MASTER

Exploiting error data from a packaging machine to reduce machine downtime

van Nieuwpoort, S.L.

Award date:
2020

Technische Universiteit Eindhoven
University of Technology

Department Industrial Engineering and Innovation Sciences
Operations Planning Accounting & Control

Exploiting error data from a packaging machine to reduce machine downtime

S.L. van Nieuwpoort

Supervisors:
dr. L. Tan, TU/e, OPAC
ir. dr. S.D.P. Flapper, TU/e, OPAC
ir. A. Philipp, Bosch Packaging Technology

Eindhoven, November 2019

Abstract

Data from industrial assets can be valuable in gaining insight in the reasons for machine downtime. In this thesis, we gathered data from a packaging machine to determine which errors are related to machine downtime. After we determined which errors are related to downtime, we determined the possible causes of the errors. We investigated one error in a case study: the marker at the end of the film. We identified the replacement of the film in the data and we found that the film replacement caused at least 65% of the downtime of the related error. We investigated whether we can use the downtime of an error to preventively replace the film, and we determined at which remaining film lengths this would save costs. We can reduce up to 7.2% of the downtime from the error and the replacement by implementing the opportunistic replacement of the film. The exact value depends on the probability of having an operator available at the moment the error occurs and the cost of downtime.

Executive summary

The research is conducted at Robert Bosch Packaging B.V. in Weert in cooperation with one of the clients from Bosch, which will stay anonymous for confidentiality reasons. Bosch sells packaging machinery for the food industry. In this project we analyzed and reduced downtime of a packaging machine of the client, of the type SVE 2520WR.

Problem statement

Since packaging machines operate in a production line, their downtime may cause downtime for the entire production line. One of the reasons for this undesirable downtime, is unexpected machine stops. Every time the packaging machines from Bosch detect a problem, they stop and generate an error message. Bosch acquired the IoT Gateway, which allows to gather the error messages over time. However, Bosch has no method to obtain an overview of the reasons for downtime and Bosch does not know how downtime can be reduced based on the reasons for downtime.

"Currently it is unclear how data from a packaging machine can be used to get insight in the reasons for downtime and how machine downtime can be reduced based on the insight."

Approach

The approach in this thesis is structured according to the five steps from the Define, Measure, Analyse, Improve and Control (DMAIC) provided by Slack & Lewis (2002). First we *defined* the objective of the process improvement as reducing downtime from errors.

In order to determine which errors we should investigate, we *measured* the downtime caused by errors. We found that the client currently uses a method to find the downtime based on the logging duration of errors. The logging duration is the time between the moment an error is generated and the moment the error is cleared. However, due to the fact that the logging duration is not exactly equal to the downtime duration, we expect the method to provide a skewed overview of the downtime per error. We determined a new method to find the downtime per error based on the data in the error log and the data in the machine status log. The error log provides the moment an error occurs and the machine status log provides the period of downtime. We compared the results of both methods categorized in downtime from errors upstream, downstream and to the machine itself. We visualized the results in figure 1. We observed that the old method overestimates the amount of downtime. We decided to continue with the newly developed method.

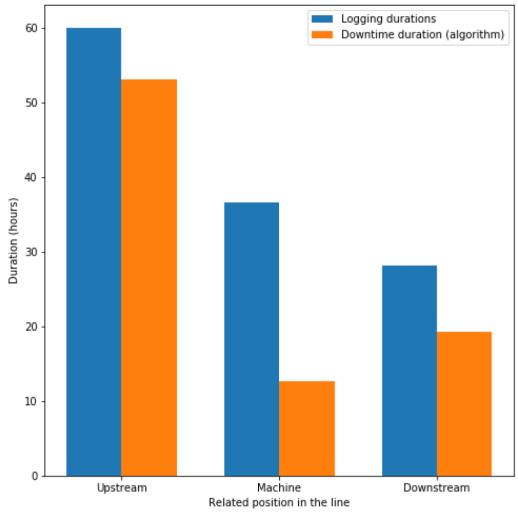


Figure 1: Comparison of the downtime overview of old method and the new method (Machine 3, May to June 2019, 93 hours of uptime)

Next, we *analyzed* the possible causes of the errors related to faults of the packaging machine. We visualize the downtime per error in figure 2. We applied Fault Tree Analysis (FTA) to find the possible causes of the three errors related to the highest amount of downtime. In order to determine whether a possible cause has caused downtime, we proposed a method that identifies possible causes in the data. This method is based on finding a machine interaction that is reflected in the data. We selected marker B as a possible cause for further investigation. Marker B is a marker at the end of the film. We could identify the occurrence of marker B in the data set through the reflections in the error log of the executed film replacements. We determined that marker B indeed caused 65% of the downtime of error 401.

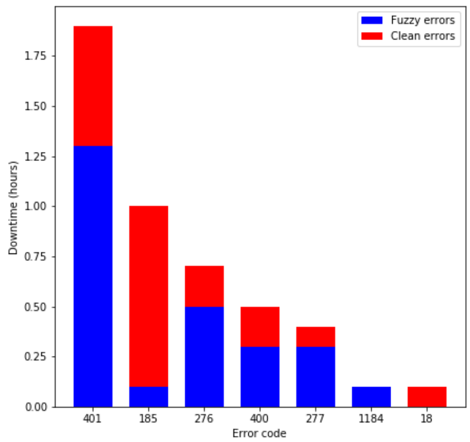


Figure 2: Downtime per error at machine 3, (production in May and June 2019, 93 hours of production)

In order to *improve* the use of the packaging machine, we decided to reduce the downtime resulting from marker B and the corresponding machine interaction: the film replacement. In order to reduce this downtime, we investigated whether the film can be replaced preventively during the downtime of an unexpected error. We selected error 160 as a suitable error for this purpose, since this error has considerable length and is not related to a problem at the machine nor related to a specific part of the film. We modeled the preventive film replacement in a Markov Decision Process (MDP). We incorporated the probability that no operator is available at the moment the film can be replaced, and we varied this probability between 0 and 1. We incorporated the relevant periods of downtime in the MDP, and we found at which remaining film lengths we should replace the film if error 160 occurs.

In order to *control* the improvement step, we determined the performance of using these boundary values in comparison to only replacing the film when it reaches its end. We incorporated the theoretical probability distributions that describe the data and we quantified the downtime reduction and the cost reduction by using discrete event simulation.

Results

We found that there exists a boundary value for the amount of bags at which it becomes optimal to decide to prematurely replace the film if error 160 occurs. If the amount of bags is equal to or less than this boundary value, it is optimal to replace the film at the occurrence of error 160. The exact boundary value depends on how costly it is if the line is down (C_{down}) and the probability of having an operator available at the moment error 160 occurs (P_{Av}). We provide the optimal boundary values for three different estimations of the downtime costs and for different values of the operator availability. Thereafter, we found that using the boundary values reduces the amount of downtime caused by the film replacement and error 160 by 0 - 7.2 %. The exact percentage depends on the actual cost of downtime and the probability that an operator is available. In absolute value, the reduction varies between 0 and 28 seconds per hour. Since we chose pessimistic parameter values, we expect that the actual reduction is slightly higher. We visualize the downtime reduction for three different cost scenarios in figure 3.

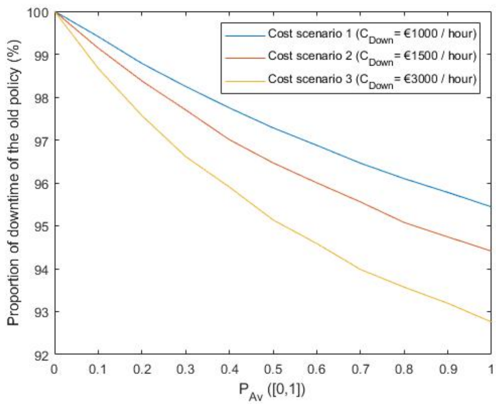


Figure 3: Downtime of the new policies as a percentage of the downtime of the old policy

Recommendations

We recommend the client to make use of preventive film replacements, using conservative boundary values provided by the model (the values for a high value for P_{Av}). The client should determine based on newly obtained data what the probability is that an operator is available at the moment error 160 occurs. Based on these values for P_{Av}, one can choose the correct boundary values.

The client should align the ordering of film with the different markers with their production schedule. Using films with markers in a production run that requires the entire film, causes unwanted machine stops, and therefore downtime. Films with these markers should only be used when the production schedule indicates that the amount of film until the marker is needed.

We recommend Bosch and the client to keep track of the reasons for manual machine stops. Manual machine stops are the most important cause of machine downtime, but currently their reasons remain unknown. In order to reduce the downtime resulting from the manual machine stops, more data on these stops is required.

Preface

This project is conducted in partial fulfillment of the Master of Science degree in Operations Management and Logistics at Eindhoven University of Technology (TU/e). The project has been conducted at Bosch Packaging in Weert, and in cooperation with one of their clients. In this preface I would like to thank the people who helped me write this thesis and who participated in the project.

The road of a project involving new technology is not always as smooth as you would like. Things are not always what they seem, and this requires a critical and flexible mindset. Several people guided me past the bumps in my road, starting with my supervisors from the TU/e. First of all, I would like to thank Lijia Tan. During the project you invested a great amount of time to improve the report and you kept me motivated to proceed until the end. Secondly, I would like to thank Simme Douwe Flapper, with whom I had thorough discussions about the mathematical part of the thesis and about the structure of the thesis.

During the execution of the project, several other people participated in the project. First of all, I would like to thank my company supervisor, André Philipp, for providing me with the opportunity to work on the project at Bosch Packaging. Secondly, I want to thank Silviu Stanimir. I am glad I could work with such a talented IoT specialist and I am grateful for your endless efforts to obtain a functioning technical setup. Lastly, I would like to thank Patrick Houbraken, who gave me useful advice during the project and kept all the project participants motivated.

Although this thesis does not mark the end of my student life, it does mark the end of the period of studying at the TU/e. During my studies, several people played an important role. Firstly, I am very grateful to have a brother with such patience and such extensive knowledge on mathematics and computer software. You have always been there when I faced challenges I could not oversee and you were always able to decompose my challenges in manageable pieces, using understandable language. Next, I want to thank my parents for their unconditional support during my studies. Lastly, I want to thank Marije, who has supported me and given me a positive attitude during the difficult parts of the project.

Simon

Contents

List of Figures

List of Tables

List of Abbreviations

DMAIC Define, Measure, Analyse, Improve and Control 2–4

FMEA Failure Mode and Effect Analysis 4, 5, *List of Concepts:* Failure Mode and Effect Analysis

FTA Fault Tree Analysis 5, 16–18, 24, 62, *List of Concepts:* Fault Tree Analysis

HMI Human-Machine Interface 6, 9, 63, 64

K-S Kolmogorov-Smirnov 46

MDP Markov Decision Process 5, 32, 36, 37, 44, 54, 75, 88

MLE Maximum Likelihood Estimation 5, 45, 46, 80–82, *List of Concepts:* Maximum Likelihood Estimation

OPC Open Platform Communications 64, *List of Concepts:* Open Platform Communications

PDF Probability Density Function 46, 48, 81, 82

List of Symbols

λ_{160} Arrival rate of error 160 during production.

b In the case study b denotes the remaining amount of film in the film roll expressed in amount of bags.

b_{bound} The value of the amount of bags at which it becomes optimal to start replacing the film if an error occurs in the next time step. This boundary value is found using the model in chapter 5 and one of the input parameters of the simulation in chapter 6.

b_{new} Total amount of bags on a new film reel.

C_{down} The cost parameter of machine downtime denoting the cost per time step.

C_{film} The cost parameter of the remaining material in the film denoting the cost per remaining bag on the film reel.

$C_{replace}$ The variable that denotes the amount of replacement costs made during the simulation in chapter 6.

C_{Total} Variable for the total cost during the simulation in chapter 6.

d The amount of time that the machine is down during the simulation.

D_{Sim} Parameter for the hours of production in the simulation in chapter 6.

m The status of the machine in the model of chapter 5. The machine can either be in production, undergoing a film replacement, down due to an error and undergoing film replacement during an error.

P_{Av} The probability that an operator is available at the moment an error occurs.

s The system state in chapter 5. The system state consists of a machine status m, and the amount of bags on the film reel, b.

$T_{Both,j}$ Stochastic random variable denoting the maximum of executing operation i during the recovering of error j.

$T_{ErrorReaction}$ Variable for the reaction time until the operator reaches the machine when the machine is down due to an error.

T_{inter} Variable for the remaining production time until the next error. The time is a draw from an exponentially distributed random number generator.

$T_{OperatorReaction}$ Variable denoting the operator reaction time. The operator reaction time is the time period starting from the moment the packaging machine reaches the end of the film until an operator starts replacing the film. Since operators may be busy at the moment the machine reaches the end of the film, the length of this period varies.

$T_{Recover,j}$ Stochastic random variable denoting the time for the machine to recover from an error.

$T_{Replace}$ Variable for the time necessary to replace the film. This is the time period starting from the moment the operator performs the first action on the machine to replace the film, until the moment the packaging machine resumes production.

T_{Sim} Variable for the amount of remaining simulation time (hours).

$T_{TotalReplace}$ Stochastic random variable denoting the time necessary to replace the film, when the end of the film has been reached. The variable consists of the operator reaction time and the replacement time.

v_m The set speed of the packaging machine.

$\overline{X}$ Mean of a sample. The sample mean is equal to the sum of the observations divided by the amount of observations. The sample mean is also referred to as the first moment.

List of Definitions

clean error occurrence Error occurrence accompanied by an automatic machine stop. During a clean error the machine does not show any indicator of the operator making mechanical adjustments to the machine or the operator pressing the stop button. 11, 16

error A fault that is detected by the machine. 6, 9, 17, 64

error logging duration The time between the moment an error message is raised by the machine and the time the error message is cleared. 9

error message The message the Human-Machine Interface displays when detecting the error. 6

Failure Mode and Effect Analysis Widely used method to list and rank failure modes. The method aims at finding the impact, frequency of occurrence and the detectability for each the failure modes of an asset. The failure modes are ranked based on a priority number. 4

fault Loss of machine functionality. In this context, the loss of functionality can be a breakdown, reduced machine speed and bags that do not meet the output requirements. 4–6, 9, 11, 16–19, 23, 29, 33, 35, 62, 64

Fault Tree Analysis A structured method to find the events leading to an undesirable event (Haasl, Roberts, Vesely & Goldberg, 1981). 5

fuzzy error occurrence Error occurrence in which the operator pressed the stop button instead of resuming production, or the machine shows indicators of a mechanical adjustment. 11, 16, 27

indicator An indicator is defined as the reflection of an interaction with the machine in the machine data. 23

IoT Gateway The gateway that is used to extract data from the packaging machine. Without going in too much detail, the IoT Gateway allows us to gather the error log, the machine status and the bag counter. 1, 3

marker A marker placed on the film by the film reel supplier. Markers are detected by the machine and cause the machine to stop. 20

Maximum Likelihood Estimation A commonly used method in literature to estimate the parameters of the sampling distribution by maximizing the likelihood of observing the set of observations in the data set. 5

new method The method we propose to determine which errors are related to downtime. The method is used to find the amount of downtime an error relates to based on the error log and the machine status log. 9

old method The method the client currently uses to determine the impact of errors. The method is based on the logging duration of errors. 9, 10

Open Platform Communications A standard for communication between different industrial control devices. 64

operator reaction time The time between the moment the machine is down and the moment the operator starts replacing the film. 28

Python Python is a programming language that is characterised by its intuitive and flexible character. 11

secondary packaging machine The secondary packaging machines, or the case packers, are the machines placed after the packaging machines that we investigate in this research. The packaging machine packs the product in a bag and the secondary packaging machine packs the bags into boxes. The bags are transported to the secondary packaging machine via a transportation belt. 33

timestamp The date and time of a data point. The data points in the error log have a timestamp that contains the year, month, date, hour, second and millisecond. 9, 10

total replacement time The time starting from the moment the machine is down because it reached the end of the film, until the film has been replaced and the machine is up and running. This time period consists of the operator reaction time and the replacement time. 28, 35

Traksys information system A line information system that is used in plants to log information from industrial machinery. 9

undetected fault Fault that is not detected by the machine and is to be detected by the operator. 6

warm-up period The period at the start of a simulation containing some sort of queuing process. In the warm-up period, the system is not yet in steady-state due to the fact that the simulation started with an empty queue. 58

Chapter 1

Introduction

When using industrial machinery, the occurrence of unexpected machine stops causes downtime. This downtime is highly undesirable since it implicitly incurs costs for personnel, material and unmet demand. In recent years, more and more industries started to see the value of analyzing data generated by industrial machinery, in order to reduce unexpected machine stops and their related downtime. In literature there are several cases in which the analysis of event logs and error logs from industrial machinery leads to a reduction of machine downtime. Botman (2017) related errors to physical defects and uses the triggering sensor data to monitor for condition monitoring and the timing of preventive maintenance. López (2017) employs error data to predict machine failure and preventively execute maintenance. In this thesis we focus on machine data from a packaging machinery for the food sector. Packaging machines are positioned in a production line, which often means that when the packaging machine is down, the entire line is down. This raises the need to reduce the causes of downtime of packaging machines. In this thesis we focus on identifying which errors are related to downtime and what their possible causes are. We select one possible cause and reduce the downtime resulting from this cause.

The research is conducted at the company that manufactures the packaging machine: Robert Bosch Packaging Technology B.V. In the remainder of this thesis, we refer to the company as Bosch. Until now, Bosch has only manufactured and sold packaging machines, without focusing on what problems their clients face in practice. On the one hand, Bosch is not involved in the daily maintenance of the packaging machines operational at clients, so little is known about clients' problems during production. On the other hand, clients lack machine knowledge that is essential to understand the relation between unexpected machine stops and their causes. This project bridges this gap by analyzing the downtime a client faces during the use of a packaging machine.

Bosch identified the need of its clients to employ data generated by their machines to get insight in the performance of the packaging machine. Bosch started to acquire the IoT Gateway, which allows to gather process data from packaging machines in the field. In this project, we use the IoT Gateway to gather data from three packaging machines operational at one of the clients of Bosch. We focus on the downtime of one specific packaging machine, since the downtime of that machine is assumed to be most costly.

The outline of the remainder of this thesis is as follows: in section 1.1 we give information about the companies involved in this project, in section 1.2 we discuss the steps to obtain process improvement, in section 1.3 we elaborate on the research design, which consists of a problem statement, the deliverables, the research questions, the scope and our approach in this project. Next in section 1.4, we discuss the outline of this thesis.

1.1 Company background

Bosch manufactures and sells vertical filling forming and sealing packaging machines in the food industry. The food packaging industry is expected to grow moderately in the coming years[1]. The packaging machines are slightly customized per customer, which means that a machine from a different customer might have a specific functionality more or less. However, the majority of the subsystems of all packaging machines do not differ. Bosch has solely been focusing on building their machines, and selling their machines. Bosch experienced that clients often do not obtain the full machine performance due to suboptimal use of the packaging machines. Bosch expects that their machine knowledge can be used to assist clients in optimizing their machine use. The first step to improving the clients' use of their packaging machine, is analyzing data from the packaging machine.

The project is executed in cooperation with a client, which will stay anonymous for confidentiality reasons. The client has its own operator trainer, which trains the operators such that they are able to operate the machines and perform daily maintenance activities. For complicated issues, the clients maintenance engineers are expected to find a solution. Only when the client finds more complex and recurrent issues, a maintenance engineer from Bosch packaging will visit the client to redress the problem.

1.2 Process improvement cycle

The steps in this thesis are based on the steps in the structured improvement cycle proposed by Slack & Lewis (2002). This cycle structures process improvement into 5 steps: Define, Measure, Analyse, Improve and Control (DMAIC). The cycle is shown in figure 1.1. We briefly explain each of the steps in the cycle.

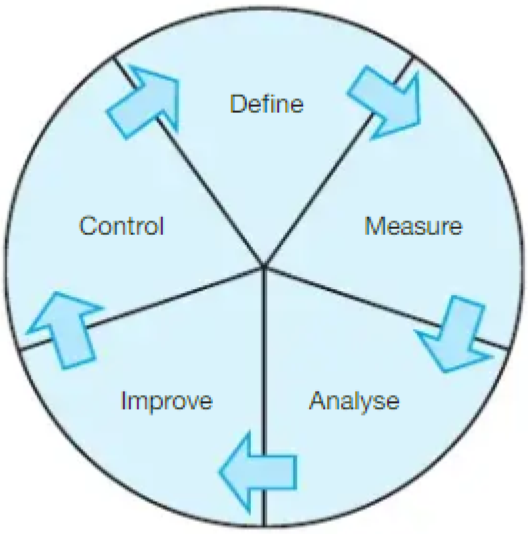


Figure 1.1: DMAIC cycle (Slack & Lewis, 2002)

Define The first step is to define the objective of the process improvement.

Measure The second step is to measure the performance of the current process. The measurements should provide evidence for the direction of the following steps.

[1]Food Packaging Market Research Report Forecast to 2023: MRFR. (n.d.). Retrieved October 8, 2019, from https://www.marketresearchfuture.com/reports/food-packaging-market-2086

Analyze The third step is to analyze the measurements of the current process. In this step one can develop hypotheses of what the root causes are of the process performance being lower than desired.

Improve Once the causes are identified, one can raise ideas to improve the process. This step consists of testing, implementing and measuring potential improvements.

Control The control step consists of continuously monitoring the improved process, such that one can determine whether the process improvement is sustaining.

1.3 Research design

In this section, we define the objective of the process improvement, which is the first step of the DMAIC cycle. We present our problem statement (subsection 1.3.1), the deliverables (subsection 1.3.2), our research questions (subsection 1.3.3), we define the scope (subsection 1.3.5) and our approach (subsection 1.3.4).

1.3.1 Problem statement

In this thesis we focus on analyzing data from a packaging machine generated during production runs, in order to reduce machine downtime. Every time the machine stops because it detects a problem, the machine stop is accompanied by the generation of an error message. Via the IoT Gateway we gather data of the occurrence of error messages. However, Bosch does not have a clear method to process the data on these messages to get an overview of downtime per error. We define downtime in the context of this project as the time during a production shift, in which the machine is not producing. Since the lifetime expectations of the parts of the packaging machines are generally longer than the time span of this project, we do not have data about part failures. Therefore we do not focus on errors related to broken parts of the machine. We know the data used in this project is coming from periods in which the machine had no defects.
When we know which errors are related to downtime, we are interested in their possible causes. Knowing the possible cause is key in identifying opportunities to avoid the downtime from errors. Based on the cause of an error, we need to revise the clients process of operating the machine in such a way that we reduce the downtime. It is unclear how we can reduce the downtime related to a cause. We define the following problem statement:

"Currently it is unclear how data from a packaging machine can be used to get insight in the reasons for downtime and how machine downtime can be reduced based on the insight."

1.3.2 Deliverables

Based on the problem statement, we identified several deliverables. We discuss the deliverables in the following categorization: deliverables for Bosch, deliverables for the client and deliverables for literature.

Firstly, we identified a deliverable for Bosch. Bosch would like to have a method to identify the errors that are related to downtime and to determine the amount of downtime the errors relate to. Since Bosch currently does not have this method, we identify the development of this method as a deliverable (i).

Secondly, we identified two deliverables for the client. For the client it is important to know what the possible causes are for their machine downtime. We identify the identification of the possible causes as the second deliverable (ii). Subsequently, we want to reduce the amount of downtime resulting from the causes. We identify the development of an approach to reduce the downtime of the causes as the third deliverable (iii).

Thirdly, we contribute to literature in two ways. We provide a structured method to use machine data to get insight in the machine downtime and the related causes (iv). Furthermore, we develop a model in which we incorporate data on the duration of errors to opportunistically schedule a maintenance operation, namely the film replacement (v). We apply the model in a case study and show the improvement numerically.

1.3.3 Research questions

Given the problem statement, we define the main research question:

How can we use machine data from a packaging machine to find causes of downtime and reduce the downtime?

The project can be divided into several research questions that help answering the main research question.

1. How can we identify the errors related to downtime?

The first research question involves defining a method to find the errors related to downtime and determining the amount of downtime an error relates to. In order to find the downtime per error, we process data coming from the machine. Answering the question results in deliverable i. Moreover, by answering the research question, we *measure* the current performance of the process, which is the second step of the DMAIC cycle. The research question is answered in chapter 2.

2. What are the possible causes of the errors related to downtime?

The second research question involves determining the possible causes of the errors related to downtime. Considering the available amount of time for this thesis, we decide to determine the possible causes of the top three errors. Answering the second research question results in deliverable ii. Furthermore, finding the possible causes is part of the *analyze* step from the DMAIC cycle. The research question is answered in chapter 3.

3. How can we reduce the downtime of the possible causes?

In subquestion 2, we identify the possible causes of the errors related to the most downtime. We find that the film replacement is one of the possible causes of downtime. This research question involves a case study in which we further examine this cause, and explore an approach to reduce its related downtime. First we further *analyze* this possible cause and verify that the possible cause has resulted in downtime (chapter 4). Next, we explore how we can *improve* the process by reducing the downtime of the film replacement (chapter 5). In maintenance, a well-known approach to reduce downtime is by opportunistic scheduling of operations (Zhu, 2015). We explore whether we can opportunistically schedule the film replacement. Lastly we quantify the reduction of downtime (chapter 6). Answering the third research question results in deliverables iii and v. Furthermore, we finish the *analyze* step and we fulfill the *improve* and *control* steps (DMAIC).

1.3.4 Approach

In this section, we elaborate on the approach to answer the research questions. Furthermore, we discuss to what extent we are compromised by the availability of data in answering the research questions.

1. How can we identify the errors related to downtime?
In an industrial setting, the machine performance is decreased by the negative impact of faults. A commonly used method to get insight in the impact of faults, is Failure Mode and Effect Analysis (FMEA) . In FMEA, one determines the impact and frequency of occurrence per failure mode.

Currently, the client uses a method to determine the downtime per error. We determined whether this old method gives a reliable overview of the relation between errors and downtime. Since this was not the case, we developed a new method. In our context, faults result in a machine stop. Every machine stop is accompanied by the generation of an error message and stored in the error log. We used this log to determine the occurrence of errors. The machine status log contains data that tells whether the machine is down or not. We used the machine status log to determine the amount of downtime related to each error message.

The FMEA technique aims at listing all possible faults in a structured way and ranking them according to their negative impact, frequency of occurrence and detectability. We only use the error log to determine the occurrence of errors, which only consists of the part of the faults detected by the machine. The client does not gather data about the faults that are not detected by the machine. This also implies that we drop the detectability factor of the errors, since each error is by definition detected by the machine. In the ideal case, we would also have considered faults that are not detected by the machine.

2. What are the possible causes of the errors related to downtime?
This research aims at determining the possible causes of the errors during production. A commonly used method to identify root causes of a fault, is by Fault Tree Analysis (FTA) . In FTA, one constructs a diagram starting with an undesirable physical state of the machine and then uses logic to make links to the possible causes leading to this state. We used the internal data from Bosch to determine what triggers the error and thereafter we used data from the operator trainer from the client to identify the possible causes.

In order to know which possible causes should be addressed, one ideally wants to know which possible cause has resulted in which proportion of the faults. However, there was no data available that directly shows which possible cause led to the occurrence of the error. We selected one of the possible causes for investigation in a case study. We decided to investigate the film replacement since it is a very probable cause for downtime and it is possible to identify the film replacement in th data .

3. How can we reduce the downtime of the possible causes?
The previous research question aims to find the possible causes of the errors related to downtime. In this research question, we aim to reduce the downtime resulting from the film replacement. Firs, we identified the film replacement using error messages in the error log. The error messages reflect particular steps of the film replacement and we used their corresponding timestamp to find the duration of the film replacement. Next, we explored the film replacement can be scheduled within error downtime. We selected a suitable error based on its frequency of occurrence and its duration. We incorporated the duration of the error and the film replacement in a MDP in order to determine at which remaining film length the client should use the error to replace the film. In order to find the expectation of all variables we use in the model, we fitted theoretical distributions to durations in the data by Maximum Likelihood Estimation (MLE) . Lastly, we quantified the process improvement and we determined whether the solution of the model performs well if distributional assumptions are relaxed. We used discrete event simulation to compare the old policy for replacing the film and the new policy with preventive replacement. We incorporated the theoretical distributions we fitted on the data.

In order to determine whether an operator should use an error to replace the film, we incorporated the operator availability at the moment an error occurs. The operator availability, determines the probability that the opportunity arising from an error can actually be seized. We did not know the availability, so we modeled and simulated different operator availabilities and we determined a robust solution.

1.3.5 Scope

The machine analyzed in this project, is the SVE 2520 WR, which is a continuous motion vertical packaging machine that is capable of packing food with a speed up to 200 bags per minute. The client offered to use one of his lines in this research, and the packaging machines in this line are of the aforementioned type. Each machine type differs slightly in possible bag sizes, bag styles. Some machine types have a functionality more or less, but the general functionalities do not differ from type to type. Inside the machine sensors provide data that is used for internal feedback loops, and generating error messages. The service department from Bosch estimates that the lifetime the components of the SVE 2520 WR are longer than the period in which this project is executed. Since we had no historical data available on the failures of parts, we did not consider the downtime related to component failure. López (2017) and Botman (2017) showed that worn out parts may cause errors, so we ensured that we did not consider machine data from a period in which a part turned out to broken or the machine had a mechanical defect.

During production of bags, faults may occur. A fault is defined as an abnormal condition in which a machine fails to perform its required function (General Services Administration, 1980). Faults can either be detected by the machine or by the operator. We show the classification of faults in figure 1.2. If the machine detects the fault through its error detection algorithms, the fault is called an error. If the machine is running, an error leads to a machine stop and an error message on the Human-Machine Interface (HMI). The machine can also detect a fault if the machine is already stopped by an error. Both errors are then shown on the HMI. If the machine does not detect the fault the machine continues working with reduced functionality. Reduced functionality in this case means that the output does not meet the requirements, e.g. the bags are not sealed or the sealing is not aligned according to the standard. This subset of faults is called undetected faults. The machine continues in a faulty state until the operator notices the malfunctioning and stops the machine manually. The machine generates a general error message indicating a manual operator stop. In this thesis, we solely focus on reducing downtime resulting from errors due to the unavailability of data on the undetected faults.

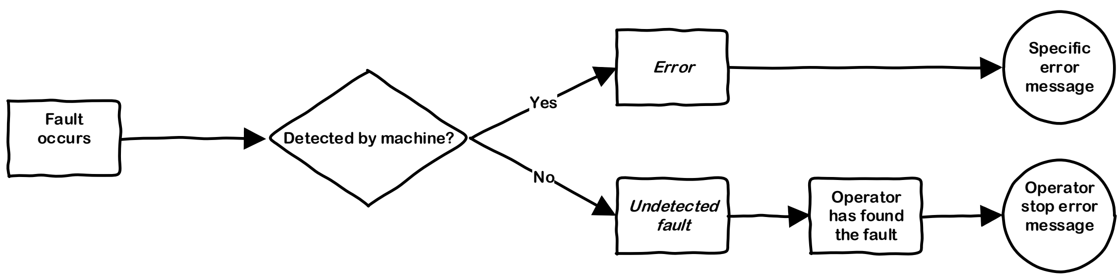


Figure 1.2: The classification of faults

During this project we gathered the data of three packaging machines in a packaging line that consists of four packaging machines. We give a schematic overview of the packaging lines with the four packaging machines in figure 1.3. For technical reasons, we could not gather the data of packaging machine 4, which is therefore out of scope. Since we are restricted in terms of time, we decided to focus on one of the remaining three packaging machines. Based on the configuration of the line, we expect the downtime of packaging machine 3 to be most costly, and therefore we only investigate the downtime of packaging machine 3 in this thesis.
All machines are connected to the same upstream belt that supplies the machines with the input product. The supply of product into the machines is regulated by a priority number. If the priority of a machine is low in comparison to another machine, the other machine will get more product. The configuration of the priority numbers is set by the operators on working on the line and is not logged.

After the packaging machine has packed the product, the product is carried away on a belt to a secondary packaging machine. Machines 1 and 2 are connected to the same belt, and are packing in parallel, while machines 3 and 4 are connected to a single belt. From the operator supervisor we know that machines 1 and 2 are usually used to produce smaller bags. Machines 1 and 2 are usually producing at 50 bags per minute (in parallel), so if both machines are operational they produce 100 bags per minute. The larger bag types are usually scheduled to be produced at machine 3, and this can also done at a set speed between 50 and 70 bags per minute. A reason to schedule the larger bags on a single machine and the smaller bags on machines 1 and 2 in parallel is that the secondary packaging machine cannot cope with larger bag types at 100 bags per minute. Machine 4 is mostly used at a low priority to catch the remaining product that is not going into machines 1,2 and 3. Since machine 1 and 2 produce in parallel, we expect that the downtime of these machines is less costly than the downtime of machine 3. The manager involved with the line supports this claim by his observations in the past. He observed that the bigger product types are produced at a lower speed and by only one machine and that therefore most pressure lies on machine 3.

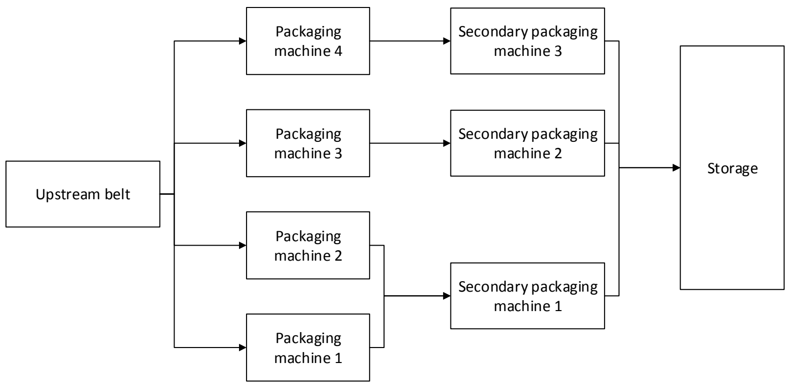


Figure 1.3: Schematic overview of the clients' line of with the four packaging machines

1.4 Thesis outline

In this first chapter, we provided a general introduction and defined the objective of the process improvement: reducing machine downtime. In chapter 2, we determine how we can measure the downtime of errors. By doing so, we obtain deliverable i and answer research question 1. In chapter 3 we analyze the possible causes of the errors related to downtime. We obtain deliverable ii and answer research question 2. In chapters 4 to 6 we conduct a case study, in which we investigate whether we can reduce the downtime of a specific cause. In chapter 4, we select one of the possible causes, namely reaching the end marker of the film. We determine the downtime that reaching the end of the film causes by identifying the replacement of the film in the data. In chapter 5, we propose a model to reduce the downtime resulting from the film replacement by using unexpected breakdowns to preventively replace the film. In chapter 6, we quantify the improvement of the policy for the timing of the film replacement using simulation. In chapters 4 to 6, we obtain deliverable iii and we answer the third research question.

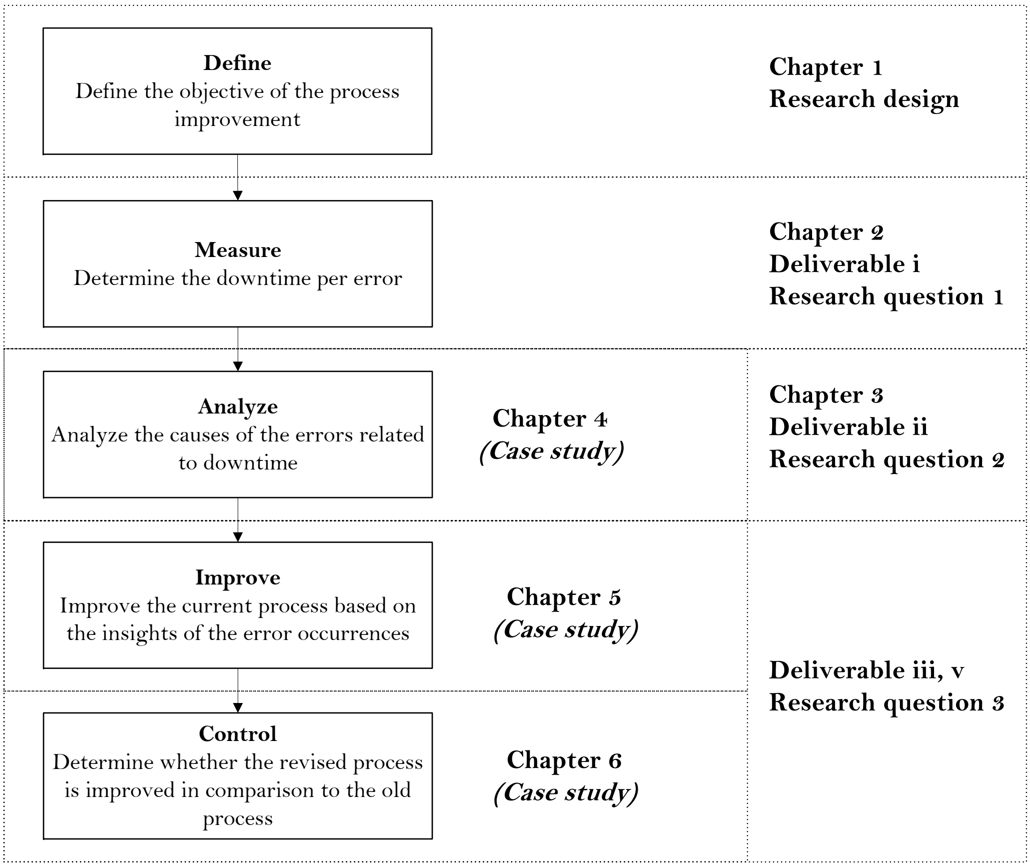


Figure 1.4: Steps in the thesis with the corresponding chapters and research questions

Chapter 2

Data Processing

In this chapter we aim to find a method to determine which errors are related to downtime and the amount of downtime. The method enables us to rank the errors according to their downtime and to investigate the errors that are related to most downtime. Note that we consider errors generated during production shifts, rather than errors outside production shifts. No mechanical defects were found during the span of this project, so we are interested in the faults that cause downtime in the production shifts. Currently the client already uses a method to determine the impact of errors based on the error logging duration. The error logging duration is the time the packaging machine is logging an error message on the HMI, which consists of the time period from generating an error until the error is automatically cleared by the machine or manually cleared by the operator. We refer to the method of the client as *the old method*. We review this method and find that the method neglects important aspects of the relation between errors and their corresponding downtime. In order to take the neglected aspects into account, we develop a new method and we compare the results obtained with both methods. We refer to the method we propose as *the new method*.

The new method combines two of the sources of machine data: the error log and the machine status log. The error log contains the codes of all error messages generated by the machine, the timestamp of occurrence, and the timestamp the error message is cleared. Error messages can be cleared by operators pressing the reset button, and a few error messages are cleared automatically. We are particularly interested in the error messages related to errors. The machine status log contains a number for the machine status, at every point in time that the status changes. We are interested in the time the machine is not producing after facing an error. The format of the error log and the machine status log are described in appendix B.

In section 2.1 we discuss the old method, and which aspects are neglected by this method. Then in section 2.2 we discuss how we can combine the error log and the machine status log using a new method. Next, we discuss the algorithm of the new method in section 2.3. Thereafter, we discuss the result of the new method in section 2.4. Lastly, we conclude in section 2.5.

2.1 The downtime measured by the old method

Currently the clients machines are installed with the Traksys information system. This information system keeps track of the error logging durations of error messages. After carefully examining the error log and the machine status log, we concluded that using the logging duration neglects the following aspects: (1) One fault may result in two errors with an overlapping duration, thus using the error logging duration may overestimate the actual downtime. (2) The logging duration of an error is not necessarily equal to the downtime. The logging stops as soon as an operator presses the reset button, while he still needs to solve the actual problem. This means that for a subset of errors, the old method may underestimate the downtime. (3) The machine generates errors when the machine is not producing. When the operator makes mechanical adjustments during a machine stop, the machine may generate error messages that reflect actions of the operator. An example is the adjustment of the film reel during a machine stop. Both the machine stop and the adjustment lead to an error message, which is logged for a certain period. This means that the sum of logging duration of the errors may be longer than the actual downtime.

Due to these neglected aspects, we expect that the overview of logging durations provides a skewed view on the total downtime. The first and the third aspects we mentioned, cause overestimation of the downtime related to errors and the second aspect causes underestimation of the downtime related to errors. Based on manual examination of the error log and the machine status log, we expect the overestimation to outweigh the underestimation. We expect the old method to overestimate the total downtime duration of errors. We decided to develop a new method which does consider the aspects the old method neglects. We know that the error log provides the timestamp of the occurrence of an error, and that the machine status log provides the machine status over time. The machine status can be used to determine the downtime of the machine. Using the machine status to measure downtime ensures that we do not overestimate the downtime. However, we have to determine how we divide the downtime over the occurring error messages in the error log. We explain how we defined the logic in the following section.

2.2 Determining the logic of the new method

In order to combine the error log and the machine status log, examined both data sources. Since the error log contains all errors, the method should iterate over the entries of this log. Then for the errors in the log, the method should determine whether the machine status log shows that the error caused a machine stop. We explain four key observations with important implications of how the method should combine the error log and the machine status log.

- **An error message can be either a notification, warning or error**
  The error logs consist of all messages that are generated by the machine over a period of time. A message can either be a notification, warning or an error. Notifications are used by the machine to inform the operator about current settings or the current machine state. Warnings do the same, but these messages indicate that the settings or machine state may result in an error. If the message is an error, the machine stops production (if it was producing). Since we machine performance is measured in downtime, we are only interested in the error messages indicating errors. The key observation is that the new method only has to determine whether these error messages have resulted in a machine stop.

- **The timestamp may deviate two seconds from the timestamp of the change in machine status**
  The packaging machine is configured such that whenever the machine runs into an error during production, the machine is stopped and the machine status is changed from executing (= producing) to either suspended or suspending (= stopped by a problem). This means that we can use the machine status log, to determine whether an error is accompanied by a machine stop. However, the packaging machine prioritizes on its operations and may generate the error up to two seconds after changing the machine status. In some cases, the error is generated slightly earlier than the machine is stopped. The key observation is that in order determine whether an error resulted in a machine stop, the method has to examine the machine status log two seconds before until two seconds after the error occurrence. This time period is based on the observed amount of deviation between the timestamp of an error and the corresponding change in machine status. If there is no change in machine status from executing to suspended, the error did not result in a machine stop.

- **Errors may require operator actions**
  When an error occurs, the related problem may require mechanical adjustment to the ma-

chine. An example of a mechanical adjustment that needs to be executed every now and then, is the adjustment of the position of the vacuum belts. Since our goal is to find the causes of downtime, it is relevant to determine whether an error required mechanical adjustment. Errors that required mechanical adjustments take longer than errors that can be fixed by starting the machine again. We can distinguish the errors during which the operator made mechanical adjustments by examining the notifications and warnings. For most mechanical adjustments the operator has to open the doors, which results in an error message. Only adjustments to the film reel do not require the opening of doors. However, adjustments to the film reel are also detected by the machine. The key observation is that the new method should determine whether the error log shows indications of mechanical adjustments during the downtime of errors. In this way, we can distinguish the errors that required mechanical adjustments from the errors that did not.

- **One fault may result in two error messages**
  It may happen that a fault triggers two different error messages. We observed that the machine may generate two error messages in less than a second interval. We examined these cases manually in the error log, and concluded that this happens when sensors monitor a similar function and both trigger a different error message. We observed that the sequence of these errors is fixed and we decided to add the downtime to the error that is detected first. The key observation is that the new methpd should verify that there are no other errors occurring close to the same machine stop. If there are other errors within two seconds of the machine stop, the method should consider the first error as the error related to the downtime period.

In the appendix A, these key points are illustrated with examples. The logic behind the observations is used to develop the algorithm in section 2.3.

2.3 The algorithm of the new method

In this section we propose the algorithm that considers the aspects neglected by the old method mentioned in 2.1. The algorithm consists of several subalgorithms. For the sake of clarity and readability, we refer to these subalgorithm by using self-explanatory names. A more detailed description of the input arguments and the output of the subalgorithm can be found in table 2.1. The algorithm combines data from the machine status log and the error log such that we obtain a list of error messages related to a machine stop with the corresponding downtime. During the downtime of an error, the machine may have generated notifications that imply that the operator undertook an action. These actions can be opening doors, manually stopping the machine, manually adjusting the film reel and activating the film clamp. We refer to errors that contain these notifications in their downtime as *fuzzy error occurrences*. If an error does not have these notifications in its downtime, we refer to it as a *clean error occurrence*. The algorithm determines whether an error occurrence is a clean error occurrences or a fuzzy error occurrence. Note that manual operator stops also result in an error message. The algorithm handles these error messages as errors, however, we can not investigate the causes of these stops due to the unavailability of data. The algorithm is implemented in Python.

Algorithm 1: Error processing algorithm

```
input  : Error log, machine status log, list of errors, warnings and notifications
output: List with a selection of the errors and their characteristics
foreach error message in the error log do
    if IsAnError=True then
        if ExecutingCheck=True then
            if ClashingErrorCheck=False then
                Save DowntimeStart, DowntimeEnd, OperatorActionCheck,
                 OtherErrorCheck
                if EndedWithOperatorStop=True then
                    Save EndedWithOperatorStop, OperatorStopStamp
                else
                    next;
                end
            else
                next;
            end
        else
            next;
        end
    else
        next;
    end
end
```

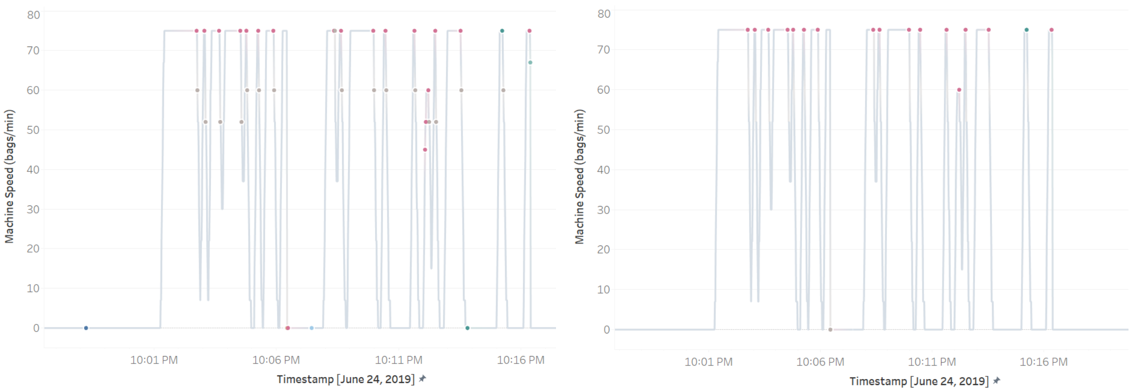
Table 2.1: Subalgorithms

Subalgorithm	**Input**	**Output**
IsAnError	The code of the error. Uses a list of errors, warnings and notifications	Returns boolean True if the code belongs to an error, and False otherwise.
ExecutingCheck	The timestamp of an error occurrence. Algorithm uses the machine status log.	Returns boolean True if the machine was executing and boolean False if the machine was not executing within a prespecified range around the input timestamp.
ClashingErrorCheck	The timestamp of an error occurrence. Algorithm uses the machine status log and the error log.	Returns boolean True if there is another error closer to the moment the downtime starts and returns boolean False if not.
DowntimeStart	The timestamp of an error occurrence. Algorithm uses the machine status log.	The timestamp that marks the start of the downtime in the machine status log if the ExecutingCheck returned boolean True.
DowntimeEnd	The timestamp of an error occurrence. Algorithm uses the machine status log.	The timestamp that marks the end of the downtime in the machine status log if the ExecutingCheck returned boolean True.
EndedWithOperatorStop	The start timestamp of the downtime and the end timestamp of the downtime. Uses the error log	Boolean True if there is an operator stop, boolean false if there is no operator stop.
OperatorStopStamp	The start timestamp of the downtime and the end timestamp of the downtime. Uses the error log.	The timestamp of the operator stop.
OperatorActionCheck	Two timestamps that mark a period of downtime. Algorithm uses the error log.	The algorithm returns list of actions. Actions can be opening doors, activating film clamp, activating vacuum bar.
OtherErrorCheck	Two timestamps that mark a period of downtime. Algorithm uses the error log.	Checks whether other errors are occurring in between two timestamps and returns a list of these errors.

2.4 Result of the new method

In this section we discuss two important aspects of the result of the application of the algorithm: the algorithm has filtered out the error messages related to a machine stop, and the overview of downtime per error altered significantly by doing so. The data we used, is coming from machine 3 from the months May and June 2019. In these months there were 3 production shifts that add up 93 hours of uptime. From a statistical point of view, it would be interesting how these production shifts relate to production shifts of other months. It may be the case that we are looking at 3 shifts in which the amount of errors or the length of errors significantly deviates from the average over the last year. We recommend to determine how the number of error occurrences and the duration of these occurrences relates to the average over a larger amount of production shifts, e.g. 10 shifts. This prevents the client from considering a deviating period as representative.

Application of the algorithm resulted in a list of errors that are related to downtime. The algorithm only took errors into account that are related to a machine stop during production, and considered the duration of the machine stop. This means that the algorithm filtered out errors that are not relevant in terms of downtime. In order to clarify this, we consider an example in figure 2.1. In the figure (a) 41 errors are shown that are either related or not related to a machine stop. The algorithm determined which errors are related to a machine stop using the machine status, and only included the 18 errors which are related to downtime. The errors related to downtime are shown in figure (b). From the 9409 errors logged in the considered production shifts, the new method finds that 5879 errors are related to a machine stop.



(a) Highlighted messages from the error log over time

(b) Highlighted messages from the error log after applying the error algorithm

Figure 2.1: Visualization of the filtering of the error algorithm

In order to compare the old and the new method, we categorized the errors in errors related to downtime caused upstream, downtime caused at the machine itself and downtime caused downstream. We did this for the production hours of May and June of 2019, in which the machine was up for approximately 93 hours. The overview is visually convenient since it shows the difference in downtime consideration of the two methods. We show the overview in figure 2.2. The overview shows that the logging duration overestimates the downtime duration. The old method found 60.0 hours of downtime caused upstream in comparison to 53.1 hours found by the new method, 36.6 hours of downtime caused at the machine in comparison to 12.7 hours, and 28.1 hours of downtime caused downstream in comparison to 19.3 hours. Using the overview provided by the old method would have caused us to focus on errors with a high logging duration, but a relatively low downtime duration. We decided to use the processed error data from the algorithm to determine the downtime of an error. In the next chapter we further elaborate on the downtime per error caused at the machine.

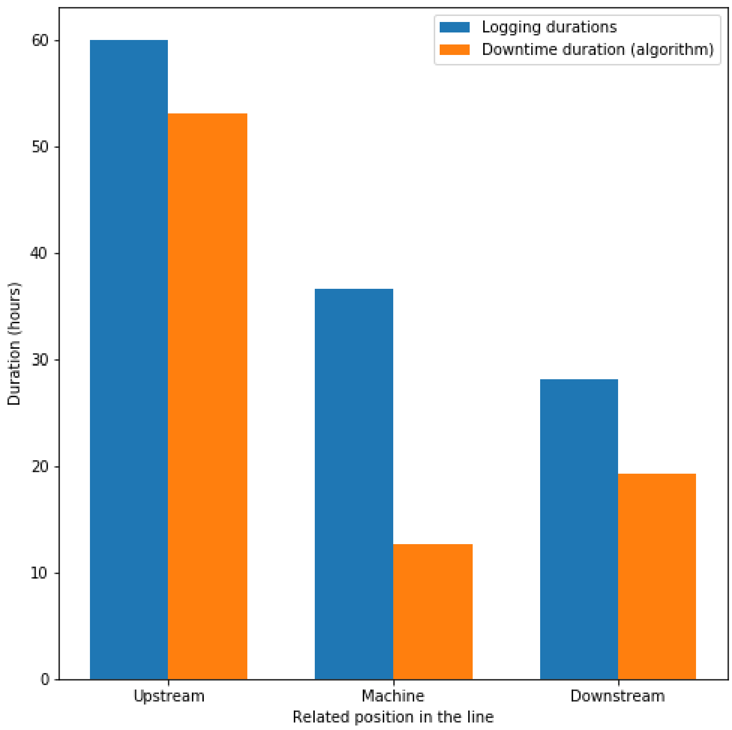


Figure 2.2: Downtime per error according to the old method and the new method categorized in upstream, at the machine and downstream. (Machine 3, May to June 2019, approximately 93 hours of uptime)

2.5 Conclusion

The goal of this chapter was to find a method to determine which errors are related to downtime. First we examined how the client currently processes the error log. We expect that the current way of using the error log incorrectly measures the amount of downtime an error relates to. We decided to develop a new method in order to correctly measure the downtime per error. We determined how the error log and the machine status log can be combined to obtain the machine downtime related to an error. We determined the logic of combining the data sources based on a few key observations of the data sources. Subsequently, we proposed an algorithm that determines the downtime per error. The algorithm consists of several subalgorithms that combine the data from the error log and the machine status log. Lastly, we showed the difference in resulting overview of the downtime of the old and the new method. We found that the old method overestimates the downtime related to errors and we decided to use the new method in the remainder of this thesis.

In this chapter we answered the first research question by developing a method to determine the downtime related to an error:
How can we identify the errors related to downtime?

Chapter 3

Errors related to downtime and their possible causes

In this chapter we discuss which errors are related to downtime. Since our focus is on the packaging machine, we consider the errors related to the packaging machine rather than the errors generated by the upstream or downstream machinery. We show the overview of downtime per error, which is the result of the data processing as described in chapter 2. We are interested in the possible causes of the errors that are related to downtime. A commonly used technique to find the possible causes of a fault is FTA. (Lee, Grosh, Tillman & Lie, 1985). We apply FTA to the three errors that are ranked highest on the list: 401, 185, and 276. We do not have data that clearly provides the frequency that a possible cause results in the error, so we do not make statements about the proportion of downtime related to the possible causes. During the project we used a test machine to verify the possible causes of error 401 and 276. Due to the limited amount of time we could not test the possible causes of error 185.

The data that we used for this chapter is the data of approximately 93 hours of uptime (estimated from the bag counter) coming from machine 3. This data is coming from the production hours from May and June of 2019.

In section 3.1, we discuss which errors are related to downtime and to which amount of downtime they relate to. In section 3.2, we elaborate FTA, which is a method that is used to identify the possible causes of faults. In sections 3.3 to 3.5, we apply FTA to the three errors errors related to most downtime. We identify the possible causes of these errors, but we do not which possible cause has resulted in which part of the downtime. We propose a method to identify possible causes in the data and to determine the proportion of downtime related to a possible cause in section 3.6. In section 3.7 we select one possible cause we will investigate in a case study. We conclude in section 3.8.

3.1 Downtime per error at the machine

The downtime in the line caused by the packaging machine consists of the downtime from manual operator stops and the downtime from errors. Since we do not have data on the underlying reasons for manual operator stops, these are out of scope [1]. We excluded the manual operator stops from the overview of most relevant errors. Subsequently, we ranked the errors in decreasing order based on their related downtime and we separated the downtime per error in the duration of fuzzy error occurrences and the duration of clean error occurrences. We show the resulting overview in figure 3.1. We conclude that the error related to the most downtime, is error 401. We decided to further

[1] Manual operator stops account for 59% of the downtime related to downtime at the machine. However, the client does not keep track of the reasons for these manual stops and we do not further investigate these stops.

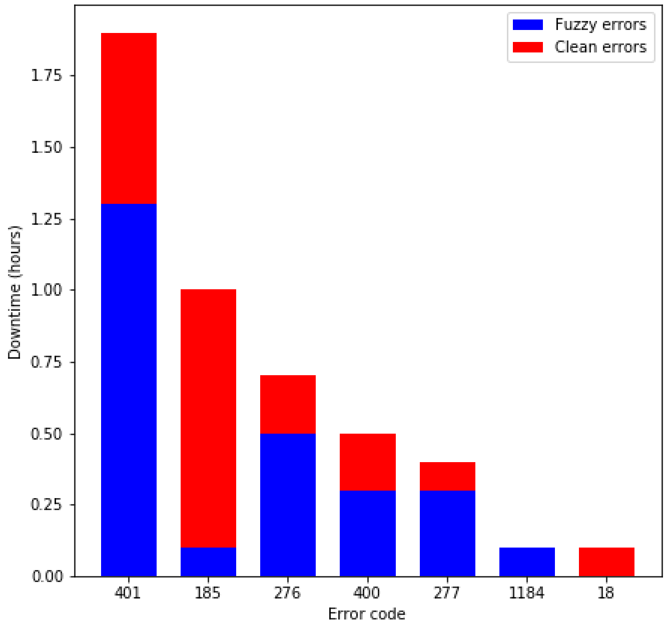


Figure 3.1: Downtime per error at machine 3, (production in May and June 2019, 93 hours of production)

examine the causes of the errors. Due to the limited availability of time, we decided to do this for the three highest ranked errors.

3.2 Fault Tree Analysis (FTA)

In order to determine the possible causes of the errors related to downtime, we will apply FTA. In this section we elaborate on the general steps of applying FTA and the data we used in order to apply each step. The steps we describe are based on Pilot (2002) and Haasl et al. (1981). Pilot (2002) provided a brief explanation of the practical application and Haasl et al. (1981) provided background on the theory and its relation to other methods. Note that we solely focused on the qualitative FTA, since we could not directly determine the frequency that a possible cause results in an error. This means that we used the structured approach of FTA to identify possible causes of a fault, but we did not investigate the probability that a possible cause leads to the error.

- **Step 1**
  In FTA, one aims at finding all possible causes of a fault in a top-down structure. This means that in the first step, one focuses on a specific fault. In our case the fault was a specific error related to downtime. One should avoid focusing on a too general fault, since that will result in an unmanageable fault tree.

- **Step 2**
  The second step is to determine the possible events leading to the fault. Note that in this step we were looking for events that are directly connected to the concerned fault. We know that the occurrence of an error is triggered by the internal algorithms of the machine. These algorithms take sensor data as input and if the sensor data appears to be outside certain

predefined bounds, the error is triggered. We also took into consideration that the sensor may have been malfunctioning, and reporting incorrect measurements that result in the error occurrence. If an event is not further split up in other events, the basic event is used and an intermediate event otherwise. The events are shown in figure 3.2.

- **Step 3**
  In the third step, one connects the events to the fault using AND and OR gates. The And and OR gates are shown in figure 3.2. If two identified reasons have to occur simultaneously in order to result in the fault, the AND operator is used. If each of the reasons in isolation results in the fault, the OR operator is used.

- **Step 4**
  In the fourth step, one determines the events that lead to the events previously found in step 2. In order to find the events for the error to occur (if the sensor is not malfunctioning), we used data from Bosch on the triggers of specific errors. Subsequently, we determined where these triggers appear in the clients' standard procedures using data from the operator trainer and our own observations of the clients' production process.

- **Step 5**
  Again, the events found in step 4, should be connected to the events of step 2 using the OR and AND operators.

- **Following steps**
  One can repeat steps two and three in an iterative manner. One can stop this procedure when a convenient level of detail on the possible causes is obtained. In the following sections, we decide to stop when the branch could be ruled out, or when we established a clear link with the clients' procedures. In the following sections, we elaborate on the resulting Fault Tree for the top three errors from figure 3.1.

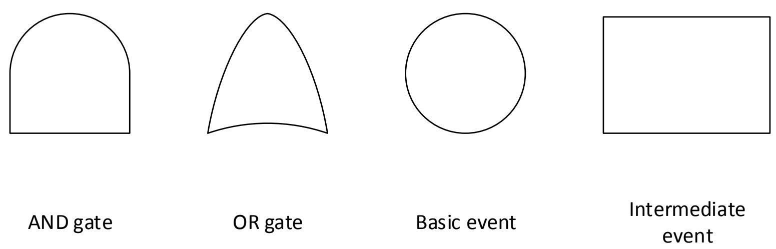


Figure 3.2: Logic-gates and events for FTA construction

3.3 Error 401: Encoder slip or too many marks on film detected

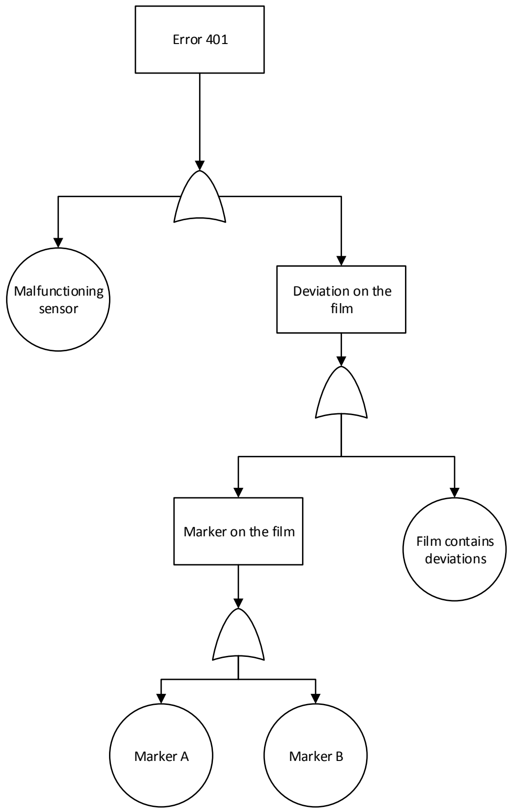


Figure 3.3: Fault analysis tree for error 401

In this section, we examine the possible causes of the error related to the most downtime; error 401. We defined this error as the fault of interest, which fulfilled *step 1*. For each of the steps, we relied on the knowledge of machine experts and the knowledge of the operator trainer from the client.

In *step 2*, we wanted to identify the events that lead to the occurrence of error 401. We know that the error is triggered when the sensor detects a certain process variable to be out of range. We know from personal communication with a machine expert, that the error is triggered if the sensor does not detect the spot in the time that a bag length of film passes the sensor, or if more than one spot passes the sensor in the time a bag length of film passes the sensor. We summarized these two possibilities as the event that the film contains a deviation. Alternatively, sensors may malfunction which may lead to the generation of error 401. We identified a malfunctioning sensor and a deviation on the film as the two direct causes of Error 401.

In *step 3*, we connected the causes from step 2 to error 401 using logic gates. Since each of the events of step 2 can lead to the occurrence of error 401, we used the OR gate.

In *step 4* we identified which events lead to the events from step 2. We know from communication with the line manager that no defects were found during the span of this project, so we expect that the sensor was not structurally malfunctioning. We did not further investigate this branch. We used personal communication with the operator trainer and our observations of the production process to determine the what events can possibly lead to a deviation on the film. We found that the supplier of the film places so called markers on the film. A marker is a line across the width of the film which is deviating in contrast. The machine detects the marker and this results in error 401. We identified these markers as an event that leads to a deviation on the film. Another possibility is that the films contain deviations due to production errors during the production of the films. The line manager stated that no deviations on the film were found during the production period. Since the operators closely examine the quality of bags in order to remove bad quality bags, we assumed that the amount of deviations on the film accounts for a negligible amount of occurrences of error 401 and we did not further examine this event.

In *step 5* we concluded that each of the identified events lead to a deviation on the film and we connected them using the OR gate.

In *step 6* we were looking for the events leading to the events from step 4. We found that the client orders films with markers halfway the film and near the end of the film. Since both markers lead to the event a marker on the film, we connect the events using an OR gate. The client orders films with a marker halfway and at the end of the film, and films with only a marker at the end of the film in order to build in automatic machine stops. We refer to a marker halfway the film as *marker* and a marker at the end of the film as *marker*.

Considering that no defects and no faulty films were found during the span of the project, we consider markers A and B as the most probable causes for error 401.

3.4 Error 185: Label device error

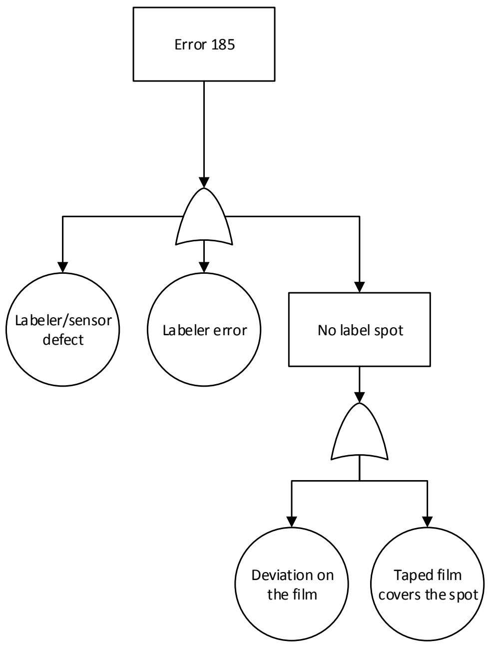


Figure 3.4: Fault analysis tree for error 185

In this section, we elaborate on the possible causes of error 185. By selecting the fault of interest, we fulfilled *step 1*.

In *step 2* we determined the events leading to the error 185. From communication with a machine expert we know that the error is triggered when the sensor located near the labeler does not detect the spot on the film on the expected position or when the labeler itself is generating an error (e.g. no ink, or device stuck). We refer to the first event as no label spot and to the second event labeler error. The appropriate action for a labeler error is described in the labeler manual and we did not further develop this event. Similarly to the previous error, a defective labeler or sensor may also lead to the generation of error 185. However, no defects have been found during the span of the project, so we do not further develop this branch.

In *step 3*, we connected the events to the occurrence of error 185. Since each of the events leads to the occurrence of the error, we used the OR gate.

In *step 4* we identified the events leading to the event that there is no label spot on the film. Based on personal communication with the operator trainer, we found that the label spot may be missing due to a deviation on the film and that the label spot may be covered when two films are taped to each other during a film replacement.

In *step 5*, we connected the events from step 4 to the event no label spot. Since each of the events leads to the occurrence of the error, we used the OR gate.

We consider the labeler error and that taped film covers the spot as the most probable reasons for error 185. Since we have not extensively tested the events leading to this error, the tree may not be conclusive.

3.5 Error 276: Splice tape detected

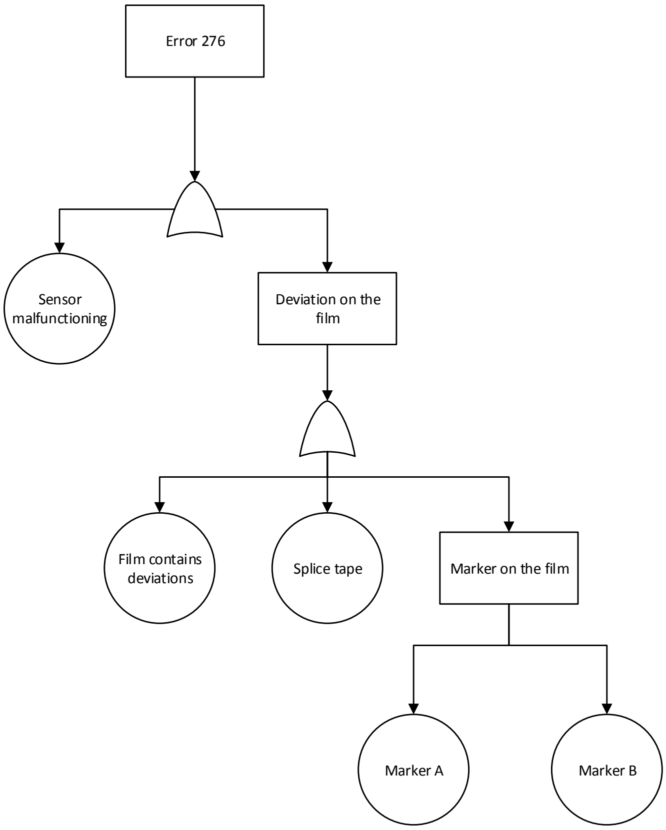


Figure 3.5: Fault analysis tree for error 276

In this section we focus on the events that lead to error 276. Selecting this error fulfilled *step 1.*

In *step 2*, we identified the events leading to error 276. From expert knowledge we know that the error is similar to error 401, only its triggering sensor is aimed at a different part of the incoming film. This part of the film generally does not contain any spots, and when the sensor does detect a difference in contrast, error 276 is generated. We summarized this difference in contrast as the event deviation on the film. It may also be possible that the sensor is malfunctioning. Since no defects were found, we expect that this possible cause is negligible and we do not further develop this branch.

In *step 3*, we connected the events to the occurrence of error 185. Since each of the events leads to the occurrence of the error, we used the OR gate.

In *step 4*, we identified the events leading to a deviation on the film. The same reasoning as for error 401 applies here: the deviations on the film due to production errors and markers on the film lead to the event a deviation on the film. Additionally, this part of the film may also deviate due to the splice tape used for taping two films to each other during a film replacement.

In *step 5*, we connected the events from step 4 to the event deviation on the film. Since each of the events leads to the occurrence of the error, we use the OR gate.

The following steps are already mentioned for the markers of error 401, so we left them out. Using the same reasoning as for error 401, we consider marker A, marker B and the splice tape on the film as the most probable causes for error 276.

3.6 Determining the frequencies of the possible causes

In the previous sections we identified the possible causes of an error. Ideally a company possesses data of the frequencies of the different causes. Unfortunately neither the client nor the machine keeps track of the occurrences of each of the possible causes. Since the error log contains data on a variety of other events during production, we reason that the error log might contain data on the possible causes. In order to identify the possible cause in the data, we need a method. In this section we propose a method that identifies the occurrence of a possible cause in the machine data, which leads to the determination of the frequency that a possible cause leads to an error. The quality of the result of the method still highly depends on the quality and extensiveness of the data available.

Determine the machine interaction
First, we determine which interactions with the machine have a one-on-one relation with the possible cause, i.e. which interaction between the environment and the machine typically occurs during or near the occurrence of the possible cause. For example, if we are trying to determine the frequency that component failure is due to the cleaning of the machine, we identify the cleaning of the machine as the interaction that has a one-on-one relation with the possible cause. Note that the interaction does not have to lead to the fault in all cases.

Determine the steps of the interaction
Subsequently, we distinguish the different steps in the interaction. For example, cleaning the machine may consist of gathering the cleaning materials, opening the doors of the machine, cleaning the machine and closing the doors of the machine.

Determine the indicators of the steps
Next, we determine how each of the steps are reflected the machine data. We refer to such a reflection as an indicator. In the ideal case, each step has a reflection in the data which. In our case, the machine is not designed to reflect interactions, so this is not the case. For the cleaning example, we have no indicators for the gathering of cleaning materials nor for the moments the cleaning starts and stops. However, we can observe the opening of the doors and the closing of the doors.

Determine the minimum required set of indicators
Lastly, we need to determine whether there are other interactions that lead to the same set of or a subset of the same indicators. If there are interactions leading to the same indicators, we need to define which indicators we mark as the interaction of interest. We have to choose between only marking the exact sequence of indicators, which may leave out occurrences of the interaction, or to allow slight deviations, which may result in falsely marking other interactions. In addition, other interactions may occur during the interaction of interest. We have to define whether we mark the indicators as the interaction if there are indicators of different interactions in between.

3.7 Selection of a possible cause

In order to identify the frequency that a possible cause results in an error, we select a possible cause and apply the method described in section 3.6. We manually select the possible cause based on three criteria: the amount of downtime the resulting error relates to, the likelihood that the possible cause causes a significant part of the downtime and the likelihood of being able to identify the possible cause in the data. We decide to select one of the possible causes of error 401, since that error has the highest impact in terms of downtime. We reason that it is very probable that marker B causes a significant amount of downtime of error 401 based on the following knowledge about the production setting. During a production shift, at most three different film types are used (three different bag sizes). Assuming that the operators do not switch back and forth between product types, the maximum amount of changes in product type during a shift is two. Based on the bag counter, we know that at least 34 films have been used in May and June and there were 3 shifts of multiple days in May and June. Based on this, we expect that marker B causes a significant proportion of of the downtime of error 401. Since marker B marks the end of the film, we know that the film has to be replaced. During production, the replacement of the film has a one-on-one relation with the occurrence of marker B. We know from machine experts that several steps during the replacement of the film lead to an indicator in the data. We expect that applying the method from section 3.6 is likely to result in the identification of marker B in the data. In the following chapter, we will apply the method to identifying the film replacement.

3.8 Conclusion

The goal of this chapter was to identify the possible causes of the errors related to downtime. We first elaborated on a method to identify the possible causes: FTA. We explained the steps to undertake and how to construct a Fault Tree. Subsequently, we applied the method on the errors 401, 185 and 276, which form 69% of the error downtime at the machine. We used knowledge of a machine expert to determine the triggering events for the errors and we used data from personal communication with the clients' operator trainer to find the events that lead to the errors during production. Lastly, we proposed a method to determine the frequency a possible cause leads to an error. Since neither the client nor the machine keeps track of the occurrence of the possible causes, we do not have data that directly provides us with the frequencies. We proposed a method that relies on the existence of an interaction with the machine that always happens near or during a possible cause. If an interaction is reflected in the error log, the method may be able to estimate the frequency of the possible cause. We selected the film replacement as a possible cause, to apply the method to.

In this chapter we answered the second research question by determining a method to find the possible causes of an error.
What are the possible causes of the errors related to downtime?

Chapter 4

Case study: The film replacement

In the previous chapter we determined the possible causes of three errors related to downtime: error 401, 185 and 276. In order to determine the frequency that a possible cause results in an error, we proposed a method to identify the possible cause in the data. We selected one of the possible causes for further investigation: marker B. We expect marker B to be an important cause of error 401, and that we can identify the occurrence of marker B in the data. The first step of the method we defined, is to find a machine interaction that occurs one-on-one with the occurrence of the possible cause. Since marker B marks the end of the film, we know that the film has to be replaced when marker B occurs. The film replacement in general does not have a one-on-one relation with marker B, since the film can also be replaced when sufficient product is produced of the current product type. However, if the film is replaced because sufficient product of the current type has been produced, the operator has to change the recipe setting of the machine. This change in recipe is reflected in the error log, which means we can distinguish the situation in which the film is replaced because sufficient product has been produced. We expect the film replacement without an indicator of a change in product type to have a one-on-one relation with marker B. In this chapter, we apply the method proposed in section 3.6 to identify the film replacement in the data. We are also interested in the average duration of the film replacement, so that we can determine the amount of downtime it causes. From now on, when we refer to the film replacement, we refer to the replacement due to reaching marker B, unless we explicitly state differently.

In section 4.1, we distinguish the steps an operator has to undertake to replace the film. In section 4.2, we determine whether these steps are reflected in the machine data, and how these steps are reflected. In section 4.3, we decide on which sequences of indicators we mark as a film replacement and which sequences we leave out. In section 4.4, we discuss which proportion of downtime of error 401 is caused by the film replacement. In section 4.5, we elaborate on the duration of the film replacement. We conclude in section 4.6.

4.1 Steps of the film replacement

In order to identify the steps of replacing the film, we used data from the operator trainer and our observations of the film replacement in practice. We constructed a list of steps starting from the moment the machine is down to the moment the film is replaced and the machine is up. We also determined whether the step has to take place, or that it is possible to fulfill the step in an alternative way and still successfully replace the film. If it is possible to fulfill the step in an alternative way, we marked the step as optional, and if not, we marked the step as required. In table 4.1 we show the identified steps of the film replacement.

Table 4.1: Steps of the film replacement

Step nr.	Process step	Required/Optional	Alternative
1	Positioning spare film reel	Required	None
2	Film reel comes near its end	Required	None
3	Sensor 1 detects marker B	Required	None
4a	Automatic machine stop	Optional	Manual stop
4b	Stopping the machine manually	Optional	Automatic stop
5	Activation of the vacuum bar	Optional	Replacement without vacuum bar
6	Activation of the gripping bar	Required	Is also activated by vacuum bar activation
7	Cutting the old film	Required	None
8	Taping the film with a sticker	Optional	Taping the film without a sticker
9	Pull down the film unwinder	Required	None
10	Start machine	Required	None
11	Sensor 2 detects marker B	Required	None
12	Start machine again	Required	None
13	Sensor 2 detects sticker on the film	Optional	Taped without sticker
14	Start machine	Optional	Only if taped with a sticker

4.2 Indicators in the error log

Several steps during the replacement of the film are reflected in the error log by an error message or in the machine status log by a change in machine status. The film can be replaced because it reaches its end, or sufficient bags have been made of the current bag type. If sufficient bags have been made of the current bag type, the operator will either stop the machine or change to another bag type. Both cases are distinguishable, since the we can observe the machine status and a notification if the operator changes the recipe.

In order to determine the indicator of a step, we used observations of the film replacement in practice and we executed the steps on a testing machine of the same type. In table 4.2 we show the identified steps, whether they are optional or required, which indicators we found and whether there is an alternative step. Due to the fact that some indicators may be absent and other opera-

Table 4.2: Process steps of a film replacement

Step nr.	Process step	Required/Optional	Indicators	Alternative
1	Positioning spare film reel	Required	None	None
2	Film reel comes near its end	Required	164	None
3	Sensor 1 detects marker B	Required	401	None
4a	Automatic machine stop	Optional	48, 276, 401	Manual stop
4b	Stopping the machine manually	Optional	163	Automatic stop
5	Activation of the vacuum bar	Optional	221, 49	Replacement without vacuum bar
6	Activation of the gripping bar	Required	49	Is also activated by vacuum bar activation
7	Cutting the old film	Required	None	None
8	Taping the film with a sticker	Optional	276	Taping the film without a sticker
9	Pull down the film unwinder	Required	None	None
10	Start machine	Required	Machine status	None
11	Sensor 2 detects marker B	Required	276, machine status	None
12	Start machine again	Required	Machine status	None
13	Sensor 2 detects sticker on the film	Optional	276, machine status	Taped without sticker
14	Start machine for production	Optional	Machine status	Only if taped with a sticker

tions than the film replacement can result in a subset of the indicators of the film replacement, we had to be careful not to confuse another operator action with a film replacement. We manually examined the different sequences of indicators of the film replacement in the error log. We explain which sequences we considered as a film replacement and which not in the following section.

4.3 Set of required indicators

In the previous step we identified the indicators in the error log related to a film replacement. We observed different variations of the indicators of the film replacement in the error log. We found that the machine not always raises a warning for coming near the end of the film (error code 164), due to the fact that the settings for this warning are not always configured correctly. Furthermore, we observed that the identified indicators may have indicators of other interactions in between, e.g. the opening of doors.

We decided to mark all sets of indicators as a film replacement that consist of the sequence of error codes 401, 276, 49, 276. We accepted the occurrence of optional indicators of a film replacement in between. However, we excluded the sequences that show indicators of other interactions unrelated to the procedure of replacing the film. Since we are interested in the duration of the film replacement, we would like to obtain film replacements in which no other operations are executed. Therefore, if there are indicators of a mechanical adjustment, have not marked the sequence as a film replacement.

We examined the production periods of May and June 2019 and we selected the film replacements manually. For each of the cases that is in line with the selection criteria, we collected the timestamp that the machine is down, the error code that marked the moment the machine is down, the timestamp at which the operator undertook a film replacement related action and the timestamp the machine was up and running again. Since we purposely left out the ambiguous cases, we expect that we have obtained a slightly biased view of the duration of the film replacement. If an operator faced troubles during the replacement of the film, we may have interpreted the resulting sequence of indicators as an ambiguous case.

4.4 Result: Downtime due to the film replacement

After identifying the film replacements in the data set, we can make statements about what proportion of downtime these film replacements relate to. Since we had to leave out the ambiguous cases, we consider the downtime of these film replacements as a lower bound for the actual amount of downtime.

We found that from the 25 fuzzy error occurrences of code 401, 19 occurrences consist of a film replacement. The sum of durations of these 19 occurrences forms 85% of the total duration of fuzzy error occurrences of 401. We show the proportion of the duration of error 401 that is caused by fuzzy errors and the film replacement in figure 4.1. Furthermore, we observed that marker B is not exactly at the end of the film. In several cases in which error 401 occurred, the operator resumed production without replacing the film. Then after roughly a minute, the operator started replacing the film. We observed that from the 36 clean occurrences 9 occurrences fall into this category. The downtime duration of these errors forms 29% of the total duration of the clean error occurrences of error 401. We show the proportion of the duration of error 401 that is caused by clean errors and the film replacement in figure 4.2. This means the identified film replacements, account for 65% of the total duration of error 401. Finally, we observed one occurrence of error 401 before a film replacement with a change of recipe. This means that this marker could have been marker A or marker B.

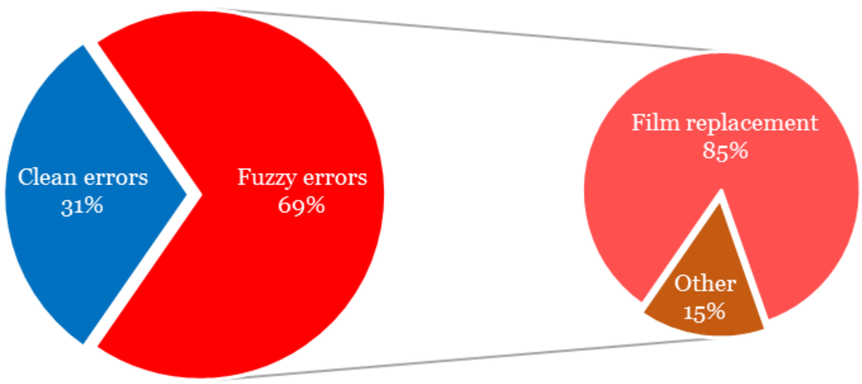


Figure 4.1: Pie charts visualizing the proportion of error 401 that is caused by fuzzy errors and the film replacement

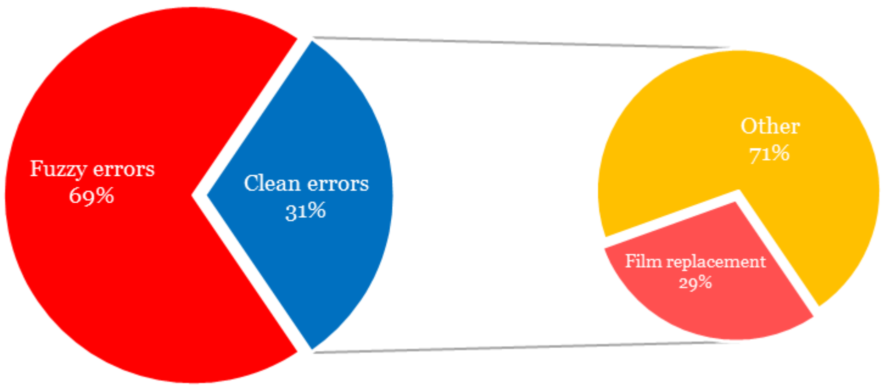


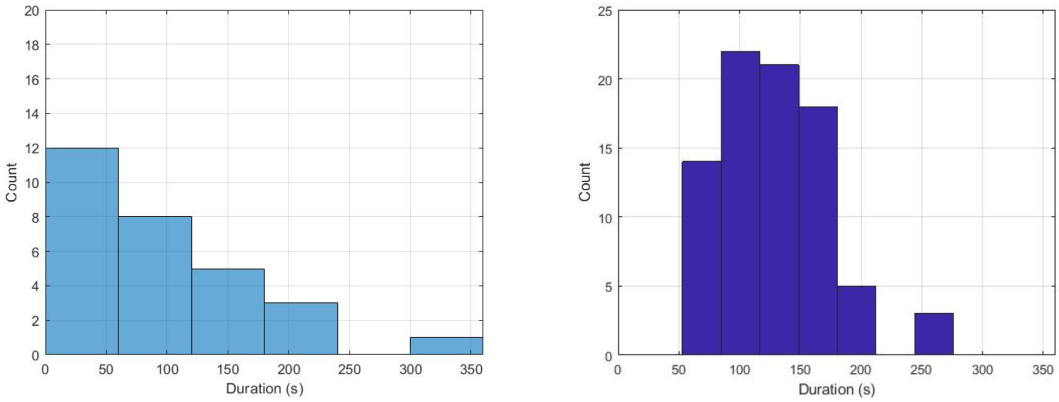
Figure 4.2: Pie charts visualizing the proportion of error 401 that is caused by clean errors and the film replacement

4.5 Duration of the film replacement

The total downtime duration of a film replacement, consists of the operator reaction time and the replacement time. We define the *operator reaction time* as the time between the timestamp the machine is down and the timestamp the operator starts replacing the film. Note that the operator can also stop the machine by starting to replace the film, which means that the operator reaction time is equal to zero. Additionally, we define the *replacement time* as the time between moment the operator starts replacing the film and starts the machine again. We consider the replacement time as the period starting from the timestamp of the first film replacement indicator resulting from the operator until the timestamp that the machine is up and running. The total time consisting of the operator reaction time and the replacement time is referred to as the *total replacement time*.

Since the four machines in the line are of the same type and replacing a film does not differ among the machines, we expect that the underlying distribution of the duration of film replacements does not differ among the machines. In order to obtain a larger data set for the replacement time, we decided to filter out the film replacements on machines 1,2 and 3. This yields a total of 81 data points for the film replacements. From these data points, 32 are coming from machine 1, 21 from machine 2 and 30 from machine 3. We expect that the same reasoning does not apply for the operator reaction time. The operator reaction time may be influenced by the fact that the downtime cost among machines may differ and the position of each machine differs. Since we focus on the downtime of machine 3, we only used the 29 data points coming from machine 3 for the operator reaction time.

The replacement time and the operator reaction time are shown in table 4.3. In figure 4.3 we show the distribution of the operator reaction time and the film replacement duration. During the manual filtering of the film replacements, we noticed that several subsets of film replacement indicators also contained indicators of other operator actions. Our procedure excluded these film replacements. This also means that we did not consider the film replacements in which the operator makes a mistake or a fault occurs during the replacement. We expect that these film replacements typically take longer, so we might underestimate the actual duration. We discuss this more extensively in chapter 7.



(a) The operator reaction time before starting to replace the empty film (n = 29)

(b) The replacement time (n = 83)

Figure 4.3: Bar charts of the operator reaction time and the replacement time

Table 4.3: Replacements time and operator reaction time to film replacements

Variable	Average (s)	Standard deviation (s)	Min (s)	Max (s)
Replacement time of M1, M2 and M3	130.2	46.7	53	276
Operator reaction time M3	88.0	80.6	0	304

4.6 Conclusion

The goal of this chapter is to find the frequency of that marker B on the film leads to error 401 by identifying the film replacement in the error log. First determined which steps a film replacement consists of. We identified the steps using personal communication with the operator trainer and observations of the film replacement in practice. Next, we determined the indicators in the data that reflect the steps of the film replacement. We identified the indicators by using expert knowledge on the machine and repeating the steps on a testing machine. Subsequently, we defined when we identify a set of indicators as a film replacement. We observed that there are several variations of the indicators in the data set and that the indicators may have indicators of other machine interactions in between. Since we are also interested in obtaining the duration of the steps of the film replacement, we left out the sequences of indicators that show indicators of other machine interactions in between. We allow for variations in the indicators of optional steps. Thereafter, we determined the proportion of error 401 occurrences that is caused by marker B and which duration consist of replacing the film. Lastly, we elaborated on the duration of the film replacement. We found that the total duration of a film replacement consists of the operator reaction time and the replacement duration. We showed the durations of both time variables.

In chapters 4 we have taken the first step in answering the third research question by determining how one of the possible causes is related to downtime.
How can we reduce the downtime of the possible causes?

Chapter 5

Opportunistic timing of the film replacement model

In the previous chapter, we found that the film replacement causes at least 65% of the downtime of error 401. In this chapter, we investigate an approach to reduce the downtime of the film replacement. In chapter 2, we observed that 59% of the downtime is caused by errors related to problems upstream and downstream. The main idea of this chapter, is that these unexpected errors may occur just before the film has to be replaced, and one can use the resulting downtime to preventively replace the film. By doing so, we may avoid downtime in two ways: 1. If there is an operator available at the moment an opportunity appears to preventively replace the film, we prevent the machine from reaching the end of the film. If the machine reaches the end of the film, it may be the case that no operator is available. In that case, we have to wait until an operator is available. This is an additional undesirable period of downtime. 2. We may save downtime by partially overlapping the downtime of the film replacement with the downtime of the error. In this chapter we determine which error the client should use and when the client should use this error for the opportunistic film replacement.

In order to find a suitable modeling approach for the decision to use the downtime of an error to replace the film, we examined the aspects that influence this decision. First of all, the moment an unexpected error occurs, there is a usable amount of film left on the film reel. If we decide to replace the film, cost is incurred which is approximately linear in the amount of remaining film. Next, there are several situations in which the machine is down for a period of time. During these periods, downtime cost is incurred. The different periods of downtime we distinguish are: the downtime of reaching marker B, the duration of the errors, and duration of replacing the film during the errors. We used the data to determine the expected duration of these time periods. Additionally, each error has a different frequency of occurrence and therefore a different probability of occurrence. The decision to replace the film should consider the probability that another error occurs before we reach the end of the film. If another error occurs, it is less costly to replace the film at this point due to the lower amount of remaining film. In order to make the model generic we decided that the model should also be applicable to clients with a different amount of errors with different expected durations. Summarizing, the model should consider the following aspects: (i) multiple errors with different durations and frequencies (ii) The cost of disposing the film. (iii) The probability that a later unexpected breakdown may occur which allows to replace the film at a lower cost. (iv) The downtime of reaching marker B, the downtime duration of the errors, the downtime duration of replacing during an error, and the downtime duration of the error. (v) The probability that an operator is not available at the moment an unexpected breakdown occurs.

In maintenance, opportunistic scheduling approaches have been applied in the context of component replacements (e.g. Zhu, 2015; de Win, 2018). However, these models are not directly applicable due to two differences between our situation and the application to component replacement. The models proposed by Zhu (2015) and de Win (2018), use renewal theory to numerically minimize the costs in a cycle divided by the time in a cycle. A cycle is the time from one renewal point to the following renewal point, i.e. the time between to consecutive component replacements. For further explanation on the use of renewal theory in preventive maintenance models, we refer to Arts (2017). The first difference is that these models assume that the time to replace the component is negligible in comparison to the lifetime of the component. In our case, this is assumption is not valid, since the duration of the replacement of the film is an important aspect of the decision to replace the film. The second difference is that the models usually incorporate one type of unexpected breakdown. In our case, we would like to make the model applicable to situations with multiple errors with a different probability of occurrence and a different expected duration. We expect that finding the expected cycle time in the case with multiple errors is complicated and we decide to avoid this by using another modeling approach. Nevertheless, we find that in situations with only one error, one can easily find the expected cycle length.

We decided to model the replacement of the film in a MDP. In the model, we define the states by the amount of remaining bags on the film reel and whether the machine is in production, undergoing a film replacement, down due to an error or undergoing a film replacement while down due to an error. Defining these states, allows us to model a separate expected duration for each of these four time periods. As long as the machine is producing, i.e. it does not face any errors, the machine produces one bag per time step. If the machine reaches the end of the reel, the film reel has to be replaced which takes in expectation a certain amount of time. With a certain probability during production, the machine faces an error in the next time step. Considering the amount of bags on the film reel and the amount of downtime one saves by replacing the film within an error, one can decide to use this opportunity to replace the film. Replacing the film incurs costs for the amount of bags that are still on the film reel that one replaces and costs for the amount of downtime. Alternatively, one can choose to wait until the error is resolved. This only incurs costs for the amount of downtime. If one decides to replace the film, it may be the case that no operator is available. This means that even when we decide to replace the film, there is a certain probability that we cannot replace the film. In that case we wait until the error is resolved to resume production with the remaining amount of film. Given the cost parameters, the duration of replacement, errors and replacement during errors, and the error probabilities, the model determines at which amount of bags one should replace the film when facing a specific error.

First we select an error with suitable characteristics for opportunistically replacing the film in section 5.1. Thereafter, we introduce the model variables and parameters in section 5.2. Then, we elaborate on the assumptions of the model in section 5.3. Thereafter, we introduce the model formulation for $|J|$ independent errors, with J denoting the set of errors, in section 5.4. Subsequently, we discuss the requirements that the structure of the model should meet in order be able to solve it in 5.5. Next, we apply the model in our case study in section 5.6. Lastly, we discuss the conclusion of this chapter in section 5.7.

5.1 Suitable errors to replace the film

In order to perform the film replacement during an error, we need to know which errors have a considerable length and frequency of occurrence. We do not have data on the time an operator needs to react on the occurrence of an error before he starts replacing the film. We expect that if an operator is available at an error occurrence, he only needs a small amount of time to react on the opportunity. Based on the distances in the production area, we expect that the operator can start replacing the film in less than 30 seconds. We decide to only consider errors with a considerable frequency of occurrence (> 100) and that a duration of at least twice the duration of the reaction time (> 60 s). By doing so, we ensure that we consider an opportunity in which we can actually reduce a significant amount of downtime. Another requirement is that the error should not be related to a fault at the machine, since that may require an operator to make mechanical adjustments during which the film cannot be replaced. Errors upstream and downstream, relate to a problem with other machines in the line. These errors cause the machine to stand still, while no other operations on the machine are performed. We decide to determine whether there are errors related to upstream and downstream problems that have a considerable length. We provide an overview of the most relevant errors in figure 5.1. There is one error that is related to upstream problems, error 11, and there are two codes related to downstream problems, 160 and 251.

Error 11 is the most frequent error (5047 occurrences in the data of May and June) and this error causes most downtime (62% of total downtime of May and June). However, the average downtime per occurrence is rather short, with 34 seconds on average and the error occurs constantly over time. We know that the error is caused by the setup of the product supply. The four packaging machines in the line of our machine of interest are connected to the same belt that supplies the input product. This means that if the sum of the speeds of the four machines exceeds the speed of product supply, the machines structurally generate error 11. The duration of the error is influenced by the priority setting for each machine, and since we do not have data on this setting, we consider the investigation of the priority settings out of scope for this project. We conclude that error 11 is not suitable for executing the film replacement since the duration is rather short and occur constantly over time. The periods in which the error has a longer duration, this is probably due to the client setting the machine to a lower priority, which presumably means that the machine performance is less important at that moment.

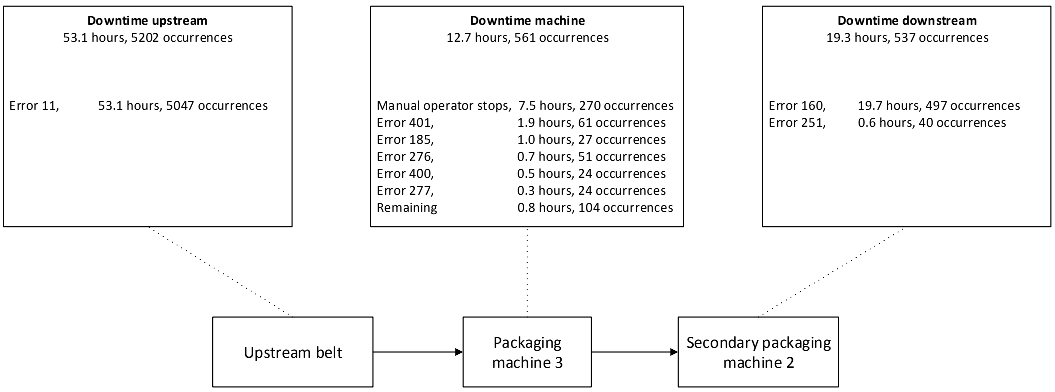


Figure 5.1: Overview of downtime in the line of machine 3, with the corresponding error codes

Error 160 is the most frequently occurring error downstream (497 occurrences in the data of May and June) and has an average duration of 133 seconds. The error is generated when the packaging machine receives a signal from the secondary packaging machine that it faces an error. We expect that it may be valuable to use the occurrence of this error to replace the film. Since the duration of error 251 is < 60 seconds, we do not consider this as an opportunity to replace the film. Therefore, we only consider error 160. The duration of error 160 is plotted in figure 5.2.

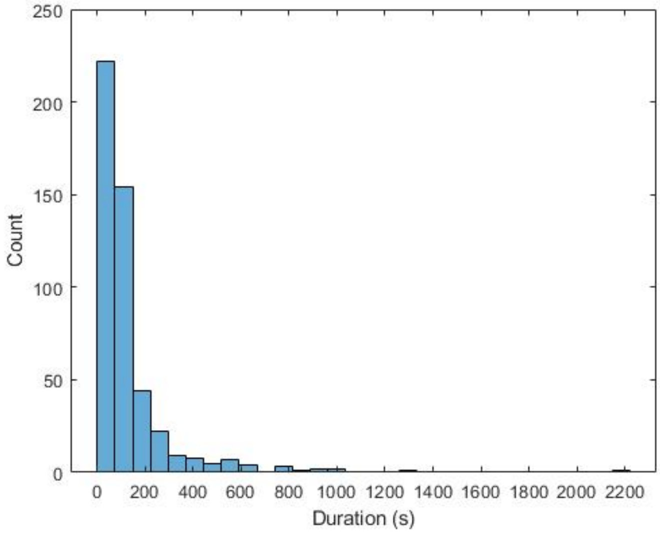


Figure 5.2: Downtime duration due to error 160 (n = 485)

5.2 Model variables and parameters

In this section we present the model variables. Note that we do not provide the transition probabilities. These are provided in subsection 5.1.

λ_j	Arrival rate of error j. With j in the set of errors J.
τ	The discrete time steps at which the model evaluates the system state.
A	The set of possible actions. The set of possible actions depends on the current state s of the system.
$a_{j,i}$	The decision to undertake action i when error j occurs. For each error, an action has to be chosen.
b	Amount of remaining bags.
b_{new}	Amount of bags on a new film.
C_{down}	The cost of downtime expressed in €per hour
C_{film}	The cost of disposing film expressed in €per bag
$C^{\pi*}$	The long-term average cost per time step for the optimal policy π^*
J	Set of independent errors that are considered an opportunity to replace the film
m	The status of the machine
M	Set of possible machine status
P	Set of all transition probabilities in the model. We denote a transition probability from state s to state s' by $P_{s,s'}$. The transition probabilities depend on the current state and possibly the action chosen in the current state.

R	Set of rewards. In this model, the set of rewards consists of costs. These costs depend on the current state and possibly the action chosen in the current state.
P_{Av}	The probability of having an operator available when an unexpected breakdown occurs in which the film can be replaced
s	System state. The state consists of a machine state m and the amount of remaining bags b: (m, b).
S	Set of possible system states.
$T_{Both,j}$	Variable for the time necessary to execute the film replacement during error j.
$T_{ErrorReaction}$	Variable denoting the error reaction time. This is the time period starting from the moment an error occurs which can be used to preventively replace the film, until an available operator has reached the machine and starts replacing the film. The time period consists of the time an available operator needs to observe the opportunity and to walk to the packaging machine.
T_{inter}	Variable for the interarrival time of errors. We assume tha the arrival rate during production is constant, which means the arrival of errors follows a Poisson process.
$T_{OperatorReaction}$	Variable denoting the operator reaction time. The operator reaction time is the time period starting from the moment the packaging machine reaches the end of the film until an operator starts replacing the film. Since operators may be busy at the moment the machine reaches the end of the film, the length of this period varies.
$T_{Recover,j}$	The time necessary to recover from the error. The errors considered are typically errors unrelated to the machine itself and unrelated to the film. The time necessary to recover from the error is the time an operator needs to resolve the underlying fault elsewhere in the line.
$T_{Replace}$	Variable for the replacement time. This variable denotes the time necessary to replace the film. This is the time period starting from the moment the operator is at the machine and starts to replace the film, until the moment the packaging machine resumes production.
$T_{TotalReplace}$	Variable for the total replacement time if the end of the film is reached. This time consists of the operator reaction time and the replacement time. This time period varies since operators may be busy with another task at the moment the end of the film is reached.

5.3 Modeling assumptions

Before we introduce the model, we elaborate on the assumptions we make by using this modeling approach.

5.3.1 Discrete time

The model evaluates the system state at n discrete points in time:

$$t_1,\ t_2,\ \ldots,\ t_{n-1},\ t_n$$

The difference between these points in time, is a constant time step called τ, which is expressed in seconds:

$$t_i - t_{i-1} = \tau$$

In each time step τ, only one transition can take place. We set the time step τ to the time necessary to produce one bag. Note that τ has to be smaller than $\mathbb{E}[T_{Recover,j}]$, $\mathbb{E}[T_{TotalReplace}]$ and $\mathbb{E}[T_{Both,j}]$. In situations with a large amount of bags on a film, setting τ equal to the time to produce one bag may result in a computationally intractable model. In such cases, we suggest to aggregate the states and use steps in which multiple bags are produced.

The assumption of discrete time steps is closely related to the assumption on the distribution of the time variables. If the time variables are exponentially distributed, one can simply calculate the transition probabilities corresponding to the length of the time step τ. However, for other disteributions, the transition probabilities are not scalable to a small time step. This means that in these cases, evaluating the system at discrete time steps is an approximation of reality. We will further discuss the implications for the results in subsection 5.3.3.

5.3.2 Constant error rate during production

The error rate of each of the errors in the set of errors J is assumed to be constant over time during production:

$$\lambda_1 = C_1,\ \lambda_2 = C_2,\ ,...,\ \lambda_{|J|} = C_{|J|}$$

With C_1, ..., C_n being constants with $C_i \in \mathbb{R}$. However, if the machine is not producing, the error rate is assumed to be 0. The error rates are assumed to be independent of the amount of bags on the film and independent of each other. This means that the interarrival times follow an exponential distribution, and the arrival of errors follows a Poisson process. A machine expert from Bosch, who has been involved with the line setup at the client, verifies that the errors from secondary packaging machines are not related to a specific part of the film and that the arrival process of errors appears to be random. In addition, Daley & Vere-Jones (2007) finds that the combination of a large amount of non-Poisson renewal processes, still has Poisson properties. In our case, error 160 consists of failure modes of the secondary packaging machine, and thus may still have Poisson properties. Unfortunately, we are not able to verify the assumption using the data. We do observe based on the timestamps that the occurrences of error 160 do not appear to be related to a specific part of the data set as the error occurs throughout the whole data set. We expect that as long as the error occurrence does not depend on the film replacement, this assumption is realistic.

5.3.3 Exponentially distributed time variables

Inherent to the decision of modeling using a MDP, is the assumption on the durations to be exponentially distributed:

$$T_1 \sim exp(\lambda_1),\ T_2 \sim exp(\lambda_2),\ T_3 \sim exp(\lambda_3), ...$$

In which T_i denote the time variables we consider in the model. We model the durations of replacing a film when it reaches its end, the duration of an error and the time to replace the film during by exponentially distributed variables. In the model, the duration of these variables is modeled as exponential distributions with the expectation set equal to the expectation we found in the data. Another assumption inherent by using a MDP, is that the process is said to be memoryless. This means that the optimal decision only depends on the current state, and not on the history of states.

We examined the distribution of the data for the time variables, and we concluded that they differ from the exponential distribution. In order to determine how the model deviates from reality, we examine the variances of the exponentially distributed time variables and the variance of the data. We find that the variance in the model is larger for the duration of an error and for the replacement of the film. In appendix C, we plot the exponential distribution against the data and we argue that the model underestimates the long-term expected costs per state. This means that the long-term optimal cost per time step found by the model, is too optimistic. We expect that the optimal policy the model provides may still be reasonable, since the model underestimates the long-term average cost per time step for each of the states.

This estimation error raises the need to ensure that the solution of the model performs well if the time variables are not exponentially distributed in reality. We ensure this in chapter 6. Since the real distributions differ from the exponential distribution, the optimal decision may depend on the time that has elapsed in a state. It may be beneficial to wait for a short amount of time in a state before making a decision. Since the model assumes exponentially distributed time variables, the model cannot incorporate this waiting decision. However, a waiting decision would result in impractical policies, which is undesirable. Therefore, we do not mind that the model neglects the elapsed time in a state.

5.3.4 Operator availability

Whenever an error occurs and we would like to replace the film during this error, an operator may either be available or not. Whether an operator is available, depends on whether the operator is busy with other tasks in the production area. For simplicity, we only consider operators available that have no task at the moment error 160 occurs and we do not consider operators available who do have a task, even if they will finish this task soon. If no operator is available, we cannot replace the film during the error. In this case, we wait until the error is resolved and we resume production with the same amount of remaining bags. If there is an available operator, the operator has to walk towards the machine and start replacing the film.

In order to incorporate the operator availability, we use a probability factor P_{Av} that an operator is available at the moment the error occurs. This means that if one decides to replace the film, there is a probability 1-P_{Av} that no operator is available. By incorporating the operator availability in this way, we assume that the probability that an operator is available is independent of the previous state. However, in reality, an operator is not available for a certain period of time. Thus, if an operator was not available at the previous time step, it will probably not be available at this time step. Based on the occurrence of error 160, we estimate the probability of observing two errors at consecutive time steps as $0.0012^2 = 1.44 \cdot 10^-6$. Since this probability is relatively small, we expect that the impact of this assumption on the model is fairly small.

5.4 The Opportunistic Film Replacement Model

A MDP is a 4-tuple: $< S, A, P, R >$. S denotes the set of all possible state of the system. The states of the system are explained in subsection 5.4.1. A is the set of possible actions. The set of actions, may depend on the state. The set of possible actions in state s is denoted by $A(s)$. The set of actions is discussed in subsection 5.4.2. Given the current state and an action, one defines the probabilities of finding the system in state s' in the next time step. The set of transition probabilities is denoted by P. We discuss the transition probabilities of the model in subsection 5.4.4. One allocates rewards to the combination of the current state, the chosen action, and the corresponding transition probability. In our model, we incur costs, thus negative rewards. The cost parameters are discussed in 5.4.5 and the one step expected costs are discussed in subsection 5.4.6. The procedure to find the optimal action in each state is discussed in subsection 5.5.1.

5.4.1 Machine State

The state of the system, s, consists of a machine state m and the amount of bags on the reel, b. The set of machine status M consists of (i) *In production*, (ii) *Replacement of the film*, (iii) *Down due to error j* and (iv) *Replacement of the film during error j with j in J*. J denotes the set of errors that can be used to replace the film. The machine state *In production* is the state in which the machine is producing bags at a constant speed. The machine state *Replacement of the film* is the machine status in which the film has reached its end and has to be replaced. The machine state *Down due to error j*, is the machine status in which an error has occurred, but this error is not used to replace the film. The machine state *Replacement of the film during error j with j in J*, is the machine state in which error j occurred and the opportunity is used to replace the film. Summarizing, the machine state m can be one of the following states:

$$m = \begin{cases} \textit{In production}; & \\ \textit{Replacement of the film}; & \\ \textit{Down due to error } j & j \in J; \\ \textit{Replacement of the film} & j \in J; \\ \textit{during error } j & \end{cases} \tag{5.1}$$

We evaluate the system state at discrete time steps τ, and we set τ equal to the time necessary to produce one bag. During production, the amount of bags on the film reel, b, decreases with one bag per time step until the film is replaced and it is set to the amount of bags on a full film reel b_{new}. That means that the following holds for b.

$$b \in \mathbb{Z} : 0 \leq b \leq b_{new} \tag{5.2}$$

The system state is denoted by s, and the state space is denoted by S. The system state is a combination of the machine state m, and the bags on the film reel b. That means that the state of the system is denoted by:

$$s = (m, b) \tag{5.3}$$

The amount of bags is set to zero if we reach the state *Replacement of the film* or *Replacement of the film during error j*. When the film is replaced, the machine starts producing again so the machine state is *In production*. The amount of bags on the film reel b, is then equal to b_{new}. In other words, the amount of remaining bags only plays a role for the machine states *In production* and *Recovering from error j*.

5.4.2 Actions and decision function

In each production state, the operator has to make a choice. The choice is to either to replace the film if error j occurs in the next time step or to do nothing if error j occurs in the next time step (resume production after the error is solved). Note that the decision alters the transition if an error occurs in the next time step and does not alter the transition in case no error occurs. In each production state, a decision has to be made for each of the errors j in J. The set of possible actions in a state s is denoted by $A(s)$. $\boldsymbol{a}$ denotes the vector of actions with length $|J|$, in which $a_{j,i}$ denotes the decision for error j. Let $D(s,j)$ be a decision function that maps the production state s to an action $a_{j,i}$, for each of the errors $j \in J$.

$$D(s,j) = \begin{cases} a_{j,1}, \textit{ If error } j \textit{ occurs, replace the film} & j \in J, \\ a_{j,2}, \textit{ If error } j \textit{ occurs do nothing} & j \in J. \end{cases} \tag{5.4}$$

Let π be a policy, which is a mapping from all combinations of production states $s \in S$ and errors $j \in J$ to a decision. In the error states, the replacement state and the states replacement during error j, there is no decision to be made.

5.4.3 Duration variables

Since we are using a MDP, we model the duration as exponentially distributed variables. Without making assumptions on the actual distribution of the durations in the real situation, we incorporate the durations by matching the expected duration of the time variable with the expected duration of the corresponding state in the model. The duration of recovering from an error, replacing the film when marker B is reached and replacing the film during an error are denoted by a stochastic random variables $T_{Recover,j}$, $T_{TotalReplace}$ and $T_{Both,j}$, respectively.

$T_{TotalReplace}$ is the total replacement time when the machine reaches marker B. As explained in chapter 4, this time period consists of the operator reaction time, $T_{OperatorReaction}$, and the replacement time, $T_{Replace}$. Since operators may be busy with another task when the end of the film is reached, the expectation of $T_{TotalReplace}$ is typically longer than the expectation of $T_{Both,j}$.

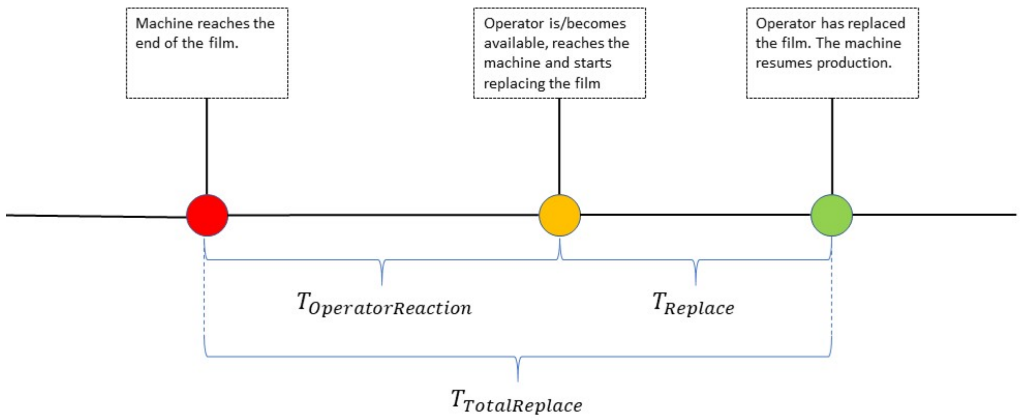


Figure 5.3: The length of $T_{TotalReplace}$

If error j occurs, we may decide to replace the film. We can only replace the film, if an operator is actually available at the moment error j occurs. If an operator is available, the time period the packaging machine is down, is $T_{Both,j}$. The length of $T_{Both,j}$ depends on whether it takes longer to walk to the machine and replace the film (situation 1), or it takes longer to resolve the error 160 (sitation 2). We visualize situation 1 in figure 5.4 and situation 2 in figure 5.5. The duration of $T_{Both,j}$ can be calculated as $T_{Both,j} = max[T_{ErrorReaction} + T_{Replace}, T_{Recover,j}]$. The length of this time period depends on whether it takes longer for the operator to replace the film, or recovering from error j takes longer. The time period an available operator needs to replace the film consists of a small amount of time an operator needs to react on the error, $T_{ErrorReaction}$, and the time necessary to replace the film, $T_{TotalReplace}$.

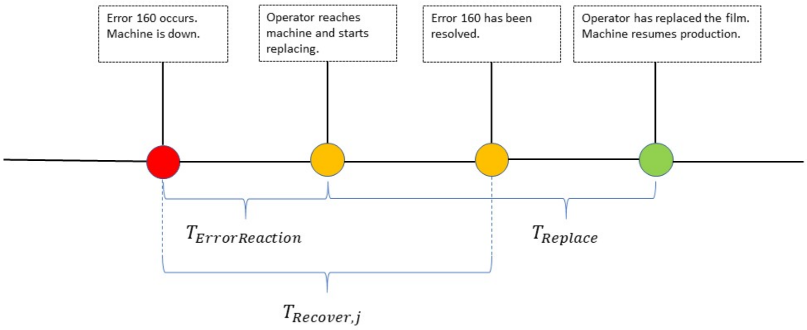


Figure 5.4: The length of $T_{Both,j}$ in the situation 1

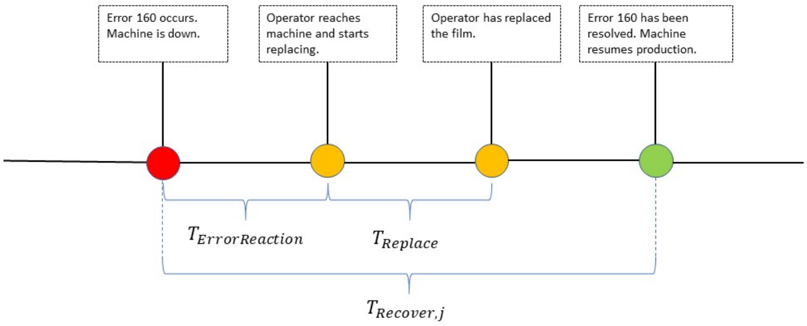


Figure 5.5: The length of $T_{Both,j}$ in situation 2

5.4.4 Transition probabilities

In this subsection we consider the transition probabilities. The transition probability from state s to state s' is the probability that the system state at the current time step is s and after one time step s'. We distinguish four different durations: the duration of producing one bag, the duration of an error, the duration of replacing the film during an error and the duration of replacing the film if the machine reached the end of the film. We set the duration of producing one bag equal to one time step. We incorporate the other durations by setting the probability of leaving the state such that the expected duration in the state corresponds to the expected duration of the corresponding time variable. As explained in the assumptions, τ needs to be smaller than the expectation of $T_{Replace}$, $T_{Recover,j}$ and $T_{Both,j}$. After introducing all transition probabilities, we visualize the one error case in figure 5.6. Note that in the case of one error, we can drop the subscript of the variables that denotes the error.

In production and no error occurs
During production, we model the time until the next error by an exponential distribution. The error rate of error j is denoted by λ_j, which is expressed in errors per time step τ. The probability of being in production state with b bags and going to the next production state with $b-1$ bags (with $b > 1$), is given by the probability that none of the errors occur. If no error occurs during the production of the last bag (b=1), the operator starts replacing the film in the next state. So this transition probability is the probability for going from a production state with $b > 1$ to the next production state $b-1$, and for going from the production state with $b = 1$ to the state with

Replacement of the film. We denote this probability by:

$$P_{NoError} = \prod_{j \in J} [e^{-\lambda_j \cdot \tau}] \tag{5.5}$$

Replacement of the film
When the system is in the state (*Replacement of the film*, *b=0*), the probability of being in state (In production, b_{new}) in the next time step is given by:

$$P_{Replace} = \frac{\tau}{\mathbb{E}[T_{TotalReplace}]} \tag{5.6}$$

And the probability of still being in the replacement state in after one time step is:

$$P_{NoReplace} = 1 - \frac{\tau}{\mathbb{E}[T_{TotalReplace}]} \tag{5.7}$$

In production and an error occurs
During production, different errors can occur. We calculate the probability that we face error j in the next time step by calculating the probability that we face an error in general, times the probability that the type of the error is j:

$$P_{Error,j} = \frac{\lambda_j}{\sum_{i \in J} \lambda_i} * (1 - P_{NoError}), \quad j \in J \tag{5.8}$$

In which λ_j denotes the arrival rate of error j. The transition probabilities at an error occurrence depend on the decision in the production state. Whenever an error j occurs and the decision is *If error j occurs, do nothing*, the system state goes from *(In production, b)* to *(Down due to error j, b-1)*.

If the decision is *If error j occurs, replace the film*, the system goes from *(In production, b)* to *(Replacement of the film during error j, 0)* with the probability that an error occurs times the probability that an operator is available at the start of the next time step:

$$P_{ErrorReplace,j} = P_{Error,j} \cdot P_{Av} \tag{5.9}$$

With probability $1 - P_{Av}$, no operator is available at the error occurrence. Thus, the probability that the system state goes from *(In production, b)* to *(Down due to error j, b-1)*, given that we decided to replace the film if error j occurs, is denoted by:

$$P_{NoErrorReplace,j} = P_{Error,j} \cdot (1 - P_{Av}) \tag{5.10}$$

Down due to error j
Whenever the machine state is *(Down due to error j, b)* with $b > 0$, the system spends in expectation $\mathbb{E}[T_{recover,j}]$ in this state before the system recovers from the error. That means that the system goes to the state (*In production*, b) with probability:

$$P_{Recover,j} = \frac{\tau}{\mathbb{E}[T_{Recover,j}]} \tag{5.11}$$

And stays in the error state with probability:

$$P_{NoRecover,j} = 1 - P_{Recover,j} = 1 - \frac{\tau}{\mathbb{E}[T_{Recover,j}]} \tag{5.12}$$

If m is down due to error j and $b = 0$, the film has to be replaced anyway, so the system state goes to *Replacement of the film* and $b = 0$ with probability $P_{Recover,j}$.

Replacement of the film during error j
If one decides to replace the film during the error occurrence, the system reaches the state *(Replacement of the film during error j, 0)*. The expected amount of time that we will spend in this state is given by $\mathbb{E}[T_{Both,j}]$. That means that the probability of reaching (*In production*, b_{new}) in the next time step is given by:

$$P_{Both,j} = \frac{\tau}{\mathbb{E}[T_{Both,j}]} \tag{5.13}$$

And the probability that the system stays in the state:

$$P_{NoBoth,j} = 1 - P_{Both,j} = 1 - \frac{\tau}{\mathbb{E}[T_{Both,j}]} \tag{5.14}$$

Table 5.1: Transition probabilities

s	s'	b	$a_{j,i}$	$P(s,s')$
(In production, b)	*(In production, b-1)*	$b > 0$	$a_{j,1}$ / $a_{j,2}$	$P_{NoError}$
(In production, b)	*(Replacement of the film, 0)*	$b = 1$	$a_{j,1}$ / $a_{j,2}$	$P_{NoError}$
(In production, b)	*(Down due to error j, b-1)*	$b > 0$	$a_{j,2}$	$P_{Error,j}$
(In production, b)	*(Down due to error j, b-1)*	$b > 0$	$a_{j,1}$	$P_{NoErrorReplace,j}$
(In production, b)	*(Replacement of the film during error j, 0)*	$b > 0$	$a_{j,1}$	$P_{ErrorReplace,j}$
(Replacement of the film, 0)	*(In production, b_{new})*	-	-	$P_{Replace}$
(Replacement of the film, 0	*(Replacement of the film, 0)*	-	-	$P_{NoReplace}$
(Down due to error j, b)	*(In production, b)*	$b > 0$	-	$P_{Recover,j}$
(Down due to error j, b)	*(Replacement the film, 0)*	$b = 0$	-	$P_{Recover,j}$
(Down due to error j, b)	*(Down due to error j, b)*	$b \geq 0$	-	$P_{NoRecover,j}$
(Replacement of the film during error j, 0)	*(In production, b_{new})*	-	-	$P_{Both,j}$
(Replacement of the film during error j, 0)	*(Replacement of the film during error j, 0)*	-	-	$P_{NoBoth,j}$

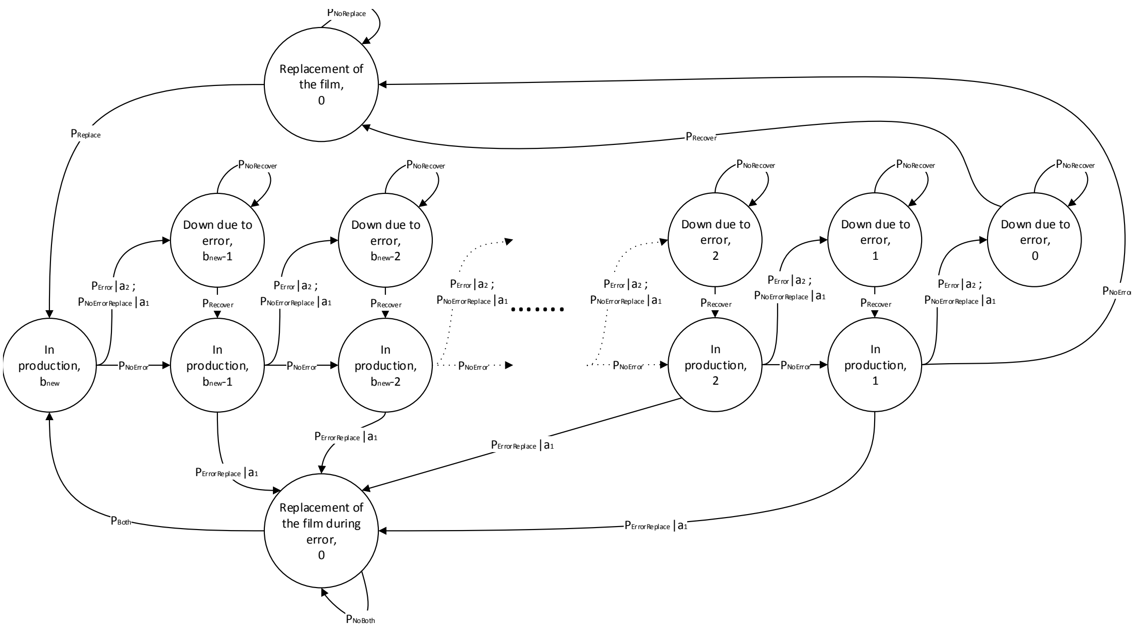


Figure 5.6: Visualization of the Opportunistic Film Replacement Model for the 1 error case

5.4.5 Cost parameters

The model considers two cost parameters: the cost of downtime, C_{down}, and the cost for film, C_{film}. The downtime cost is the cost one pays per time step being in a state in which one cannot produce. The cost for the film is the cost one pays per remaining bag on the reel, when one decides to preventively replace the film. First we define the one step expected costs in subsection 5.4.6 and then the long run average costs in 5.4.7.

5.4.6 One step expected costs

The one step expected costs consist of the costs that are incurred between the current time step and the next time step. In the states in which the machine is not producing, the costs incurred is C_{down}. In the production states, the one step expected cost depends on the decision made. If one decides to replace the film if error j occurs, the one step expected cost consist of the probability that an error occurs, times the probability of having an operator available, times the cost of disposing the remaining film. In each of the production states we make a decision for each error j in J, so we have a direct cost term for all errors in J. We denote the direct cost of the decision for error j as $c_j(a)$. The total one step expected costs in a production state is then:

$$c_{\boldsymbol{a}}(s) = \sum_{j \in J} c_j(s,a) \tag{5.15}$$

In which $c_j(s,a)$ is denoted by:

$$c_j(s,a) = \begin{cases} P_{Av} \cdot P_{Error,j} \cdot (b-1) \cdot C_{film} & \text{if } a_{j,i} = a_{j,1} \\ 0 & \text{if } a_{j,i} = a_{j,2} \end{cases} \tag{5.16}$$

In the states in which there is no production, the direct cost is equal to C_{down}.

5.4.7 Long-run average costs

Let $V_n(s)$ denote the minimum average cost per time unit when in state s with n time steps left on the production horizon. We know from Tijms (2003) (page 259) that:

$$V_n(s) = \min_{a \in A} c_a(s) + \sum_{s' \in S} P_{s,s'}(a) V_{n-1}(s') \tag{5.17}$$

In which $P_{s,s'}$ is the transition probability of going from state s to s'. We are looking for the policy π, that minimizes the long term average cost. We denote the optimal policy by π^*. Then the optimal long term average cost can be denoted by:

$$C^{\pi*} = \lim_{n \to \infty} \frac{V_n^{\pi*}(s)}{n} \tag{5.18}$$

5.5 Requirements of the MDP

In this section we introduce some definitions from Markov Chain theory, that we need to determine that we can find the long-term (undiscounted) cost per time step of our model. Tijms (2003) proposes an algorithm that converges to the long-term (undiscounted) cost per time step, if the model is unichain. We first introduce five definitions are necessary to understand the unichain definition. We further elaborate on the optimal policy in section 5.4.

Definition 1. *Two states are said to communicate under a probability π if there is a positive probability of reaching state each state from the other with a positive number of transitions.*

Definition 2. *A state is recurrent if the probability of re-entering that state is 1 in the long-term.*

Definition 3. *If a recurrent state x communicates with another state y, then y has to be recurrent also.*

Definition 4. *An ergodic or recurrent class of states is a set of recurrent states that all communicate with each other, and do not communicate with any state outside the this class.*

Definition 5. *A non-recurrent state is called transient, since at some finite point in time the state will never be visited again.*

Definition 6. *A Markov Decision Process (MDP) is unichain if the transition matrix corresponding to every policy contains a single recurrent class, and a (possibly empty) set of transient states.*

Now we show that our model is a unichain MDP. We observe that the production states communicate with each other and with the replacement state. These states will be reentered in the long-term, so these states are part of a recurrent class. Error states can either communicate with a production state or not communicate at all, depending on P_{Av} and the optimal decision in the production state that lies before the error state. The set of error states that are communicating with a production state, are therefore also part of the recurrent class. Error states that are not communicating, are not reentered in the long-term and are therefore transient states. Lastly, if it is at some point optimal to decide to replace the film during error j, then the state replacement during error j is also communicating with a production state and thus part of the recurrent class. If it is never optimal to replace the film during an error j, then the state replacement during error j is a transient state.

To conclude, our model consists of a recurrent class and a set of transient states. The recurrent class consists of: production states, the replacement state, a possibly empty set of error states, and possibly the replacement during an error state. The set of transient states consists of a possibly empty set of error states and possibly the replacement during an error states. Therefore we can conclude, based on definition 6, that our model is unichain.

5.5.1 Optimal policy

We are interested in minimizing the long term average cost per time unit. Tijms (2003) proposes an algorithm that converges to the optimal solution for unichain MDPs. In the previous section, we showed that the stronger unichain assumption holds, which implies that also the weaker case of the unichain assumption is satisfied. Since we have an unichain MDP, we know that the value iteration algorithm converges to the optimal average long term cost. The algorithm starts with $V_0(s) = 0$, and then recursively computes $V_n(i)$ for n = 1, 2,... until an arbitrarily large n. The stopping criterion can be defined as using the percentage we are allowed to deviate from the minimal average costs. The expression for $V_n(i)$ is given by:

$$V_n(s) = \min_{\forall a \in A(s)} c_a(s) + \sum_{j \in J} p_{s,s'}(a) V_{n-1}(s'), \quad s, s' \in S \tag{5.19}$$

We implemented the value iteration algorithm in Matlab 2019 for the one error and the two error case. The use of the script is explained in appendix H. We show a numerical example for two errors in appendix E.

5.6 Case study: Opportunistic Film Replacement

In this section we apply the model to the situation of the client. In section 5.1, we selected error 160 as a suitable error to replace the film, so we consider this error as the opportunity to preventively replace the film. Since we are considering error 160, we denote the variables with a subscript for the error with the subscript 160. We first fit a distribution to the data in order to find a theoretical distribution that describes the data. We use the theoretical distributions to construct a distribution for the time of a replacement during error 160 ($T_{Both,160}$) and find the expected duration. Then we elaborate on the parameter values and we summarize the expected durations we found. Using the expected durations, we calculate the transition probabilities. Subsequently, we show the results and we conduct a sensitivity analysis.

5.6.1 Distribution of the duration variables

The model incorporates the durations of the errors, the film replacement and the minimum of the errors and the film replacements. In order to get a closed form expression for all of the variables, we fit a distribution to the data. In chapter 4, we discussed the data of the duration variables and observed implications for their distributions:

- The distributions are strictly non-negative
- The film replacements consist of the reaction time of the operator and the time to replace the film (chapter 4).
- We have very few data points for the operator reaction time (< 30)
- The replacement time has a clear minimum duration, i.e. we should consider distributions that have a very low probability to take on the values lower than this minimum.

We take these implications into account in the fitting procedure. In literature the most common methods to fit a distribution are the Ordinary Least Squares method and the MLE method. Although we do not find clear evidence of one method outperforming the other, we decide to apply maximum likelihood estimation to fit several candidate distributions. The distributions that are commonly used in literature and comply with our mentioned non-negativity requirement commonly used in literature are: the Weibull distribution, the Lognormal distribution and the uniform distribution.

Total replacement duration $T_{TotalReplace}$
The total replacement duration in the model, $T_{TotalReplace}$, is the sum of $T_{Replace}$ and $T_{OperatorReaction}$. We assume that these variables are independently distributed. This is a reasonable assumption considering the fact that the durations of these variables depend on different factors. The operator reaction time mainly depends on whether the operator is busy with another task at the moment the machine reaches the end of the film and his position in the production area, and the replacement time consists of the execution of a fixed set of actions.

Since we have very few data points for the operator reaction time (< 30), we expect that fitting a distribution using MLE has a high risk of finding a coincidental fit. We do observe the following properties of the reaction time: (1) the mean is very close to the standard deviation, (2) plotting the data in a histogram shows that higher durations have a lower probability of occurring, which rules out a uniform distribution. After visual examination, we decide that an exponential distribution has a visually reasonable fit. We use the method of moments to fit an exponential distribution. The exponential distribution has one parameter, so we can estimate the distribution using the first moment, which is the sample mean $\overline{X}$. From chapter 4 we know that the $\overline{X}$ is equal to 88.0 s. We show the histogram of the data with exponential distribution in figure C.1. Since the replacement duration has a minimum duration of 53 seconds, we decide to fit a model with a location parameter. A location parameter shifts the probability density function to the right. The uniform distribution has two parameters that naturally define the location of the distribution and for this distribution the procedure does not differ. For the Weibull and the Lognormal distribution, the addition of a location parameter increases the complexity of finding the MLEs of the parameters, since we now have to estimate three parameters instead of two. For the sake of readability, we do not elaborate on the details of the procedure used to find the parameters. The general procedure of finding the MLEs is described in appendix F and the detailed descriptions of the procedure is described in appendix G.

For the Weibull distribution we try a location parameter using the standard estimation of taking

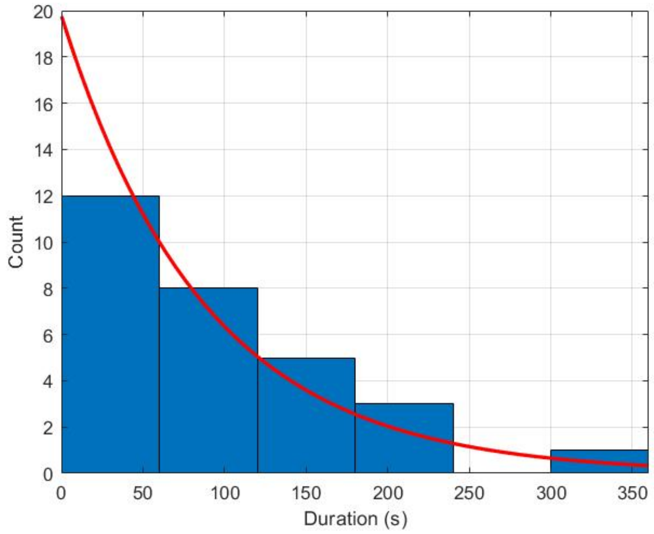


Figure 5.7: Fitted theoretical distribution for $T_{OperatorReaction}$ (n = 29)

the minimum of the data and we try the estimation suggested by Shifley & Lentz (1985). The resulting values are 52.95 and 52.60, respectively. Subsequently, we apply MLE to fit a two parameter Weibull to the data after we have shifted the data points to the left by the estimation of the location parameter. For the Lognormal distribution we follow the practical procedure suggested by Cohen & Whitten (1980) and we observe that this does not provide a solution. We find that the likelihood function is increasing as the location parameter approaches the minimum of the data and we set the location parameter equal to this minimum (= 52.95).

In order to measure the goodness of fit, there are numerous test statistics available. A commonly used statistic to measure goodness of fit is the Pearson χ^2 statistic. However, this statistic requires the user to define the bins of the data, and this choice influences the statistic. Aslan & Zech (2002) suggested to use the Kolmogorov-Smirnov (K-S) test to avoid the binning decision and we decided to use the K-S test. Based on the value of the statistic, we conclude that the Weibull distribution with the parameters $\hat{\gamma} = 52.60$, $\hat{\eta} = 86.45$, $\hat{\beta} = 1.67$ has the best fit. We plot the histogram with the data and the theoretical distribution in figure 5.8. The Probability Density Function (PDF) of the 3-parameter Weibull distribution is:

$$f(x;\eta;\beta;\gamma) = \frac{\beta}{\eta}(\frac{x-\gamma}{\eta})^{\beta-1}e^{-(\frac{x-\gamma}{\eta})^{\beta}} \tag{5.20}$$

Table 5.2: Results of the MLE procedure for the duration of the film replacement

Weibull		**Lognormal**	**Uniform**
$\hat{\gamma} = 52.95$	$\hat{\gamma} = 52.60$	$\hat{\gamma} = 52.95$	$\hat{a} = 276$
$\hat{\eta} = 85.52$	$\hat{\eta} = 86.45$	$\hat{\mu} = 4.07$	$\hat{b} = 53$
$\hat{\beta} = 1.62$	$\hat{\beta} = 1.67$	$\hat{\sigma} = 1.04$	
$p = 0.51$	$p = 0.59$	$p = 0.06$	$p = 3.8 \cdot 10^{-26}$

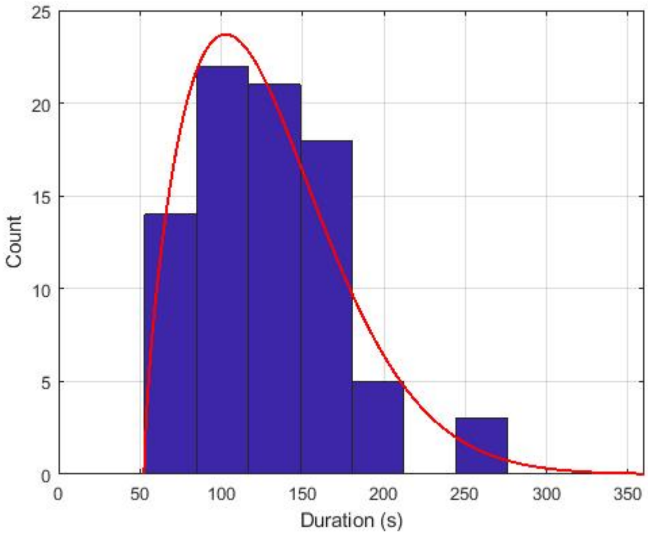


Figure 5.8: Fitted theoretical distribution for $T_{Replace}$

Downtime duration of error 160, $T_{Recover,160}$

The duration of error 160 does not need a location parameter, as the minimum duration is close to zero. We attempted to fit the 2-parameter Weibull distribution, the 2-parameter Lognormal distribution and the Uniform distribution. However, the fitting procedure did not result in an acceptable fit. Visualizing the data in a boxplot, shows that the data contains outliers, which are known to diminish model fit (Yuan & Bentler, 2001). We show the data set in a boxplot in figure 5.9. The boxplot shows 88 datapoints as outliers, which means that the value of these data points deviates a lot from the other data points. We examined the deviating data points more closely. We observed that the ten data points with the highest durations are either coming from the same morning shift or occur at the end of the production. We deleted these data points as we expect them to be related to an unusual problem. The other 78 outliers do not show clear evidence of an unusual event. We decides to delete these data points in order to obtain a better model fit, but we acknowledge that these cases need more thorough investigation. Since we leave out multiple data points with a long duration, we expect that we slightly underestimate of the error duration.

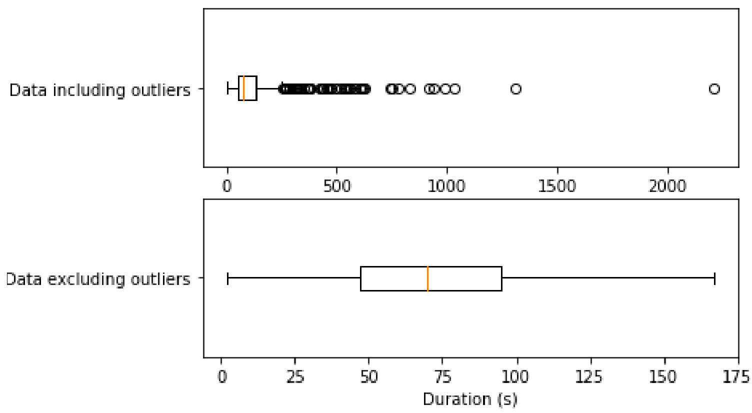


Figure 5.9: Boxplots of the error durations of error code 160, including and excluding outliers

Subsequently, we apply the fitting procedure for the 2-parameter Weibull distribution, the 2-parameter Lognormal distribution and the Uniform distribution. The resulting parameters with their corresponding p-value is shown in table 5.3 We conclude that the Weibull distribution has the best fit. The histogram of the data and the theoretical distribution is shown in 5.10. The PDF of the 2-parameter Weibull distribution is:

$$f(x;\eta;\beta) = \frac{\beta}{\eta}(\frac{x}{\eta})^{\beta-1}e^{-(\frac{x}{\eta})^{\beta}} \tag{5.21}$$

Table 5.3: Results of the MLE procedure for the duration of error 160

Weibull	**Lognormal**	**Uniform**
$\hat{\eta} = 81.56$	$\hat{\mu} = 4.12$	$\hat{a} = 1$
$\hat{\beta} = 2.05$	$\hat{\sigma} = 0.65$	$\hat{b} = 166$
$p = 0.52$	$p = 3 \cdot 10^{-3}$	$p = 9.1 \cdot 10^{-14}$

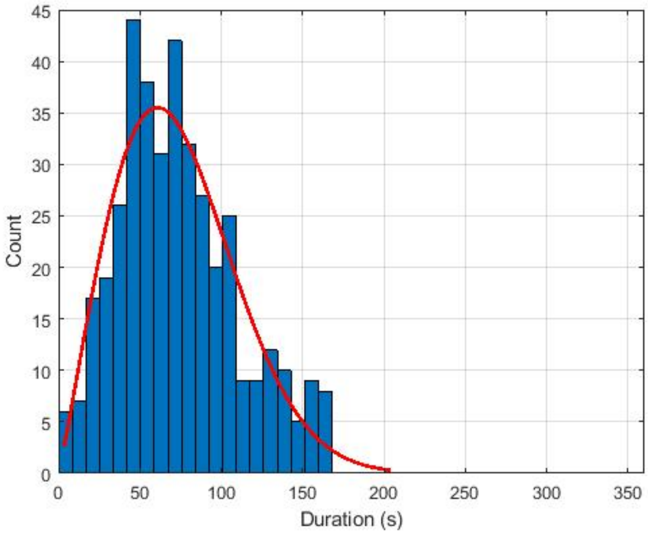


Figure 5.10: Fitted theoretical distribution for $T_{Recover,160}$

Error 160 arrival rate, λ_{160}
As explained in section 5.3, we assume that the error interarrival times of error 160 are independent and follow an exponential distribution. During the process of fitting a theoretical distribution to the duration of error 160, we excluded 89 data points from the data set. We do not consider these data points for the calculation of the interarrival time. That means that we consider 396 arrivals in 93 hours of production only considering the up time of the machine. We calculated the arrival rate λ_{160} as the amount of arrivals divided by the production time. This means we set λ_{160} equal to 4.26 arrivals per hour. We evaluated the model in time steps of 1 second, so λ_{160} is equal to 0.0012 arrivals per second.

Duration of replacing the film during error 160, $T_{Both,160}$
When we decide to replace the film, and an operator is available at the moment error 160 occurs, the time the machine is down depends on $T_{ErrorReaction}$, $T_{Replace}$ and $T_{Recover,160}$. The duration of T_{Both} is the maximum of $T_{Recover,160}$ and $T_{ErrorReaction} + T_{Replace}$. Since we did not have data on $T_{ErrorReaction}$, we had to make an assumption. Based on the distances in the production area,

we assumed that $T_{ErrorReaction}$ is at most 30 seconds. In order to get a pessimistic estimation, we set $T_{ErrorReaction}$ equal to 30 seconds. There is a probability that when the operator reaches the machine, the error is already resolved. This means that part of the errors with a duration less than 30 seconds can not be used to start the film replacement. However, we ignored this possibility and we calculated the duration of T_{Both} as if the film is also replaced in these cases. By ignoring the fact that we can have a longer error reaction time than the error duration, we slightly overestimated the duration of T_{Both}. Note that the probability of missing an error due to a long error reaction time can be incorporated in the value for the operator availability P_{Av} by lowering the value for P_{Av}.

5.6.2 Parameter values

Now that we have approximations of the underlying distributions of the data, we can calculate the parameters of the model. We set the time steps to one second, since the average machine speed is close to 60 bags per minute over the production period. This means that $\tau = 1$ s.

The cost parameters are C_{down} and C_{film}. The film cost of the remaining film is based on the cost of a new film and divided by the amount of remaining bags on the film reel. For the larger bag type, which is usually produced at machine 3, the film consists of roughly 6700 bags. We set b_{new} equal to 6700. A new film costs roughly €400, so we set C_{film} equal to €0.06 €/bag. The cost of downtime depends on whether downtime results in unmet demand or there is still some slack in the production schedule. Since we do not have data of the amount of slack in a production schedule, we estimate the costs for three scenarios: 1. There is still slack in the production schedule and downtime does not immediately result in unmet demand. The cost of downtime is in this case based on the cost of personnel. We estimate that the cost of downtime is €1000 per hour. 2. In the second scenario there is no slack left, but unmet demand can still be prevented by continuing production after the scheduled time. We set this downtime cost equal to €1500 per hour. 3. In the third scenario, downtime results in unmet demand and we set this cost equal to €3000 per hour.

Table 5.4: Parameter and variable values of the case study

Variable/Parameter	Value	Unit
τ	1	seconds
b_{new}	6700	bags
C_{down}	{1000, 1500, 3000}	euro/hour
C_{film}	0.06	euro/bag
λ_{160}	4.26	errors/hour
$\mathbb{E}[T_{Recover,160}]$	72.3	seconds
$\mathbb{E}[T_{TotalReplace}]$	217.8	seconds
$\mathbb{E}[T_{Both,1}]$	161.4	seconds
P_{Av}	{0.1, 0.2, ..., 1.0}	

5.6.3 Calculation of the transition probabilities

We calculated the transition probabilities using the expectations of the distributions we fitted and the expressions introduced in section 5.4.4. The calculation is shown below. We summarize the resulting transition probabilities in table 5.5.

First we calculated the probability of error occurrence in a time step. We found that λ_{160} is equal to 0.0012 arrivals per second so $P_{NoError}$ is equal to:

$$P_{NoError} = 1 - e^{-\lambda_{160}} = 1 - e^{-0.0012} = 0.9988$$

This probability gives us two transition probabilities. Firstly the transition probability of being in *(In production, b)*, with b¿1, and going to the next production state (independent of the choice made) *(In production, b-1)* and secondly the transition probability of being in *(In production, b)*, with b=1, and going to *(Replacement of the film, 0)* (independent of the choice made). We calculate the probability that error 160 occurs in a time step as:

$$P_{Error,160} = 1 - (1 - e^{-\lambda_{160}}) = 1 - P_{NoError} = 0.0012$$

Now we can determine the transition probability of going from the state *(In production, b)* with b¿1, and given that one chooses to replace the film if error 160 occurs, to *(Replacement of the film during error 160, 0)*. As mentioned, this transition probability depends on the operator availability which we will vary. We defined the transition probability as a function of P_{Av}:

$$P_{ErrorReplace,160} = P_{Error,160} \cdot P_{Av} = 0.0012 P_{Av}$$

The transition probability of going from the state *(In production, b)* with $b > 1$ and given that one chooses to replace the film if error 160 occurs, to *(Down due to error 160, 0)* then is:

$$P_{Error,160} \cdot (1 - P_{Av}) = 0.0012 \cdot (1 - P_{Av})$$

On the other hand, one can choose to do nothing if error 160 occurs in state *(In production, b)* with $b > 0$. The transition probability of going from the state *(In production, b)* to *(Down due to error 160, b-1)*, given that one chooses to do nothing if error 160 occurs and that $b > 0$, is equal to $P_{Error,160}$.
Whenever the system is in the state down due to error 160, the system spends on average $\mathbb{E}[T_{Recover,j}]$ in this state. We derive this expectation from the fitted distribution of the error duration and find that the mean of this distribution is 72.3 s. Then we calculate the transition probability of leaving the state *(Down due to error 160, b)* in the next time step. This is the transition probability of going from *(Down due to error 160, b)* with $b > 0$ to *(In production, b)* and the transition probability of going from *(Down due to error 160, 0)* to *(Replacement of the film, 0)*. The probability is calculated as follows:

$$P_{Recover,160} = \frac{\tau}{\mathbb{E}[T_{Recover,j}]} = \frac{1}{72.3} = 0.0138$$

The probability that the system state stays in *(Down due to error 160, b)* in the next time step is equal to 1 - $P_{Recover,160}$:

$$P_{Recover,160} = 1 - \frac{\tau}{\mathbb{E}[T_{Recover,j}]} = 1 - \frac{1}{72.3} = 0.9862$$

Similarly, when the system is in the state *(Replacement of the film, 0)*, the transition probability to *(In production, b_{new})* is calculated by the sum of the reaction time we found in the data and the replacement duration. We assume that these variables are independently distributed. This means that the expected duration of replacing the film including the reaction time is 88.0 + 129.8 = 217.8 seconds.

$$P_{Replace} = \frac{\tau}{\mathbb{E}[T_{TotalReplace}]} = \frac{1}{217.8} = 0.0046$$

The probability that the system state stays in *(Replacement of the film, 0)* in the next time step is equal to 1 - $P_{Recover,160}$:

$$P_{Replace} = 1 - \frac{\tau}{\mathbb{E}[T_{TotalReplace}]} = 1 - 0.0046 = 0.9954$$

Lastly, when the system state is in the state *(Replacing the film during error 160, 0)* the expected time spend in this state is the expectation of the maximum of the replacement time plus 30 seconds, and the error duration. Since we fitted Weibull distributions to the $T_{TotalReplace}$ and the

$T_{Recover,j}$, we have to take the expectation of the sum of two Weibull distributions. We assume these distributions to be independent. Then we calculate the expectation in Matlab using draws from Weibull distributed random number generators with the parameters we found in subsection 5.6.1. We find that the expectation is equal to 161.4 seconds. Thus the probability of going from *(Replacement of the film during error 160, 0)* to *(In production, b_{new})* is:

$$P_{Both,160} = \frac{\tau}{\mathbb{E}[T_{Both,i}]} = \frac{1}{161.4} = 0.0062$$

Then the probability of staying in *(Replacement of the film during error 160, 0)* is:

$$P_{NoBoth,160} = 1 - P_{Both,160} = 0.9938$$

Table 5.5: Transition probabilities of the case study

Transition probability	Value
$P_{NoError}$	0.9988
$P_{Error,1}$	0.0012
$P_{ErrorReplace,1}$	$P_{Av} \cdot 0.0012$
$P_{NoErrorReplace,1}$	$(1 - P_{Av}) \cdot 0.0012$
$P_{Replace}$	0.0046
$P_{NoReplace}$	0.9954
$P_{Recover,1}$	0.0138
$P_{NoRecover,1}$	0.9862
$P_{Both,1}$	0.0062
$P_{NoBoth,1}$	0.9938

5.6.4 Results

We solved the model for the aforementioned parameter settings. We consider a solution that deviates at most 1% from the optimal long-term cost per time step acceptable. That means we specify the stopping criterion as $\epsilon = 1\%$. We varied the availability of the operator, P_{Av}, from 1 to 0.1 with steps of 0.1. The calculation time varies between 60 seconds for the higher values of C_{down} and P_{Av} to 180 seconds for lower values of C_{down} and P_{Av}.

The optimal policy has a nice structure in terms of the amount of bags b. For all solutions there is a boundary value, which we call b_{bound}. b_{bound} defines that if one faces error 160, and there there are b_{bound} or less bags, one should decide to replace the film during the error. We show the resulting values for b_{bound} and the corresponding $C^{\pi*}$ for all three cost scenarios in table 5.6. We observe that b_{bound} and $C^{\pi*}$ alter as we vary P_{Av} and C_{down}. Thus, the client should estimate the value of P_{Av} and use a value for b_{bound} depending on the C_{down}. Since currently the P_{Av} is not known, the client could use a conservative value, which means using a value for b_{bound} for a P_{Av} close to 1.

5.6.5 Sensitivity analysis

In this section we investigate how a change in input parameter affects the solution of the model. We are interested in the robustness of the solution, i.e. how much the optimal solution alters if the input is different in reality. Since we cannot vary all parameters simultaneously, we vary one parameter at a time together with P_{Av}. We are most interested in the sensitivity to the parameters that we had to estimate, as these parameters are prone to incur an estimation error. We decide to analyze the difference in optimal policy and optimal cost when our estimations incur an error.

Firstly, we consider the effect of an estimation error in the C_{down}. We estimated C_{down} for three

Table 5.6: Results for b_{bound} of the three scenarios for C_{down}

	Cost scenario 1 (C_{down} = €1000/hour)		Cost scenario 2 (C_{down} = €1500/hour)		Cost scenario 3 (C_{down} = €3000/hour)	
P_{Av}	b_{bound} (bags)	$C^{\pi*}$ (€/hour)	b_{bound} (bags)	$C^{\pi*}$ (€/hour)	b_{bound} (bags)	$C^{\pi*}$ (€/hour)
1.0	381	104,1	502	154,7	752	300,5
0.9	388	104,3	516	155,2	779	301,6
0.8	395	104,6	530	155,6	804	302,8
0.7	403	104,9	543	156,2	835	304,0
0.6	413	105,2	558	156,7	865	305,4
0.5	420	105,5	570	157,3	899	307,0
0.4	429	105,8	588	157,9	930	308,7
0.3	441	106,2	608	158,6	982	310,7
0.2	450	106,5	627	159,4	1028	312,9
0.1	462	106,9	644	160,2	1082	315,4

different scenarios during production: the scenario in which there is still slack in the production schedule, the scenario in which the client can compensate for the downtime by producing over time and the scenario in which the downtime causes loss of demand. The scheduling department stated that the first two scenarios are most common and the third scenario rarely happens (once in the last six months). We decided to consider the sensitivity to an error in the estimation of the downtime cost for scenarios 1 and 2. Based on the fact that the estimations could not be verified by data, we consider an estimation error of at most 20%. That means that we expect that the estimation of scenario 1 lies between €800 and €1200, and the estimation for scenario 2 lies between €1200 and €1800. We show the resulting optimal policies in figure 5.11. The figure shows that if the actual downtime cost is lower than estimated, the value for b_{bound} is also lower. The other way around, if the actual downtime costs are higher, the optimal b_{bound} is also higher. This is according to the expectations, since bags become relatively more expensive as the downtime costs become less, and the other way around. The figure shows that for a 20% error, the optimal b_{bound} shifts with 60 to 110 bags, depending on the direction of the error and the operator availability. For the client this implies that it is important to use a value for b_{bound} that is in line with the C_{down}.

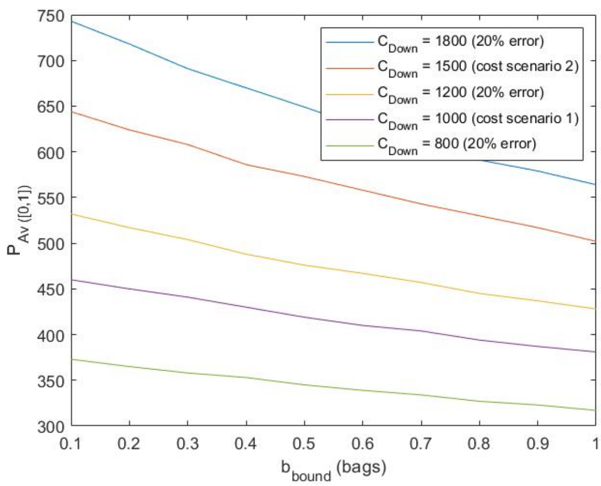


Figure 5.11: Values for b_{bound} for a 20% error in the C_{down} estimation

In figure 5.12 we show the effect of a 20% error on the long term average cost per hour. We observe that the reduction of the (long-term average) costs per hour becomes smaller as the C_{down} decreases. This is according to the expectations, since decreasing the cost of downtime is similar to increasing the cost of the film. In figure 5.12, we show the corresponding long-term cost per hour. We observe that for lower C_{down}, the cost savings are negligible (<€5 per hour). This implies that the client should only apply the opportunistic film replacements in situations with a relatively high cost of downtime. For example, a C_{down} of 1500 euro per hour can be used as a boundary. Secondly, we vary the operator reaction time, $T_{OperatorReaction}$. In order to reduce

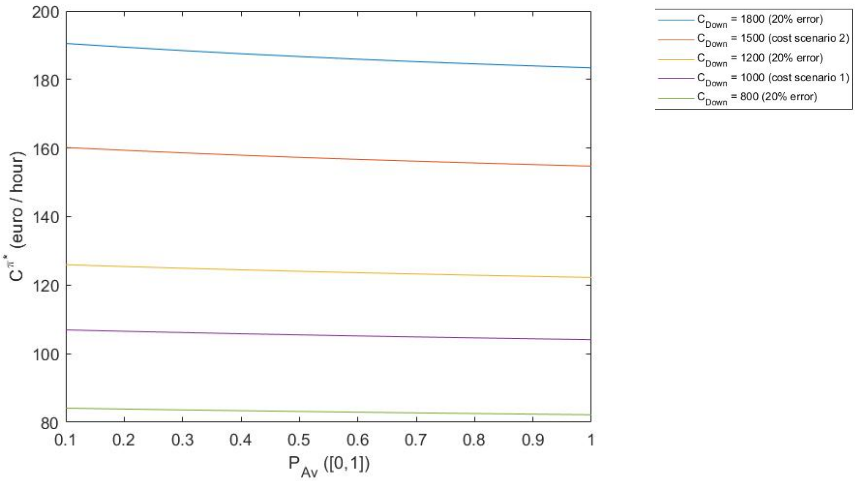
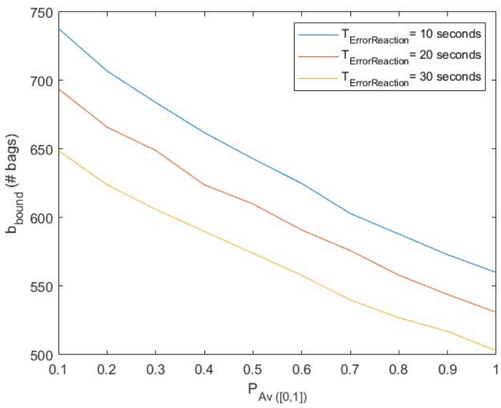
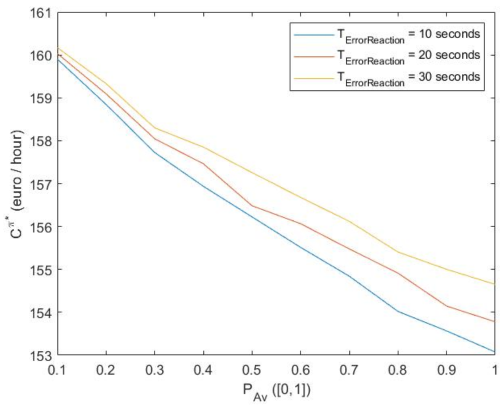


Figure 5.12: Effect of a 20% error in C_{down} on the long-term optimal cost per hour

the amount of different parameters we have to consider, we decided to focus on cost scenario 2 (C_{down} = €1500 per hour). We used a pessimistic estimation of the operator reaction time of 30 seconds. We expect the actual time to vary over the values below 30 seconds. We calculated the optimal boundaries and the optimal costs using more optimistic estimations for $T_{Both,j}$, namely 30, 20 and 10 seconds. Using the draws of random number generators for $T_{Replace}$ and $T_{Recover,160}$, we calculated the corresponding values of $\mathbb{E}[T_{Both,160}]$: 161.4, 152.1, 143.1 seconds, respectively. We show the resulting values for b_{bound} for the different values of P_{Av} in figure 5.13 (a). We observe that the optimal values for b_{bound} increase as $T_{ErrorReaction}$ decreases. This is according to our expectations, since the expected savings in terms of downtime become larger, it becomes beneficial to replace the film during error 160 at a higher amount of bags. In figure 5.13 (b), we show the corresponding change in the optimal cost per hour. We observe that the cost per downtime decreases slightly as $T_{ErrorReaction}$ decreases. The magnitude of the decrease depends on the P_{Av}. This is as expected, since a decrease in operator reaction time leads allows to avoid more downtime and the value of P_{Av} determines how much downtime can be avoided. This implies that it is important to have an accurate estimation of $T_{ErrorReaction}$, as this changes the amount downtime and cost that can be saved.



(a) Effect of a lower $T_{ErrorReaction}$ on b_{bound}



(b) Effect of a lower $T_{ErrorReaction}$ on $C^{\pi*}$

Figure 5.13: Evaluation of the effect of a lower $T_{ErrorReaction}$ on b_{bound} and $C^{\pi*}$

5.7 Conclusion

The goal of this chapter is to investigate whether we can reduce the downtime caused by the film replacement, by preventively replacing the film during unexpected breakdowns. First we selected an error, unrelated to the film nor the machine, that has a convenient duration and frequency of occurrence: error 160. Then we proposed a MDP in which we model the decision to use the error to replace the film preventively. The model considers the relevant periods of downtime, the remaining amount of film, the frequency of occurrence of errors and the probability that an operator is not available at the moment error 160 occurs. We solve the model with varying operator availability and for three different cost scenarios. We find the optimal boundary value at which it becomes optimal to replace the film. In a sensitivity analysis we vary the cost of downtime and the time an operator needs to react on an error. We find that small differences do not cause a large shift in the outcome. However, we recommend to use a conservative value for the boundary value (e.g. $P_{Av} = 0.9$), since the operator availability is currently unknown.

In chapters 5 we proceeded with answering the third research question by investigating an approach to reduce the amount of downtime caused by the film replacement.
How can we reduce the downtime of the possible causes?

Chapter 6

Quantification of the improvement

In this chapter, we quantify the current and the new procedure for replacing a film without changing to another product type. The current standard procedure consists of replacing the film as soon as it reaches its end. We refer to the current policy of not using preliminary replacements, as the *old policy*. The model in the previous chapter shows that we can reduce downtime and costs by using machine stops to preliminary replace the film. The model provides us with boundary values for the amount of bags. This boundary value tells us to replace the film if we observe error 160 and the amount of bags on the reel is equal or less than this boundary value. We refer to the policy using these boundary values as the *new policy*. We evaluate the performance of the new policy in comparison to the performance of not using preliminary replacements by discrete event simulation. By using simulation, we can incorporate the theoretical distributions that describe the data and we can verify whether the solution from the model still performs well if the assumed exponentially distributed time variables follow another distribution. We incorporate the fitted distributions for the operator response time, the replacement time and the error duration. Then we simulate the old policy and the new policy and compare the resulting downtime and costs. For the new policy, we use the boundary values we found using the model of the previous chapter.

In section 6.1, we present the variables and the parameters of the simulation. Then in section 6.2 we describe the simulation model. In 6.3 we elaborate on the settings of the model. In section 6.4 we present the results. Thereafter, we reflect on the model in section 6.5. And lastly, we conclude in section 6.6.

6.1 Variables and parameters of the simulation

Firstly, we introduce the variables of the simulation. The values of the variables changes during the simulation.

b	Variable for the amount of bags on the film reel.
C_{Total}	Variable for the total amount of costs made. This is the sum of the downtime costs and the costs for disposing the film
$C_{replace}$	Variable for the amount of costs made for disposing film
d	Variable for the amount of downtime during the simulation.
T_{Both}	Variable for the amount of time necessary to replace the film during an error occurrence. The variable is given by $max[T_{ErrorReaction}+T_{TotalReplace},\ T_{error}]$.
T_{inter}	Variable for the remaining production time until the next error. The time is a draw from a exponentially distributed random number generator.

$T_{OperatorReaction}$ — Variable denoting the operator reaction time. This is the time period starting from the moment the packaging machine reaches the end of the film until the operator reaches the machine to replace the film. This includes the time an operator is busy with another task. The value for the variable is a draw from the random number generator corresponding to the theoretical distribution found in chapter 5.

$T_{Recover}$ — Variable that denotes the duration of error 160. The value is drawn from a random number generator corresponding to the theoretical distribution found in chapter 5.

$T_{Replace}$ — Variable that denotes the replacement time. The value is drawn from a random number generator corresponding to the theoretical distribution found in chapter 5.

T_{Sim} — Variable for the remaining length of the simulation. At the start of a simulation run, the variable is set equal to D_{Sim}.

$T_{TotalReplace}$ — Variable for the time the operator needs to replace the film. The time is the sum of the draws from the random number generator corresponding to the theoretical distributions for $T_{OperatorReaction}$ and $T_{Replace}$ found in chapter 5.

Secondly, we introduce the parameters of the simulation. The parameters are fixed during the simulation.

b_{bound} — Parameter for the boundary value of the film replacement. If an error occurs and b is equal to or less than the b_{bound}, we decide to replace.

b_{new} — Parameter for the amount of bags on a new film.

C_{down} — Parameter for the cost of downtime.

C_{film} — Parameter for the cost of disposing a bag on the film reel.

D_{Sim} — Parameter for the length of the simulation.

P_{Av} — Parameter for the probability that an operator is available.

$T_{ErrorReaction}$ — Parameter for the time an available operator needs to react before he reaches the machine and starts replacing the film during error 160 (and $b \leq b_{bound}$).

v_m — Parameter for the set speed of the machine.

6.2 Simulation description

In this section we first describe the simulation, and we then give a visual representation in figure 6.1.

We start producing with a new film. In order to determine when the next error will occur, we take a draw from a random number generator for the interarrival time T_{inter}. Now we can determine what will be *the next event*. If the error occurs before we reach the end of the film and before the simulation ends, i.e. $T_{Sim} > T_{inter} \leq \frac{b}{v_m}$, (i) an error occurs before the reaching the end of the film. If the error occurs later than we reach the end of the film and the end of the simulation, i.e. $T_{Sim} > \frac{b}{v_m} < T_{inter}$, (ii) we produce until the end of the film and start replacing the film. If the error occurs later than the end of the simulation and the remaining film divided by the machine speed is longer than the time until the end of the simulation, i.e. $T_{inter} \geq T_{Sim} \leq \frac{b}{v_m}$, (iii) the simulation ends. To summarize, we determine which event happens next by looking at the minimum of T_{inter}, $\frac{b}{v_m}$ and T_{Sim} and then start to handle the event as explained in the following.

If the next event is the occurrence of an error (i), we update b and determine whether b is less than b_{bound}. If this is the case, we start replacing the film with the probability that an operator is available, P_{Av}. If we start replacing the film, we draw a sample from random number generators to determine the duration of the film replacement during the error. The duration is calculated as $T_{Both} = max[T_{ErrorReaction} + T_{TotalReplace},\ T_{error}]$. Then we again, determine whether we first finish replacing the film during the error, or that we reach the end of the simulation. We update the variables d, b, T_{Sim} accordingly. If we have not reached the end of the simulation, we draw a number from the random number generator for the T_{inter} to determine the occurrence of the next error. We resume production and determine *the next event.*
In the other case, if we do not replace the film ($b<b_{bound}$, or no available operator), we determine $T_{Recover}$ by a draw from the corresponding random number generator. Then we determine whether the simulation time, T_{Sim}, is still longer than the error duration, $T_{Recover}$. If so, we update the variable d. Then we draw a number from the random number generator for the T_{inter}, to determine the occurrence of the next error. We resume production and determine *the next event.*

If the next event is a regular film replacement (ii), we determine the total time to replace the film, which is $T_{Replace} + T_{OperatorReaction}$, by drawing from the corresponding random number generators. Then we determine whether we first finish replacing the film, or that we reach the end of the simulation. We update the variables d, b, T_{Sim} accordingly. If the simulation has not yet ended, we draw a number from the random number generator for T_{inter}. We resume production and determine *the next event.*

If the next event is the end of the simulation (iii), we stop the simulation.

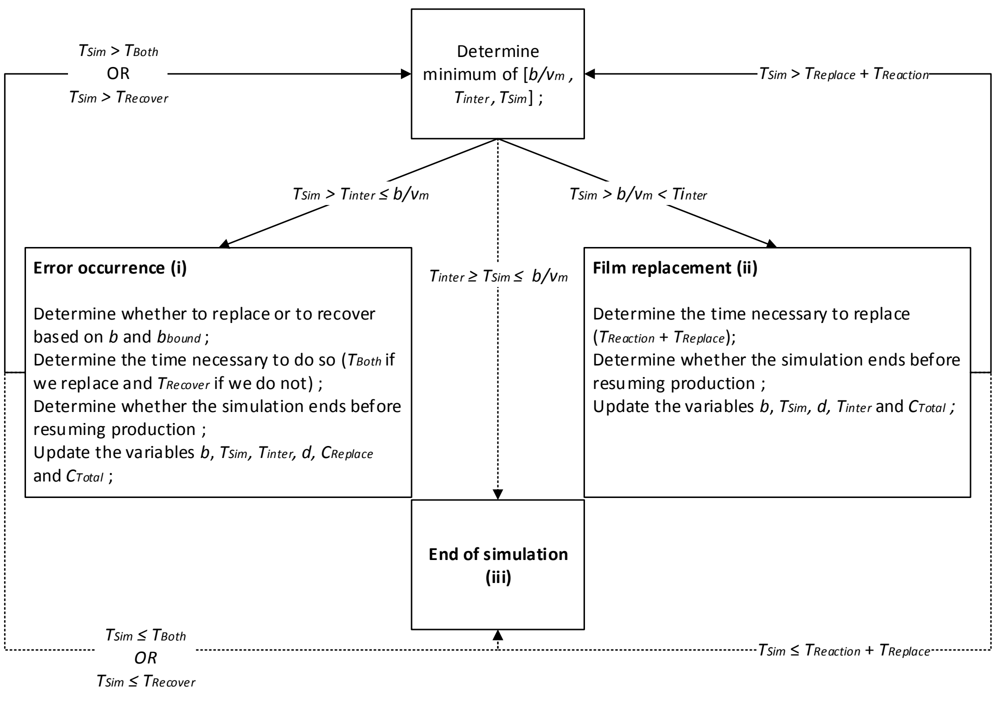


Figure 6.1: Visual representation of the simulation

6.3 Parameter settings

In simulations that contain queues for the arrival process, e.g. for incoming demand, it is common to cut-off the first period of the simulation. This first period is called the warm-up period. This is useful when the system starts with an empty queue, which is not representative for the real situation the simulation aims to describe. However, in our process no queuing takes place and the process regularly resets when the film is replaced.

In order to determine the length and the amount of simulation runs, we followed Boon, van Leeuwaarden, Mathijsen, van der Pol & Resing (2017). This procedure consists of estimating the standard deviation of the variable of interest, and then calculating the amount of runs necessary to obtain the desired confidence interval. The procedure is based on the implications of the central limit theorem. We followed the procedure and found that we need 23 runs of 10^4 hours to obtain our desired bounds. The detailed calculation and validation is provided in appendix I. Before we determined the amount of necessary simulation runs, we used a simulation length of 10^5 hours with 25 runs. In retrospect, we can conclude that this simulation length was unnecessarily long. However, the confidence intervals are tighter and we used the results of this simulation. We provide the parameter settings of the simulation in 6.1.

We want to compare the downtime and the costs for replacing the film prematurely in comparison to only replacing the film when it reaches its end. We refer to replacing the film prematurely as *the new policy*, and to replacing the film when it reaches its end as *the old policy*. For the new policy we use the values for b_{bound} we found in the previous chapter by varying the value for P_{Av} (0.1 - 1) and by varying the C_{down} for the three scenarios (1000 €/hour, 1500 €/hour and 3000 €/hour). For the old policy, we set b_{bound} to zero, which means that we only replace when we reach the end of the reel. These parameters, and the remaining parameter values per scenario are shown in 6.1.

Table 6.1: Parameter settings for the simulation

	Scenario 1		**Scenario 2**		**Scenario 3**	
	New policy	*Old policy*	*New policy*	*Old policy*	*New policy*	*Old policy*
P_{Av} ([0,1])	0.1 - 1.0		0.1 - 1.0		0.1 - 1.0	
C_{down} (€/hour)	1000	1000	1500	1500	3000	3000
$T_{ErrorReaction}$ (sec)	30	30	30	30	30	30
b_{bound} (# bags)	*See table 5.6*	0	*See table 5.6*	0	*See table 5.6*	0
D_{Sim} (hours)	10^5	10^5	10^5	10^5	10^5	10^5
b_{new} (# bags)	6700	6700	6700	6700	6700	6700
C_{film} (€/film)	400	400	400	400	400	400
v_m (bags/min)	60	60	60	60	60	60

6.4 Results

In this section we discuss the results of the described simulation. We consider the total amount of downtime and the total costs. In order to calculate the confidence intervals, we follow the approximation of the confidence intervals from Boon et al. (2017) (page 23). This approximation is based on implications of the central limit theorem. We decided to use a 95% confidence level, since we consider this an acceptable confidence level. We calculate the bounds for the 95% confidence interval by: $\bar{Z} - 1.96 \cdot \sqrt{\frac{S^2}{n}}$ and $\bar{Z} + 1.96 \cdot \sqrt{\frac{S^2}{n}}$, in which $\bar{Z}$ denotes the mean of the 25 simulations, S^2 the sample variance and n the number of repetitions of the simulation.

First we consider the reduction in downtime. In figure 6.2 we show the downtime of the new policy as a percentage of the downtime of the old policy. We varied P_{Av} and we considered the three different downtime cost scenarios ($C_{down} = \{1000$ €/hour, 1500 €/hour, 3000 €/hour$\}$). The downtime consists of the downtime due to the film replacement and due to error 160. The results are shown per scenario, per policy (new/old) and per value for the availability P_{Av}. The old policy resulted in 0.114 hours of downtime per hour of production. We observe that the client can save between 0 - 7.2% depending on the operator availability at the moment error 160 occurs and the cost scenario. In absolute value, this is roughly half a minute. One can find the values for the total downtime with the corresponding confidence intervals in appendix K.

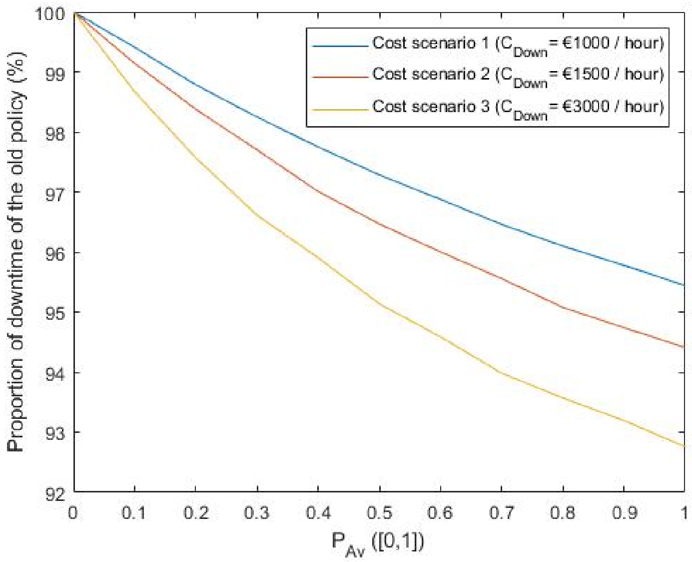


Figure 6.2: Downtime of the new policy as a percentage of the downtime of the old policy

Then we consider the total costs of the new policy in comparison to the total cost of the old policy. The total cost consist of the costs of disposing film and the costs of downtime related to the film replacement and error 160. In figure 6.3 we show the total cost as a percentage of the total cost of the old policy. The total cost resulting from the old policy for cost scenario 1, 2 and 3 is €114.2, €171.3 and €342.7, respectively. We observe that the cost reduction lies between 0% - 5.8 %. In absolute value that is a maximum reduction of €17.7 per hour of production. One can find the values for the total downtime with the corresponding confidence intervals in appendix K.

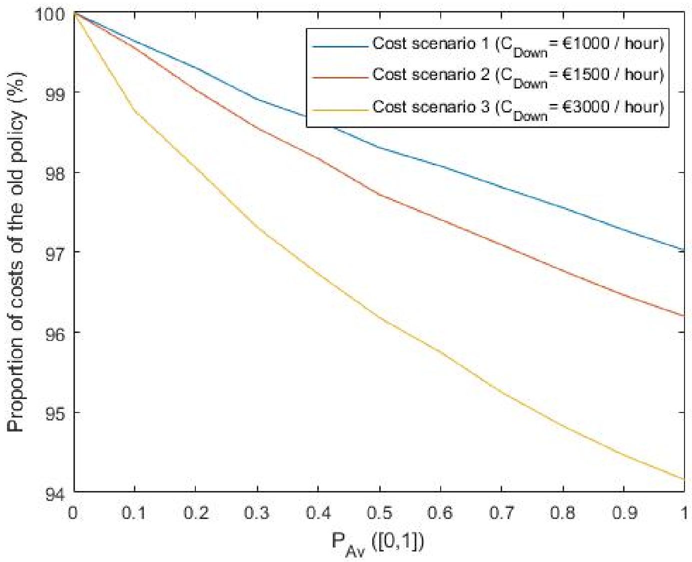
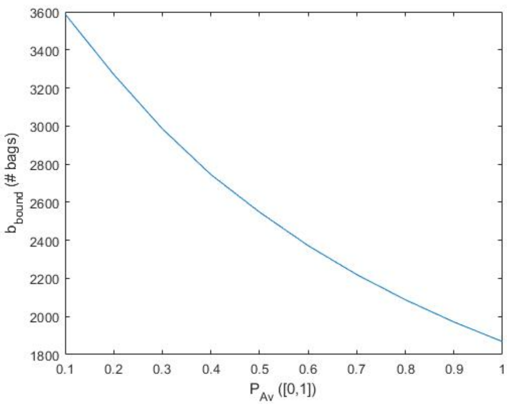


Figure 6.3: Cost of the new policy as a percentage of the downtime of the old policy

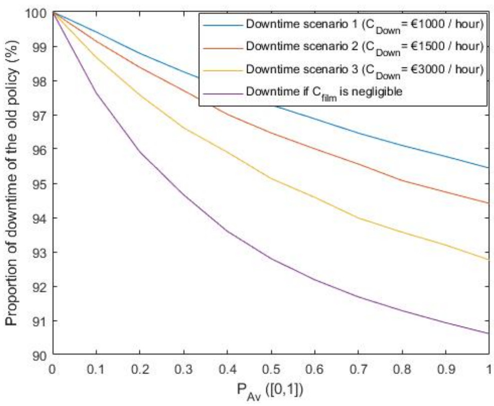
6.5 Reflection on the results

In the previous section, we found that the new policy slightly reduces the amount of downtime and the amount of costs per hour of production. For cost scenario 1 the cost reduction is negligible (<€5). However, the amount of downtime and the amount of costs one can reduce by applying opportunistic film replacements increases as the cost of downtime increases. We know that the client of interest has a relatively low utilization and a relatively low cost of downtime in comparison to other clients from Bosch. This means that although we obtained a small improvement, the method can be more valuable for other clients. In order to determine the performance for clients with a higher downtime cost, we quantified the improvement using a negligible C_{film} (equal to zero) in comparison to the C_{down}. We show the resulting values for b_{bound} and the resulting downtime as a percentage of the downtime resulting from the old policy in figure 6.4. For each of the values for $P_{Av} > 0$ the amount of downtime can be slightly reduced by using higher values for b_{bound}. The downtime reduction for the policy with negligible C_{film} lies between 0% - 9.4%. This means a absolute reduction of 0 - 38 seconds of downtime per production hour. This result shows that the method would have resulted in slightly better results in a case with a higher cost of downtime. However, we observe that the relative cost reduction becomes smaller as the cost in downtime increases. This implies that for clients with similar time variables and higher downtime cost, the percentual reduction in downtime an cost will be similar to the percentual reduction of the specific client.

In the previous chapters, we determined the parameters and the variables. We expect that the method we used to identify the film replacements in the data (chapter 4), might have left out film replacements in which problems occurred. This might have resulted in a underestimation of the expectation of $T_{Replace}$. In addition, in the fitting procedure for the duration of error 160, $T_{Recover,160}$, we had to leave out data points with a relatively high duration. According to the boxplot, these data points lied at an abnormal distance from the remaining data points. We could not verify whether these data points were related to a very unusual problem, so we might have underestimated the duration of $T_{Recover,160}$. Considering the possible underestimations of $T_{Replace}$ and $T_{Recover,160}$ the results may be slightly pessimistic, i.e. the downtime reduction and the cost reduction might be slightly higher.



(a) The values for b_{bound} if the C_{film} is negligible



(b) The downtime of the new policy and the policy with a negligible C_{down} as a percentage of the downtime of the old policy

Figure 6.4: Evaluation of the downtime of the policy with negligible C_{film}

6.6 Conclusion

The goal of this chapter is quantify the downtime reduction of the new policy in comparison to the old policy by discrete event simulation. Using discrete event simulation, we could validate that the solution of the model in chapter 5 performs well if the time variables do not follow an exponential distribution. We explained the variables and parameters we used, we described the simulation, and we showed which parameter settings used. The results showed that the new policy reduces the downtime of error 160 and the film replacement up to 7.2% and costs up to 5.8%. This is a downtime reduction of 28 seconds per hour and a cost reduction of €17.7 per hour. We also determined what the downtime reduction would be if the cost of the film would be negligible in comparison to the downtime cost. In this case, we found a downtime reduction up to 9.4% of the downtime of error 160 and the film replacement.

In chapters 6 we took the final step in answering the third research question by quantifying the improvement in the clients' film replacement policy.
How can we reduce the downtime of the possible causes?

Chapter 7

Conclusions and recommendations

In this chapter we discuss the main findings of this research. First we summarize the answers on the research questions in section 7.1. Then we make recommendations for improvement for Bosch and the client in section 7.2. Finally, we discuss the limitations of this thesis and we make recommendations for future research in section 7.3.

7.1 Conclusions

How can we identify the errors related to downtime?
First we determined that the current method (old method) the client uses to identify the downtime related errors, neglects important aspects of the relation between errors and the corresponding machine downtime. We proposed a new method that takes these aspects into account and finds the amount of downtime an error relates to. We find that the old method overestimates the downtime per error due to the neglected aspects.

What are the possible causes of the errors related to downtime?
By answering the previous research question, we obtained an overview of which errors are related to downtime. We are interested in the possible causes of the errors, such that we can reduce the causes from resulting in downtime. We applied FTA to the top three errors, which is a structured approach to find possible causes of a fault. In this procedure we combined the knowledge from Bosch on the triggers of errors and the knowledge of the client on the way they operate the packaging machine. Subsequently, we proposed a method to find the frequency that a possible cause results in an error using the machine data. In order to apply this method, we selected one of the possible causes based on amount of downtime of the error, the likelihood that the possible cause actually relates to downtime and the likelihood of identifying the possible cause in the data set. We selected marker B as the possible cause, which is a marker at the end of the film. We decided to identify the possible cause by identifying the film replacements in the data set.

How can we reduce the downtime of the possible causes?
In the previous research question, we proposed a method to determine the frequency a possible cause has resulted in an error and we selected one of the possible causes for investigation. In this chapter we apply the proposed method to the selected possible cause. We distinguish the steps of a film replacement and find which of these steps are reflected in the error log. Then we determine which set of indicators we consider as a film replacement. We manually filtered out the film replacements and found that the film replacement accounts for 65% of the total downtime of error 401. We use the identified film replacements to determine the average duration of the film replacement.
In order to reduce the downtime, we propose to opportunistically schedule the film replacement during unexpected machine stops. We selected error 160 as a suitable error to prematurely replace

the film, based on its duration and frequency of occurrence. We incorporate the data on error 160 and the data on the film replacement in an opportunistic model, that tells at which remaining film length it is beneficial to use the downtime of error 160 to prematurely replace the film. We find that the boundary value at which this becomes beneficial, depends on the operator availability and the cost of downtime. The model suggests different boundary values for different cost scenarios. In order to quantify the downtime reduction and the cost reduction of prematurely replacing the film, we simulate both policies: replacing the film when it reaches its end, and prematurely replacing the film. We find that prematurely replacement can reduce the downtime of error 160 and the film replacement 7.2% and reduce the costs 5.8%.

7.2 Recommendations for Bosch and the client

Keep track of the reasons for manual machine stops
In chapter 2 and 3, we found that a significant amount of downtime is caused by manual machine stops. However, the client currently does not keep track of the reasons for these stops. for both Bosch and the client, it is useful to keep track of these reasons. We recommend to find a way to keep track of these reasons without bothering the operators. A possibility is to let operators select a reason from a shortlist on the HMI, after they have started the machine again. By doing so, it would be possible to keep track of the main reasons for machine stops. Such a shortlist on the HMI requires adjustment of the HMI software.

Order films with markers according to the production schedule
In chapter 3, we found that one of the possible causes of downtime are markers on the film. We found that marker B caused a subset of the occurrences of error 401. Based on the amount of remaining error occurrences of 401, we expect that the client currently uses films with marker A when it is not necessary to do so. We recommend the client to incorporate the production schedule in the decision to order films with markers A and B. Using films with a marker halfway while one needs to use the entire film, will result in unnecessary downtime during production of the film.

Preventive replacement of the film
The model in chapter 5 shows the values for b_{bound} at which it becomes beneficial to replace the film if error 160 occurs. We observed that the cost reduction is negligible for low downtime cost. We therefore recommend to only use opportunistic film replacement when the downtime cost is €1500 or higher (scenario 2 or 3). The value for b_{bound} depends on the cost of downtime, C_{down} and the operator availability, P_{Av}. The latter is currently unknown, so we recommend to use a conservative value for b_{bound} for each of the cost scenarios, e.g. the value for a P_{Av} of 0.9. In addition, we advise to gather data of the opportunistic replacements to estimate the value of the operator availability P_{Av} and the time the operator needs to react on the opportunity arising from error 160, $T_{ErrorReaction}$. Based on these estimated values, one can update the parameters used in the model to find a more accurate value for the optimal replace boundary b_{bound}.

In order to use error 160 to prematurely replace the film, the operators should be notified when the machine passes the b_{bound} value. The most practical implementation is to adjust the limit of the pre-warning of reaching the end of the film according to the b_{bound} value. The machine raises a warning once it passes this limit. However, this limit is known to be inaccurate due to small changes in film thickness. In order to obtain an accurately timed notification for passing the b_{bound}, we recommend to use the bag counter on the machine. If the amount of bags before the b_{bound} is known, the bag counter can provide the exact moment the operators should replace the film if error 160 occurs. The implementation requires adjustment of the current software of the HMI, since there is currently no function to generate a notification at a specified number of bags on the bag counter. Ideally, passing the b_{bound} should generate a notification that is visible even if the operator is not standing in front of the packaging machine. We recommend to use the disposed films for product tests.

7.3 Limitations and future research

We identified a number of limitations of the research and an interesting direction for future research. First of all, we only considered a subset of all faults, namely the faults detected by the packaging machine. This means that we only considered part of the problems the client faces in practice. It remains to be investigated what the impact is of undetected faults and what their possible causes are. In addition, we only measured the impact of the errors in terms of downtime, due to the unavailability of data on the quality of bags. It remains a topic of interest, to determine which errors are related to quality problems and to incorporate this into the overview of impact per error.

In chapter 4, we proposed a method to identify the marker at the end of the film using the machine data. We found that several steps during the replacement of the film are reflected in the error log, but the exact sequence of indicators varies and may have indicators of other machine interaction in between. We expect that our method filtered out film replacements in which operators made mistakes, since these replacements result in deviating indicator sequences. We consider this is an important limitation of our method, and we suggest to overcome this limitation by investigating the sensor data of the packaging machine. We investigated data at the HMI level of the machine, i.e. the processed data that is used for messages on the HMI. In the future, it will be possible to gather data from the Open Platform Communications (OPC) level, which allows to gather the unprocessed sensor data. This development step allows to gather the angular velocity, which in turn allows to make an approximation of the diameter of the film on the machine. One can estimate the diameter of the reel by using:

$$v = \omega \cdot r$$

In which r is the diameter of the film roll, v is the speed of the incoming film and ω is the angular velocity.

Another limitation of this research is that we could not control all variables that influence the machine performance. We used the production shifts from May and June, which comprise approximately 93 hours of uptime. During these shifts, several different operators have been operating the machines. Due to confidentiality reasons, we could not obtain data on the different operators during these shifts. In addition, the machine has produced three different bag types, but we could not gather data on which bag type the machine was producing over time. It remains unknown how the machine performance is influenced by difference in operators and difference in bag type. The initial plan of this research was to gather data of six months, however, due to technical issues the first four months of data were not usable. Due to the limited amount of time we had left, we were not able to cross-validate whether the data of these months is representative. However, it is a small effort to apply the data processing method from chapter 2 again, once a larger data set has been obtained.

In chapter 5, we proposed a model in which we investigated the opportunity to preventively replace the film during the downtime of error 160. We determined the duration of replacing the film during an error given that an operator is available at the moment the error occurs. However, we did not investigate the possibility to let operators stop with their current task to replace the film, nor the possibility to let them finish their task and replace the film afterwards. It may be beneficial to incorporate both possibilities based on data of the operator tasks during a production run.

In this project, we solely focused on the machine data of the packaging machine. The machine is notified when the upstream or downstream equipment faces a problem and we used downstream error 160 to prematurely replace the film. It could be interesting to incorporate the error logs of the downstream and upstream equipment into this analysis, since the errors contain data on the duration and frequency. This data could be useful in determining whether errors related to the upstream or downstream equipment have a long duration or a short duration. From the perspective of the packaging machine, all errors from the secondary packaging machine are summarized by errors 160 and 251. However, the secondary packaging machines error may reveal more information about the actual problem. This data is very useful in determining whether one should use the opportunity to replace the film.

An interesting direction for future research, is investigating the influencing factors of error occurrence and error duration. As stated, we think that the skill level of operators and the bag type can be of influence. One could investigate the influence of these factors by comparing the machine performance, while controlling for other possible factors of influence. The resulting relations may implicate how clients from Bosch can improve their machine performance.

References

Arts, J. (2017). *Maintenance modeling and optimization.* Technische Universiteit Eindhoven.

Aslan, B. & Zech, G. (2002). Comparison of different goodness-of-fit tests. *arXiv preprint math/0207300*.

Bain, L. & Engelhardt, M. (1992). *Introduction to probability and mathematical statistics.* Brooks/-Cole Cengage Learning. Retrieved from `https://books.google.nl/books?id=MkFRIAAACAAJ`

Boon, M., van Leeuwaarden, J., Mathijsen, B., van der Pol, J. & Resing, J. (2017, February). *Stochastic simulation (lecture notes).* TU/e.

Botman, L. A. (2017). *Machine failure diagnostics and prognostics using error information and sensor data* (Master Thesis). Technische Universiteit Eindhoven.

Cohen, A. C. & Whitten, B. J. (1980). Estimation in the three-parameter lognormal distribution. *Journal of the American Statistical Association*, *75*(370), 399–404.

Daley, D. J. & Vere-Jones, D. (2007). *An introduction to the theory of point processes: volume ii: general theory and structure.* Springer Science & Business Media.

de Win, M. M. (2018). *Preventive maintenance planning for additive manufacturing systems* (Master Thesis). Technische Universiteit Eindhoven.

General Services Administration. (1980). Telecommunications: Glossary of telecommunication terms [Standard].

Griffiths, D. A. (1980). Interval estimation for the three-parameter lognormal distribution via the likelihood function. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *29*(1), 58–68.

Haasl, D., Roberts, N., Vesely, W. & Goldberg, F. (1981, Jan). Fault tree handbook.

Kowarik, A. (n.d.). *Estimate parameters of a three-parameter lognormal distribution (log-scale).* `https://www.rdocumentation.org/packages/EnvStats/versions/2.3.1/topics/elnorm3`. (Accessed: 20-09-2019)

Lee, W.-S., Grosh, D. L., Tillman, F. A. & Lie, C. H. (1985). Fault tree analysis, methods, and applications: A review. *IEEE transactions on reliability*, *34*(3), 194–203.

López, J. P. G. (2017). *Condition-based maintenance employing system errors* (Master Thesis). Technische Universiteit Eindhoven.

Pilot, S. (2002, Mar). *What is a fault tree analysis?* ASQ. Retrieved from `http://asq.org/quality-progress/2002/03/problem-solving/what-is-a-fault-tree-analysis.html` (Accessed: 2019-25-10)

Reliability HotWire. (2013). *Discussion of maximum likelihood estimation for the 3-parameter weibull distribution.* `https://www.weibull.com/hotwire/issue148/hottopics148.htm`. (Accessed: 10-09-2019)

Shifley, S. & Lentz, E. (1985). Quick estimation of the three-parameter weibull to describe tree size distributions. *Forest Ecology and Management*, *13*(3-4), 195–203.

Slack, N. & Lewis, M. (2002). *Operations strategy.* Pearson Education.

Smith, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, *72*(1), 67–90.

Teimouri, M. & Gupta, A. K. (2013). On the three-parameter weibull distribution shape parameter estimation. *Journal of Data Science*, *11*(3), 403–414.

Thacker, B., S.W.Doebling, Hemez, F., Anderson, M., Pepin, J. & Rodriguez, E. (2004, 10). Concepts of model verification and validation.
doi: 10.2172/835920

Tijms, H. C. (2003). *A first course in stochastic models.* John Wiley and sons.

Yuan, K.-H. & Bentler, P. M. (2001). Effect of outliers on estimators and tests in covariance structure analysis. *British Journal of Mathematical and Statistical Psychology*, *54*(1), 161–175.

Zhu, Q. (2015). *Maintenance optimization for multi-component systems under condition monitoring* (Unpublished doctoral dissertation). Department of Industrial Engineering Innovation Sciences.

Appendix A

Key observations for impact determination

In this appendix, the key observations are illustrated that were used to find the actual downtime caused by error messages.

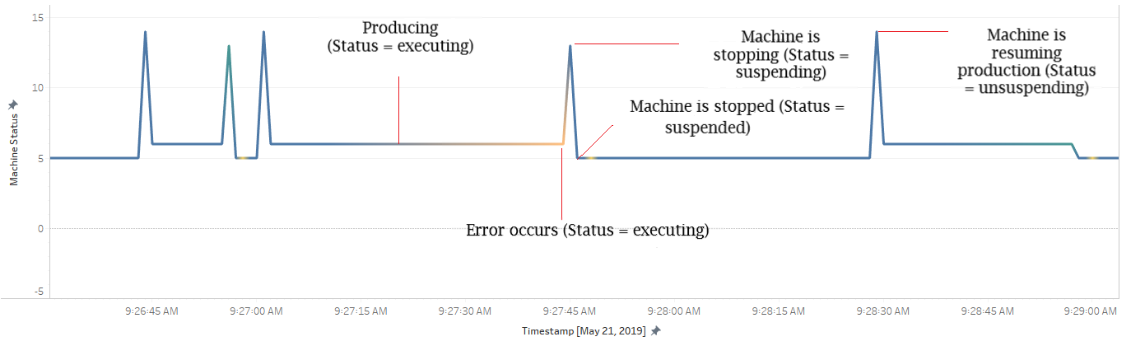


Figure A.1: Plot of the machine status. The plot shows the observation that whenever the machine runs into an error the machine is stopped and the machine status is changed from executing to suspending

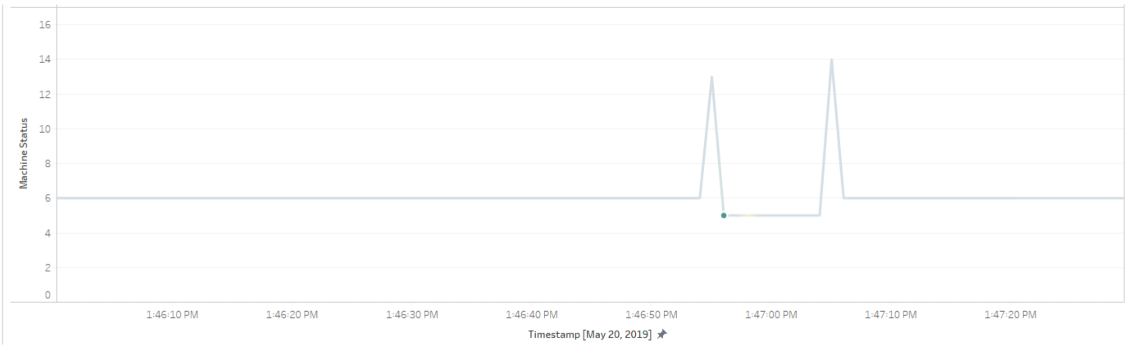


Figure A.2: Plot of the machine status with highlighted error occurrence. The timestamp of the error occurrence deviates 2 seconds from the last timestamp at which the machine is executing (machine status = 6)

Appendix B

Format of the machine data

In this chapter, we show the format of the data.

Format error log

The error log contains the codes of the messages raised by the machine with a millisecond timestamp. When the message is cleared, the machine logs the same number with a minus sign. The number can correspond to an error, a warnings or a notification. We know from an internal document what message corresponds to the number and whether this is an error, warning or notification. There are 1690 different messages so we do not provide information on all messages, but we provide in-text explanation of errors necessary.

Table B.1: Format of the error log

Error code	**Timestamp**
1	2019-07-01 03:40:27:889000
-1	2019-07-01 03:48:01:464000
3	2019-07-01 03:55:15:048000
-3	2019-07-01 04:01:10:111000

Format machine status

The machine status log consists of a number with a timestamp. The machine logs this number when the machine status is changed to a new machine status. From the software department we received the integer value that denotes the machine status, corresponds to an actual status. The status can be Clearing, Stopped, Starting, Idle, Suspended, Executing, Stopping, Aborted, Holding, Held, Unholding, Suspending, Unsuspending, Resetting, Completing and Complete. The machine starts almost immediately after pressing start and stops almost immediately when an error occurs. We mainly use the machine status to determine whether the machine was executing or not executing.

Table B.2: Format machine status data

Machine status	**Timestamp**
5	2019-07-01 03:40:27:889000
6	2019-07-01 03:48:01:464000
9	2019-07-01 03:55:15:048000
5	2019-07-01 04:01:10:111000
6	2019-07-01 04:01:10:243000

Format bag counter

The bag counter counts the amount of filled bags that are produced. The bag counter logs an integer with the new amount of bags and the corresponding timestamp.

Table B.3: Format of the bag counter

Bag count	**Timestamp**
16826280	2019-07-01 03:40:27:889000
16826281	2019-07-01 03:48:01:464000
16826282	2019-07-01 03:55:15:048000
16826284	2019-07-01 04:01:10:111000
16826285	2019-07-01 04:01:10:243000

Appendix C

Assumption of exponentially distributed time variables

In this appendix, we examine the difference between the empirical distributions of the data and the exponential distribution the model assumes. This examination is part of the discussion of the model assumptions in section 5.3. The data that is used, is coming from May and June 2019. In figure C.2, we plot the exponential distribution assumed by the model and the empirical distribution based on the observed data of the replacement time. In figure C.3, we plot the exponential distribution assumed by the model and the empirical distribution based on the observed data of the error 160 duration. In figure C.4, we plot the exponential distribution assumed by the model and the distribution of the duration of replacing the film during an error, constructed using sampling from random number generators. We calculated this distribution by repeatedly drawing random numbers from the distributions for the error duration and the replacement time. Then we took the maximum of the replacement time plus the assumed 30 seconds of response time of an available operator and the error duration.

We observe that each of the exponential distributions, overestimates the probability of having a very short duration. In order to determine what this means for the solution the model provides, we examine the value iteration algorithm we used of Tijms (2003). This algorithm starts by determining the optimal decision in each state with one time step to go. Since the expected costs over one time step can be calculated directly, this results in the one step expected costs, which is denoted by $V_1(s)$. In this step, the model overestimates the probability of leaving the error states, the replacement state and the state in which we replace during an error. This means that in each of these states the value for $V_1(s)$ is lower than the actual value. Then the same holds for the subsequent time steps until the exponential distribution no longer overestimates the probability, i.e. where the red dotted line crosses the empirical distribution. Since $V_n(s)$ is computed recursively by using:

$$V_n(s) = \min_{a \in A} c_a(s) + \sum_{s' \in S} P_{s,s'}(a) V_{n-1}(s')$$

we expect that the model underestimates the expected costs of the states in which the machine is down. It is not possible to compute the error from the graphs and we do not make further approaches in quantifying the difference. However, we quantify the performance of the solution in a simulation in chapter 6, and we find that the solutions perform generally well.

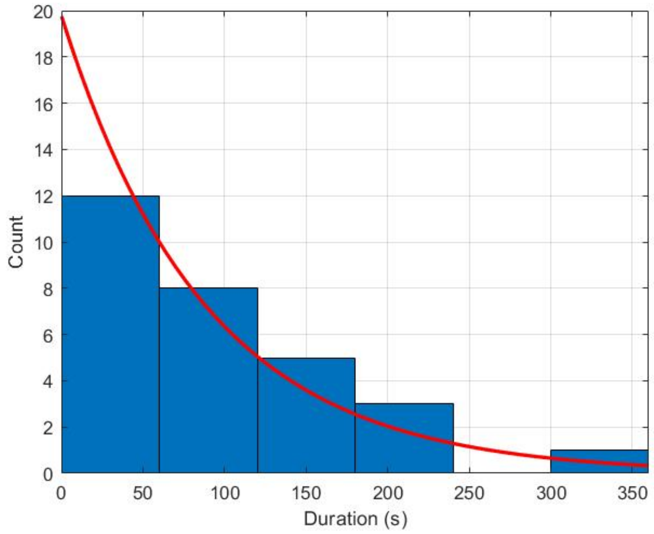


Figure C.1: Comparison of the empirical distribution and the exponential distribution assumed by the model for $T_{OperatorResponse}$ (n = 29)

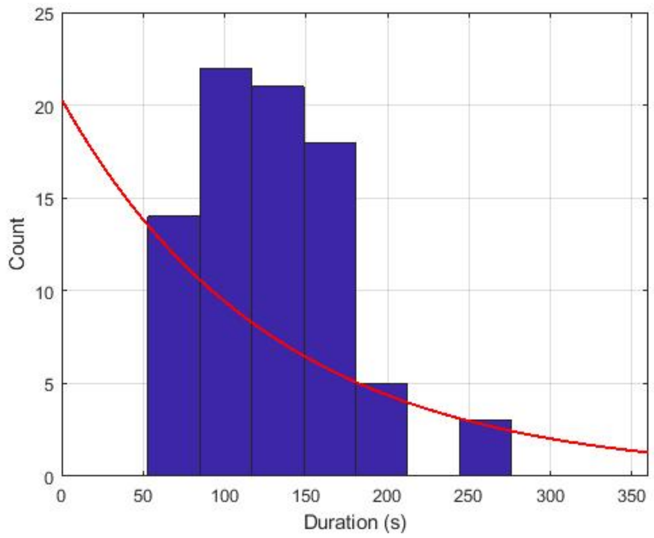


Figure C.2: Comparison of the empirical distribution and the exponential distribution assumed by the model for $T_{Replace}$ (n = 83)

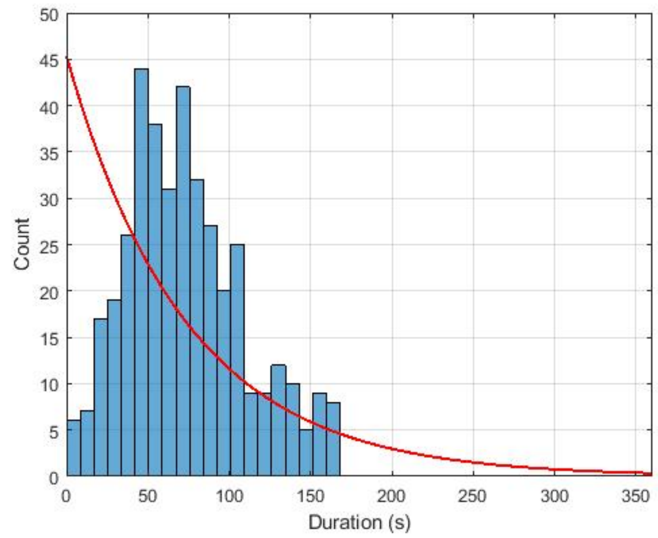


Figure C.3: Comparison of the empirical distribution and the exponential distribution assumed by the model for $T_{Recover,160}$ (n = 396)

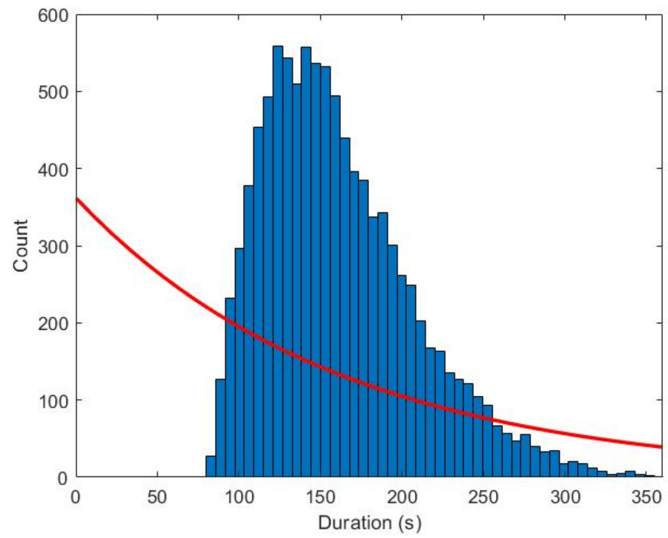


Figure C.4: Comparison of the empirical distribution and the exponential distribution assumed by the model for $T_{Both,160}$ (constructed from 10,000 samples)

Appendix D

Validation and verification of the model

In this appendix, we elaborate on the validation and the verification of the MDP. Model validation is the task of determining to which degree the model represents the real world from the perspective of the intended use of the model. Model verification is the task of determining whether the model implementation accurately represents the developer's conceptual description of the model and the solution to the model (Thacker et al., 2004).

Model validation

Part of the validation of the model is done by discussing the modeling assumptions in section 5.3. We concluded that the time steps of the model should be relatively small in comparison to the time variables, the error rate should be approximately constant during production, the distribution of the time variables should rule out that it is optimal to take the elapsed time in a state into account and the probability of observing two opportunities to replace the film should be fairly small. We discussed that each of these assumptions seems reasonable based on the data we have. Furthermore, the model assumes that the time variables are exponentially distributed. We observe that this is assumption is not in line with the real distribution of the time variables. We therefore expect that the model finds a slightly optimistic long-term average cost of a policy. The difference in distribution raises the need to validate the results in a more general setting in which we incorporate distributions that describe the actual data. This is done in chapter 6 and we concluded that the results perform well if the assumption of exponentially distributed time variables is relaxed.

Model verification

We implemented the model for two errors in Matlab 2019. If only one error is considered, one can easily set the arrival rate of the second error to zero, which simplifies the model to the one error case. Since there are no results available that we can replicate, we validate the model by considering several settings with extreme parameter values. In these extreme cases, we can make statements about the optimal solution, and we can verify if the solution of the model is in line with these statements. We came up with the following extreme parameter situations with a $P_{Av} > 0$:

- $\mathbf{E}[T_{Both}]$ is longer than the sum of $\mathbf{E}[T_{Recover}]$ and $\mathbf{E}[T_{TotalReplace}]$.
  Given that C_{down} and C_{film} are positive, it is never optimal to replace the film during an error since it incurs extra downtime.

  We tested the extreme case for five different realistic values of C_{down} and we observed in none of the cases an optimal policy with the decision to preventively replace. In each of the cases, we verified whether the model provided an optimal policy with the decision to replace the film during the error, if we altered the expectations of the time variables such that $\mathbf{E}[T_{Both}] < \mathbf{E}[T_{Recover}] + \mathbf{E}[T_{TotalReplace}]$. The optimal policy changed as expected to an optimal policy with the decision to replace the film during the error from a certain amount of bags on the reel.

- A negligible C_{down} in comparison to the C_{film}.
  If C_{down} is negligible in comparison to the C_{film}, it is never optimal to dispose bags. That means that it is never optimal to decide to replace the film during an error, except for when the film is empty, i.e. one faces an error during producing the last bag.

  We tested this case for five different realistic combinations of the time variables with $\mathbf{E}[T_{Both}]$ $< \mathbf{E}[T_{Recover}] + \mathbf{E}[T_{TotalReplace}]$. We observed in each of the cases that the optimal policy is to only replace when there is one bag left on the film reel, which is in line with the expected solution.

- A negligible C_{film} in comparison to the C_{down}.
  If the C_{film} is negligible in comparison to the C_{down}, the absolute value of the C_{down} should have no influence on the optimal policy. Without knowing the optimal policy, we can test this by setting C_{film} equal to zero and varying the C_{down}.

  We tested this case with five different combinations of expected values for the time variables with $\mathbf{E}[T_{Both}] < \mathbf{E}[T_{Recover}] + \mathbf{E}[T_{TotalReplace}]$. We observed that each of the cases the model provides an optimal policy that does not change if we alter the value of C_{down}.

Besides these situations, we determine whether the optimal policy changes as expected when we alter the cost parameters or the expectation of the time variables. The policy changes according to our expectations. Since the solutions of the model are in line with the solutions we expect, we can conclude with a probability bordering on certainty that the model is implemented correctly.

Appendix E

Numerical example

In this appendix, we show a numerical example of the model of chapter 5 with two different errors, i.e. $|J| = 2$.

Numerical example with $|J| = 2$

Now we consider a fictitious example with two independent errors that can be used to prematurely replace the film, i.e. $|J| = 2$. That means that the set of machine status M consists of:

$$m = \begin{cases} \textit{In production}; \\ \textit{Replacement of the film}; \\ \textit{Down due to error 1}; \\ \textit{Down due to error 2}; \\ \textit{Replacement of the film during error 1}; \\ \textit{Replacement of the film during error 2}; \end{cases}$$

The machine is set at a speed of 30 bags per minute, so we set τ equal to 2 seconds. The amount of bags on a new film is equal to 1000 bags, so $b_{new} = 1000$. The cost of disposing one bag on the film is €0.10, so $C_{film} = 0.10$. The cost of 1 hour of downtime is €2000, so the $C_{down} = 2000$. The error rates of error 1 and error 2 are 1 λ_1 and λ_2. λ_1 is equal to 1 error per hour and λ_2 is equal to 10 errors per hour. The expected duration of error 1, $\mathbb{E}[T_{recover,1}]$, is equal to 150 seconds and the expected duration of error 2, $\mathbb{E}[T_{recover,2}]$, is equal to 200 seconds. The expected replacement duration when the end of the film is reached is 250 seconds, so $\mathbb{E}[T_{TotalReplace}] = 250$. Since the machine may reach the end of the film at a moment that no operator is available, this number is typically higher than the replacement during an error at which an operator is available. With probability 0.7, there is an available operator at the moment error 1 or 2 occurs, so P_{Av}=0.7. If an operator is available at the occurrence of an error 1 and replaces the film, this takes in expectation 250 seconds, so $\mathbb{E}[T_{Both,1}] = 200$. If an operator is available at the occurrence of an error 2 and replaces the film, this takes in expectation 300 seconds, so $\mathbb{E}[T_{Both,2}] = 300$. The values are summarized in table E.1. We calculate the transition probabilities of the model, using the given parameter values and the equations described in subsection 5.4.4. The detailed calculation is described below. We summarize the resulting transition probabilities in table E.2. We solve the corresponding model using the value iteration algorithm from Tijms (2003) using a stopping criterion allowing for a 1% deviation from the long-term minimal cost. We find that the optimal solution is to decide to replace if error 1 occurs if we have 192 bags or less, and to decide to replace if error 2 occurs if we have 132 bags or less. Since the machine is set at a speed of 30 bags per minute, the time necessary to produce a bag is equal to 2 seconds. We set the time steps of the model equal to 2, so τ=2. We calculate the probability that no error occurs during a time step by

Table E.1: Parameter and variable values of the numerical example

Variable/Parameter	Value	Unit
τ	2	seconds
b_{new}	1000	bags
C_{down}	2000	euro/hour
C_{film}	0.10	euro/bag
λ_1	1	errors/hour
λ_2	10	errors/hour
$\mathbb{E}[T_{Recover,1}]$	150	seconds
$\mathbb{E}[T_{Recover,2}]$	200	seconds
$\mathbb{E}[T_{TotalReplace}]$	250	seconds
$\mathbb{E}[T_{Both,1}]$	200	seconds
$\mathbb{E}[T_{Both,2}]$	300	seconds
P_{Av}	0.7	

Table E.2: Transition probabilities of the numerical example

Transition probability	Value
$P_{NoError}$	0.9939
$P_{Error,1}$	$5.5386 \cdot 10^{-04}$
$P_{Error,2}$	0.0055
$P_{ErrorReplace,1}$	$3.8818 \cdot 10^{-04}$
$P_{ErrorReplace,2}$	0.0043
$P_{NoErrorReplace,1}$	$1.6636 \cdot 10^{-04}$
$P_{NoErrorReplace,2}$	0.0016
$P_{Replace}$	0.0080
$P_{NoReplace}$	0.9920
$P_{Recover,1}$	0.0133
$P_{NoRecover,1}$	0.9867
$P_{Recover,2}$	0.0100
$P_{NoRecover,2}$	0.9900
$P_{Both,1}$	0.0100
$P_{Both,2}$	0.0067
$P_{NoBoth,1}$	0.9900
$P_{NoBoth,2}$	0.9933

using equation 5.5 and the values for λ_1 and λ_2 expressed in errors per second:

$$P_{NoError} = \prod_{j \in J} [e^{-\lambda_j \cdot \tau}] = e^{-\frac{2}{3600}} \cdot e^{-\frac{20}{3600}} = 0.9939$$

Then we calculate the probability of facing error 1 after a time step using equation 5.8:

$$P_{Error,1} = \frac{\lambda_1}{\sum_{i \in J} \lambda_i} \cdot (1 - P_{NoError}) = \frac{1}{11} \cdot (1 - P_{NoError}) = 5.5455 \cdot 10^{-04}$$

Likewise, the probability for error 2 is calculated:

$$P_{Error,2} = \frac{\lambda_1}{\sum_{i \in J} \lambda_i} \cdot (1 - P_{NoError}) = \frac{10}{11} \cdot (1 - P_{NoError}) = 0.0055$$

Now that we know the probability of error occurrence, we can calculate the transition probability of going from a production state to the state replacement during error 1, given that one decides to replace if error 1 occurs:

$$P_{ErrorReplace,1} = P_{Error,1} \cdot P_{Av} = 3.8818 \cdot 10^{-04}$$

And likewise the probability of from a production state to the state replacement during error 2, given that error 2 occurs:

$$P_{ErrorReplace,2} = P_{Error,2} \cdot P_{Av} = 0.0043$$

Then the transition probability of going from a production state to the error 1 state, given that we decide to replace at the error occurrence:

$$P_{NoErrorReplace,1} = P_{Error,1} \cdot (1 - P_{Av}) = 1.6636 \cdot 10^{-04}$$

And likewise the probability of going from a production state to the error 2 state, given that we decide to replace at the error occurrence:

$$P_{NoErrorReplace,2} = P_{Error,2} \cdot (1 - P_{Av}) = 0.0016$$

When error 1 occurs and we decide to not replace the film, the system reaches one of the error 1 states. The probability of leaving this error state after one time step is:

$$P_{Recover,1} = \frac{\tau}{\mathbb{E}[T_{Recover,1}]} = \frac{2}{150} = 0.0133$$

Likewise, if we decide to not replace the film if error 2 occurs, the system reaches one of the error 2 states. We calculate the probability of leaving this state after one time step as:

$$P_{Recover,2} = \frac{\tau}{\mathbb{E}[T_{Recover,1}]} = \frac{2}{200} = 0.01$$

Complementary, the probability of still being in the error 1 state after one time step is:

$$P_{NoRecover,1} = 1 - P_{Recover,1} = 0.9867$$

Likwise, the probability of still being in the error 2 state after one time step:

$$P_{NoRecover,2} = 1 - P_{Recover,2} = 0.9900$$

$$P_{Replace} = \frac{\tau}{\mathbb{E}[T_{TotalReplace}]} = \frac{2}{250} = 0.0080$$

Then the probability of not leaving the replacement state is:

$$P_{NoReplace} = 1 - P_{Replace} = 0.9920$$

When system is in the state replacing the film during error 1, the probability of leaving this state is given by equation 5.13:

$$P_{Both,1} = \frac{\tau}{\mathbb{E}[T_{Both,1}]} = \frac{2}{200} = 0.01$$

The probability of not leaving the state is then:

$$P_{NoBoth,1} = 1 - P_{Both,1} = 0.99$$

Likewise, when system is in the state replacing the film during error 2, the probability of leaving this state is given by

$$P_{Both,2} = \frac{\tau}{\mathbb{E}[T_{Both,2}]} = \frac{2}{300} = 0.0067$$

And the probability of not leaving the state:

$$P_{NoBoth,2} = 1 - P_{Both,2} = 0.9933$$

Appendix F

Maximum Likelihood Estimation (MLE)

In this appendix, we explain the the general procedure of finding the parameters of a sampling distribution by MLE. For more explanation of the procedure, we refer to Bain & Engelhardt (1992). Let θ be the vector of parameters of the underlying distribution of our data points (also called observations) $x_1, x_2, ..., x_n$ and $f(x, \theta)$ denote the probability density function of the underlying distribution. Then the likelihood function of observing data $x_1, x_2, ..., x_n$ is given by:

$$L(\theta) = f(x_1;\theta)f(x_2;\theta)\cdots f(x_n;\theta) \tag{1}$$

We want to find the parameters that maximize the likelihood. It turns out that taking the log of the likelihood generally simplifies maximizing the function:

$$\ln\left(L(\theta)\right) = \sum_{i=1}^{n} \ln\left(f(x_i;\theta)\right) \tag{2}$$

Then we maximize the function with respect to θ. This is often done by taking the derivative with respect to θ and equating to zero:

$$\frac{d}{d\theta} ln(\ L(\theta)) = 0 \tag{3}$$

The last term yields a closed-form expression for the maximum likelihood estimators of the parameters in most cases with one or two parameters to estimate. However, we are considering distributions with a location parameter, and therefore we generally have one parameter more to optimize. This means that we are less likely to obtain closed form expressions for each of the expressions. In order to verify that we are maximizing, one should look at the second derivative. In the cases we discuss, we already know from literature that we may end up with local maxima instead, which may still have a reasonable and acceptable fit.

Appendix G

Detailed MLE procedure

In this appendix, we elaborate on the details of the fitting procedure of the replacement duration. The distribution clearly has a minimum duration, which implies that we look for a distribution with a location parameter. We decide to use MLE to fit several candidate distributions that are commonly used in literature: the Weibull distribution, the Lognormal distribution and the Uniform distribution. The estimators of the parameters of the Uniform distribution are found by taking the minimum value and the maximum value of the data set. 1. We first consider the 3 parameter Weibull distribution with a positive location parameter. The PDF of the 3 parameter Weibull distribution is given by:

$$f(x;\eta;\beta;\gamma) = \frac{\beta}{\eta}(\frac{x-\gamma}{\eta})^{\beta-1}e^{-(\frac{x-\gamma}{\eta})^{\beta}} \tag{G.1}$$

With $x \geq \gamma$, $0 \geq \gamma < \infty$, $\eta > 0$ and $\beta > 0$, where η, β and γ are called the scale, shape and location parameter respectively.
Following the general procedure of maximizing the likelihood function with respect to each of the parameters, yields the following expressions for the partial derivatives with respect to the parameters (from Teimouri & Gupta, 2013):

$$\frac{\partial ln(L(\theta))}{\partial \gamma}L(\theta) = \frac{n}{\beta} + \sum_{i=1}^{n} \ln\frac{x_i-\gamma}{\eta} - \sum_{i=1}^{n}(\frac{x_i-\gamma}{\eta})^{\beta}\ln(\frac{x_i-\gamma}{\eta}), \tag{G.2}$$

$$\frac{\partial ln(L(\theta))}{\partial \beta} = -\frac{n\beta}{\eta} + \frac{\beta}{\eta}\sum_{i=1}^{n}(\frac{x_i-\gamma}{\eta})^{\eta}, \tag{G.3}$$

$$\frac{\partial ln(L(\theta))}{\partial \eta} = -(\beta-1)\sum_{i=1}^{n}\frac{1}{x_i-\gamma} + \frac{\beta}{\eta}\sum_{i=1}^{n}(\frac{x_i-\gamma}{\eta})^{\beta-1}. \tag{G.4}$$

With n denoting the total number of observations and x_i the i-th observation.
Equating to zero does not result in closed-form expressions for the estimators that maximize the likelihood function. In literature several approaches to find these parameters are proposed. Two commonly used approaches are (i) optimization of the three parameters simultaneously through search algorithms, (ii) estimating the location parameter with an alternative approach, whereafter the MLEs for the scale and shape parameter can be found through numerically optimization. We decide to follow the latter since this approach can be implemented very practically. For the estimation of the location parameter, we decide to use two estimations: the standard estimation in literature and an estimation suggested by Shifley & Lentz (1985). The standard estimation used in literature is setting γ equal to the minimum of the observations: $min(x_1, .., x_n)$. Smith (1985) shows that this estimator is consistent and no other estimator converges at a faster rate if $\beta < 2$. Using exactly the minimum of the sample results in the natural log of 0, which is undefined. Reliability HotWire (2013) recommends to choose a number arbitrarily close to the minimum of

the sample. We therefore take $\hat{\gamma_1} = 0.999 * 53 \approx 52.95s$. The second method acknowledges that there may be unrecorded observations smaller than the smallest observation and suggests $\gamma = \frac{x_1 x_n - x_2^2}{x_1 + x_n - 2x_2}$ (Shifley & Lentz, 1985). This results in $\hat{\gamma_2} = 52.6s$.
Now that γ is assumed to be known, we are essentially looking for the MLEs of the parameters of a two parameter Weibull with the data points shifted to the left with the value $\hat{\gamma}$. That means that we calculate $x_i' = x_i - \hat{\gamma}$ for each individual data point x_i, and use the x_i' to calculate the MLEs for the two parameter Weibull. The PDF simplifies to:

$$f(x;\eta;\beta) = \frac{\beta}{\eta}(\frac{x}{\eta})^{\beta-1}e^{-(\frac{x}{\eta})^{\beta}} \tag{G.5}$$

Repeating the general procedure of finding the MLEs of the 2 parameter Weibull yields the following expressions for the partial derivatives of the scale and shape parameter:

$$\frac{\partial ln(L(\theta))}{\partial \eta} = n\eta - \sum_{i=1}^{n} \ln {x_i}^{\beta} \tag{G.6}$$

$$\frac{\partial ln(L(\theta))}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^{n} \ln x_i - \frac{1}{\eta}\sum_{i=1}^{n} x_i^{\beta} \ln x_i \tag{G.7}$$

Equating to zero and rearranging then yields a closed-form expression for $\hat{\eta}$ in terms of β and a expression for β:

$$\hat{\eta} = \frac{\sum_{i=1}^{n} x_i^{\hat{\beta}}}{n} \tag{G.8}$$

$$\frac{n}{\hat{\beta}} + \sum_{i=1}^{n} \ln x_i = \frac{1}{\hat{\eta}}\sum_{i=1}^{n} x_i^{\hat{\beta}} \ln x_i \tag{G.9}$$

Then filling in the expression for $\hat{\eta}$ in the second expression and multiplying by $\hat{\beta}$ yields:

$$n + \hat{\beta}\sum_{i=1}^{n} \ln x_i = n\hat{\beta}\frac{\sum_{i=1}^{n} x_i^{\hat{\beta}} \ln x_i}{\sum_{i=1}^{n} x_i^{\hat{\beta}}} \tag{G.10}$$

The last expression can be computed numerically for $\hat{\beta}$, which determines $\hat{\eta}$. We do the computation in Matlab 2019a.

2. Secondly, we elaborate on the MLEs of the 3 parameter Lognormal distribution. The PDF is given by:

$$f(x;\mu;\sigma;\gamma) = \frac{1}{(x-\gamma)\sigma\sqrt{2\pi}}e^{-\frac{(\ln(x-\gamma)-\mu)^2}{2\sigma^2}} \tag{G.11}$$

Now following the general procedure yields (from Cohen & Whitten (1980)):

$$\frac{\partial \ln(L(\theta))}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}[\ln(x_i - \gamma) - \mu] \tag{G.12}$$

$$\frac{\partial ln(L(\theta))}{\partial \sigma} = \frac{n}{\sigma} + \frac{1}{\sigma^2}\sum_{i=1}^{n}[ln(x_i - \gamma) - \mu]^2 \tag{G.13}$$

$$\frac{\partial ln(L(\theta))}{\partial \gamma} = \frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{[ln(x_i - \gamma) - \mu]}{(x_i - \gamma)} + \sum_{i=1}^{n}(x_i - \gamma)^{-1} \tag{G.14}$$

Differently from the 3 parameter Weibull distribution, equating to zero yields a closed form expression for μ and σ in terms of γ:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} ln(x_i - \hat{\gamma}) \tag{G.15}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} ln^2(x_i - \hat{\gamma}) - [\frac{1}{n}\sum_{i=1}^{n} ln(x_i - \hat{\gamma})]^2 \tag{G.16}$$

Substituting the μ and σ for their closed-form expressions in the equation for γ, yields an expression for γ that can be evaluated numerically:

$$\begin{aligned}&[\sum_{i=1}^{n}((x_i) - \gamma)^{-1}][\sum_{i=1}^{n} ln(x_i) - \gamma) - \sum_{i=1}^{n} ln^2(x_i) - \gamma)\\ &+ (\frac{1}{n}(\sum_{i=1}^{n} ln(x_i - \gamma))^2] - n\sum_{i=1}^{n}\frac{ln(x_i - \gamma)}{x_i - \gamma} = 0\end{aligned} \tag{G.17}$$

For practical use, Cohen & Whitten (1980) recommend to find an estimator for γ by evaluating this expression over the left hand side of $min(x_1, ..., x_n)$. We do this for the range of $\gamma = 0$ until $\gamma = min(x_1, ..., x_n)$, with steps of 0.1 and observe that we do not obtain a solution. Griffiths (1980) as cited in Kowarik (n.d.), provides an expression for the log likelihood as a function of γ, which we evaluate over the same interval. This shows that the likelihood is increasing in γ. We expect the optimal value for γ to be arbitrarily close to $\gamma = min(x_1, ..., x_n)$, thus we decide to take $\hat{\gamma} = 0.999 * 53 \approx 52.95s$. Then μ and σ follow from equations G.15 and G.16.

Appendix H

Implementation of the Opportunistic Film Replacement Model

In this appendix, we elaborate on the implementation of the opportunistic film replacement model of chapter 5 and we explain the use of the model. The model is implemented in Matlab 2019. The model is implemented for two errors, which means that it is applicable to situations in which the film can be replaced during one and two errors. The model is also applicable to situations with only one error by setting the arrival rate of the second error to 0. In the following sections, we discuss the input parameter section and the output section.

Parameter section

The input parameters that have to be specified are shown in figure H.1. The parameter section is marked with a comment ‘Parameter section’. In order to get the correct model output, one needs to specify each of the parameters of the model in the correct unit of measurement. The unit of measurement is described in a comment on the same line as the parameter. In table H.1 we show for each of the parameters (i) the corresponding variable of the model in chapter 5, (ii) the unit of measurement and (iii) a brief description of the parameter. If one wants to model a situation with only one error that allows to preventively replace the film, the Lambda_2 should be set to zero.

```matlab
%---------------------- Parameter section ---------------------%

B_new = 6700;           %Amount of bags on a new film; In bags
C_down = 2000;          %Downtime cost per hour; In euros
C_bag = 0.10;           %Film cost per bag; In euros
replaceTime = 250;      %Time necessary to replace the film; In seconds
recoverTime1 = 150;     %Downtime duration of error 1; In seconds
recoverTime2 = 200;     %Downtime duration of error 2; In seconds
bothTime1 = 200;        %Duration of film replacement during error 1; In seconds
bothTime2 = 300;        %Duration of film replacement during error 1; In seconds
Lambda1 = 1;            %Arrival rate of error 1; In errors per hour
Lambda2 = 10;           %Arrival rate of error 2; In errors per hour
Availability = 0.7;     %Probability that an operator is available at occurrence error 1 or 2; (Unitless)

tau = 2;                 %Timesteps, set equal to time to produce 1 bag; In seconds
error = 0.01;           %Allowed deviation from the optimal long-term cost per timestep (Default 0.01); (Unitless)
```

Figure H.1: Input parameters of the model

Parameter	Model notation	Unit	Description
B_new	b_{new}	Bags	Amount of bags
C_down	C_{down}	Euro	Cost per hour of downtime
C_bag	C_{film}	Euro	Cost per bag
replaceTime	$T_{Replace}$	Seconds	Expectation of the total replacement time
recoverTime1	$T_{Recover,1}$	Seconds	Expectation of the time necessary to resolve error 1
recoverTime2	$T_{Recover,2}$	Seconds	Expectation of the time necessary to resolve error 2
bothTime1	$T_{Both,1}$	Seconds	Expectation of the downtime if an available operator replaces the film during error 1
bothTime2	$T_{Both,2}$	Seconds	Expectation of the downtime if an available operator replaces the film during error 2
Lambda1	λ_1	Errors per hour	Arrival rate of error 1
Lambda2	λ_2	Errors per hour	Arrival rate of error 2
Availability	P_{Av}	Unitless	Probability that an operator is available at the moment an error occurs
tau	τ	Seconds	Discrete timesteps of the model
error	ϵ	Unitless	Proportion the solution may deviate at most from the optimal long-term cost per time step

Table H.1: Parameter names in the model implementation

Output section

As soon as the model finds the solution at the prespecified deviation from the optimal cost per time step, the model stops. The console shows boundary values for error 1 and error 2. If one faces the corresponding error with an amount of bags on the film less than or equal to this boundary value, one should decide to replace the film. Furthermore, the model finds an upper bound and a lower bound for the optimal long-term average cost per time step. The model takes the average of the upper and lower bound and shows the long-term average cost per hour in the console. The upper bound, M_n and the lower bound, m_n, can be found in the Matlab workbench. We show an example of the output in figure H.2.

```
Elapsed time is 58.111911 seconds.
b_bound1 is equal to 247 bags
b_bound2 is equal to 165 bags
Long-term average cost per hour is 757.9902 euro
>>
```

Figure H.2: Output of the model implementation

Appendix I

Validation of the number of simulation runs

In this appendix we elaborate on the determination of the amount of runs of the simulation. The simulation is used to determine the amount of cost per time unit and the amount of downtime per time unit. In order to determine the necessary amount of runs we follow the procedure suggested by Boon et al. (2017). In this procedure, one first estimates the standard deviation of the variable of interest, by using a short initial run with a relatively small amount of runs. Assuming that this estimation is representative for the real standard deviation of the variable, one can then use the estimation of the standard deviation to calculate how much runs are necessary to get the desired confidence intervals. One can calculate the desired confidence intervals by using the implications of the central limit theorem, which states that as you take k samples from a random variable Z and calculate the mean of these samples $\bar{Z}$, the distribution of the sample means, $\bar{Z}$, approaches a normal distribution. How closely the distribution of $\bar{Z}$ approaches a normal distribution depends on the size of k and the number times you repeat the drawing of k samples. Using the implications of this theorem, we can calculate the amount of times we have to repeat the drawing of k samples to get our half-width confidence interval, given that we know our desired confidence level and the desired length of the confidence interval:

$$\left(\frac{Z_{a/2} \cdot \sigma}{\epsilon}\right)^2 < n \tag{I.1}$$

In which $Z_{a/2}$ is the factor for the confidence level, σ the standard deviation of the variable of interest, n the number of times we take k samples from Z and ϵ the desired half-width of the confidence interval. The confidence level is expressed as a percentage. This percentage can be interpreted as if you would repeat this procedure many times, the mean would lie in the interval for this percentage of the repetitions. In literature, a 95% confidence level is often considered acceptable, so we decided to use this level. The value for $Z_{a/2}$ for a 95% confidence level is 1.96. In the following, we elaborate on how we estimated the σ. For more explanation of the central limit theorem and its implications for sampling, we refer to Boon et al. (2017) and Bain & Engelhardt (1992).

In order to estimate the standard deviation of the downtime per time unit and the cost per time unit, we simulated a long production period. By manually varying the length of this production period, we found that 10^4 hours of production can be simulated within a couple of seconds per run, which we consider convenient. In order to estimate the standard deviation of the variables of interest, we chose an initial amount of runs. If one uses a very low amount of runs, the standard deviation might vary much from the actual standard deviation. If this amount is too high, the computation will take too long. We decided to estimate the standard deviations for each of the parameter settings, using 25 runs. We observed that the standard deviation of the downtime per hour did not differ much between the different b_{bound} and their corresponding P_{Av} and C_{down}. The standard deviation of the total cost per hour increases approximately linear in the C_{down}. For both the standard deviation of the downtime per hour and the total cost per hour, we took the highest value. By doing do, we have a pessimistic estimation of the real value. This results in $\sigma^{Downtime} = 0.6 \cdot 10^{-3}$ (hours downtime per hour production) and $\sigma^{TotalCost} = 1.5$ euro. In order to determine a reasonable length for the confidence intervals, we consider how close the means of the cost per hour and the downtime per hour are. We find that the smallest difference for the total cost per hour is approximately 2 €/hour and the smallest difference for the downtime per hour is approximately $5 \cdot 10^{-4}$ hours per hour. In order to prevent the confidence intervals from overlapping, we would need the bound of the interval on a distance of $\epsilon_1 = 1$ €/hour for the cost per hour and a bound of $\epsilon_2 = 2.5 \cdot 10^{-4}$ hours for the bound on the downtime per hour.

Now that we obtained our σ, the desired ϵ and we know the confidence level, we can fill in the equation I.1. This results in a number of runs for the downtime, n_{down} of:

$$\left(\frac{Z_{a/2} \cdot \sigma^{DownTime}}{\epsilon_1}\right)^2 < n_{down} = 22.13... \approx 23$$

And for the number of runs for the cost per downtime, n_{cost}:

$$\left(\frac{Z_{a/2} \cdot \sigma^{TotalCost}}{\epsilon_2}\right)^2 < n_{cost} = 15.36... \approx 16$$

We conclude that the amount of 25 runs with a simulation of 10^4 production hours is sufficient to get the desired bounds. Due to the non-sequential order of validating the length and the amount of runs of the simulation, we already ran the simulation with a length of 10^5 hours for 25 runs. Considering the results of this validation, we could have used less runs. However, the results have tighter confidence intervals and are valid, so we use the results of this more lengthy simulation.

Appendix J

Validation and verification of the discrete event simulation

In this appendix, we elaborate on the validation and the verification of the simulation model. The simulation model is implemented in Matlab 2019, since this software is available at Bosch. In addition, Matlab is convenient for incorporating random number generators that follow specified theoretical distributions.

Validation of the simulation model

The discrete simulation model of chapter 6, can be seen as a more general form of the model in chapter 5. Similarly to the MDP model, the simulation model assumes constant error rate during production and that the operator availability can be modeled as a fixed probability of having an operator available at the moment an error occurs. We already discussed the validation for the MDP model, and we do not repeat the same arguments. The discrete simulation model also evaluates the model at discrete time steps. However, the length of a time step is no longer fixed, but depends on the first upcoming event. The time between events are based on draws from random number generators that follow the theoretical distributions found in the data. We expect that the durations in the simulation model represent reality better, since the durations are based on theoretical distribution that describe the data.

Verification of the simulation model

We verify the correct implementation of the simulation model in two ways. First, we use the debugging function of Matlab to follow the code of the simulation line by line. In this way, we verify that each step is followed by the correct subsequent step and that the corresponding calculations of the variables are correctly implemented. Second, we save the interim results of the total costs, C_{Total}, the costs of the disposed film, $C_{Replace}$, and the amount of downtime, d, per film. We verify that the model never disposes more bags than the b_{bound} and the values for the downtime per film appear to be reasonable. With reasonable values we mean that the values vary around an approximation of the expected downtime. For this approximation we take the expected amount of errors between b_{new} and b_{bound}, times the expected duration of the errors $\mathbf{E}[T_{Recover}]$, plus the average of $\mathbf{E}[T_{TotalReplace}] + \mathbf{E}[T_{Both}]$.

$$\hat{T} = (b_{new} - b_{bound}) * P_{Error,160} * \mathbf{E}[T_{Recover}] + \frac{\mathbf{E}[T_{TotalReplace}] + \mathbf{E}[T_{Both}]}{2}$$

We use this approximation instead of the exact value since the approximation is more convenient to calculate manually.

Appendix K

Simulation results

In this appendix, we show the results of the simulation procedure of chapter 6. Considering the small differences in downtime and cost, we show the results per 100 hours of simulation.

Table K.1: Results of the simulation: the average downtime resulting from the new policy in hours per 100 hours of simulation

	Cost scenario 1		Cost scenario 2		Cost scenario 3	
P_{Av}	d	CI_{95}	d	CI_{95}	d	CI_{95}
0.1	11.35	[11.35 - 11.35]	11.32	[11.32 - 11.33]	11.27	[11.26 - 11.28]
0.2	11.28	[11.28 - 11.28]	11.24	[11.23 - 11.24]	11.14	[11.14 - 11.15]
0.3	11.22	[11.22 - 11.22]	11.16	[11.15 - 11.17]	11.04	[11.03 - 11.04]
0.4	11.16	[11.16 - 11.16]	11.08	[11.07 - 11.09]	10.95	[11.95 - 11.96]
0.5	11.11	[11.11 - 11.11]	11.02	[11.01 - 11.02]	10.87	[11.86 - 11.87]
0.6	11.07	[11.07 - 11.07]	10.97	[11.96 - 11.97]	10.80	[11.80 - 11.81]
0.7	11.02	[11.02 - 11.02]	10.91	[11.91 - 11.92]	10.73	[11.73 - 11.74]
0.8	10.98	[10.98 - 10.98]	10.86	[10.85 - 10.87]	10.69	[10.68 - 10.69]
0.9	10.94	[10.94 - 10.94]	10.82	[10.82 - 10.83]	10.64	[10.64 - 10.65]
1	10.90	[10.90 - 10.90]	10.78	[10.78 - 10.79]	10.59	[10.59 - 10.60]

Table K.2: Results of the simulation: the average total costs resulting from the new policy in euro per 100 hours of simulation

	Cost scenario 1		Cost scenario 2		Cost scenario 3	
P_{Av}	C_{Total}	CI_{95}	C_{Total}	CI_{95}	C_{Total}	CI_{95}
0.1	11484	[11479 - 11490]	17197	[17187 - 17207]	33885	[33867 - 33903]
0.2	11446	[11441 - 11452]	17130	[17120 - 17139]	33667	[33653 - 33681]
0.3	11415	[11409 - 11420]	17067	[17059 - 17075]	33468	[33450 - 33486]
0.4	11383	[11377 - 11389]	16991	[16983 - 16999]	33309	[33293 - 33324]
0.5	11354	[11349 - 11358]	16936	[16927 - 16944]	33127	[33109 - 33144]
0.6	11329	[11321 - 11336]	16891	[16885 - 16898]	33000	[32980 - 33019]
0.7	11298	[11292 - 11304]	16843	[16834 - 16851]	32838	[32820 - 32855]
0.8	11273	[11267 - 11278]	16783	[16774 - 16792]	32730	[32712 - 32748]
0.9	11251	[11246 - 11257]	16745	[16737 - 16754]	32632	[32614 - 32649]
1	11226	[11221 - 11232]	16703	[16694 - 16712]	32498	[32479 - 32517]

Table K.3: Results of the simulation of the old policy (average of 100 hours of production

Variable	Expectation	$CI_{95\%}$
d	11.42 hour	[11.42 - 11.42]
C_{Total} (C_{down} = 1000 €/ hour)	€11422	[11419 - 11425]
C_{Total} (C_{down} = 1500 €/ hour)	€17133	[17129 - 17137]
C_{Total} (C_{down} = 3000 €/ hour)	€34266	[34258 - 34274]