THESIS FOR THE DEGREE OF LICENTIATE OF ENGINEERING

Enhancing aftermarket demand planning with product-in-use data

JOAKIM ANDERSSON

Department of Technology Management and Economics

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2019

Enhancing aftermarket demand planning with product-in-use data

JOAKIM ANDERSSON

© JOAKIM ANDERSSON, 2019

Licentiate thesis
ISSN 1654-9732
Technical report L2019:115

Department of Technology Management and Economics
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Printed by Chalmers Digital Print
Gothenburg, Sweden 2019

Enhancing aftermarket demand planning with product-in-use data

Joakim Andersson
Department of Technology Management and Economics
Chalmers University of Technology

ABSTRACT

Aftermarket demand planning, consisting of forecast and known demand, is a critical activity for both the uptime of customers' products and the supply chain-related costs. Traditional aftermarket forecasting methods use historical demand as the only input to the statistically based forecasts, usually combined with judgmental modifications. Since the underlying mechanisms of customer demand are the need for maintenance or repair, these methods fail to deliver good forecast accuracy in various contexts, e.g. phase-in/out of spare parts or for spare parts with intermittent demand patterns

However, rapid increase in access to new types of data created by on-board sensors, improved algorithms for predicting future events with massive amount of data and increased Information and Communications Technology (ICT) capabilities are generating new and potential improved demand planning methods for aftermarket services.

The purpose of this thesis is to investigate how product-in-use data can be used in, and improve the performance of, demand planning processes for aftermarket services, by describing and explaining related performance effects and challenges.

The conceptual framework developed consists of three components (potential demand planning methods, challenges for implementation and development of these methods and confirmation of the results). Each component is attached to a research question which is dealt with in three separate studies. The two first studies are single-case studies which address the potential improved demand planning methods and challenges in aftermarket supply chain planning processes. The third study combined qualitative and quantitative methods, of which the latter was a regression method with exogenous variables (ARX).

Results show how product-in-use data is best utilised in planning spare parts with different attributes, e.g. different life-cycle phases and demand frequencies. Eight potentially relevant interventions, i.e. proposed methods, using product-in-use data in the demand planning process are identified. Challenges are explored in relation to the process complexity and data complexity of aftermarket supply chains, and underlying reasons and interdependencies are identified. Forecast accuracy is proven to be better for phase-in spare parts with medium/high demand frequency in a quantitative study.

The practical contributions of the thesis are insights how to enhance the aftermarket demand planning by new causal-based methods using product-in-use data, as well as awareness of the complexity and challenges of development and implementation of such data-driven processes. The thesis contributes to theory by proposing how product-in-use driven demand planning methods can be used and for which types of context the methods can create value. The process focus contributes to how to apply and develop methods for aftermarket causal-based demand planning

Key words: Demand planning, Forecasting, Aftermarket, Big data, Digitalisation, Product-in-use.

Acknowledgements

After a very long break from the academic world, I got the opportunity to become an Industrial PhD student, thanks to my employer, Volvo GTO, Service Market Logistics. It has been almost three very exciting, interesting and challenging years since then. So, thank you Johan, my current manager and Thomas, my previous manager, that gave me this great opportunity and for all support and collaboration.

To Patrik and Arni, my excellent supervisors and co-writers, thank you for your great support, advices, discussions and that you always find time for me even when your calendars are full.

I´m also very grateful for all great discussion and collaborations with colleagues, both with people from the advanced analytics department, as well as others.

Many thanks also to my roommates Lars-Erik and Klas for all good advices and interesting discussions. Thanks also to all other colleagues from Supply and Operations Management.

Last but not least, I would like to express my gratitude to my wife, Daniella, for all your support and understanding, this wouldn´t been possible without you. And to my kids, Samuel and Agnes, thank you for your curios and interesting questions.

List of appended papers

Paper I:

Andersson, J. and 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 & Logistics Management,* Vol. 48, No. 5, pp. 524-544.

Paper II:

Andersson, J., Halldorsson, A and Jonsson, P. (2019). "Improving supply chain planning processes for aftermarket services: Challenges of using product-in-use data".

An earlier version of this paper was published in the Proceedings of the 30thNOFOMA Conference, June 14-15, 2018, Kolding.

Paper III:

Andersson, J. (2019). "Causal based spare parts forecasting exploiting product-in-use data in a heavy vehicle aftermarket context".

An earlier version of this paper was published in the Proceedings of of the 26th Eoruma Conference, June 17-19, 2019, Helsinki.

PAPER	FIRST AUTHOR	SECOND AUTHORS	RESPONSIBILITIES
I	Joakim Andersson	Patrik Jonsson	The first author collected all data and wrote the larger part of the paper initially. The paper was improved jointly. Initial analysis was done by the first author and developed jointly.
II	Joakim Andersson	Patrik Jonsson and Arni Halldorsson	The first author collected all the data and made initial analysis and wrote a first draft of the paper. The paper was developed jointly regarding analysis and writing.
III	Joakim Andersson		Single author.

Table of Contents

1 Introduction

The aim of this thesis is to establish knowledge regarding utilization of the data created by connected vehicles for aftermarket demand planning. The term product-in-use data has been chosen for this data, stemming from the finished product, transmitted by telecom networks or by connected wired devices (Hussain et al., 2012). In an aftermarket context this data can be divided into operational data (e.g. mileage and operating hours), sensor data (e.g. vibrations, temperature, voltage and pressure), and sensor data processed by embedded Electronic Control Units (ECUs), referred to below as fault codes. Additional data, such as repair history, external data (for example weather, road conditions, macro-economic) are also considered to be within the scope of this research, together with information traditionally used in aftermarket supply chain planning, such as demand history, lead times, product cost etc. The overall research objective is to study current practices and to explore and suggest new and enhanced principles and methods regarding aftermarket demand planning using product-in-use data, as well as evaluating these opportunities for the various product classes, expressed for example by life cycle stages and demand frequency. This chapter concerns the background of the research problem, followed by the problem definition and discussion, from which the research questions and scope are derived. Finally, the chapter ends with an outline of the contents of the thesis.

1.1 Background

Aftermarket supply chain planning encompasses the activities required to ensure the uptime of finished products as well as striving to balance the costs of operational activities (e.g. ordering, transportation, picking etc.), inventory and costs related to back orders, e.g. expediting costs and lost sales (Vollman et al., 1991).

The main arguments for a firm to focus on an efficient supply chain planning are threefold, 1) supporting customers' requirement for up-time has a great impact on revenue as well as on customer satisfaction (Cohen et al., 2006), 2) strengthening firms' competitiveness which contributes to customer loyalty, boosting sales of both finished products and aftermarket products (McAlexander et al., 2002), and 3) increasing the ability of firms to handle the complexity of the aftermarket supply chain activities at a competitive cost level, for example (Cohen et al., 2006). Hence, the overall challenge to remain competitive in the long term is to fulfill customer requirements (up-time, quality, operating cost, services etc.) at an acceptable cost for the customer and, at the same time, maintain the firm's own costs in pursuing the necessary supply chain activities at an optimal level. A prerequisite for efficient aftermarket supply chain planning is the demand planning process. From an aftermarket perspective, the demand planning process lays the foundation for the remaining planning processes (e.g.) inventory planning and distribution requirements planning), i.e. without a high-performing demand planning process, the performance of the whole supply chain planning process will be poor. Hence, poor demand planning will have negative effects on service levels for end-customers, as well as drive higher capital and operational costs. Traditional aftermarket demand planning and forecasting has been using customer demand history as the only source for this process (Bacchetti and Saccani, 2012), which do not consider the underlying demand (Romeijnders et al., 2012; Dekker et al., 2013). Analytical advancements, increased number of smart connected vehicles and machines create new opportunities for aftermarket demand planning (Frowein et al., 2014).

1.1.1 Revenue and customer relations

The service provided by the aftermarket supply chain supports the customers during the life-cycle of the finished products, which is critical in ensuring customer loyalty, but also from both a competitive point of view, as well as from a profitability perspective (Cohen et al., 2006, Xiao and Yang, 2008). Profit margins generated by the aftermarket are often higher than those achieved by new product sales. The aftermarket may generate at least three times the turnover of the original purchase during a given product lifecycle (McAlexander et al., 2002; Wise and Baumgartner, 1999). An example supporting this claim comes from Cavalieri et al. (2007), who argue that the aftermarket may generate three times more revenue than sales of finished products. Furthermore, original equipment manufacturers (OEMs) should enhance the value of the finished products where value-adding aftermarket services need to be provided to customers through the product lifecycle to gain a long-term relationship and create loyalty (Cavalieri et al., 2007). So, from income and customer relations perspectives, aftermarket services are extremely important for firms manufacturing durable goods, which also makes new demands on the existing business models for manufacturing companies, i.e. a transformation from product-centric business models to more service-centric business models is required. One prerequisite to maintain an efficient aftermarket supply chain operation is to perform satisfactory demand planning. Poor demand planning drives both excessive inventory and poor availability, which, in turn, drives both increased cost and poor customer satisfaction (Kilger and Wagner, 2008).

1.1.2 Competition from other OEMs and new competitors

Due to rapid technology development, information and communication technology (ICT) and market globalization, competition in the aftermarket is fierce. Today's marketplace is competitive and dynamic. In the unpredictable environment that characterizes the aftermarket, the main competing factors are price and service. A competitive price is needed to attract customers, and service capabilities are essential to meet the customers' requirements of up-time in terms of delivery times and availability (Xiao and Yang 2008). In order to remain competitive, the supply chain needs to be managed in a cost-efficient fashion. One important contribution in this respect is to reduce demand uncertainty, another is to improve forecast accuracy. According to Porter and Heppelman (2014), this digital revolution also creates opportunities for new competitors in the market. An example is the entry of Amazon into the aftermarket business (Abbema 2018).

1.1.3 Aftermarket supply chain complexity

For several reasons, the aftermarket supply chain processes are considered to be complex. This thesis has a supply chain planning perspective, hence the complexity is looked at from the OEM point of view.

Firstly, one driver of aftermarket supply chain complexity is that many actors are involved in the aftermarket supply chain planning processes and are able to disrupt the planning (Gebauer et al., 2013). These actors are, for example, suppliers, haulers, third-party logistic providers, OEM managed distribution centers and workshops (which may be OEM-owned or independent).

Secondly, according to Cohen et al. (2006), for example, the unpredictable nature of demand also contributes to complexity. Since the underlying reason for the demand is a need for service or repair, the demand is difficult to predict. Demand fluctuates depending on life cycle, launch dates, varying usage and operational environments of the products, marketing activities and

competitors' behavior. Moreover, spare parts vary greatly in price, criticality and specificity (Huiskonen, 2001), which adds to the complexity.

A third factor explaining the complexity of aftermarket supply chain planning is the supply chain structure, with a large number of suppliers and spare parts delivery points (Cohen et al., 2006, Holmqvist and Pessi, 2006). The supply chain structure could also be a cause of demand variation, creating a bullwhip effect (Lee et al., 1997).

The fourth factor driving the complexity is cost-related, driven by very high cost of down-time for end-users and high cost related to backorders (Bachetti and Saccani, 2012). The difficult problem of balancing the cost of the activities in the supply chain against customers' requirements for high availability is thus a key challenge.

The fifth factor regarding complexity is the fierce competition in aftermarket services, mainly regarding price and service levels, which requires high availability of spare parts, leading to a high number of stocked items and therefore a high inventory cost. On the other hand, if availability is too low, with too many stock-outs, backorder costs in terms of emergency transport and lost sales of spare parts will increase, and in the long term will reduce the goodwill of the company, leading to a risk of decreased sales of new products as well.

An essential task in order to address these complexities is to make sure that the demand planning process is efficient and effective (Boylan and Syntetos, 2010). A summary of the characteristics leading to the previously described problems for an aftermarket supply chain, as well as a corresponding comparison with the characteristics of a manufacturing supply chain, is presented in Table 1.1 below.

Table 1.1: Manufacturing and spare parts supply chain characteristics

	OEM supply chain	Aftermarket supply chain
Demand driver	Customer need	Needed for maintenance/repair
Demand pattern	Stable, high-frequency	Intermittent and, in many cases, low frequent
Number of SKUs	Scant amount	10-100 times more
Type of products	Complete units/Sub-assemblies	Components broken down to lower level
Supply chain design	Convergent	Divergent
SCP methods	MRP/JIT	DRP/ROP
Inventory turnover	High	Low

1.1.4 Enabling data for improved spare parts demand planning

With emerging technology in modern equipment and use of the data produced by software and sensors in machines or vehicles (Internet of Things (IoT) or smart connected products), there is great potential in monitoring and collecting more causal data from connected vehicles in use (Porter and Heppelmann, 2014). Examples of these product-in-use data are operational data (e.g. mileage operating hours), sensor data (e.g. oil pressure, temperature) and fault codes (e.g. malfunction of certain components). A related concept that deals with what to do with the data, e.g. analysis, predictions and modeling, is known as big data. Enhancements in recent years regarding storage and computing of large amount of data, combined with advanced analytical

methods (e.g. machine learning and AI) and advances in information and telecommunication technology (ICT) have great potential to enable improved demand planning methods (Arribas-Bel, 2014).

1.2 Motivation for the research

The three areas mentioned above: revenue share from the aftermarket, competition in the aftermarket, and supply chain planning complexity provide the motivation for the research toward a licentiate, together with the enabling factors: data from smart connected products combined with advanced analytical methods. All in all, the usage of data deriving from connected vehicles has the potential to enable more efficient demand planning methods (Manyika et al., 2011). The shortcomings of traditional planning methods, combined with new opportunities with regard to retrieving and analyzing product-in-use data, strengthen competitiveness in the aftermarket and create new business opportunities, e.g. new business models, pay-as-you-drive contracts, lease contracts, capacity rental, insurance based on usage, and opportunities for the OEM to improve predictions of future demand, and, consequently enhance the planning of the aftermarket supply chain (Cohen et al., 2006). Figure 1.1 summarizes the drivers and enablers that motivate the research.

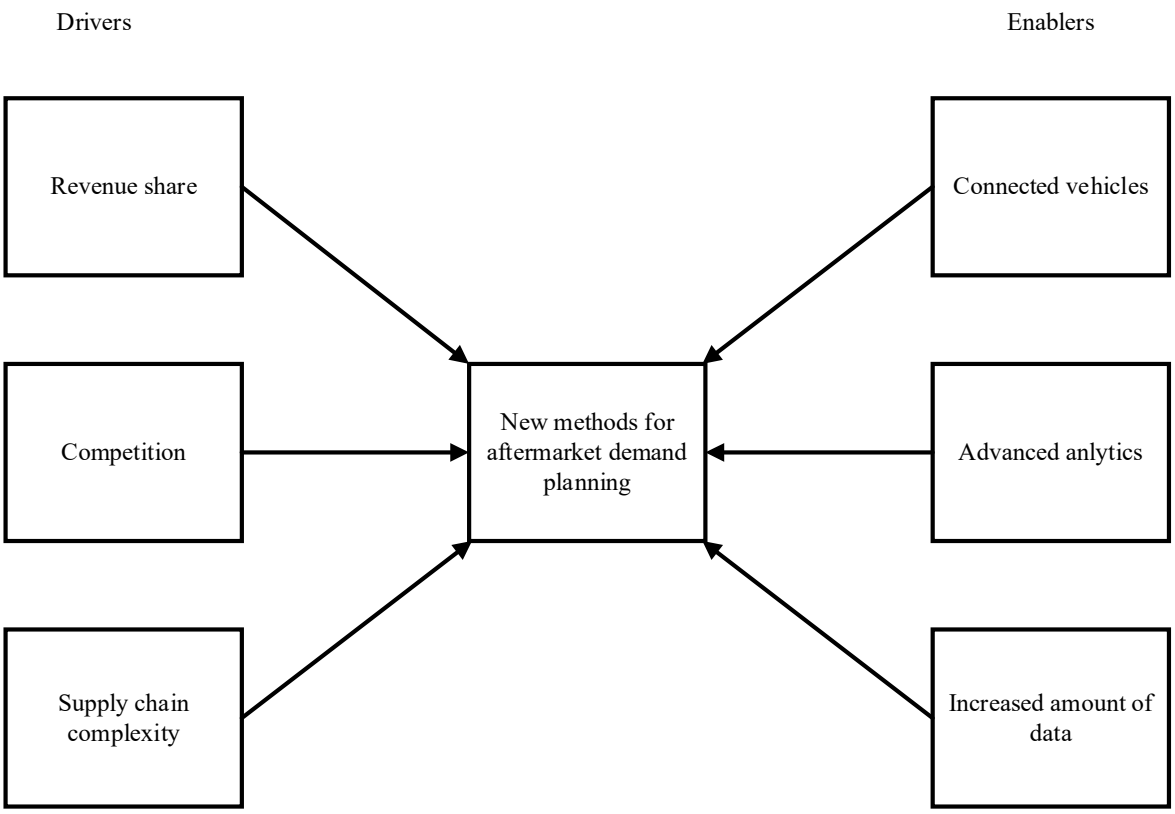


Figure 1.1 Motivation for the research (drivers and enablers)

1.3 Purpose and scope

The problems and opportunities described in the background, i.e. the significance of an important aftermarket supply chain due to the large revenue share, competition from existing and forthcoming competitors, the specific complexity challenges of aftermarket supply chain planning and the opportunities to take advantage of the product-in-use data and advanced analytics, lead to the purpose of the licentiate thesis: to investigate how product-in-use data can

be used in, and improve the performance of, demand planning processes for aftermarket services, by describing and explaining related performance effects and challenges.

Furthermore, the appropriate methods, given the contextual factors, such as spare parts usage, life cycle and demand patterns, were evaluated and six propositions were developed, one of which was tested in a quantitative study.

In order to fulfill this purpose, a first step was to understand the underlying reasons for the customer demand for spare parts, i.e. what causes the demand for different types of spare parts and contexts, as well as analyzing the potential effects on the demand planning processes, resulting in eight proposed methods described as interventions for an improved aftermarket demand planning process. This was the focus in the first study while study 2 went deeper into explaining and evaluating the challenges of developing and applying product-in-use data driven methods in aftermarket supply chain processes. Finally, some of the propositions in study 1 were tested in a confirmatory study, by applying a quantitative regression methodology.

1.4 Research questions

The initial overarching research problem, how can the aftermarket planning process benefit from data generated by connected vehicles (product-in-use data), prompted a thorough literature search regarding aftermarket demand planning methods and product-in-use data usage in an aftermarket demand planning context. The use and effect of connected vehicle data in causal-based demand planning were studied by attempting to answer the first comprehensive research question:

RQ1: How can product-in-use data be used in (1A), and improve (1B), the performance of the demand planning process for aftermarket services?

During the first study it became evident that two more detailed studies were required; 1) a qualitative study in order to analyze the challenges of developing and applying causal-based demand planning methods using product-in-use data and 2) a quantitative study of the effects of the propositions developed in study 1. Making a separate study of the challenges is motivated by the fact that without developing the company's capabilities in handling and processing this new type of data, the improved demand planning results will not occur. The quantitative study is motivated by the need to confirm the findings in study 1.

Study 2 was guided by RQ2**:** What are the challenges regarding the use of product-in-use data in supply chain planning for aftermarket services? This study was carried out as a case study involving three different data-driven analytical applications at a single-case company where the unique aftermarket planning processes were analyzed, as well as their interdependencies. RQ2 was derived from the first study since much of the empirical data concerned problems and challenges. This research question also approaches the problem identified during the literature search for the second study where it was evident that the main data-driven challenges can be viewed in two dimensions: 1) data complexity and 2) process complexity

Study number 3 is a quantitative study performed as a machine learning study, supported by qualitative interview data in order to create the most appropriate analytical scenarios. The propositions from the first study served as a foundation for developing a testable hypothesis for study 3. Thus, study 3 was guided by RQ3: What are the effects on forecast accuracy when applying a causal-based regression methodology for spare part forecasting using explanatory factors as exogeneous variables?

1.5 Conceptual framework

The above research questions and analysis of the results from the literature review resulted in the research model below (Figure 1.2). The model is adapted from the CIMO framework (Jonsson and Holmström, 2016), where C stands for context, I for interventions, M for mechanisms and O for outcome. Many mechanisms have an impact on the outcome of the interventions, proposed methods, studied in relation to RQ1. These mechanisms imply that a firm that wants to benefit from processes utilizing product-in-use data needs to build up capabilities that most firms currently does not have. These capability gaps, expressed as challenges, need to be mitigated to obtain a positive outcome from such processes and are related to RQ2. Hence, in this research model, the mechanisms are replaced by challenges.

Categories of causal-based forecasting methods and categories of spare part classification attributes have been identified. Furthermore, types of connected vehicle data with potential use in causal-based forecasting has been identified, as well as potential challenges in utilizing product-in-use data in analytical models. The first study relates to RQ1 and has an explorative approach and sets out to discover and evaluate proposed methods with potential to improve the aftermarket demand planning process, while the second study relates to RQ2, i.e.to challenges regarding utilization and development of product-in-use data enabled supply chain planning processes.

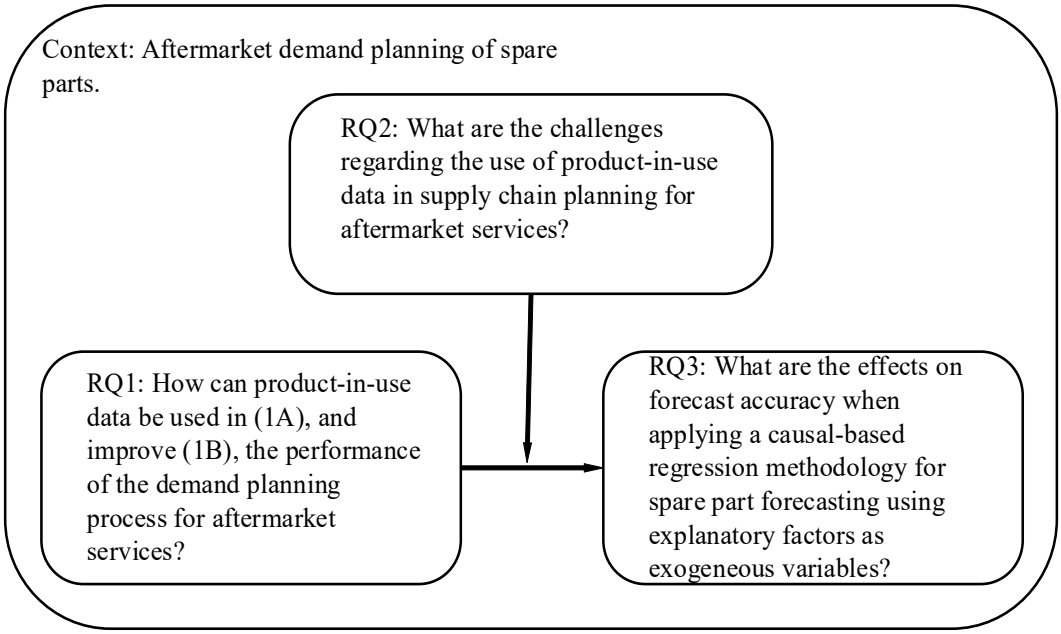


Figure 1.2: Conceptual research model

Finally, the qualitative propositions generated by the first two studies, required a confirmatory analysis to verify the outcome. This study also faced some of the challenges reported in study 2, mainly the challenges related to developing causal-based demand planning methods (e.g. data quality, large data volumes and technology-related challenges. This explains how the three guiding, comprehensive, research questions with the associated studies are related to each other and to the purpose.

Furthermore, since this research area is understudied (Cavalieri et al., 2008), the studies went from pure exploration to being explanatory, and finally, confirmatory. This concept results in a combination of studies that both deals with the opportunities and challenges (capability gaps) and (partly) evaluates the outcome of a causal-based demand planning process for aftermarket

services. See Figure 1.2 for how the different studies guided by the research questions are connected and conceptualized.

1.6 Outline of the thesis

The next chapter describes the theoretical framework that was developed followed by a chapter describing the research methodology. That is followed by a summary of the three attached papers, including the connection and fit between them, results and analysis with regard to the overall research questions and, finally, discussion and conclusions.

2 Theoretical framework

This chapter describes the theoretical foundation guiding the thesis. Section 2.1 briefly defines aftermarket supply chain followed by a section focusing on the aftermarket supply chain planning process, concentrating on the demand planning process. After that, sections describe the underlying mechanisms of spare part demand, the data needed for causal-based demand planning processes and related challenges in order to develop an enhanced casual-based demand planning process.

2.1 Supply chain planning for the aftermarket

An aftermarket supply chain has different characteristics than a manufacturing supply chain and is considered to be more complex due for example to highly unpredictable demand, higher number of SKUs, shorter response times, life cycle responsibilities and high number of distribution points (Cohen et al., 2006). As Stadtler and Kilger (2008) state: there are endless variants of supply chain set-ups which depend on various variables, e.g. location of the suppliers, type of products, customer requirements, lead times, etc. The main difference between production supply chains and aftermarket supply chains for durable goods is that aftermarket supply chains need to plan and stock components using a make-to-stock concept, whereas companies selling durable goods mainly produce to order.

2.2 Aftermarket supply chain planning process

Supply chain planning for spare parts has other challenges and needs to be handled separately from supply chain planning for finished products. Spare parts can be produced in the manufacturing plant like the components of the primary product, or by external suppliers, although sales and distribution of the spare parts are usually organized within a separate organization and/or process (Johnson and Mena, 2009). Hence, the aftermarket organization focus is on distributing spare parts supporting the uptime of the customers products in a cost-efficient way. All spare parts are considered to be externally procured, either via scheduling agreements or via contracts (Dickersbach and Passon, 2015).

A common model defining SCM processes is the SCOR model which encompasses the process below (Supply-Chain Council, 2002 and Supply-Chain Council, 2007b): Plan: Plan covers processes to balance resources with demand, as well as the communication of demand within the supply chain.

Source: Source addresses the selection and follow-up of suppliers

Make: Production planning, production, test, quality etc.

Deliver: Deliver covers processes such as order reception, reservation of inventories, generation of shipping documents and invoicing.

Return: In the scope of return are processes for returning defective or excess supply chain products as well as MRO products.

This thesis focuses on the plan-process, particularly demand planning. The process type is further developed into process categories (level 2) and process elements (level 3). According to Stadtler and Kilger (2008) the latter level describes the detailed process steps and best practices. An example of a demand planning process on level 3 is illustrated in Figure 2.1.

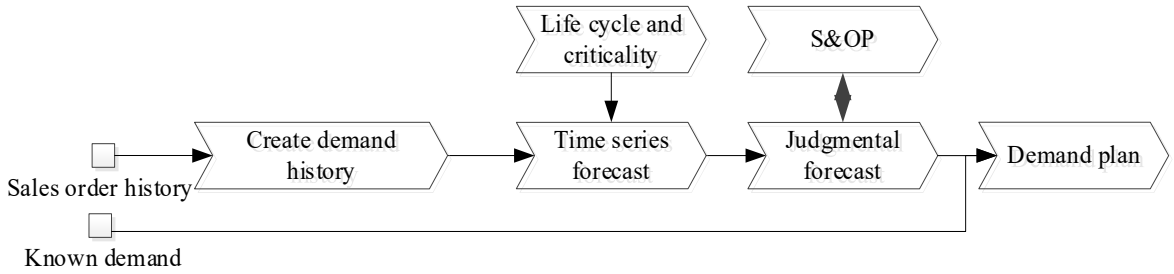


Figure 2.1: Demand planning process (adopted and simplified from the aftermarket of the case company).

The demand plan consists of the forecast and demand known beforehand, such as planned distribution demand and campaign demand. Planned distribution demand is derived demand from a child location of the delivering distribution center, e.g. spare parts needed for a pre-planned maintenance activity or a net requirement from the DRP planning at another DC. A forecast is created for each SKU per planning period (usually weeks or months) in the supply chain. This is performed during the planning horizon, which for the automotive industry is usually 12-18 months. The forecast uses the corrected demand history as input, which is built up by the forecast relevant historical customer orders, i.e. not the planned distribution demand or other type of dependent demand from child locations that was ordered on beforehand. The raw demand history is aggregated into time intervals and cleansing for example of. untypical demand is performed (outlier correction).

2.2.1 Classification of spare parts

According to Jouni et al. (2011) the main purpose of classifying spare parts is to create a manageable number of categories of spare parts to be able to effectively manage them. In order to make this classification successful, it is important to consider the specific context a company is operating in, so that the categorization framework supports the business requirements (ibid.). The overall context in this thesis is the aftermarket usage of commercial vehicles and machines, although within that context there are several underlying (operational) contexts, such as vehicles/machines in the mining industry, long-haul vehicles, distribution trucks and others. Even within a sub-context there are variations in demand depending for example on the usage of the vehicle. For example, vehicles in US are operated for more hours and greater mileage than in Europe, on average. Moreover, different markets vary widely regarding spare part demand, for example due to different road and weather conditions.

The complexity of spare part control lies in controlling the nature of demand; unpredictable and sporadic demand of spare parts requiring responsive actions, wide variety of products; large product portfolio with a high number of SKUs (Cohen et al., 2006, Jouni et al., 2011). In a multinational company with a wide product range, for which it needs to provide aftermarket service, an extensive spare parts range needs to be provided. Furthermore, these spare parts usually have great variety when it comes to attributes such as price, demand frequency, usage, etc. In order to apply the best intervention in each case in the demand planning of these parts, hence a classification framework that considers the context related to each firm is needed in order to support planning of spare parts (Jouni et al., 2011).

2.2.2 Forecasting of spare parts

<u>Time series forecasting</u>

On a high level, forecasting methods can be divided into: quantitative methods and qualitative methods. See for example; Makridakis et al. (1998) and Boylan and Syntetos (2010). Most

common quantitative methods are the traditional time series methods, such as moving average and single exponential smoothing (Bacchetti and Saccani, 2012), which can be relevant to use for fast-moving parts.

Many researchers agree that these traditional methods are not sufficient when it comes to forecasting of spare parts with intermittent demand (Dekker et al., 2013, Romejinders et al., 2012, Cavalieri et al., 2008, Hellingrath and Cordes, 2014 and Wang and Syntetos, 2011).

There are several proposals regarding forecasting for slow moving and intermittent demand. The most well-known and commonly used methods are the ones developed by Croston (1972) and a modified variant developed by Syntetos and Boylan (2005), the Syntetos–Boylan approximation (SBA). Other methods in this category are various bootstrapping methods, e.g. Willemain et al. (2004) and Efron (1979). These methods require large computational effort and lack empirical evidence (Boylan and Syntetos, 2010). Other forecasting methods are based on the ARIMA methodology. Makridakis et al. (1998) describes the ARIMA method as a three-step approach: 1) data collection and analysis, 2) selection of forecasting method and 3) forecasting and validation. A different approach is the forecasting method using demand aggregation described by Bartezzaghi and Kalchschmidt (2011). The aggregation could be done at product level, market level or time bucket level. Since time series are not considering the underlying demand, the recent advancement in product-in-use data and analytical capabilities, several calls for research on causal-based methods has been made (e.g. Syntetos et al., 2016).

Causal-based forecasting

The literature study resulted in three categories of causal-based forecasting methods, with potential use for spare parts forecasting. The first category represents methods based on reliability. Cavalieri et al. (2008), for example, describe the failure rate of components combined with the installed base and the number of a specific item installed in each finished product instead of the demand history. Cavalieri et al. (2008) propose two methods of retrieving the data used in reliability forecasting: 1) using databases that carry information about the failure rate for different types of products and 2) Life Data Analysis (LDA), i.e. collecting the failure rate dynamically by monitoring critical components. Dekker et al. (2013) also mention the possibility of using this methodology to predict the demand for returned products used for remanufacturing of used spare parts.

The second category of methods uses data from on-board sensors in vehicles to predict maintenance actions. This category is known as condition-based maintenance (CBM), see e.g. Jardine et al. (2006). The purpose of this method is to improve the planning of maintenance activities on single machines/vehicles/equipment. However, a forecast for a total population could be produced by aggregating the result derived for a single vehicle. The sensor data consists of various types of information such as pressure data, temperature, oil condition and different type of degradation data. Newer vehicles also have on-board diagnostics features that can interpret the sensor data and produce warnings and fault codes.

Djurdjanovic et al. (2003) describe the construct of an intelligent maintenance system (IMS) and its elements. An IMS is an embedded diagnostic/prognostic system that can forecast failures, by monitoring the degradation of machines or components. The assumption is that by analyzing this data, the failure of certain components can be predicted beforehand and the service provider can place the order with the OEM in advance. This type of pre-planning can reduce the average forecast error and/or the total cost and service consequence of forecast errors, since the demand for some items will be considered as a non-forecast-affecting demand. These items will be ordered and delivered on-time for the planned maintenance. Alternatively,

by using this methodology, the lead time from customer order to delivery from the central warehouse of the OEM will be decreased and, hence, less safety stock will be needed.

The third category of causal-based forecasting methods contains methods based on regression analysis, including multivariate variants using more than one explanatory variable. Other causal-based regression models are the ARIMAX and ARX variants. These methods use explanatory factors that correlate with the demand and can also be called leading indicator forecasting. Possible applications of regression analysis can be item demand forecast based on one or more dependent variables generated from the vehicle in use, e.g. fault codes and sensor data (Makridakis et al., 1998). In recent years machine and deep learning methods have been explored and have shown potential benefits. Combinations of methods are also considered as a means for improving the accuracy of demand forecast (Makridakis et al., 1998).

Judgmental forecasting

Makridakis et al. (1998) propose that qualitative forecasting methods could be useful in combination with quantitative methods, especially for medium- and long-term forecasts. According to Makridakis et al. (1998), qualitative methods consist of expert opinions, market research, focus groups and Delphi methods. Boylan and Syntetos (2010) also includes judgmental forecasting in the last step in their threefold approach, known as post-processing, although their finding indicates certain problems in applying these methods, especially manual increases in the forecast caused by bias from individuals.

Mixed methods forecasting

According to e.g. Makridakis et al. (2008) and De Menezes et al. (2000) there is great potential in enhancing the forecast accuracy by mixing methods. For example, by mixing time series methods with causal-based methods in quantitative forecasting, or, by combining quantitative results with judgmental methods.

2.2.3 Pre-planned demand

If a failure/maintenance need for certain components can be predicted beforehand and the service provider can place the order with the OEM in advance. This type of pre-planning can reduce the average forecast error and/or the total cost and service consequence of forecast errors, since the demand for some items will be considered as a non-forecast-affecting demand. These items will be ordered and delivered on time for the planned maintenance. Alternatively, by using this methodology, the lead time from customer order to delivery from the central warehouse of the OEM will be reduced and, hence, less safety stock will be needed.

By monitoring the product-in-use, statuses and faults can be detected well in advance before a service or repair activity is needed, which will enhance the potential number of pre-planned maintenance operations and un-planned breakdowns. Except for the customer benefits in terms of increased up-time it will also contribute to improved capacity planning for the service provider.

2.3 Spare part demand drivers

The causal factors for the underlying demand generating item demand in the aftermarket are mainly driven by a need for maintenance or repair of the finished product, in this case vehicles. Maintenance activities are either planned preventive maintenance or unplanned corrective maintenance and repair caused by a breakdown (e.g. Romeijnders et al., 2012). Preventive maintenance can either be time-based or predictive, condition-based maintenance being a specific variant. For item forecasting, a first classification is therefore to distinguish between

item demands generated from preventive or corrective maintenance activities. These demand drivers are summarized in table 2.1.

Table 2.1: Summary of spare part demand drivers

Spare parts Demand driver	**Description**
Preventive maintenance	Service intervals (time and/or mileage) Reduce need for corrective maintenance Requires high degree of customer loyalty
Predictive maintenance	Based on status of monitored components Predicts failure in advance
Corrective maintenance	After break-down High costs of repair and down-time
Condition based maintenance	Based on status Real-time based

2.3.1 Preventive maintenance

Preventive maintenance is usually based on service intervals (Kennedy et al.,2002), either using mileage or time intervals, and aims to reduce the need for corrective maintenance and consequently reduce maintenance costs. In theory it is easy to predict the item demand caused by this type of activity. Each vehicle has a service plan, and if the maintenance activities and mileage are logged and if this data is combined with the components needed for the next planned maintenance simple logic could be used to create a demand plan with good quality. To be effective, this method requires a high degree of customer loyalty and requires the customer to follow the service plan. A disadvantage with this method is that it is not optimal from a cost perspective and often results in over maintenance.

2.3.2 Predictive maintenance

An advancement in service interval-based maintenance is to predict the maintenance need based on the status of monitored components. Sensor data analyzed by on-board ECUs identifies and evaluates the need for replacement of the monitored components (Yang et al., 2008) in advance, i.e. the maintenance/repair can be planned in advance.

2.3.3 Corrective maintenance

(Yang et al., 2008) defines corrective maintenance as an operation done after a component in a machine, e.g. a vehicle, has broken down. Corrective maintenance, or repair, is usually much costlier than preventive maintenance and results in longer down-time. Hence, researchers (e.g. Romeijinders et al., 2012)) argue the importance of moving from corrective maintenance to predictive or condition-based maintenance.

2.3.4 Condition based maintenance (CBM)

CBM is related to the predictive maintenance approach, with the difference that CBM uses real-time sensor data. Fritzsche et al. (2014) developed a method aimed at optimizing the total maintenance cost, striving to find the best timing for replacing components by balancing the cost of premature replacement (over-maintenance) and the breakdown cost (under-

maintenance). Fritzsche et al. (2014) describe their concept, known as prognostic health management (PHM), by using data such as performance degradation, physical degradation, vibration and usage-based factors.

2.4 Product-in-use data

At least four layers of data sources regarding spare parts demand prediction and forecasting using product-in-use data can be identified. These range from data created by use of the vehicle itself to external data that can be used to analyze the conditions in which the vehicle operates: (1) vehicle on-board or connected-vehicle generated data (operational data, fault codes and sensor data); (2) manufacturing system generated data (installed base); (3) maintenance generated data (item usage) (4) external data sources (external data). See Table 2.2 for summary.

Table 2.2: Data categories for aftermarket SCP (Andersson and Jonsson, 2018)

Data category	**Example of data items**	**Source of the data**
Operational data	Mileage driven, running hours, location, average speed, RPM	Vehicle on-board
Fault codes	Malfunction of critical components and alerts	Vehicle on-board
Sensor data	Oil quality, pressure, temperature, dimensions of brake disks, vibrations	Vehicle on-board
Install base	Current and future number of vehicles per vehicle type and region	Manufacturing system
Item usage	Items included in specific maintenance repair operations. Vehicle repair maintenance history. Failure rate and service intervals	Maintenance system
External factors	Weather, macro-economic factors, competitors	External data sources

2.5 Big data analytics challenges

The term big data is sometimes used without any in-depth knowledge of what it means, however the most common definition refers to volume, variety and velocity (see e.g. Hazen et al., 2014).

Volume simply describes the size of the data used for analytics in bytes, usually terabytes or petabytes where big data is concerned. There is no exact definition, although Russom (2011) refers to users storing between 3 and 10 terabytes. Another example in Markopoulos (2012) states: ‘it is calculated that a personal car manufacturer generates 5000 data samples every 33 ms, resulting in 152,000 samples per second, 9 million per minute, 13 billion per day, and 4 trillion samples per year.

Variety stems from the fact that firms retrieve data from more sources than previously. These sources keep data with different kinds of information and with different structures and underlying technologies. Examples of data sources are: sensor data (e.g. from connected vehicles or smart devices), sales transactions, geographical data from connected vehicles or route planning systems, inventory data and delivery schedules (Russom, 2011).

Velocity relates to the fact that data collection and updating has also undergone major changes in most companies regarding frequency, from weekly/monthly to daily/hourly, and sometimes on-line (Waller and Fawcett, 2013).

What are the benefits of big data analytics in Supply Chain Management (SCM) in general and Supply Chain Planning (SCP) in particular? Ozgur et al. (2015) claims the possibility of transformation of the business. In a study by McAfee and Brynjolfsson (2012) it was reported that data-driven companies are more productive and profitable than non-data driven competitors. Others consider big data analytics within SCM as a competitive advantage (e.g. Waller and Fawcett, 2013).According to Sanders (2014) it is the combination of big data and analytics that creates value (see Figure 2.2).

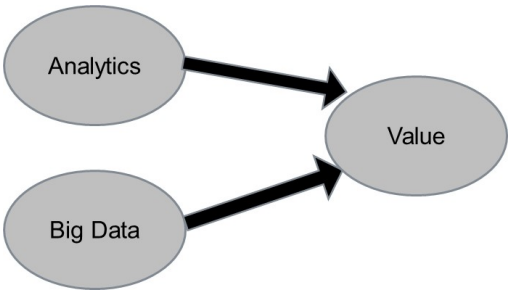


Figure 2.2 Interaction of big data and analytics: Sanders (2014)

McAfee and Brynjolfsson (2012) claim that big data usage has the potential to revolutionize how management decisions are made through the use of data-driven decisions. Schoenherr and Speier-Pero (2015) suggest that big data analytics leads to improved decision-making capabilities, ability to improve supply chain efficiency, enhanced demand planning capabilities, decreased supply chain costs, and increased visibility.

Despite the potential benefits that big data analytics can bring to business in general mentioned above and the supply chain in particular, there are obstacles and challenges to overcome. Kache and Seuring (2017) identified the following key challenges for exploiting big data on the supply chain level: supply chain governance, integration and collaboration, IT capabilities and infrastructure, information security, and human skills. These types of challenges could also be expected when implementing big data analytics in supply chain planning.

Governance concerns the challenge of setting common goals, defining the direction for future direction of big data analytics in a supply chain context, coordinating activities and ensuring that rules are followed in order to enable efficient collaboration. Integration and collaboration relate to the willingness of the supply chain parties to share information through the network. Kache and Seuring (2017) argue the importance of sharing the data in the supply chain from an end-to-end perspective. IT capabilities and infrastructure are key resources regarding the opportunities for effective usage of the available supply chain data, and contain the software, hardware and communication capabilities, including common definitions of standards and interfaces. Information security relates to the management of sensitive data, e.g. customer data. If the customers do not trust how their data is secured, it might lead to reduced willingness to share data with other supply chain parties. Human skills can be described as the companies'

ability to attract people with the needed skills in big data analytics as well as to develop their human resources in this domain. Without a strong mass of analytical skilled personnel, it will be difficult to produce valuable outcome based on big data in the supply chain.

McAfee and Brynjolfsson (2012) describes five challenges regarding big data usage in a supply chain context: Leadership is the reason for being successful with big data application is not having access to the best data, it is rather a result of how executive leadership acts and communication. Important issues regarding leadership are setting clear goals, understand the opportunities that big data can enable and how the market develops. Talent management regards the critical aspect of big data has switched from the cost of data towards the critical skills to analyze the data (McAfee and Brynjolfsson, 2012). Data scientists, statistical analysts, resources dealing with cleaning and organizing data are critical. The resources that are both skilled data scientists and have a good understanding of the business needs are especially valuable. See "Data Scientist: The Sexiest Job of the 21st Century," by Davenport and Patil (2012). Technology refers to the tools available to handle the volume, velocity, and variety of big data have improved greatly in recent years, and the cost of these tools has fallen dramatically. Furthermore, much of the software is open source. However, using these tools requires skills that currently most companies lack (McAfee and Brynjolfsson, 2012). Decision making emphasize that relevant decisions are made by people who understand big data-related problems. Moreover, McAfee and Brynjolfsson (2012) also argue the importance of a flexible organization where people with relevant skills are brought together and are involved in the decision-making process. Finally, Company culture regards how to be successful in the transformation to becoming a data-driven organization, companies need to move away from decisions based on pure instinct and convert the company culture to a fact-based, data-driven culture.

Other researchers have also recognized big data analytics challenges, even though their research had a different purpose than exclusively exploring challenges regarding big data analytics within SCM. The following examples are important to mention: Arribas-Bel (2014), emphasize that a major challenge for the future of big data in forecasting is the advanced skills needed to understand statistics and implement new models. Zhong et al. (2016), mention the following big data challenges from a supply chain perspective: data collection, data transmission, data storage, processing technology, decision making based on big data and interpretation and applications of big data in SCM. Hazen et al. (2014) focus solely on data quality. According to LaValle et al. (2011), the major challenges regarding big data analytics are the management and cultural challenges. Manyika et al. (2011) postulated that there would be a huge lack of skilled data scientists, as well as skilled business analysts and managers who can facilitate company's use of big data analytics in the future. In 2018 there would be a shortage of 140,000 to 190,000 data scientists and 1.5 million business analysts and managers with these skills.

In order to synthesize the literature based on big data challenges, the following categories were common: data, organization, technology, and people skills. Data-driven challenges are related to 3V, and concern activities stemming from the volume, variety and velocity of the data. The following aspects are included in the organizational dimension: the ability to take effective and efficient decisions, transformation of the company culture to become a data-driven organization, coordination and directives for big data initiatives and applications and leadership in terms of communicating a supportive and clear vison, creating an efficient organization and taking data-driven decisions. From a technological perspective, an organization should invest in tools that manage the data itself (data storage, communication, data mining data collection etc.) Regarding people skills, the ability of business analysts, managers and data scientists to

meet the big data challenges are crucial. In this respect, it is not only the analytical skills that are critical, communication, decision making, and business domain knowledge are also of utmost importance.

2.6 Summary of the literature

The fundamental findings from the literature review are presented in alignment with the conceptual CIMO model, introduced in section 1.6 (Figure 2.3).

Given the context of the thesis; aftermarket demand planning using product-in-use data as explanatory factors as underlying demand drivers, the literature review resulted in three overall causal based demand planning methods, showed in figure 2.3 on the left hand side (potential utilization of product-in-use data in demand planning). These methods, used in spare parts forecasting and/or pre-planning of maintenance or repair, requires some capabilities regarding the dimensions: data, organization, technology and people skills (figure 2.3 under challenges regarding utilization of product-in-use data in demand planning), which drives required challenges in order to receive a positive outcome (figure 2.3, right hand side, i.e. potential outcome). This CIMO logic indicates that for the given context, a positive outcome (e.g. increased forecast accuracy and reduced demand un-certainty), can only be achieved by using the potential methods if the related challenges are overcome.

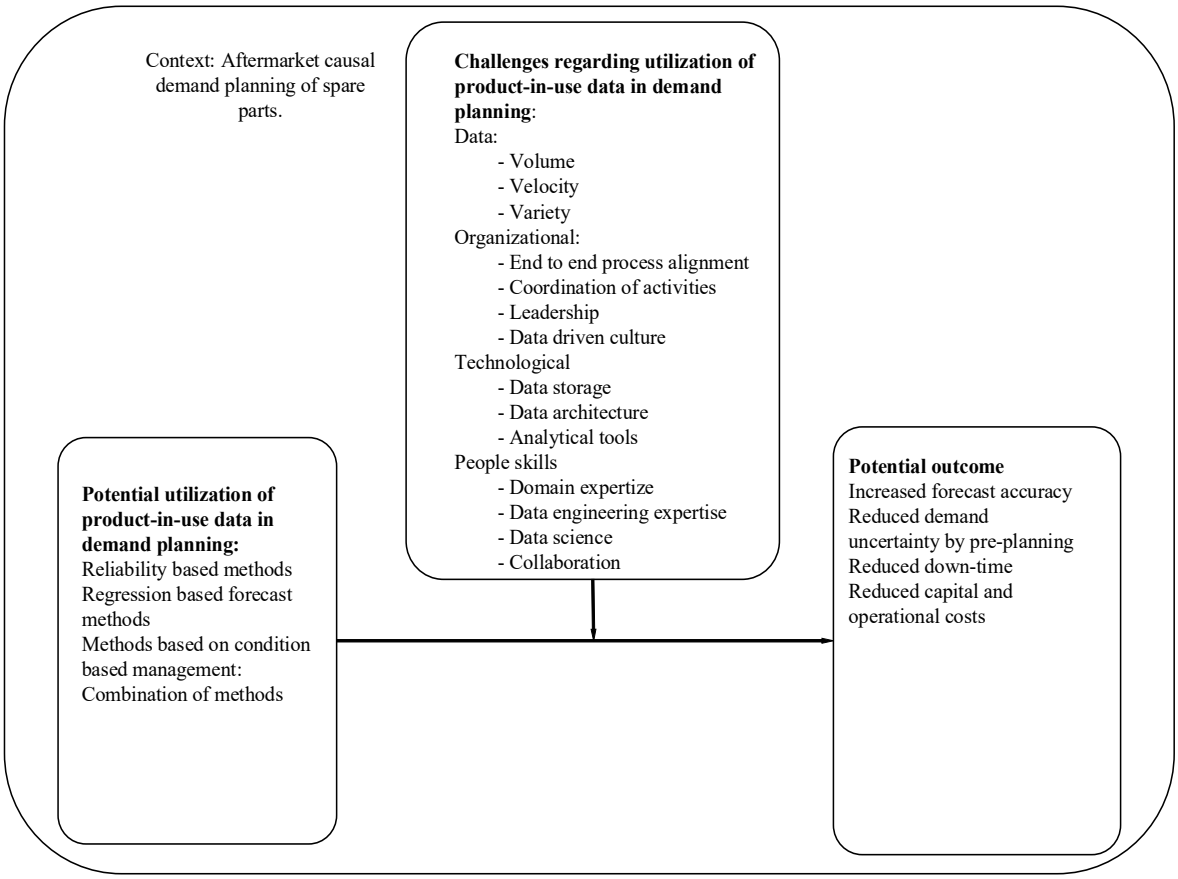


Figure 2.3: Literature summary in relation to the conceptual model.

3 Research methodology

3.1 Research process

The research project was initiated as a collaborative project between Service Market Logistics (SML) at Volvo Group Truck Operations (GTO) and Chalmers, Technology Management and Economics (TME), some overall directives and guidelines being given from the start. The project was defined with 4 different work packages:

1. Segmenting the automotive after-market portfolio
   The potential of utilizing connected vehicle information for simplified and innovative aftermarket demand planning is not the same for all markets, components, service events, etc. We therefore have to understand how to differentiate the planning and utilize new information differently for various segments of planning objects and decisions
2. Innovating demand planning utilizing connected vehicle information
   This work package focuses on exploring the effect and developing new forecasting models for 'critical' segments, by combining or exchanging the present use of historical sales data with connected vehicle information. Another area of this research addresses how to aggregate data, find methods with which to analyze and make the data relevant, and to establish implementable systematic processes for demand planning the supply chain.
3. Advancing the sales and operations planning
   This work package focuses on assessing how the use of connected vehicle information and innovations in the demand planning process enables the development of an advanced aftermarket S&OP process in the heavy vehicle industry. This includes assessing how demand planning and supply planning over the 1-2 years horizon in the entire supply chain can be improved.
4. Piloting, future projects and knowledge dissemination
   This final work package is concerned with piloting developed demand tools, formulating important and relevant strands for further research, and dissemination. As outlined in the first sub-section of this research plan, there are several supply chain planning issues which could benefit from utilizing connected vehicle information. This project focuses on demand planning and S&OP, but another important aim of the project is to identify strands of future research related to how to utilize connected vehicle data in processes and decisions related to supply chain design, planning and execution.

The first two work packages above were transformed into a more precise research problem with the purpose and aligned research questions (described in the conceptual model in section 1.6). The research process has been carried out using an iterative approach, i.e. first an overall literature study was performed to gain insights in the research area, followed by detailed literature reviews for each topic related to the individual papers. In general, the following activities has been carried out through the research process (Figure 3.1), which are discussed in detail in the research design section (section 3.2)

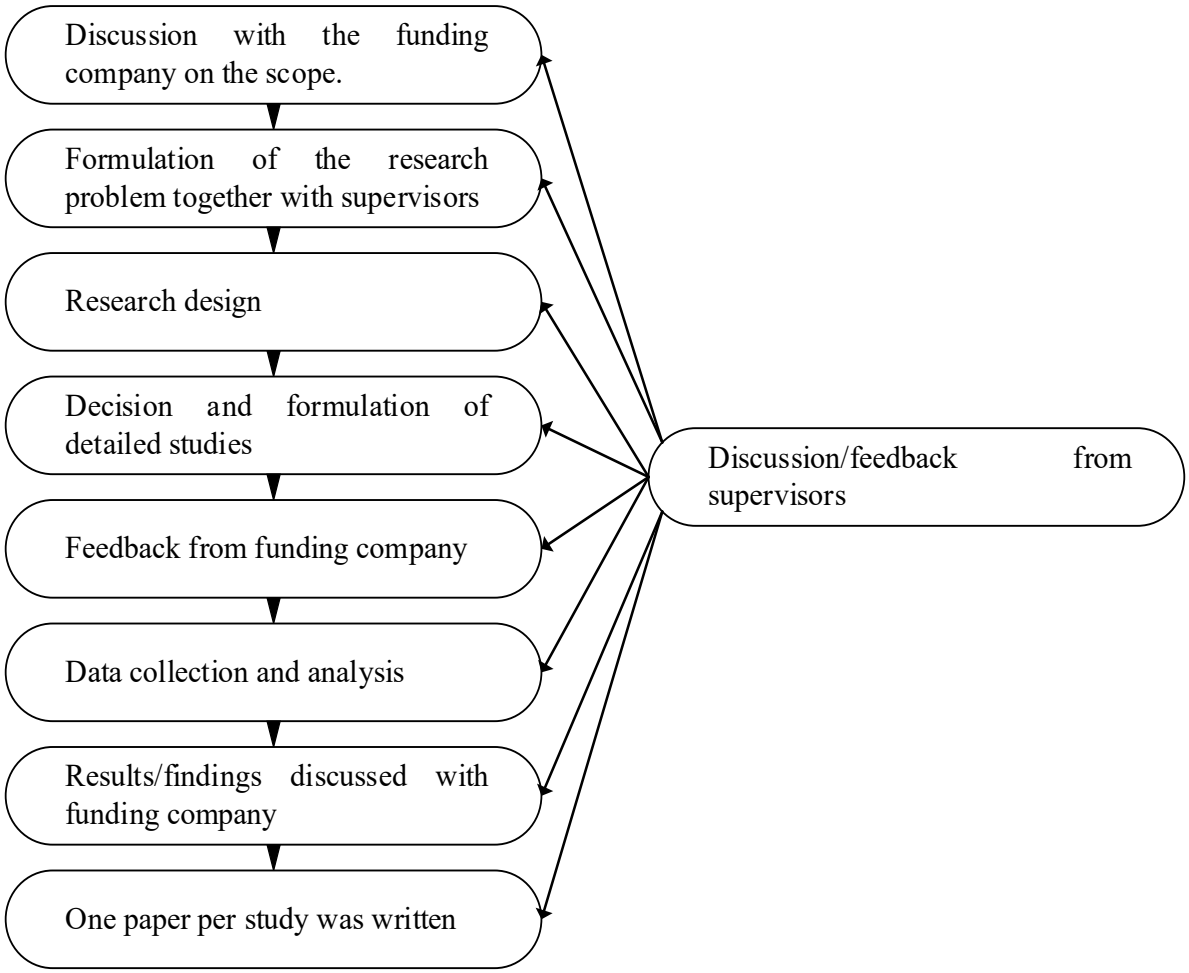


Figure 3.1: Overview of research activities

3.2 Research design

In order to understand a phenomenon in-depth and when the research questions are open-ended a qualitative research design is advisable (Voss et al., 2002). Bryman and Bell (2015) argue that qualitative research is relevant when contextual understanding is required. The under-researched nature of the field of aftermarket supply chain planning utilizing product-in-use data and the fact that these studies cover an emerging phenomenon requires the research problem to be studied closely in its context (Flick, 2014). Furthermore, RQ1 and RQ2 are of an open-ended nature and require in-depth understanding of the study object (the aftermarket demand planning process). Hence, a qualitative research approach was considered appropriate for the first two studies.

This research approach relates to design science (Holmström et al., 2009), which is an approach for proposing and developing new operational practices and processes. In order to really understand and propose how to develop new practices and process configurations, this approach normally requires close interaction with a real-world case during a longer time horizon. Therefore, design science is often combined with a single case study methodology, both when focusing on the early development phase (Öhman et al., 2015) and on the later theory elaboration phase (e.g. Jonsson and Ivert, 2015) of new and innovative practices. Causal-based forecasting methods are emerging but not completely new. From the case company and from participation in industry conferences, pilots regarding using product-in-use data for causal-based forecasting have been examined. However, existing usage of product-in-

use data in aftermarket demand planning processes is not documented, understood or implemented on a large scale. The qualitative components in the research in these theses can be defined as a later design science approach focusing on case-based theory elaboration (Ketokivi and Choi, 2014). How to use product data in causal-based forecasting is considered to be existing knowledge, but not how to relate various data and methods to different item categories and demand characteristics, and what outcome to expect from usage. The above argument is also applicable to the study regarding analytical challenges using product-in-use data in the same context as above. Existing studies in that area mostly focus on either big data and/or advanced analytics challenges as a high-level phenomenon or as, for example, in Kache and Seuring (2017) at company or supply chain levels. No research has been found studying challenges for this contemporary phenomenon in a real-world context.

To reinforce the findings of the propositions in study 1, a mixed method research approach was selected for the third study. A confirmatory study was designed with the purpose of testing some of the propositions from the first study. Together with guidance from relevant literature on causal-based forecasting, a machine learning approach was selected. The data sample was determined using a qualitative, collaborative approach, in close cooperation with data scientists and domain experts at the case company.

Figure 3.2 presents an overview of the first two (qualitative) studies.

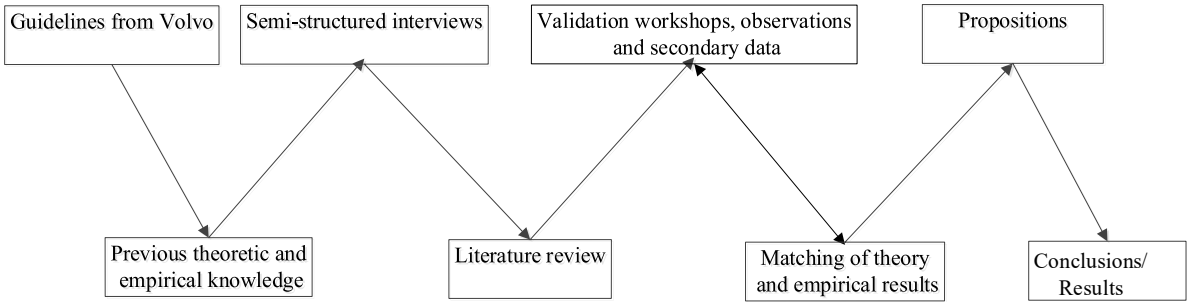
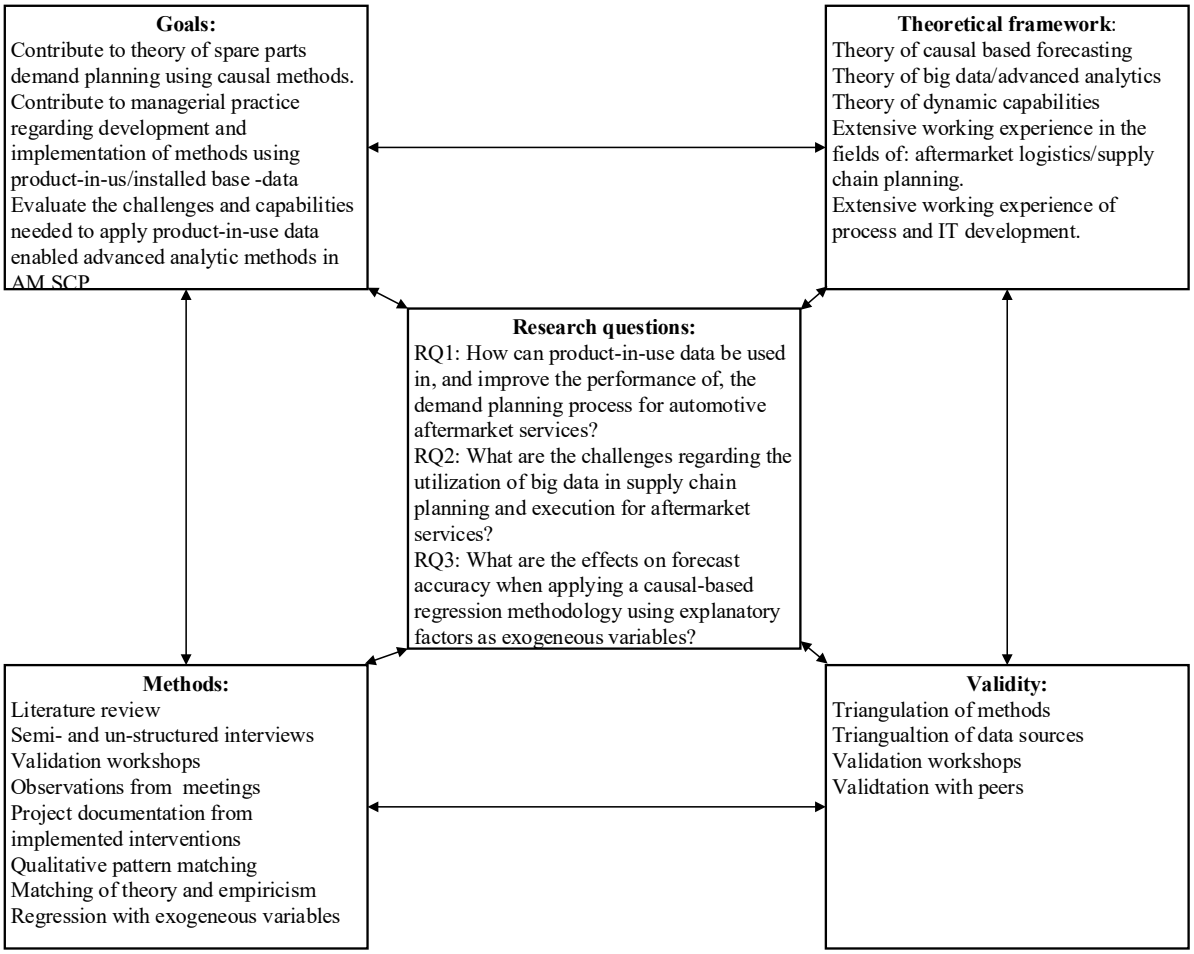


Figure 3.2: Research process Studies 1 and 2

The bidirectional arrow in Figure 3.2 illustrates the abductive approach (Dubois and Gadde, 2002) leading to the propositions in the qualitative studies. This abductive approach implies an iterative process where theory was matched with empirical findings (Spens and Kovacs, 2006). In consideration of the fact that this research area is under-studied, theory from several fields was used in the theoretical framework and new practical and scientific knowledge was developed through this abductive reasoning approach.

The conceptual model adopted from Maxwell (2013), described in Figure 3.3 summarizes the different components of the research design adopted in this thesis. As illustrated in the figure, the components interact with each other, meaning that a change in one component impacts other components. Furthermore, this conceptual research design framework should be viewed as a guide for the thesis, i.e. each study has an individual detailed research design.

Figure 3.3: Conceptual research model, (Maxwell, 2013)



The starting point for all three studies was the research goals, developed from the guidelines provided by the company funding this research. From these goals initial research questions were developed, but since the model is interactive, these questions have undergone several modifications. In turn, the research questions have been the foundation for the other components of the model, i.e. methodology, theoretical framework and validity. In the research process, there have been many opportunities to improve the contents of the different components of the model and, as stated previously, the model constitutes an interactive process, hence a change in one part of the model has often led to changes in other parts.

3.3 Case selection

Given the lack of empirical research regarding aftermarket demand planning (Bacchetti and Saccani, 2012) and the need for in-depth understanding of a new phenomenon, a single case study approach was chosen (Eisenhardt and Graebner, 2007). Due to access to rich empirical data, which strengthens the opportunities for theory building (Eisenhardt and Graebner, 2007), the chosen case company is Volvo Service Market Logistics. Other motives for choosing this case company are: 1) the case company has launched several initiatives in the area of advanced analytics using exogeneous data in the aftermarket context, and 2) the case company manages the aftermarket logistics for several brands operating in different business segments in a global environment (heavy vehicles, construction equipment, buses and marine and industrial engine applications). Study 1 is connected to and motivated by RQ1, which in turn corresponds to Work Package 1 in Section 3.1: the potential of utilizing connected vehicle information for simplified and innovative aftermarket demand planning. study 2 was initially a subset of study

1, although it became a separate study due to recommendations made by reviewers of the original version of Paper 1. This study addresses the challenges of developing and implementing causal-based methods based upon product-in-use data in an aftermarket demand planning process. Hence this study is related to RQ2. Study 3 is a quantitative study (with an initial qualitative assessment) adopting regression analysis with exogeneous variables and is motivated by the need for confirmation of the qualitatively based results in study 1. Study 3 is guided by RQ3, which was derived from the initial Work Package 2 (Section 3.1)

3.3.1 Study 1

At the start of study 1, the case company conducted some initial tests of utilizing product-in-use data in their spare parts forecasting process. Additionally, some studies had been carried out regarding enhancing the demand planning process by boosting the utilization of pre-planned demand. Since this emerging phenomenon is high on the agenda at the case company, and based on its initial endeavor and the fact that accessibility to empirical data was very good, the decision to study the utilization of product-in-use data in the aftermarket context by choosing Volvo SML was expected to provide the best basis for this study. Hence the unit of analysis for this study is the demand planning process.

3.3.2 Study 2

As mentioned previously, the decision to carry out study 2 was a consequence of the results from study 1. Since the characteristics of the problem in this study (challenges in aftermarket supply chain processes) are of a similar nature as in study 1 (understudied, new phenomena and utilization of product-in-use data in a new context), it made sense to continue with the same case company. A search for relevant cases already started or ongoing initiatives was performed for this study. The criteria for selection was that these processes should contribute to the aftermarket supply chain planning (or execution), should have been tested or used as applications in a process and use product-in-use data. The selected processes which met the selection criteria are spare parts forecasting, pre-planning and vehicle monitoring.

3.3.3 Study 3

Following the propositions in the first study, the choice was to perform a confirmatory study on a sub-set of these propositions. The selected criteria for testing improved forecast accuracy by applying a regression-based forecasting method with explanatory variables were phase in spare parts (between one and five years since launch), medium and high frequency spare parts (average demand per period > 1.5). By having access to expertise in the fields of data science and demand planning, as well as access to rich historical data of demand and explanatory factors (fault codes, driving distance and operating hours), the decision was made to choose the same case company as for the first two studies.

3.4 Data collection

For Studies 1 and 2, the collected data consist mainly of qualitative components, whereas study 3 constitutes of both qualitative and quantitative data.

3.4.1 Study 1

The data regarding study 1 was collected between August 2016 and July 2017 and was collected in several different ways (Table 3.1). The data collection process for the first paper started with a set of seven interviews.

Table 3.1: Data collection summary, study 1

Data collection method	**Set-up**
Interviews	Semi-structured
Observations	Participation in daily work of demand and inventory planners.
Internal documentation	Reports from tests regarding causal-based forecasting
Informal discussion	Unstructured meetings and discussions with various colleagues
Workshops	Validation meetings with various management teams

The interviewees were purposefully selected and are considered key informants, i.e. people with deep knowledge in their respective domain (Yin, 2014). For each of the interviews, extensive notes were taken, and the respondents were contacted to give their views on the results. After the first analysis phase, certain key informants were contacted for complementary questions and clarifications.

Furthermore, observations, internal documentation, informal discussion and workshops were also used for data collection, and later for validation purposes. Internal documentation mostly contains results from small data mining tests, where causal-based demand planning methods have been explored. Informal discussion took place on numerous occasions, mostly based on the role I play and the fact that I am involved in the development of the advanced analytics function at the case company. Observations were mainly carried out by sitting next to demand and inventory planners when they were using the causal-based methods in real-life situations. Workshops with various management teams and together with specialists were a valuable tool for validation of results and confirmation of results. One workshop was held together with an external company (also in the vehicle industry). The purpose of that workshop was to increase generalizability.

3.4.2 Study 2

This study too is in general based on qualitative data. Since the original purpose of Paper 1 was to include challenges, considerable data from interviews from the first study already existed. However, this data was not sufficient to fulfil the requirements of data to perform this study. Hence, additional seven interviews were performed (Table 3.2).

Table 3.2: Data collection summary, study 2

Data collection method	**Set-up**
Interviews	Semi-structured
Informal discussion	Weekly management meetings
Workshops	Validation meetings with various management teams. Cross-functional workshops

To gain a comprehensive understanding of the challenges using product-in-use data in an aftermarket context, individuals from multiple functional areas were interviewed (LaPlaca et al., 2018), such as data scientists, demand and inventory planners, and managers in functions representing the customer interface, including logistics and aftermarket services.

Moreover, participating in weekly management meetings in an advanced analytics department and several cross-functional workshops (e.g. between service market logistics and truck monitoring, and between service market logistics and up-time development) provided an additional opportunity to collect empirical data. Validation workshops regarding the findings and conclusions of this study were also held.

3.4.3 Study 3

The third study used a mixed-method approach. First qualitative data was collected in order to determine the scope of the project, i.e. which attributes of the spare parts should be selected, and which explanatory factors should be selected. This work was done during several workshops. The next step was to retrieve a raw, temporary data set, which implied matching of several tables from various databases. The data set was then prepared (cleansing and imputation), resulting in the data set used for the quantitative autoregressive method with exogeneous variables (ARX). The resulting data set contains the following fields: spare part number, period number, demand, fault code frequency 1-3, average distance per vehicle and average operating time per vehicle. The number of spare parts is 611, and the number of periods is 65.

3.5 Data analysis

This section addresses the data analysis, describing the analysis approach for each paper.

3.5.1 Study 1

Analysis in the first study was done in five iterative steps. Step one consisted in understanding current empirical usage of causal-based methods at the case company and how that relates to demand planning theory. The second step was to understand the potential data usage for contributing to the causal-based methods. This was aligned with the literature. In the next step, step three, the findings from the previous steps were matched and resulted in a gross list of potential causal-based methods. This was followed by the fourth step, where the potential effects were evaluated and finally, step five, when six propositions were developed. These propositions were analyzed with regard to the spare parts characteristics from the literature review, i.e. utilization, life cycle and demand pattern. Moreover, the propositions and their potential outcome related to combinations of utilization, lifecycle and demand pattern were presented and discussed with the management team of materials management and a team consisting of demand and inventory specialists for validation purposes. An interview with an external automotive manufacturer was held as well, resulting in increased generalizability, considering that these interviewees confirmed the result from the analysis.

3.5.2 Study 2

Interviewees were analyzed to gain an understanding of three supply chain planning application areas utilizing product-in-use data that was selected as cases for this study. The result of this initial analysis provides a comprehensive view of the three different cases regarding data usage and actor involvement. The data was coded and analyzed within each case, according to categories of challenges regarding the four dimensions of challenges (data, human skills, organization and technology), application areas and type of challenge. To understand how the identified challenges relates to each other and to the data and process complexity, a cross functional analysis was carried out.

3.5.3 Study 3

The analysis of study three was initiated by a qualitative assessment of the data gathered from workshops aiming to select the criteria for the sample for the quantitative analysis and the explanatory factors which will most likely produce improved forecast accuracy. This step was followed by retrieval and matching of data which met the selection criteria (age, demand quantity and price). The first criterion is due to the proposition in paper 1 regarding phase-in spare parts, whereas the second and third criteria were used to filter out many low-frequency items where no significant result is expected. The last criterion (standard cost of sales) was used due to no causality being expected on standard items, such as nuts, bolts, fasteners etc.

The resulting data set (part number, demand history, history of explanatory factors and some categorical data) was analyzed using a python algorithm, applying an autoregressive method with the explanatory variables and previous demand as regressors, in order to predict the demand one period ahead.

3.6 Research quality

An overview of the research quality assessment for each study follows. This section first explains how research quality was dealt with in the qualitative studies (Studies 1 and 2), followed by study 3 (mixed method study).

3.6.1 Studies 1 and 2

The traditional research quality criteria are internal validity, external validity, reliability and objectivity (Guba and Lincoln, 1989; Yin, 2013). These criteria are important in quantitative research, although their relevance to qualitative research has been questioned (Bryman and Bell, 2015). Guba and Lincoln (1989) introduced the term trustworthiness as quality criteria, which was used for the quality assessment for the qualitative studies, which are of an interpretive nature. The corresponding criterion for internal validity in qualitative research is credibility (Guba and Lincoln 1989). According to Halldorsson and Aastrup (2003), credibility is determined by how well the constructed findings are in line with the data gathered from the respondents. In the first two studies the risk of lack of credibility was mitigated by several recurrent dialogs with respondents as well as with supervisors and was low since I am very familiar with the problem and the field. The interview guide was also checked with supervisors before the interviews. Furthermore, triangulation of data sources was used to strengthen the credibility (interviews, project reports and observations).

External validity in quantitative research can be compared with transferability in qualitative research. The latter concerns whether the result of a study is generalizable outside its context, which in the case of this thesis is the automotive aftermarket. However, validations have been done with another automotive manufacturer, which agreed with the findings. Moreover, to strengthen generalizability, use of thick descriptions and peer reviews was carried out (Creswell and Miller, 2000). Furthermore, the automotive context contains some industry-unique product data items, but the general research focus on product data in causal-based forecasting should not be case-specific. Transferability of the results to other cases and contexts is enabled since the study design, data collection and analysis are based on conceptual definitions presented in the literature section (Yin, 2014).

Reliability is in parallel with dependability, which concerns trackability. According to Halldorsson and Aastrup (2003), trackability can be handled by documenting the research process and its decisions. Both the qualitative studies have adopted these principles by documenting each step in the research process (research design, data collection and data analysis).

Finally, objectivity is comparable to confirmability, and concerns the degree to which the researcher's bias influences the result. To avoid bias as far as possible, the result was discussed with peers from academia as well as from Volvo. Multiple workshops were held to avoid research bias.

3.6.2 Study 3

The qualitative components of study 3 lead to the selection of data, i.e. the independent variables were chosen for the forecast comparison. Credibility for this phase was supported by interviews with respondents carried out on causal-based forecasting analysis on a minor scale, as well as secondary evidence, such as documentation from such tests. Regarding transferability, the same reasoning as for Studies 1 and 2 is valid, i.e. because the research process is based on conceptual definitions from theory, transferability can be considered to be achieved. To achieve reliability in the quantitative phase of the study, the regression analysis, several measures was taken. The methodological choice and usage of the method (multiple regression in Python), was evaluated by an expert in data science. The data was examined in detail and action was taken to replace obviously erroneous data, as well as for imputation of missing data. Since the sample size was large (~ 700 spare parts), not all individual results were checked visually, but spot checks were performed. To validate the quantitative results, the goodness of fit was measured by the adjusted R^2 value, using the principle of stepwise regression, i.e. after adding a new explanatory factor the adjusted R^2 value was examined. If adjusted R^2 value increases, then the model performs better than the previous one (without that variable) and the variable is kept. The result was also compared to traditional time series forecasts, as well as the actual implemented forecast values. To further strengthen validity, a comparison of the result (with the same variables) was carried out on another (equally large) sample of spare parts, with the same characteristics (age, demand pattern and criticality).

3.7 Summary of the studies and methodological fit

Table 3.3 summarizes the studies with regard to data collection, analysis and research quality

Table 3.3: Methodological fit

RQ	Data sources	Data analyses	Research quality
How can product-in-use data be used in, and improve the performance of, the demand planning process for automotive aftermarket services?	Interviews Secondary data, e.g. documentation and observations	Analysis of patterns in the qualitative data Triangulation the interview data with field test data Matching of empirical data with theory	Peer review Research documentation Validation workshops
What are the challenges regarding the utilization of big data in supply chain planning and execution for aftermarket services?	Qualitative data to guide the simulation study Quantitative data (forecast quality)	Analysis of patterns in the qualitative data Triangulation the interview data with field test data Matching of empirical data with theory	Peer review Research documentation Validation workshops Goodness of fit
What are the effects on forecast accuracy when applying a machine learning methodology using explanatory factors as exogeneous variables?	Qualitative data in terms of interviews and focus groups	Data selection in collaboration with peers Methodological selection by studying previous projects Multiple regression analysis	Peer review Triangulation of qualitative vs quantitative results Validation with second sample

3.8 Research approach/my role as Industrial PhD student

Being an Industrial PhD student with long experience in the automotive industry, mainly working in the field of supply chain management with development activities and projects, creates considerable opportunities for research, as well as there being challenges and risks to be mitigated.

During these 20 + years I have been in contact with and shared experiences, best practices and information with many firms (both in the automotive industry and in other industries), as well as with consultants and researchers. Hence, this has been an advantage in defining the research problem, interview questions, but also, in the case of interpreting the data and in that way creating knowledge. This experience has led to a strong belief that a phenomenon is usually not very easy to study, i.e. there are many dimensions and dependencies to consider. Thus, a simple 'cookbook' or one-size-fits-all methodology is not feasible in researching an emerging phenomenon, especially for generating new theory (Bryman and Bell, 2015). On the contrary, there is also a risk in being closely related to the research object. According to Coughlan and Coghlan (2002), one risk in doing collaborative research is that of researcher impartiality, i.e. having too close a relationship with the case company can result in biased case results. This risk has been an obvious potential risk from the first day of my doctorate and, it is dealt with in section 3.7 (research quality).

Both the positive and negative aspects of this type of research have been dealt with in a thorough manner, e.g. by peer reviews and by holding validation workshops. Another ethical consideration experienced in the doctoral journey so far is the difference between the somewhat conflicting goals between the funding company and the academic requirements in terms of requirements of rigor in research vs the major requirement from a practitioner point of view of practical relevance. In order to deliver both rigor and relevant research, several prevalent measures have been taken. Already mentioned are the actions regarding feedback loops and continuous reviews together with peers in academia and key participants from the case company. Further considerations have been to be aware of my role as researcher, e.g. be aware of bias risk and document and pay careful attention to the chain of evidence and confirm empirical findings with theory (Näslund et al., 2010).

Without claiming that action research methodology has been performed thoroughly, it is still fair to claim that the collaborative nature of the research represented by this thesis has borrowed some principles from this methodology, for example, the cyclic approach where the planning phase has been done together with key participants from the case company, the feedback loops in the analysis phase and implementation of proposed actions.

3.9 Limitations

As an Industrial PhD candidate, I have access to rich empirical data at my employer. At the same time, a limiting factor is that I cannot collaborate with competitors. Two main risks in being employed in the industry doing research in the academia are: 1) the risk of being biased and 2) conflicts of interest. These risks were mitigated in the following way:

1) Close cooperation and many discussions regarding the data collection results with my supervisors and validation sessions of the outcome with several management teams at the case company. A validation workshop with some supply chain specialists at another vehicle company was also held.

2) Conflicts of interest may concern what to publish in academic journals vs which findings to keep internally within the company. This was dealt with through detailed discussions with managers at the case company.

4 Summary of the results of the appended papers

In this section a summary of the results of each paper is presented. This section focuses on the results of each study (each represented by a paper).

4.1 Paper 1 (Big Data in spare parts supply chains: The potential of using product-in-use data in aftermarket demand planning)

The first paper addresses the potential usage of product-in-use data in an aftermarket demand planning context, identifies and proposes relevant causal-based methods and investigates the applicability of each method in relation to different types of spare parts, grouped by identified contextual attributes. The identified attributes that have a potential impact on the outcome of these causal-based methods are life cycle phase, utilization (repair or maintenance spare part) and demand type (intermittent or high frequency). RQ1 is revisited to describe the agreement between the guiding research question and the results of this paper.

RQ1: How can product-in-use data be used in (1A), and improve (1B), the performance of the demand planning process for aftermarket services?

Literature matched with empirical data resulted in data sources with a potential impact on future demand. This data comprises four levels with varying sources: 1) connectivity data (sensor data, fault codes and operational data), 2) installed base data, 3) item usage and 4) external factors (Table 2.1). Moreover, the literature review resulted in three comprehensive groups of causal methods for predicting future spare parts: 1) reliability-based methods, 2) methods based on condition-based maintenance and 3) methods based on regression between the response variable (demand) and the explanatory factors (product-in-use data). For a more extensive description of these categories, see section 2.2.3. These categories where matched with the empirical data in several iterative steps, including reverting to the interviewees for detailed discussions and clarifications. The results of this analysis resulted in eight different proposed causal-based demand planning methods, each based on either one of the overall causal methods (reliability, regression or condition-based management). Furthermore, the proposed methods were evaluated regarding their feasibility for the contextual variables (life cycle, utilization and demand pattern), as well as for the usage of data elements from the gross list of product-in-use data elements. This is summarized in Figure 4.1.

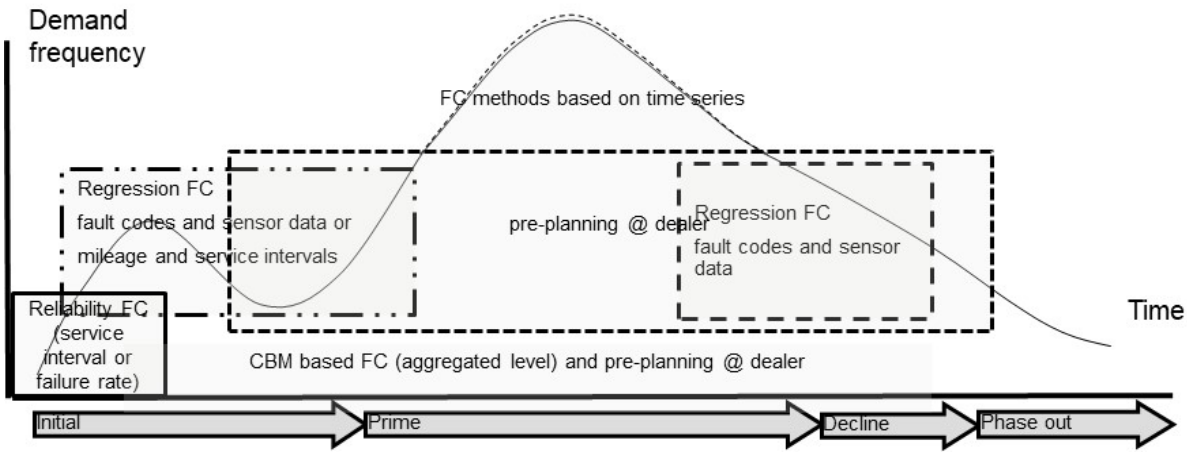


Figure 4.1: Demand planning methods vs spare part categories

The analysis of these proposed methods motivated six propositions on how to use each method for improved forecast accuracy and/or reduced demand uncertainty, including combining methods and feasible contexts (e.g. for which life cycle phase and for which demand categories.

For spare parts, with limited on no demand history, the reliability-based method, using failure rate for repair parts and service intervals for maintenance parts is considered as most appropriate. Further in the life cycle, regression-based methods using fault codes and sensor-based data as explanatory factors for spare parts with medium and high frequency. For components with a higher product cost and are critical for uptime, condition-based maintenance can give an improved forecast, mainly for intermittent and lumpy demand. Alternatively, conditional based maintenance methods can also be used in order to reduce the demand uncertainty by increase the number of pre-planned maintenance operations and, hence, reduce the demand un-certainty. For medium frequency spare parts, in the middle of the lifetime, pre-planning based on preventive maintenance also have a potential impact and can reduce the demand un-certainty. In the later phases, also regression-based methods using fault codes and sensor data has a potential to increase the forecast accuracy. For relatively high frequency spare parts in the stable mature life-cycle phase, the result was that the most appropriate method would be to use traditional, time series-based forecasting methods.

4.2 Paper 2 (Improving supply chain planning processes for aftermarket services: Challenges of using product-in-use data).

This paper focus on the challenges in early phases of developing and implementing applications utilizing product-in-use data from an aftermarket perspective. The paper is guided by the following research question:

RQ2: What are the challenges regarding the use of product-in-use data in supply chain planning for aftermarket services?

Literature findings resulted in four main dimensions in which the challenges can be categorized: data, organization, technology and people skills. The main driver of organizational challenges is the need to create a data-driven culture, including management actions and directions. Regarding technology challenges, the main issues are capabilities for data collection, storage and processing of the large amount of varying types of data with different velocity (hardware, software and data communication). For the data as such, the main challenges are volume, variety and velocity (3V). Regarding people skills, the main challenges are to attract skilled resources with different competences (mainly data scientists and domain experts) and to create collaborative, end-to-end working procedures. Furthermore, the literature review also concluded that these challenges are driven by data and process complexity.

4.2.1 Context-related challenges

Three different processes were studied: causal-based forecasting, pre-planning of maintenance and vehicle monitoring. The first process (causal-based forecasting) was considered to be complex from a data perspective but not from a process perspective, whereas the two remaining processes in the study (pre-planning and vehicle monitoring) were considered to be complex from a process perspective but less complex from a data point of view, Figure 4.2).

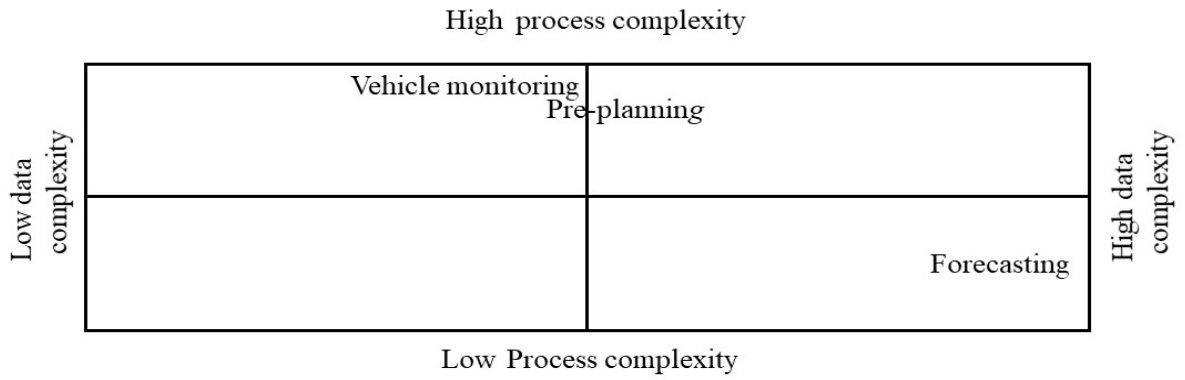


Figure 4.2:Studied processes relations to data and process complexity

The result of the within-case analysis shows the individual challenges and their interdependencies for the three different processes: causal-based forecasting (figure 4.3), pre-planning (Figure 4.4 and vehicle monitoring (Figure 4.5).

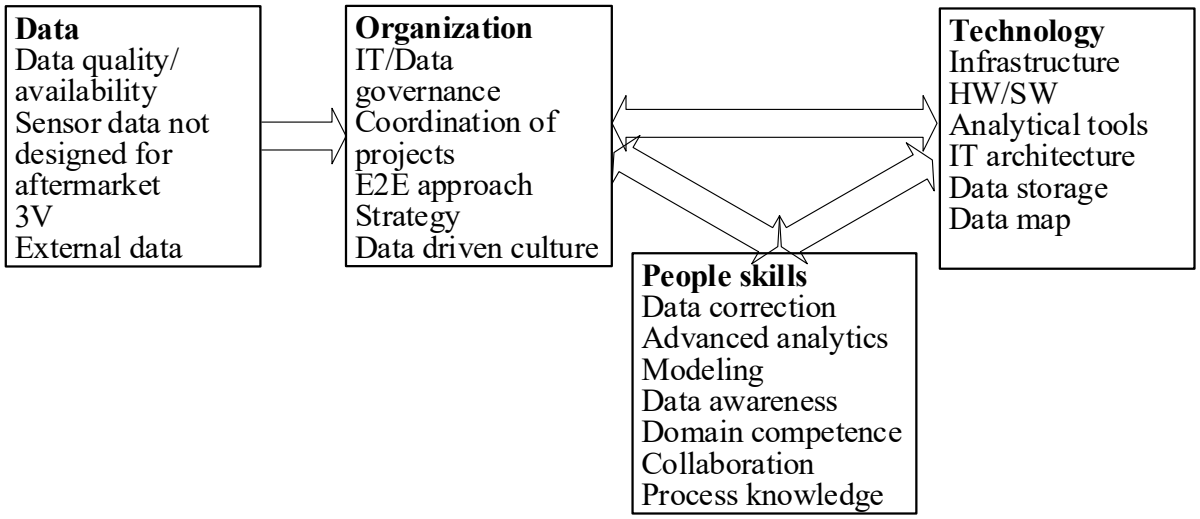


Figure 4.3: Challenges and interdependencies regarding causal-based forecasting

Due to the data complexity of the causal-based forecasting process, the most severe challenges for this process belong to the data and technology dimensions, even though there are also challenges in the organizational and people skills dimensions. Furthermore, an important finding is that the root cause of the challenges is the data and that the other challenge dimensions have a mutual impact. The many varying data sources can be used in many different ways, depending on the large number of different possible causal-based demand planning methods. This data complexity creates challenges in the technology dimension, especially related to IT-architecture. The technology challenges in turn have an impact on the challenges in the people skills dimension, e.g. for data scientists and domain experts regarding challenges of an analytical and data preparation nature. The latter challenges put requirements in the organizational dimension, requiring a high degree of cross-functional ang agile way of work.

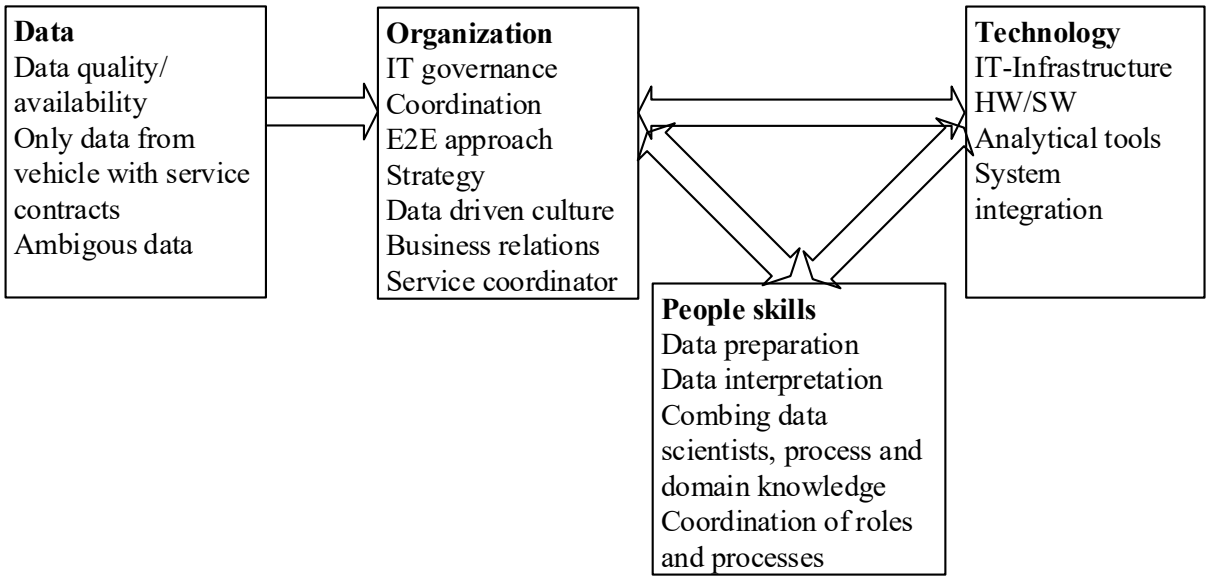


Figure 4.4 Challenges and interdependencies regarding pre-planning

The pre-planning process has considerable challenges regarding people skills and organizational dimensions due to the process complexity. Also, the origin of all other challenges derives from the data challenges in this process as well. Main challenges here are challenges of an integrative and collaborative nature, e.g. actor involvement and willingness to share data. These business relations challenges belong to the organizational dimension but have mutual impacts with the people skills dimension in form of interaction and coordination requirements. This process also requires a high degree of system integration which is mutually related to the coordination challenges. One promising attempt to handle these issues are the establishment of the service coordinator role, that acts as an integrator for the entire process.

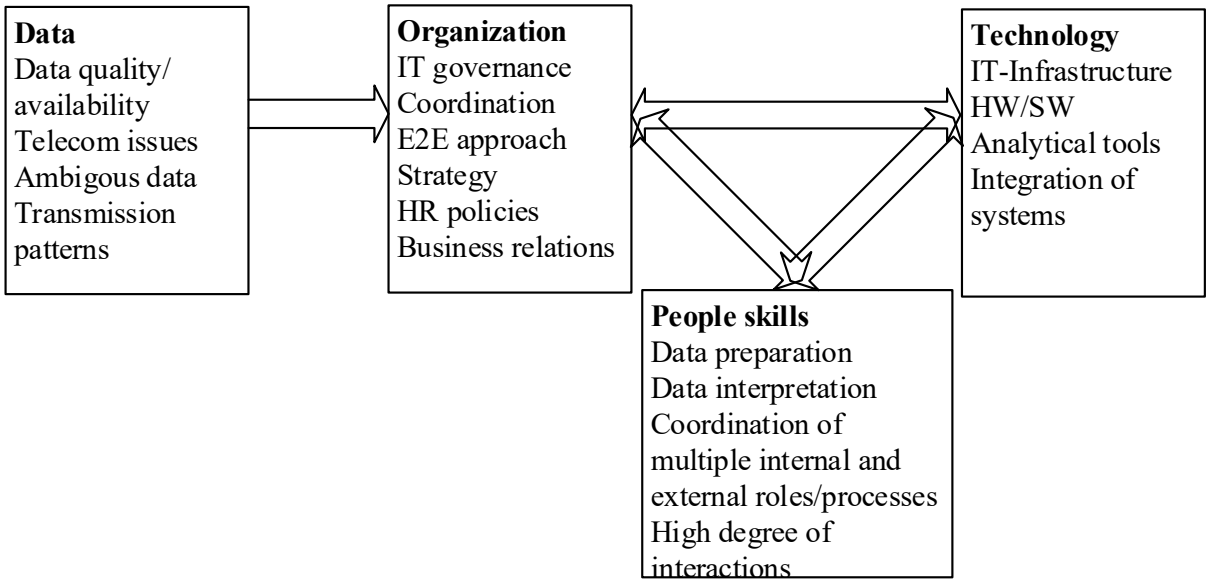


Figure 4.5: Challenges and interdependencies regarding vehicle monitoring

For the last process in this study (vehicle monitoring, the main challenges are in the organizational and people skills dimensions. The starting point, also for this process, is the data-driven challenges. Similar to the pre-planning process, the vehicle monitoring process is also considered to be complex from a process point of view, due to its interactions between actors and activities. Therefore, the main challenges and interactions of the challenges are of the same nature as for the pre-planning process. However, some differences are present. Due

to the urgent nature of this process the timeliness in a large number of interactions creates unique challenges regarding people skills and organization.

4.2.2 Common challenges across the processes

A cross-case analysis was also performed. This analysis resulted in common challenges existing in all three processes (Table 4.1).). A common result within all three processes is that the trigger of the identified challenges is the data itself. One example is that poor data quality creates challenges for people regarding data preparation (e.g. data correction, imputation and understanding). Empirical data indicates that between 80 and 90 % of a data scientist working time is spent on data preparation issues. Other common challenges are technology related and regards capabilities for IT infrastructure, data storage and processing, e.g. analytical tools. Common organizational challenges are, for example, increased capabilities regarding data governance and coordination of big data analytical projects. Hence an E2E approach and changes in the company's strategic processes causes new type of challenges. Regarding the skills required to develop and use big data analytical applications, common challenges are to achieve an efficient, collaborative working environment between people with different competences, e.g. data scientists, data engineers and domain experts.

Table 4.1: Common challenges for all three application areas

Challenge dimension			
Data	Technology	Organization	People skills
Data quality Data availability	IT infrastructure HW/SW Analytical tools	IT/Data governance Coordination of projects E2E approach Strategy HR policies Data driven culture	Data science Combining analytical/domain knowledge

4.3 Paper 3 (Causal-based spare parts forecasting exploiting product-in-use data in a heavy vehicle aftermarket context)

The aim of this paper is not to contribute to new theory, but rather to bring existing methods from statistics/machine learning and test these methods into a real-world setting for spare parts forecasting by using new types of data.

Drawing on results from study 1 and with reference to literature, the aim of this paper is to test the effect of explanatory factors together with demand history in a multiple regression analysis, known as an ARX model (autoregression with exogenous variables). A dataset of ~ 700 spare parts, using fault codes, mileage, operating time and demand history as regressors contributed to improved forecast accuracy on three different phases in the life cycle (after one year, after three years and after five years) compared to traditional time series forecasting methods. The analysis was performed as a stepwise regression which started with using historical demand and fault code 1 as regressors, after which an additional regressor was added, until all six regressors were tested. Forecast errors are measured as Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The result regarding forecast error divided into age-based classes (MAPE) is showed in Figure 4.6.

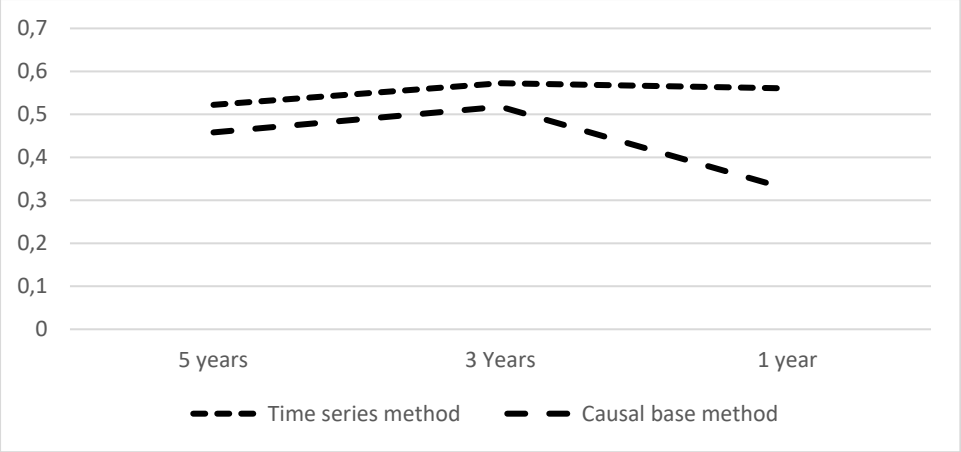


Figure 4.6: Forecast accuracy based on MAPE (age classes)

Forecast error is reduced for every added explanatory factor when it is measured as MAE, for all three phases of the early life cycle. When measuring the forecast error with MAPE, the same effect is seen for the first year. However, after three and five years of history, the forecast error was reduced until the last fault code was added (fault code 3), although no increased performance was seen by adding average distance and time for this metric. A plausible explanation for the different outcome due to metrics is that MAE is more sensitive to large deviations, for example more common for high-frequency spare parts. All in all, it is fair to say that the model using demand and fault codes outperforms the traditional time series methods.

An analysis of the model performance broken down by demand frequency class was carried out as well. This result is shown in Figure 4.6 and shows the values from the best time series method for each frequency class, as well as the best performing causal-based method. From this point of view too, the multiple regression model outperforms the traditional methods, with one exception. For slow-moving spare parts (average demand per period < 5) both single and double exponential methods outperform the regression method with explanatory factors.

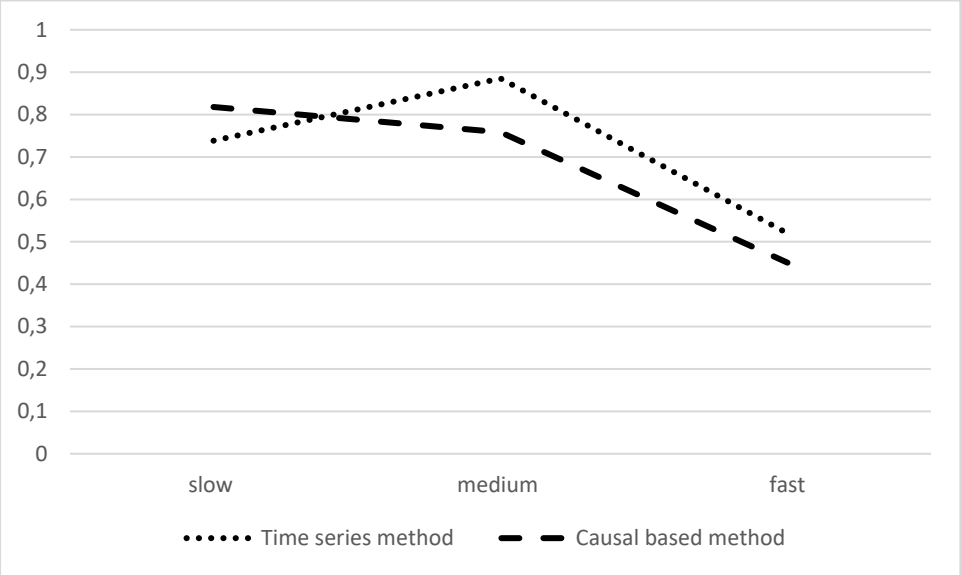


Figure 4.7: Forecast accuracy based on MAPE (frequency classes)

Medium frequency spare parts are defined by an average demand per period between 5 and 25 and fast by an average demand per period > 25. As shown in Figure 4.7, the largest improvement concerns fast-moving spare parts.

The overall conclusion from this study is the improved forecasting accuracy obtained by applying regression analysis with product-in-use data as explanatory factors, in early life cycle phases, mainly for medium/high frequency parts.

5 Discussion

This section addresses a discussion on how the results contribute to the overall purpose of the research presented in this thesis, as well as to the theoretical framework.

To do so, the purpose is stated again: to investigate how product-in-use data can be used in, and improve the performance of, demand planning processes for aftermarket services, by describing and explaining related performance effects and challenges. First the results from each study are discussed with regard to the overall purpose and literature, followed by a discussion of the research limitations and future research.

5.1 Contribution to the purpose

The first main contribution to the purpose describes and proposes which causal-based demand planning methods are relevant in an aftermarket demand planning context. The majority of the research regarding spare parts forecasting and demand planning focuses on intermittent and low-frequency demand. Reliability-based methods (e.g. Cavalieri et al., 2008) and CBM-based methods (e.g. Jardine et al., 2006) predict the failure of single components in single machines. Regression-based methods are applied at aggregated demand level, although there are very few papers with real-world data. Furthermore, these are limited in terms of the data used as explanatory variables. Moreover, the existing literature is mostly on a conceptual level, proposing forecasting methods, but without suggestions regarding which spare parts classes to apply them to. Since the spare parts demand planning literature regarding causal based methods lacks empirical based results, the research in this thesis fills an important gap regarding which type of data to be used for the different methods, as well as for which type of spare parts each method should be used. Furthermore, the proposed methods can both improve the forecast accuracy by the proposed causal-based forecasting methods or, for spare parts where condition based maintenance are feasible (see figure 4.1), reduce the demand uncertainty and predict the demand as known demand, which will also assist the vehicle owner due to fewer unplanned stops and the service provider with improved workshop planning. In contrast to the existing research regarding casual-based demand planning in an aftermarket context, these results take all relevant demand drivers into account. This problem is addressed for example by Van der Auweraer et al. (2019).

A second main result, from a process perspective, is the potentially improved demand planning performance of utilizing the product-in-use data in three process steps: 1) automatic forecast calculation, 2) increasing the share of demand that bypasses the forecast calculation and is handled as known demand in the DRP calculation and, 3) usage of the product-in-use data in the judgemental forecasting activity. Previous research regarding this topic has mainly been either high-level conceptual studies (e.g. Cohen et al., 2006) or focused on very specific conditions and are not considering the full demand planning process or all potential explanatory factors (Van der Auweraer et al., 2019). This thesis generates an overall understanding of how causal-based methods in demand planning in an aftermarket demand planning process can be used, as well as potential results.

The third major contribution describes how different causal-based and time series methods can be combined, e.g. at the intersection of life cycle phases or between maturity steps when implementing causal-based forecasting. This is in line with previously findings, e.g. Makridakis et al. (2008), although results in this thesis propose combination of causal-based methods and traditional, time series methods. Multi-method forecasts may generate higher forecast accuracy than single-method forecasts (De Menezes et al., 2000), and an appropriate approach would be to combine any of the causal-based methods with time-series methods. As different causal-based methods are likely to have different impact on spare parts in different

life cycle phases, it is proposed to combine methods when a spare part is in the transition between two life cycle phases. Mixed methods could also be beneficial in an early implementation phase of causal-based forecasting by mixing automatic causal-based methods with judgmental methods until the causal methods are proven. By doing so, some of the challenges could be mitigated. Hence, this research extends the scope of potential forecasting methods that could be combined and address other usage areas where it is applicable, e.g. in intersections of life cycle phases or maturation phases regarding implementation of causal based forecasting.

Challenges in using big data for supply chain planning applications, in terms of capability gaps, are identified and grouped into their respective challenge dimension. Furthermore, there are challenges in relation to development vs implementation phases, the interdependency between challenges are analyzed, as well as their relation to process and data complexity. This has been reported in existing literature by e.g. McAfee and Brynjolfsson (2012)

The fourth main result regards the empirical results regarding big data analytical challenges in an aftermarket context and how these challenges are inter-related. The main results regarding challenges from using product-in-use data in an aftermarket process relates to data quality, which stems from a complex systems environment (3V) and the process complexity in the various supply chain processes. The above statements suggest that the driver of all challenges is usage of the data produced by on-board sensors. The data itself contains challenges related to the 3 Vs (Hazen et al., 2014), although the data affects actors in the supply chain from a technology point of view (e.g. requirements for software and hardware, analytical tools etc.) and from a competence point of view (for example combining data science skills and supply chain domain knowledge). Organizational challenges also appear as a consequence of the data. Overall, the main challenge (broken down in Paper 2) for a company concerns how to become data driven and create such a company culture (McAfee and Brynjolfsson, 2012).

The fifth main contribution is the results regarding challenges from a capability perspective and create knowledge on what challenges needs to be overcome to be able to achieve the potential results from a causal-based aftermarket demand planning process. That is to say, without strong organizational, technology and human capabilities the advantage of using the product-in-use data will not be realized. Previous research has either been at a rather high level (e.g. Kache and Seuring, 2015 or McAfee and Brynjolfson, 2012) or focused on specific challenges, for example Hazen et al. (2014), who focused on data quality, Russom (2011) or Richey et al. (2016), who adressed the 3Vs or Schoenherr and Speier-Pero (2015) whose main focus was on people and organization. The research in this thesis put the challenges both in an aftermarket demand planning context, but also identifies common and unique challenges per process.

The sixth major issue regards the confirmation of some of the potential performance effects proposed by study 1. Adopting a regression model with historical demand as well as fault codes, driving distance and operating hours as independent variables, show an improved forecast accuracy for spare parts in early life cycle phase (between one and five years since launch with medium and high demand frequency. Compared to traditional time series methods, the causal-based regression methods perform better with a lower forecast error.

Furthermore, some of the challenges regarding data (handling large data volumes with different periodicity and varying data types) were experienced and mitigated during the study. In consideration of the conceptual model in Section 1.7, this study therefore confirms and reinforces the findings from the previous studies and tests some of the previous findings with data from samples from a real-world context. In addition, the findings from study 3 reinforce the contribution towards the overall purpose.

5.2 Synthesis of the results

As a basis for the discussion of the synthesized result, the results are presented here according to their relationship with the conceptual model first presented in section 1.6. (Figure 5.1).

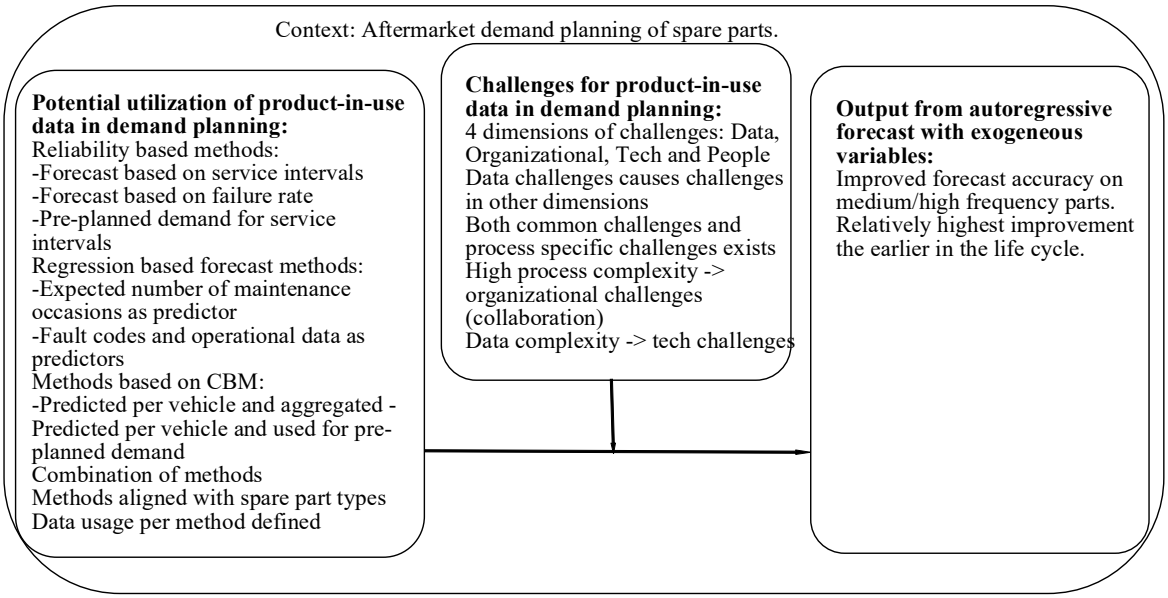


Figure 5.1: Conceptual model with empirical results

The context of this thesis is aftermarket demand planning and the analysis and results are based on the case company (Volvo Group Truck Operations) which produce and maintains heavy vehicles for industrial applications. Aftermarket demand planning in general, has some specific contextual attributes described in existing literature (e.g. intermittent and/or low frequency demand, underlying reasons are the need for maintenance and repair and usage specific attributes). This is also valid for the studies built upon Volvo specific data, although some specific attributes in this case are: a large number of configurations of the machines and vehicles, a large variation of the application of the machines and vehicles (e.g. mining, construction, long haul and distribution), different conditions in different installed base regions and the fact that most of the installed base is moving. The potential causal based demand planning methods are stated on the left-hand side of the model. These methods can also be of value for other industries. A validation workshop with a car manufacturer and other discussions/meetings with representatives from other industries (warehouse equipment and medical equipment) confirmed the proposed methods. Even though the proposed methods are transferable for most aftermarkets of different industries, different types of products need to be analyzed regarding how these methods can be used regarding the contextual variables, e.g. life cycle, demand variability and criticality.

Beyond the contextual factors, it is also important to consider the business environment. For example, the case company operates in a B2B environment with different requirements then, for example, consumer goods. These requirements, e.g. high demand for uptime, varying length of the life cycle and involvement of many actors, creates unique challenges, e.g. regarding interaction and collaboration, while other challenges are common, for example the data and technology related challenges.

For these proposed methods to be effective, i.e. to be able to produce improved demand planning accuracy, the corresponding challenges need to be mitigated, since the challenges are dependent on both data and process complexity, as well as context dependent. The result (on

the right-hand side of Figure 5.1) is very much a consequence on how the challenges are met for each specific intervention. In this research, it has been confirmed (study 3) that a positive outcome has been achieved for a subset of the potential causal-based methods (regression-based forecasting with fault codes and sensor data). To achieve that positive outcome, data, technology, organizational and people skills challenges have been faced. With extended data, technology, organizational and people skills capabilities, an important reflection is that the result would have been better, it would have been easier and faster to produce the result, and it would have been easier to implement new methods in the operational process.

6 Concluding remarks

This chapter contains the conclusion of the research and relates back to the research questions, as well as sections describing, theoretical contribution and practical implications. The chapter ends with limitations and further research.

6.1 Conclusions

This thesis has studied the potential usage of product-in-use data in an aftermarket demand planning context with related challenges and evaluation of the performance. Three studies were performed, each one guided by a research question.

Research question 1: How can product-in-use data be used in (1A), and improve (1B), the performance of the demand planning process for aftermarket services? To answer the first part (1A), three overall causal based demand planning methods were identified in the literature study and eight detailed causal-based demand planning methods, with connected requirements of specific data elements and for which type of spare parts each given method is feasible, were proposed (1B). This classification encompasses contextual factors, such as demand frequency and variability, life cycle phases and spare parts usage (maintenance or repair).

Regarding RQ1, eight causal-based forecasting methods and are proposed, with aligned spare part types (as illustrated in Figure 4.1): The two first are reliability prediction methods based on service interval and failure rate respectively. Methods three and four are regression methods based on service intervals and fault codes, sensor data and operational data, respectively. The fifth method on aggregated forecast (condition-based management) predicts the need for maintenance or future repair for individual components in individual machines/vehicles by compare specific values from sensors and apply diagnostic logic. The sixth and seventh methods are pre-planning (condition-based management and service intervals, respectively), where maintenance is scheduled in advance. The eighth method uses any method to generate alerts

Related to RQ2: What are the challenges regarding the use of product-in-use data in supply chain planning for aftermarket services? Three processes were studied with regards to the data-driven challenges; causal-based forecasting, pre-planned maintenance operations and vehicle monitoring using product-in-use data. Unique and common challenges for using and implementing causal based processes were studied and resulted in challenges in four dimensions (data, organization, technology and people skills). Moreover, the interrelations between the challenges were studied and as a result of that, it was found that the data is the trigger of all other challenges, as well as the fact that the process complexity has a large impact (mostly on the organization and people skills dimensions). Data complexity has a large impact on the technology dimension

The findings consist of common challenges across all three studied processes and context dependent challenges, i.e. they are unique for a specific process. Common data related challenges are data quality and data availability, stemming from the data that originates from the vehicles ECU's. This data cause challenges in the technology dimension, e.g. regarding hardware for storage and processing and for software regarding data preparation and analytics. Furthermore, these challenges cause challenges in the people skills dimension regarding and organizational related challenges, e.g. regarding data governance, coordination of projects and to adopt an E2E-approach. Although, the most important challenge with regards to organization belongs to the creation of a data-driven culture.

The performance effect of a subset of the proposed methods from study 1 was carried out. This quantitative study used an autoregressive method with product-in-use data as exogenous

variables and aimed at answering the third research question, RQ3: What are the effects on forecast accuracy when applying a causal-based regression methodology for spare part forecasting using explanatory factors as exogeneous variables? This study resulted in a confirmation of the proposition that regression-based forecasting methods using fault codes and sensor data have a positive outcome, i.e. increased forecast accuracy, for spare parts in early life cycle phases combined with medium/high demand frequency.

Since this research area is understudied, both from a detailed process perspective using empirical data in an empirical context, but also from the perspective of combining context, interventions (proposed methods), mechanisms (challenges) and outcome (the so called CIMO model), This approach takes an approach of all aspects of a successful usage/implementation of causal based demand planning methods for aftermarket services.

6.2 Theoretical contribution

Due to the emerging nature of the research problem, knowledge was borrowed from several fields, including big data analytics, traditional and causal based forecasting and demand planning, maintenance literature and literature regarding aftermarket challenges and opportunities.

Previous research has either focused on generic explorations of using big data on rather high-level supply chain problems, and there has been little research with empirical data with real-world examples. Some examples of causal-based forecasting have been carried out, but almost exclusively on limited data and/or considering very specific cases, e.g. predicting failures of certain components. The research in this thesis contributes to theory by adopting an early approach to take a process focus and combine explorative, explanatory and confirmatory approaches to deliver insights on this emerging research field. Moreover, the process focus contributes to how to both apply and develop methods for aftermarket causal-based demand planning methods. Furthermore, this research contributes to the literature of demand planning methods in an aftermarket context by utilization of product-in-use data, produced and communicated by the finished product.

6.3 Practical implications

Regarding practical contribution, managers in aftermarket supply chain planning can use the proposed methods from this research to enhance their aftermarket planning performance and connect the methods to the spare parts classes regarding life cycle phases, demand characteristics and criticality (or spare part usage). This research can also provide insights and guidance on how to build up analytics around product-in-use data, as well as what data to explore for potential results. At the case company, the results from this thesis has been used as guidance for development of causal based demand planning and several projects exists to further enhance this process. Other managerial implications are awareness of the complexity and the challenges of development and implementation of such data-driven processes. This awareness can be used for planning and prioritization of projects involving product-in use data by evaluating the potential effects together with the data and process complexity and, by that, foresee hinders early in the development phase and to avoid problems before implementation.

6.4 Limitations and Further research

Because there is only one case company, the generalizability of the results is somewhat limited. However, measures were taken in each study to address these limitations. For example, the authors' a priori knowledge and experience together with opportunities for informal

discussions with representatives of other companies in various industries to a great extent confirm the findings.

Several options for future research may be relevant based on the research presented in this thesis. Two main strands for further research are 1) to expand the research to other contexts and 2) further development regarding methods for causal-based demand planning. Regarding the context, a relevant consideration would be to study other processes (e.g. inventory planning, distribution planning or sales and operations planning) for the aftermarket, another is to study potential performance effects for various industries and, a third direction for further research could be to study how other aftermarket supply chains can benefit from using product-in-use data, e.g. consumer supply chains based on data from external sources such as social media. Regarding development of analytical demand planning methods, two main angles can be proposed. One is to expand the scope and study the contribution of other types of data, e.g. external data such as economic factors, weather and road conditions and social media data. Another potential research idea is to quantitatively study improved machine learning methods for various contextual factors, e.g. life cycle phases, demand attributes, regional aspects. Other possible research topics could focus on what the data and analytical capabilities can bring by study servitization opportunities for the aftermarket.

7 References

Andersson, J. and 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 & Logistics Management,* Vol. 48, No. 5, pp. 524-544.

Arribas-Bel, D. (2014), "Accidental, open and everywhere: Emerging data sources for the understanding of cities", *Applied Geography,* Vol. 49, No., pp. 45-53.

Bacchetti, A. and Saccani, N. (2012), "Spare parts classification and demand forecasting for stock control: Investigating the gap between research and practice", *Omega,* Vol. 40, No. 6, pp. 722-737.

Bartezzaghi, E. and Kalchschmidt, M. (2011), "The Impact of Aggregation Level on Lumpy Demand Management" in Altay.N, Litteral.A (Eds.) , *Service Parts Management*, Springer-Verlag,London(2011), pp. 89-104

Boylan, J. E. and Syntetos, A. A. (2010), "Spare parts management: A review of forecasting research and extensions", *IMA Journal of Management Mathematics,* Vol. 21, No. 3, pp. 227-237.

Bryman, A. and Bell, E. (2015), *Business research methods*, Oxford University Press.

Cavalieri, S., Garetti, M., Macchi, M. and Pinto, R. (2008), "A decision-making framework for managing maintenance spare parts", *Production Planning & Control,* Vol. 19, No. 4, pp. 379-396.

Cohen, M. A., Agrawal, N. and Agrawal, V. (2006), "Winning in the aftermarket", *Harvard business review,* Vol. 84, No. 5, pp. 129.

Coughlan, P. and Coghlan, D. (2002), "Action research for operations management", *International Journal of Operations & Production Management,* Vol. 22, No. 2, pp. 220-240.

Croston, J. D. (1972), "Forecasting and Stock Control for Intermittent Demands", *Operational Research Quarterly (1970-1977),* Vol. 23, No. 3, pp. 289-303.

Davenport, T. H. and Patil, D. (2012), "Data scientist", *Harvard business review,* Vol. 90, No. 5, pp. 70-76.

De Menezes, L. M., Bunn, D. W. and Taylor, J. W. (2000), "Review of guidelines for the use of combined forecasts", *European Journal of Operational Research,* Vol. 120, No. 1, pp. 190-204.

Dekker, R., Pinçe, Ç., Zuidwijk, R. and 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,* Vol. 143, No. 2, pp. 536-545.

Dickersbach, J. T., Passon, M. F.,(2015), *Service Parts Planning with SAP SCM: Processes, Structures, and Functions*. SpringerVerlag Berlin Heidelberg.

Djurdjanovic, D., Lee, J. and Ni, J. (2003), "Watchdog Agent—an infotronics-based prognostics approach for product performance degradation assessment and prediction", *Advanced Engineering Informatics,* Vol. 17, No. 3, pp. 109-125.

Dubois, A. and Gadde, L.-E. (2002), "Systematic combining: an abductive approach to case research", *Journal of Business Research,* Vol., No. 7, pp. 553.

Efron, B. (1979), "Bootstrap Methods: Another Look at the Jackknife", *The Annals of Statistics,* Vol. 7, No. 1, pp. 1-26.

Flick, U. (2014). *The SAGE handbook of qualitative data analysis*, SAGE.

Fritzsche, R., Gupta, J. N. D. and Lasch, R. (2014), "Optimal prognostic distance to minimize total maintenance cost: The case of the airline industry", *International Journal of Production Economics,* Vol. 151, No., pp. 76-88.

Frowein, B., Lang, N., Schmieg, F. and Sticher, G. (2014). “Returning to Growth: A Look at the European Automotive Aftermarket”, Marktstudie, The Boston Consulting Group, Berlin and München

Gebauer, H., Tennstedt, F., Elsässer, S. and Betke, R. (2010), "The Aftermarket in the Automotive Industry–How to Optimize Aftermarket Performance in Established and Emerging Markets", https://www.capgemini.com/wp-content/uploads/2017/07/tl_The_Aftermarket_in_the_Automotive_Industry.pdf, accessed March 28, 2018.

Halldórsson, Á. and Aastrup, J. (2003), "Quality criteria for qualitative inquiries in logistics", *European Journal of Operational Research,* Vol. 144, No. 2, pp. 321-332.

Hazen, B. T., Boone, C. A., Ezell, J. D. and Jones-Farmer, L. A. (2014), "Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications", *International Journal of Production Economics,* Vol. 154, pp. 72-80.

Hellingrath, B. and Cordes, A.-K. (2014), "Conceptual approach for integrating condition monitoring information and spare parts forecasting methods", *Production & Manufacturing Research,* Vol. 2, No. 1, pp. 725-737.

Holmqvist, M. and Pessi, K. (2006), "Agility through scenario development and continuous implementation: a global aftermarket logistics case", *European Journal of Information Systems,* Vol. 15, No. 2, pp. 146-158.

Holmström, J., Ketokivi, M. and Hameri, A. P. (2009), "Bridging practice and theory: A design science approach", *Decision Sciences,* Vol. 40, No. 1, pp. 65-87.

Huiskonen, J. (2001), "Maintenance spare parts logistics: Special characteristics and strategic choices", *International Journal of Production Economics,* Vol. 71, No. 1–3, pp. 125-133.

Hussain, R., Lockett, H. and Vasantha, G. V. A. J. (2012), "A framework to inform PSS Conceptual Design by using system-in-use data", *Computers in Industry,* Vol. 63, No. 4, pp. 319-327.

Jardine, A. K., Lin, D. and Banjevic, D. (2006), "A review on machinery diagnostics and prognostics implementing condition-based maintenance", *Mechanical systems and signal processing,* Vol. 20, No. 7, pp. 1483-1510.

Johnson, M. and Mena, C. (2008), "Supply chain management for servitised products: A multi-industry case study", *International Journal of Production Economics,* Vol. 114, No. 1, pp. 27-39.

Jonsson, P. and Holmstrom, J, (2016), "Future of supply chain planning: closing the gaps between practice and promise", *International Journal of Physical Distribution & Logistics Management,* Vol. 46, No. 1, pp. 62-81.

Jonsson, P., Ivert, L. K (2015), "Improving performance with sophisticated master production scheduling", *International Journal of Production Economics,* Vol. 168, pp. 118-130.

Jouni, P., Huiskonen, J. and Pirttilä, T, (2011), "Improving global spare parts distribution chain performance through part categorization: A case study", *International Journal of Production Economics,* Vol. 133, No. 1, pp. 164-171.

Kache, F. and Seuring, S. (2017), "Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management", *International Journal of Operations & Production Management,* Vol. 37, No. 1, pp. 10-36.

Kennedy, W., Patterson, J. W. and Fredendall, L. D. (2002), "An overview of recent literature on spare parts inventories", *International Journal of production economics,* Vol. 76, No. 2, pp. 201-215.

Ketokivi, M. and Choi, T. (2014), "Renaissance of case research as a scientific method", *Journal of Operations Management,* Vol. 32, No. 5, pp. 232-240.

Kilger, C. and Wagner, M. (2008). “Demand planning”, IN: Stadtler H., Kilger C. (eds), *Supply Chain Management and Advanced Planning*, Springer, Berlin, Heidelberg, pp 133-160.

Lavalle, S., Lesser, E., Shockley, R., Hopkins, M. S. and Kruschwitz, N. (2011), "Big Data, Analytics and the Path From Insights to Value", *MIT Sloan Management Review,* Vol. 52, No. 2, pp. 21-32.

Lee, H. L., Padmanabhan, V. and Whang, S. (1997), "The bullwhip effect in supply chains", MIT *Sloan management review,* Vol. 38, No. 3, pp. 93.

Makridakis, S., Hyndman, R. J. and Wheelwright, S. C. (1998), *Forecasting: Methods and Applications,* John Wiley & Sons, Inc.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Byers, A. H. (2011), "Big data: The next frontier for innovation, competition, and productivity”, http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation, Accessed September 14, 2017.

Markopoulos, J. (2012). “Ways The Industrial Internet Will Change Manufacturing”, Forbes, https://www.forbes.com/sites/ciocentral/2012/11/29/5-ways-the-industrial-internet-will-change-manufacturing/.

Maxwell, J. A. (2013). *Qualitative research design : an interactive approach*. Thousand Oaks, SAGE Publications.

McAfee, A. and Brynjolfsson, E. (2012), "Big Data: The Management Revolution. (cover story)", *Harvard business review,* Vol. 90, No. 10, pp. 60-68.

McAlexander, J. H., Schouten, J. W. and Koenig, H. F. J. (2002), "Building Brand Community", *Journal of Marketing,* Vol. 66, No. 1.

Näslund, D., Kale, R. and Paulraj, A. (2010), "Action research in supply chain management—a framework for relevant and rigorous research", *Journal Of Business Logistics,* Vol. 31, No. 2, pp. 331-355.

Ozgur, C., Kleckner, M. and Li, Y. (2015), "Selection of Statistical Software for Solving Big Data Problems", *SAGE Open,* Vol. 5, No. 2.

Porter, M. E. and Heppelmann, J. E. (2014), "How smart, connected products are transforming competition", *Harvard business review,* Vol. 92, No. 11, pp. 64-88.

Richey Jr, R. G., Morgan, T. R., Lindsey-Hall, K. and Adams, F. G. (2016), "A global exploration of big data in the supply chain", *International Journal of Physical Distribution & Logistics Management,* Vol. 46, No. 8, pp. 710-739.

Romeijnders, W., Teunter, R. and van Jaarsveld, W. (2012), "A two-step method for forecasting spare parts demand using information on component repairs", *European Journal of Operational Research,* Vol. 220, No. 2, pp. 386-393.

Russom, P. (2011), "Big data analytics", *TDWI best practices report, fourth quarter,* Vol. 19, No.4, pp. 1-34.

Sanders, N. R. (2014). *Big data driven supply chain management: A framework for implementing analytics and turning information into intelligence*, Pearson Education.

Schoenherr, T. and Speier-Pero, C. (2015), "Data Science, Predictive Analytics, and Big Data in Supply Chain Management: Current State and Future Potential", *Journal Of Business Logistics,* Vol. 36, No. 1, pp. 120-132.

Spens, K. M. and Kovács, G. (2006), "A content analysis of research approaches in logistics research", *International Journal of Physical Distribution and Logistics Management,* Vol. 36, No. 5, pp. 374-390.

Stadtler, H. and Kilger, C. (2008), "Supply chain management and advanced planning", *Concepts, Models, Software and Case Studies,* Springer Verlag, Berlin.

Syntetos, A. A., Babai, Z., Boylan, J. E., Kolassa, S. and Nikolopoulos, K. (2016), "Supply chain forecasting: Theory, practice, their gap and the future", *European Journal of Operational Research,* Vol. 252, No. 1, pp. 1-26.

Syntetos, A. A. and Boylan, J. E. (2005), "The accuracy of intermittent demand estimates", *International Journal of Forecasting,* Vol. 21, No. 2, pp. 303-314.

Waller, M. A. and Fawcett, S. E. (2013), "Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management", *Journal o Business Logistics,* Vol. 34, No. 2, pp. 77-84.

Van der Auweraer, S., Boute, R. N. and Syntetos, A. A. (2019), "Forecasting spare part demand with installed base information: A review", *International Journal of Forecasting,* Vol. 35, No. 1, pp. 181-196.

Wang, W. and Syntetos, A. A. (2011), "Spare parts demand: Linking forecasting to equipment maintenance", *Transportation Research Part E: Logistics and Transportation Review,* Vol. 47, No. 6, pp. 1194-1209.

Willemain, T. R., Smart, C. N. and Schwarz, H. F. (2004), "A new approach to forecasting intermittent demand for service parts inventories", *International Journal of Forecasting,* Vol. 20, No. 3, pp. 375-387.

Wise, R. and Baumgartner, P. (1999), "Go Downstream: The New Profit Imperative in Manufacturing", *Harvard business review,* Vol. 77, No. 5, pp. 133-141.

Vollmann, T. E. (2005). *Manufacturing planning and control systems for supply chain management*, McGraw-Hill, New York.

Voss, C., Tsikriktsis, N. and Frohlich, M. (2002), "Case research in operations management", *International Journal of Operations & Production Management,* Vol. 22, No. 2, pp. 195-219.

Xiao, T. and Yang, D. (2008), "Price and service competition of supply chains with risk-averse retailers under demand uncertainty", *International Journal of Production Economics,* Vol. 114, No. 1, pp. 187-200.

Yang, Z. M., Djurdjanovic, D. and Ni, J. (2008), "Maintenance scheduling in manufacturing systems based on predicted machine degradation", *Journal of intelligent manufacturing,* Vol. 19, No. 1, pp. 87-98.

Yin, R. K. (2014). *Case study research : design and methods*, SAGE, London

Zhong, R. Y., Newman, S. T., Huang, G. Q. and Lan, S. (2016), "Big Data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives", *Computers & Industrial Engineering,* Vol. 101, No., pp. 572-591.

Öhman, M., Finne, M. and Holmström, J. (2015), "Measuring service outcomes for adaptive preventive maintenance", *International Journal of Production Economics*, Vol. 170, Part B, No., pp. 457-467.