MASTER

Improving Spare Parts Demand Forecasting at ASML with Installed Base Information

van den Oord, W.D.

Award date:
2022

Kerkdriel, December 2021

Improving Spare Parts Demand Forecasting at ASML with Installed Base Information

BSc. W.D. van den Oord
1002271

In partial fulfilment of the requirements for the degree of

Master of Science
In Operations Management and Logistics

University Supervisors
Dr. K.H. van Donselaar, Eindhoven University of Technology, OPAC
Dr. Ir. R.J.I. Basten, Eindhoven University of Technology, OPAC

Company Supervisors
Ir. R. van Hugten, ASML, Service Inventory Management
Ir. J. Stip, PDEng, ASML, Service Inventory Management

TU/e School of Industrial Engineering

Series Master Theses Operations Management and Logistics

Keywords: Spare Parts, Demand Forecasting, Installed Base Information, Reliability-Based Forecasting, Weibull Distribution

Abstract

Spare parts demand forecasts play a vital role in the planning of spare part basestock levels across the service supply chain of ASML. However, ASML faces a challenge to accurately forecast changing spare parts demand patterns caused by installed base changes in this supply chain. Moreover, the time series-based forecasting benchmarks currently available in the literature are also unable to correctly anticipate these changing demand patterns.

Demand for spare parts generally follows the demand for installed systems, but with a delay. This delay between the moment of install and failure can be described by the failure behaviour of a spare part. In this research, spare part failure observations from the past are used to generalize this failure behaviour with a Weibull distribution. This distribution can be used to estimate the failure probability of an active part, during a given forecast horizon. A reliability-based forecasting model is proposed to translate these failure probabilities of all active parts into a spare parts demand forecast.

Both a simulation and an empirical study show that the reliability-based forecasting model can outperform the benchmarks from the literature when confronted with changing demand patterns. Therefore, this research came up with a forecasting methodology that can more accurately predict the occurrence of spare part demand, based on information derived from the installed base.

Executive Summary

This master thesis project is conducted at ASML, the world's leading provider of advanced lithography systems for the semiconductor industry. In this executive summary, an overview is provided of the most important research outcomes for ASML and its spare parts demand forecasting processes.

Problem Definition

ASML and its customers draw up service contracts to limit the costly consequences of machine downtime. The Service Inventory Management (SIM) team at ASML is concerned with the tactical planning of spare parts needed to meet these service contracts in a cost-efficient way. Spare parts demand forecasts are an important driver for these planning activities and have a large impact on the customer's material availability. Currently, spare parts demand forecasting at ASML mostly depends on historic demand data. However, Dekker et al. (2013) conclude that spare parts control requires a level of forecast accuracy that cannot be provided by the extrapolation of historic demand only. In addition, Van der Auweraer et al. (2019) denote that the use of information about the installed base has the potential to considerably increase the forecasting accuracy of spare parts demand. The installed base is defined as the set of machines for which ASML provides after-sales services. Over the years, ASML gathered and registered information about this installed base in its role as Original Equipment Manufacturer (OEM). This information now has the potential to improve the customer's material availability. Therefore, the main research question of this master thesis is defined as:

How can the forecast accuracy in the current framework for forecasting and planning of spare parts demand at ASML be improved by using installed base information?

To answer this research question, the current practices are analysed and a new forecasting methodology is proposed based on the installed base information available at ASML. Moreover, the value of this methodology is tested in terms of forecast accuracy by a simulation and empirical study.

Analysis & Diagnosis

The current forecasting process produces spare parts demand forecasts based on the demand over the last three years. First, this process generates periodical demand rates per system, which describe the average demand per machine during a given period in these three years. Next, a demand rate forecast is derived by exponentially weighting these demand rates. Finally, a spare part demand forecast is produced by multiplying this demand rate with the current number of active machines. This process is referred to as an Enriched Exponentially Weighted Moving Average (Enriched EWMA).

The development of spare parts demand at ASML is strongly related to changes in the installed base. In general, spare parts demand follows the demand for installed systems, but with a delay (Dekker et al., 2013). Thus, the constant growth of the installed base at ASML results in service parts with an increasing demand pattern. The forecast performance at ASML has been analysed to understand the impact of this growth. This analysis distinguished spare parts with and without a strictly increasing demand pattern. In addition, time series-based forecasting methods from the literature were included as a benchmark. The forecast performance in this analysis is measured by the scaled mean error (SME) and mean absolute scaled error (MASE), which measure the forecast bias and error respectively. The forecast bias describes the tendency of a forecasting method to be consistently higher or lower than the actual demand, whereas the forecast error expresses the absolute difference between a forecast and the actual demand. The combination of forecast bias and error gives a representative overview of the overall forecast performance when considering spare parts demand (Teunter & Duncan, 2009). Table 1 summarizes the forecast performance of the Enriched EWMA and three of the benchmarks included: Simple Moving Average (SMA), Double Exponential Smoothing (DES) and the Syntetos-Boylan Approximation (SBA).

Table 1: Out-of-sample forecast performance for service parts with and without strictly increasing demand

	Strictly Increasing		Not Strictly Increasing	
	SME	**MASE**	**SME**	**MASE**
SMA	0.207	1.037	0.015	0.907
DES	0.012	1.150	0.064	1.399
SBA	0.280	0.931	0.169	0.671
Enriched EWMA	0.146	1.070	-0.068	1.032

On the one hand, 19% of the service parts at ASML currently deal with a strictly increasing demand pattern. Table 1 indicates that no time series-based forecasting method is currently able to correctly handle both forecast bias and error when dealing with an increasing demand pattern. On the other hand, 81% of the service parts at ASML still deal with a more stable demand pattern. In this case, the Enriched EWMA slightly over-forecasts the actual demand. The SMA indicates that an equal division of weights could already deliver significant improvements both in terms of forecast bias and error.

Solution Design

A reliability-based forecasting model (RFM) has been derived to more accurately capture changing demand patterns. The RFM is based on the relationship of the delay between the installed base and spare parts demand changes, as stated by Dekker et al (2013). This delay between the moment of install and failure can be described by the failure behaviour of a spare part. The RFM uses spare part failure observations from the past to generalize this failure behaviour with a Weibull distribution. This distribution is used to estimate the failure probability of an active service part during a forecast horizon. Service part registrations at ASML help to create an overview of all active service parts in the field, in combination with their age. In the end, the RFM derives a spare part demand forecast by aggregating the failure probabilities of these installed service parts.

An important element of the RFM is the estimation of the Weibull parameters. This research introduced two estimation procedures. First, the lifetime-driven approach traditionally fits the parameters based on the lifetimes of (failed) service parts. Moreover, the demand-driven approach is a more robust alternative, based on the timing of service part failures and the demand it triggered.

Implementation

The forecast performance of the RFM has been tested during both a comparative simulation and an empirical study. The simulation study showed that the RFM outperformed the time series-based benchmarks when confronted with changing demand patterns. This result applies to the RFM for both Weibull parameter estimation procedures. Furthermore, the empirical study tested to what extent these results could be replicated in practice. A representable subset of service parts with a strictly increasing demand pattern has been created to evaluate the forecast performance in practice. In addition, the Enriched EWMA, DES and SBA were included as benchmarks. Figure 1 gives an overview of the forecast performance in terms of forecast error and bias. In general, the demand-driven RFM outperforms the benchmarks from the literature in terms of forecast error and bias. Most importantly, this demand-driven RFM delivers significant forecast error and bias improvements in comparison with the current practices of the Enriched EWMA. The lifetime-driven approach struggles with data availability in practice, whereas the demand-driven approach proves to be more robust.

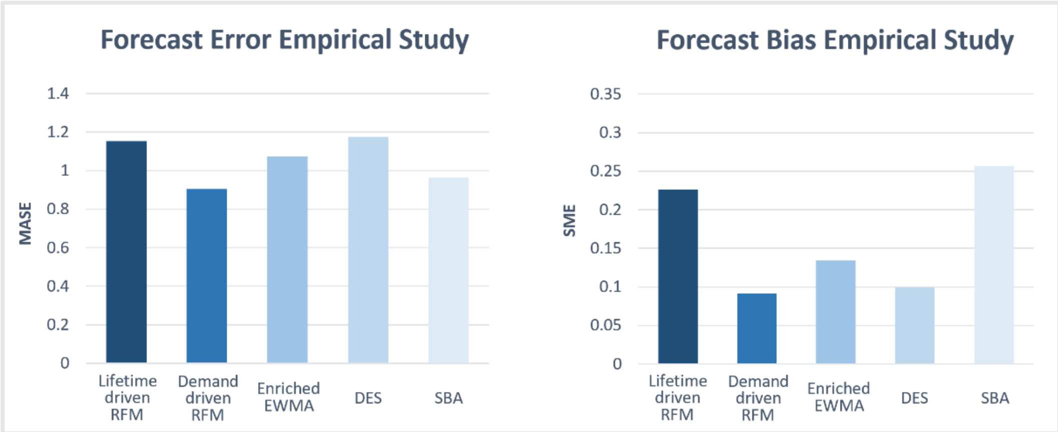

Figure 1: Forecast error and bias during the comparative empirical study

Conclusions & Recommendations

The RFM provides an answer to the main research question. This reliability-based model can improve the forecast accuracy of spare parts demand forecasting at ASML by the use of installed base information. On average, the RFM reduces the average forecast error for service parts with an increasing demand pattern by 15.8%, whereas the forecast bias is reduced by 31.9%. Therefore, it is recommended to implement the RFM for service parts with an increasing demand pattern. The potential value of this model increases when more service parts start to deal with increasing demand patterns due to installed base changes at ASML. Moreover, the RFM offers plenty of opportunities to build on this work to extract even more value from installed base information. In specific, the RFM creates valuable insight into the installed base lifecycle and part failure behaviour. These insights can enhance tactical and strategical decision-making beyond the scope of spare parts planning at ASML.

In addition, this research also showed how traditional time series-based forecasting techniques can still realize significant forecast performance improvements for the remaining set of service parts with a more stable demand pattern (Table 1). In this case, on average, the forecast error and bias of these service parts can be decreased by at least 12.1% and 77.9% respectively.

Preface

The report in front of you marks the end of my master's study Operations Management & Logistics at the Eindhoven University of Technology. ASML offered me the opportunity to perform my master thesis at the cutting edge of technology and innovation. It was a truly great experience to graduate from a world-leading company where everything revolves around progress. I am proud to say I could contribute to this progress by writing this master thesis, named: *'Improving Spare Parts Demand Forecasting at ASML with Installed Base Information'*. However, this would not have been possible without the help of numerous people.

First, I would like to express my gratitude to Karel van Donselaar, my first supervisor, for his guidance, support and enthusiasm throughout these exciting times. Moreover, thanks to Rob Basten for his assistance in the role of second supervisor. The feedback from both of you is greatly appreciated and really helped me to further improve the quality of my research.

Furthermore, I would like to thank all people at ASML that helped me during my stay. In specific, I want to express my gratitude to my company supervisors Roy van Hugten and Joan Stip. I really appreciated all of your help and support during the project. Moreover, a special thanks to all members of the Service Inventory Management team, who made me feel welcome.

Last but not least, I would like to thank my family and friends, who always supported me and allowed me to get the most out of it. With this master thesis there comes an end to my academic career. Now it is time to put all of the knowledge into practice and say goodbye to my student life. It has been a pleasure.

Ward van den Oord

Kerkdriel, December 2021

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Definition
ADI	Average Demand Interval
ADS	Average Demand Size
BOM	Bill of Material
CAM	Configuration as Maintained
CDF	Cumulative Distribution Function
CRO	Croston's Method
CSCM	Customer Supply Chain Management
CV^2	Squared Coefficient of Variation
DES	Double Exponential Smoothing
DUV	Deep Ultraviolet light
EUV	Extreme Ultraviolet light
EWMA	Exponentially Weighted Moving Average
FSD	Field Service Defect
IFR	Initial Failure Rate
MAD	Mean Absolute Deviation
MASE	Mean Absolute Scaled Error
ME	Mean Error
MLE	Maximum Likelihood Estimator
MSE	Mean Squared Error
OEM	Original Equipment Manufacturer
PDF	Probability Density Function
PHM	Proportional Hazard Model
RFM	Reliability-based Forecasting Model
RW	Random Walk
SBA	Syntetos – Boylan Approximation
SCM	Supply Chain Management
SES	Single Exponential Smoothing
SIM	Service Inventory Management
SMA	Simple Moving Average
SME	Scaled Mean Error
TSB	Teunter Syntetos Babai method
ZF	Zero Forecast

1 Introduction

Every system built by humans is unreliable in the sense that it degrades with age and/or usage. Moreover, a system fails when it is no longer able to deliver the designed outputs. These failures can be catastrophic with serious economic losses as a result, depending on the complexity and value of the capital assets involved (Boylan & Syntetos, 2008; Öner et al., 2007). Complex and high-value assets are crucial elements in the production processes of the semiconductor industry. This industry is the aggregate of companies involved with the design and production of semiconductors. System breakdowns in this industry generally cause large production problems. Spare parts management is essential to minimize the length of costly downtime periods (Van Houtum & Kranenburg, 2015).

This master thesis considers the role of spare parts demand forecasting and the potential of including explanatory factors derived from the installed base. The research is carried out at ASML, a world-leading player in the semiconductor industry. First, this chapter gives a brief introduction to the company and its business activities. The remainder of this chapter respectively describes the problem definition, research objective, research questions, scope and research methodology.

1.1. Company Description

ASML originated in 1984 as a joint venture between Philips and Advanced Semiconductor Materials International (ASMI). Throughout the years ASML managed to become the world's leading provider of advanced lithography systems for the semiconductor industry. The company develops, produces and sells complex machines that are crucial for the production of integrated circuits for microchips. Moreover, ASML has multiple types of advanced systems that use Deep Ultraviolet (DUV) and Extreme Ultraviolet (EUV) lithography. The semiconductor industry is driven by affordable scaling, which is the ability to make smaller transistors at the right price. Moore's law predicts the number of transistors on an integrated circuit doubles every two to three years, meaning the computing power increases exponentially while keeping the costs affordable. The goal of ASML is to help the industry continue Moore's law (ASML, 2021).

ASML is headquartered in Veldhoven, the Netherlands, and operates globally with activities in Europe, Asia and the United States. More specifically, ASML has over 60 locations in 16 countries. This ranges from manufacturing and customer support sites to R&D facilities. However, most of the ASML locations across the world serve as a (local) warehouse containing spare parts and service tools for lithography systems. Among its customers, there are major global semiconductor players like Samsung, Intel and Taiwan Semiconductor Manufacturing. In 2020, ASML realized a turnover of about €14 billion (ASML, 2021). ASML's business strategy is reflected by close cooperation with suppliers and customers. Given this strategy, aftersales events become increasingly relevant for ASML to acquire a competitive advantage. An important part of the aftersales activities at ASML concerns service operations and system maintenance. This is described as installed base management and accounted for €3.6 billion of the €14 billion turnover in 2020 (ASML, 2021). Therefore, after-sales services are not only valuable for competitive advantages but also generate remarkably high direct revenues. With service contracts, ASML provides a flexible solution to meet the customers' needs. The Service Inventory Management (SIM) team at ASML is concerned with the tactical planning of spare parts and service tools needed to meet these service contracts.

1.2. Problem Definition & Research Objective

For ASML to maintain its leading global technology position in the semiconductor industry, it keeps on innovating and needs to provide its customers with a high level of service. The implementation of advanced lithography technology requires large investments of the customers. Moreover, the unavailability of lithography systems usually brings an entire production process to a standstill, resulting in serious economic losses. Therefore, when a system breaks down, corrective maintenance action is needed as soon as possible. Part of the unavailability time consists of waiting for the required spare parts and service tools to become available. Considering the high downtime costs, the relevance of having the right materials available at the right time for the right costs becomes evident. ASML and its customers draw up performance-based service contracts to limit the negative consequences of downtime. This makes ASML responsible for obtaining the service level agreements as denoted in the contract. Failure to meet these agreements results in penalty costs and reputational damage for ASML. Currently, spare parts demand forecasts are an important driver for the planning activities of the SIM team. In this way, the forecasts have a large impact on the customer's material availability. The more accurate the forecasts, the more cost-efficient the service agreements with the customers can be met.

There are several aspects of spare parts which make demand management a complex task (Bacchetti & Saccani, 2012). Most spare parts have to deal with a slow and intermittent demand pattern, which makes it hard to predict. Additionally, the number of parts that need to be managed is in general very high. These aspects result in difficulties for service logistics to find the right balance between inventory holding, stock-out and obsolescence costs, while offering competitive service contracts (Dekker et al., 2013). Currently, spare parts demand forecasting at ASML mainly depends on historic demand data. Dekker et al. (2013) conclude spare parts control requires a level of forecast accuracy that cannot be provided by the extrapolation of historic demand only. The use of installed base information, like part failure behaviour and machine details, has the potential to considerably increase the forecasting accuracy of spare parts demand (Van der Auweraer, Boute & Syntetos, 2019). Currently, only the installed base size is included in the forecasts, while other underlying demand-causing factors are ignored. Over the years, ASML gathered and registered information about the installed base in its role as Original Equipment Manufacturer (OEM). This information now has the potential to improve the customer's material availability. Therefore, this master thesis looks into this potential by finding the best way to include installed base information in the current framework of forecasting and planning at ASML. Thus, the research objective of this master thesis is:

'Develop a forecasting method that improves the forecast accuracy in the current framework for forecasting and planning of spare parts demand at ASML, using installed base information.'

1.3. Research Questions

In this section, the main research question and sub-research questions of this study are presented. In accordance with the research objective, the main research question is defined as:

How can the forecast accuracy in the current framework for forecasting and planning of spare parts demand at ASML be improved by using installed base information?

Before improvements based on installed base information can be made, first it is needed to understand and outline the current framework for forecasting and planning at ASML. Subsequently, the forecast performance of the current framework can be defined. Therefore, the first and second research questions are described as:

1. ***How does the current framework for forecasting and planning spare parts at ASML operate?***
2. ***What is the forecast performance in the current framework for forecasting and planning spare parts at ASML?***

In general, spare parts are highly varied, with different service requirements, costs and demand patterns (Boylan & Syntetos, 2008). Classification of spare parts can facilitate spare parts demand forecasting and help identify areas for improvement (Bacchetti & Saccani, 2012). Therefore, the third research question is denoted as:

3. ***How can spare parts be classified to facilitate spare parts demand forecasting at ASML?***

ASML has been registering information about the installed base in its role as OEM for years. This information has the potential to considerably improve the forecasting accuracy of spare parts demand according to, among others, Dekker et al. (2013) and Van der Auweraer et al., (2019). The fourth research question looks into this potential by finding a way to include relevant elements of installed base information into a new forecasting methodology:

4. ***How can installed base information be used in a new forecasting method for spare parts demand at ASML?***

The last step is to find out whether a new forecasting method, based on installed base information, can fulfil its potential. The forecast performance of a new methodology can be directly compared to the current forecasting method and other benchmarks from the literature. Therefore, the last research is denoted as:

5. ***How does the new forecasting method based on installed base information perform in the framework for forecasting and planning of spare parts at ASML?***

1.4. Research Scope

The duration of this master thesis is limited, which implies the importance of a realistically bounded scope in order to fulfil the research objective. A clear scope is essential to answer the research questions and make a meaningful contribution to the existing literature within the time given. However, the scope should be broad enough to maintain an acceptable level of relevance for ASML. Driessen et al. (2015) present a research framework for the planning and control of spare parts in organizations that maintain high-value capital assets. This framework is used to define what processes are considered in-scope in this research and is visualised by Figure F.1 in Appendix F. Demand forecasting takes both a central position in this framework and this research. Moreover, when considering the input processes of demand forecasting, only the tracking and maintaining of spare part demand and installed base information is considered to be in-scope. In addition, the direct output of demand forecasting is used for tactical planning and inventory control decisions at ASML. However, the functioning of these processes in relation to demand forecasting is considered to be out-of-scope.

In general, the complex machines of ASML consist out of thousands of parts. However, not all parts are of interest when forecasting and controlling the use of spare parts. Demand for spare parts is caused by failures of service parts used in the installed base. ASML distinguishes service parts as parts that can potentially fail or degrade and need replacement in the field. Parts that are damaged or lost during service actions are also classified as service parts. Moreover, the installed base is defined as the set of machines for which ASML provides after-sales services (Dekker et al., 2013). In total, thousands of service parts are under the direct control of the SIM team. In practice, not all demand for these service parts is caused by failures in the field. Demand can also be triggered by the install of a new machine or the upgrade of an old part. However, these planned types of demand do not require corrective service actions and are considered to be out-of-scope. Finally, the SIM team is also responsible for the planning and forecast of service tools. These are tools needed to carry out any service action on the installed base of ASML. However, service tools are considered to be out-of-scope.

1.5. Research Methodology & Outline

The structure of this master thesis is based on the regulative cycle by Van Strien (1997), as depicted in Figure 1.1. This regulative cycle consists of five different phases: (1) problem definition, (2) analysis & diagnosis, (3) solution design, (4) intervention and (5) evaluation. This first chapter already gave an overview of the problem situation and the role of spare parts planning in addressing it. During the analysis & diagnosis phase, this problem situation is examined on different levels. First, the literature is reviewed for developments and relevant concepts in the field of spare parts demand forecasting during Chapter 2. Furthermore, the current practices are analysed by answering Research Question 1 in Chapter 3. Research Questions 2 and 3 are used to diagnose the current forecast performance and find areas for improvement. Both questions are discussed in Chapter 4, which concludes the analysis & diagnosis phase. Subsequently, Chapter 5 starts the solution design phase and derives a conceptual forecasting model based on installed base information, which provides an answer to Research Question 4. Next, this model is tested in a simulation study during Chapter 6, which starts the intervention phase. Afterwards, the model can be applied and reviewed under real-time conditions in Chapter 7. The combination of both chapters helps to answer Research Question 5. Besides, Chapter 8 provides conceptual solutions to relax assumptions underlying the new forecasting model.

In the end, the whole process is evaluated in the last phase, described by Chapter 9. This chapter formulates an answer to the main research question, states recommendations and indicates future research directions. Figure 1.2 summarizes the report outline by linking the chapters of this thesis to the phases of the regulative cycle.

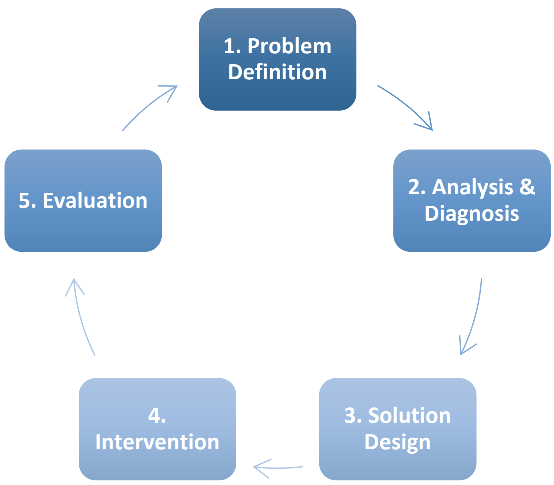

Figure 1.1: Regulative cycle by Van Strien (1997)

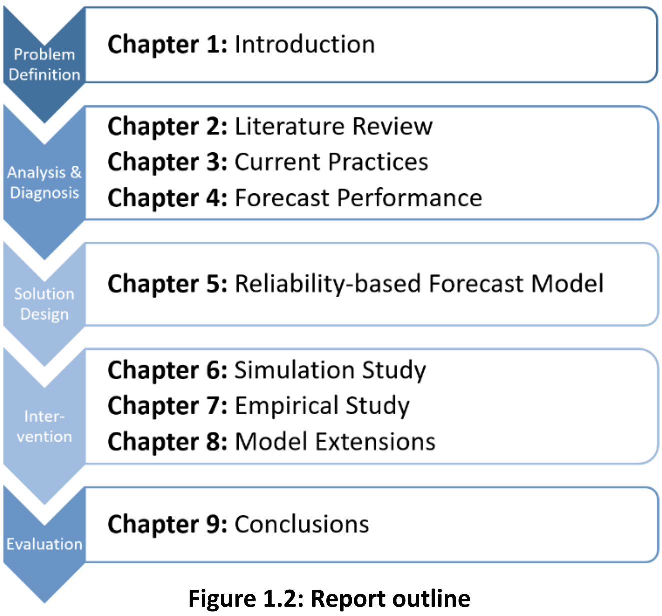

Figure 1.2: Report outline

2 Literature Review

In this chapter, the relevant literature on spare part demand forecasting is discussed and summarized. This involves the literature on spare parts classification (Section 2.1), spare parts demand forecasting techniques (Section 2.2) and performance metrics (Section 2.3). Moreover, all relevant gaps in the literature, which will be addressed by this thesis, are discussed (Section 2.4).

2.1. Spare Part Classification

Spare parts are highly varied, with different service requirements, costs and demand patterns (Boylan & Syntetos, 2008). These are characteristics that make spare parts demand forecasting both relevant and complex. Bacchetti & Saccani (2012) argue that spare part classification could provide a helpful foundation for forecasting the demand for these parts. Demand forecasting may be driven by part data collected for different classes. The most popular criteria used in literature are part cost, part criticality, supply characteristics, demand volume/value and demand variability (Bacchetti & Saccani, 2012). Part criticality indicates the importance of the part for sustaining production. The literature on spare parts demand forecasting proposes both qualitative and quantitative options to define criticality for classification purposes. Qualitative criticality classification requires expert knowledge and can be impractical when considering thousands of parts (Bacchetti & Saccani, 2012). Quantitative methods try to classify based on demand usage rates, part costs or average downtime (Cavalieri et al., 2008).
Another popular way of classifying spare parts is by using specific demand characteristics. Syntetos, Nikolopoulos & Boylan (2005) argue that the categorization of alternative demand patterns facilitates the selection of a forecasting method. The four main categories of demand patterns are described by Cavalieri et al. (2008) in the following way:

- **Smooth demand:** Demand has stable sizes and there are only a few periods with no demand.
- **Erratic demand:** There are few periods with zero demand but the demand sizes are highly variable.
- **Intermittent demand:** Appears randomly with many periods having no demand but when it occurs the demand sizes are stable.
- **Lumpy demand:** Demand for which there are many periods with zero demand and demand sizes will be highly variable when it occurs.

Spare parts mainly have to deal with an intermittent or lumpy demand pattern (Bacchetti & Saccani, 2012). Syntetos et al. (2005) categorize the four main demand pattern categories based on two criteria: the average inter-demand interval (ADI) and the squared coefficient of variation of demand sizes (CV^2). The ADI measures the average time between two consecutive periods with non-zero demand, whereas the CV^2 considers the variability in demand sizes during all periods with non-zero demand. Syntetos et al. (2005) derive cut-off values for both criteria to optimize the combination between spare parts demand forecasting techniques and the four demand patterns categories: $ADI = 1.32$; $CV^2 = 0.49$. Figure 2.1 visualizes the four main categories of demand patterns derived from these cut-off values (Costantino et al., 2018).

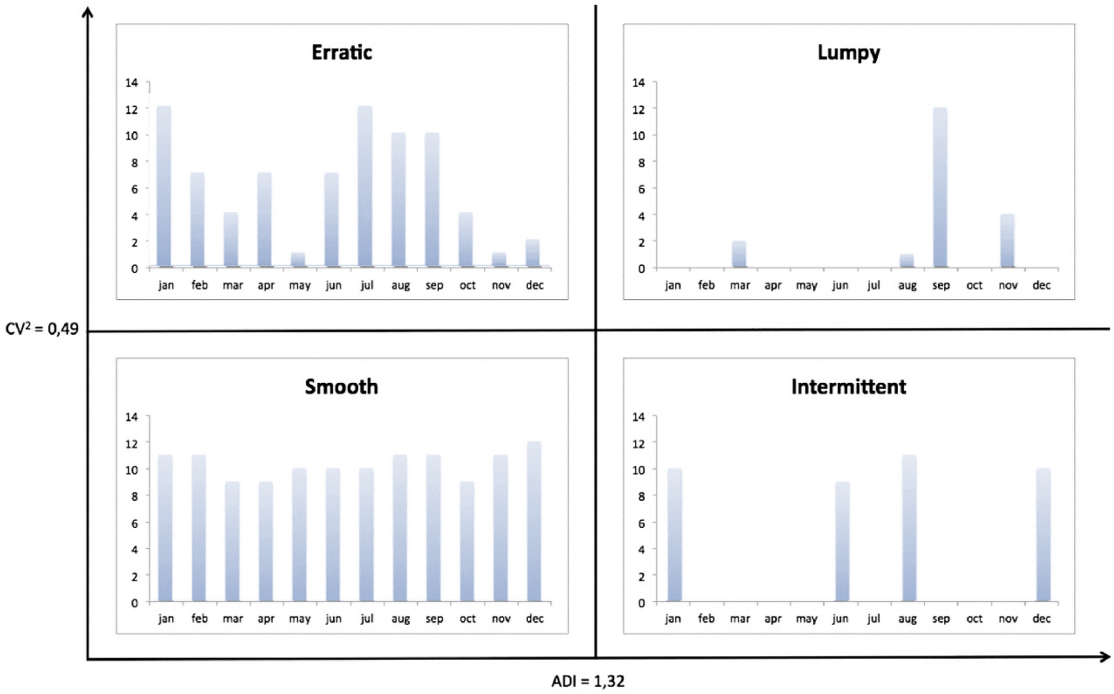

Figure 2.1: Classification of spare part demand patterns (Costantino et al., 2018)

2.2. Spare Part Demand Forecasting Techniques

In general, demands forecasting techniques can be roughly classified into two categories: qualitative and quantitative. Qualitative methods are based on subjective judgements and become impractical when considering thousands of spare parts. Quantitative methods can be sub-divided into three categories: time series-based, feature-based and hybrid (Bacchetti & Saccani, 2012). Time series-based techniques often only make use of historical demand data, whereas feature-based methods estimate demand as a function of explanatory variables. Moreover, hybrid methods combine both approaches. This section discusses the literature on time series-based forecasting methods (Section 2.2.1) and feature-based forecasting techniques using installed base information (Section 2.2.2).

2.2.1. Time Series-based Methods

The literature on spare parts demand forecasting has been dominated by time series-based forecasting techniques over the last decades (Van der Auweraer et al., 2019). Moving average and single exponential smoothing are two often used time series-based methods for predicting intermittent demand in practice (Syntetos & Boylan, 2005). The moving average method computes a (weighted) average of the last n data points in the time series. Furthermore, single exponential smoothing takes a weighted average of the last forecast and the current value of demand. Both methods are easy to use and have shown satisfactory performance on real intermittent data (Syntetos & Boylan, 2005). Nevertheless, Croston (1972) introduced an alternative method to handle time series with many zero demand periods. This method takes both the demand size and inter-arrival time of demand into account. Croston's method is seen as a benchmark in the field of intermittent demand forecasting (Syntetos et al., 2005). However, Croston's method turns out to have a positive bias when estimating spare parts demand (Syntetos & Boylan, 2005). This means that Croston's method systematically under-estimates the actual demand. The Syntetos-Boylan Approximation was introduced to approximately correct Croston's method. A large comparative study by Syntetos & Boylan (2005) shows that this approximation outperforms (simple) moving average, single exponential smoothing and Croston's method regarding intermittent demand. However, the Syntetos-Boylan Approximation still is not considered flawless. This approximation and the above-mentioned methods are not performing well when sudden demand drops occur. Teunter, Syntetos & Babai (2011) solved this problem by introducing the TSB method. This approach considers the demand probability instead of the inter-arrival times of demand. At last, Appendix B summarizes the equations and logic behind all the time series-based methods mentioned in this section.

2.2.2. Feature-based Methods with Installed Base Information

Over the last decade, the development of data collection, storage and analytics took a great leap forward (Van der Auweraer et al., 2019). This created possibilities to improve spare parts demand forecasting by including underlying factors that actually generate this demand. Those factors are referred to as installed base information (Van der Auweraer et al., 2019). Spare parts control requires a level of forecast accuracy that cannot be provided by time series-based forecasting techniques only (Dekker et al., 2013). Subsequently, Andersson & Jonsson (2018) identify three categories of feature-based forecasting using installed base data: reliability-based forecasting, regression-based forecasting and condition-based forecasting.

The category of reliability-based forecasting is based on the failure behaviour of parts, size and usage of the installed base, and the maintenance policy in place (Van der Auweraer et al., 2019). Additionally, Cavalieri et al. (2008) recommend predicting future failures by using part life data. Different distributions, like exponential and Weibull, can be fitted to this data, depending on the degradation behaviour of the part. The failure behaviour of parts in combination with accurately installed base registrations can potentially be used to derive forecasts.

Regression-based forecasting can use installed base information in the role of an explanatory variable when it correlates with spare part demand. In this case, possible explanatory variables are historical failures, the size of the installed base and the part failure behaviour (Van der Auweraer et al., 2019). In the literature, multiple regression-based models have been tested on spare parts demand before, without specifically including installed base information. The most prominent options using regression and data mining are neural networks and support vector machines. Neural networks are data-driven and self-adaptive models that have the advantage of dealing with an assumption-free nature (Zhang et al., 1998). They have the potential to outperform the time series-based techniques on spare parts demand, but are prone to overfitting and work as a black box (Kourentzes, 2013; Zhang et al., 1998). Support vector machines carry out regression tasks by constructing hyperplanes. However, they underperform when dealing with intermittent demand (Boukhtouta & Jentsch, 2018).

At last, condition-based forecasting involves information gathered by periodic or continuous monitoring of the components' condition. Wang & Syntetos (2011) were among the first to touch upon this category and developed a delay-time-based model to forecast spare parts demand. This model and other condition-based techniques show promising results but require a reliable way of periodic or continuous monitoring of all parts under control.

2.3. Performance Metrics

When comparing forecasting techniques, performance metrics are essential. In general, there are two types of relevant performance metrics: forecasting accuracy measures and stock control metrics. The literature on forecast accuracy metrics provides a range of options, all with their advantages and disadvantages. However, when comparing forecasting methods, having a higher forecast accuracy does not immediately result in a better inventory control performance (Teunter & Duncan, 2009; Kourentzes, 2013). The Mean Error (ME) metric measures the forecasting bias, while the Mean Absolute Deviation (MAD) and Mean Squared Error (MSE) describe the variability of the forecasting error. All three methods are simple in use but scale-dependent. Subsequently, the Mean Absolute Percentage Error (MAPE) is easy to interpret and scale-independent, but biased when periods of zero demand occur (Byrne, 2012). This is a problem for intermittent and lumpy spare part demand, which is characterised by many periods with zero demand according to Section 2.1. In the literature, several performance metrics are proposed to address this problem. Kim & Kim (2016) came up with a modified version of the MAPE, the Mean Arctangent Absolute Percentage Error (MAAPE). Another solution is the Adjusted Mean Absolute Percentage Error (A-MAPE), which is one of the most used metrics when comparing spare parts demand forecasting (Boukhtouta & Jentsch, 2018). At last, the Mean Absolute Scaled Error (MASE) prevents biases on intermittent demand by scaling the forecasting error. Most comparative studies on spare parts demand forecasting use multiple metrics to account for the problems and exploit the advantages. This is confirmed by Walström & Segerstedt (2010), who state that using a single performance metric does not represent all dimensions of the forecast performance when considering intermittent demand.

2.4. Literature Gap

This section discusses the most relevant gaps in the current scientific literature on spare parts demand forecasting in combination with installed base information. Those gaps are addressed by this research. In recent years, the development of data collection, storage and analytics stimulated research into the direction of more feature-based forecasting techniques using installed base information. Dekker et al. (2013) emphasised as one of the first the importance of more research into the direction of forecasting with installed base information. This conclusion has been drawn again years later by Van der Auweraer et al. (2019). Both papers introduce causal relationships between the installed base, maintenance policies and spare parts demand. However, the literature is currently lacking a method that can comprise these high-level relationships into a demand forecast on a service part level. Most current approaches in the literature consider the installed base on machine level, which means a fixed number of service parts per machine is assumed (Van der Auweraer et al., 2019). This assumption does not fit in perfectly for the semiconductor industry, where almost every machine has a unique configuration of service parts. Therefore, this research attempts to contribute to the scientific literature by developing a new methodology that addresses these gaps.

Furthermore, it is important not to lose sight of spare parts management in practice when attributing to the scientific literature. Bacchetti & Saccani (2012) suggest that very scarce empirical research has addressed the field of spare part management yet. Subsequently, in this field, there exists a gap between what has been largely investigated and proposed in the scientific literature and the lagging industrial practices, according to Cavalieri et al. (2008). This applies to a large extent to the literature on forecasting with installed base information. Most recent research in this direction focuses on conceptually testing new ideas, without evaluating the applicability in practice (Van der Auweraer et al., 2019). This research attempts to address this gap by conceptualizing a new forecasting model based on the data available in practice. Moreover, the performance of this model in practice is evaluated by an empirical study with data from the semiconductor industry.

In the end, Chapter 9 considers to what extent this research managed to address the literature gaps discussed in this section.

3 Current Practices

In this chapter, the current framework of forecasting and planning at ASML is outlined. It is important to understand how spare parts demand forecasts are currently made and used at ASML. These insights are required before any improvements with installed base information can be suggested and implemented. In this way, this chapter finds an answer to Research Question 1. First, the spare parts planning processes are explained in Section 3.1. This helps to understand the role of spare parts demand forecasting in the customer service supply chain of ASML. Afterwards, Section 3.2 describes in more detail how these spare part demand forecasts are currently derived and reviewed.

3.1. Spare Parts Planning Processes

The Supply Chain Management (SCM) department at ASML is responsible for ensuring material availability to ASML's factories and customers. Moreover, the Customer Supply Chain Management (CSCM) department is an important branch of SCM and is specifically devoted to the material availability of the customers. This project is conducted within the CSCM department, as part of the Service Inventory Management (SIM) team. The SIM team is concerned with the tactical planning and inventory control of spare parts and service tools all over the world. ASML operates a worldwide service network of central and local warehouses to meet the service levels agreed with the customers. The central warehouses supply the local warehouses, which are located closer to the customers. Local warehouses can also supply each other by using lateral transshipments. Demand at the customers can also be satisfied by so-called emergency shipments from the central warehouse. A simplified overview of the customer supply chain is depicted in Figure 3.1.

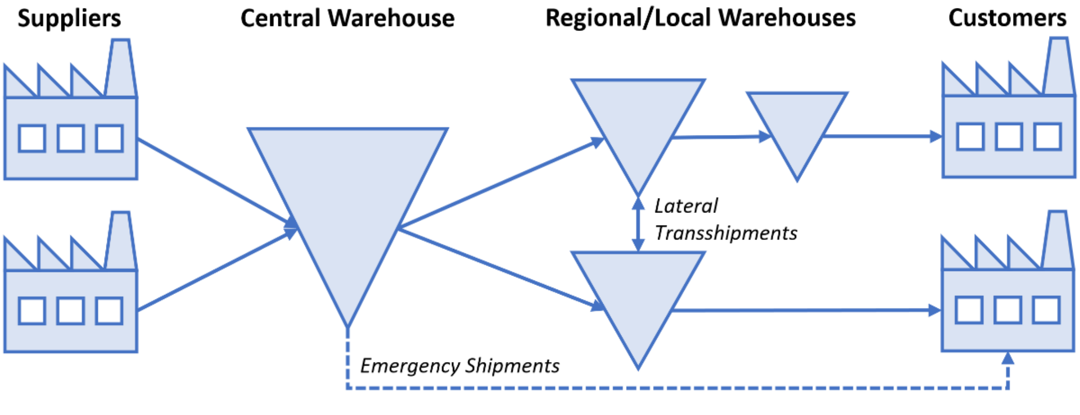

Figure 3.1: ASML customer service supply chain

The objective of the SIM team is to ensure material availability to the customer by using this service network in a cost-efficient way. The main task involved is to set the spare part basestock levels at the central and local warehouses. This decision-making includes a trade-off between obtaining the agreed service levels and keeping the inventory, stock-out and obsolescence costs as low as possible. The SIM team deals with this trade-off by applying the spare parts inventory model as derived by Van Houtum & Kranenburg (2015). Their logic is used in a multi-item multi-echelon algorithm which helps to propose basestock levels for all service parts in scope. The most relevant input parameters for this algorithm are the service network structure, service level agreements, spare part cost prices and spare part demand rates. A part of the SIM team is devoted to generating spare parts demand forecasts, as they are the most important driver of the planning processes. The demand rates in the spare parts inventory model by Van Houtum & Kranenburg (2015) are assumed to follow independent Poisson processes. In practice, this means only a single demand rate parameter has to be estimated. The Poisson process assumption can be justified in two possible ways, according to Van Houtum & Kranenburg (2015). First, the assumption holds when the lifetimes of the service parts are exponentially distributed. Moreover, the Poisson process assumption is also sustained when the installed base of machines is sufficiently large, while the lifetimes of service parts are generally (arbitrary) distributed. An additional assumption in the inventory model states that the demand rate is constant for each service part. This assumption is reasonable when the machine downtimes are relatively short or rarely occur (Van Houtum & Kranenburg, 2015).

3.2. Spare Parts Demand Forecasting Process

Spare parts demand forecasting plays an important role in the planning processes of the SIM team. This section explains in more detail how the current forecasting process works. First, Section 3.2.1 clarifies how spare parts demand forecasts are derived from past demand usage rates. Moreover, Section 3.2.2 denotes how the installed base information underlying this process is currently obtained. Afterwards, Section 3.2.3 discusses which steps can be taken to enrich the forecast from Section 3.2.1 with initial failure rates and expert knowledge. At last, Section 3.2.4 considers how the performance of the current forecasting process is monitored by tracking signals.

3.2.1. Current Forecast Method

In the current configuration, spare parts demand forecasts are based on demand usage rates from the past. The demand usage rate is denoted by the parameter u_t^k and describes the demand for service part k of a single machine during period t. Thus, the main idea is to relate the spare parts demand of service part k in a given period to the average number of machines containing this part during the same period. In the current configuration, all periods are measured in quarters. Parameter x_t^k denotes the aggregated demand for service part k during period t. This only includes the demand related to corrective service actions, as described by the research scope in Section 1.4. Moreover, ASML counts the number of machines that contained service part k during any period t in the past by the parameter IB_t^k. This parameter is referred to as the average applicable installed base size of service part k during period t. The average applicable installed base size during a past period can include machines that are not active anymore nowadays. Section 3.2.2 describes the derivation of parameter IB_t^k in more detail. The average demand usage rate u_t^k of a single machine during period t can be found by equation 3.1.

$$u_t^k = \frac{x_t^k}{IB_t^k} \tag{3.1}$$

The demand usage rates from the past are used to forecast the demand usage rate during the upcoming periods. The current configuration requires the demand usage data of a service part k over the last three years, aggregated into twelve quarters. Based on this data, it is possible to predict the demand usage rate in any upcoming quarter. In specific, at time t a forecast can be derived for any upcoming quarter $t+i$, where parameter i represents the forecasting horizon in quarters. This demand usage rate forecast is denoted by the parameter $\hat{u}_{t+i}^k$ and can be found by applying an exponentially weighted moving average on the usage rates u_t^k of the last twelve quarters. More recent usage rates receive a higher weight, following an exponential pattern. The weight given to each quarter t is found by taking the power of a so-called recency factor p, in which the power depends on how long ago the usage occurred. The recency factor has been optimized in the past and is the same for every service part: $p = 5/7$. Parameter $w_{t;\,t-q}$ represents the weight given to the usage rate from q quarters ago, when making a forecast during quarter t. Equation 3.2 describes how all weights are computed when considering n quarters in total. In the current configuration, n is equal to twelve, the equivalent of three years.

$$w_{t;\,t-q} = p^q, \quad q \in \{0, n-1\} \tag{3.2}$$

Subsequently, equation 3.3 denotes how these weights can be used to find the demand usage rate forecast $\hat{u}_{t+i}^k$. This expression can be formally described as an exponentially weighted moving average (EWMA). The EWMA closely approaches the single exponential smoothing (SES) method (Nahmias & Olsen, 2015). This SES method has already been introduced and discussed in Section 2.2.1 of the literature review. In this case, the only difference between both methods is that the EWMA uses a limited set with three years of demand data, whereas SES does not exclude any past demand.

$$\hat{u}_{t+i}^k = \frac{\sum_{q=1}^{n} w_{t;\,t-q}\, u_{t-q}^k}{\sum_{q=1}^{n} w_{t;\,t-q}} = \frac{\sum_{q=1}^{n} p^q u_{t-q}^k}{\sum_{q=1}^{n} p^q} \tag{3.3}$$

ASML does not only maintain the number of applicable machines during the last three years. The number of machines that will be installed during the time horizon $[t, t+i]$ is also known when a forecast is produced during quarter t. Therefore, it is possible to determine the average applicable installed base size IB_{t+i}^k in period $t+i$, for which a forecast is produced. The total expected demand $\hat{x}_{t+i}^k$ for service part k during quarter $t+i$ can be found by multiplying the forecasted demand usage rate $\hat{u}_{t+i}^k$ with the future installed base IB_{t+i}^k. This computation is denoted by equation 3.4.

$$\hat{x}_{t+i}^k = \hat{u}_{t+i}^k IB_{t+i}^k \tag{3.4}$$

Although the current forecasting process aggregates demand data quarterly, the final forecasts are produced on a monthly level. Simply dividing the outcome of equation 3.4 by three gives a monthly forecast based on the quarterly forecast. When making a forecast on a monthly basis, all twelve quarters shift one month forward every time a forecast is produced. Furthermore, the outcome of equation 3.4 presents a forecast of the worldwide demand for a service part under consideration. However, the planning processes might require forecasts on a more local level, which could be per region, customer or machine type. A local-level forecast can be found by taking the product of the forecasted demand usage rate $\hat{u}_{t+i}$ and the installed base size applicable to this aggregation level. This research focuses on the forecast performance at the highest aggregation level.

3.2.2. Applicable Installed Base Size

The current method uses the average applicable installed base size of a service part to create demand usage rates. This section explains the derivation of the parameter IB_t^k, which has been defined as the average applicable installed base size of service part k during quarter t. This derivation requires insight into the service BOM (Bill of Material) registrations of machines at ASML. The service BOM indicates which service parts are integrated into a specific machine. The configuration of every ASML machine is unique, which means every machine has its own service BOM. The number of applicable machines per service part can be found when the service BOM is known for every machine in the field.

The composition of the service BOM of a machine starts with the introduction of so-called service collections. These are groups of service parts that apply to one or more machine types, also called object types. However, this does not mean that all parts in a service collection apply to every machine of a corresponding object type. In general, service collections consist of two types of service parts: parts that are applicable on every machine in the object type and parts that are only used in a limited set of machines in the object type. The always applicable parts are directly added to the service BOM. The limited applicable parts require an extra validation step to confirm whether they are actually used in the machine. This step requires insight into the current configuration of the entire machine. The Configuration as Maintained (CAM) describes the whole configuration of a machine as it is currently maintained in the field. This involves all changes made to the machine after it was installed. When service parts, classified as limited in the service collection, are mentioned in the CAM of a machine, they are added to its service BOM. Another possible way to end up in the service BOM is to be part of an applicable option collection. Customers of ASML have the option to enhance their machine by selecting from several upgrades. The service parts involved are combined in option collections, which are not directly linked to object types. At last, Figure 3.2 summarizes the composition of a service BOM for a machine. The number of applicable machines per service part can be derived when the service BOM is completed for every active machine in the field.

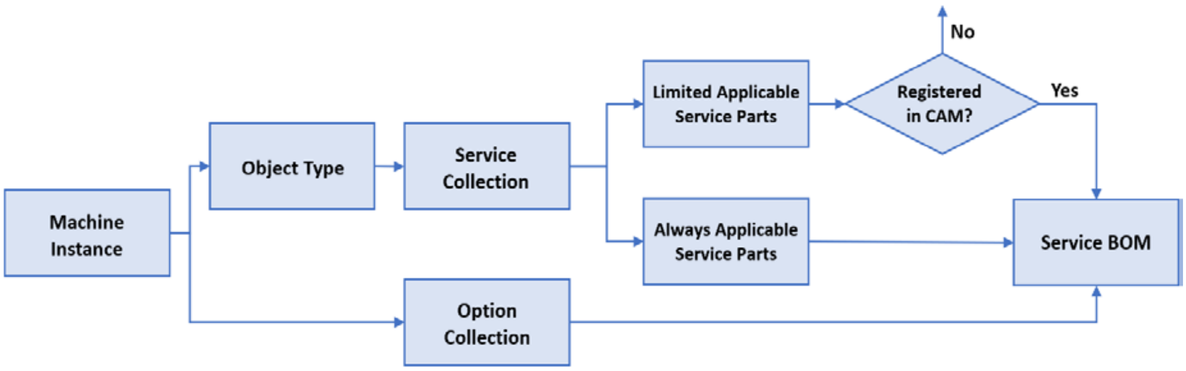

Figure 3.2: Service BOM composition for a machine instance

3.2.3. Initial Failure Rates & Expert Knowledge

The current forecasting method depends on demand usage rates from the past. The SIM team becomes responsible for the planning of service parts when they passed through the initial life phase. In general, this phase takes about three years. However, for these parts only a limited amount of usage data is available. Therefore, it is hard to create accurate forecasts in the first years of a service part by only using demand usage rates. ASML tries to bridge this period by a concept called the Initial Failure Rate (IFR). This rate is a rough approximation of the failure behaviour of a service part, estimated by the Development & Engineering (D&E) department at ASML. The IFR is especially useful in the initial phase of a service part, but its value decreases as more usage data becomes available over the years. ASML uses an IFR transition framework based on two criteria: usage during the last three years and machine years. The number of machine years indicates how many years all machines together have been running while using the service part. Figure 3.3 shows how the forecast gradually transits from IFR to a usage forecast (UF). This figure indicates that the IFR is initially leading in the forecasting process for service parts. In some cases, an extra review step is included to resolve potential unrealistic differences between the IFR and usage forecast.

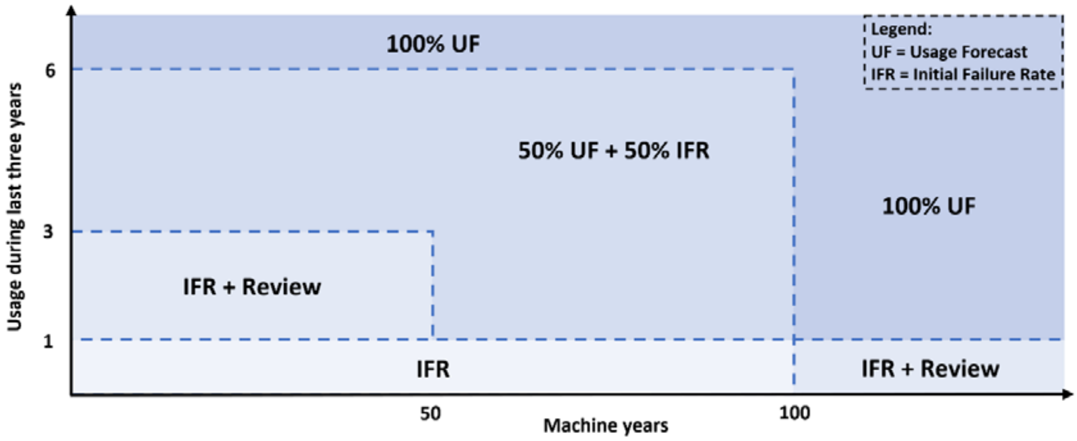

Figure 3.3: IFR transition threshold framework

The forecast retrieved from the transition diagram is not always the final forecast. In the end, the forecasting process contains a review moment during which expert knowledge can be used to adjust the forecast. Expert knowledge can involve all types of information that impacts the usage or failure behaviour in the near future but is not yet reflected in the historical usage. However, the remainder of this research does not take the effect of IFRs and expert knowledge into account. Therefore, this research focuses on service parts with a 100% usage forecast. The outcome of the usage forecast, as presented by equation 3.4, is referred to as the Enriched EWMA during the remainder of this research.

3.2.4. Tracking Signals

An important part of the forecasting process is the assessment of the forecast performance over time. Currently, the SIM team uses a tracking signal to monitor this performance. Tracking signals are based on statistical process control theory and generate alerts whenever the process is statistically out of control. In specific, the SIM team uses the tracking signals derived by Trigg (1964). A tracking signal for service part k in month t describes the relationship between the smoothed error SE_t^k and the smoothed mean absolute deviation $SMAD_t^k$. First, the error E_t^k is introduced to describe the difference between forecasted demand $\hat{x}_t^k$ and actual demand x_t^k for service part k in month t. In addition, smoothing parameter α is introduced to exponentially weight these past error observations to derive the SE_t^k. The smoothed error SE_t^k can be computed for $t > n$ by using equation 3.5. Besides, equation 3.6 is required to initialise SE_t^k for period n.

$$SE_t^k = (1-\alpha)SE_{t-1}^k + \alpha E_t^k = (1-\alpha)SE_{t-1}^k + \alpha\left(x_t^k - \hat{x}_t^k\right) \tag{3.5}$$

$$SE_n^k = \sum_{t=1}^{n} \left(x_t^k - \hat{x}_t^k\right) \tag{3.6}$$

The computation of the $SMAD_t^k$ follows the same line of reasoning as for the SE_t^k. However, in this case, the mean absolute error MAD_t^k is introduced to describe the absolute difference between forecasted demand $\hat{x}_t^k$ and actual demand x_t^k of service part k in month t. The smoothed mean absolute error $SMAD_t^k$ can be computed for $t > n$, by using equation 3.7. In addition, equation 3.8 describes how to initialise the $SMAD_t^k$ for period n.

$$SMAD_t^k = (1-\alpha)SMAD_{t-1}^k + \alpha MAD_t^k = (1-\alpha)SMAD_{t-1}^k + \alpha\left|x_t^k - \hat{x}_t^k\right| \tag{3.7}$$

$$SMAD_n^k = \frac{1}{n}\sum_{t=1}^{n}\left|x_t^k - \hat{x}_t^k\right| \tag{3.8}$$

At last, the tracking signal TS_t^k of service part k in month t can be found by dividing the SE_t^k with the $SMAD_t^k$, as described by equation 3.9. When the SE_t^k deviates from zero, the forecasting method deals with a bias. This bias is scaled by the smoothed mean absolute deviation. Therefore, a tracking signal TS_t^k indicates how large the forecast bias is relative to the absolute error.

$$TS_t^k = \frac{SE_t^k}{SMAD_t^k} \tag{3.9}$$

In general, tracking signals can generate reactive and proactive alerts. The reactive alerts are generated when a tracking signal is located outside the two sigma control limits (Trigg, 1964). A signal above the upper control limit indicates an under-forecast, while a signal under the lower control limit points at an over-forecast. Trigg (1964) derived an approximation for the two sigma control limits:

$$CL = \pm\frac{2.4\alpha}{\sqrt{2\alpha - \alpha^2}} \tag{3.10}$$

The SIM team adopted a value of 0.1 for the smoothing parameter α, as recommended by Trigg (1964). Furthermore, the tracking signals currently require an initialisation of twelve periods $(n = 12)$. Figure 3.4 gives an example of a spare part demand forecast by the Enriched EWMA for a service part at ASML. The corresponding tracking signals are depicted on the right. The upper and lower control limits are derived by equation 3.10, which results in limits of 0.55 and -0.55 respectively. This example indicates that the forecast is statistically out of control in the ninth month, which points to an under-forecast. In the months before this reactive alert, the forecast is already trending towards this situation. A proactive alert is generated when three subsequent tracking signals are trending upwards.

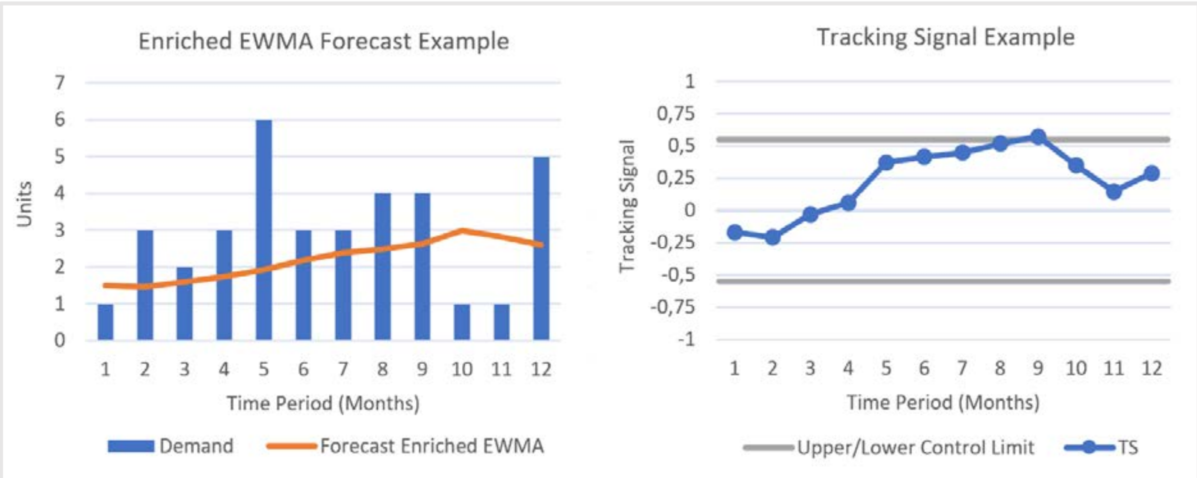

Figure 3.4: Enriched EWMA forecast example with the corresponding tracking signals

A disadvantage of tracking signals is that they mainly measure the forecast bias and not the forecast error. A forecast can have a high level of error but still be completely unbiased. The opposite can also be true, as parts with relatively small errors can be biased as well. This can result in parts classified as biased while they are no direct threat to the planning processes. The next section considers the overall forecast performance of the Enriched EWMA for all relevant service parts remaining in scope.

4 Current Forecast Performance

In this chapter, the forecast performance of the current forecast method, called Enriched EWMA, is analysed in multiple ways. First, the performance metrics and the descriptive statistics of the demand data are specified in Sections 4.1 and 4.2 respectively. Afterwards, the performance metrics are used to track the forecast performance of the Enriched EWMA for this demand data (Section 4.3). Besides, spare parts classification (Section 4.4) and time series-based benchmark methods from the literature (Section 4.5) are used to put this forecast performance into perspective. In this way, this chapter provides an answer to Research Questions 2 and 3, and closes the diagnostic and analysis phase.

4.1. Performance Metrics

The assessment of the current forecast performance requires performance metrics. Walström & Segerstedt (2010) advise the use of multiple complementary metrics, to account for the different shortcomings that most performance metrics have. Moreover, Teunter & Duncan (2009) claim that the combination of forecast bias and error measures gives a representable overview of the overall forecast performance. The current forecasting process only uses the tracking signals from Section 3.2.4 for this purpose. However, these tracking signals are useful for performance control over time but less informative in a performance evaluation study. Sections 4.1.1 and 4.1.2 discuss which complementary bias and error measures are included in this research respectively.

4.1.1. Forecast Bias Metrics

The forecast bias describes the tendency of a forecasting method to be consistently higher or lower than the actual value. In this research, the Mean Error (ME) is introduced to express the size and direction of the forecast bias. The ME is defined by Equation 4.1.

$$ME = \frac{1}{n}\sum_{t=1}^{n}(x_t - \hat{x}_t) \tag{4.1}$$

In this case, parameter n represents the number of monthly out-of-sample demand datapoints. The ME metric is widely used in the literature on spare part demand forecasting. However, this metric is scale-dependent, which can cause problems when comparing the forecast bias between parts with different demand level scales. This shortcoming is solved by scaling the ME with the average demand per time unit (Syntetos & Boylan, 2005). This derivation is named the Scaled Mean Error (SME) metric and can be derived from the ME to allow forecast bias comparisons between parts (Equation 4.2).

$$SME = \frac{\sum_{t=1}^{n}(x_t - \hat{x}_t)}{\sum_{t=1}^{n} x_t} = \frac{ME}{\frac{1}{n}\sum_{t=1}^{n} x_t} \tag{4.2}$$

4.1.2. Forecast Error Metrics

The forecast error expresses the absolute difference between a forecast and the actual value, independent of the direction. The Mean Squared Error (MSE) is included to describe the forecast error by taking the square of this difference. In this way, this metric reflects the variance of the forecast errors and gives more weight to the larger deviations. The MSE is defined by Equation 4.3.

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(x_t - \hat{x}_t)^2 \tag{4.3}$$

The MSE is widely used in the literature on spare parts demand forecasting in combination with the ME. However, the MSE is also scale-dependent, just like the ME. The literature review in Section 3.2.3 provided several scale-independent forecast error metrics that could be complementary to the MSE. In general, the MAPE is the most intuitive scale-independent measure available in the literature. However, this metric is not able to handle intermittent demand, due to the occurrence of many periods with zero demand (Byrne, 2012). The MASE, proposed by Hyndman & Koehler (2006), gives a scale-free measurement of the forecast error and is unbiased regarding intermittent spare parts demand. In this case, this measure is preferred over other scale-independent measures like the A-MAPE and MAAPE due to its more intuitive interpretability (Section 2.3). In more detail, the MASE scales the forecast error by the in-sample Mean Absolute Deviation (MAD), which is obtained by using a naïve forecasting method, called Random Walk (RW). This naïve method takes the actual demand during the last month as a forecast for the coming month. The MASE is defined by Equation 4.4.

$$MASE = \frac{\frac{1}{n}\sum_{t=1}^{n}|x_t - \hat{x}_t|}{\frac{1}{n'-1}\sum_{t=2}^{n'}|x_t - x_{t-1}|} \tag{4.4}$$

In this case, n and n' represent the number of monthly out-of-sample and in-sample demand data points respectively. In line with the research by Babai, Syntetos & Teunter (2014), all performance metrics are arithmetically averaged to enhance performance comparisons with subsets of service parts. Finally, an inventory control performance measure would be another relevant addition next to these forecast performance metrics, as a higher forecast performance does not immediately result in a better inventory control performance (Kourentzes, 2013). However, the inventory control processes are considered to be out-of-scope in this research.

4.2. Demand Data Overview

Historical demand data has been obtained for a 5-year period for all service parts in scope (May 2016 – April 2021). Monthly aggregates are created from this demand data because forecasts are produced every month. The data history for each service part has been split into two samples: in-sample (36 months) and out-of-sample (24 months). The in-sample months are used for initialisation and parameter-fitting, whereas the out-of-sample months serve as a test and evaluation set. All service parts without demand during the in-sample or out-of-sample months are excluded from the study. The number of remaining service parts in the set is referred to as 100% due to confidentially reasons. Table 4.1 gives an overview of the descriptive statistics for this group of service parts.

Table 4.1: Descriptive statistics of the 5-year demand data history for all service parts in scope

	Monthly Demand		Demand Size		Demand Intervals	
	Mean	**St. Dev.**	**Mean**	**St. Dev.**	**Mean**	**St. Dev.**
Minimum	0.033	0.181	1.000	0.000	1.000	0.000
25th percentile	0.083	0.280	1.000	0.000	2.318	1.455
Median	0.222	0.485	1.211	0.419	5.643	4.375
75th percentile	0.694	0.710	1.731	0.827	9.834	6.227
Maximum	136.889	79.005	136.889	79.005	19.667	13.421

Table 4.1 indicates that the average monthly demand is below one for most of the service parts. These service parts only sporadically have to deal with demand. In addition, when demand occurs, the demand size is close to one most of the time. The standard deviations of the demand sizes also point at a low degree of lumpiness among the parts. On the other hand, most parts have to deal with relatively high levels of intermittency, given that most average demand intervals are well above one. The literature review in Chapter 2 described how Syntetos et al. (2005) derived threshold values to recognize intermittent and lumpy demand behaviour. The threshold between high frequency and intermittent demand is given by $ADI = 1.32$, which means 89% of all service parts in scope can be classified as intermittent. The threshold between low and high demand size variability ($CV^2 = 0.49$) is only met by 5% of the service parts.

However, what Table 4.1 does not show, is that the average demand levels have steadily increased over the past years. One of the explanations for this is related to the increase in the average applicable installed base size per service part. Figures 4.1 and 4.2 show the increase over the last three years for both the average spare parts demand and the average applicable installed base size per quarter respectively. The scale of the average applicable installed base size has been normalized due to confidentially reasons. Dekker et al. (2013) argue that spare parts demand follows the demand for installed systems, but with a delay. In both cases the increase is linear, but the demand for spare parts is increasing faster than the applicable installed base size. The remainder of this research elaborates on the implications of this relationship.

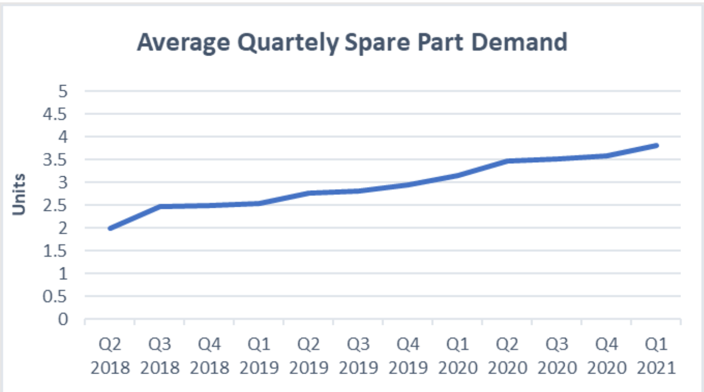

Figure 4.1: Overview of the last three years of the average quarterly spare parts demand

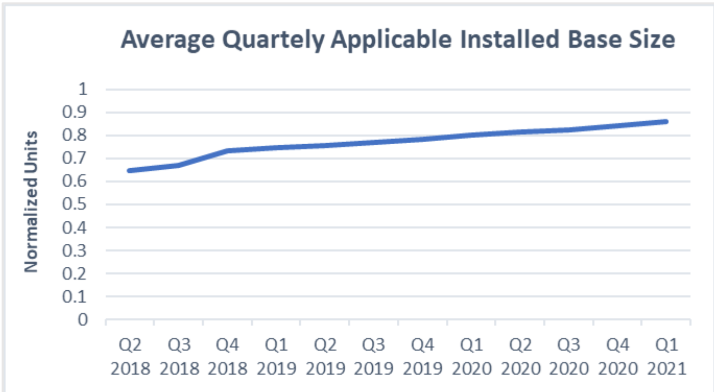

Figure 4.2: Overview of the last three years of the average quarterly applicable installed base size

4.3. Enriched EWMA Performance

In this section, the performance of the Enriched EWMA is assessed by the performance metrics that have been introduced in Section 4.1. First, Section 4.3.1 considers the overall performance of all service parts in scope. Afterwards, Section 4.3.2 analyses the forecast performance for service parts with an increasing demand pattern.

4.3.1. Overall Forecast Performance

The Enriched EWMA requires three years of demand data for a single forecast. Therefore, the in-sample period of three years has been used for initialisation, whereafter a monthly forecast is computed during the last two years with out-of-sample months. Table 4.2 shows how the Enriched EWMA performed during the last two years in terms of SME, MSE and MASE. The MASE is initialised by computing the MAD of the RW method during the in-sample period. According to Table 4.2, all three metrics strictly increased over the 2-year time horizon. This overall increase coincides with a steady increase in overall demand for spare parts during the most recent years (Figure 4.1).

Table 4.2: Overall out-of-sample forecast performance Enriched EWMA

	SME	MSE	MASE
Year 4	0.022	4.337	0.967
Year 5	0.055	6.499	1.111
Overall	**0.041**	**5.418**	**1.039**

When using the tracking signals metric, the percentage of biased parts in the last month of years 4 and 5 can be found. In this case, this percentage decreased from 24.7% to 17.8% between both points in time, which is not in line with the results in Table 4.2. Section 3.2.4 already explained that tracking signals can have problems with recognizing the service parts with the most problematic forecasts. It now seems that only using tracking signals has caused a distorted picture of the forecast performance over the years. This is confirmed when directly comparing the biased and unbiased parts with the MASE. It turns out that the biased group at the end of year 5 scores better in terms of forecast error (1.055) than the unbiased group (1.122). In the end, this validates that only using a single bias measure does not represent all dimensions of the forecast performance (Walström & Segerstedt, 2010).

4.3.2. Forecast Performance with Increasing Demand Pattern

This section analyses the impact of an increasing demand pattern on the forecast performance more closely. First, all service parts with an increasing demand pattern are identified. This is done by computing the average demand over each of the last three years per service part. Only the parts with a strictly increasing mean demand over these three years are selected. From these service parts, 91% already dealt with a strictly increasing demand pattern during the years of the in-sample period. Table 4.3 splits the performance for the strictly and not strictly increasing parts. Both groups with service parts take up 19% and 81% of the remaining service parts in scope respectively.

Table 4.3: Out-of-sample forecast performance Enriched EWMA with and without strictly increasing demand

	Strictly Increasing			Not Strictly Increasing		
	SME	MSE	MASE	SME	MSE	MASE
Year 4	0.133	10.698	0.835	-0.070	2.845	0.998
Year 5	0.154	17.868	1.305	-0.066	3.833	1.066
Overall	**0.146**	**14.292**	**1.070**	**-0.068**	**3.339**	**1.032**

The SME metric in Table 4.3 indicates that all service parts with a strictly increasing demand pattern are dealing with a large positive bias. This bias indicates that the Enriched EWMA is systematically under-predicting the actual demand levels. The increase in SME between years 4 and 5 shows that the increase in forecast bias is disproportional to the average demand growth between both years. Besides, the MASE also points at a disproportional increase in forecast error on the same time horizon. Thus, the overall forecast performance of the Enriched EWMA is deteriorating disproportionately when considering service parts with a strictly increasing demand pattern.

The performance of the group of service parts without a strictly increasing demand pattern puts the results further into perspective. First, Appendix A.1 presents more detailed descriptive statistics about both subsets. Most importantly, the demand levels of the subset without strictly increasing demand are stable over time and about 4.5 times lower than for the strictly increasing group (Table A.2). The SME metric in Table 4.3 points out that the Enriched EWMA currently systematically over-forecasts this type of demand. Thus, the Enriched EWMA, which uses the same parameter configuration for all service parts, struggles to specifically capture both increasing and stable demand patterns. The next section considers additional classification criteria to find out what other features have a significant impact on the forecast performance as well.

4.4. Spare Parts Classification

In this section, spare part classification is used to distinguish differences in forecast performance for service parts at ASML. In general, service parts are known for their highly varied characteristics (Boylan & Syntetos, 2008). First, Section 4.4.1 explains which classification criteria can be derived from the literature and installed base information. Afterwards, these criteria are used in Section 4.4.2 to empirically test their effect on the forecast performance of the Enriched EWMA.

4.4.1. Classification Criteria

The literature on spare parts demand forecasting already provided a range of potential classification criteria. First, the demand usage rate is a commonly used criterion to define the criticality of a service part (Cavalieri et al., 2008). However, Boylan et al. (2008) found that the ADI and CV^2 are the most distinctive spare part characteristics for classification. Van Wingerden et al. (2014) confirm this and conclude that the average demand size is less distinctive. According to Bacchetti & Saccani (2012), the part cost is the most popular spare part control characteristic, next to these demand characteristics. Classifiers derived from installed base information, on the other hand, are hardly represented in the literature (Bacchetti & Saccani, 2012). Given the explanatory role of installed base information, both the applicable installed base size per service part (Section 3.2.2) and the overall part age are included. The overall part age describes the time that has passed since the first-ever registration of a part.

4.4.2. Performance Classification

The next step is to empirically test the impact of the discussed classification criteria on the forecast performance of the Enriched EWMA. In line with the research by Van Wingerden et al. (2014), the set of parts has been divided into three equal groups, which score lowest, middle and highest on each of the characteristics mentioned, during the out-of-sample months. Next, it was possible to find both the forecast error and bias for each of the groups per classification criteria. In this case, the MASE and SME are used to define this forecast error and bias respectively. The scaling of both methods reduces the impact of differences in demand levels between the categorized groups. Sections 4.4.2.1 and 4.4.2.2 discuss the results of the forecast error and bias classification respectively.

4.4.2.1. Forecast Error Classification

The effect of the classification criteria from Section 4.4.1 on the forecast error is visualized by Figure 4.3. All criteria, except part cost, have one group scoring arguably worse in terms of forecast error than the others. Syntetos et al. (2005) found that the standard EWMA generally struggles in terms of forecast error when dealing with intermittent demand, which is known for its large ADI level. This is only partly reflected by the results in Figure 4.3. High intermittence would generally mean a time series-based forecasting method struggles to deliver accurate forecasts. However, too much intermittence ensures the forecast quickly adapts to values close to zero, which is close to the actual demand most of the time. Figure 4.3 also indicates that demand size variability starts to cause problems only for higher levels of variability. This corresponds to the finding that only 5% of the service parts can be seen as highly variable, following the thresholds of Syntetos et al. (2005). At last, the overall age and applicable installed base size classifiers clearly show the negative impact of only having a brief demand history.

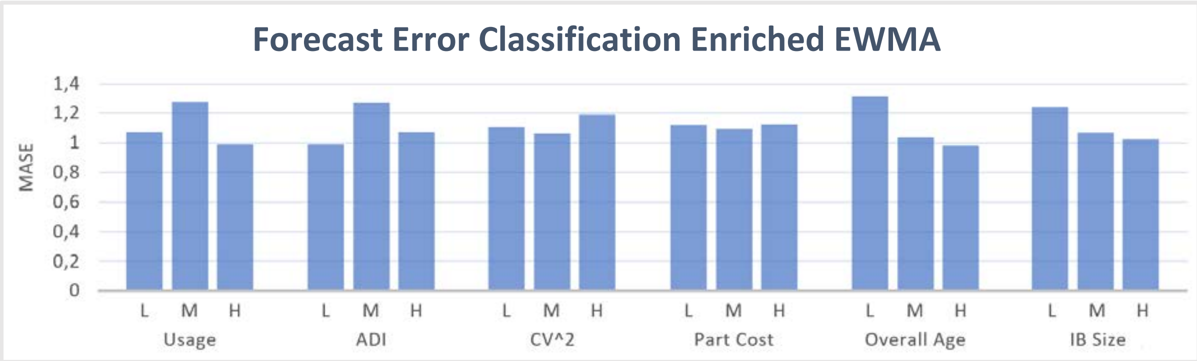

Figure 4.3: Out-of-sample forecast error classification of the Enriched EWMA

4.4.2.2. Forecast Bias Classification

The impact of classification on the forecast bias is depicted in Figure 4.4. In this case, clear trends can be recognized for most classifiers. Parts with low usage and/or high intermittence levels generally deal with over-forecast, which is caused by many periods with zero demand. When the usage increases and/or intermittence levels decrease, the level of under-forecasting gradually increases. The demand size variability, overall part age and applicable installed base size follow the same trend, but to less extent. Finally, it can be concluded that part cost has no significant effect on the overall forecast performance. Both the forecast error in Figure 4.3 and the forecast bias in Figure 4.4 remain fairly stable when classifying based on part costs. Nevertheless, part cost still is a relevant property to consider, as forecasting becomes more important when parts are more expensive (Van Wingerden et al., 2014). When using this initial classification, it turns out that the high usage parts, on average, are worth 50% more than the low usage parts.

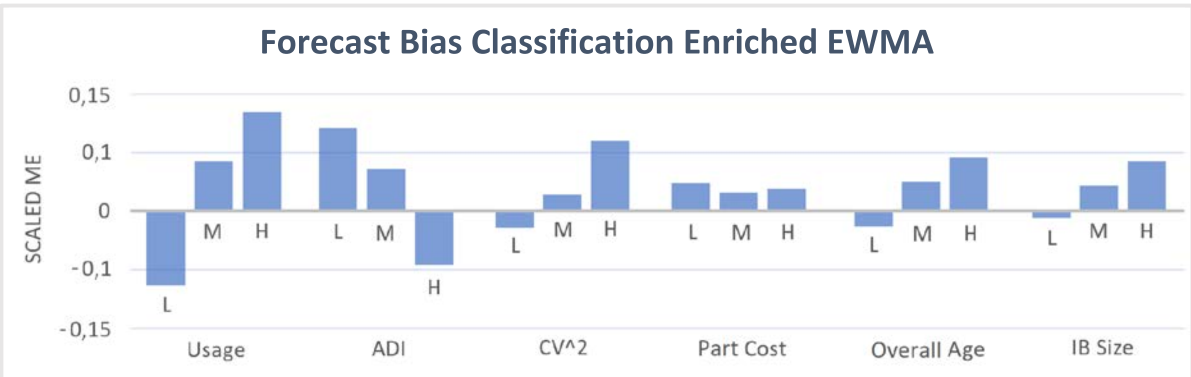

Figure 4.4: Out-of-sample forecast bias classification of the Enriched EWMA

4.5. Benchmark Performance

In this section, benchmark methods from the literature are introduced to put the forecast performance of the Enriched EWMA into perspective. First, Section 4.5.1 discusses which benchmark methods are included for analysis. Moreover, Section 4.5.2 clarifies the model-fitting and optimization procedures for those methods. Afterwards, all outcomes are presented and evaluated in Section 4.5.3.

4.5.1. Benchmark Methods

The literature review in Section 2.2.1 gave a short overview of the most prominent time series-based methods that have dominated the literature on spare parts demand forecasting over the last decades. These methods are used as a benchmark during this research. Table 4.4 presents an overview of all methods involved and classifies them as traditional, Croston-based or naïve. The Croston-based methods are specifically designed to capture the slow-moving and intermittent behaviour of spare parts, unlike the traditional methods. Furthermore, Random Walk (RW) is the naïve method already used by the MASE, whereas the Zero Forecast (ZF) predicts there will be no demand at all. Appendix B elaborates on the assumptions and equations behind the traditional and Croston-based methods.

Table 4.4: Time series-based forecasting methods used as a benchmark

Forecasting Method	Abbreviation	Parameters	Reference
Traditional Methods			
Simple Moving Average	SMA	n	Nahmias & Olsen (2015)
Single Exponential Smoothing	SES	α	Nahmias & Olsen (2015)
Double Exponential Smoothing	DES	α, β	Nahmias & Olsen (2015)
Croston-Based Methods			
Croston's Method	CRO	α, β	Croston (1972)
Syntetos - Boylan Approximation	SBA	α, β	Syntetos & Boylan (2005)
Teunter Syntetos Babai Method	TSB	α, β	Teunter et al. (2011)
Naïve Methods			
Random Walk	RW	-	Hyndman & Koehler (2006)
Zero Forecast	ZF	-	Teunter & Duncan (2009)

4.5.2. Model-fitting & Optimization

This section explains how the in-sample and out-of-sample demand history of every service part are used to derive the forecast performance of all benchmark methods. The first 12 in-sample months serve the initialisation of the first forecast per method (initialise demand interval for Croston's method e.g.). Appendix B explains this initialisation for the traditional and Croston-based methods in more detail. The remaining 24 in-sample months are used for the warming-up and optimization of all parameters involved. A distinction can be made between SMA and the remaining methods that use smoothing parameters. The SMA depends on parameter n, which describes the number of periods included for a forecast. In general, the more periods n taken into account, the more the forecast lags behind any change (Nahmias & Olsen, 2015). Furthermore, Johnston et al. (1999) suggest the use of a

n between the 3 and 12 data points when dealing with monthly spare parts demand data. In line with this suggestion, parameter n can take the following values during the optimization: $\{3, 6, 9, 12\}$.

The remaining traditional and Croston-based methods are based on smoothing parameters. The literature on spare parts demand forecasting recommends low smoothing parameter values when operating in an intermittent demand context (Syntetos & Boylan, 2005). Over the years, Croston (1972), Willemain et al. (1994) and Gutierrez, Solis & Mukhopadhyay (2008) specifically recommended the use of smoothing parameter values in the range of 0.05 – 0.20 when dealing with intermittent demand. In accordance with these findings, a subset of the following five smoothing parameters is considered: $\{0.05, 0.10, 0.15, 0.20, 0.25\}$. This means there are 25 possible parameter combinations for all methods with two smoothing parameters.

In this empirical evaluation study, all parameters are optimized in two different ways: by using the same configuration across all service parts in scope and on an individual service part basis. In both cases, the optimal configuration is defined by the parameters that minimize the MASE over the last 18 months of the in-sample months (November 2017 – May 2019). Month 13 until 18 are used as a warm-up period for the parameter configuration (May 2017 – October 2017). In this way, the impact of the initialisation on the parameter selection is reduced and the optimization considers more recent data points. At last, the 24 out-of-sample months are used for evaluation and comparison (May 2019 – April 2021). Figure 4.5 summarizes the process of initialising, optimizing and evaluating for the five years of demand data.

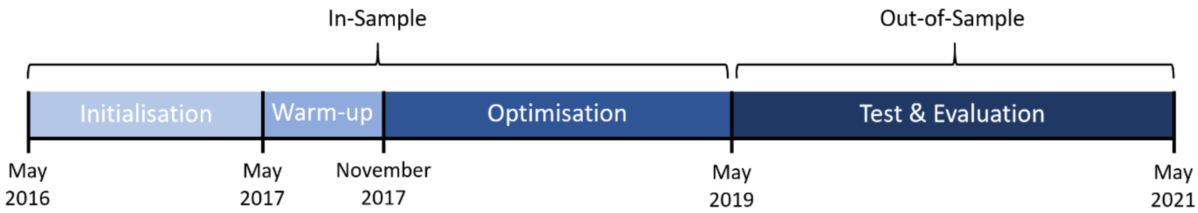

Figure 4.5: Five-year timeline of the spare parts demand data

4.5.3. Empirical Results

In this section, the demand data is used to find the optimal forecast performance for all benchmark methods. Furthermore, this performance is evaluated during the out-of-sample period. First, Section 4.5.3.1 considers the benchmark forecast performance for all service parts in scope. Afterwards, Section 4.5.3.2 analyses the effect of an increasing demand pattern, following the logic of Section 4.3.

4.5.3.1. Overall Forecast Performance

This section evaluates the forecast performance of both part parameter optimizations for all service parts in scope. Table 4.5 reports the optimal out-of-sample results of all benchmark methods, with a single optimal parameter configuration across all service parts. Besides, Table 4.6 presents the out-of-sample results with all parameters individually optimized. A comparison of the results in both tables indicates that for all methods the individual part optimization performs slightly better in terms of MSE and MASE. This is an interesting outcome, as the research by Babai et al. (2014) showed an opposite result. However, the individual part optimization performing worse in the study of Babai et al. (2014) was attributed to a low number of demand data points. The forecasting performance for the ZF and RW is the same in both tables because both methods have no parameters to optimize.

Table 4.5: Out-of-sample benchmark forecast performance with single parameter configuration optimization

Method	Optimal Parameters	SME	MSE	MASE
SMA	n = 12	0.099	7.170	0.994
SES	α = 0.05	0.287	8.832	0.989
DES	α = 0.25; β = 0.10	0.064	9.578	1.367
CRO	α = 0.20; β = 0.05	0.246	4.707	0.757
SBA	α = 0.20; β = 0.05	0.260	4.728	0.756
TSB	α = 0.20; β = 0.05	0.142	4.896	0.797
RW	-	0.015	5.673	1.059
ZF	-	1.000	31.389	0.830

Table 4.6 indicates that the three Croston-based methods perform closely to each other. TSB scores slightly better in terms of forecast bias, whereas SBA comes out best when considering the forecast error. In general, SMA, SES, TSB and RW are theoretically unbiased. However, only RW has a bias actually close to zero. Nevertheless, the naïve RW is among the worst when considering the MSE and MASE. As expected, the ZF has the largest positive bias. Despite a reasonable MASE, ZF suffers from a large MSE, which gives extra weight to large errors. In addition, Teunter & Duncan (2009) found that ZF is always outperformed by the other methods when used in an inventory control setting. The performance over the last two years can be directly compared with the results of the Enriched EWMA. Just like RW, the Enriched EWMA has a small (positive) bias but falls behind in terms of forecast error. Appendix C gives an extended overview of the forecast results per year of the out-of-sample period.

Table 4.6: Out-of-sample benchmark forecast performance with individual part parameter optimization

Method	SME	MSE	MASE
SMA	0.114	5.508	0.932
SES	0.198	6.645	0.982
DES	0.078	5.925	1.174
CRO	0.259	4.635	0.731
SBA	0.226	4.623	0.721
TSB	0.188	4.677	0.769
RW	0.015	5.673	1.059
ZF	1.000	31.389	0.830
Enriched EWMA	0.041	5.418	1.039

4.5.3.2. Forecast Performance with Increasing Demand Pattern

The next step is to find out how all benchmark methods perform when faced with strictly increasing demand, following the logic from Section 4.3. Table 4.7 presents the results of the individual part parameter optimization for the service parts with strictly increasing demand. According to Table 4.7, DES and RW come closest to incorporating the increasing trend of the demand, given the relatively small overall SME (bias) levels. However, both methods are among the worst when it comes to forecasting error. The opposite is true for the Croston-based methods, which deal with relatively high biases but are among the best in terms of forecast error. Moreover, the Enriched EWMA has a lower (positive) bias than SES and the Croston-based methods but is still relatively far away from zero. The combination of incorporating the latest installed base fluctuations and giving more weight to more recent observations does not prevent the Enriched EWMA from lagging behind the trend. In the end, the combination of intermittent and increasing demand causes problems for all methods. Currently, no method is able to successfully handle both bias and error when encountering this demand pattern.

Table 4.7: Out-of-sample benchmark forecast performance with strictly increasing demand

Method	SME	MSE	MASE
SMA	0.207	14.904	1.037
SES	0.239	16.578	1.054
DES	0.091	15.013	1.150
CRO	0.283	13.609	0.924
SBA	0.280	13.659	0.931
TSB	0.238	13.669	0.952
RW	0.032	15.480	1.225
ZF	1.000	88.513	1.331
Enriched EWMA	0.146	14.292	1.070

Table 4.8 presents the results for the service parts without strictly increasing demand. The demand levels of this set of service parts are stable over time and relatively low. Appendix A.1 presents more details about the average monthly demand levels and descriptive statistics for the subset of service parts without strictly increasing demand. Table 4.8 indicates that the Enriched EWMA even slightly over-forecasts this type of demand, but is still among the best in terms of forecast bias (SME). Nevertheless, the results in Table 4.8 give options to improve the forecast bias and/or error of the Enriched EWMA. The best option is presented by SMA, which scores better on all three performance metrics than the Enriched EWMA. Moreover, Croston-based methods like SBA and TSB come with a slightly higher (positive) bias but can generate large improvements in terms of forecast error. Thus, it is possible to already improve the current forecast performance when considering service parts with low and stable demand.

Table 4.8: Out-of-sample benchmark forecast performance without strictly increasing demand

Method	SME	MSE	MASE
SMA	0.015	3.304	0.907
SES	0.155	4.315	0.965
DES	0.064	3.794	1.399
CRO	0.235	2.530	0.675
SBA	0.169	2.504	0.671
TSB	0.136	2.568	0.726
RW	-0.003	3.373	1.020
ZF	1.000	17.993	0.713
Enriched EWMA	-0.068	3.339	1.032

5 Reliability-Based Forecasting Model

This chapter starts the solution design phase and comes up with a new forecasting methodology. The analysis and diagnosis phase found that all available methods have their problems with estimating an increasing demand pattern. Currently, the Enriched EWMA attempts to incorporate installed base size fluctuations in its forecast but still lags behind the trend. This chapter derives a new forecasting method in an attempt to capture this trend with additional installed base information. In this way, an answer to Research Question 4 is provided. Deriving a new method and answering this question consists of several steps. First of all, Section 5.1 uses the explanatory role of installed base information to conceptualize a reliability-based forecasting model (RFM). Afterwards, Section 5.2 turns this conceptual plan of the RFM into a detailed design. At last, Section 5.3 and 5.4 describe how installed base information is integrated into the RFM by deriving an overview of the installed service parts and the failure behaviour of service parts respectively.

5.1. Conceptual Plan

In this section, the main ideas behind a so-called reliability-based forecasting model are examined. First, Section 5.1.1 discusses the explanatory role and potential value of installed base information. Moreover, Section 5.1.2 summarizes how the current forecasting process uses installed base information in an attempt to capture this potential. Afterwards, Section 5.1.3 conceptualizes a new forecasting model which expands these current beliefs on forecasting with installed base information.

5.1.1. Explanatory Role Installed Base Information

The installed base of ASML has been growing gradually for years (Figure 4.2). In the meanwhile, the average demand for spare parts has been following this trend (Figure 4.1), which is in line with the findings by Dekker et al. (2013). They described the relationship between the evolution of the installed base size and spare parts demand. Figure 5.1 visualizes this relationship by indicating the delay between installed base fluctuations and spare parts demand changes.

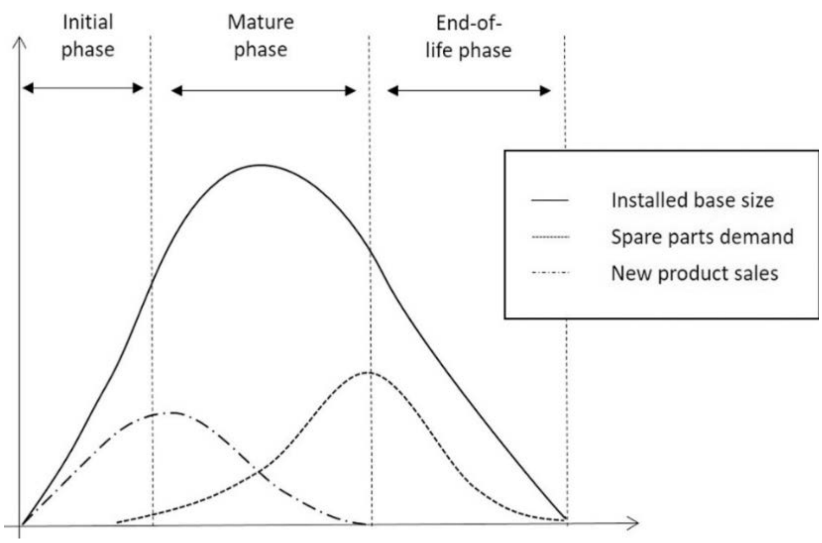

Figure 5.1: Relationship between the installed base size and spare parts demand (Dekker et al., 2013)

The combination of growing global demand for microchips and ASML's leading role in the microchip supply chain offers no reason to doubt any further increase in the overall installed base in the coming years. On the other hand, older machine types reach their end-of-life phase, causing a decline in the installed base for these machine types. Thus, the composition and size of the overall installed base will start to change even more over time. According to Dekker et al. (2013) and Van der Auweraer & Boute (2019), forecasting with installed base information has the most potential when the size of the installed base does change over time. Therefore, it can be concluded that the potential value of forecasting with installed base information increases over time for ASML.

5.1.2. Current Application Installed Base Information

The current forecasting process already attempts to capture installed base changes by including install base size fluctuations. Section 3.2 explained how ASML currently estimates the applicable installed base size of a service part to enrich the time series data. This makes the Enriched EMWA a hybrid method that combines time series data with feature-based elements. In more detail, the number of active machines in a given period is now directly related to the spare parts demand coming from the same period. In this way, the forecasting process does not consider any delayed effect between changes in installed base and spare parts demand. Furthermore, the current applicable installed base size estimation is done on machine-level, which does not necessarily reflect the active number of service parts in the field. The analysis in Section 4.3.2 indicated that the current forecasting process still lags behind an increasing demand pattern, with the current application of installed base information. Thus, this process is not able to fully capture the potential of forecasting with installed base information.

5.1.3. Reliability-Based Forecasting Model Concept

In order to capture the potential of installed base information, the concept of a reliability-based forecasting model (RFM) is introduced. The main idea behind this model is to more accurately describe the relationship between the installed base and spare parts demand. Currently, ASML registers more information for a service part than only the applicable installed base size. In general, all material flows of a service part within ASML have to be registered. From these registrations, it is possible to deduce when a specific service part failed and how long it lasted. The RFM can use these service part failure observations to estimate the failure behaviour of a service part. This failure behaviour directly describes the expected delay between installed base changes and spare parts demand. Moreover, the RFM can use the service part registrations to create an overview of the current active installed base and its age distribution. The combination of a set of installed service parts and their failure behaviour allows the RFM to estimate the conditional probability of failure for each of these service parts in a given time horizon. This line of reasoning forms the foundation under the RFM and uses installed base information in a different way than the current forecasting process. The RFM considers the active installed base on a service part level instead of the current machine level. In addition, the current age distribution of this active installed base can be recognized. Therefore, the RFM uses a more representable and detailed overview of the actual installed base than the current forecasting process. Section 5.2 works out the RFM concept in more detail and explains how it can be used to derive forecasts of spare parts demand.

5.2. Reliability-Based Forecasting Model Design

In this section, the RFM is built by capturing the conceptual ideas devised in Section 5.1.3. The main objective of reliability-based forecasting is to estimate the required number of spare parts coming from the installed base, during a given forecast horizon. The design of the RFM is done step-by-step, starting with an introduction of the required theoretical assumptions and model notation in Section 5.2.1. Afterwards, the ideas behind the RFM are conceptualized in Section 5.2.2. At last, Section 5.2.3 describes how the RFM can produce forecasts at any desired aggregation level.

5.2.1. Assumptions & Notation

Theoretical assumptions are required when deriving a forecasting model based on service part failure behaviour. First, it is assumed that the operational environment of the installed base has no impact on the failure behaviour of a service part. Therefore, differences in machine types or machine usage rates are of no importance in this model. Moreover, ASML released newer part versions for most of the service parts over the years. In the current forecasting process, no distinction is made between these different versions, as described in Section 3.2. The RFM assumes that these different service part versions all have an identical failure behaviour. Furthermore, the failure behaviour is also identical for service parts that are installed at different points in time. At last, when a service part is built-in, it is always considered to be as-good-as-new, independent of any usage or repairs in the past. Table 5.1 denotes the notation that is used throughout this research to describe the RFM.

Table 5.1: Reliability-based forecasting model notation

Notation	Description
a	The service part age.
β	The Weibull shape parameter.
η	The Weibull scale parameter.
i	The forecast horizon.
k	An in-scope service part.
$p_{a,i}^{k}$	The conditional probability of a service part k with age a to fail in forecast horizon i.
t	The current point in time when a forecast is made.
$n_{a,t}^{k}$	The number of installed service parts k with age a at time t.
$\hat{x}_{t,t+i}^{k}$	The demand forecast at time t during the forecast horizon i for service part k.
A_t^k	The age of the oldest service part k at time t.
N_t^k	The total number of installed service parts k at time t.
T^k	The lifetime random variable of service part k.

5.2.2. Detailed Model Description

In this section, the RFM is further conceptualized by using the model notation from Table 5.1. The conditional probability $p_{a,i}^{k}$ describes the chance that a service part k with age a will fail during a forecast horizon of length i. Given the assumptions, this conditional probability can be found independent of the current point in time t. Therefore, this probability only depends on the current age of the service part and the length of the forecast horizon. Equation 5.1 formally presents the conditional probability $p_{a,i}^{k}$. Later on, Section 5.4 explains how failure behaviour modeling can be used to derive these conditional probabilities based on service part lifetime observations.

$$p_{a,i}^{k} = P\left(T^{k} \le a + i \mid T^{k} > a\right) = \frac{P\left(T^{k} \ge a\right) - P\left(T^{k} \ge a + i\right)}{P(T^{k} > a)} \tag{5.1}$$

Thus, when making a forecast at time t, it is important to know the age distribution of all installed service parts at this point. Parameter $n_{a,t}^{k}$ is introduced to count the number of service parts with age a at time t for service part k. Furthermore, parameter A_t^k defines the age of the oldest service part k at time t. Summing over all ages up to and including A_t^k equals the total number of installed service parts k at time t, as described by Equation 5.2.

$$N_t^k = \sum_{a=1}^{A_t^k} n_{a,t}^{k} \tag{5.2}$$

Two additional assumptions are required before the conditional probabilities and the installed part age distribution can be combined into a spare parts demand forecast. First, it is assumed that any service part newly installed during the forecast horizon $[t, t+i]$ cannot fail during the same forecast horizon. This leaves a set of N_t^k parts that can either fail or survive during the forecast horizon $[t, t+i]$. Additionally, for each of these service parts, it is assumed that their time-to-failure is independent. This assumption generally holds for most parts, as they are installed in different machines. Subsequently, the conditional probability of surviving the forecast horizon is equal to one minus the conditional probability of failing. In this way, the demand resulting from a specific installed service part during the time horizon $[t, t+i]$ is Bernoulli distributed. However, the forecast of the probability distribution of the total demand for this type of service part on the same time horizon does not follow a binomial distribution (Van der Auweraer & Boute, 2019). The Bernoulli random variables are independent of each other, but not identically distributed. The service parts do not all have the same age, which means the expected values are also different. In this case, the sum of independent Bernoulli trials can be described by a Poisson binomial distribution (Hong, 2013). Therefore, the spare parts demand of service part k during the time horizon $[t, t+i]$ is equal to a Poisson binomial distributed variable. The expected spare parts demand for service part k is then simply equal to the sum of the mean of N_t^k independent Bernoulli trials.

When the age distribution of the currently installed service parts at time t is known, the conditional probability $p_{a,i}^{k}$ can be determined for all N_t^k parts. This conditional probability $p_{a,i}^{k}$ describes the expected value of the Bernoulli trial for each service part. Thus, the expected number of failures is equal to the sum of the conditional probabilities. In order to ease the computation, the number of parts $n_{a,t}^{k}$ with age a at time t can be multiplied with conditional probability $p_{a,i}^{k}$. This gives the expected number of failures for this group of service parts k with age a during the time horizon $[t, t+i]$. The total expected number of failed service parts k over a time horizon $[t, t+i]$ can be found by summing over all applicable ages a of the service parts installed at time t. This expected number of failures is directly used as a spare parts demand forecast of service part k for the given time horizon. The spare parts demand forecast of the RFM is summarized by equation 5.3.

$$\hat{x}_{t,t+i}^{k} = \sum_{a=1}^{A_t^k} n_{a,t}^{k} p_{a,i}^{k} \tag{5.3}$$

The RFM manages to combine three different types of installed base information: the active number of installed service parts, the current service part age distribution and observations of service parts that failed. In this way, any change in the installed base of service parts is directly reflected by the forecast. Besides, considering the current age distribution allows for capturing differences in failure behaviour over a service part's lifetime. Section 5.3 and 5.4 describe how this installed base information can be used to derive the parameters $n_{a,t}^{k}$ and $p_{a,i}^{k}$ respectively, but first Section 5.2.3 explains how the RFM can be applied at any desired aggregation level.

5.2.3. Forecast Aggregation Level

In the current configuration, the RFM represents the total worldwide demand for a given service part. However, Section 3.2.1 explained that the planning processes at ASML also require forecasts of spare part demand at more local levels. This section also described how the current forecasting process generates forecasts at a local level. In summary, the Enriched EWMA first derives a forecast of the worldwide demand usage rate. Subsequently, this demand usage rate and the local installed base size are used to break down the forecast to local levels. With the RFM, ASML can also derive forecasts for any desired aggregation level. The model assumes that all installed service parts are subjected to an independent Bernoulli trial. Moreover, the expected spare parts demand from any aggregation of service parts can be described by a Poisson binomial distribution. Therefore, a local level forecast can be found by summing the conditional probabilities of all installed service parts that meet the aggregation requirements. This allows the RFM to make spare parts demand forecasts for specific regions, machines or customers. The RFM then considers detailed information about the age distribution of a local group of service parts when deriving a local level forecast.

5.3. Installed Service Parts Overview

The RFM requires an overview of the installed service parts and their ages at the moment of forecasting. The registration of service parts at ASML resolves around the Configuration as Maintained (CAM), which has been introduced in Section 3.2.2. The CAM describes the whole configuration of a machine as it is currently maintained in the field. In general, this CAM registers all service parts that are currently installed in any machine under the guidance of ASML. Moreover, for each of these installed service parts, the moment of built-in is registered. Thus, the CAM registrations can be used to derive an overview of the active service parts and their current ages. In this way, it is possible to determine parameter $n_{a,t}^{k}$ for service part k at time t for all ages a up to and including A_t^k.

The configuration of the set of installed service parts changes almost constantly. On the one hand, new service parts are added to the installed base when new machines are installed. On the other hand, an installed service part can be replaced by a new service part. In the last case, this can have multiple reasons. The most prominent reason concerns a service part defect that directly requires a replacement. Therefore, a service part defect directly triggers spare parts demand at a warehouse in the service network of ASML. This is the type of demand the RFM tries to forecast. However, not every replacement follows a failure of its predecessor. Service parts could also have been replaced by an upgraded service part version, on customer request for example. Demand for these types of upgrades is planned and not forecasted by the RFM. Nevertheless, independent of the replacement reason, any change in the installed base is relevant for the RFM. New service part installs result in additional service parts that can potentially trigger spare parts demand. Furthermore, service part replacements always trigger a change in the age distribution. In the end, every built-out of a specific service part affects the RFM in some way. Therefore, it is important to check on the latest CAM changes every time the RFM has to produce a forecast.

5.4. Service Part Failure Behaviour

The failure behaviour of a service part is another crucial element of the RFM. An estimation of the failure behaviour is required to find the conditional failure probability $p_{a,i}^{k}$ for a service part k with age a during the forecast horizon $[t, t+i]$. This section discusses how service part failures from the past can be used to find this estimation and the conditional probabilities for a service part k. First, Section 5.4.1 discusses why the Weibull distribution is the most suitable for modelling the failure behaviour. Moreover, Section 5.4.2 derives a lifetime-driven Weibull parameter estimation procedure, whereas Section 5.4.3 comes up with a demand-driven alternative. This demand-driven procedure provides estimations that are more robust in comparison with the traditional lifetime-driven approach.

5.4.1. Failure Behaviour Modelling

An appropriate failure behaviour modelling approach is required to translate service part failure observations from the past into conditional failure probability estimates. Sections 5.4.1.1 and 5.4.1.2 discuss the choices for a parametric modelling approach and the use of the Weibull distribution respectively.

5.4.1.1. Parametric Modelling Approach

The literature on reliability and survival analysis contains both parametric and non-parametric approaches to model the failure behaviour of service parts. A commonly used non-parametric approach is the Kaplan-Meier estimate. This estimate tries to find a fitting reliability function without assuming any distribution (Kaplan & Meier, 1958). In general, the reliability function of the Kaplan-Meier estimate is a series of declining horizontal steps. The more failure observations become available, the closer this function approaches the actual reliability function. On the other hand, a parametric approach tries to fit a univariate non-negative distribution to model the time-to-failure.

The non-parametric approach by Kaplan-Meier (1958) delivers a piecewise constant reliability function. This can give problems with estimating the conditional probabilities when having a limited set of failure observations and/or a short forecast horizon. However, when fitting the right parametric model, this yields a smooth reliability function with possibilities to extrapolate beyond the range of the data. This extrapolation can help to make some estimation at ranges for which limited or no failure observations are available yet. This especially applies to service parts that have not yet been in use for a very long time. Additionally, a parametric approach makes it possible to derive a closed-form expression for conditional probability $p_{a,i}^{k}$. In general, parametric models provide greater efficiency, because fewer parameters have to be estimated. Given this reasoning, a parametric approach is preferred over a non-parametric approach because precise and efficient estimations are required.

5.4.1.2. Weibull Distribution

This section provides the reasoning behind the use of the Weibull distribution to model the failure behaviour of service parts when following a parametric approach. The survival probability distribution of a service part is always strictly positive and in most cases (highly) skewed to the right (Turrini & Meissner, 2019). This description indicates that a normality assumption is not appropriate for the RFM. Nevertheless, the exponential and Weibull distribution do comply with this finding and are most used during past research on service parts failure behaviour (Van der Auweraer et al., 2019). The exponential distribution deals with a constant failure rate due to its memory-less property. Having a constant failure rate is a simplifying assumption which facilitates calculations. However, the exponential distribution overestimates or underestimates the failure behaviour of a part when the actual failure rate changes over its useful life (Wessels, 2007). Furthermore, the assumption that the failure rate does not change over its useful life does not hold in practice most of the time. On the other hand, the Weibull distribution is very flexible and can more accurately describe this complex failure behaviour of service parts shown in practice (Wessels, 2007; Starling, Mastrangelo & Choe, 2021). The Weibull distribution includes the effect of aging and can generalize an exponential distribution whenever needed. Therefore, the Weibull distribution is preferred over the exponential distribution.

In general, the Weibull distribution makes use of two parameters: shape parameter β and scale parameter η. The shape parameter describes the slope of the probability distribution function, whereas the scale parameter ensures this function is scaled to the right proportions. The probability density function (PDF) of the Weibull distribution is given by equation 5.4.

$$f_T(t) = \left(\frac{\beta}{\eta}\right)\left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^{\beta}}, \qquad t \geq 0, \eta > 0, \beta > 0 \tag{5.4}$$

When shape parameter β is set to 1, the Weibull distribution reduces to an exponential distribution. Furthermore, the reliability function $R(t)$ and failure rate function $\lambda(t)$ of Weibull's model are given by equations 5.5 and 5.6 respectively (Arts, 2017).

$$R(t) = 1 - F_T(t) = e^{-\left(\frac{t}{\eta}\right)^{\beta}} \tag{5.5}$$

$$\lambda(t) = \lim_{\varepsilon \to 0} \frac{P(T \leq t + \varepsilon | T \geq t)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{F_T(t+\varepsilon) - F_T(t)}{\varepsilon} \frac{1}{R(t)} = \frac{f_T(t)}{R(t)} = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \tag{5.6}$$

The failure rate describes the number of expected failures per time unit. In the case of $\beta \neq 1$, the failure rate changes over time. In specific, with $\beta < 1$, the probability of failing during a given forecast horizon decreases over time, whereas $\beta > 1$ means this probability increases with time. This property allows the Weibull distribution to more accurately compute the required conditional probabilities $p_{a,i}^k$. The derivation of conditional probability $p_{a,i}^k$ in equation 5.1 can be translated into a closed-form expression by using the Weibull reliability function from equation 5.5. The derivation of this closed-form expression is presented by equation 5.7.

$$p_{a,i}^k = \frac{P(T^k \geq a) - P(T^k \geq a+i)}{P(T^k > a)} = \frac{F_{T^k}(a+i) - F_{T^k}(a)}{1 - F_{T^k}(a)} = \frac{e^{-\left(\frac{a}{\eta}\right)^{\beta}} - e^{-\left(\frac{a+i}{\eta}\right)^{\beta}}}{e^{-\left(\frac{a}{\eta}\right)^{\beta}}} = 1 - e^{\left(\frac{a}{\eta}\right)^{\beta} - \left(\frac{a+i}{\eta}\right)^{\beta}} \tag{5.7}$$

In the end, the RFM requires two Weibull parameters to describe service part failure behaviour and find estimates for conditional probabilities $p_{a,i}^k$. In this research, two different procedures have been derived to estimate both Weibull parameters based on failures seen in the past. Section 5.4.2 provides a lifetime-driven estimation procedure that uses the lifetimes of (failed) service parts to fit both parameters. Moreover, Section 5.4.3 introduces a more robust alternative estimation procedure based on the timing of service part failures and the demand it triggered instead of their lifetimes.

5.4.2. Lifetime-Driven Weibull Parameter Estimation Procedure

One option to describe the failure behaviour of a service part is by generalizing past observations of service part lifetimes. Section 5.4.2.1 explains which type of lifetime observations are available at ASML for this purpose. Furthermore, Section 5.4.2.2 defines how a maximum likelihood estimator can translate these observations into Weibull parameters.

5.4.2.1. Service Part Lifetime Observations

The service part registrations in the CAM can do more than only describe the set of installed service parts. These registrations can also be used to define how long a service part lasted before it failed. In this case, there are two types of relevant lifetime observations available: uncensored and right-censored observations.

Uncensored lifetime observations describe the time that passed before a service part failed and triggered demand. When a service part defect occurs, a service order is created to request all required materials from the warehouse. This service order contains information about what has been built in and built out during the service action and when this took place. The defect service part that has been built out is classified as a Field Service Defect (FSD). All changes described by the service orders are registered in the CAM eventually. The set of uncensored lifetime observations contains all service parts that have been classified as FSD in the past. Their time-to-failure is found by linking the FSD registration back to the moment of install.

Right-censored lifetime observations describe a lower bound of the actual time-to-failure of a service part. These are observations of service parts that did not fail (yet) but did already reach some age. In general, all current ages of the installed service parts, derived by Section 5.3, can be denoted as right-censored lifetime observations. Figure 5.2 visualizes an example of uncensored and right-censored service part lifetimes observations over time. This type of data collection is referred to as generalized Type-I right censoring because all observations have varying introduction dates (Starling et al., 2021). The combination of uncensored and right-censored lifetime observations serves as input to derive the Weibull parameters with a maximum likelihood estimator. Only including uncensored observations can lead to a wrong estimation of the actual failure behaviour. The example in Figure 5.2 can be used to briefly explain this. In this case, two uncensored observations followed upon a failure. However, all three right-censored observations of currently installed parts surpass the lifetimes of both failed parts. When only including the uncensored observations, the expected lifetime of the service part is underestimated, which can result in a demand overestimation.

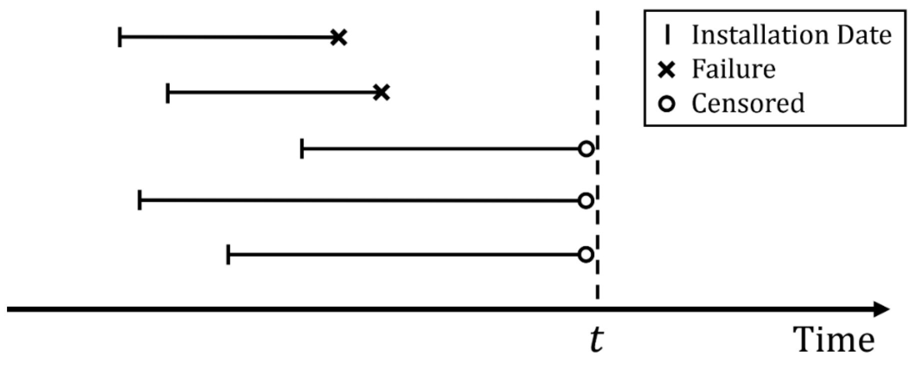

Figure 5.2: Example of uncensored and right-censored service part lifetimes

5.4.2.2. Maximum Likelihood Estimator

The lifetime-driven Weibull parameter estimation procedure uses a maximum likelihood estimator (MLE) to derive both Weibull parameters for a service part. The maximum likelihood estimator is the most often used method when dealing with the generalized Type-I right censoring data (Starling et al., 2021). The MLE estimates probability distribution parameters by maximizing a likelihood function. The likelihood function is defined as the product of the densities of each data point. A distinction is made between uncensored observations u and right-censored observations c, which are assumed to be independent of each other. In total, the MLE considers U uncensored and C right-censored lifetime observations. The likelihood function for fitting the Weibull parameters with the combination of both types of observations is given by equation 5.8 (Nelson, 2003). The likelihood of right-censored observations c_j, with $j \in C$, is the probability that its lifetime is larger than c_j.

$$\begin{aligned} L(u_1, \dots, u_U;\ c_1, \dots, c_C) &= \prod_{i=1}^{U} f_T(u_i) \cdot \prod_{j=1}^{C} \left(1 - F_T(c_j)\right) \\ &= \prod_{i=1}^{U} \left(\left(\frac{\beta}{\eta}\right) \left(\frac{u_i}{\eta}\right)^{\beta-1} e^{-\left(\frac{u_i}{\eta}\right)^{\beta}} \right) \cdot \prod_{j=1}^{C} \left(e^{-\left(\frac{c_j}{\eta}\right)^{\beta}} \right) \end{aligned} \tag{5.8}$$

The objective is to find the optimal (β, η)-combination which maximizes this likelihood function. In order to achieve this, equation 5.8 must be differentiated with respect to both β and η. Afterwards, both equations have to be set equal to zero. The solution of the combination of both equations results in the best fitting Weibull parameter configuration. Equations 5.9 and 5.10 show the result when trying to solve for η and β respectively. All steps required to go from equation 5.8 to equations 5.9 and 5.10 are described in detail in Appendix D.

$$\eta^{\beta} = \frac{1}{U} \left(\sum_{i=1}^{U} \left(u_i^{\beta}\right) + \sum_{j=1}^{C} \left(c_j^{\beta}\right) \right) \tag{5.9}$$

$$\frac{1}{U} \sum_{i=1}^{U} \ln(u_i) + \frac{1}{\beta} - \ln(\eta) = \left(\frac{1}{\eta^{\beta} U}\right) \left(\sum_{i=1}^{U} u_i^{\beta} \ln(u_i) + \sum_{j=1}^{C} c_j^{\beta} \ln(c_j) - \ln(\beta) \left(\sum_{i=1}^{U} u_i^{\beta} + \sum_{j=1}^{U} c_j^{\beta} \right) \right) \tag{5.10}$$

Replacing parameter η in Equation 5.10 by the estimation in Equation 5.9 delivers Equation 5.11:

$$\frac{1}{\beta} = \frac{\sum_{i=1}^{U} u_i^{\beta} \ln(u_i) + \sum_{j=1}^{C} c_j^{\beta} \ln(c_j)}{\sum_{i=1}^{U} u_i^{\beta} + \sum_{j=1}^{C} c_j^{\beta}} - \frac{1}{N} \sum_{i=1}^{U} \ln(u_i) \tag{5.11}$$

Unfortunately, there is no closed-form solution for the derivation in equation 5.11. An iterative procedure is required to find an estimation for shape parameter β. Afterwards, this estimation can be used to find a value for parameter η, by using equation 5.9. Programming language R (RStudio) is used to perform this iterative procedure with additional R-package 'fitdistrplus' (Delignette-Muller & Dutang, 2015). This package makes use of the numerical method by Nelder and Mead to solve equation 5.11 and optimize the likelihood function given by equation 5.8. In this way, the Weibull parameters β and η can be found for a service part. At last, Nelson (2003) points out that the MLE can produce reliable Weibull parameter estimations with at least 15 uncensored observations.

5.4.3. Demand-Driven Weibull Parameter Estimation Procedure

The lifetime-driven Weibull parameter estimation procedure requires the lifetimes of failure observations from the past to fit the Weibull parameters. However, these uncensored lifetime observations are not always available in practice. These observations are derived with the help of FSD registrations. In general, the number of FSD registrations should be in line with the demand for spare parts. Unfortunately, this is not always the case in practice. Without an FSD registration, it is currently not possible to derive an accurate uncensored lifetime observation for a defect service part that triggered spare part demand. Moreover, the MLE is slightly biased when the sample of uncensored observations is low and/or the censoring level is high (Starling et al., 2021). There can be multiple reasons for missing FSD registrations. A service engineer in the field is responsible for all registrations on the service order. This engineer could simply have forgotten to register it at the service order. A more structural reason is that some built-out parts are immediately scrapped, which makes an FSD registration redundant.

A demand-driven Weibull parameter estimation procedure is proposed to approximate the Weibull parameters when accurate lifetime observations are unavailable. This more robust approach requires an overview of the installed service parts and their age distribution during the past. This overview indicates for a specific service part which equivalent service parts were installed with a certain age during a given month. Moreover, this approach needs the aggregated demand produced by this installed base of service parts during these months. This monthly demand data is currently available for all service parts and does not depend on FSD registrations. Next, the demand-driven approach estimates the Weibull parameters by linking the overview of installed service parts to the monthly aggregates of spare parts demand produced by these service parts. In more detail, the failure behaviour of this set of installed service parts is determined by finding the Weibull parameters that minimize the forecast error of the RFM on a monthly basis. In this way, the shape and scale parameters approximate the overall failure behaviour that is most applicable to the entire set of installed service parts. Thus, the demand-driven parameter estimation is done on an installed base level, whereas the lifetime-driven approach from Section 5.4.2 operates on a service part level. Chapters 6 and 7 help to estimate to what extent data availability and accuracy affect the forecast performance of the RFM for both estimation procedures.

6 Simulation Study

In this chapter, simulation experiments are set up to estimate and validate the potential of the reliability-based forecasting model (RFM). This simulation study uses the Weibull distribution to generate failures and spare parts demand from a set of installed service parts. The objective of this study is to test how accurately the RFM can predict this demand while using the lifetime-driven and demand-driven Weibull parameter estimation procedures. Moreover, this simulation study allows the RFM to forecast the demand while using the same Weibull parameters that generated this demand in the first place. In this way, this study helps to estimate the effect of data availability and the Weibull parameter estimation procedures on the forecast performance of the RFM. First, the design of this study is addressed in Section 6.1, which includes a discussion on the simulation conditions and setup. Afterwards, the results of all simulation experiments are presented and evaluated in Section 6.2.

6.1. Simulation Study Design

In this simulation study, the performance of the RFM is tested under varying conditions. Therefore, this simulation study makes use of several parameters. Different combinations of these parameters help to recreate these varying conditions. Section 6.1.1 explains which relevant parameters are considered to generate spare parts demand. Afterwards, Section 6.1.2 discusses which parameter configurations are put to a test during separate simulation experiments. Finally, Section 6.1.3 addresses in more detail how these simulation experiments are carried out and evaluated.

6.1.1. Simulation Parameters

This section introduces all parameters and discusses their relevance. First of all, the Weibull distribution is used to generate service part failures in this simulation. This requires Weibull parameters that capture the typical failure behaviour of spare parts. Turrini & Meissner (2019) denoted that spare part demand is typically (highly) skewed to the right. The shape parameter β of the Weibull distribution describes how the slope of the probability distribution function behaves over time. In this simulation, three shape parameters are used which comply with a strictly positive and decreasing probability distribution function: $\beta \in \{0.9, 1.1, 1.5\}$. This range of values is typical for spare parts failure behaviour and includes both decreasing ($\beta < 1$) and increasing failure functions ($\beta > 1$) (Van der Auweraer & Boute, 2019). Furthermore, scale parameter η determines the scale of the time-to-failure in the Weibull distribution. An important property of the Weibull distribution states that the scale parameter defines the 63.2 percentile of the data:

$$F_T(\eta) = 1 - e^{-\left(\frac{\eta}{\eta}\right)^{\beta}} = 1 - e^{-1} = 0.632 \tag{6.1}$$

Therefore, a higher scale parameter increases the average lifetime of a service part. In this study, the service part lifetime is measured in months and the scale parameter can adopt three different values: $\eta \in \{100, 500, 1000\}$. This range of values typically scales the Weibull distribution to the proportions of actual spare parts and includes both slow-moving and fast-moving spare parts (Van der Auweraer & Boute, 2019). In combination with the shape parameters, this results in a set of nine possible options to examine. The Weibull parameters do not vary over time and are the same for all service parts involved in a simulation experiment, which is in line with the assumptions stated in Section 5.1.1.

The failures in the simulation are generated from a set of installed service parts. Every simulation experiment starts with an initial number of installed service parts, which are considered as good as new. This number is represented by parameter Γ. The values of this parameter reflect the actual size of ASML's installed base and have been standardized for confidentiality reasons: $\Gamma \in \{100, 400\}$.

In practice, the total number of installed service parts can change over time, caused by new installations or discards. In this simulation study, it is assumed that the number of installations and discards of service parts both follow a linear process over time. This assumption is justified by the linear behaviour of the average applicable installed base size in Figure 4.2 of Chapter 4. Moreover, linear increasing and decreasing patterns in applicable installed base size can be recognized on an individual level for most service parts. This study allows for different monthly installation and discards rates to be tested. Both rates are defined as a fraction of the initial installed base size Γ. The installation and discard rates of service parts are described by parameters γ and ρ respectively. In this study, both parameters can adopt two standardized values: $\gamma, \rho \in \{2, 4\}$ per month. Combinations of both parameters allow for the replication of the three phases in the installed base lifecycle, as depicted in Figure 5.1. In conclusion, Table 6.1 presents an overview of all parameters involved.

Table 6.1: Simulation study parameter overview

Parameter	Description	Values
β	Weibull Shape Parameter	$\{0.9, 1.1, 1.5\}$
η	Weibull Scale Parameter	$\{100, 500, 1000\}$
Γ	Initial Installed Base Size	$\{100, 400\}^*$
γ	Service Part Installation Rate	$\{2, 4\}^*$ per month
ρ	Service Part Discard Rate	$\{2, 4\}^*$ per month

**Standardized due to confidentiality reasons.*

6.1.2. Simulation Experiment Parameter Configurations

In this section, a set of parameter configurations is derived to test the potential of the RFM under varying circumstances. Different combinations of parameters allow for a controlled comparison between those circumstances. In total there is a set of five parameters that can vary: $\{\beta, \eta, \Gamma, \gamma, \rho\}$. Every parameter configuration $\{\beta, \eta, \Gamma, \gamma, \rho\}$ is considered as an individual simulation experiment. Parameters γ and ρ are used to recreate demand patterns caused by the three phases from the installed base lifecycle: the initial, mature, end-of-life phases (Figure 5.1). These phases describe an increasing, stable and decreasing installed base size respectively. The initial and mature phases use an initial installed base size of 100 service parts, whereas the end-of-life phase starts with 400. Subsequently, the failure behaviour of the service part involved is described by the nine possible Weibull parameter combinations from Table 6.1. In total, this gives 27 simulation experiments that give insight into the effect of the installed base fluctuations and failure behaviour on the RFM. Table 6.2 presents an overview of the parameter configurations.

Table 6.2: Parameter configuration overview of all simulation experiments

Installed Base Lifecycle	$\boldsymbol{\beta}$	$\boldsymbol{\eta}$	$\boldsymbol{\Gamma}$	$\boldsymbol{\gamma}$	$\boldsymbol{\rho}$	Simulation Experiments
Initial Phase	$\{0.9, 1.1, 1.5\}$	$\{100, 500, 1000\}$	100	4	2	$1-9$
Mature Phase	$\{0.9, 1.1, 1.5\}$	$\{100, 500, 1000\}$	100	2	2	$10-18$
End-of-Life Phase	$\{0.9, 1.1, 1.5\}$	$\{100, 500, 1000\}$	400	2	4	$19-27$

6.1.3. Simulation Experiment Setup

The setup of the simulation experiments from Table 6.2 has to be addressed in more detail before any result can be obtained. First, Section 6.1.3.1 briefly discusses how the simulation experiments generate and handle service part failures. Afterwards, the application of the RFM with different Weibull parameter estimation procedures is introduced in Section 6.1.3.2. Finally, Section 6.1.3.3 examines which benchmark methods and performance metrics are used.

6.1.3.1. Service Part Failure Simulation

In general, every simulation experiment has a duration of ten years, which is equivalent to 120 months. The first eight years are used to initialise and warm up the simulation. Furthermore, the forecast performance is only evaluated on the spare parts demand generated during the last two years of the experiment. Thus, the first eight years are considered in-sample, whereas the last two years serve as an out-of-sample set. In these experiments, every new service part is assigned a Weibull distributed time-to-failure. When an active service part fails, it is assumed that a new replacing service part is installed within one month. The simulation carries out the monthly discards by removing the service parts that have been installed the longest, which includes the time potential predecessors have been installed. Each of the parameter configurations from Table 6.2 is tested in a simulation experiment consisting of 1000 independent runs. These experiments are carried out in Microsoft Excel. The Weibull distributed lifetime of a service part is generated by applying the inverse transform method on the Weibull cumulative distribution function (CDF). The procedure of the inverse transform method with the Weibull CDF is summarized by the following three steps:

1. Generate a random number from the standard uniform distribution: $U \sim Uniform[0,1]$.
2. Take the inverse of the Weibull CDF: $F_T(t) = U \rightarrow F_T^{-1}(U) = \eta(-ln(1-U))^{1/\beta}$.
3. Return a Weibull distributed time-to-failure: $T = F_T^{-1}(U)$.

6.1.3.2. Reliability-based Forecasting Models

The simulation experiments test three different configurations of the RFM. First, the RFM is tested for the lifetime-driven and demand-driven Weibull parameter estimation procedures, which are derived in Sections 5.4.2 and 5.4.3 respectively. Moreover, the RFM is assessed while using the same Weibull parameters that are used to generate the failures, meaning there is a perfect parameter fit. In this section, it is briefly explained how the lifetime-driven and demand-driven approaches are initialised.

The lifetime-driven approach uses the MLE to estimate the Weibull parameters based on uncensored and right-censored lifetime observations (Section 5.4.2). These lifetime observations are based on all built-in and built-out registrations during the in-sample period. All service parts with a built-in and built-out registration during this period are included as uncensored observations. Moreover, the service parts already built in during the in-sample period but not built out before the end of this period, are included as right-censored observations. The length of the right-censored lifetime observations is equal to the difference between the built-in timestamp and the end of the in-sample period. Every simulation run applies the MLE to those lifetime observations.

The demand-driven approach estimates the Weibull parameters based on the demand generated during the in-sample period. This procedure uses the overview of installed service parts in combination with the monthly demand produced by these service parts during the in-sample period. Every simulation run, the Weibull parameters are estimated by running the RFM during this in-sample period while minimizing the average forecast error, which is defined by the MASE. The Weibull parameters estimated by both procedures are fixed during the entire out-of-sample period.

6.1.3.3. Benchmark Methods & Performance Metrics

Every simulation experiment includes three forecasting methods that help to benchmark the performance of the RFM. First, the Enriched EWMA is involved to directly compare the RFM with the current practices. Moreover, two additional time series-based benchmarks are included: Double Exponential Smoothing (DES) and the Syntetos - Boylan Approximation (SBA). DES and the SBA have been selected because they scored best in terms of forecast bias and error respectively during the empirical benchmark evaluation in Section 4.5. DES is preferred over the Random Walk (RW) method because RW by definition lags behind the actual demand. This empirical study found that for both methods the smoothing parameter combination of $\alpha = 0.20$ and $\beta = 0.05$ came out as optimal when using a single optimal parameter configuration (Table 4.5). Therefore, this combination is applied throughout all simulation experiments for both methods. In addition, the methods use the in-sample period to initialise and warm-up, following the logic from Appendix B.

Finally, the forecast performance during the out-of-sample period is assessed by the SME and MASE. The scaling aspect of both methods allows for comparisons between simulation experiments that deal with different demand levels. The MASE uses the last three in-sample years for initialization. The overall forecast performance during a simulation experiment is found by taking the arithmetic mean of the performance metrics over the 1000 runs.

6.2. Simulation Study Results

In this section, the outcomes of all simulation experiments are evaluated. This is done in two steps. First, the demand for spare parts generated during the two out-of-sample years is compared between the simulation experiments (Section 6.2.1). This gives more insight into the role of the service part failure behaviour and installed base fluctuations on the development of spare parts demand. Afterwards, it is evaluated how accurately the RFM forecasts this demand (Section 6.2.2). Finally, the differences between reliability-based and time series-based forecasting are discussed (Section 6.2.3).

6.2.1. Simulated Spare Parts Demand

This section evaluates how service part failure behaviour and installed base fluctuations are related to the spare parts demand generated during the out-of-sample period. Table 6.3 presents the average monthly demand levels throughout the out-of-sample period for all parameter configurations from Table 6.2, based on 1000 independent simulation runs.

Table 6.3: Out-of-sample monthly spare parts demand for all parameter configurations

		Monthly Demand Initial Phase			Monthly Demand Mature Phase			Monthly Demand End-of-Life Phase		
β	η	Year 9	Year 10	Overall	Year 9	Year 10	Overall	Year 9	Year 10	Overall
0.9	**100**	2.604	2.923	**2.763**	0.603	0.595	**0.589**	1.511	1.331	**1.421**
	500	0.504	0.574	**0.539**	0.118	0.120	**0.119**	0.319	0.265	**0.292**
	1000	0.261	0.305	**0.283**	0.059	0.060	**0.060**	0.163	0.140	**0.152**
1.1	**100**	5.432	6.224	**5.828**	1.424	1.401	**1.413**	3.169	2.891	**3.030**
	500	1.123	1.258	**1.191**	0.301	0.306	**0.303**	0.704	0.648	**0.676**
	1000	0.545	0.620	**0.582**	0.150	0.151	**0.151**	0.374	0.328	**0.351**
1.5	**100**	17.726	19.936	**18.846**	3.634	3.717	**3.676**	9.015	8.251	**8.633**
	500	4.849	5.552	**5.201**	1.516	1.579	**1.547**	3.185	2.831	**3.008**
	1000	2.291	2.612	**2.452**	0.842	0.920	**0.881**	1.624	1.462	**1.543**

The results in Table 6.3 indicate how the nine Weibull parameter configurations affect the average monthly demand for spare parts during both out-of-sample years. In general, the average monthly demand decreases when the scale parameter increases, independent of the shape parameter and installed base fluctuations, which is in line with the derivation of equation 6.1. Moreover, the shape parameter is also positively related to the average monthly demand, independent of the scale parameter and installed base fluctuations. Service parts tend to fail earlier when the failure rate increases at a higher pace over time.

Furthermore, Table 6.3 indicates that the installed base fluctuations in the initial and end-of-life phase are directly translated into an increasing and decreasing spare part demand pattern respectively. In the mature phase, some other interesting demand behaviour can be recognized. In this case, the simulation experiments with shape parameters $\beta = 0.9$ and $\beta = 1.1$ deal with fairly stable demand during the out-of-sample period. However, the situations with $\beta = 1.5$ point at a small increasing demand pattern in Table 6.3. This finding shows that failure behaviour can also directly affect spare parts demand growth. Shape parameters $\beta = 0.9$ and $\beta = 1.1$ have a failure rate very close to being constant, while parameter $\beta = 1.5$ results in an increasing failure rate over time. Therefore, the increase in demand for $\beta = 1.5$ could be caused by a gradual increase in the average installed base age. The rate at which the spare parts demand changes does depend on the failure behaviour of a service part. The combination of installed base fluctuations and service part failure behaviour replicates spare parts demand patterns that are representative of the three different lifecycle phases. The next section analysis how accurately the RFM can predict these demand patterns.

6.2.2. Forecast Performance Simulation

The out-of-sample forecast performance of the RFM is tested for three Weibull parameter estimation approaches: the lifetime-driven, demand-driven and the perfect fit approach (Section 6.1.3.2). In addition, the Enriched EWMA, DES and SBA are included as benchmarks (Section 6.1.3.3). Table 6.4 indicates the forecast error, in terms of MASE, for the simulation experiments with Weibull scale parameter $\eta = 500$. Furthermore, Table 6.5 presents the forecast bias, in terms of SME, for the same experiments. The outcomes of the remaining experiments with $\eta = 100$ and $\eta = 1000$ are listed in Appendix E.3. The experiments in Tables 6.4 and 6.5 vary between the Weibull shape parameters and the installed base lifecycle phases.

Table 6.4: Out-of-sample simulation experiment forecast error with parameter η = 500

Lifecycle Phase	Parameters			RFM (MASE)			Benchmarks (MASE)		
	γ	ρ	β	Lifetime-Driven	Demand-Driven	Perfect Fit	Enriched EWMA	DES	SBA
Initial	4	2	0.9	0.972	0.971	0.913	1.026	1.087	0.978
			1.1	0.827	0.848	0.779	0.867	0.919	0.856
			1.5	0.740	0.768	0.714	0.811	0.880	0.772
Mature	2	2	0.9	1.121	1.088	1.030	1.082	1.123	1.027
			1.1	0.927	0.911	0.845	0.898	0.976	0.843
			1.5	0.817	0.842	0.783	0.843	0.934	0.790
End-of-Life	2	4	0.9	0.977	1.013	0.948	1.088	1.178	1.021
			1.1	0.829	0.866	0.799	0.911	0.967	0.870
			1.5	0.754	0.797	0.717	0.833	0.873	0.805

Table 6.5: Out-of-sample simulation experiment forecast bias with parameter η = 500

Lifecycle Phase	Parameters			RFM (SME)			Benchmarks (SME)		
	γ	ρ	β	Lifetime-Driven	Demand-Driven	Perfect Fit	Enriched EWMA	DES	SBA
Initial	**4**	**2**	**0.9**	0.159	0.153	0.095	0.193	0.163	0.295
			1.1	0.127	0.148	0.082	0.188	0.155	0.221
			1.5	0.142	0.161	0.094	0.206	0.177	0.309
Mature	**2**	**2**	**0.9**	0.191	0.162	0.112	0.154	0.111	0.151
			1.1	0.183	0.159	0.114	0.158	0.101	0.175
			1.5	0.141	0.152	0.099	0.153	0.104	0.162
End-of-Life	**2**	**4**	**0.9**	-0.121	-0.144	-0.078	-0.175	-0.151	-0.243
			1.1	-0.117	-0.137	-0.071	-0.176	-0.149	-0.222
			1.5	-0.107	-0.122	-0.062	-0.166	-0.132	-0.256

The results in Tables 6.4 and 6.5 allow for a forecast performance comparison between the different Weibull estimation approaches of the RFM. As expected, the RFM with a perfect Weibull parameter fit outperforms the lifetime-driven and demand-driven approaches for every scenario. In most scenarios, the lifetime-driven approach comes closest to the forecast performance of this perfect fit approach. The results indicate how much forecast accuracy is lost when an estimation procedure is required to obtain the Weibull parameters. However, the lifetime-driven approach does not beat the demand-driven approach during every scenario. The results for all scenarios with $\eta = 500$ indicate that the demand-driven approach outperforms the lifetime-driven approach during the three scenarios with the lowest average monthly demand (Table 6.3). This trend could also be observed for the scenarios with $\eta = 100$ and $\eta = 1000$ (Appendix E.3). In general, a lower average monthly demand directly translates to fewer uncensored observations during the in-sample period. In the end, the results show that that the demand-driven approach is affected less by the data availability. The decreasing relationship between data availability and the performance of the lifetime-driven approach can be explained by the role of the MLE. The accuracy of the MLE depends on the number of uncensored observations (Section 5.4.2.2). In the end, the demand-driven approach proves to be more robust. However, there is one important remark to the simulation experiments. This study separated the three lifecycle phases to test the relationship between data availability and forecast performance. In practice, the mature and end-of-life phase generally follow after a phase during which already uncensored observations are gathered. In this case, it is expected that the forecast accuracy of the lifetime-driven approach improves for the low demand level scenarios of both phases.

Moreover, the simulation outcomes in Tables 6.4 and 6.5 also help to recognize under which conditions the RFM can provide the most value relative to the benchmarks. The RFM manages to outperform the time series-based techniques during the experiments that deal with spare parts demand changes from the initial and end-of-life phase. This result is independent of the Weibull parameter estimation approach. The initial and end-of-life phase generally create spare demand patterns that increase and decrease over time respectively (Table 6.3). However, when the demand levels are more stable over time, like during the mature phase, the RFM struggles to provide additional value. The scenarios in the mature phase with $\beta = 0.9$ and $\beta = 1.1$ show that the RFM cannot beat the best performing benchmarks while using the perfect fit approach. Nevertheless, the RFM with the lifetime-driven and demand-driven approach still performances closely to the Enriched EWMA during the mature phase.

Furthermore, in most scenarios, DES outperforms the SBA in terms of forecast bias (SME) and the other way around for the forecast error (MASE). The Enriched EWMA generally scores in between both methods for both forecast bias and error. These findings are in line with the empirical benchmark evaluation in Chapter 4. In addition, all conclusions can be generalized across the remaining 18 simulation experiments in which the Weibull scale parameter is changed to $\eta = 100$ and $\eta = 1000$ (Appendix E.3). In the end, the main findings are in line with the conclusions of Dekker et al. (2013) and Van der Auweraer & Boute (2019). They stated that forecasting with installed base information has the most potential when the size of the installed base changes over time.

6.2.3. Difference Reliability-based and Time Series-based Forecasting

The last step is to elaborate on the simulation results by giving a closer look at the differences between reliability-based and time series-based forecasting. A single simulation run is used as an example to visualize how the RFM tries to capture the demand for spare parts. Figure 6.3 visualizes the demand generated during the out-of-sample years while using parameter setting $\{\beta, \eta, \Gamma, \gamma, \rho\} = \{1.1, 100, 100, 4, 2\}$. Moreover, the figure depicts how the RFM, DES and SBA try to estimate the spare parts demand during both years. In this case, the RFM uses the perfect fit approach.

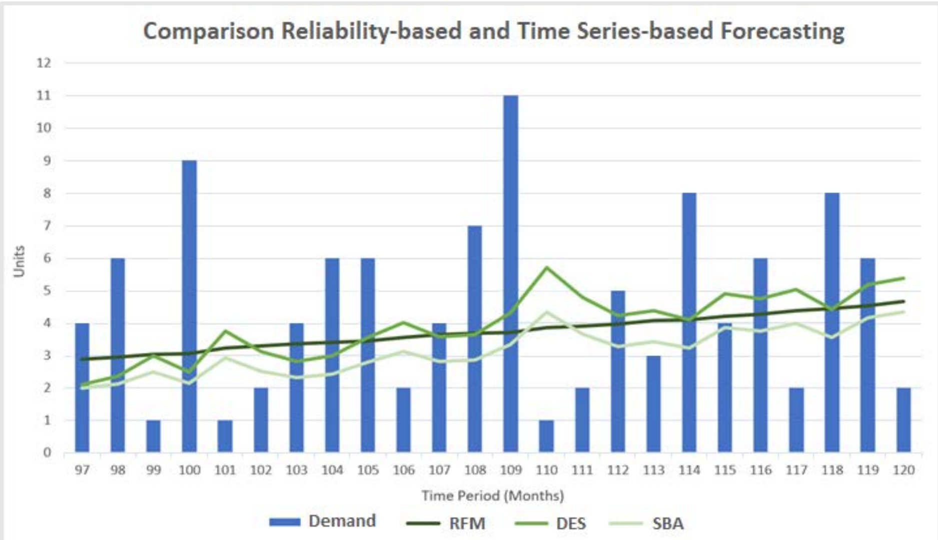

Figure 6.3: Example of a single out-of-sample simulation run

Figure 6.3 shows that DES and the SBA directly adapt their forecast for next month by including the most recent observation. In this way, both methods try to anticipate any change in demand in the short term. In Figure 6.3, months 99 till 105 give a clear example of the DES and SBA lagging behind an increase in demand, which is caused by anticipating short-term demand variability. Moreover, Figure 6.3 indicates that the RFM is less reactive when demand occurs. More specifically, the RFM only experiences a shift in age distribution, because failed service parts are replaced by new ones. In this way, the RFM manages to capture the overall increase in spare parts demand in the long term, without lagging behind any short-term demand variability. In conclusion, this visualization explains why reliability-based forecasting can outperform time series-based forecasting when encountering changing demand patterns.

7 Empirical Study

In this chapter, an empirical study is set up to evaluate the forecast performance of the RFM for service parts at ASML. The simulation study in Chapter 6 already indicated to what extent the RFM can outperform benchmark methods when using Weibull parameter estimation procedures. This chapter assesses how the RFM holds up in practice, with the lifetime-driven and demand-driven Weibull parameter estimation procedure. First, the design of this empirical study is explained in Section 7.1. Afterwards, all results are presented and evaluated in Section 7.2.

7.1. Empirical Study Design

The design of the empirical study consists of several important aspects, which are discussed in this section. First of all, Section 7.1.1 considers the initialisation of the RFM based on the service part registrations and spare part demand data available. Afterwards, Section 7.1.2 derives a representable subset of service parts for which the RFM is initialised and evaluated. In the end, Section 7.1.3 explains how the forecast performance of the RFM is tested for each of these service parts.

7.1.1. Reliability-based Forecasting Model Initialisation

In this section, it is clarified how this empirical study initialises the RFM to derive spare parts demand forecasts. The objective of this empirical study is to evaluate the forecast performance of the RFM during the last two years for which demand data is available (May 2019 – April 2021). All service parts registrations and spare parts demand that occurred before this out-of-sample period is used for the initialisation of the RFM. Chapter 5 introduced the two most important elements of the RFM: An overview of the installed service parts over time and the failure behaviour of service parts. Sections 7.1.1.1 and 7.1.1.2 discuss the initialisation of both elements for this empirical study respectively.

7.1.1.1. Installed Service Part Overview Initialisation

The RFM uses parameter $n_{a,t}^{k}$ to describe the number of active service parts with age a per month t, when considering service part k. The service part registrations at ASML denote which service parts were active during a given month since the first-ever introduction of this service part k. Their age in this given month is equal to the time that has passed since their built-in registration. In this way, an overview of the service part age distribution is created for every month in which service part k was active. Built-in and/or built-out registrations between two consecutive forecasts give an incentive to add and/or remove service parts from this installed base overview respectively. Service parts without any change only become older.

7.1.1.2. Service Part Failure Behaviour Initialisation

The RFM requires an estimation of the failure behaviour for a service part k. In specific, this is needed to determine the conditional failure probability $p_{a,i}^{k}$ for the active service parts with age a during the forecast horizon $[t, t+i]$. Sections 5.4.2 and 5.4.3 introduced the lifetime-driven and demand-driven Weibull parameter estimation procedure respectively. This section briefly explains the initialisation for both procedures, which shows many similarities with the simulation study set up in Section 6.1.3.2.

The lifetime-driven approach uses the MLE to estimate the Weibull parameters based on uncensored and right-censored lifetime observations, as explained in Section 5.4.2. These lifetime observations are based on all built-in and built-out registrations of a service part k from its first-ever introduction to the start of the out-of-sample period. Therefore, the length of the period over which registrations can be found differs among the different service parts k. This period is referred to as the in-sample period. All service parts with a built-in and built-out registration during this period are included as uncensored observations. Moreover, the service parts built in during this in-sample period but not built out before the end of this period, are included as right-censored observations. The length of the right-censored lifetime observations is equal to the difference between the built-in timestamp and the end of the in-sample period.

The demand-driven approach estimates the Weibull parameters based on the actual demand generated by installed service parts, as derived in Section 5.4.3. Thus, this procedure uses the overview of the installed service parts from Section 7.1.1.1 in combination with the monthly demand they produced since the first-ever introduction of the service part. For every service part k the last two years are used as an out-of-sample period (May 2019 – April 2021). The years between the first-ever introduction of service part k and the start of the out-of-sample period define the in-sample period. The Weibull parameters are approximated by running the RFM during this in-sample period while minimizing the average forecast error. This forecast error is defined by the MASE. The Weibull parameters estimated by both procedures are fixed during the entire out-of-sample period.

7.1.2. Service Part Subset

The empirical study is carried for a subset of service parts at ASML. A tool has been developed at ASML to extract the required built-in and built-out data for a service part. However, this tool is not able to process large numbers of service parts. Therefore, this section describes how a representative subset of service parts has been created by applying multiple criteria to the total set of service parts in scope.

First of all, the subset only contains service parts with a strictly increasing demand pattern during the out-of-sample period. Section 4.3.2 concluded that the methods currently available struggled the most with this category of service parts. In addition, the simulation study in Chapter 6 pointed out that the RFM provides the most value when capturing changing demand patterns. The subset of service parts with a strictly increasing demand pattern is further classified to ensure there are enough failure observations available. Nelson (2003) found that the MLE can produce reliable Weibull parameter estimations when there are at least 15 uncensored observations available (Section 5.4.2.2). This applies to 89% of the set of service parts with a strictly increasing demand pattern. The service parts in the subset of this empirical study are drawn from this group.

In order to end up with a representative subset, the remaining group of service parts is categorized by the overall age classifier. Section 4.4.1 introduced this classification criterium as the time that has passed since the first-ever registration of a service part. Moreover, Section 4.4.2 found that this criterium can distinct significant differences in terms of forecast error and bias. In Section 4.4.2, the entire scope of service parts has been classified into equal-sized subsets which score lowest, middle and highest on overall age respectively. The remaining set of service parts from which the subset is drawn deals with a ratio of 41%, 29% and 30% for service parts with a low, middle and high overall age respectively. Therefore, the subset of this empirical study is equally distributed with randomly drawn service parts classified as low, middle and high overall age respectively. This representative subset also allows for an analysis of the effect of the service part history length on the forecast performance. In total, this study extracted service part registrations for 3% of all service parts in scope, which is equivalent to 18% of all service parts with a strictly increasing demand pattern and at least 15 uncensored observations. Appendix A.2 presents the descriptive statistics of this subset.

7.1.3. Benchmark Methods & Performance Metrics

The evaluation of the forecast performance of both configurations of the RFM requires benchmark methods and performance metrics. This section briefly discusses which benchmark methods and performance metrics are included in this empirical study. First, this study uses the same benchmark methods as the simulation study in Chapter 6: The Enriched EWMA, DES and SBA. All three methods have already been tested on the same 2-year out-of-sample set in Chapter 4. In this case, DES and the SBA use individually optimized smoothing parameters. Section 4.5.3 pointed out that this individual smoothing parameter optimization on average outperforms a single optimal parameter configuration, which has been optimized across all service parts in scope. Thus, all service parts in the subset use their optimal smoothing parameters found by the empirical optimization in Chapter 4. Moreover, the forecast performance during the out-of-sample period is assessed by the SME and MASE. The scaling of both methods allows for a more equitable comparison between different categorized groups of service parts. The demand during the last three years of the in-sample period of every service part (May 2016 – April 2019) is used to initialise the MASE, which is in line with the analysis in Chapter 4.

7.2. Empirical Study Results

Both RFM configurations and the three benchmark methods have been used to forecast the demand during the out-of-sample period. This section presents and evaluates the forecast performance of these five options throughout this period. First, Sections 7.2.1 and 7.2.2 assess the forecast performance in terms of forecast error and bias respectively. Afterwards, Section 7.2.3 draws conclusions about the overall forecast performance of the RFM in practice. In the end, the general failure behaviour of the service parts in this empirical study is reflected in Section 7.2.4.

7.2.1. Forecast Error

The forecast error of all five options under consideration is evaluated by the MASE. Figure 7.1 visualizes how the lifetime-driven RFM, demand-driven RFM, Enriched EWMA, DES and SBA perform on forecast error for the subsets which score lowest, middle and highest on overall age. The average age of these three groups is 6.56, 12.58 and 20.47 years respectively.

Figure 7.1: Overall age classification with the out-of-sample forecast error of the empirical study

In Figure 7.1, a clear difference in performance can be observed between both RFMs with a lifetime-driven and demand-driven Weibull parameter estimation procedure. The demand-driven approach outperforms the lifetime-driven approach for all three categories. The lifetime-driven RFM struggles the most for the service part with only a limited history. However, when the overall age of the service part increases, the forecast error decreases and approaches the forecast error of the demand-driven RFM more closely. The demand-driven RFM does not only outperform its lifetime-driven equivalent but also manages to score better in terms of forecast error than the Enriched EWMA, DES and SBA. The SBA has been considered the best time series-based alternative when it came to forecast error during earlier analyses in Chapter 4, which is confirmed by the results in Figure 7.1. In this case, the difference in forecast error between the demand-driven RFM and the SBA slightly increases when the overall age decreases. Nevertheless, all five methods show some decrease in forecast error when the overall age increases. This result is in line with the empirical classification analysis on the overall age classifier in Section 4.4.2.

7.2.2. Forecast Bias

The SME is used to evaluate the forecast bias of all five methods. In this case, Figure 7.2 depicts the effect of the overall age on the forecast bias. First of all, the lifetime-driven RFM is clearly affected by the overall age, as the bias significantly decreases as the overall age increases. When the overall age is low, a significant share of the lifetime observations is censored. According to Starling et al. (2021), the MLE struggles to accurately describe the failure behaviour in this case. Nevertheless, the demand-driven approach is able to more accurately capture this behaviour, independent of the overall age. This results in a forecast methodology that is better at keeping up with the trend in comparison with the Enriched EWMA, DES and SBA as well.

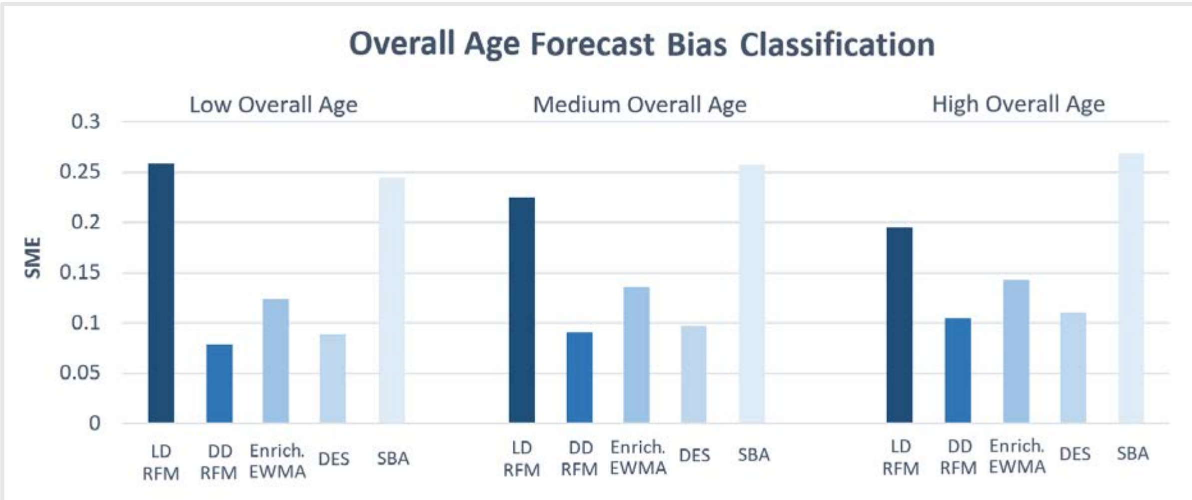

Figure 7.2: Overall age classification with the out-of-sample forecast bias of the empirical study

7.2.3. Overall Forecast Performance

The combination of forecast bias and error gives a representable overview of the overall forecast performance when considering spare parts demand (Teunter & Duncan, 2009). The analysis on service parts with a strictly increasing demand pattern in Chapter 4 pointed out that no method is currently able to score accurately both in terms of forecast bias and error. Time series-based benchmarks like DES and the SBA performed best on forecast bias and error respectively. However, in both cases, this was at the expense of their forecast performance counterpart. This empirical study showed that the demand-driven RFM can combine excellence in terms of forecast bias and error by outperforming both DES and the SBA. Thus, in practice, the demand-driven RFM can capture a significant share of the potential value attributed to this method by the simulation study in Chapter 6. Most importantly, the RFM manages to deliver significant forecast performance improvements for ASML in comparison with its current forecasting process. On average, the RFM reduces the average forecast error of the subset of service parts by 15.8%, whereas the forecast bias is reduced by 31.9%.

7.2.4. Service Part Failure Behaviour

The RFM produced the most accurate forecasts when the Weibull parameters are estimated by a demand-driven estimation procedure, as derived in Section 5.4.3. The distribution of the Weibull shape parameters found by this procedure is visualized in Figure 7.3.

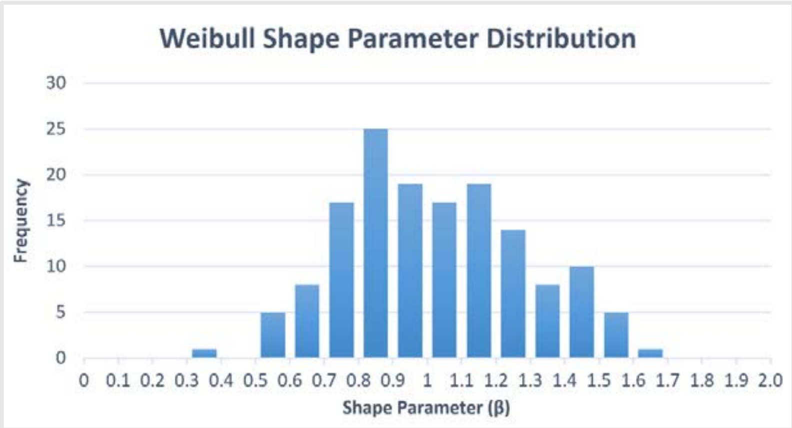

Figure 7.3: Weibull shape parameter distribution with the demand-driven RFM

Figure 7.3 contains relevant information about the failure behaviour of the service parts used in this empirical study. These service parts are known for their increasing demand pattern. The distribution in Figure 7.3 is centered around one and slightly skewed to the right. Moreover, this distribution validates the choice for the three Weibull parameters in the simulation study of Chapter 6 ($\beta \in \{0.9, 1.1, 1.5\}$) Section 5.4.1.2 explained how the shape parameters determine whether a service part deals with an increasing ($\beta > 1$) or decreasing ($\beta < 1$) failure rate when time passes. The shape parameter distribution indicates that none of both is dominating. Information about the failure behaviour of service parts can enhance decision-making at ASML beyond the ranges of spare parts planning. Service parts with an increasing failure rate function could for example be eligible for a preventive maintenance policy. This means those service parts are preventively replaced because the probability of failure increases with time. Moreover, the failure behaviour of service parts could be of interest to the Development & Engineering department at ASML, which is responsible for service part developments and improvements.

Finally, the finding of a shape parameter distribution centered around one is interesting for the spare part planning processes at ASML. These processes currently assume that the demand rates follow independent Poisson processes (Section 3.1). This assumption can be justified either when the failure behaviour of service parts is exponentially distributed or when the installed base is sufficiently large while the service part lifetimes are generally distributed (Van Houtum & Kranenburg, 2015). This Poisson process assumption generally holds when planning on a higher aggregation level, because this leads to a sufficiently large installed base. However, on a local level, this applies to a lesser extent. Nevertheless, Figure 7.3 shows that the failure behaviour of most service parts with an increasing demand pattern is (close-to) exponentially distributed ($\beta = 1$). In this case, it is still reasonable to assume a Poisson failure process when planning on lower aggregation levels.

8 Model Extensions

The RFM has been built on a combination of theoretical assumptions. In general, the fewer assumptions needed, the closer this forecasting model comes to describing reality. The objective of this chapter is to conceptually derive solutions to relax some of the assumptions behind the RFM. First, Section 8.1 conceptualizes a proportional hazard model which can include the effect of service part features, like part version and machine type, on the failure behaviour. Moreover, Section 8.2 derives a way to include spare parts demand in the RFM from service parts that are installed during the forecasting horizon.

8.1. Proportional Hazard Model

The RFM currently assumes that the failure behaviour of a service part is independent of any part feature or environmental factor. However, in practice, this assumption does not easily hold up. In reality, relevant features like applicable machine type, machine usage rate and part version can cause differences in failure behaviour for a service part. In this section, a proportional hazard model is proposed to extend the RFM model from Chapter 5.

In general, a proportional hazard model (PHM) is a statistical technique to explore and describe the relationship between service part failure behaviour and relevant explanatory variables. The PHM is based on the proportional hazard assumption. This assumption states that every individual version of a service part has the same failure rate, which is combined with a unique covariate-based scalar. In this case, the covariates are assumed to be time-independent. This means that the ratio between the failure rate functions of two individual service parts is constant over time. Thus, the covariates are multiplicatively related to the failure rate. Furthermore, the PHM assumes that all observations are independent of each other. Section 5.2.1 already validated this assumption for the RFM by stating that service parts at ASML are generally installed in different machines. The general form of the PHM in equation 8.1 describes the failure rate as a relation between a baseline failure rate $\lambda_0(t)$ and an exponential function with covariates Z_j. The effect of a covariate Z_j on this baseline failure rate is expressed by the coefficient α_j. In this way, a combination of v covariates can influence the baseline failure rate.

$$\lambda(t) = \lambda_0(t) \cdot e^{\left(\sum_{j=1}^{v} \alpha_j Z_j\right)} \tag{8.1}$$

A significant advantage of the PHM is that the coefficients α_j can be estimated by regression independently of the baseline failure rate function λ_0 (Cox, 1972). The Cox proportional hazard model is the most commonly used PHM for reliability data. The baseline failure function of the Cox PHM is a non-parametric piecewise constant failure rate function. Therefore, the Cox PHM reduces to a Kaplan-Meier estimate when no significant covariates can be found. Another advantage of the PHM is its flexibility in taking different forms. The PHM allows the baseline failure rate function to be replaced by a parametric failure rate. Moreover, the exponential covariate term is based on parametric assumptions and can take both different linear and non-linear forms. In this research, the Weibull distribution is used to describe the time-to-failure of service parts (Section 5.4.1). Equation 8.2 presents the Weibull-based PHM, which uses the Weibull failure rate function from equation 5.6.

$$\lambda(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{\left(\sum_{j=1}^{v}\alpha_j z_j\right)} \tag{8.2}$$

Equation 8.3 explains how a failure rate function can be translated into a reliability function when applying the Weibull-based PHM (Arts, 2017).

$$R(t) = 1 - F_T(t) = e^{-\int_0^t \lambda(x)dx} = e^{-\int_0^t \left(\frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} e^{\left(\sum_{j=1}^{v}\alpha_j z_j\right)}\right)dx} \tag{8.3}$$

The failure rate function $\lambda(x)$ can be simplified by replacing the covariation term with parameter y:

$$\lambda(t) = \frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} e^{\left(\sum_{j=1}^{v}\alpha_j z_j\right)} = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \cdot y = \frac{\beta}{\eta}\left(\frac{t}{\eta} \cdot y^{\left(\frac{1}{\beta-1}\right)}\right)^{\beta-1} \tag{8.4}$$

Next, the expression $y^{\left(\frac{1}{(\beta-1)}\right)}$ is substituted by parameter d:

$$\lambda(t) = \frac{\beta}{\eta}\left(\frac{td}{\eta}\right)^{\beta-1} \tag{8.5}$$

This expression for the failure rate can be used to derive the reliability function, following the same line of reasoning as for equation 8.3. Subsequently, solving the integral and further simplifying the expression gives equation 8.6.

$$R(t) = e^{-\int_0^t \lambda(t)dt} = e^{-\int_0^t \left(\frac{\beta}{\eta}\left(\frac{xd}{\eta}\right)^{\beta-1}\right)dx} = e^{-\frac{1}{d}\left(\frac{td}{\eta}\right)^{\beta}} = e^{\left(-\frac{td}{\eta d^{(1/\beta)}}\right)^{\beta}} \tag{8.6}$$

The last step is to directly compare this expression with the general reliability function of the Weibull distribution, given by equation 5.5. In this way, the effect of the covariates on the probability parameters of the Weibull distribution can be described by equations 8.7 and 8.8.

$$\beta_{cov} = \beta \tag{8.7}$$

$$\eta_{cov} = \frac{1}{\eta d^{(1/\beta)}} = \eta y^{-(1/\beta)} = \eta e^{\left(\sum_{j=1}^{v}\alpha_j z_j\right)^{-(1/\beta)}} \tag{8.8}$$

In the end, the covariates do influence scale parameter η, while shape parameter β remains unchanged (Ghodrati & Kumar, 2005). This result is in line with the proportional hazard assumption. Moreover, this derivation allows the Weibull-based PHM to compute the conditional failure probabilities in the RFM with equation 5.7. In this way, it is possible to include any relevant covariate into the RFM. The Weibull parameters of the baseline failure rate can be determined independently of these covariates. Therefore, the parameters can still be estimated by the Weibull parameter estimation procedures from Chapter 5. Only the scale parameter has to be updated afterwards when all covariate coefficients α_j are found by regression. The regression coefficients can be derived with the censored lifetime observations, which are also used to derive the Weibull baseline failure rate function by the lifetime-driven RFM. Covariates should only be included in the RFM when they have a significant effect on the failure behaviour of a service part. Otherwise, this extension only increases the model complexity and data requirements, without improving the forecast accuracy in the end.

8.2. Additional Install Failures

The simulation study in Chapter 6 indicated that the RFM with the perfect fit approach still has to deal with a small positive bias when confronted with spare parts demand from the initial and mature phase.

The current RFM only considers the N_t^k service parts that are installed at the moment of forecasting t. However, in practice, service parts can be installed and fail during the same forecast horizon. This is generally the case with service parts that deal with infant mortality. The failure behaviour of these service parts is generally described by a decreasing failure rate over time. In this case, including new install failures in the RFM could make a difference. Therefore, this section explores the options to include the expected failures from newly installed service parts in a certain forecast horizon.

In general, there can be multiple reasons for the install of a new service part during a forecast horizon. The most prominent cause is an unexpected failure of a service part. However, the install of a new service part could also be planned when a service part receives an upgrade or a new machine is installed. Especially the last reason is relevant because the applicable installed base of a service part at ASML continuously grows over time (Figure 4.2). In most cases, the timing of event-based installs is known beforehand. This information can be used to directly estimate the additional demand coming from new service parts with infant mortality. The potential point in time at which a service part is built into a new machine is described by $t + j$, for which applies: $i > j$. When installed at the time $t + j$, this leaves a period of length $i - j$ during which the service part can still fail in the given forecast horizon. Subsequently, parameter m_{t+j}^k counts the number of service parts k that are installed in a new machine at the time $t + j$. Probability $p_{0,i-j}^k$ then describes the chance that a new service part, installed in a new machine at the time $t + j$, will fail during the remainder of the forecast horizon with length $i - j$. Demand coming from newly installed service parts is also Bernoulli distributed, whereas the total expected demand follows a Poisson Binomial distribution again. Equation 8.9 denotes how this additional expected demand can be added to the basis RFM from equation 5.3.

$$\hat{x}_{t,t+i}^k = \sum_{a=1}^{A_t^k} n_{a,t}^k p_{a,i}^k + \sum_{j=1}^{i-1} m_{t+j}^k p_{0,i-j}^k \tag{8.9}$$

In the case of an unexpected failure during the forecast horizon, the exact timing is of course not known beforehand. The simulation study in Chapter 6 assumed that every broken service part is replaced within one month. This assumption is in line with current practices, which means that service part replacements become relevant when the forecast horizon length increases. The unknown timing of these failures makes it impossible to include those expected failures with an exact term. Nevertheless, this number of expected failures can be approximated when using the forecast of the initial RFM model (equation 5.3). This forecast describes the expected number of failures during the forecast horizon. Parameter $l_{t,t+i}^k$ then denotes the number of new service parts k installed between time t and $t + i$ caused by a failure in this time horizon. The number of infant failures coming from this set of service parts can be approximated when assuming every service part was replaced halfway through the forecast horizon. Equation 8.10 presents how this approximation can expand equation 8.9. Another option is to assume that the failures occur equally distributed during the forecast horizon.

$$\hat{x}_{t,t+i}^k = \sum_{a=1}^{A_t^k} n_{a,t}^k p_{a,i}^k + \sum_{j=1}^{i-1} m_{t+j}^k p_{0,i-j}^k + l_{t,t+i}^k p_{0,i/2}^k \tag{8.10}$$

Finally, this section focused on the inclusion of additional service part failures. However, it is also possible that service parts are discarded during the forecast horizon. In this case, the RFM slightly over-forecast the expected spare parts demand. Following the line of reasoning for equation 8.9, the remaining time during which a discarded service part can still fail during the forecast horizon could be included. However, it is debatable whether this failure will trigger demand shortly before discarding.

9 Conclusions

Up to this point, four out of five phases of the regulative cycle by Van Strien (1997) have been completed. This chapter discusses the final phase of this cycle, which includes an evaluation of all results from the previous phases. In this way, conclusions and managerial insights can be derived. First, this chapter denotes the conclusions and answers the main research question in Section 9.1. Next, Section 9.2 translates the outcomes into recommendations for ASML. Moreover, Section 9.3 discusses to what extent this research filled gaps in the literature on spare parts demand forecasting. Finally, Section 9.4 examines the limitations of this study and derives directions for future research.

9.1. Conclusions

This research has been conducted within the CSCM department at ASML. Based on the potential of installed base information to improve the accuracy of spare parts demand forecasting, the following main research question has been derived in the problem definition phase:

How can the forecast accuracy in the current framework for forecasting and planning of spare parts demand at ASML be improved by using installed base information?

The diagnosis and analysis phase indicated that demand forecasts play a vital role in the planning of spare part basestock levels at warehouses across the customer service supply chain of ASML. The development of spare parts demand in this supply chain is strongly related to changes in the installed base. In general, spare parts demand follows the demand for installed systems, but with a delay (Dekker et al., 2013). Therefore, the constant growth of the installed base at ASML results in more service parts with an increasing demand pattern over time. The current forecasting process attempts to incorporate installed base size changes in its forecasts but still lags behind the trend. The same applies to the time series-based forecasting benchmarks currently available in the literature, which all struggle to accurately capture an increasing demand pattern.

A reliability-based forecasting model (RFM) has been derived to capture the delay and more accurately predict changing demand patterns. The RFM converts installed base information, like service part registrations, into spare parts demand forecasts by generalizing the failure behaviour of service parts with a Weibull distribution. In this way, any change in the installed base is directly reflected by a forecast of the RFM. A comparative simulation study pointed out that the RFM provides the most value in comparison with time series-based benchmarks when confronted with changing demand patterns. In this case, the RFM can outperform the current forecasting process and the time series-based benchmarks, independent of the Weibull parameter estimation procedure. In general, the RFM manages to capture the overall change in spare parts demand in the long-term, without lagging behind short-term demand variability. The comparative empirical study for service parts at ASML validated the potential of the RFM in practice. However, in this case, only the demand-driven RFM outperformed the current forecasting process and the time series-based benchmarks both in terms of forecast bias and error.

In conclusion, the RFM provides an answer to the main research question. This model can improve the forecast accuracy of spare parts demand at ASML by the use of installed base information. This conclusion applies most to service parts that deal with a changing demand pattern. Installed base growth at ASML in the coming years only increases the potential value of this model for the entire scope of service parts. Moreover, this research also showed how traditional time series-based forecasting techniques can still realize significant forecast performance improvements for the large set of service parts with a more stable demand pattern, in comparison with the current practices.

9.2. Recommendations

Given all research outcomes, a number of recommendations can be provided regarding the processes related to service parts at ASML:

- **Implement the RFM for service parts with an increasing demand pattern.**
  This research found that an RFM can significantly improve the forecast performance for service parts with an increasing demand pattern in comparison with the current practices. On average, the RFM reduces the average forecast error of the subset of service parts by 15.8%, whereas the forecast bias is reduced by 31.9%. Therefore, it is recommended to implement the RFM for this group of service parts. An important requirement for the application of the RFM is the availability of accurate service part registrations. This research has shown how valuable accurate service part registrations can be.

- **Improve the forecast performance of service parts with a stable demand pattern.**
  In the current scope of service parts at ASML, a significant share has to deal with stable demand. However, this research showed that the RFM is not able to provide significant forecast performance improvements for these service parts. Nevertheless, improvements can still be realized for service parts with stable demand by adjusting the parameters of the current forecasting methodology. The simple moving average method indicated that an equal division of weights, instead of an exponential division, could already significantly improve the forecast performance. In this way, the forecast error and bias for the current scope of service parts with stable demand on average can be decreased by at least 77.9% and 12.1% respectively.

- **Implement multiple complementary performance metrics to assess the overall forecast performance.**
  In the current forecasting process, the overall forecast performance is only analysed by tracking signals. This metric provides insight into the forecast bias but does not necessarily measure the forecast error. Waldström & Segerstedt (2010) state that only using a single measure does not represent all dimensions of the forecast performance when considering intermittent demand. This finding has been validated throughout this research. Therefore, it is recommended to add complementary performance metrics, which reflect both the forecast bias and error. Both SME and MASE were used throughout this research as scaled measures and would be good complementary candidates to fill this vacancy.

- **Use the information derived by the RFM beyond the scope of spare part demand forecasting.**
  The RFM does not only provide forecast accuracy improvements but also creates insights into the failure behavior of service parts. This is valuable information that can be used to decide on, among others, maintenance strategies and service part development. Moreover, insights on the delay between installed base changes and spare part demand can support the tactical and strategic decision-making on multiple levels in the organization.

9.3. Contribution to Scientific Literature

This research adds to the current scientific literature in multiple ways. Chapter 2 provided an overview of the current scientific literature on spare parts demand forecasting. Moreover, Section 2.4 identified several gaps regarding this topic. In this section, it is examined to what extent this research managed to address these gaps in scientific literature.

- **A new forecasting methodology has been conceptualized that captures the potential of installed base information.**
  Dekker et al. (2013) describe the relationship of delay between installed base fluctuations and spare parts demand. The RFM is now able to accurately capture this relationship in a forecast methodology by describing the service part failure behaviour with a Weibull distribution. Furthermore, considering the installed base on a service part level is an important step forward in comparison with the most recent approaches in the literature (Van der Auweraer et al., 2019). These approaches examined the installed base on machine level, which means a fixed number of service parts per machine is assumed. This extension is particularly valuable in the semiconductor industry, where almost every machine has a unique service part configuration.

  In addition, Kim, Dekker & Heij (2017) claims that customer interactions and maintenance contracts could be an important source of installed base information. This claim is confirmed by this research, as the installed base information used by the RFM directly results from the close cooperation and service contracts between ASML and its customers.

- **The gap between scientific literature and industrial practices has been addressed.**
  Very scarce empirical research has addressed the combination of spare part management and installed base information, as pointed out in Section 2.4. Recent research on forecasting with installed base information mostly focused on conceptually testing new models, without considering the applicability in practice (Van der Auweraer et al., 2019). This research examined the performance of the RFM both in a simulation and empirical study. The empirical study has been carried out by using actual data from the semiconductor industry. The comparison between both studies indicated how data availability can affect the performance of the RFM.

  An important element of the RFM is the estimation procedure of the Weibull parameters. Currently, the maximum likelihood estimator is the most often used procedure in the scientific literature on spare part failure behaviour modeling (Starling et al., 2021). However, this research showed that the forecast accuracy of the RFM in combination with this procedure is very sensitive to data availability. Therefore, a more robust Weibull parameters estimation procedure has been derived. In contrast to the traditional procedure, the more robust approach is able to outperform the benchmarks in both the simulation and empirical study. In the end, the introduction of this new approach also helps to address the gap between scientific literature and industrial practices.

In the end, this research makes some important steps forward in the scientific literature on spare parts demand forecasting. However, there are still plenty of opportunities to build on this work and extract even more value from installed base information. The next section discusses the limitations of this research and provides directions for future research.

9.4. Limitations & Future Research

In this section, the limitations of this research are discussed in more detail. Moreover, based on those limitations, directions for future research are proposed.

- **The main challenge for the implementation of an RFM is the collection of accurate data.**
  This research indicated that the potential of the RFM is affected by the availability and quality of the installed base information. Wagner & Lindemann (2008) denote that most engineering companies only have a 'cloudy view' of their installed base. This finding harms the generalizability of the RFM. Nevertheless, this research does give an incentive to keep track of the complete composition of the installed base. Spare parts management should become an inter-organizational process, where collaboration and the distribution of data play an important role.

- **The RFM assumes that every equivalent service part has an identical failure behaviour.**
  This assumption implies that the failure behaviour of equivalent service parts is independent of the operational environment. However, in practice, this failure behaviour could be dependent on how and under which circumstances a service part is used. The same applies to the effect of different service part versions on the failure behaviour. Section 8.1 already conceptually derived a regression-based solution direction, called PHM, that can include the effects of relevant explanatory variables into the RFM. The effect of the PHM can be the subject of future research.

- **The RFM only considers failures of service parts installed at the moment of forecasting.**
  The current configuration of the RFM does not include the potential failures of service parts that are installed during the forecast horizon. Therefore, the RFM becomes less accurate when forecasting on a longer time horizon. Section 8.2 conceptually derived an extension that includes this additional demand into the RFM. The effect of this extension can be tested by a future study.

- **The performance in this research has been solely based on forecast bias and error measures.**
  In general, a higher forecast performance does not immediately result in a better inventory control performance (Teunter & Duncan, 2009; Kourentzes, 2013). This is the main reason that recent literature on spare parts demand forecasting started testing methods relative to inventory control performance. The relatively complex service network of ASML and the limited time given were the main reasons for sticking to forecast performance measures during this research. Teunter & Duncan (2009) claim that this could still be of great value when bias and error measures are used alongside each other. Nevertheless, a relevant next step would be to quantify the potential of the RFM when applied in a multi-item multi-echelon service network.

- **The research mainly focused on the potential of installed base forecasting in combination with increasing spare parts demand.**
  The simulation study in Chapter 6 briefly pointed out that the RFM could be of great value for situations with decreasing installed base and spare parts demand as well. The RFM directly anticipates when machines and their service parts are discarded, which can prevent over-forecasting and obsolescence costs at ASML. Nevertheless, the empirical study of Chapter 7 only validates the value of installed base forecasting for the group with increasing demand in practice.

- **There are no strict boundaries between the regions of the superior performance for all methods.**
  This research only comes up with broad categories of service parts for which the RFM can make an impact. However, classification criteria could be used to find the exact boundaries of superior performance during a large-scale comparison between the RFM and the time series-based benchmarks. In this case, the role of data availability is particularly interesting.

Bibliography

Andersson, J., & Jonsson, P. (2018). Big data in spare parts supply chains: The potential of using product-in-use data in aftermarket demand planning. *International Journal of Physical Distribution and Materials Management*, *48*(5), 524-544.

Arts, J. (2017). Maintenance modeling and optimization. *BETA Working Paper 526*. Eindhoven University of Technology.

ASML (2021). *Annual Report 2020*. Retrieved from: https://www.asml.com/en/investors/annual-report/2020

Babai, M. Z., Syntetos, A., & Teunter, R. (2014). Intermittent demand forecasting: An empirical study on accuracy and the risk of obsolescence. *International Journal of Production Economics*, *157*, 212-219.

Bacchetti, A., & Saccani, N. (2012). Spare parts classification and demand forecasting for stock control: Investigating the gap between research and practice. *Omega*, *40*(6), 722-737.

Boukhtouta, A., & Jentsch, P. (2018). Support Vector Machine for Demand Forecasting of Canadian Armed Forces Spare Parts. In *2018 6th International Symposium on Computational and Business Intelligence*, 59-64.

Boylan, J. E., & Syntetos, A. A. (2008). Forecasting for inventory management of service parts. In *Complex system maintenance handbook*. Springer, London.

Boylan, J. E., Syntetos, A. A., & Karakostas, G. C. (2008). Classification for forecasting and stock control: a case study. *Journal of the operational research society*, *59*(4), 473-481.

Byrne, R. F. (2012). Beyond Traditional Time-Series: Using Demand Sensing to Improve Forecasts in Volatile Times. *Journal of Business Forecasting*, *31*(2), 13.

Cavalieri, S., Garetti, M., Macchi, M., & Pinto, R. (2008). A decision-making framework for managing maintenance spare parts. *Production planning & control*, *19*(4), 379-396.

Costantino, F., Di Gravio, G., Patriarca, R., & Petrella, L. (2018). Spare parts management for irregular demand items. *Omega*, *81*, 57-66.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, *34*(2), 187-202.

Croston, J. D. (1972). Forecasting and stock control for intermittent demands. *Journal of the Operational Research Society*, *23*(3), 289-303.

Dekker, R., Pinçe, Ç., Zuidwijk, R., & Jalil, M. N. (2013). On the use of installed base information for spare parts logistics: A review of ideas and industry practice. *International Journal of Production Economics*, *143*(2), 536-545.

Delignette-Muller, M. L., & Dutang, C. (2015). fitdistrplus: An R package for fitting distributions. *Journal of Statistical Software 64.4*, 1-34.

Driessen, M., Arts, J., van Houtum, G. J., Rustenburg, J. W., & Huisman, B. (2015). Maintenance spare parts planning and control: a framework for control and agenda for future research. *Production Planning & Control*, *26*(5), 407-426.

Ghodrati, B., & Kumar, U. (2005). Reliability and operating environment-based spare parts estimation approach: A case study in Kiruna Mine, Sweden. *Journal of Quality in Maintenance Engineering*, *11*(2), 169-184.

Gutierrez, R. S., Solis, A. O., & Mukhopadhyay, S. (2008). Lumpy demand forecasting using neural networks. *International journal of production economics*, *111*(2), 409-420.

Hong, Y. (2013). On computing the distribution function for the Poisson binomial distribution. *Computational Statistics & Data Analysis*, *59*, 41-51.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International journal of forecasting*, *22*(4), 679-688.

Johnston, F. R., Boylan, J. E., Shale, E., & Meadows, M. (1999). A robust forecasting system, based on the combination of two simple moving averages. *Journal of the Operational Research Society*, *50*(12), 1199-1204.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, *53*(282), 457-481.

Kim, S., & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, *32*(3), 669-679.

Kim, T. Y., Dekker, R., & Heij, C. (2017). Spare part demand forecasting for consumer goods using installed base information. *Computers and Industrial Engineering*, *103*, 201–215.

Kourentzes, N. (2013). Intermittent demand forecasts with neural networks. *International Journal of Production Economics*, *143*(1), 198-206.

Nahmias, S., & Olsen, T. (2015). *Production and Operations Analysis* (7th ed.). Long Grove, Ill.: Waveland Press, Inc.

Nelson, W. B. (2003). *Applied life data analysis* (Vol. 521). New York: John Wiley & Sons.

Öner, K. B., Franssen, R., Kiesmüller, G., & van Houtum, G. J. (2007). Life cycle costs measurement of complex systems manufactured by an engineer-to-order company. In *The 17th international conference on flexible automation and intelligent manufacturing*, 569-589.

do Rego, J. R., & De Mesquita, M. A. (2015). Demand forecasting and inventory control: A simulation study on automotive spare parts. *International Journal of Production Economics*, *161*, 1-16.

Starling, J. K., Mastrangelo, C., & Choe, Y. (2021). Improving Weibull distribution estimation for generalized Type I censored data using modified SMOTE. *Reliability Engineering & System Safety*, *211*, 107505.

Suijkerbuijk, C., Boon, M., Kapodistria, S., & Mavroeidis, D. (2017). Integration of preventive maintenance and inventory management for healthcare systems at Philips Healthcare.

Syntetos, A. A., & Boylan, J. E. (2005). The accuracy of intermittent demand estimates. *International Journal of forecasting*, *21*(2), 303-314.

Syntetos, A. A., Nikolopoulos, K. & Boylan J. E. (2005). Judging the judges through accuracy implication metrics: The case of inventory forecasting. *International Journal of Forecasting*, *26*(1), 134-143.

Teunter, R. H., & Duncan, L. (2009). Forecasting intermittent demand: a comparative study. *Journal of the Operational Research Society*, *60*(3), 321-329.

Teunter, R. H., Syntetos, A. A., & Babai, M. Z. (2011). Intermittent demand: Linking forecasting to inventory obsolescence. *European Journal of Operational Research*, *214*(3), 606-615.

Trigg, D. W. (1964). Monitoring a forecasting system. *Operation Research Society*, *15*(3), 271-274.

Turrini, L., & Meissner, J. (2019). Spare parts inventory management: New evidence from distribution fitting. *European Journal of Operational Research*, *273*(1), 118-130.

Van der Auweraer, S., & Boute, R. (2019). Forecasting spare part demand using service maintenance information. *International journal of production economics*, *213*, 138-149.

Van der Auweraer, S., Boute, R. N., & Syntetos, A. A. (2019). Forecasting spare part demand with installed base information: A review. *International Journal of Forecasting*, *35*(1), 181-196.

Van Houtum, G. J., & Kranenburg, B. (2015). *Spare parts inventory control under system availability constraints*. New York: Springer.

Van Strien, P. (1997). Towards a methodology of psychological practice. *Theory and Psychology, 7*(5), 683-700.

Van Wingerden, E., Basten, R. J. I., Dekker, R., & Rustenburg, W. D. (2014). More grip on inventory control through improved forecasting: A comparative study at three companies. *International journal of production economics, 157*, 220-237.

Wagner, S. M., & Lindemann, E. (2008). A case study-based analysis of spare parts management in the engineering industry. *Production planning & control*, *19*(4), 397-407.

Walström, P. & Segerstedt, A. (2010). Evaluation of forecasting error measurements and techniques for intermittent demand. *International Journal of Production Economics*, *128*(2), 625-636.

Wang, W., & Syntetos, A. A. (2011). Spare parts demand: Linking forecasting to equipment maintenance. *Transportation Research Part E: Logistics and Transportation Review*, *47*(6), 1194-1209.

Wessels, W. R. (2007). Use of the Weibull versus exponential to model part reliability. In *2007 Annual Reliability and Maintainability Symposium*, 131-135.

Willemain, T. R., Smart, C. N., Shockor, J. H., & DeSautels, P. A. (1994). Forecasting intermittent demand in manufacturing: a comparative evaluation of Croston's method. *International journal of forecasting*, *10*(4), 529-538.

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International journal of forecasting*, *14*(1), 35-62.

Appendices

Appendix A: Detailed Descriptive Statistics

A.1. Descriptive Statistics Full Scope

In this appendix, an overview is presented of the descriptive statistics for all relevant sets of service parts used throughout Chapter 4. This includes the entire set of service parts and the subsets of parts with and without a strictly increasing demand pattern. A distinction between in-sample and out-of-sample is made for each set of parts. The in-sample set consists of 36 months (May 2016 – April 2019), whereas the out-of-sample set contains 24 months (May 2019 – April 2021). First, Tables A.1 and A.2 present the average demand levels for the in-sample and out-of-sample periods respectively.

Table A.1: Average monthly in-sample demand

Service Part Set	Number of Service Parts	Average Monthly Demand			
		Year 1	Year 2	Year 3	Y1 – Y3
Full Scope	100%	0.451	0.623	0.803	**0.626**
Strictly Increasing	19%	0.756	1.035	1.492	**1.094**
Not Strictly Increasing	81%	0.379	0.526	0.642	**0.516**

Table A.2: Average monthly out-of-sample demand

Service Part Set	Number of Service Parts	Average Monthly Demand		
		Year 4	Year 5	Y4 – Y5
Full Scope	100%	0.992	1.221	**1.106**
Strictly Increasing	19%	2.413	3.573	**2.993**
Not Strictly Increasing	81%	0.658	0.669	**0.664**

Moreover, Table A.3 presents more detailed descriptive statistics for the full scope of service parts.

Table A.3: Descriptive statistics of all service parts in scope

		Monthly Demand		Demand Size		Demand Intervals	
		Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
In-Sample	**Minimum**	0.028	0.167	1.000	0.000	1.000	0.000
	25th percentile	0.111	0.319	1.000	0.000	2.692	1.675
	Median	0.222	0.422	1.077	0.272	5.833	4.614
	75th percentile	0.611	0.698	1.533	0.920	11.667	5.508
	Maximum	102.389	67.051	102.389	67.051	17.500	6.364
Out-of-Sample	**Minimum**	0.042	0.204	1.000	0.000	1.000	0.000
	25th percentile	0.125	0.338	1.000	0.000	1.917	1.123
	Median	0.292	0.550	1.200	0.422	4.600	3.209
	75th percentile	0.833	0.924	1.824	1.015	7.667	3.786
	Maximum	205.556	39.871	205.556	39.871	11.500	4.950

Furthermore, Tables A.4 and A.5 present more detailed descriptive statistics for the subsets of service parts with and without a strictly increasing demand pattern respectively.

Table A.4: Descriptive statistics of all service parts with an increasing demand pattern

		Monthly Demand		Demand Size		Demand Intervals	
		Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
In-Sample	Minimum	0.028	0.167	1.000	0.000	1.000	0.000
	25th percentile	0.111	0.319	1.000	0.000	1.346	0.849
	Median	0.444	0.616	1.385	0.650	3.182	2.234
	75th percentile	1.389	1.722	2.182	1.471	7.000	3.162
	Maximum	71.000	24.966	71.000	24.966	17.500	10.607
Out-of-Sample	Minimum	0.083	0.338	1.000	0.000	1.000	0.000
	25th percentile	0.417	0.654	1.222	0.441	1.045	0.213
	Median	1.042	1.398	1.875	0.835	1.438	0.919
	75th percentile	2.542	1.351	3.188	1.515	3.286	1.942
	Maximum	205.556	39.871	205.556	39.871	7.667	2.121

Table A.5: Descriptive statistics of all service parts without an increasing demand pattern

		Monthly Demand		Demand Size		Demand Intervals	
		Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
In-Sample	Minimum	0.028	0.167	1.000	0.000	1.000	0.000
	25th percentile	0.056	0.232	1.000	0.000	2.692	1.874
	Median	0.167	0.380	1.000	0.000	7.000	3.391
	75th percentile	0.444	0.695	1.462	0.660	11.667	4.041
	Maximum	102.389	67.051	102.389	67.051	17.500	9.192
Out-of-Sample	Minimum	0.042	0.204	1.000	0.000	1.000	0.000
	25th percentile	0.083	0.282	1.000	0.000	3.286	2.089
	Median	0.167	0.380	1.167	0.389	4.600	3.104
	75th percentile	0.500	0.655	1.571	0.787	7.667	4.619
	Maximum	99.389	40.003	99.389	40.003	11.500	7.778

A.2. Descriptive Statistics Service Part Subset

The empirical study in Chapter 7 derived a representable subset of service parts with a strictly increasing demand pattern. Table A.6 presents the descriptive statistics of this subset.

Table A.6: Descriptive statistics of the subset of service parts with an increasing demand pattern

		Monthly Demand		Demand Size		Demand Intervals	
		Mean	**St. Dev.**	**Mean**	**St. Dev.**	**Mean**	**St. Dev.**
In-Sample	**Minimum**	0.028	0.167	1.000	0.000	1.000	0.000
	25th percentile	0.139	0.351	1.000	0.000	1.346	0.831
	Median	0.417	0.713	1.295	0.621	3.500	2.394
	75th percentile	1.333	1.657	2.134	1.368	7.000	3.278
	Maximum	65.528	23.674	65.528	23.674	17.500	8.788
Out-of-Sample	**Minimum**	0.083	0.338	1.000	0.000	1.000	0.000
	25th percentile	0.375	0.647	1.250	0.421	1.095	0.221
	Median	1.125	1.297	1.800	0.721	1.353	0.989
	75th percentile	2.417	1.403	2.162	1.463	3.286	1.876
	Maximum	174.917	37.852	174.917	37.852	7.667	2.121

Appendix B: Time Series-Based Forecasting Methods

Forecasting methods based on historical demand have been divided into traditional and Croston-based methods (Table 4.4). Methods from both groups are used as a benchmark throughout this research. This appendix discusses and explains the methods in both groups in more detail. Moreover, the initialisation of each method is clarified.

B.1. Traditional Methods

Simple moving average (SMA) comes up with a forecast by taking the average of the last n data points in the time series. Equation B.1 describes how to compute an SMA forecast for period t, with $t > n$:

$$\hat{x}_{t+i} = \frac{1}{n}\sum_{j=1}^{n} x_{t-n+j} \tag{B.1}$$

Where $\hat{x}_{t+i}$ is the prediction for period $t + i$ and x_t represents the actual demand during period t. This method is mostly used in combination with stationary and slow-moving demand (Nahmias & Olsen, 2015). SMA requires an initialization period of n months to come up with the first forecast. Single exponential smoothing (SES) is a method for stationary time series which uses a weighted average of both the last forecast and the current value of demand (Nahmias & Olsen, 2015). Just like SMA, SES is derived from the assumption that the underlying demand process is stationary. The formula for single exponential smoothing is given by equation B.2.

$$\hat{x}_{t+i} = \alpha x_t + (1 - \alpha)\hat{x}_t \tag{B.2}$$

Where α represents the smoothing parameter, which can take any value between 0 and 1. In general, the closer this value is to zero, the more focus there is on the past values and vice versa. The SES method is initialised by taking the average over the first 12 in-sample months:

$$\hat{x}_{13} = \frac{1}{12}\sum_{t=1}^{12} x_t \tag{B.3}$$

Double exponential smoothing (DES) includes another smoothing equation and smoothing parameter β to account for linear trends. A forecast for period $t + i$ can be computed by equations B.4, B.5 and B.6. In this case, S_t represents the value of intercept at time t, while G_t gives the value of the slope at time t.

$$\hat{x}_{t+i} = S_t + iG_t \tag{B.4}$$

$$S_t = \alpha x_t + (1 - \alpha)(S_{t-1} + G_{t-1}) \tag{B.5}$$

$$G_t = \beta(S_t - S_{t-1}) + (1 - \beta)G_{t-1} \tag{B.6}$$

The values of intercept and slope of the DES method are initialised by using the first 12 in-sample months in the following way:

$$S_{13} = \frac{1}{12}\sum_{t=1}^{12} x_t \tag{B.7}$$

$$G_{13} = \frac{1}{6}\sum_{t=7}^{12} x_t - \frac{1}{6}\sum_{t=1}^{6} x_t \tag{B.8}$$

B.2. Croston-Based Methods

Croston's method (CRO) estimates the average demand size and average inter-arrival time between demands separately by exponential smoothing. Therefore, when no demand occurs during a review interval, the forecast will not change. Initially, Croston's method used a single smoothing parameter for both equations, whereas the current applications use two parameters, α and β, to update the demand size (Z_t) and demand intervals (I_t). Furthermore, parameter q_t indicates the time interval since the last nonzero demand occurred. Equations B.9, B.10 and B.11 summarize Croston's method.

$$\hat{x}_{t+1} = \frac{Z_t}{I_t} \tag{B.9}$$

$$\begin{aligned} if\ x_t > 0 \quad then \quad & Z_t = \alpha x_t + (1-\alpha) Z_{t-1}; \\ & I_t = \beta q_t + (1-\beta) I_{t-1}; \\ & q_{t+1} = 1 \end{aligned} \tag{B.10}$$

$$\begin{aligned} if\ x_t = 0 \quad then \quad & Z_t = Z_{t-1}; \\ & I_t = I_{t-1}; \\ & q_{t+1} = q_t + 1 \end{aligned} \tag{B.11}$$

The Syntetos-Boylan Approximation (SBA) was introduced to approximately correct Croston's method, which turned out to be biased (Syntetos et al., 2005). This correction is given by equation B.12.

$$\hat{x}_{t+1} = \left(1 - \frac{\beta}{2}\right) \frac{Z_t}{I_t} \tag{B.12}$$

Teunter et al. (2011) came up with the Teunter Syntetos Babai Method (TSB), which is able to react more quickly to sudden drops in demand. This method is a derivative of Croston's but updates the demand probability (p_t) instead of the inter-arrival times of demand (I_t).

$$\hat{x}_{t+1} = p_t Z_t \tag{B.13}$$

$$\begin{aligned} if\ x_t > 0 \quad then \quad & Z_t = \alpha x_t + (1-\alpha) Z_{t-1}; \\ & p_t = p_{t-1} + (1-\beta) p_{t-1} \end{aligned} \tag{B.14}$$

$$\begin{aligned} if\ x_t = 0 \quad then \quad & Z_t = Z_{t-1}; \\ & p_t = p_{t-1} + \beta(0 - p_{t-1}) \end{aligned} \tag{B.15}$$

The initialisation of the Croston-based methods requires the introduction of variable τ, which denotes the number of periods with non-zero demand during the first 12 in-sample months.

$$\begin{aligned} if\ \tau > 0 \quad then \quad & Z_{13} = \frac{1}{\tau} \sum\nolimits_{t \in \tau} x_t \\ & I_{13} = \frac{12}{\tau} \\ & p_{13} = \frac{\tau}{12} \end{aligned} \tag{B.16}$$

$$\begin{aligned} if\ \tau = 0 \quad then \quad & Z_{13} = 1 \\ & I_{13} = 12 \\ & p_{13} = 1/12 \end{aligned} \tag{B.17}$$

Appendix C: Empirical Benchmark Performance Results

In this appendix, the extended empirical forecast results from the benchmark evaluation analysis in Section 4.5 are presented. Tables C.1, C.2 and C.3 show the results for the full scope of service parts and the subsets with and without a strictly increasing demand pattern respectively.

Table C.1: Extended benchmark forecast performance with individual part parameter optimization

	SME			MSE			MASE		
Method	**Year 4**	**Year 5**	**Overall**	**Year 4**	**Year 5**	**Overall**	**Year 4**	**Year 5**	**Overall**
SMA	0.118	0.111	**0.114**	4.583	6.433	**5.508**	0.883	0.981	**0.932**
SES	0.209	0.189	**0.198**	6.095	7.194	**6.645**	0.907	1.057	**0.982**
DES	0.089	0.068	**0.078**	5.670	6.181	**5.925**	0.996	1.352	**1.174**
CRO	0.270	0.251	**0.259**	4.525	4.745	**4.635**	0.665	0.780	**0.731**
SBA	0.238	0.216	**0.226**	4.528	4.717	**4.623**	0.651	0.791	**0.721**
TSB	0.191	0.186	**0.188**	4.579	4.774	**4.677**	0.699	0.840	**0.769**
RW	0.019	0.012	**0.015**	4.476	6.870	**5.673**	0.964	1.154	**1.059**
ZF	1.000	1.000	**1.000**	25.681	37.097	**31.389**	0.736	0.924	**0.830**
Enriched EWMA	0.022	0.055	**0.041**	4.337	6.499	**5.418**	0.967	1.111	**1.039**

Table C.2: Extended benchmark forecast performance with strictly increasing demand

	SME			MSE			MASE		
Method	**Year 4**	**Year 5**	**Overall**	**Year 4**	**Year 5**	**Overall**	**Year 4**	**Year 5**	**Overall**
SMA	0.235	0.188	**0.207**	11.974	17.834	**14.904**	0.812	1.262	**1.037**
SES	0.257	0.226	**0.239**	13.890	19.266	**16.578**	0.799	1.308	**1.054**
DES	0.104	0.080	**0.091**	11.421	18.605	**15.013**	0.935	1.365	**1.150**
CRO	0.300	0.272	**0.283**	11.432	15.787	**13.609**	0.687	1.161	**0.924**
SBA	0.288	0.275	**0.280**	11.486	15.832	**13.659**	0.692	1.170	**0.931**
TSB	0.259	0.223	**0.238**	11.431	15.908	**13.669**	0.701	1.202	**0.952**
RW	0.048	0.022	**0.032**	12.748	18.212	**15.480**	0.938	1.512	**1.225**
ZF	1.000	1.000	**1.000**	51.724	125.301	**88.513**	0.973	1.689	**1.331**
Enriched EWMA	0.133	0.154	**0.146**	10.698	17.868	**14.292**	0.835	1.305	**1.070**

Table C.3: Extended benchmark forecast performance without strictly increasing demand

	SME			MSE			MASE		
Method	**Year 4**	**Year 5**	**Overall**	**Year 4**	**Year 5**	**Overall**	**Year 4**	**Year 5**	**Overall**
SMA	0.017	0.013	**0.015**	2.850	3.759	**3.304**	0.899	0.915	**0.907**
SES	0.167	0.143	**0.155**	4.267	4.363	**4.315**	0.933	0.998	**0.965**
DES	0.078	0.052	**0.064**	4.321	3.267	**3.794**	1.010	1.788	**1.399**
CRO	0.245	0.224	**0.235**	2.905	2.156	**2.530**	0.660	0.691	**0.675**
SBA	0.196	0.142	**0.169**	2.896	2.110	**2.504**	0.641	0.702	**0.671**
TSB	0.132	0.139	**0.136**	2.972	2.163	**2.568**	0.698	0.754	**0.726**
RW	-0.006	0.000	**-0.003**	2.536	4.210	**3.373**	0.970	1.070	**1.020**
ZF	1.000	1.000	**1.000**	19.574	16.412	**17.993**	0.681	0.745	**0.713**
Enriched EWMA	-0.070	-0.066	**-0.068**	2.845	3.833	**3.339**	0.998	1.066	**1.032**

Appendix D: Derivation Lifetime-driven Weibull Parameter Estimation Procedure

This appendix presents the derivation of the Maximum Likelihood estimator (MLE). In this case, the MLE considers U uncensored and C censored lifetime observations. The likelihood function for fitting the Weibull parameters with the combination of both types of observations is given by equation D.1 (Nelson, 2003).

$$L(u_1, \dots, u_U;\ c_1, \dots, c_C) = \prod_{i=1}^{U} f_T(u_i) \cdot \prod_{j=1}^{C} \left(1 - F_T(c_j)\right)$$
$$= \prod_{i=1}^{U} \left(\left(\frac{\beta}{\eta}\right) \left(\frac{u_i}{\eta}\right)^{\beta-1} e^{-\left(\frac{u_i}{\eta}\right)^{\beta}} \right) \cdot \prod_{j=1}^{C} \left(e^{-\left(\frac{c_j}{\eta}\right)^{\beta}} \right) \quad \text{(D.1)}$$

The next step is to take the natural logarithm of both sides. In this way, the expression is simplified, because the exponential components are removed, and the products are changed into sums.

$$\ln\left(L(u_1, \dots, u_U;\ c_1, \dots, c_C)\right) = \sum_{i=1}^{U} \left((\beta - 1)\ln(u_i) - \left(\frac{u_i}{\eta}\right)^{\beta} \right) + \ln(\beta) - k\ln(\eta) - \sum_{j=1}^{C} \left(\frac{c_j}{\eta}\right)^{\beta} \quad \text{(D.2)}$$

The objective is to find the optimal (β, η)-combination which maximizes this likelihood function. In order to achieve this, equation D.2 must be differentiated with respect to both β and η. Next, both resulting equations have to be set equal to zero. The solution of the combination of both equations results in the best fitting Weibull parameter configuration. Equations D.3 and D.4 present both derivatives.

$$\frac{\partial \ln\left(L(u_1, \dots, u_U;\ c_1, \dots, c_C)\right)}{\partial \eta} = \left(\frac{\beta}{\eta}\right) \sum_{i=1}^{U} \left(\frac{u_i}{\eta}\right)^{\beta} - U\left(\frac{\beta}{\eta}\right) + \left(\frac{\beta}{\eta}\right) \sum_{j=1}^{C} \left(\frac{c_j}{\eta}\right)^{\beta} \quad \text{(D.3)}$$

$$\frac{\partial \ln\left(L(u_1, \dots, u_U;\ c_1, \dots, c_C)\right)}{\partial \beta} = \sum_{i=1}^{U} \left(\frac{1}{\beta} + \ln\left(\frac{u_i}{\eta}\right) - \left(\frac{u_i}{\eta}\right)^{\beta} \ln\left(\frac{u_i}{\eta}\right) \right) + \sum_{j=1}^{C} \left(-\left(\frac{c_j}{\eta}\right)^{\beta} \ln\left(\frac{c_j}{\eta}\right) \right) \quad \text{(D.4)}$$

Both derivatives are set equal to zero. Equations D.5 and D.6 show the result when trying to solve for η and β respectively.

$$\eta^{\beta} = \frac{1}{U} \left(\sum_{i=1}^{U} \left(u_i^{\beta}\right) + \sum_{j=1}^{C} \left(c_j^{\beta}\right) \right) \quad \text{(D.5)}$$

$$\frac{1}{U} \sum_{i=1}^{U} \ln(u_i) + \frac{1}{\beta} - \ln(\eta) = \left(\frac{1}{\eta^{\beta} U}\right) \left(\sum_{i=1}^{U} u_i^{\beta} \ln(u_i) + \sum_{j=1}^{C} c_j^{\beta} \ln(c_j) - \ln(\beta) \left(\sum_{i=1}^{U} u_i^{\beta} + \sum_{j=1}^{U} c_j^{\beta} \right) \right) \quad \text{(D.6)}$$

Replacing parameter η in Equation D.6 by the estimation in Equation D.5 delivers Equation D.7:

$$\frac{1}{\beta} = \frac{\sum_{i=1}^{U} u_i^{\beta} \ln(u_i) + \sum_{j=1}^{C} c_j^{\beta} \ln(c_j)}{\sum_{i=1}^{U} u_i^{\beta} + \sum_{j=1}^{C} c_j^{\beta}} - \frac{1}{N} \sum_{i=1}^{U} \ln(u_i) \quad \text{(D.7)}$$

Appendix E: Simulation Experiment Results

In this appendix, all the simulation experiment outcomes from the simulation study in Chapter 6 are listed. In total, 27 simulation experiments have been carried out. First, an extended overview of the parameter configurations during each simulation experiment is given in Section E.1. Moreover, Section E.2 presents the average spare parts demand levels during the two out-of-sample years of every simulation experiment. At last, the forecast performance metrics of every single experiment are displayed in Section E.3.

E.1. Extended Overview Simulation Experiment

Table E.1 gives an extended overview of the parameter configuration of every simulation experiment.

Table E.1: Extended overview of all simulation experiments

Simulation Experiment	β	η	Γ^*	γ^*	ρ^*
1	0.9	100	100	4	2
2	0.9	500	100	4	2
3	0.9	1000	100	4	2
4	1.1	100	100	4	2
5	1.1	500	100	4	2
6	1.1	1000	100	4	2
7	1.5	100	100	4	2
8	1.5	500	100	4	2
9	1.5	1000	100	4	2
10	0.9	100	100	2	2
11	0.9	500	100	2	2
12	0.9	1000	100	2	2
13	1.1	100	100	2	2
14	1.1	500	100	2	2
15	1.1	1000	100	2	2
16	1.5	100	100	2	2
17	1.5	500	100	2	2
18	1.5	1000	100	2	2
19	0.9	100	400	2	4
20	0.9	500	400	2	4
21	0.9	1000	400	2	4
22	1.1	100	400	2	4
23	1.1	500	400	2	4
24	1.1	1000	400	2	4
25	1.5	100	400	2	4
26	1.5	500	400	2	4
27	1.5	1000	400	2	4

**Standardized due to confidentiality reasons.*

E.2. Service Parts Demand Levels Simulation Experiments

This section gives an overview of the average spare parts demand levels obtained during the simulation experiments that are listed in Table E.1. In total, the statistics are derived for all 27 simulations experiments. Table E.2 presents the average monthly demand for each of these scenarios during their two out-of-sample years. All these outcomes are based on 1000 independent simulation runs. The standard deviation of the average monthly demand is given in between brackets for each scenario.

Table E.2: Monthly out-of-sample spare parts demand of all simulation experiments

		Monthly Demand Initial Phase			Monthly Demand Mature Phase			Monthly Demand End-of-Life Phase		
β	η	Year 9	Year 10	Overall	Year 9	Year 10	Overall	Year 9	Year 10	Overall
0.9	**100**	2.604 *(0.441)*	2.923 *(0.499)*	**2.763** ***(0.333)***	0.603 *(0.231)*	0.595 *(0.220)*	**0.589** ***(0.158)***	1.511 *(0.365)*	1.331 *(0.317)*	**1.421** ***(0.239)***
	500	0.504 *(0.209)*	0.574 *(0.230)*	**0.539** ***(0.155)***	0.118 *(0.098)*	0.120 *(0.104)*	**0.119** ***(0.072)***	0.319 *(0.175)*	0.265 *(0.146)*	**0.292** ***(0.109)***
	1000	0.261 *(0.144)*	0.305 *(0.161)*	**0.283** ***(0.109)***	0.059 *(0.071)*	0.060 *(0.072)*	**0.060** ***(0.050)***	0.163 *(0.121)*	0.140 *(0.117)*	**0.152** ***(0.079)***
1.1	**100**	5.432 *(0.631)*	6.224 *(0.669)*	**5.828** ***(0.469)***	1.424 *(0.338)*	1.401 *(0.327)*	**1.413** ***(0.234)***	3.169 *(0.515)*	2.891 *(0.473)*	**3.030** ***(0.344)***
	500	1.123 *(0.302)*	1.258 *(0.316)*	**1.191** ***(0.221)***	0.301 *(0.154)*	0.306 *(0.165)*	**0.303** ***(0.112)***	0.704 *(0.243)*	0.648 *(0.214)*	**0.676** ***(0.159)***
	1000	0.545 *(0.207)*	0.620 *(0.237)*	**0.582** ***(0.151)***	0.150 *(0.108)*	0.151 *(0.114)*	**0.151** ***(0.079)***	0.374 *(0.181)*	0.328 *(0.172)*	**0.351** ***(0.113)***
1.5	**100**	17.726 *(1.038)*	19.936 *(1.094)*	**18.846** ***(0.714)***	3.634 *(0.495)*	3.717 *(0.438)*	**3.676** ***(0.311)***	9.015 *(0.774)*	8.251 *(0.673)*	**8.633** ***(0.502)***
	500	4.849 *(0.594)*	5.552 *(0.660)*	**5.201** ***(0.439)***	1.516 *(0.323)*	1.579 *(0.338)*	**1.547** ***(0.221)***	3.185 *(0.489)*	2.831 *(0.439)*	**3.008** ***(0.309)***
	1000	2.191 *(0.401)*	2.612 *(0.464)*	**2.452** ***(0.296)***	0.842 *(0.252)*	0.920 *(0.263)*	**0.881** ***(0.172)***	1.624 *(0.353)*	1.462 *(0.326)*	**1.543** ***(0.339)***

E.3. Forecast Performance Simulation Experiments

In this section, the out-of-sample forecast performance of the Enriched EWMA, DES, SBA and RFM is listed for every simulation experiment. In the case of the RFM, the forecast performance is denoted for the lifetime-driven, demand-driven and perfect fit approach. All tables in this section display the average of the ME and MASE based on 1000 simulation runs. Furthermore, the standard deviations of the performance metrics over these 1000 different runs are given in between brackets. The SME can be obtained by dividing the ME by the average out-of-sample monthly overall demand from the corresponding parameter configuration in Table E.2.

Table E.3: Out-of-sample simulation experiment results of Enriched EMWA

		Initial Phase		Mature Phase		End-of-Life Phase	
β	***η***	**ME**	**MASE**	**ME**	**MASE**	**ME**	**MASE**
0.9	**100**	0.511 *(0.177)*	0.843 *(0.138)*	0.093 *(0.081)*	0.958 *(0.198)*	-0.233 *(0.126)*	0.919 *(0.125)*
	500	0.104 *(0.087)*	1.026 *(0.164)*	0.018 *(0.046)*	1.082 *(0.243)*	-0.051 *(0.067)*	1.088 *(0.178)*
	1000	0.054 *(0.091)*	1.213 *(0.264)*	0.010 *(0.062)*	1.232 *(0.265)*	-0.030 *(0.061)*	1.219 *(0.198)*
1.1	**100**	1.052 *(0.191)*	0.899 *(0.169)*	0.203 *(0.156)*	0.826 *(0.111)*	-0.478 *(0.214)*	0.833 *(0.129)*
	500	0.224 *(0.141)*	0.867 *(0.109)*	0.048 *(0.109)*	0.898 *(0.153)*	-0.119 *(0.116)*	0.911 *(0.197)*
	1000	0.113 *(0.076)*	0.919 *(0.193)*	0.025 *(0.079)*	1.214 *(0.213)*	-0.061 *(0.113)*	1.196 *(0.095)*
1.5	**100**	4.431 *(0.309)*	0.943 *(0.214)*	0.594 *(0.181)*	0.846 *(0.143)*	-1.802 *(0.356)*	0.849 *(0.143)*
	500	1.071 *(0.232)*	0.811 *(0.147)*	0.237 *(0.153)*	0.843 *(0.134)*	-0.499 *(0.164)*	0.833 *(0.099)*
	1000	0.462 *(0.186)*	0.833 *(0.124)*	0.130 *(0.087)*	0.853 *(0.165)*	-0.273 *(0.121)*	0.821 *(0.153)*

Table E.4: Out-of-sample simulation experiment results of DES

		Initial Phase		Mature Phase		End-of-Life Phase	
β	***η***	**ME**	**MASE**	**ME**	**MASE**	**ME**	**MASE**
0.9	**100**	0.409 *(0.149)*	0.865 *(0.159)*	0.062 *(0.071)*	1.060 *(0.224)*	-0.187 *(0.116)*	0.996 *(0.172)*
	500	0.088 *(0.095)*	1.087 *(0.189)*	0.013 *(0.078)*	1.123 *(0.249)*	-0.044 *(0.071)*	1.178 *(0.245)*
	1000	0.043 *(0.087)*	1.336 *(0.293)*	0.006 *(0.051)*	1.291 *(0.253)*	-0.026 *(0.064)*	1.389 *(0.243)*
1.1	**100**	0.814 *(0.174)*	0.957 *(0.159)*	0.152 *(0.138)*	0.920 *(0.195)*	-0.403 *(0.184)*	0.865 *(0.160)*
	500	0.185 *(0.074)*	0.919 *(0.172)*	0.031 *(0.123)*	0.976 *(0.210)*	-0.101 *(0.183)*	0.967 *(0.216)*
	1000	0.089 *(0.064)*	0.973 *(0.207)*	0.018 *(0.081)*	1.319 *(0.236)*	-0.048 *(0.103)*	1.119 *(0.336)*
1.5	**100**	3.575 *(0.243)*	0.988 *(0.229)*	0.476 *(0.160)*	0.927 *(0.214)*	-1.496 *(0.317)*	0.928 *(0.239)*
	500	0.921 *(0.197)*	0.880 *(0.157)*	0.161 *(0.149)*	0.934 *(0.209)*	-0.397 *(0.201)*	0.873 *(0.165)*
	1000	0.378 *(0.154)*	0.839 *(0.129)*	0.091 *(0.082)*	0.948 *(0.250)*	-0.218 *(0.103)*	0.872 *(0.172)*

Table E.5: Out-of-sample simulation experiment results of SBA

		Initial Phase		Mature Phase		End-of-Life Phase	
β	η	ME	MASE	ME	MASE	ME	MASE
0.9	100	0.799 *(0.188)*	0.827 *(0.121)*	0.095 *(0.079)*	0.938 *(0.217)*	-0.342 *(0.181)*	0.893 *(0.112)*
	500	0.159 *(0.119)*	0.978 *(0.179)*	0.018 *(0.091)*	1.027 *(0.225)*	-0.071 *(0.099)*	1.021 *(0.171)*
	1000	0.080 *(0.101)*	1.154 *(0.245)*	0.014 *(0.065)*	1.182 *(0.243)*	-0.049 *(0.068)*	1.143 *(0.203)*
1.1	100	1.613 *(0.244)*	0.826 *(0.152)*	0.200 *(0.152)*	0.756 *(0.102)*	-0.713 *(0.279)*	0.829 *(0.123)*
	500	0.263 *(0.162)*	0.856 *(0.115)*	0.053 *(0.112)*	0.843 *(0.146)*	-0.150 *(0.136)*	0.870 *(0.166)*
	1000	0.161 *(0.101)*	0.883 *(0.191)*	0.029 *(0.094)*	1.147 *(0.198)*	-0.091 *(0.138)*	1.087 *(0.298)*
1.5	100	6.803 *(0.398)*	0.906 *(0.194)*	0.601 *(0.192)*	0.781 *(0.125)*	-2.914 *(0.393)*	0.838 *(0.114)*
	500	1.607 *(0.314)*	0.772 *(0.151)*	0.251 *(0.162)*	0.790 *(0.108)*	-0.770 *(0.141)*	0.805 *(0.111)*
	1000	0.644 *(0.219)*	0.832 *(0.133)*	0.128 *(0.096)*	0.803 *(0.129)*	-0.399 *(0.143)*	0.803 *(0.112)*

Table E.6: Out-of-sample simulation experiment results of RFM with the lifetime-driven approach

		Initial Phase		Mature Phase		End-of-Life Phase	
β	η	ME	MASE	ME	MASE	ME	MASE
0.9	100	0.323 *(0.138)*	0.902 *(0.106)*	0.099 *(0.083)*	0.993 *(0.212)*	-0.150 *(0.105)*	0.861 *(0.119)*
	500	0.086 *(0.105)*	0.972 *(0.184)*	0.023 *(0.069)*	1.121 *(0.234)*	-0.035 *(0.087)*	0.977 *(0.164)*
	1000	0.047 *(0.085)*	1.198 *(0.254)*	0.024 *(0.093)*	1.311 *(0.264)*	-0.043 *(0.073)*	1.181 *(0.211)*
1.1	100	0.651 *(0.154)*	0.835 *(0.121)*	0.195 *(0.148)*	0.806 *(0.111)*	-0.345 *(0.179)*	0.793 *(0.113)*
	500	0.151 *(0.067)*	0.827 *(0.115)*	0.055 *(0.097)*	0.922 *(0.154)*	-0.079 *(0.154)*	0.829 *(0.159)*
	1000	0.071 *(0.059)*	0.918 *(0.179)*	0.037 *(0.097)*	1.229 *(0.213)*	-0.054 *(0.121)*	1.086 *(0.252)*
1.5	100	2.872 *(0.219)*	0.889 *(0.175)*	0.554 *(0.165)*	0.823 *(0.121)*	-1.131 *(0.253)*	0.791 *(0.109)*
	500	0.739 *(0.198)*	0.825 *(0.156)*	0.218 *(0.129)*	0.817 *(0.119)*	-0.322 *(0.091)*	0.754 *(0.116)*
	1000	0.307 *(0.136)*	0.862 *(0.137)*	0.139 *(0.101)*	0.836 *(0.194)*	-0.165 *(0.082)*	0.768 *(0.102)*

Table E.7: Out-of-sample simulation experiment results of RFM with the demand-driven approach

		Initial Phase		Mature Phase		End-of-Life Phase	
β	η	ME	MASE	ME	MASE	ME	MASE
0.9	100	0.367	0.808	0.094	0.987	-0.174	0.887
		(0.161)	*(0.111)*	*(0.078)*	*(0.236)*	*(0.107)*	*(0.102)*
	500	0.082	0.971	0.019	1.088	-0.042	1.013
		(0.101)	*(0.198)*	*(0.072)*	*(0.245)*	*(0.092)*	*(0.163)*
	1000	0.038	1.139	0.015	1.282	-0.023	1.142
		(0.073)	*(0.231)*	*(0.079)*	*(0.253)*	*(0.062)*	*(0.188)*
1.1	100	0.767	0.814	0.199	0.831	-0.378	0.822
		(0.168)	*(0.149)*	*(0.159)*	*(0.104)*	*(0.177)*	*(0.115)*
	500	0.176	0.911	0.048	0.911	-0.093	0.866
		(0.079)	*(0.119)*	*(0.076)*	*(0.162)*	*(0.170)*	*(0.153)*
	1000	0.081	0.873	0.028	1.203	-0.045	1.069
		(0.060)	*(0.188)*	*(0.084)*	*(0.194)*	*(0.109)*	*(0.276)*
1.5	100	3.605	0.892	0.575	0.849	-1.346	0.815
		(0.214)	*(0.186)*	*(0.164)*	*(0.133)*	*(0.278)*	*(0.103)*
	500	0.837	0.768	0.235	0.842	-0.367	0.797
		(0.214)	*(0.173)*	*(0.141)*	*(0.120)*	*(0.112)*	*(0.099)*
	1000	0.348	0.823	0.126	0.839	-0.193	0.799
		(0.138)	*(0.139)*	*(0.096)*	*(0.139)*	*(0.091)*	*(0.114)*

Table E.8: Out-of-sample simulation experiment results of RFM with the perfect fit approach

		Initial Phase		Mature Phase		End-of-Life Phase	
β	η	ME	MASE	ME	MASE	ME	MASE
0.9	100	0.255	0.771	0.065	0.921	-0.092	0.831
		(0.135)	*(0.098)*	*(0.076)*	*(0.204)*	*(0.099)*	*(0.108)*
	500	0.051	0.913	0.013	1.030	-0.023	0.948
		(0.084)	*(0.165)*	*(0.064)*	*(0.223)*	*(0.069)*	*(0.165)*
	1000	0.023	1.131	0.009	1.202	-0.014	1.101
		(0.069)	*(0.221)*	*(0.065)*	*(0.248)*	*(0.051)*	*(0.198)*
1.1	100	0.437	0.806	0.157	0.753	-0.205	0.769
		(0.128)	*(0.101)*	*(0.127)*	*(0.099)*	*(0.158)*	*(0.099)*
	500	0.098	0.779	0.035	0.845	-0.048	0.799
		(0.064)	*(0.109)*	*(0.089)*	*(0.157)*	*(0.158)*	*(0.157)*
	1000	0.048	0.852	0.020	1.154	-0.027	1.045
		(0.051)	*(0.160)*	*(0.077)*	*(0.198)*	*(0.078)*	*(0.262)*
1.5	100	1.902	0.862	0.445	0.778	-0.713	0.762
		(0.147)	*(0.171)*	*(0.151)*	*(0.111)*	*(0.203)*	*(0.111)*
	500	0.489	0.714	0.153	0.783	-0.186	0.717
		(0.183)	*(0.143)*	*(0.119)*	*(0.108)*	*(0.081)*	*(0.108)*
	1000	0.209	0.791	0.082	0.789	-0.114	0.734
		(0.113)	*(0.118)*	*(0.093)*	*(0.185)*	*(0.059)*	*(0.102)*

Appendix F: Research Framework Spare Parts Management

Driessen et al. (2015) derived a research framework for the planning and control of spare parts in organizations that maintain high-value capital assets. Figure F.1 visualizes this framework, which is used to define the scope of this research in Section 1.4.

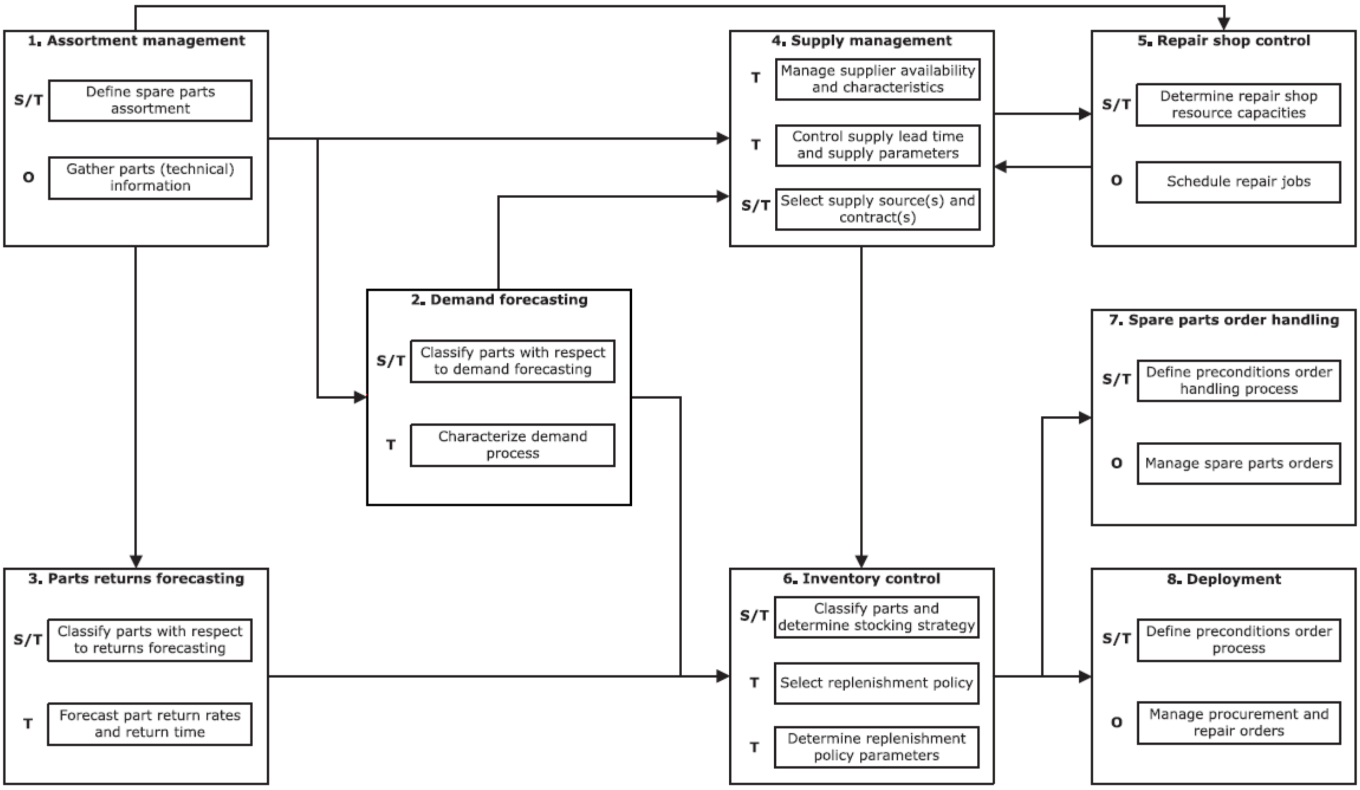

Figure F.1: Research framework for the planning and control of spare parts