hot Encoded features can help model these distinct effects more precisely. Introducing these One-Hot Encoded features increases our initial four features to a total of 59 features.

4.3.6 Model Training

Section 3.5.1 identified a number of machine learning models useful for predicting production amounts. To validate these models it is necessary to split the dataset into a training set and a testing set. Then the models can train themselves on the training set, and show their performances based on the testing set. The evaluation metrics from Section 3.5.2 will provide the performances of the models. Gholamy, Kreinovich, and Kosheleva (2018) shows why the best results are attained if datasets allocate 20% to 30% percent of the original data points for testing, and use the remaining 70% to 80% for training. The non aggregated dataset has eight full years of weekly values resulting in 416 data points. The aggregated set only has 404 data points because the last 12 weeks cannot be aggregated. This research chooses to split the data into a training set with data from the years 2016 to 2021, and a testing set with data from the years 2022 and 2023. This sets the division for both prediction models to about 30% to 70%.

To further improve the prediction models, this research performs parameter tuning to the applicable models. These are the DT, RF, GBM, and the KNN model. For these models, the research chooses to perform five folds cross-validation. It uses the *TimeSeriesSplit* tool, which correlates to the top part of Figure 3.3 in Section 3.5.4. In this manner the folds maintain the temporal structure, ensuring that future values are not used to predict past values. Appendix C shows the parameter distributions for each model, which all lie around the standard values of the python model functions.

Additionally, to make a valid comparison between the models, this research adds a baseline prediction to the assessment. The purpose of the baseline model is to provide a benchmark. If a complex model cannot outperform the baseline model, it suggests that the complexity is unnecessary, and simpler methods might be preferable. It assumes a constant production rate, making predictions easy to calculate. This method starts with finding the average amount of production per week/data-point over the whole training set. Each week/data-point of the test set is given this average, resulting in a very simple prediction. If the baseline predictions perform the best, it would mean that this research advises the company to not use one of the prediction models and simply use an average to predict the future values of expected production amounts.

To conclude, the research aims to predict the amount of production of the machines from the packaging filling line 81. With an aggregated and non aggregated dataset, evaluation metrics of seven prediction models (including the baseline) will show how to achieve this goal. To further clarify the method of prediction modelling, Figure 4.2 displays the approach.

Figure 4.2 shows that depending on the type of the model and the possibility of parameter tuning, the approach starts either with defining the parameters, or just fitting the model. In the case of tuning it continues to find the optimal set of parameters and fitting that model. For all models after fitting, they perform predictions on the test set and their performances are stored. After this procedure for all machines is done, the final average performance can be evaluated.

4.4 Planning Model

This section delivers the method of modelling the planning of preventive maintenance activities. Section 3.3 contains a thorough discussion of planning models in the context of maintenance. The planning model of this section is mostly inspired by the PMSP and RPMSP of Budai-Balke (2009). Section 4.4.1 gives a description of the model, and Section 4.4.2 delivers its objective. Then, Section 4.4.3 discusses the restrictions and assumptions, and Section 4.4.4 provides the mathematical formulation of the model. Finally, Section 4.4.5 delivers a short explanation of the rolling horizon.

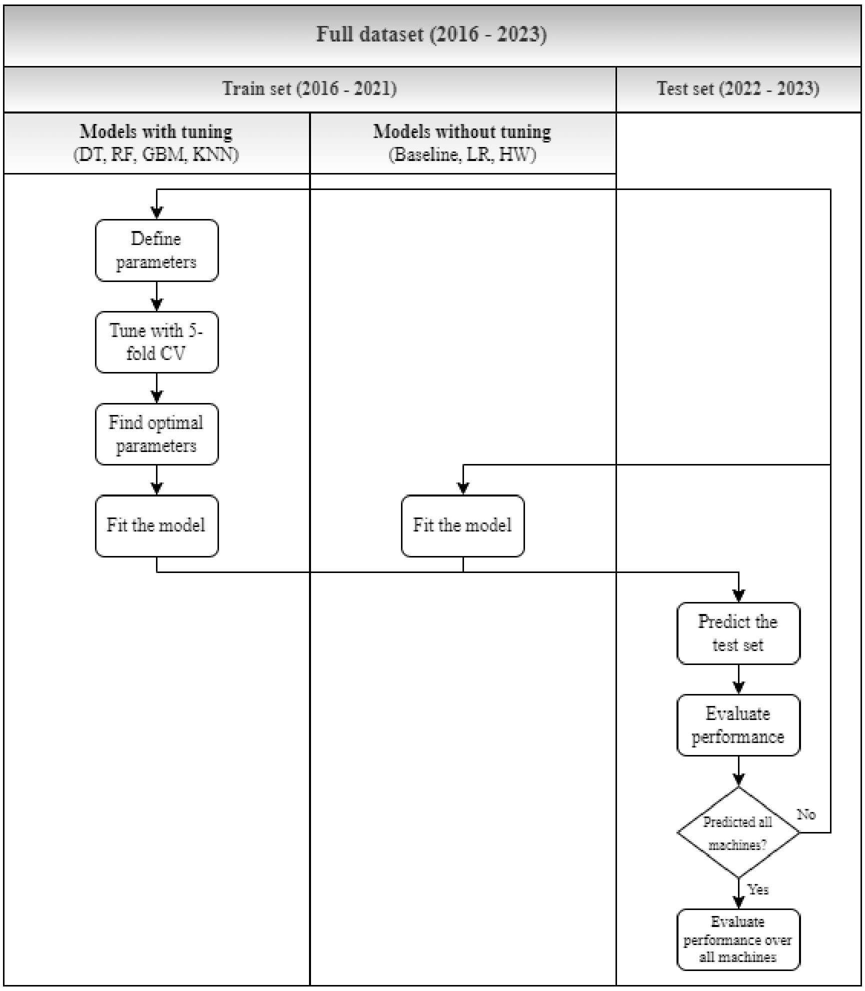


FIGURE 4.2: Approach of prediction modelling

4.4.1 Model Description

The definition of the model is as follows. Given a set of routine activities called PM-plans, the goal is to schedule them in a way that minimizes the amount of over- and under-maintenance. The intend is to plan them such that the past production since the last execution is as close to the counter-based interval as possible. The model uses a rolling horizon which Section 4.4.5 discusses. The schedule has a finite planning horizon of 52 weeks, containing several moments for the execution of these PM-plans, known as stop-days. For each PM-plan, the counter-based interval, which represents the maximum amount of production between consecutive executions, is known with the method from Section 4.2. With the non aggregated predicted amount of production by the model of Section 4.3, it is known how much production the company can expect between the stop-days. Additionally, the amount of production that has passed since the last execution of each PM-plan is provided.

A PM-plan must only be executed within the planning period before the amount of production exceeds the counter-based interval. For each PM-plan, the first moment, before the expected production reaches the counter-based interval is known. Additionally, the last moment before the remaining expected production becomes lower than the counter-based interval is known. Given the expected production between the stop-days, it is also known for each moment, when the amount of production between the moments that follow exceed the counter-based interval.

The planning model of this thesis appears to be related to the machine scheduling problem, similar to the models of Budai-Balke (2009). The connection lies in the fact that jobs, with some specifications and time windows between consecutive executions, need to be scheduled within a certain time frame. However, the objective of this thesis' model differs, focusing on planning the PM-plans as close to their most optimal moment (considering over- and under-maintenance) as possible, depending on the restrictions. Moreover, PM-plans are repetitive (mainly the plans with a short interval), which is not the case for the machine scheduling problem.

4.4.2 Objective

The objective of the planning model is to schedule the PM-plans across the planning horizon while minimizing over- and under-maintenance. To minimize over-maintenance, means to keep the amount of scheduled activities of a PM-plan at a minimum. To minimize under-maintenance, means to schedule the activities as close to the optimal moment before the amount of production exceeds the counter-based interval. Due to restrictions it might not be possible to plan all PM-plans on their most optimal moments in the planning horizon. The distance from these best options is what the model should reduce as much as possible.

4.4.3 Restrictions and Assumptions

There are several aspects that result in a restriction of the planning model. On any given stop-day there is a restriction on the amount of work. The company distinguishes stop-days into two categories: short and long stop-days. Consequently, short stop-days can perform less PM-plans in comparison to long stop-days. In discussion with the Maintenance Planner, restrictions are the following:

- Execution duration per PM-plan
- Skill-set requisites of mechanics per PM-plan
- Availability of mechanics per Stop-day
- Spare-part availability per PM-plan per stop-day

For now, the company fails to extract this information for a planning model. Therefore, this research makes the assumption that it is sufficient to combine the restrictions into one that states an allowed number of PM-plans per moment; a stop-day.

Additionally, the Maintenance Planner provided a set of combinations of PM-plans that have to be clustered. Clustering is another restriction, and is important for the set up of maintenance. For example, whenever a mechanic takes apart a section of a machine, in some specific cases it is best to perform all PM-plans considering that section.

Further restrictions are the results of the planning model depending on the production amounts during the planning horizon as an input. There is definitely no possible way that the company can know these production amounts exactly. Therefore, this research assumes that the predictions of the models from Section 4.3 are adequate enough to use as parameters for the planning model. The main argument originates from the fact that currently there are no planning models in use, and even a model with parameters from the output of a prediction model could improve this situation.

Another problem is the preference of having maintenance too early or too late. The question is how late a PM-plan can be scheduled, if the company wishes them to be scheduled late at all. Considering this subject, this research assumes that all PM-plans are not allowed to be scheduled any later than the moment when the amount of production has surpassed its counter-based interval. This relates to a safe, risk-averse and conservative choice, which falls in line with the current decisions of the Maintenance Engineering Team.

4.4.4 Mathematical Formulation

The aim is to give a schedule for preventive maintenance activities in a finite horizon, such that the amount of production before the execution of PM-plans is as close to their counter interval as possible. The mathematical formulation is as follows.

Indices

T	Set of moments/stop-days for preventive maintenance
P	Set of PM-plans
A	$\{(m,n)\vert$ m must be combined with n, $\forall m,n \in P\}$ Set of grouped PM-plans

In comparison to Budai-Balke (2009), this research does not divide the activities into multiple sets. There is a single set of activities: the PM-plans. A PM-plan can be seen as a routine activity of maintenance. If necessary the model will plan multiple activities of a single PM-plan in the planning horizon. Similar to Budai-Balke (2009), there is a single set of discrete moments in which the model schedules PM-plans. Additionally, there is a set of multiple PM-plans that should be clustered due to requirements by the maintenance coordinator and the mechanics. This could originate from efficiency purposes during the activity, for example performing two tasks together because it has a very similar set up.

Parameters

S_p	Start state of the counter of PM-plan $p \in P$
C_t	Capacity of moment/stop-day $t \in T$
FM_p	The first moment after the expected production reaches the counter-based interval of PM-plan $p \in P$
M_{tp}	The moment after the expected production of following moments after $t \in T$ reaches the counter-based interval of PM-plan $p \in P$
LM_p	The last moment after the remaining expected production becomes lower than the counter-based interval of PM-plan $p \in P$
mc	The maintenance activity costs for each moment and for every PM-plan
pc	The penalty costs for scheduling an activity away from their optimal moment considering over-maintenance

The parameters differ from the model of Budai-Balke (2009). Here the focus lies on planning the PM-plans as best considering their counter-based intervals. One similarity is the use maintenance and penalty costs. This helps to generate insights in finding a schedule that keeps the costs at a minimum. Note that there are only penalty costs for over-maintenance because under-maintenance is minimized with the constraints.

Decision variable

x_{tp}	binary variable that denotes whether PM-plan $p \in P$ is assigned to moment $t \in T$ (=1) or not (=0)

The binary variable generates the schedule for every PM-plan during the planning horizon. Budai-Balke (2009) makes use of another binary variable that denotes if a moment is in use for preventive maintenance such that the model can assess the "possession costs". In the context of this research this is not necessary due to the narrow scope. Section 2.3.1 arguments that not all PM-plans need counter-based maintenance. Therefore, all moments are still needed for other types of preventive maintenance activities, meaning assessing the possession costs of a stop-day would not generate valid results.

Objective function

Minimize:

$$mc * \sum_{p=1}^{P}\sum_{t=1}^{T} x_{tp} + pc * \sum_{p=1}^{P}\left(FM_p - \sum_{t=1}^{FM_p} t * x_{tp}\right) + pc * \sum_{p=1}^{P}\left(T - \sum_{t=(LM_p)}^{T} t * x_{tp}\right) \tag{4.3}$$

The first expression is the maintenance costs times the sum of the total number of executed PM-plans across the moments. A minimum of the total executed PM-plans across the planning horizon relates to the main goal to minimize over-maintenance and maintenance costs. The second expression calculates for all PM-plans the difference between the best possible moment and the planned first moment of the solution, and multiplies this to the penalty costs. Coincidentally, the last expression calculates for all PM-plans the difference between the last possible moment and the last planned moment of the solution, and multiplies this to the similar amount of penalty costs. These expressions guarantee that the first and last planned moment are as close to the best possible moments as possible, considering the difference between passed expected production and their corresponding counter-based intervals.

Constraints

$$\sum_{t=1}^{FM_p} x_{tp} = 1, \qquad \forall p \in P \tag{4.4}$$

$$\sum_{s=1}^{M_{tp}-t} x_{(t+s)p} \geq 1, \qquad \forall p \in P, 1 \leq t \leq LM_p \tag{4.5}$$

$$\sum_{t=LM_p}^{T} x_{tp} = 1, \qquad \forall p \in P \tag{4.6}$$

$$\sum_{p=1}^{P} x_{tp} \leq C_t, \qquad \forall t \in T \tag{4.7}$$

$$x_{tm} = x_{tn}, \qquad \forall (m,n) \in A, \forall t \in T \tag{4.8}$$

$$x_{tp} \in \{0,1\}, \qquad \forall p \in P, \forall t \in T \tag{4.9}$$

Constraints 4.4 ensure that the first moment for each PM-plan is scheduled before the passed expected production reaches their counter-based interval. Constraints 4.5 guarantee that after a PM-plan is scheduled, it is scheduled again before the passed expected production reaches the counter-based interval, until the end of the planning horizon. Constraints 4.6 provide that the last moment for each PM-plan is scheduled before the expected production of the following moments reaches the counter-based interval. Constraints 4.7 ensure that the number of scheduled PM-plans on any moment/stop-day does not exceed its capacity, e.g. no more than 30 PM-plans allowed on a stop-day. Constraints 4.8 provide that PM-plans that must be clustered will always be grouped in the schedule. Finally, Constraints 4.9 define the decision variables as binary.

4.4.5 Rolling Horizon

A rolling horizon is common for problems in dynamic and unpredictable environments. This approach involves making decisions based on available information about the future over multiple periods. For instance, at the beginning of a project, plans are made for the upcoming months or years. As time progresses and the project enters subsequent phases, these plans are revised based on new information. Figure 4.3 shows how the rolling horizon scheduling strategy operates.

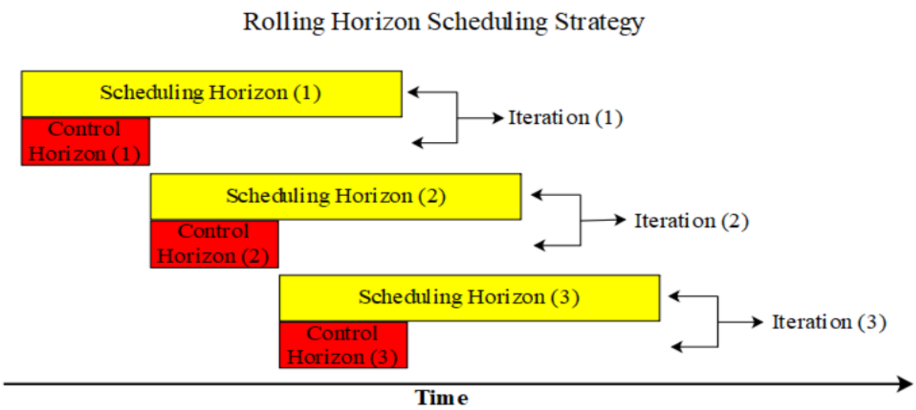


FIGURE 4.3: Rolling horizon scheduling strategy operation (Ott, Almuhaini, and Khalid, 2019)

It is an iterative process, where the model finds a solution for the entire scheduling horizon, which can be adjusted after the control horizon. Each iteration the model reschedules to account for new information. This provides flexibility and adaptability by allowing for continuous adjustments based on the latest information, ensuring plans remain relevant and effective. This is extra beneficial in the context of our goal, because our parameters are based on the results of a prediction model which will have some level of errors. The inclusion of a rolling horizon benefits this by reducing the amount of uncertainty in the scheduling process.

For this planning model, this research sets the planning horizon at 52 weeks, and the control horizon at 13 weeks. Do note that although the control horizon is similar to the aggregation level of the first prediction model in Section 4.3, this planning model uses the predictions of the non aggregated values. The planning model should observe the specific values of production between the moments for preventive maintenance, hence the predictions of aggregated values are invaluable for this. The definitions of the planning and control horizon are sufficient because there is no need to schedule further than a year into the future because the uncertainties would increase too much. The control horizon also coincides with the time taken for preparation of maintenance activities.

4.5 Conclusion

This chapter describes the methods of collecting, selecting, and processing data to create the design of the solutions. The first solution delivers the method of determining the usage counters, with the help of equations 4.1 and 4.2. The usage counter indicates how close the PM-plan is to its interval, making it possible to monitor its use. The company can obtain counter-based intervals through a simple formula that translates time-based intervals into production-hour intervals. Given that preventive maintenance plans are typically scheduled 13 weeks in advance, the team should perform these calculations on a monthly basis to ensure timely preparations. Moreover, this research suggests further investigation on automatically tracking these counters to further improve the monitoring of maintenance activities.

For the Maintenance Engineering Team to know when the amount of production in the upcoming 13 weeks exceeds the counter-based interval, they must gather data from various software systems, particularly the Manufacturing Execution System (MES) and the Tactical Planning Team's central database. The MES data provides daily production states per machine, while the Tactical Planning Team's records offer weekly production hours. This chapter suggests several predictive models using the dataset for line 81 from 2016 to 2024. For production forecasting, this chapter suggests two prediction model objectives: one that aggregates production data over 13 weeks to provide maintenance engineers with insights on when PM-plans should be activated, and another that predicts the expected production each week for the upcoming 52 weeks to aid in planning multiple PM-plans. The first is useful for the Maintenance Engineering Team at any moment during the year if they wish to check if any specific PM-plan will exceed its counter-based interval, while the latter is beneficial for knowledge on weekly production in a planning model of all PM-plans.

To continue with this planning model, the planning of preventive maintenance activities should follow a mathematical model that minimizes over- and under-maintenance by scheduling PM-plans as close to their optimal moments as possible. The mathematical model of this thesis takes inspiration from both machine scheduling problems, as well as the PMSP and RPMSP of Budai-Balke (2009). It uses concepts such as maintenance costs, and penalty costs for over-maintenance. Constraints on clustering, available capacity per stop-day and the counter-based intervals for each PM-plan ensure that maintenance is planned as effective as possible, reducing both the risk of equipment failure and unnecessary maintenance. The model follows a rolling horizon approach, used in dynamic and unpredictable environments, which is beneficial in reducing the scheduling uncertainty from the prediction model errors. The resulting planning model provides a robust framework that aligns preventive maintenance with predicted production hours, ensuring optimal scheduling.

Chapter 5

Results

This chapter addresses the fourth research question:

"What are the results of the solution design and how do they affect the performance of the maintenance activities?"

It explores the performances of the solutions of this research. The research question unfolds into several sub-questions, including:

- *How do the new usage counters improve the current situation?*
- *What are the results of the predictive models?*
- *What are the results of the planning model?*
- *How do the solutions improve the current situation?*
- *How sensitive is the model to new situations?*

Section 5.1 examines the results of the new usage counters, and Section 5.2 addresses the results of the predictive models. Then, Section 5.3 provides the results of the planning model, and Section 5.4 discusses how these solutions improve the current situation. Section 5.5 performs a sensitivity analysis, and Section 5.5 concludes the chapter.

5.1 Usage Counters Results

This research has the aim to deliver a clear step-by-step methodology for the Maintenance Engineering Team to transition to a counter-based strategy. This method should be applicable for all production filling lines. Appendix B.1 delivers the step-by-step method to create the spreadsheet that shows the counters of the PM-plans for any production filling line. Appendix B.2 gives an example screenshot of such a spreadsheet. This spreadsheet shows both the progression of the time-based strategy and the counter-based strategy, where the time-based progress entails the proportion of calendar days passed to the time-based interval in days, and the counter-based progress is equal to the usage counter from equations 4.1 and 4.2. Note that the time-based numbers relate to the current maintenance strategy, while the counter-based numbers are what this research suggests as a solution.

TABLE 5.1: Five results from counters spreadsheet for line 81 (made 24-6-2024)

Code	Description	Last execution	Time-based progress	Interval (hours)	Counter (hours)	Counter-based progress
PM1	SIXO812.1, 1J	2-11-2023	64%	3,375	1,942	58%
PM1	Losdok81, 6M	26-1-2024	82%	2,447	2,165	88%
PM1	TRANS-VL-DS81, 6M	30-1-2024	80%	1,688	1,266	75%
PM2	DOZO81, 1J	24-8-2023	84%	3,375	2,331	69%
PM4	ETIMA, 3J	31-3-2022	75%	15,098	12,176	81%

Table 5.1 reveals that the counter-based progressions are lower or higher than the time-based progressions, meaning that considering the real production times of the machines, maintenance should not be planned as early or late as with the time-based strategy. Take for example the fourth PM-plan in the table. Regarding its time-based interval, it will be scheduled in the near future because it is already at 84%, while the counter-based progress is only at 69%. Executing this PM-plan too early is an example of over-maintenance. These examples in the table indicate that the company could perform less over- and under-maintenance if they consider the counter-based percentages from these spreadsheets for the packaging filling lines, instead of their current time-based systems.

5.2 Prediction Models Results

This section delivers the results of the prediction models. To be clear, there are two types of predictions, each having their own goal: (i) Provide knowledge on when to plan any PM-plan with enough time (13 weeks) for preparations. (ii) Improve the planning activities of multiple PM-plans in the upcoming period by providing the expected amount of production in between stop-days. The first goal uses the dataset with the aggregated data, and the second with no aggregation. Section 5.2.1 delivers the results of the aggregated predictions, and Section 5.2.2 provides the results of non aggregated predictions. Then, Section 5.2.3 continues discussing learning curves. Section 5.2.4 discusses the importance of the features of the best model types, and Section 5.2.5 gives some insights on graphs that show the performances of the best model types. Note that Appendix D includes the remaining performance metrics per machine.

5.2.1 Prediction Results of the Aggregated Data

To find the best prediction method for the aggregated values, this section evaluates the performances based on the evaluation metrics. Either the RMSE, MAD, sMAPE, or the Bias (see Section 3.5.2) give an indication on which prediction model outperforms the others, in addition to stating if there are over- or underestimations. These metrics are based on how the model performed on the test data set. Table 5.2 shows the RMSE performances of the model types on the test data set. The cells that are coloured indicate the lowest value out of all the model types.

TABLE 5.2: RMSE Values per Machine for Aggregated Predictions

Machine	Baseline	LR	DT	RF	GBM	KNN	HW
Machine 1	16962.28	15652.42	12562.92	12749.03	13824.30	14092.99	9828.33
Machine 2	12111.09	10104.90	12271.43	9511.08	11640.18	10407.77	9888.67
Machine 3	16748.30	15326.59	16642.06	12170.41	13588.41	17084.87	14927.61
Machine 4	15848.00	14742.76	14468.52	13393.70	12704.25	17009.05	14958.46
Machine 5	15996.98	15370.80	12968.91	12316.56	12943.69	17188.76	15734.29
Machine 6	12252.22	10747.84	12990.92	10566.26	10654.10	11003.18	10122.87
Machine 7	12545.61	10878.37	13545.70	10933.23	11145.53	11383.73	10211.33
Machine 8	12890.92	11078.07	13313.91	10959.58	11125.92	11627.22	10997.85
Machine 9	11319.29	10416.53	11181.43	9576.71	10229.83	11047.92	12043.69
Machine 10	11303.45	9044.17	10926.83	8563.96	9588.10	9992.25	10314.83
Machine 11	18766.12	16740.88	14835.50	13798.97	15045.69	19042.01	19422.50
Machine 12	11005.63	14284.13	10995.94	9423.34	9483.66	9948.40	15866.67
Machine 13	15061.18	13238.92	13763.90	12651.37	12596.22	13343.21	13036.14
Machine 14	11578.96	10956.42	14685.63	11872.06	11511.25	11526.54	10630.38
Machine 15	12129.14	10241.82	12402.09	9659.85	9657.13	9866.88	9270.64
Machine 16	11751.22	9682.54	9825.24	8875.62	9196.19	9561.40	8894.34
Machine 17	12812.09	10365.74	12652.41	9746.79	9251.68	12091.73	9844.29
Machine 18	14077.78	9927.78	10723.97	9912.74	9513.15	11337.03	12281.75
Machine 19	10741.07	9958.27	10272.35	10814.90	12174.48	10539.99	11665.11
Machine 20	11772.32	10188.67	10791.49	10143.33	9714.60	10903.29	11571.03
Machine 21	15962.12	15519.97	17536.11	12501.19	13075.00	17382.65	16183.20
Machine 22	20474.81	12003.11	20726.20	9718.33	11881.64	20202.12	15925.29
Machine 23	21115.02	12354.30	21418.71	10141.56	12190.40	20746.78	15719.86
Machine 24	11745.63	10385.27	10970.52	9952.01	10358.18	10816.97	11660.15
Machine 25	12571.42	11925.60	15176.20	12224.11	12188.88	12184.96	11216.86
Machine 26	12074.02	11156.35	13351.63	10577.59	11166.84	11295.15	10552.90

The values in Table 5.2 indicate the RMSE value of a model type per machine, representing the closeness of the predicted values to the actual values. For example, the first machine has a RMSE value of about 15,500 considering the Linear Regression model, while the Holt-Winters model has a value of about 10,000, which is the lowest and therefore better value for this machine. Table 5.2 shows that for all machines the model types that use prediction modelling techniques outperform the baseline, which assumes constant production.

Another observation is the fact that some machines have significantly worse performances in comparison to others. For example, all models have explainable difficulty in predicting the aggregated values of Machine 3, the Box Packer, since not every product goes through this machine. The expected production amount of this specific machine depends on the product type. Therefore, the prediction model would need extra input to possibly account for these variations. The company lacks data on these values, which makes it difficult at the present moment to generate accurate predictions for these machines. The models do however perform well for Machine 20, the Pasteurizing Machine. This machine simply processes all products, which makes for better predictions.

In the context of the problem, it is not wise to recommend a different prediction method per machine, since the implementation of such a system is too complex for the company. Therefore, this section demands an extra evaluation of the average performances over all machine, which also provides a more comprehensive overview of all performance metrics. Table 5.3 shows these average performances on the test data set, including the RMSE, MAD, sMAPE, and the Bias. The cells that are coloured indicate the best value out of all the model types.

TABLE 5.3: Average Performance Metrics of Model Types for Aggregated Data

Model	RMSE	MAD	sMAPE	Bias
Baseline	13908.33	11873.16	19.65%	-9351.49
LR	12011.24	9929.66	16.25%	-3829.79
DT	13500.02	11229.74	18.99%	-7084.88
RF	10875.16	9466.41	15.43%	-5224.77
GBM	11401.90	10029.58	16.40%	-6301.94
KNN	13139.49	10926.30	18.68%	-6478.31
HW	12414.19	9970.86	15.14%	7103.20

Considering the RMSE and the MAD, the Random Forest model performs the best out of all the models, while observing the sMAPE, it shows that the Holt-Winters model performs the best. All models except for Holt-Winters seem to be underestimating, with the lowest deviation from 0 by the Linear Regression model. Comparing the metrics of RF with HW, it shows that the HW model is worse considering the RMSE and MAD, while the difference of the sMAPE is relatively low. The LR model also performs worse on most metrics in comparison to the RF model. From these facts, the research can conclude that the Random Forest model performs the best out of all the model types, for the aggregated predictions.

5.2.2 Prediction Results of the Non Aggregated Data

Remaining is the evaluation of the performance metrics to find the best model for predicting the amount of production per week for the upcoming 52 weeks. Table 5.4 shows the RMSE performances of the model types on the test data set. The cells that are coloured indicate the lowest value out of all the model types.

TABLE 5.4: RMSE Values per Machine for Non Aggregated Predictions

RMSE	Baseline	LR	DT	RF	GBM	KNN	HW
Machine 1	2291.99	2139.76	2041.42	2054.18	2141.48	2210.02	2328.47
Machine 2	2004.56	1652.51	2125.72	1725.27	1756.21	1826.33	2175.16
Machine 3	2159.42	2050.17	2067.05	1998.93	2042.99	2274.05	2060.22
Machine 4	2146.53	2026.35	2066.56	2002.95	2036.08	2276.77	2053.13
Machine 5	2208.95	2110.48	2218.05	2041.36	2104.31	2352.77	2146.52
Machine 6	1840.98	1550.50	1809.86	1601.17	1583.21	1671.17	1987.74
Machine 7	1857.20	1566.39	1617.92	1608.13	1597.88	1692.34	1954.11
Machine 8	1982.00	1604.28	1857.57	1696.47	1698.44	1780.31	2069.51
Machine 9	1976.50	1596.96	1607.78	1713.43	1688.04	1780.99	1972.99
Machine 10	1916.84	1528.53	1633.60	1597.83	1594.41	1693.62	1765.38
Machine 11	2520.83	2354.77	2597.84	2325.92	2383.50	2650.89	2959.89
Machine 12	2290.10	2060.21	1937.74	2074.49	2100.07	2116.24	3492.55
Machine 13	1980.15	1686.16	1742.17	1678.74	1688.15	1762.64	2079.82
Machine 14	1575.15	1662.66	1648.45	1566.44	1557.85	1750.91	1806.73
Machine 15	1969.76	1601.94	1678.42	1643.78	1634.75	1693.68	1913.47
Machine 16	1941.36	1545.18	1795.29	1601.03	1582.52	1667.28	1885.31
Machine 17	2053.83	1751.97	2073.64	1832.49	1763.79	2170.56	3369.32
Machine 18	2161.78	1693.60	1977.95	1840.45	1888.01	1954.42	2181.30
Machine 19	1932.13	1610.98	1948.38	1898.62	1864.35	1860.33	2053.51
Machine 20	2000.86	1623.76	1664.88	1718.20	1716.88	1888.19	2106.85
Machine 21	2160.16	2082.65	2185.14	2005.62	2063.88	2305.45	2116.52
Machine 22	2388.07	1977.22	2101.21	1924.81	2021.50	2366.14	2120.57
Machine 23	2445.88	2021.35	2152.24	1968.99	2069.95	2407.69	2186.23
Machine 24	1968.19	1580.75	1607.36	1674.95	1670.20	1764.98	1986.18
Machine 25	1657.87	1742.22	1693.24	1636.47	1631.58	1825.18	1843.36
Machine 26	1827.31	1573.99	1553.94	1601.20	1578.59	1650.83	1982.10

The values in Table 5.4 indicate the RMSE value of a model type per machine, again representing the closeness of the predicted values to the actual values. Initially it shows that these values are much lower in comparison to the aggregated predictions, because these non aggregated values are simply lower. Now mostly the LR, DT, and RF models perform best across the machines. For similar reasons as in the previous section, some machines perform better in comparison to others. Another interesting fact is the worsening of the HW performances. It seems that aggregating data increases the benefits of a HW model.

Again, an extra evaluation of the average performances per machine provides a more comprehensive overview over all performance metrics. Table 5.5 shows these average performances of the test data set, including the RMSE, MAD, sMAPE, and the Bias. The cells that are coloured indicate the lowest value out of all the model types.

TABLE 5.5: Average Performance Metrics of Model Types for Non Aggregated Data

Model	RMSE	MAD	sMAPE	Bias
Baseline	2048.40	1610.90	38.65%	-615.61
LR	1784.44	1406.62	36.66%	-431.28
DT	1900.13	1464.80	37.62%	-402.50
RF	1808.92	1437.25	36.18%	-391.02
GBM	1825.33	1453.03	36.87%	-478.37
KNN	1976.68	1539.52	39.44%	-430.09
HW	2176.81	1693.78	40.71%	-276.97

The sMAPE values show that these performances are worse comparing to the predictions of the aggregated data set, showing that increasing the aggregation level of variable outcomes could improve performances. Considering the RMSE and the MAD, the Linear Regression model performs the best out of all the models, while observing the sMAPE, it shows that the Random Forest model performs the best. All models seem to be underestimating, with the HW model having a Bias value closest to zero. Although all metrics are very similar, the research can conclude that the Linear Regression model performs the best out of all the model types, for the aggregated predictions.

5.2.3 Learning Curves

This subsection discusses the *learning curves* of some of the outcomes of the prediction models. Brownlee (2019) reviews learning curves in machine learning, which are graphical representations that show the performance of a model as a function of the amount of training data. He further states that they help in diagnosing the behavior of machine learning models, especially in terms of understanding how well a model is learning and how it might perform on new data. A typical learning curve figure shows two curves, one for the training error and another for the test error. All machines show similar learning curves, some with a little more variation than the other, meaning that this section selects a single machine that displays the curve in a clear manner. Figure 5.1 shows the learning curves of the RF model of the aggregated values for Machine 22.

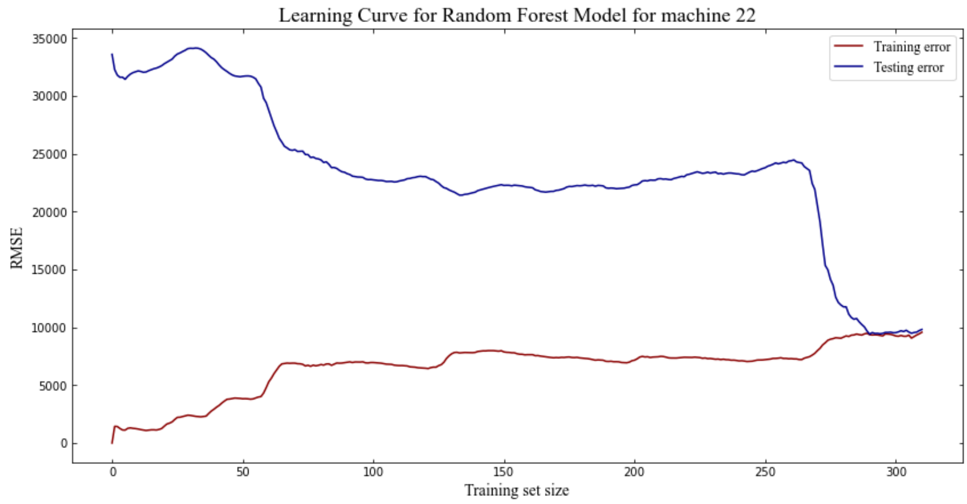


FIGURE 5.1: Learning Curve of RF Model (aggregated values) for Machine 22

Figure 5.1 shows how increasing the training set size reduces the RMSE value of the testing data set. There is little to gap at the end of the curve, indicating that the model has a suitable fit. Still, the convergence of the curves is very short, meaning that bringing more data could help in further establishing the performance of the model. Subsequently, Figure 5.2 shows the learning curves of the LR model of the non aggregated values for Machine 26.

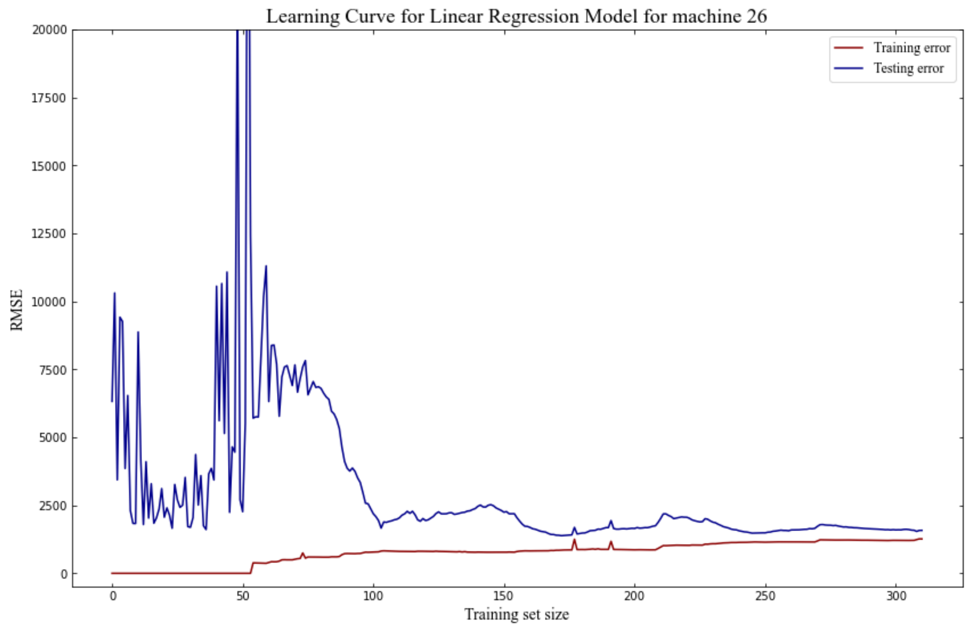


FIGURE 5.2: Learning Curve of LR Model (non aggregated values) for Machine 26

In Figure 5.2, the training error starts low and remains relatively flat as the training set size increases. This indicates that the model has a good fit with the training data. The occasional spikes in training error suggest there might be specific subsets of data that are particularly difficult for the model to fit. There are fluctuations in the testing error, particularly at smaller training sizes, which stabilize as the training set size increases. These fluctuations may be due to the variability in the small training sets or specific challenges within the dataset. Both curves show do not exhibit signs of over-fitting or under-fitting, since they converge with a small gap.

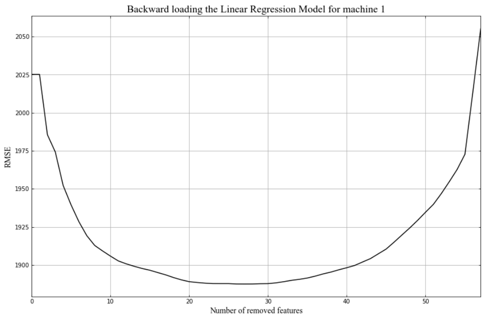
The two figures of this section indicate that for these specific cases, the prediction modelling techniques show sufficient predictions without an excess of over- or under-fitting. Both figures show the validation error decreasing, converging, and approaching the training error as more data is added. This suggests that adding more data will not significantly improve model performance. This indicates that the dataset is sufficient for the current model. Note that all machines have somewhat similar learning curves, but concluding that all are sufficient would be unscientific. These two examples prove that the prediction method is sufficient for these specific datasets.

5.2.4 Feature Importance

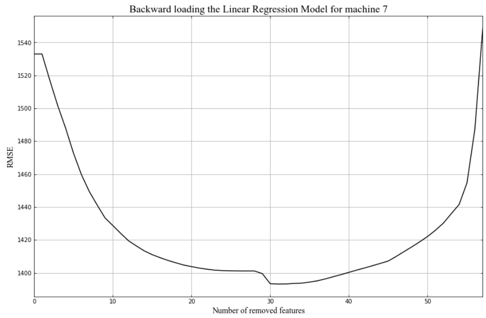
This section aims to find information on the importance of the features in the dataset. This helps to better understand how the models use the information of the features and which attributes have the biggest impact on the performances of the models. One method of evaluating feature importance is *backward loading*. James et al. (2013) calls it backward selection, which in the context of a prediction modelling involves evaluating the impact of each feature on the target variable by systematically removing features from the model and observing the changes in model performance. This method can help identify the most important features.

According to James et al. (2013), backward loading starts by training a prediction model type with all features and evaluating its performance using a metric, for which this section chooses the RMSE. For each feature, it temporarily removes the feature, retrains the model, and evaluates the performance. The feature whose removal results in the smallest increase in the RMSE is permanently removed. This process repeats iteratively, with the model being retrained and evaluated after each removal. At each step, the current set of features, the performance metric, and the removed feature are recorded. This continues until no features remain, resulting in a detailed list of feature importance based on their impact on model performance.

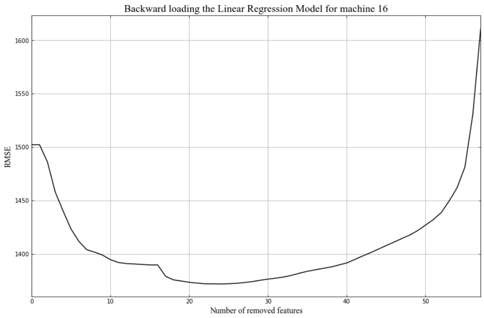
The results from backward loading will show how the model's RMSE changes as features are removed. Features that, when removed, result in a significant increase in RMSE can be considered important. It turns out that the progression of the RMSE values as features are removed is somewhat similar for each machine, but the order of features removed are very different. A random selection of four different machines delivers sufficient insights. Figure 5.3 shows the backward loading progression of the LR model for the non aggregated values of four randomly selected machines.



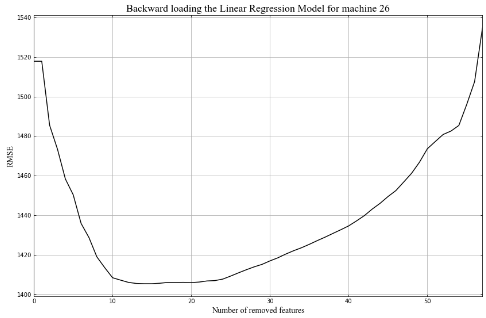
(A) Machine 1



(B) Machine 7



(C) Machine 16



(D) Machine 26

FIGURE 5.3: Backward Loading progression of the LR Model (non aggregated values)

Figure 5.3 shows that there is the initial phase of a decreasing RMSE value, leading up to an optimal point after which it starts increasing. At the beginning the RMSE is relatively high, indicating the model might be over fitting. As it removes features, the RMSE decreases, reaching a minimum point. This indicates the optimal number of features for the model, where it has the best performance with the lowest RMSE. Beyond the optimal point, removing more features causes the RMSE to increase again, suggesting that the model is losing important information and starting to under fit.

Figure 5.3 also demonstrates that backward loading can be an effective method for feature selection in linear regression models. However, it is crucial to identify the optimal point of feature removal to ensure the best model performance. The specific number of features that should be removed to achieve this optimal point varies between different machines. The selection of features could change as the dataset grows, making it very complex and time extensive to find the optimal combination for each machine in the future.

The order of the removed features differs between the machines. For some, the feature that represents the year is the least important, while for others the different categorical week features are removed first. For all machines it seems that the tactical planning feature is always kept until the very end, meaning it can be seen as the most important feature. Unsurprisingly, the tactical planning feature gives the model a general direction of what amount of production can be expected.

Based on these findings, this research chooses not to remove any features to improve the performances of the model because of the following reason: The company wishes to find a sufficient, general, and simple manner of predicting production hours per machine, and finding the optimal set of features is a complex and time extensive progress, if this is done for all machines individually. Regarding the final objective of this solution, the RMSE values of the models with all features are not significantly higher in comparison to the RMSE values of the models with an optimal set of features. In simpler words: The models that include all features get the job done. Still, the backwards loading method does deliver insights in the possibility of further improving the predictions, which the company should not neglect.

5.2.5 Visual Performance

This section discusses the insights to gain from the graphs that represent the performances of the best prediction models. The aim is to create a feeling on what the prediction methods are trying to achieve and how they are accomplishing this. With two datasets (aggregated and non aggregated), seven model types, and 26 machines in total, there are a lot of graphs to analyze. Some machines have better results in comparison to others, and this section selects a few to discuss.

After the models learn from the training data, they try to predict the values of the testing data. To remind: the testing data contains the weekly values of the years 2022 and 2023. These values represent the amount of production per machine, measured in minutes. Graphs that show the actual values next to the predicted values provide a visual performance of a prediction model. Ideally, a good performance correlates to a graph where the two lines closely match and stay around the same values. For some machines these values are highly variable. Therefore, representing cumulative values increases the clarity of these figures.

Aggregated Data

For the predictions of the aggregated data, this section chooses to evaluate Machine 3 and Machine 20. These two are chosen because Section 5.2.1 already concludes that Machine 3, the Box Packer, has a bad performance based on the evaluation metrics, while Machine 20, the Pasteurizing Machine, has a good performance. This difference originates from the fact that the production of the Box Packer depends on the type of product through the line, while this is not the case for the Pasteurizing Machine. Now, this section can confirm these performances with visual graphs. Figure 5.4 shows the graphs of the aggregated actual and predicted values of the RF model for Machine 3.

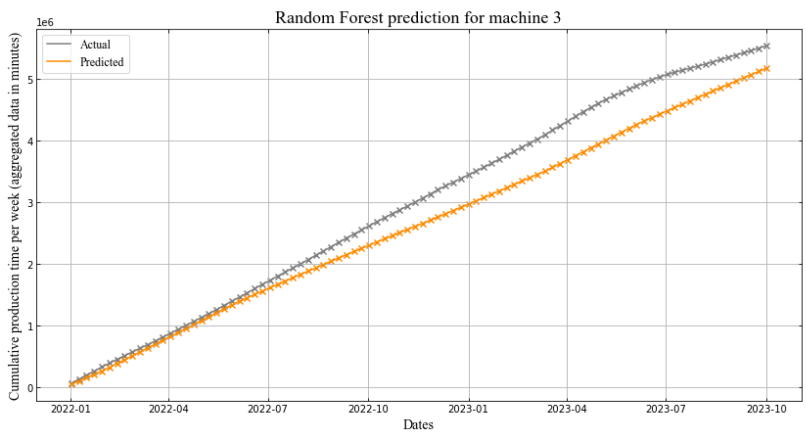


FIGURE 5.4: Visual Performance of RF for Machine 3

Figure 5.4 shows how the cumulative production time per week for machine 20 from January 2022 to October 2023, with actual times in grey and Random Forest model predictions in orange. Figure 5.4 reveals that the Random Forest model's predictions closely follow the actual trend but makes some underestimations, which results in the orange being under the grey line. This is also observed in the table that gives the Bias values in Appendix D, which are negative. The underestimation is more present when predicting further into the future due to the cumulative aspect. This would argument to make these predictions on a relatively regular basis to update the training data of the model. Figure 5.5 shows a better graph of the aggregated actual and predicted values of the RF model, now for Machine 20.

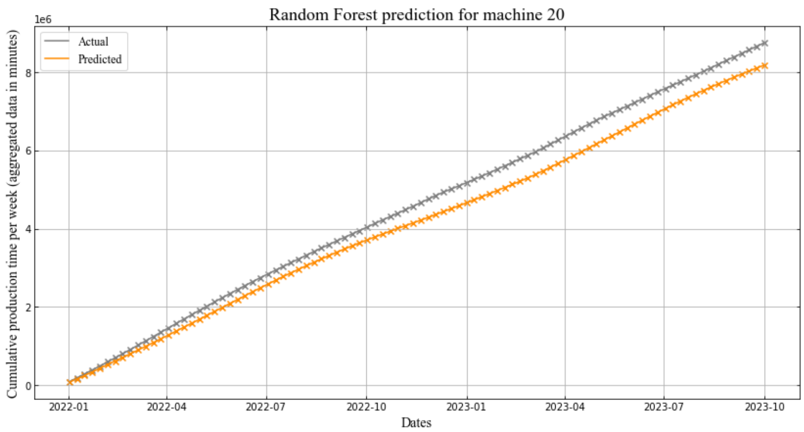


FIGURE 5.5: Visual Performance of RF for Machine 20

Again, Figure 5.5 demonstrates that the Random Forest model is slightly underestimating the real values of production. It shows that the model does perform better for this machine, since the lines lie closer together, but after a year from the start it becomes increasingly difficult to accurately predict, due to the cumulative error. An updating mechanism in a time series prediction model could be a solution to this problem. It refers to the process by which the model incorporates new data to update its parameters and forecasts. This mechanism is crucial for maintaining the model's accuracy and relevance over time, especially in dynamic environments where patterns and trends can change.

Non Aggregated Data

For the predictions of the non aggregated data, this section chooses to evaluate Machine 14 and Machine 19. These two are chosen because Section 5.2.2 already concludes that Machine 14, the Multi Packer, has a bad performance based on the RMSE, while Machine 19, the Palletiser Machine, has a good performance. Figure 5.6 shows the graph of the non aggregated actual and predicted values of the LR model for Machine 14. Figure 5.6 displays the cumulative production time per week for machine 14 from January 2022 to January 2024, with the actual production times shown in grey and the predicted times using a Linear Regression model shown in orange. Both the actual and predicted values increase, where the predicted values show a consistent line, while the actual values are more variable.

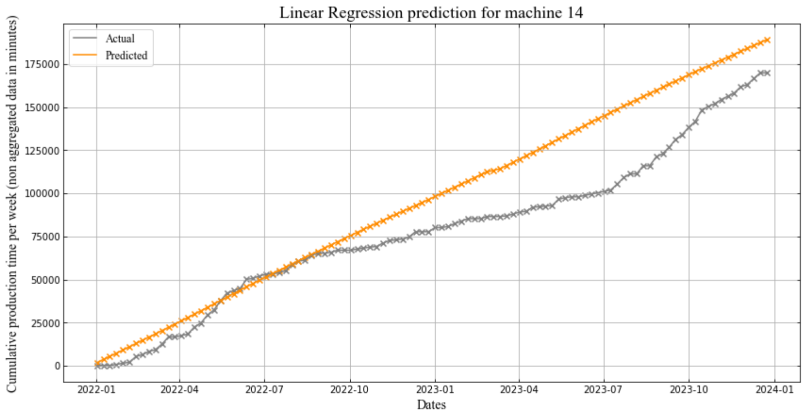


FIGURE 5.6: Visual Performance of LR for Machine 14

There is a clear discrepancy between the two lines: the Linear Regression model consistently overestimates the production time compared to the actual values, meaning it should have a positive Bias, which Appendix D confirms. This suggests that the Linear Regression model may not accurately capture the variability and trends in the actual production data, leading to an overestimation for machine 14. This indicates that a linear model may be too simplistic for this machine. To improve the prediction accuracy, and knowing from Section 5.2.2 that more complex model types did not perform better, incorporating more relevant variables might help in capturing the patterns better. Nonlinear models, e.g. neural networks, often outperform linear models due to their ability to capture complex relationships, recognize patterns, handle variability, and adapt to new data. These models are flexible and can accommodate intricate interactions between multiple factors. They could be better suited for the predictions for machines that have highly variable production.

To continue, Figure 5.7 show a better graph of the aggregated actual and predicted values of the LR model, now for Machine 19. Figure 5.7 shows for both the actual and predicted production times a steady line upwards. The Linear Regression model makes only slight underestimations of the actual production time, as evidenced by the predicted values being lower than the actual values. This time a simple Linear Regression model seems to be well suited for the predictions.

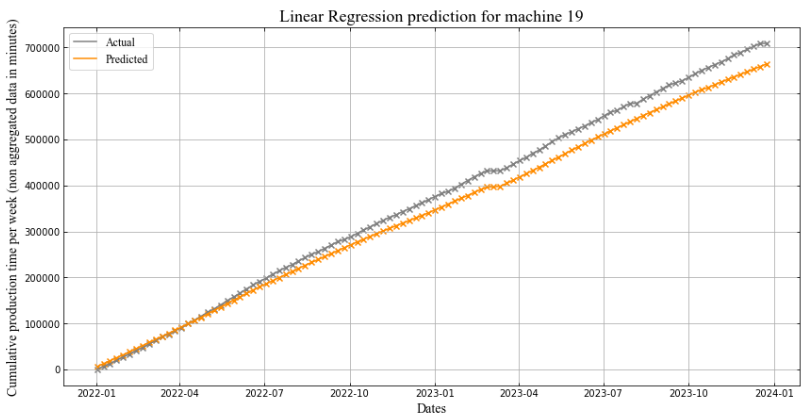


FIGURE 5.7: Visual Performance of LR for Machine 19

Additionally, Appendix E contains the visual performances of this section with non-cumulative values. They display the values on a weekly basis, highlighting any discrepancies or patterns in the predictions, whereas the cumulative figures are useful for understanding the long-term/cumulative accuracy of the model. The non-cumulative figures are intended for the Maintenance Engineering Team of the company. The figures are more difficult to understand, due to the variability of the production, but the team members can recognize aspects such as seasonality or revisions. If the figures of this section fail to help in understanding how they represent the production time of a machine, take a look at Appendix E.

Overall, evaluating the visual performances of the prediction models, there is one main observation: There are some machines with a very variable amount of production during the years due to a dependence on product types that go through the packaging filling line, resulting in bad performances, while other machines have better results. Although this section handles only some model types considering four machines, it creates a general interpretation of the sufficiency of the prediction method. With these results, this research concludes that the prediction method is sufficient enough and can continue with the results of the planning model.

5.3 Planning Model Results

This section provides the results of the planning model from Section 4.4. The objective of the planning model is to schedule the PM-plans across the planning horizon while minimizing over- and under-maintenance. For the sake of demonstration, this section chooses to apply the model on the year 2023 for the packaging filling line 81. The python program Gurobi Optimizer solves the mathematical model, which is readily available for Heineken, depending on the expertise of the employee. The model uses input parameters to generate the schedule for the year 2023 as the output. Section 5.3.1 discusses the input parameters, while Section 5.3.2 delivers the output of the planning model.

5.3.1 Input

The input parameters originate from several sections of this research. Firstly, the PM-plan parameters (see Section 4.4.4) consist of: the ID, description, correlated machine, the start state (S_p), the moment for planning the first execution (FM_p), and the moment for planning the last execution (LM_p). The different software systems deliver the given PM-plans and their corresponding parameter values. Correlated machines per PM-plan are necessary for the calculations that involve production amounts. The start state is found by adding the amount of production from the last execution of the PM-plan up until the start of the planning horizon. Table 5.6 shows some of the in total 169 PM-plans parameters. Note that only PM-plans for which the counter-based interval was constant in the year 2023 are included because otherwise the deviation from this interval is justified.

To explain Table 5.6, take for example the first row. This PM-plan has the ID 21060035 with the corresponding description. It is an inspection activity an its time-based interval is three months. The machine that corresponds to this plan is Machine 13, the Defoiling Machine. Since the last execution of this PM-plan this machine has been running about 270 hours as indicated by S_p. With the expected amount of production from the prediction model (the LR model) of the previous section, in

TABLE 5.6: The first four PM-plan parameters

ID	Description	Machine	S_p (hours)	FM_p	LM_p
21060035	PM1 Codeer App. Programma keuze list 3M	13	271.76	7	18
21056979	PM1 DEFO81 6M	26	821.40	10	12
21053588	PM1 DEFO81 E & Î 6M	26	1052.68	9	12
21053589	PM1 DEFO810 Wtb 6M	26	2136.34	4	12
	...				

addition to the calculated counter-based intervals of Section 5.1, it finds that the first execution of this PM-plan should be scheduled before the seventh moment, while the last execution should be scheduled after the 18th moment.

Secondly, there are moment parameters to define, which are opportunities for maintenance during the year 2023, also known as stop-days. They consist of: the moment number, week number, and capacity (C_t). The week numbers are necessary for the rolling horizon. During 2023, there were in total 25 moments/stop-days for maintenance. Researching past historical data gave an impression on adequate capacity values. For now all moments get a capacity of 30 PM-plans at a maximum, except for the single moment when there was a revision (the fifth moment), which gets a capacity of 80. Table 5.7 shows some of the moment parameters.

TABLE 5.7: The first four moment parameters

Moment	Week Number	C_t
1	2	30
2	4	30
3	6	30
4	8	30
	...	

Another parameter is the moment before the expected production of following moments after t reaches the counter-based interval (M_{tp}). These are calculated in the python script. The packaging manager and the software system gave confidential information on historical definitions for down-time and maintenance costs. This research will not use the exact values, but defines them close to the historical values. Therefore, the maintenance costs at €200.-, and the penalty costs at €50.-.

Additionally, in a meeting with the maintenance planner, information was gathered about which PM-plans were always clustered when scheduling the activities during the year. Currently, clustering of activities is entirely dependent on the knowledge and skill of the maintenance planner. He states that for two PM-plans to be clustered, they must satisfy three requirements: (i) both PM-plans are part of the same machine (except for the filler); (ii) both PM-plans must have the same interval; (iii) the last execution of both PM-plans must be around a similar point in time. Following these requirements a set of PM-plans that should be clustered were found. Take for example the second and third row in Table 5.6 and observe how they have a similar machine, a similar interval, and a start state around the same value. These two PM-plans are one out of 31 sets of PM-plans that should be clustered.

Another aspect is the rolling horizon. The scheduling horizon is set at one year (2023), while the control horizon is set at 13 weeks. This means there are four executions of the planning model, at the beginning of each quarter of the year. Before the execution of the planning model all parameters are recalculated given the results of the previous quarters.

5.3.2 Output

With the given input parameters, finally the planning model can deliver a solution that delivers the optimal schedule that reduces under- and over-maintenance. This section discusses the results of the planning model, how to evaluate the performance of a schedule, and what insights to gather from these outputs.

Solution Schedule

During the year 2023, there were four executions of the planning model. After the four executions and storing each control horizon, the program stores a single schedule. Table 5.8 shows some values of this schedule. For example, the PM-plan with ID 21060035 (see second column) is scheduled on the 6th, 10th, 15th, and 21st stop-day during the year 2023.

TABLE 5.8: Planning model solution for the year 2023

	PM-plan IDs					
Moment	**21060035**	**21056979**	**21053588**	**21053589**	**...**	**21059544**
1	0	0	0	0	...	0
2	0	0	0	0	...	1
3	0	0	0	1	...	0
4	0	0	0	0	...	0
5	0	0	0	0	...	0
6	1	0	0	0	...	0
7	0	0	0	0	...	0
8	0	1	1	0	...	0
9	0	0	0	0	...	0
10	1	0	0	0	...	0
11	0	0	0	0	...	0
12	0	1	1	1	...	0
13	0	0	0	0	...	0
14	0	0	0	0	...	0
15	1	0	0	0	...	0
16	0	0	0	0	...	0
17	0	0	0	0	...	0
18	0	0	0	0	...	0
19	0	0	0	0	...	0
20	0	0	0	0	...	0
21	1	0	0	0	...	0
22	0	0	0	0	...	0
23	0	0	0	0	...	0
24	0	0	0	0	...	0
25	0	0	0	0	...	0

Table 5.8 shows how the Planning Model is able to generate a schedule for the maintenance activities of the packaging filling line 81. This schedule considers all restrictions such as clustering aspects, capacity constraints, maintenance costs, and penalty costs for over- and under-maintenance. Each moment represents a stop-day during

the year, and the values (if the value is 1) state if a PM-plan should be executed on that given moment.

Performance Metrics

The program only delivers a schedule and objective function, which are difficult to interpret and compare with the current situation. Therefore, this research provides methods of analyzing the schedule with the real production amounts during the year. Why analyze the amount of over- and under-maintenance with the objective function (some amount of costs), when it is possible to calculate the precise amount of production hours deviating from the counter-based intervals. This section introduces the following equations:

$$NPM = \frac{\sum_{p=1}^{P} \sum_{t=1}^{T} x_{tp} * |PP_{tp} - CI_p|}{\sum_{p=1}^{P} \sum_{t=1}^{T} x_{tp}} \tag{5.1}$$

where, NPM is the Numerical Performance Metric indicating the average amount of hours deviating from the counter-based interval per scheduled PM-plan. x_{tp} is the binary decision variable, PP_{tp} is the passed production since the previous execution of the PM-plan, and CI_p is the counter-based interval of PM-plan p. Subtracting the counter-based interval from the values that represent the amount of production between executions of the PM-plans gives a representation of accuracy of the schedule. These values should be as close to zero as possible, meaning that the amount of production between executions is equal to the counter-based interval, which subsequently means the amount of over- and under-maintenance is low. Summing these values, and dividing it by the total number of executed PM-plans, generates one numerical performance metric for a schedule.

$$OM = \frac{\sum_{p=1}^{P} \sum_{t=1}^{T} x_{tp} * Max\{0, CI_p - PP_{tp}\}}{\sum_{p=1}^{P} \sum_{t=1}^{T} x_{tp} * |PP_{tp} - CI_p|} \tag{5.2}$$

where, OM is the Over-Maintenance indicating the percentage of hours where PM-plans were scheduled too early. x_{tp} is the binary decision variable, PP_{tp} is the passed production since the previous execution of the PM-plan, and CI_p is the counter-based interval of PM-plan p. The formula is similar to the NPM, but here it sums up all the hours that the passed production was lower than the counter-based interval, and divides it by the total amount of hours deviating to arrive at the percentage of over-maintenance.

$$UM = \frac{\sum_{p=1}^{P} \sum_{t=1}^{T} x_{tp} * Max\{0, PP_{tp} - CI_p\}}{\sum_{p=1}^{P} \sum_{t=1}^{T} x_{tp} * |PP_{tp} - CI_p|} \tag{5.3}$$

where, UM is the Under-Maintenance indicating the percentage of hours where PM-plans were scheduled too late. The formula is similar to the OM, but here it sums the hours that the passed production was higher than the counter-based interval.

$$PD = \frac{\sum_{t=1}^{T} m_t}{T} \tag{5.4}$$

where, PD is the Planned Downtime indicating the percentage of moments where the stop-day was in use for preventive maintenance. The formula simply sums up the binary decision variable m_t that denotes if moments are in use, and divides it by the total number of moments T.

$$AvgL = \frac{\sum_{p=1}^{P} \sum_{t=1}^{T} PW_{tp}}{\sum_{p=1}^{P} \sum_{t=1}^{T} x_{tp}} \tag{5.5}$$

where, AvgL is the Average Length indicating the average length in weeks that were between the executions of PM-plans. x_{tp} is the binary decision variable, PW_{tp} is the passed number of weeks since the previous execution of the PM-plan. The formula sums up the length that was between all executions of PM-plans and divides it by the total number of executed PM-plans in the schedule.

$$MC = mc * \sum_{p=1}^{P} \sum_{t=1}^{T} x_{tp} \tag{5.6}$$

where, MC is the total amount of Maintenance Costs. The formula sums up the total number of scheduled PM-plans in the planning horizon, and multiplies it with the cost of one scheduled maintenance activity. This equation makes it possible to evaluate the costs aspect of different schedules.

Solution Schedule Results

Table 5.9 delivers the results of all performance metrics for both the solution of the planning model with the use of a rolling horizon, and without.

TABLE 5.9: Results of Planning Model Solution Schedule of 2023

	PM Schedule with RH	**PM Schedule** without RH
NPM	200.60	223.99
OM/UM	25%/75%	37%/63%
PD	92%	96%
AvgL	17.25	16.80
MC	€74,200.-	€76,600.-

Table 5.9 shows that the use of a Rolling Horizon (RH) is beneficial for the reduction of over- and under-maintenance because of the lower NPM value. This is the direct result of consistently predicting the future production in each quarter, instead of once at the start of the year, reducing the cumulative error during the year. For the schedule with the use of rolling horizon, the average amount of hours deviating from the counter-based interval per scheduled PM-plan (NPM), lies at around 200 hours. The planning model fails to reduce this number to zero because it is only possible to schedule PM-plans on integer intervals, which will always result in some over-and

under-maintenance. The deviation is about 25% of over-maintenance (OM), meaning that the PM-plans are mostly scheduled too late. Without the rolling horizon the deviation is slightly higher, and the division of over- and under-maintenance is more equal. The moments in use for preventive maintenance (PD), and the average length between the executions (AvgL) lie around similar values for both schedules. The total amount of maintenance costs is higher for the schedule without the rolling horizon due to a larger total of scheduled PM-plans. Overall, the method that uses the Rolling Horizon performs better considering all performance metrics.

5.4 Improvement over Current Situation

This section explores how the planning model performs compared to the current situation. With the data of real real executions during the year 2023, it is possible to find the differences between reality and the planning model solution schedule, based on the amount of production between executions. Additionally, it is possible to adjust some parameters to create a solution schedule with a time-based strategy. Section 5.4.1 discusses the results from the historical data, while Section 5.4.2 provides the results from the time-based solution.

5.4.1 Results from Historical Data

Historical data gives us the real moments when PM-plans were scheduled and what effect this had on the performance metrics. This makes it possible to compare these values to the results of both the schedule from the planning model with and without the rolling horizon. Table 5.10 delivers the results of all performance metrics based on historical data, in addition to both the solution of the planning model with the use of a rolling horizon, and without.

TABLE 5.10: Results of Historical Data compared to the Solution Schedule of 2023

	Reality	**PM Schedule** with RH	**PM Schedule** without RH
NPM	359.62	200.60	223.99
OM/UM	48%/52%	25%/75%	37%/63%
PD	100%	92%	96%
AvgL	15.67	17.25	16.80
MC	€74,600.-	€74,200.-	€76,600.-

What the results show, is that the average amount of hours deviating from the counter-based interval per scheduled PM-plan (NPM) is significantly higher in comparison to the other schedules. This could be due to the extra complexity that arises in reality, such as irregularities during the year. Mechanics becoming sick, or unpredictable shortages in spare parts are examples of situations that happen in real life, while the planning model schedule does not consider these irregularities. Also the division of over- and under-maintenance is more evenly distributed in reality, and all moments for preventive maintenance were in use during the entire year. The total amount of maintenance costs of scheduled PM-plans is similar in reality, indicating that the planning model could reduce over- and under-maintenance without increasing the costs.

5.4.2 Results from Time-Based Solution

Another method of evaluating the improvement of the current situation is comparing the current time-based strategy to the newly introduced counter-based strategy. This will deviate from the performances of reality because the Maintenance Planner sometimes deviates a little from the schedule due to irregularities such as mechanics becoming sick, or unpredictable shortages in spare parts. Simply changing a few input parameters creates a planning model that generates a schedule with a time-based strategy. For example, define the start state in weeks since its last execution instead of hours. Use the time-based intervals in weeks instead of the newly calculated counter-based intervals. Finally, instead of predicting the expected amount of production in between moments for preventive maintenance, define these values as the amount of weeks in between these moments. All other parameters will change according to these new input parameters. This creates the possibility to compare the performance of the solution schedule for both the time-based as the counter-based strategy. Table 5.11 delivers the results of all performance metrics of this time-based solution schedule in addition to both the solution of the planning model with the use of a rolling horizon, and without.

TABLE 5.11: Results of Time-Based Solution compared to the Counter-Based Solution Schedule of 2023

	Time-Based Schedule	**PM Schedule** with RH	**PM Schedule** without RH
NPM	295.08	200.60	223.99
OM/UM	22%/78%	25%/75%	37%/63%
PD	92%	92%	96%
AvgL	17.99	17.25	16.80
MC	€71,800.-	€74,200.-	€76,600.-

As expected, it seems that the time-based solution schedule performs worse on most aspects in comparison to the newly introduced counter-based strategy. The average deviation from the counter-based intervals is higher, because this model does not consider any production hours at all. The amount of over- and under-maintenance is also less evenly distributed. The maintenance costs are lower because the time-based strategy schedules less executions of PM-plans in total, which is also why the average length is higher. The company could evaluate if the extra costs are worth the improvements considering over- and under-maintenance.

The entire planning model's purpose was to introduce a method of scheduling PM-plans such that the amount of over- and under-maintenance is minimized. It seems that for the year 2023, this goal has been achieved. Still, this is only an analysis of the year 2023, and this was a normal year with the brewery in production for most of the time. It could be that different years have different performances based on the amount of production in that year.

5.5 Sensitivity Analysis

A sensitivity analysis is a vital tool for enhancing the reliability and applicability of models across various fields by systematically exploring the effects of variability in input parameters. It helps in understanding the robustness of the model and identifying what the effects are on the final outcome. This section chooses to increase the variability in real production amounts across the year 2023 to investigate how this changes the performance outcomes of the planning model. It is expected that a decrease in production in the future could be detrimental regarding over-maintenance. The counter-based planning model could be useful in such a situation. A decrease in production across the year would increase the benefits of the solution designs. This section specifically chooses not to increase the amount of production, because the brewery is expected to lower its production with about 30% in the upcoming years.

To decrease the real production amounts, this section subtracts a random amount to each weekly data point of the year 2023. The random number originates from a uniform distribution with a lower and upper bound. The boundaries of this distribution gives an indication of the extra variability during the year. For example, if these random numbers are between -10 and 0 hours of production per week, the amount of production during the year is slightly decreased, while the boundaries of -100 and 0 hours have a significantly bigger impact. For each week in the year, after subtracting the random number, the value cannot be lower than 0 because negative production hours are not possible. Figure 5.8 shows how two different selections of boundaries of the uniform distribution have different effects on the initial amount of production in the year 2023.

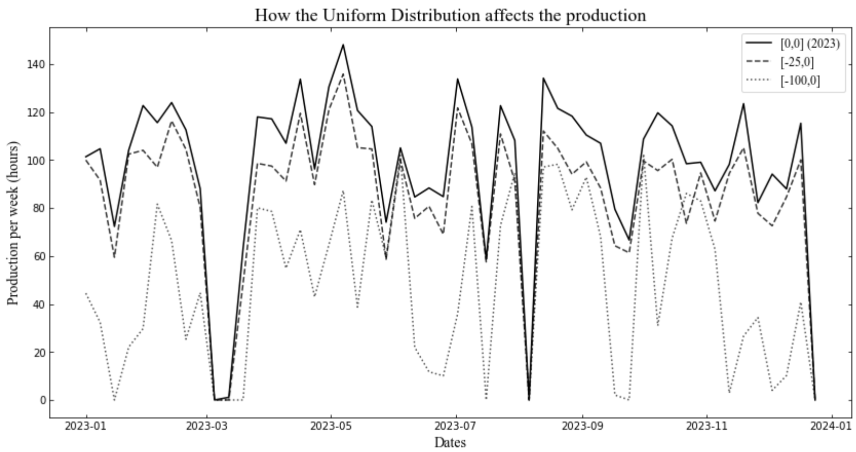


FIGURE 5.8: Visualisation of the effects of the Uniform Distribution

Figure 5.8 illustrates how different uniform distributions affect weekly production over a one-year period. The x-axis represents dates from January 2023 to January 2024, while the y-axis measures production per week in hours. There are three lines representing different uniform distributions. The solid line corresponds to a range of [0, 0], indicating no added variation or randomness in production. The dashed line represents a range of [-25, 0], introducing some (but not much) negative variability. The dotted line shows a range of [-100, 0], introducing significant negative variability. The more negative the range, the lower the overall production per week, and the higher the variability across the year. This type of sensitivity analysis is useful for understanding the robustness of the solution design under different conditions of variability.

With the same parameter settings of Section 5.3, the sensitivity analysis evaluates the performance metrics of the planning model applied to different fabricates years with increasing variability in production. Due to the random distribution, the average of 50 iterations gives a sufficient representation of the performance metrics over the randomly generated years. Table 5.12 shows the NPM and the division of OM with UM, of the different fabricated years with decreasing production.

TABLE 5.12: Results Sensitivity Analysis based on NPM and OM/UM

	NPM (hours)			**OM/UM** (%)		
U. Distr.	**TBM**	**CBM** with RH	**CBM** without RH	**TBM**	**CBM** with RH	**CBM** without RH
[0,0]	295.08	200.60	223.99	22%/78%	25%/75%	37%/63%
[-10,0]	256.07	208.91	210.89	31%/69%	31%/69%	42%/58%
[-25,0]	218.44	202.11	245.93	50%/50%	29%/71%	57%/43%
[-50,0]	278.71	178.58	270.69	78%/22%	35%/65%	70%/30%
[-75,0]	390.48	196.82	336.56	90%/10%	49%/51%	84%/16%
[-100,0]	505.03	188.12	383.99	93%/7%	52%/48%	89%/11%

Table 5.12 provides clear observations of the benefits of the counter-based planning model. As the variability increases and the overall production in the fabricated years decreases, only the schedule of the planning model with the use of the rolling horizon achieves relatively good performances. Tale for example the NPM values, indicating the average deviation from the counter-based interval per scheduled PM-plan. As the boundaries of the Uniform Distribution increases, the NPM value of the time-based schedule also increase. The planning model schedule without the use of the rolling horizon also increase but with lesser effect. Table 5.12 also shows how the schedule of the planning model with the use of a rolling horizon converges to a division of over- and under-maintenance that is evenly distributed. For the other schedules, the decreasing demand during the year leads to a large increase in the percentage of over-maintenance. Table 5.13 continues to show the AvgL and MC values.

TABLE 5.13: Results Sensitivity Analysis based on AvgL and MC

	AvgL (weeks)			**MC** (€)		
U. Distr.	**TBM**	**CBM** with RH	**CBM** without RH	**TBM**	**CBM** with RH	**CBM** without RH
[0,0]	17.99	17.25	16.80	€71,800.-	€74,200.-	€76,600.-
[-10,0]	17.99	17.97	17.33	€71,800.-	€71,016.-	€75,200.-
[-25,0]	17.89	18.71	17.50	€71,600.-	€63,316.-	€72,600.-
[-50,0]	17.80	20.55	18.58	€71,400.-	€55,784.-	€67,000.-
[-75,0]	17.70	21.09	19.55	€71,200.-	€46,016.-	€64,800.-
[-100,0]	17.29	23.19	20.31	€70,400.-	€38,920.-	€59,200.-

Table 5.13 shows similar results. Only the solution schedule of the planning model that uses a rolling horizon properly reacts to the decrease in production during the year. Table 5.13 reveals how the average length in between scheduled PM-plans increases from 17 weeks to 23 weeks, while for the time-based maintenance schedule it stays at around 17 weeks. Additionally, due to the decrease in production, the counter-based maintenance schedules react and reduce the total number of PM-plans scheduled, subsequently reducing the total amount of maintenance costs. Take for example the fifth row. The uniform distribution indicates that on average

each week will have 37.5 hours less production (because the boundaries are [-75,0]), which is somewhat equivalent to a reduction of 5 shifts, a reduction of about 30%. Observing the maintenance costs in this situation, the newly introduced counter-based strategy could save €25,000.- in a year compared to the time-based strategy.

Overall, the sensitivity analysis evaluates the impact of variability in production amounts on the performance of a planning model. It decreases real production amounts across 2023 to investigate the effects on the model's outcomes. The results demonstrate that the counter-based planning model with a rolling horizon adapts better to decreased production, maintaining balanced over- and under-maintenance levels. It also shows reduced costs compared to a time-based maintenance strategy. This suggests the counter-based model's effectiveness under varying production conditions. On the contrary, it is expected that a general large increase of production from the status quo would increase the total maintenance costs due to the extra effort needed to keep over- and under-maintenance at a minimum. The counter-based maintenance strategy adapts to the big changes in production. In other words, if the company expects that the production will be lower in the upcoming years, the benefits of counter-based maintenance and specifically the use of this planning model are noticeable.

5.6 Conclusion

This chapter delivers the results of the solution design and shows how it affects the performance of the maintenance activities. From a step-by-step method, it provides a spreadsheet where the Maintenance Engineering Team could evaluate the counter-based progression of each PM-plan. Counter-based progressions are often times lower or higher than the time-based progressions, meaning that with the real production times of the machines, maintenance should not be planned as early or late as with the time-based strategy. This is an indication that the company could perform less over- and under-maintenance if they monitor the counter-based progression of the PM-plans on a frequent basis.

Furthermore, results of the prediction models show that when evaluating the average performance metrics over all machines, both the Random Forest and the Holt-Winters model show beneficial prediction performances to provide knowledge on when to plan any PM-plan with enough time (13 weeks) for preparations. Based on the performance metrics (RMSE, MAD, sMAPE), the Random Forest models outperform the Holt-Winters models. Linear Regression delivers the best predictions to improve the planning activities of multiple PM-plans in the upcoming period by providing the expected amount of production in between stop-days. In the Linear Regression model, it was found that the tactical planning feature was the most important features, having the most impact on the predicted values. Overall, there are some machines with a very variable amount of production during the years due to a dependence on product types that go through the packaging filling line, resulting in bad prediction performances, while other machines have better results. The company could either ignore these bad predictions and continue transitioning to counter-based maintenance or investigate additional input variables to enhance predictability for machines with variable production amounts.

Considering the planning model, the mathematical programming environment can find an optimal solution with the current parameter settings of packaging filling line 81. In view of many performance metrics, the use of a rolling horizon seems to benefit the final solution to the problem, which can be explained by the reduced cumulative error. The planning model solutions also create schedules that perform better in comparison to the real values gathered from historical data, but the complexities of reality must also be considered in this comparison. Additionally, parameters were adjusted to create a planning model solution schedule that implements a time-based strategy. The counter-based planning model, with or without the use of a rolling horizon, perform better than the time-based maintenance solutions. To further validate the solution design, the sensitivity analysis examines how production variability affects the planning model's performance by reducing production amounts in 2023 with variable amounts. Results show that the counter-based planning model with a rolling horizon adapts better to decreased production, maintains balanced maintenance levels, and reduces costs compared to a time-based strategy. This highlights the counter-based model's effectiveness, especially if production is expected to decrease in the coming years.

Chapter 6

Conclusion, Discussion, and Recommendations

This final chapter presents the conclusion, discussion, and recommendations of this research. Section 6.1 offers the conclusion, followed by the discussion in Section 6.2. Lastly, Section 6.3 outlines the recommendations for the company, based on the findings of the previous chapters.

6.1 Conclusion

At the present moment, Heineken experiences difficulties with the current maintenance strategy. This strategy being time-based, is a maintenance plan that uses fixed time intervals, e.g. once every two weeks. A counter-based maintenance plan is of another type used for planned maintenance based on asset counter registrations. A time-based maintenance plan for the packaging filling lines would be adequate if the interval of the maintenance activities contains a constant amount of production hours. However, for the packaging filling lines in, the production between two consecutive maintenance activities currently differs in the thousands of hours. A clear and concise statement of the core problem is the following:

"In the current situation, the maintenance planning activities at the packaging filling lines in the brewery at Zoeterwoude of Heineken are time-based, while the production activities differ throughout the year and change dynamically, which indicates that a counter-based maintenance strategy is more suitable."

Chapter 1 structures the research by introducing several research questions with correlating sub-research questions to provide a clear pathway through this report. This section discusses the conclusions of the research to deliver a final answer to the following main research question: *"How can the maintenance planning activities of the packaging filling lines in the brewery at Zoeterwoude of Heineken be adjusted to change from a time-based to a counter-based strategy*?"

The first step was to gather information and gain knowledge on the context of the problem. Chapter 2 thoroughly explores the many concepts that are involved when one discusses the maintenance activities of the packaging filling lines in the brewery. The most interesting finding correlates to the quantitative evidence that emphasizes the drawback of the current time-based strategy. Additionally, an analysis of two multipacker machines justifies that a counter-based strategy is superior to a time-based strategy, due to a decrease of activities and a more constant amount of production between them.

Chapter 3 explores the literature regarding time- and counter-based maintenance, interval determination, planning modelling in maintenance, production lines, and prediction modelling. A significant gap is the transition from time-based to counter-based maintenance strategies. There are no methods for using knowledge and data from time-based systems to adjust to usage counters, nor are there guidelines for determining adequate maintenance intervals in the absence of failure data. While prediction models are extensively theorized, their application in counter-based maintenance remains unexplored, particularly in predicting future usage amounts. Additionally, current literature lacks specific planning models that account for maintenance activities based on usage counters. The research fills these gaps in the literature.

Regarding the gaps in the literature, Chapter 4 continues with the design of the solutions to the problems of the company. The first solution involves determining usage counters with two newly introduced equations. Furthermore, data from multiple software systems and databases provide a prediction of production for scheduling maintenance. Two prediction models are suggested: one for aggregated production data over 13 weeks and another for forecasting weekly production for the upcoming 52 weeks. The goal of the first is to provide knowledge on when to plan any PM-plan with enough time (13 weeks) for preparations. The second improves the planning activities of multiple PM-plans in the upcoming period by providing the expected amount of production in between stop-days, which are moments to perform preventive maintenance. Additionally, a mathematical model aims to minimize over- and under-maintenance by optimally scheduling PM-plans, inspired by machine scheduling problems and incorporating constraints on capacity, counter-based intervals, and clustering. This model follows a rolling horizon approach, improving scheduling accuracy in dynamic environments.

Chapter 5 presents the results of the solution design and its impact on maintenance activities. It introduces a method for the Maintenance Engineering Team to evaluate counter-based progressions of PM-plans, which suggest less or more frequent maintenance than the current time-based strategies, potentially reducing over- and under-maintenance. Prediction models show that a Random Forest and Linear Regression model are the best methods to provide knowledge for scheduling PM-plans. The company could either choose to use the current models, or investigate additional input variables to enhance predictability for machines with very variable production amounts. The planning model, using a mathematical programming environment, optimizes scheduling to minimize under- and over-maintenance, performing better than the current time-based strategy despite variability. A sensitivity analysis shows that the counter-based planning model with a rolling horizon adapts better to decreased production, maintains balanced maintenance levels, and reduces costs compared to a time-based strategy. This highlights the counter-based model's effectiveness, especially if production is expected to decrease in the coming years.

Overall, to adjust the maintenance planning activities of the packaging filling lines in the brewery at Zoeterwoude of Heineken from a time-based plan to a counter-based plan, begin by determining the usage counters, based on previously determined time-based intervals and production amounts. Then, use a Random Forest model to forecast production for planning any PM-plan 13 week beforehand for preparations. Additionally, optimize scheduling with a mathematical model that minimizes over- and under-maintenance by scheduling preventive maintenance plans optimally, incorporating constraints on capacity, clustering, and counter-based intervals. If the brewery aims to decrease the production in the upcoming years, there is no doubt that counter-based maintenance will be beneficial for the efficiency and effectiveness of maintenance, in addition to a large reduction of costs.

6.2 Discussion

This section delves into the limitations of the research, suggesting potential improvements, and proposing directions for future research. The discussion helps to interpret the impact of the research, its contribution to the field, and its practical applications. Section 6.2.1 discusses the limitations and potential improvements of this research. Then, Section 6.2.2 and Section 6.2.3 provide the contribution to theory and practice, respectively. Section 6.2.4 finalizes the discussion with suggestions for further research

6.2.1 Limitations and Potential Improvements

First, there are some limitations due to some practical difficulties. For example, the last execution dates are not being registered in the main software management system. The counter-based calculations depend on the date of the last execution of an activity, so it is critical that these dates are logged and extracted in a correct and valid manner. Currently, the team cannot access the data that contains these dates and for some PM-plans data is missing. It is urgent that Heineken and IBM Maximo (the company that oversees the main software management system) sit together to solve this restriction.

Moreover, the calculation of the counters needs a connection between the data from multiple software systems. All PM-plans have a certain functional location code for its asset that indicates its presence in a machine on the filling line. These codes should make a connection to the machine production data. Currently, this research has made some assumptions on the connection of certain assets to machines that simplify the calculations. This demands a thorough validation by a Maintenance Engineer to further improve the effectiveness of counter-based maintenance.

On top of that, there are possibilities considering the automation of the calculations and integration of the software systems. Due to start up difficulties with the new management software system, the program still lacks the possibility to make these simple calculations accurately and automatically. In many discussions with multiple stakeholders the conclusion remains that a Maintenance Engineer would be responsible for uploading updated counter values to the software system to update the calculations. This is definitely not ideal. The Data and Analytics Team is already instructed (on the date of publishing this thesis) to investigate how different systems could automatically update the calculations.

Finally, specifically about the planning model, there is a big assumption on the restrictions of scheduling preventive maintenance. The assumption is the following: Per moment for preventive maintenance, there is a maximum number of activities allowed, while in reality, this depends on the duration of activities, skill-set requisites, workforce and spare-part availability, and the schedules of activities that were excluded (PM-plans that keep applying a time-based maintenance strategy). This assumption makes the implementation of the current planning model difficult. To further improve the research, the planning model requires more extensive capacity constraints, creating a more holistic approach to better the planning and scheduling activities of maintenance. Luckily, the manager of the Maintenance Engineering Team has hired a new intern that will continue the research on a solution to this problem.

6.2.2 Contribution to Theory

Regarding the contribution to the scientific body of knowledge, this thesis is an case study about the transition from a calendar time-based tot a counter-based strategy. Current literature contains a lot of knowledge on the differences between these strategies, and the benefits and disadvantages of them in specific situations, but it clearly lacks any method of transitioning from one method to the other. The research delivers a clear method of incorporating useful data and knowledge from a time-base maintenance system, to arrive at a counter-based system that integrates useful sources from different stakeholders. No other source discusses using time-based settings to arrive at a counter-based preventive maintenance plan.

A specific example to explain the relevance of the research considers the problem of predicting production hours to be able to schedule activities in advance for preparations. Literature discusses many benefits of counter-based (often called usage-based) maintenance, but it fails to specify the difficulties that arise with transitioning to such a strategy. One advantage of a time-based system is scheduling activities with little effort, because the time between activities is constant, but in a counter-based system this instantly becomes problematic. How could a large organization know their amount of production in the future with great accuracy? This research delivers the method of finding an adequate prediction method, by assessing multiple machine learning and forecasting methods, subsequently incorporating this in a planning model to fully embody the benefits of counter-based maintenance.

Additionally, theory lacks a specific mathematical model that schedules activities based on usage counters to minimize over- and under-maintenance. There are many existing scheduling models, but none consider the context of counter-based maintenance. Difficulties arise in such models due to the repetitive activities, all depending on unknown amounts of production. This research delivers a clear modelling method of incorporating predictions as a parameter, and using a rolling horizon to account for the error of these predictions. The mathematical model might seem trivial, but there are several complex aspects. It is not trivial due to the combinatorial nature of the binary decision variables and the elaborate, interdependent constraints such as clustering and capacity. It also provides a set up for a more complex application with more specific capacity constraints such as available workforce, spare parts, or task duration.

6.2.3 Contribution to Practice

This research provides several practical contributions to Heineken and potentially to other companies with similar maintenance challenges. Obviously, the main contribution relates to the step-by-step approach, applicable for maintenance teams to systematically shift their maintenance planning, thereby improving the efficiency and effectiveness of their operations. By aligning maintenance activities more closely with actual usage, organizations could excel in resource allocation, reducing downtime, and streamlining maintenance processes. The development of a mathematical model that minimizes over- and under-maintenance presents a practical tool for maintenance planners. This model, which incorporates constraints on capacity, clustering, and counter-based intervals, can be used to create optimal maintenance schedules. The rolling horizon approach ensures that the model remains adaptive to changes in production, further enhancing its utility in dynamic environments.

A specific example to explain the relevance in practice considers the cost reductions. An organization that applies a time-based maintenance strategy and plans to reduce the amount of production in future years can reduce a lot of costs by applying a counter-based planning model. The sensitivity analysis indicates that the counter-based model is particularly effective in reducing costs when production levels decrease. This cost efficiency is crucial for maintaining competitiveness in the brewing industry. On top of that, the practical contributions extend beyond Heineken to any organization looking to transition to a counter-based maintenance strategy. The research provides a clear framework and set of guidelines that other companies can adapt to their specific contexts. This includes the steps to gather necessary data, the development of prediction models, and the implementation of an optimized scheduling model.

6.2.4 Suggestions for Further Research

While this research provides a solid foundation for transitioning from a time-based to a counter-based maintenance strategy, several areas warrant further investigation to enhance the effectiveness and applicability of the findings. First, while this research identified Random Forest and Linear Regression as effective models for predicting production, future studies could explore additional machine learning techniques, such as neural networks or ensemble methods, to improve prediction accuracy. Additionally, incorporating more diverse and extensive datasets, including extra features such as product types, could refine these models.

Secondly, there is a lot of potential in a planning model that is more holistic, including preventive maintenance activities that are time-based, counter-based, and possibly even condition-based. The current planning model includes basic capacity constraints. Future research could develop more sophisticated models that account for varying duration's of maintenance activities, workforce skill sets, spare part availability, and the interaction with other maintenance activities. These models could use mixed-integer programming or other advanced optimization techniques to better represent the complex nature of maintenance scheduling.

Finally, conducting case studies in different industries and settings would extend the applicability of the proposed counter-based maintenance strategy. Industries such as automotive, aerospace, or manufacturing, which have high variability in production, could provide valuable insights into the adaptability of the model. These studies could also assess the long-term impacts of transitioning to a counter-based maintenance strategy on, for example, equipment lifespan, operational efficiency, and cost savings. This would provide a deeper understanding of its benefits and potential drawbacks.

6.3 Recommendations

Based on the findings and conclusions of this research, this section delivers several recommendations for Heineken to effectively transition from a time-based to a counter-based maintenance strategy for the packaging filling lines at the brewery in

Zoeterwoude. These recommendations aim to improve maintenance effectiveness and efficiency, reduce over- and under-maintenance, and prevent unnecessary costs in the future.

- **Transition to a counter-based strategy for a selection of the preventive maintenance plans.**

This research does not recommend to transform all preventive maintenance activities to a counter-based system. First of all, the company should not alter the maintenance strategy of preventive maintenance activities with a current time-based interval of one or two weeks. The current time-based maintenance intervals for the preventive maintenance tasks vary from one week to eight years. These tasks are always planned on a stop-day, which occur once every other week, meaning the shortest amount of time between maintenance activities possible is one week for short tasks and two weeks for longer tasks. Therefore it is not necessary to investigate if a PM-plan should be planned on a shorter notice than one or two weeks. Secondly, the company should not alter the maintenance strategy of preventive maintenance activities that consider condition monitoring, cleaning, or safety and compliance (Codes PM6 to PM8). Condition Monitoring means that the proactive decision if a task should be performed, depends on the state of what the sensors measure, not on the counter of the machine. The necessity of Cleaning tasks does not depend on the amount of production time, but depends on the total time that has passed. Safety and Compliance tasks are necessary whenever the regulations state that they are necessary, and this should not depend on the amount of production of a line. Overall, it would not be beneficial for the company to adjust the maintenance strategy for these tasks to a counter-based maintenance strategy.

- **Continuously monitor and evaluate the usage counters and the counter-based strategy.**

Heineken should continuously monitor the usage counters of the preventive maintenance plans. To begin, this research recommends applying the method of creating a spreadsheet that visualizes the usage counters (see Appendix B). Next to this, a 20 page methodology document has been delivered to the team that explains every step of the way for any Maintenance Engineer to apply counter-based maintenance on any of the packaging filling lines. In the upcoming period, the Maintenance Engineering Team should incorporate the usage counters into the main software management system. Eventually, the effort of monitoring will decrease as the counter-based strategy is fully implemented and the software system automatically updates the counters.

- **Utilize predictive models for maintenance planning and scheduling activities.**

Heineken should implement the Random Forest model for forecasting aggregated production amounts. This model should be used to plan preventive maintenance activities 13 weeks in advance, allowing sufficient time for preparations. This research has proven that this model type is the best choice for this specific purpose. Additionally, the company should continuously refine the model by incorporating more data and exploring other predictive techniques, or introduce new features to improve forecasting accuracy.

- **Further optimize scheduling with the planning model.**

To minimize over- and under-maintenance, Heineken should adopt the mathematical model developed in this research for scheduling preventive maintenance activities. The maintenance planner could generate the schedule with this planning model and integrate the results in his work. This model incorporates constraints on capacity, clustering, and counter-based intervals. It is strongly recommended to further investigate a more holistic mathematical model. Another intern will continue the research, incorporating the duration of activities, skill-set requisites, workforce and spare-part availability, and the schedules of time-based preventive maintenance activities. A holistic mathematical model could omit a lot of work for the maintenance planner, subsequently saving a lot of costs and generating time for this employee to focus his efforts on other problems.

- **Collaborate with the Data & Analytics Team.**

This research recommends to create a data model that integrates the different sources of data. Currently, the maintenance engineers lack the skill to combine data from the Manufacturing Execution System to link the production hours of machines to the preventive maintenance plans. Luckily the Data & Analytics Team is able to design and validate such a model. They could refine the predictive models and explore additional input variables that could enhance the accuracy. Regular collaboration between both teams will ensure that the models remain relevant and effective in predicting production amounts and scheduling maintenance activities.

- **Share best practices across the organization.**

Finally, Heineken should document and share the best practices and lessons learned from the implementation of the counter-based maintenance strategy at the Zoeterwoude brewery with other facilities within the organization. The brewery in Den Bosch already performs some preventive maintenance actions with usage counters, but their operations are not perfect. Additionally, during this research there was contact with breweries in Poland who also made a short start with counter-based systems, but they ran into similar difficulties that were present in this thesis. Zoeterwoude has the opportunity to continue the implementation and create efficient and effective counter-based operations. If they continuously share their practices globally, this will promote a culture of continuous improvement and operational excellence across the company.

Bibliography

Aghezzaf, E and N M Najid (2008). "Integrated production planning and preventive maintenance in deteriorating production systems". In: *Information Sciences* 178, pp. 3382–3392.

Ahmad, Rosmaini and Shahrul Kamaruddin (2012). "An overview of time-based and condition-based maintenance in industrial application". In: *Computers and Industrial Engineering* 63 (1), pp. 135–149.

Alardhi, M. and A. W. Labib (2008). "Preventive maintenance scheduling of multi-cogeneration plants using integer programming". In: vol. 59. Palgrave Macmillan Ltd., pp. 503–509. DOI: 10.1057/palgrave.jors.2602386.

Axsäter, Sven (2006). *INVENTORY CONTROL*. Vol. 2. DOI: 10.1007/0-387-33331-2.

Ayvaz, Serkan and Koray Alpay (July 2021). "Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time". In: *Expert Systems with Applications* 173. ISSN: 09574174. DOI: 10.1016/j.eswa.2021.114598.

Basker, B A and T M Husband (1978). *DETERMINATION OF OPTIMAL OVERHAUL INTERVALS AND INSPECTION FREQUENCIES - A CASE STUDY*. DOI: doi.org/10.1016/0026-2714(78)90731-X.

Ben-Daya, M and M A Rahim (2001). *INTEGRATED PRODUCTION, QUALITY AND MAINTENANCE MODELS: AN OVERVIEW*. DOI: 10.1007/978-1-4615-1635-4_1.

Bishop, C.M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Brownlee, Jason (Aug. 2019). *How to use Learning Curves to Diagnose Machine Learning Model Performance*.

Budai-Balke, Gabriella (2009). "Operations research models for scheduling railway infrastructure maintenance". In: pp. 1–190.

Cadi, Abdessamad Ait El et al. (2015). "Generalizing of the integration of noncyclical preventive maintenance scheduling and production planning for a series-parallel production line". In.

Christer, A H and T Doherty (1970). "Scheduling Overhauls of Soaking Pits". In: *Operational Research Quarterly* 28 (4), pp. 915–926. URL: https://www.jstor.org/stable/3008865.

Christer, A H and W M Waller (1984). "Reducing Production Downtime Using Delay-Time Analysis". In: *Source: The Journal of the Operational Research Society* 35 (6), pp. 499–512. URL: https://about.jstor.org/terms.

Cichelli, James A (1977). "An evaluation of calendar based preventive maintenance system versus usage based preventive maintenance system." In: URL: http://preserve.lehigh.edu/etd.

Dedopoulos, L.T. and N Shah (1995). "PREVENTIVE MAINTENANCE POLICY OPTIMIZATION FOR MULTIPURPOSE PLANT EQUIPMENT". In: *Computers and Chemical Engineering* 19, pp. 693–698.

Dekker, R (1996). "Applications of maintenance optimization models: a review and analysis". In: *Reliability Engineering and System Safety*, pp. 229–240. DOI: https://doi.org/10.1016/0951-8320(95)00076-3.

Deloux, Estelle, Mitra Fouladirad, and Christophe Bérenguer (2016). "Health-and-usage-based maintenance policies for a partially observable deteriorating system". In: *Proceedings of the Institution of Mechanical Engi-neers* 230 (1), pp. 120–129. DOI: 10.1177/1748006X15609863Ãŕ. URL: https://hal.science/hal-01397367.

Dijkhuizen, Gerhard Van (2000). "MAINTENANCE GROUPING IN MULTI-STEP MULTI-COMPONENT PRODUCTION SYSTEMS". In: Springer, pp. 283–306.

Dijkhuizen, Gerhard Van and Aart Van Harten (1997). *Optimal clustering of frequency-constrained maintenance jobs with shared set-ups*.

Donoho, David (Oct. 2017). *50 Years of Data Science*. DOI: 10.1080/10618600.2017.1384734.

Ebrahimipour, V., A. Najjarbashi, and M. Sheikhalishahi (Feb. 2015). "Multi-objective modeling for preventive maintenance scheduling in a multiple production line". In: *Journal of Intelligent Manufacturing* 26 (1), pp. 111–122. ISSN: 15728145. DOI: 10.1007/s10845-013-0766-6.

Efron, Bradley (2020). "Prediction, Estimation, and Attribution". In: *Statistical Association* 115 (530), pp. 636–655.

Fokkert, J. I. Van Zante-De et al. (2007). "The Netherlands schedules track maintenance to improve track workers' safety". In: *Interfaces* 37 (2), pp. 133–142. ISSN: 00922102. DOI: 10.1287/inte.1060.0246.

Frost, Daniel and Rina Dechter (1998). "Optimizing with constraints: a case study in scheduling maintenance of electric power units". In: *Lecture Notes in Computer Science* 1520, pp. 469–488.

Gholamy, Afshin, Vladik Kreinovich, and Olga Kosheleva (2018). "A Pedagogical Explanation A Pedagogical Explanation Part of the Computer Sciences Commons". In: *Departmental Technical Reports*. URL: https://scholarworks.utep.edu/cs_techrephttps://scholarworks.utep.edu/cs_techrep/1209.

Graham, R.L. et al. (1979). "Optimization and Approximation in Deterministic Sequencing and Scheduling: a Survey". In: *Annals of discrete mathematics* 5, pp. 287–326.

Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media.

Haghani, Ali and Yousef Shafahi (2002). "Bus maintenance systems and maintenance scheduling: model formulations and solutions". In: *Transportation Research Part A* 36, pp. 453–482. URL: www.elsevier.com/locate/tra.

Hand, David J. (Feb. 2006). *Classifier technology and the illusion of progress*. DOI: 10.1214/088342306000000060.

Heerkens, Hans and Arnold Van Winden (2017). *Solving Managerial Problems Systematically*. 1st ed. Noordhoff Uitgevers bv, pp. 1–135.

Henriksson, Lovisa (2019). *Optimal Maintenance and Operation Scheduling Using Mixed Integer Linear Programming - A Method to Automatize Maintenance Planning*.

Holland, Charles W. and Robert A. McLean (1975). "Applications of replacement theory". In: *AIIE Transactions* 7 (1), pp. 42–47. ISSN: 05695554. DOI: 10.1080/05695557508974983.

Hu, Jueming et al. (Nov. 2023). "Knowledge-enhanced reinforcement learning for multi-machine integrated production and maintenance scheduling". In: *Computers and Industrial Engineering* 185. ISSN: 03608352. DOI: 10.1016/j.cie.2023.109631.

James, Gareth (Gareth Michael) et al. (2013). *An introduction to statistical learning : with applications in R*, pp. 1–426. ISBN: 9781461471370.

Jonge, Bram de et al. (2015). "Optimum maintenance strategy under uncertainty in the lifetime distribution". In: *Reliability Engineering and System Safety* 133, pp. 59–67.

Kang, Ziqiu, Cagatay Catal, and Bedir Tekinerdogan (Feb. 2021). "Remaining useful life (RUL) prediction of equipment in production lines using artificial neural networks". In: *Sensors (Switzerland)* 21 (3), pp. 1–20. ISSN: 14248220. DOI: 10.3390/s21030932.

Kim, Jeongyun, Yongjun Ahn, and Hwasoo Yeo (Dec. 2016). "A comparative study of time-based maintenance and condition-based maintenance for optimal choice of maintenance policy". In: *Structure and Infrastructure Engineering* 12 (12), pp. 1525–1536. ISSN: 17448980. DOI: 10.1080/15732479.2016.1149871.

Koomson, Solomon Evans Kweku, Victor Amoako Temeng, and Yao Yevenyo Ziggah (Apr. 2024). "A Novel Approach in Predicting Dump Truck Tyre Life in a Mine Based on Multilayer Perceptron Neural Network Optimised with Particle Swarm Optimisation". In: *Mining, Metallurgy and Exploration* 41 (2), pp. 769–786. ISSN: 25243470. DOI: 10.1007/s42461-024-00954-y.

Koulinas, Georgios K., Panagiotis D. Paraschos, and Dimitrios E. Koulouriotis (2024). "A machine learning framework for explainable knowledge mining and production, maintenance, and quality control optimization in flexible circular manufacturing systems". In: *Flexible Services and Manufacturing Journal*. ISSN: 19366590. DOI: 10.1007/s10696-024-09537-x.

Li, Guofa et al. (Sept. 2018). "Development of a preventive maintenance strategy for an automatic production line based on group maintenance method". In: *Applied Sciences (Switzerland)* 8 (10). ISSN: 20763417. DOI: 10.3390/app8101781.

Li, Ning et al. (July 2024). "Combination prediction of underground mine rock drilling time based on seasonal and trend decomposition using Loess". In: *Engineering Applications of Artificial Intelligence* 133. ISSN: 09521976. DOI: 10.1016/j.engappai.2024.108064.

Liu, Peng, Guanjun Wang, and Zhong-Heng Tan (2024). "Calendar-time-based and age-based maintenance policies with different repair assumptions". In: *Applied Mathematical Modelling* 129, pp. 592–611.

N.V., Heineken (2023). *Heineken N.V. Annual Report 2023*. URL: www.theheinekencompany.com/our-company/our-2023-annual-report.

Ott, Moheeb, Mohammad Almuhaini, and Muhammad Khalid (2019). "A MILP-Based Restoration Technique for Multi-Microgrid Distribution Systems". In: *IEEE Access* 7, pp. 136801–136811. ISSN: 21693536. DOI: 10.1109/ACCESS.2019.2942633.

Pan, Shih Jie et al. (Feb. 2023). "Investigation of Machine Learning Methods for Predictive Maintenance of the Ultra-High-Pressure Reactor in a Polyethylene-Vinyl Acetate Production Process". In: *Electronics (Switzerland)* 12 (3). ISSN: 20799292. DOI: 10.3390/electronics12030580.

Percy, David (2008). "Preventive Maintenance Models for Complex Systems". In: Springer, pp. 179–208. DOI: 10.1007/978-1-84800-011-7.

Purnomo, Muhammad Ridwan Andi, Dzuraidah Abdul Wahab, and Salvinder Singh (2023). "Optimisation of Capacitated Planned Preventive Maintenance in Multiple Production Lines Using Optimisation-in-the-Loop Simulation". In: *International Journal of Integrated Engineering* 15 (5), pp. 263–272. ISSN: 2229838X. DOI: 10.30880/ijie.2023.15.05.028.

Qing, L et al. (2010). *Scheduling Problems with Preventive Maintenance on Two Parallel Production Lines*.

Rausand, Marvin, Anne Barros, and Arnljot Høyland (2021). *System reliability theory: models, statistical methods, and applications*. Vol. 3, pp. 1–803. ISBN: 9781119373520.

Schnaubelt, Matthias et al. (2019). "A comparison of machine learning model validation schemes for non-stationary time series data". In: *Institute for Economics*. DOI: 10.13140/RG.2.2.29545.24168. URL: https://www.researchgate.net/publication/340092030.

Sharma, Garima and Rajiv Nandan Rai (Jan. 2021). "Age Based Overhaul Policy for Multiple Repairable Systems with Imperfect Maintenance: Case Study of Aero Engines". In: *International Journal of Mathematical, Engineering and Management Sciences* 6 (1), pp. 193–206. ISSN: 24557749. DOI: 10.33889/IJMEMS.2021.6.1.012.

Shmueli, Galit (Aug. 2010). "To explain or to predict?" In: *Statistical Science* 25 (3), pp. 289–310. ISSN: 08834237. DOI: 10.1214/10-STS330.

Siswanto, A. and N. Kurniati (Apr. 2018). "Determining optimal preventive maintenance interval for component of Well Barrier Element in an Oil and Gas Company". In: vol. 337. Institute of Physics Publishing. DOI: 10.1088/1757-899X/337/1/012066.

Skachkova, Elena, Artem Alenin, and Vladimir Mokshin (2022). "Study of Machine Learning Methods in the Analysis of Production and Maintenance of Oil Wells". In: Institute of Electrical and Electronics Engineers Inc. ISBN: 9781665485715. DOI: 10.1109/ITNT55410.2022.9848554.

Steyerberg, Ewout W et al. (2001). *Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis*.

Tan, Jonathan and Mark Kramer (1997). "A general framework for preventive maintenance optimization in chemical process operations". In: *Computers chemical Engineering* 21 (12), pp. 1451–1469.

Tiddens, Wieger, Jan Braaksma, and Tiedo Tinga (Feb. 2023). "Decision Framework for Predictive Maintenance Method Selection". In: *Applied Sciences (Switzerland)* 13 (3). ISSN: 20763417. DOI: 10.3390/app13032021.

Tinga, T. (2013). *Principles of Loads and Failure Mechanisms*. Springer, pp. 1–308. DOI: 10.1007/978-1-4471-4917-0.

Tinga, Tiedo (Oct. 2010). "Application of physical failure models to enable usage and load based maintenance". In: *Reliability Engineering and System Safety* 95 (10), pp. 1061–1075. ISSN: 09518320. DOI: 10.1016/j.ress.2010.04.015.

Tucci, Mauro, Antonio Piazzi, and Dimitri Thomopulos (May 2024). "Machine Learning Models for Regional Photovoltaic Power Generation Forecasting with Limited Plant-Specific Data". In: *Energies* 17 (10). ISSN: 19961073. DOI: 10.3390/en17102346.

Vatn, J., P. Hokstad, and L. Bodsberg (1996). "An Overall Model for Maintenance Optimization". In: *Reliability Engineering and System Safety* 51, pp. 241–257.

Vaurio, J (1999). "Availability and Cost Functions for Periodically Inspected Preventively Maintained Units". In: *Reliability Engineering and System Safety* 63, pp. 133–140.

Wang, Hongzhou (June 2002). "A survey of maintenance policies of deteriorating systems". In: *European Journal of Operational Research* 139 (3), pp. 469–489. DOI: https://doi.org/10.1016/S0377-2217(01)00197-7.

Wang, Q, D Li, and Z Zhang (2016). "A New Model of Preventive Maintenance Scheduling for Parallel Production Line". In: *Research Institute of CETC* 28.

Xu, L et al. (2017). *Preventive Maintenance Planning in a Unreliable Production Line with a Branch Buffer*.

Appendix A

Box Plots

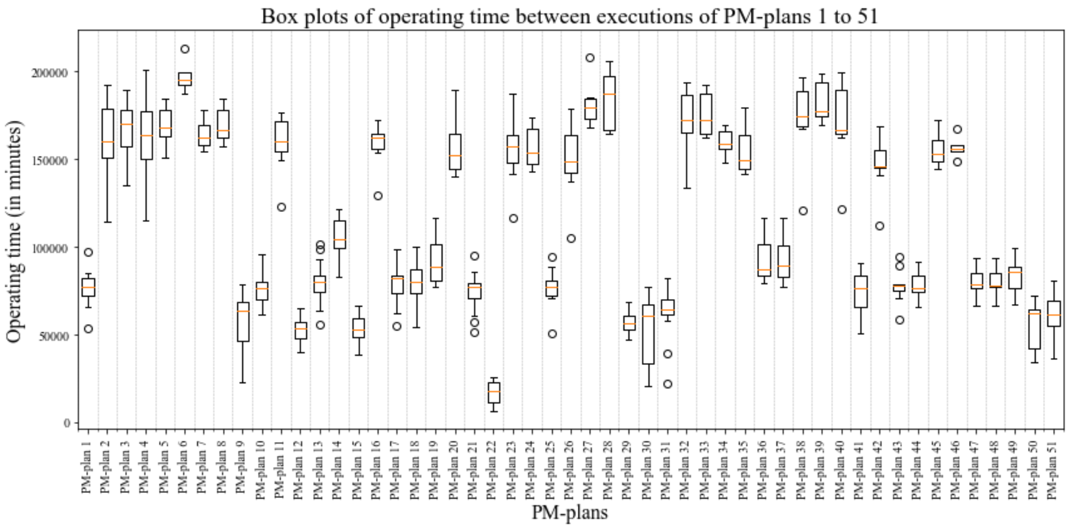


FIGURE A.1: Box plots showing differences in Production time between executions of PM-plans

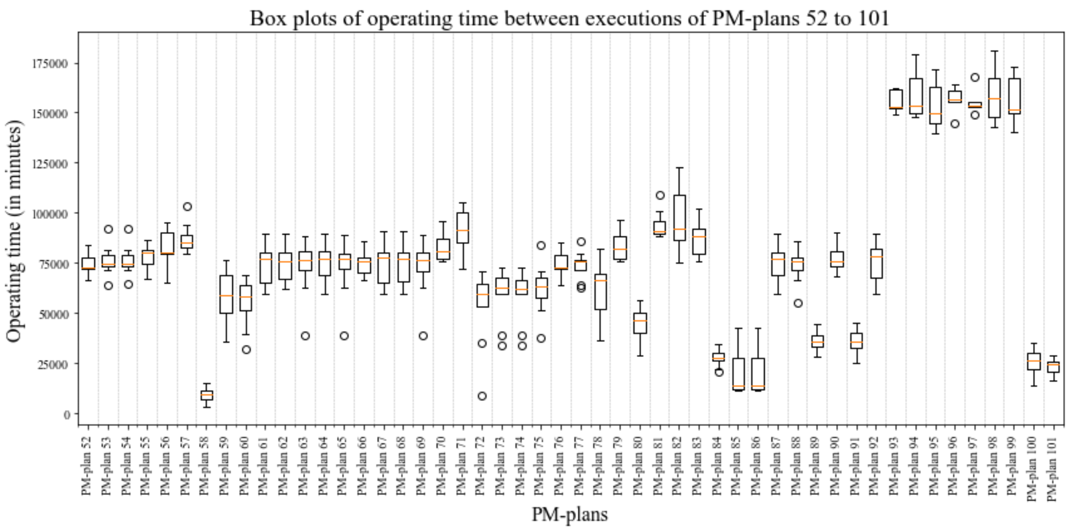


FIGURE A.2: Box plots showing differences in Production time between executions of PM-plans

Appendix B

Methodology for Counter-Based Maintenance

B.1 Step-by-step method

1. Search all PM plans in Maximo
   (a) Go to the Maximo Preventive Maintenance Plans list view.
   (b) Select all records, and use a where clause like the following:

   ((status = 'ACTIVE' and siteid = 'NL01')) and (((location like '%NL01-2-2081%')) OR (exists (select 1 from maximo.asset where ((location like '%NL01-2-2081%')) and (assetnum = pm.assetnum and siteid=pm.siteid) and (plustisconsist=0))))

   (c) Download the list to a excel file.
   (d) Copy the data into the counters template excel file (values only).
2. Clean up the data
   (a) Select all Description cells and use Control H to replace all commas with a space.
   (b) Check weird intervals (e.g. 6 jr instead of 6J).
   (c) Delete the rows of the PM plans that:
      i. Are PM6 to PM8
      ii. Have interval lower or equal to 2 weeks
      iii. Are OFF-LINE
      iv. Are already based on counters (DRU)
      v. Are not a PM plan (e.g. JWO or Revision)
3. Gather the data on last executions of the PM-plans (either from SAP or Maximo) Note: Use SAP or Maximo. For Maximo it is not possible yet for the Maintenance Engineering Team to extract this information. Discuss this with Maximo employee how to gather the information.
   (a) Make sure to extract the description of executed maintenance plans and their scheduled dates from 2016 to present. See the columns in the template to download the correct format.

(b) Copy the data into the "SAP" or "Maximo" sheet in the counters excel file.

(c) Replace all commas in the descriptions with spaces.

(d) Check for irregularities in column T.

4. Gather the data on production amounts from MES

(a) Open the downtime data box from MES in excel.

(b) Create the pivot table with in the following form:

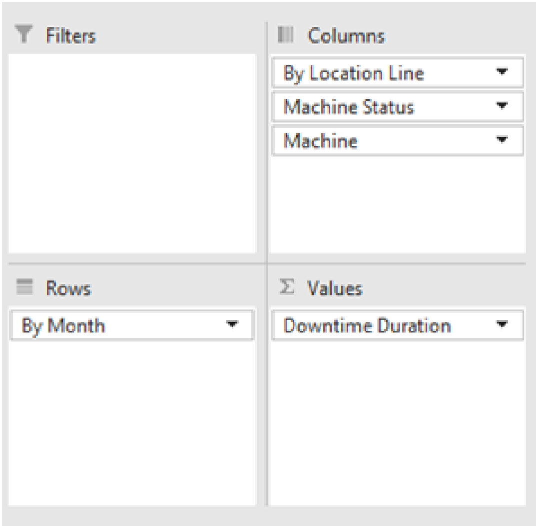


FIGURE B.1: Example of pivot table formation for MES data

(c) Expand all rows to see all dates.

(d) Show in tabular form.

(e) Set filter to the necessary line.

(f) Extract the data and sum the volloop, leegloop, productie as this is the Real production time.

(g) Use the Control H function to change the dates to real dates (e.g. replace Jan with -1-).

(h) Create an extra column with the average of the Etimas.

5. Insert the correct data connection in the sheet "DATA connection"

(a) Fill all unique SAP location values from Column W in the "Counters" sheet in the location column in the "DATA connection" sheet.

(b) Fill in the corresponding machine from the MES output in the machine column in the "DATA connection" sheet.

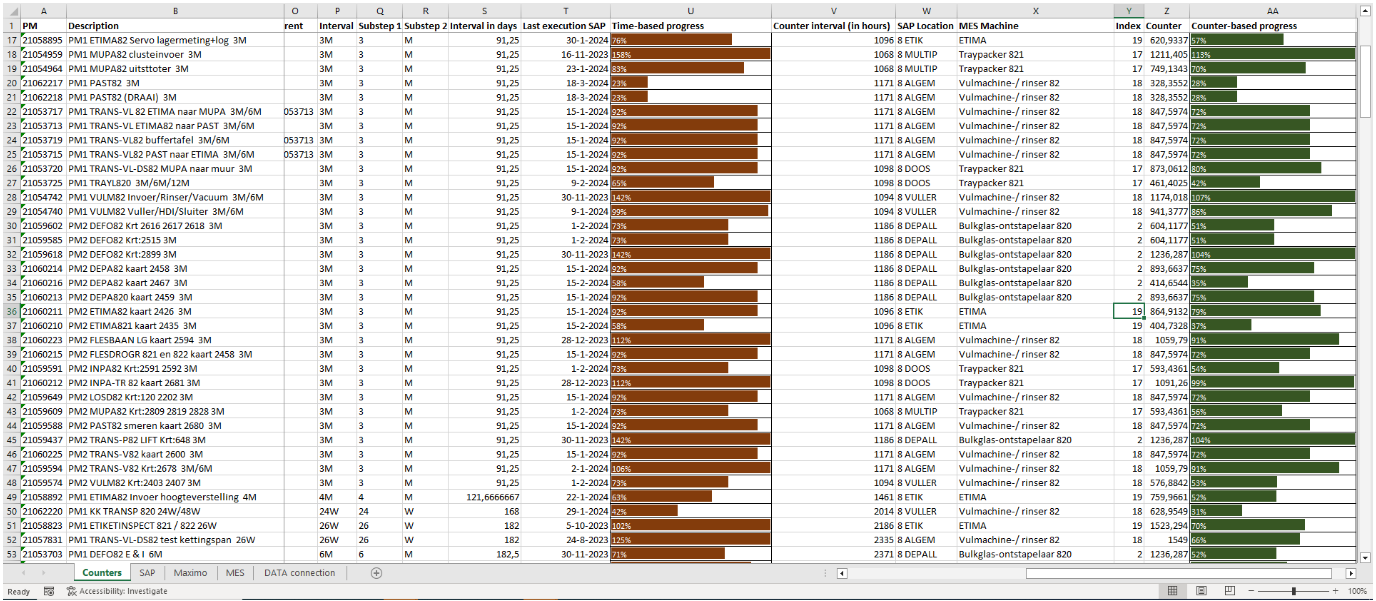
(c) Fill in the corresponding index for only the Real production time values from the "MES" sheet.

(d) Calculate the average production hours per year for the corresponding machines from the MES data and fill it in the Counter column in the "DATA connection" sheet.

6. Adjust the formula of counter calculations
   (a) In the "Counters" sheet change the formula of column z such that the correct section of the "MES" sheet is used. Set the rows correct to the MES data.
   (b) Check Data connection formulas if still correct.

B.2 Example spreadsheet

When a maintenance engineer completes the step-by-step method, the spreadsheet looks like the following:



	A	B	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA
1	PM	Description	rent	Interval	Substep 1	Substep 2	Interval in days	Last execution SAP	Time-based progress	Counter interval (in hours)	SAP Location	MES Machine	Index	Counter	Counter-based progress
17	21058895	PM1 ETIMA82 Servo lagermeting+log 3M		3M	3	M	91,25	30-1-2024	76%	1096	8 ETIK	ETIMA	19	620,9337	57%
18	21054959	PM1 MUPA82 clusteinvoer 3M		3M	3	M	91,25	16-11-2023	158%	1068	8 MULTIP	Traypacker 821	17	1211,405	113%
19	21054964	PM1 MUPA82 uitsttoter 3M		3M	3	M	91,25	23-1-2024	83%	1068	8 MULTIP	Traypacker 821	17	749,1343	70%
20	21062217	PM1 PAST82 3M		3M	3	M	91,25	18-3-2024	23%	1171	8 ALGEM	Vulmachine-/ rinser 82	18	328,3552	28%
21	21062218	PM1 PAST82 (DRAAI) 3M		3M	3	M	91,25	18-3-2024	23%	1171	8 ALGEM	Vulmachine-/ rinser 82	18	328,3552	28%
22	21053717	PM1 TRANS-VL 82 ETIMA naar MUPA 3M/6M	053713	3M	3	M	91,25	15-1-2024	92%	1171	8 ALGEM	Vulmachine-/ rinser 82	18	847,5974	72%
23	21053713	PM1 TRANS-VL ETIMA82 naar PAST 3M/6M		3M	3	M	91,25	15-1-2024	92%	1171	8 ALGEM	Vulmachine-/ rinser 82	18	847,5974	72%
24	21053719	PM1 TRANS-VL82 buffertafel 3M/6M	053713	3M	3	M	91,25	15-1-2024	92%	1171	8 ALGEM	Vulmachine-/ rinser 82	18	847,5974	72%
25	21053715	PM1 TRANS-VL82 PAST naar ETIMA 3M/6M	053713	3M	3	M	91,25	15-1-2024	92%	1171	8 ALGEM	Vulmachine-/ rinser 82	18	847,5974	72%
26	21053720	PM1 TRANS-VL-DS82 MUPA naar muur 3M		3M	3	M	91,25	15-1-2024	92%	1098	8 DOOS	Traypacker 821	17	873,0612	80%
27	21053725	PM1 TRAYL820 3M/6M/12M		3M	3	M	91,25	9-2-2024	65%	1098	8 DOOS	Traypacker 821	17	461,4025	42%
28	21054742	PM1 VULM82 Invoer/Rinser/Vacuum 3M/6M		3M	3	M	91,25	30-11-2023	142%	1094	8 VULLER	Vulmachine-/ rinser 82	18	1174,018	107%
29	21054740	PM1 VULM82 Vuller/HDI/Sluiter 3M/6M		3M	3	M	91,25	9-1-2024	99%	1094	8 VULLER	Vulmachine-/ rinser 82	18	941,3777	86%
30	21059602	PM2 DEFO82 Krt 2616 2617 2618 3M		3M	3	M	91,25	1-2-2024	73%	1186	8 DEPALL	Bulkglas-ontstapelaar 820	2	604,1177	51%
31	21059585	PM2 DEFO82 Krt:2515 3M		3M	3	M	91,25	1-2-2024	73%	1186	8 DEPALL	Bulkglas-ontstapelaar 820	2	604,1177	51%
32	21059618	PM2 DEFO82 Krt:2899 3M		3M	3	M	91,25	30-11-2023	142%	1186	8 DEPALL	Bulkglas-ontstapelaar 820	2	1236,287	104%
33	21060214	PM2 DEPA82 kaart 2458 3M		3M	3	M	91,25	15-1-2024	92%	1186	8 DEPALL	Bulkglas-ontstapelaar 820	2	893,6637	75%
34	21060216	PM2 DEPA82 kaart 2467 3M		3M	3	M	91,25	15-2-2024	58%	1186	8 DEPALL	Bulkglas-ontstapelaar 820	2	414,6544	35%
35	21060213	PM2 DEPA820 kaart 2459 3M		3M	3	M	91,25	15-1-2024	92%	1186	8 DEPALL	Bulkglas-ontstapelaar 820	2	893,6637	75%
36	21060211	PM2 ETIMA82 kaart 2426 3M		3M	3	M	91,25	15-1-2024	92%	1096	8 ETIK	ETIMA	19	864,9132	79%
37	21060210	PM2 ETIMA821 kaart 2435 3M		3M	3	M	91,25	15-2-2024	58%	1096	8 ETIK	ETIMA	19	404,7328	37%
38	21060223	PM2 FLESBAAN LG kaart 2594 3M		3M	3	M	91,25	28-12-2023	112%	1171	8 ALGEM	Vulmachine-/ rinser 82	18	1059,79	91%
39	21060215	PM2 FLESDROGR 821 en 822 kaart 2458 3M		3M	3	M	91,25	15-1-2024	92%	1171	8 ALGEM	Vulmachine-/ rinser 82	18	847,5974	72%
40	21059591	PM2 INPA82 Krt:2591 2592 3M		3M	3	M	91,25	1-2-2024	73%	1098	8 DOOS	Traypacker 821	17	593,4361	54%
41	21060212	PM2 INPA-TR 82 kaart 2681 3M		3M	3	M	91,25	28-12-2023	112%	1098	8 DOOS	Traypacker 821	17	1091,26	99%
42	21059649	PM2 LOSD82 Krt:120 2202 3M		3M	3	M	91,25	15-1-2024	92%	1171	8 ALGEM	Vulmachine-/ rinser 82	18	847,5974	72%
43	21059609	PM2 MUPA82 Krt:2809 2819 2828 3M		3M	3	M	91,25	1-2-2024	73%	1068	8 MULTIP	Traypacker 821	17	593,4361	56%
44	21059588	PM2 PAST82 smeren kaart 2680 3M		3M	3	M	91,25	15-1-2024	92%	1171	8 ALGEM	Vulmachine-/ rinser 82	18	847,5974	72%
45	21059437	PM2 TRANS-P82 LIFT Krt:648 3M		3M	3	M	91,25	30-11-2023	142%	1186	8 DEPALL	Bulkglas-ontstapelaar 820	2	1236,287	104%
46	21060225	PM2 TRANS-V82 kaart 2600 3M		3M	3	M	91,25	15-1-2024	92%	1171	8 ALGEM	Vulmachine-/ rinser 82	18	847,5974	72%
47	21059594	PM2 TRANS-V82 Krt:2678 3M/6M		3M	3	M	91,25	2-1-2024	106%	1171	8 ALGEM	Vulmachine-/ rinser 82	18	1059,79	91%
48	21059574	PM2 VULM82 Krt:2403 2407 3M		3M	3	M	91,25	1-2-2024	73%	1094	8 VULLER	Vulmachine-/ rinser 82	18	576,8842	53%
49	21058892	PM1 ETIMA82 Invoer hoogteverstelling 4M		4M	4	M	121,6666667	22-1-2024	63%	1461	8 ETIK	ETIMA	19	759,9661	52%
50	21062220	PM1 KK TRANSP 820 24W/48W		24W	24	W	168	29-1-2024	42%	2014	8 VULLER	Vulmachine-/ rinser 82	18	628,9549	31%
51	21058823	PM1 ETIKETINSPECT 821 / 822 26W		26W	26	W	182	5-10-2023	102%	2186	8 ETIK	ETIMA	19	1523,294	70%
52	21057831	PM1 TRANS-VL-DS82 test kettingspan 26W		26W	26	W	182	24-8-2023	125%	2335	8 ALGEM	Vulmachine-/ rinser 82	18	1549	66%
53	21053703	PM1 DEFO82 E & I 6M		6M	6	M	182,5	30-11-2023	71%	2371	8 DEPALL	Bulkglas-ontstapelaar 820	2	1236,287	52%

FIGURE B.2: Example of the spreadsheet

Appendix C

Distributions for Parameter Tuning

LISTING C.1: Parameter distributions per model

```
# Parameter distributions for the Decision Tree model
param_dist_DT = {
    'max_depth': [10, 20, 30],
    'min_samples_split': [2, 5, 10],
    'min_samples_leaf': [1, 2, 4],
    'max_features': [None, 'sqrt', 'log2']
}

# Parameter distributions for the Random Forest model
param_dist_RF = {
    'n_estimators': [100, 200, 300, 400, 500],
    'max_depth': [10, 20, 30],
    'min_samples_split': [2, 5, 10],
    'min_samples_leaf': [1, 2, 4],
    'max_features': [None, 'sqrt', 'log2']
}

# Parameter distributions for the Gradient Boosting Machine model
param_dist_GBM = {
    'n_estimators': [100, 200, 300, 400, 500],
    'learning_rate': [0.01, 0.05, 0.1, 0.2],
    'max_depth': [3, 4, 5, 6],
    'min_samples_split': [2, 5, 10],
    'min_samples_leaf': [1, 2, 4],
    'subsample': [0.7, 0.8, 0.9, 1.0],
    'max_features': [None, 'sqrt', 'log2']
}

# Parameter distributions for the K-Nearest Neighbor model
param_dist_KNN = {
    'n_neighbors': [3, 5, 7, 9, 11],
    'weights': ['uniform', 'distance'],
    'algorithm': ['auto', 'ball_tree', 'kd_tree', 'brute'],
    'leaf_size': [10, 20, 30, 40, 50],
    'p': [1, 2]
}
```

Appendix D

Performance Metrics per Machine

TABLE D.1: MAD Values per Machine for Aggregated Predictions

MAD	Baseline	LR	DT	RF	GBM	KNN	HW
Machine 1	14697.69	14016.37	10063.16	11256.67	12404.34	12447.74	6731.69
Machine 2	9141.32	8161.49	9383.63	8244.64	9859.64	8597.42	7541.24
Machine 3	15900.54	13512.06	13973.13	10742.18	12497.94	14478.69	11161.58
Machine 4	14984.70	12443.75	11633.77	12347.91	11343.42	14336.77	11565.42
Machine 5	15185.04	12994.49	11465.35	11092.43	11674.45	14430.01	12326.63
Machine 6	9903.67	8636.87	10871.16	9141.77	9371.44	8940.21	8228.22
Machine 7	10444.33	8650.17	11259.73	9535.53	9868.24	9289.83	8385.01
Machine 8	10254.61	8771.03	11325.76	9727.60	9900.63	9509.37	9012.02
Machine 9	8955.21	8120.54	9428.09	8152.38	8833.55	8826.68	10058.35
Machine 10	8451.64	7395.80	8849.70	7542.94	8482.12	7964.35	8642.82
Machine 11	17532.31	13864.29	12468.64	11876.81	13050.14	16189.34	14917.52
Machine 12	10012.75	11530.35	9404.61	8295.88	8487.98	8974.31	13283.24
Machine 13	12134.04	10987.19	12133.65	11430.34	11484.54	11298.24	10708.18
Machine 14	10691.53	10031.81	10461.36	10181.17	10229.79	10118.11	9318.02
Machine 15	9222.32	7919.57	9902.65	8316.05	8242.19	7826.18	8034.91
Machine 16	8852.10	7662.13	8201.17	7589.21	7812.79	7574.33	7545.77
Machine 17	10967.01	8306.09	10323.61	7916.07	7555.20	9485.42	7758.36
Machine 18	10958.58	8467.68	9005.82	8587.13	8256.41	9501.78	9834.71
Machine 19	8017.01	7933.61	8504.46	9118.21	10094.67	8262.18	9569.11
Machine 20	8949.31	7977.93	9179.50	8822.50	8447.41	8458.81	9565.77
Machine 21	15184.39	13592.79	14828.98	11270.90	11866.15	14627.13	12073.10
Machine 22	18446.94	9488.34	18404.45	7924.88	10320.68	16873.08	12601.66
Machine 23	19093.68	9868.90	19119.40	8451.49	10740.62	17445.73	12170.03
Machine 24	9090.43	8074.00	9233.84	8652.57	9041.96	8552.34	9588.10
Machine 25	11618.95	11012.28	11094.41	10679.88	11032.94	10932.03	9817.51
Machine 26	10011.99	8751.76	11453.23	9229.61	9869.78	9143.67	8803.29

TABLE D.2: sMAPE Values per Machine for Aggregated Predictions

sMAPE	Baseline	LR	DT	RF	GBM	KNN	HW
Machine 1	14.89%	14.55%	10.26%	11.35%	12.51%	12.64%	6.78%
Machine 2	10.60%	9.78%	11.62%	9.93%	12.12%	10.55%	8.41%
Machine 3	29.02%	24.85%	25.71%	19.39%	22.51%	27.34%	18.84%
Machine 4	26.72%	22.41%	20.32%	21.58%	20.01%	26.27%	19.07%
Machine 5	26.50%	23.05%	19.38%	19.07%	20.21%	26.00%	19.75%
Machine 6	12.29%	10.98%	14.05%	11.70%	11.99%	11.54%	9.64%
Machine 7	13.00%	11.02%	14.55%	12.28%	12.68%	12.04%	9.74%
Machine 8	11.62%	10.13%	13.54%	11.39%	11.59%	11.25%	9.66%
Machine 9	9.85%	9.12%	10.87%	9.10%	10.09%	10.14%	10.37%
Machine 10	9.30%	8.34%	10.21%	8.61%	9.74%	9.24%	9.14%
Machine 11	26.97%	21.33%	18.64%	17.99%	19.86%	25.64%	20.85%
Machine 12	8.68%	9.83%	8.20%	7.20%	7.36%	7.79%	11.05%
Machine 13	14.17%	13.21%	14.73%	13.58%	13.69%	13.79%	11.56%
Machine 14	50.14%	47.49%	50.08%	48.85%	48.66%	47.92%	44.06%
Machine 15	10.01%	8.84%	11.15%	9.36%	9.24%	8.82%	8.45%
Machine 16	9.59%	8.54%	9.29%	8.45%	8.72%	8.53%	7.97%
Machine 17	10.65%	7.89%	10.17%	7.68%	7.28%	9.36%	7.38%
Machine 18	11.23%	8.71%	9.54%	8.91%	8.51%	9.99%	9.74%
Machine 19	8.91%	9.11%	9.97%	10.57%	11.82%	9.65%	10.14%
Machine 20	9.49%	8.66%	10.22%	9.79%	9.26%	9.44%	9.68%
Machine 21	26.95%	24.48%	26.74%	19.79%	20.92%	26.96%	19.75%
Machine 22	43.00%	20.14%	42.98%	16.89%	22.14%	39.79%	24.16%
Machine 23	43.53%	20.50%	43.71%	17.66%	22.59%	40.28%	23.08%
Machine 24	9.93%	9.06%	10.54%	9.80%	10.25%	9.79%	9.92%
Machine 25	51.15%	48.93%	52.04%	48.33%	49.51%	48.85%	43.99%
Machine 26	12.81%	11.50%	15.32%	11.98%	13.04%	12.17%	10.49%

TABLE D.3: Bias Values per Machine for Aggregated Predictions

Bias	Baseline	LR	DT	RF	GBM	KNN	HW
Machine 1	-14697.69	-14016.37	-9216.51	-11250.06	-12404.34	-12447.74	-2572.93
Machine 2	-8489.53	-3852.24	-7024.76	-5519.92	-4229.17	-5472.56	4785.67
Machine 3	-12225.30	-7308.64	-7048.62	-3977.64	-7565.73	-6760.44	8271.95
Machine 4	-11174.46	-5828.50	2390.98	-2243.33	-6079.02	-5251.03	8202.09
Machine 5	-11001.79	-6336.65	1239.89	-1519.43	-6058.72	-5232.17	8960.28
Machine 6	-9499.89	-6187.10	-9495.16	-7517.88	-7998.49	-7227.90	6597.10
Machine 7	-10183.55	-6605.30	-8613.63	-8111.99	-8649.00	-7824.68	6741.34
Machine 8	-9845.17	-4789.71	-9964.07	-7754.66	-7951.71	-7147.22	7126.53
Machine 9	-8489.88	-3653.90	-7192.44	-6817.24	-6834.72	-5814.64	8694.12
Machine 10	-7727.40	-2450.66	-5362.70	-5639.51	-5609.26	-4776.49	7247.49
Machine 11	-12804.83	-4647.23	2294.49	-1939.50	-6643.36	-6423.10	12174.41
Machine 12	-4603.87	5952.63	-2812.75	-1263.18	-1622.63	-2260.71	10655.80
Machine 13	-12034.39	-8386.73	-10657.01	-10244.98	-10316.79	-8874.39	9376.74
Machine 14	2964.74	1915.13	-9225.36	-2462.76	-426.76	355.17	318.72
Machine 15	-8412.22	-6224.53	-7158.32	-6661.70	-7602.54	-6237.01	6392.93
Machine 16	-7890.47	-5388.54	-5583.06	-6416.93	-6920.85	-5758.42	5832.05
Machine 17	-10323.44	2444.70	-6805.16	-5563.93	-5574.05	-6825.38	2891.45
Machine 18	-9872.68	-1856.69	-6543.20	-6065.57	-6100.02	-7344.94	8459.28
Machine 19	-7263.22	-3261.15	-5744.70	-3901.65	-3258.20	-3883.84	8380.27
Machine 20	-8257.60	-2725.89	-6755.60	-6186.12	-6378.16	-5121.07	8100.67
Machine 21	-10744.44	-6351.61	-3024.76	-1829.38	-5866.93	-5146.74	9279.11
Machine 22	-18053.57	-1154.44	-18003.70	-2965.83	-7072.40	-15198.25	11779.89
Machine 23	-18609.99	-1244.25	-18555.13	-3549.62	-7286.51	-15720.56	11153.69
Machine 24	-8511.44	-4085.76	-6138.60	-6546.36	-7192.30	-5872.16	8260.19
Machine 25	4312.52	3357.36	-8871.82	-1375.80	544.65	1503.91	420.77
Machine 26	-9699.24	-6888.37	-10335.09	-8520.00	-8753.41	-7673.72	7153.63

TABLE D.4: MAD Values per Machine for Non Aggregated Predictions

MAD	Baseline	LR	DT	RF	GBM	KNN	HW
Machine 1	1864.86	1766.46	1846.54	1736.11	1745.63	1802.07	1692.27
Machine 2	1495.18	1270.58	1240.85	1338.32	1321.44	1382.83	1647.68
Machine 3	1807.63	1743.52	1766.03	1693.36	1685.09	1821.51	1681.90
Machine 4	1789.55	1711.65	1731.52	1749.65	1680.06	1813.19	1641.60
Machine 5	1839.58	1799.87	1833.67	1782.33	1751.27	1856.60	1730.11
Machine 6	1407.00	1249.25	1273.33	1263.95	1313.48	1334.30	1619.82
Machine 7	1427.23	1265.56	1263.40	1287.38	1290.98	1359.58	1562.01
Machine 8	1494.48	1262.26	1275.35	1324.86	1393.56	1376.19	1652.96
Machine 9	1440.89	1206.17	1353.92	1311.94	1345.41	1357.41	1459.32
Machine 10	1424.24	1122.75	1358.80	1210.23	1195.22	1246.47	1185.27
Machine 11	2138.93	1937.47	2213.07	2000.54	1937.19	2082.74	2249.00
Machine 12	1544.00	1226.04	1312.02	1278.24	1195.97	1455.32	2622.45
Machine 13	1552.07	1363.91	1674.07	1348.59	1366.21	1394.29	1734.45
Machine 14	1350.77	1365.60	1375.38	1316.52	1309.83	1308.28	1277.53
Machine 15	1481.94	1252.56	1177.24	1261.66	1247.49	1295.31	1479.35
Machine 16	1446.02	1198.99	1389.18	1218.49	1245.22	1269.82	1452.31
Machine 17	1634.68	1186.42	1390.19	1348.47	1358.46	1444.00	2651.48
Machine 18	1651.35	1213.22	1242.03	1397.98	1345.21	1464.88	1612.14
Machine 19	1415.09	1213.35	1410.90	1264.62	1336.34	1349.60	1616.65
Machine 20	1471.80	1199.58	1292.39	1300.51	1360.48	1356.55	1655.96
Machine 21	1796.94	1750.53	1829.27	1742.71	1845.15	1818.91	1702.64
Machine 22	2037.71	1635.49	1548.78	1628.24	1659.18	1835.85	1790.45
Machine 23	2079.31	1669.63	1606.38	1664.05	1683.93	1876.72	1836.24
Machine 24	1437.93	1208.66	1143.35	1298.73	1341.72	1331.63	1503.57
Machine 25	1432.81	1460.56	1549.97	1393.69	1390.65	1386.29	1318.91
Machine 26	1421.34	1292.14	1226.61	1298.56	1295.62	1334.32	1662.26

TABLE D.5: sMAPE Values per Machine for Non Aggregated Predictions

sMAPE	Baseline	LR	DT	RF	GBM	KNN	HW
Machine 1	28.62%	29.75%	26.29%	26.65%	27.76%	28.77%	27.64%
Machine 2	28.13%	27.38%	35.64%	27.55%	28.05%	28.77%	33.60%
Machine 3	49.21%	47.97%	47.64%	46.62%	46.84%	53.12%	45.69%
Machine 4	47.92%	48.51%	46.73%	44.99%	45.40%	52.26%	44.17%
Machine 5	48.25%	50.26%	48.78%	45.12%	46.27%	52.62%	45.40%
Machine 6	27.72%	28.45%	30.05%	27.56%	27.96%	28.59%	33.60%
Machine 7	28.05%	28.70%	28.75%	28.05%	28.50%	29.02%	32.43%
Machine 8	27.09%	25.96%	29.32%	26.50%	27.48%	27.40%	31.69%
Machine 9	25.56%	23.77%	24.05%	24.49%	25.54%	26.31%	27.45%
Machine 10	24.99%	23.33%	23.57%	23.20%	23.63%	24.09%	22.54%
Machine 11	48.73%	45.93%	51.56%	44.28%	45.32%	50.37%	45.09%
Machine 12	21.97%	19.04%	20.07%	20.44%	20.62%	20.85%	30.41%
Machine 13	28.34%	27.29%	26.21%	26.49%	27.32%	28.34%	33.94%
Machine 14	90.13%	90.69%	90.11%	90.03%	90.22%	93.23%	103.94%
Machine 15	25.62%	25.33%	24.90%	23.85%	23.98%	24.65%	27.32%
Machine 16	25.07%	24.47%	25.06%	23.48%	23.49%	24.29%	27.10%
Machine 17	23.92%	18.28%	24.00%	20.62%	20.58%	21.70%	31.29%
Machine 18	26.97%	22.05%	26.51%	24.63%	25.25%	26.14%	27.83%
Machine 19	25.51%	24.29%	27.08%	28.15%	27.76%	27.03%	31.01%
Machine 20	25.20%	22.79%	22.39%	24.46%	24.54%	25.91%	30.02%
Machine 21	47.83%	47.51%	49.53%	44.63%	45.93%	52.30%	45.14%
Machine 22	68.41%	53.54%	54.28%	53.35%	55.92%	65.94%	58.40%
Machine 23	68.23%	53.29%	54.27%	53.25%	55.74%	65.85%	58.61%
Machine 24	25.40%	24.19%	23.80%	24.97%	25.60%	25.83%	28.09%
Machine 25	89.46%	90.49%	88.86%	89.53%	89.49%	92.67%	100.51%
Machine 26	28.61%	29.89%	28.55%	27.72%	29.38%	29.27%	35.46%

TABLE D.6: Bias Values per Machine for Non Aggregated Predictions

Bias	Baseline	LR	DT	RF	GBM	KNN	HW
Machine 1	-954.76	-1124.31	-712.24	-857.75	-957.50	-990.31	-583.80
Machine 2	-533.09	-470.41	-536.02	-396.33	-269.61	-398.14	-807.29
Machine 3	-870.30	-751.32	-480.39	-565.50	-650.83	-504.01	-139.44
Machine 4	-791.04	-637.46	-250.52	-382.86	-568.06	-408.78	-37.38
Machine 5	-771.97	-704.65	-281.57	-326.00	-557.10	-403.01	-9.71
Machine 6	-615.04	-628.70	-499.89	-511.17	-590.78	-499.52	-940.79
Machine 7	-665.01	-667.44	-530.35	-557.64	-626.77	-549.78	-861.77
Machine 8	-631.21	-527.75	-530.22	-525.19	-587.84	-490.21	-889.69
Machine 9	-512.04	-422.89	-395.39	-391.44	-443.94	-369.15	-545.59
Machine 10	-467.10	-320.96	-118.83	-368.30	-372.14	-289.08	-5.26
Machine 11	-931.00	-589.27	-840.13	-458.39	-582.43	-504.65	1686.13
Machine 12	-110.77	448.30	-77.31	28.70	8.38	-50.20	2430.14
Machine 13	-804.66	-806.78	-712.68	-687.24	-778.02	-658.66	-1118.31
Machine 14	260.28	187.79	-488.27	128.53	116.15	88.49	-749.93
Machine 15	-506.20	-595.04	-476.89	-411.38	-503.59	-395.24	-654.86
Machine 16	-465.40	-529.88	-414.57	-417.87	-429.36	-348.83	-646.47
Machine 17	-667.89	182.11	-467.64	-335.79	-411.59	-84.56	2531.48
Machine 18	-620.08	-305.01	-547.65	-479.11	-534.17	-529.72	-608.67
Machine 19	-428.36	-426.10	-333.66	-273.23	-345.54	-274.65	-806.31
Machine 20	-521.20	-378.25	-164.49	-425.63	-456.27	-361.93	-834.11
Machine 21	-747.37	-679.70	-196.87	-321.27	-553.98	-404.33	17.21
Machine 22	-1420.51	-302.55	34.44	-431.96	-706.71	-991.92	-629.33
Machine 23	-1458.65	-307.04	20.63	-451.35	-731.87	-1025.95	-701.08
Machine 24	-516.59	-461.66	-406.70	-440.80	-489.95	-381.66	-634.09
Machine 25	371.55	296.68	-425.77	228.88	235.74	174.38	-668.55
Machine 26	-627.35	-691.00	-632.00	-536.34	-649.86	-531.02	-993.78

Appendix E

Non-cumulative Visual Performances

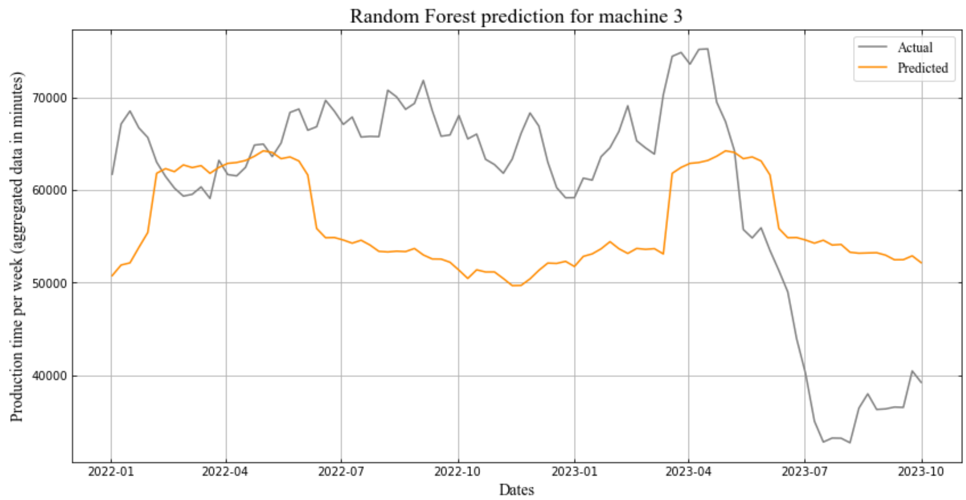


FIGURE E.1: Visual Performance of RF for Machine 3 (non-cumulative)

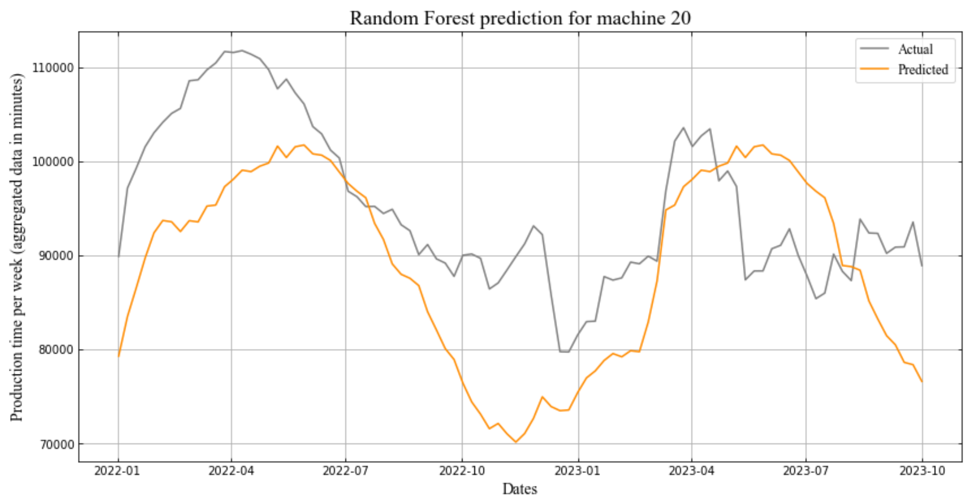


FIGURE E.2: Visual Performance of RF for Machine 20 (non-cumulative)

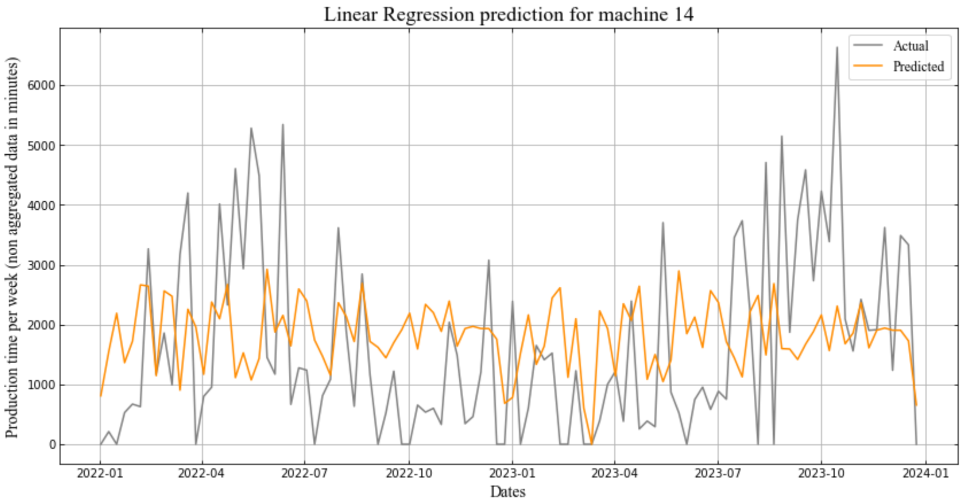


FIGURE E.3: Visual Performance of LR for Machine 14 (non-cumulative)

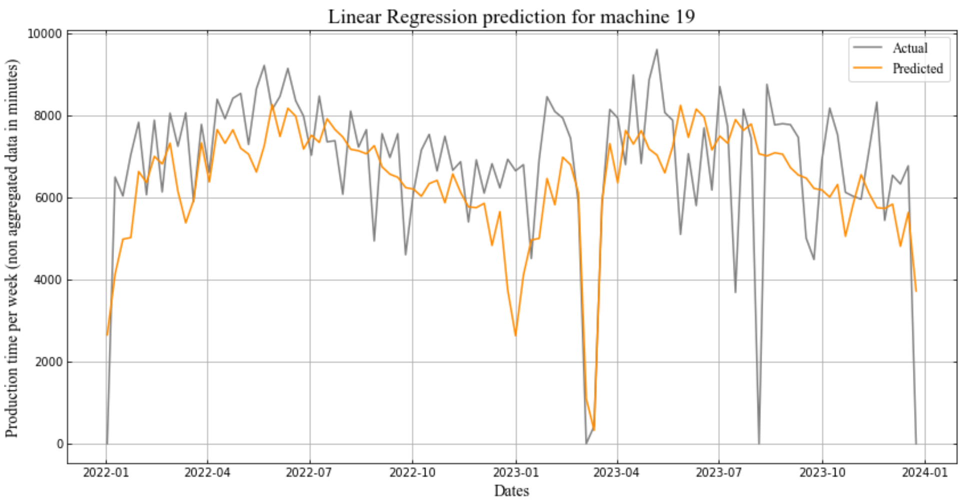


FIGURE E.4: Visual Performance of LR for Machine 19 (non-cumulative)

Appendix F

Declaration of Artificial Intelligence

"During the preparation of this work, I used ChatGPT from OpenAI to assist in writing and coding activities. After using this tool/service, I thoroughly reviewed and edited the content as needed, taking full responsibility for the final outcome."

-

Nils Meulenbroek, August 30, 2024