Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Engineering
Department of Mechanical and Industrial Engineering

Torkild Sandnes Grøstad
Eivind Stangebye Larsen

Machine Learning-based Time Series Forecasting for Dynamic Reorder Points

A Case Study for Logistics Center Helse Midt-Norge

Master's thesis in Engineering & ICT
Supervisor: Fabio Sgarbossa
Co-supervisor: Aili Biriita Bertnum
June 2023

NTNU
Norwegian University of Science and Technology

ST. OLAVS HOSPITAL
UNIVERSITETSSYKEHUSET I TRONDHEIM

Torkild Sandnes Grøstad
Eivind Stangebye Larsen

Machine Learning-based Time Series Forecasting for Dynamic Reorder Points

A Case Study for Logistics Center Helse Midt-Norge

Master's thesis in Engineering & ICT
Supervisor: Fabio Sgarbossa
Co-supervisor: Aili Biriita Bertnum
June 2023

Norwegian University of Science and Technology
Faculty of Engineering
Department of Mechanical and Industrial Engineering

NTNU
Norwegian University of
Science and Technology

Preface

This paper is written as a master thesis for the Department of Mechanical and Industrial Engineering at the Norwegian University of Science and Technology, in collaboration with Logistics Center Helse Midt-Norge (LC HMN). The paper was written in the spring of 2023.

We would like to thank our supervisor Professor Fabio Sgarbossa for his input and encouragement throughout the thesis. We will also thank the section leader of Logistics Center Helse Midt-Norge, Aili Biriita Bertnum, for supporting us with the knowledge and data necessary to write this thesis in collaboration with Logistics Center Helse Midt-Norge.

Abstract

The demand for health care in Norway is increasing due to demographic changes, including a growing and aging population. This increase in demand is expected to result in higher healthcare spending and increase the pressure on the Norwegian economy. To address these challenges, the Norwegian government is seeking to utilize digital transformation and big data to improve the efficiency of the public sector, including the healthcare industry. Municipal executives and IT professionals believe that the use of machine learning and other technological solutions can help streamline tasks and improve the quality of processes in the healthcare sector. However, many municipalities have not yet begun implementing IT projects that involve machine learning. In general, IT investments have been reported to improve process quality and efficiency in the public sector, but there is still room for improvement using data analytics. In light of the background, the objective of this study was to *"Make a contribution to the current literature by investigating the effects of implementing demand forecasting and a dynamic reorder point policy for Logistics Center Helse Midt-Norge"*. The following research questions were introduced to guide the study:

- RQ1: What is the state-of-the-art within demand forecasting for inventory management?
- RQ2: How can the inventory be classified?
- RQ3: How can the AS-IS fixed reorder point be improved through a dynamic reorder point?
- RQ4: What is the impact of implementing advanced forecasting methods for the dynamic reorder point?

The research methodology in this study integrated various methods, including a systematic literature review, an empirical case study, data analysis, and simulation. The systematic literature review is employed to investigate and address RQ1. On the other hand, RQ2, RQ3, and RQ4 were addressed through an empirical case study. Within the case study, RQ2 utilized data analysis, while RQ3 and RQ4 were addressed using simulation methods. The results show that by categorizing Stock Keeping Unit (SKU) based on demand variability, valuable insights can be gained to enhance inventory management operations. The implementation of dynamic reorder point policies proves to be highly effective in reducing average inventory while maintaining a high service level, resulting in an impressive 42% inventory reduction while ensuring a service level of 98.9%. However, it was observed that implementing advanced forecasting methods for the dynamic reorder point policy does not necessarily outperform basic demand forecasting models in terms of average inventory and service level. This underscores the importance of considering the trade-off between complexity and performance when selecting forecasting methods for the dynamic reorder point policy.

Keywords— hospital warehouse, inventory management, dynamic reorder point, demand forecasting, machine learning

Sammendrag

Etterspørselen etter helsetjenester i Norge øker på grunn av demografiske endringer, inkludert en voksende og aldrende befolkning. Denne økningen i etterspørsel forventes å føre til høyere helseutgifter og øke presset på norsk økonomi. For å takle disse utfordringene, ønsker den norske regjeringen å utnytte digital transformasjon og stordata (Big Data) for å forbedre effektiviteten i offentlig sektor og helseindustrien. Kommunale ledere og fagfolk innen IT-bransjen mener at bruk av maskinlæring og andre teknologiske løsninger kan bidra til å effektivisere oppgaver og forbedre kvaliteten på prosesser i helsesektoren. Imidlertid har mange kommuner ennå ikke startet implementeringen av IT-prosjekter som involverer maskinlæring. Generelt sett har det blitt rapportert at IT-investeringer forbedrer prosesskvalitet og effektivitet i offentlig sektor, men det er fortsatt rom for forbedring ved bruk av dataanalyse. Med bakgrunn i dette var målet med denne oppgaven å *"Bidra til den nåværende litteraturen ved å undersøke effektene av å implementere etterspørselsprognoser og dynamisk bestillingspunkt for Logistikksenteret Helse Midt-Norge"*. Følgende forskningsspørsmål ble introdusert for å veilede studien:

- RQ1: Hva er "State-Of-The-Art" for etterspørselsprognoser innen lagerstyring?
- RQ2: Hvordan kan lageret klassifiseres?
- RQ3: Hvordan kan den nåværende faste bestillingspunktet forbedres ved bruk av dynamisk bestillingspunkt?
- RQ4: Hva er virkningen av å implementere avanserte prognosemetoder for det dynamiske bestillingspunktet?

Forskningsmetodikken i denne studien benyttet ulike metoder, inkludert en systematisk gjennomgang av litteratur, et empirisk case-studie, dataanalyse og simulering. Den systematiske gjennomgangen av litteratur ble brukt for å undersøke og svare på det første forskningsspørsmålet. Resterende forskiningspørsmål (RQ2, RQ3 og RQ4) ble besvart gjennom en empirisk case-studie. I det empiriske case-studiet benyttet det andre forskningsspørsmålet dataanalyse, mens det tredje og fjerde forskningsspørsmålet ble besvart ved hjelp av simulering. Resultatene viser at ved å kategorisere varene til Logistikksenteret Helse Midt-Norge basert på etterspørselens variabilitet, kan verdifulle innsikter oppnås for å forbedre lagerstyringsoperasjoner. Implementeringen av dynamiske bestillingspunkter viser seg å være svært effektivt for å redusere gjennomsnittlig lagerbeholdning samtidig som en høy servicenivå opprettholdes, noe som resulterer i en imponerende 42% reduksjon i lagerbeholdning og et servicenivået på 98,9%. Imidlertid ble det observert at implementering av avanserte prognosemetoder for de dynamiske bestillingspunktene ikke nødvendigvis presterer bedre enn grunnleggende etterspørselsprognosemetoder når det gjelder gjennomsnittlig lagerbeholdning og servicenivå. Dette understreker viktigheten av å vurdere avveiningen mellom kompleksitet og ytelse ved valg av prognosemetoder for de dynamiske bestillingspunktene.

Contents

Preface . . . . **i**

List of Figures . . . . **vi**

List of Tables . . . . **ix**

List of Acronyms . . . . **xi**

1 Introduction **1**

1.1 Background and Motivation 1

1.2 Literature Gap 3

1.3 Research Objective and Questions 3

1.4 Research Scope 5

1.5 Thesis Structure 7

2 Theoretical Background **8**

2.1 Inventory Management 8

2.2 Time-Series Forecasting 12

2.3 Descriptive Statistics 15

2.4 Data Preprocessing 19

2.5 Artificial Intelligence 20

3 Methodology **28**

3.1 Systematic Literature Review 28

3.2 Empirical Case Study 31

3.3 Data Analysis 32

3.4 Simulation 35

3.5 Research Overview 36

4 Systematic Literature Review **38**

4.1 Descriptive Analysis 38

4.2 Content Analysis 40

5 Empirical Case Study: Logistics Center Helse Midt-Norge . . . 48

5.1 Characteristics of The Case Company . . . 48

5.2 Data Analysis . . . 54

5.3 Multi-Scenario Analysis . . . 74

6 Discussion . . . 111

6.1 The State-of-The-Art in Demand Forecasting for Inventory Management . . . 112

6.2 Material Categorization Methods . . . 114

6.3 The Python-Based Simulation Model . . . 116

6.4 Improving the AS-IS Fixed Reorder Point . . . 118

6.5 Exploring the Impact of Forecast Intervals on Basic Dynamic ROP Strategy . . . 122

6.6 The Effectiveness of Advanced Forecasting Models for The Dynamic ROP . . . 127

7 Conclusion . . . 134

7.1 Implementation Guidelines for Dynamic Reorder Point Based on Demand Forecasting . . . 136

7.2 Contribution . . . 139

7.3 Limitations and Further Work . . . 139

References . . . 141

Appendix . . . 152

A Service Level . . . 152

B SLR Result Tables . . . 153

List of Figures

1 The elderly development index in Norway from 2000 to 2020 and the projected future development to 2040. The year 2020 equals an index of 100 (Mellbye and Gierløff, 2018). . . . 1

2 Potential contribution area illustrated 3

3 Established cooperation structures between municipalities and health care organizations in Norway (Asperud, 2020) 5

4 Graphical representation of the holding cost. 9

5 Graphical representation of the order cost. 9

6 Example of Periodic Review Policy (S. Chopra, 2016). 11

7 Example of Continuous Review Policy (Arnold, 2017). 11

8 Demand patterns resulting from trend and seasonal components (Moroff et al., 2021b). . . . 12

9 Common trends of time series 13

10 Syntatic transformation, (Chu and Ilyas, 2019) 19

11 Deep learning is a subset of machine learning, which in turn is a subset of artificial learning. 20

12 Under- and overfitting examples 21

13 Graphical representation of four selected activation Functions 23

14 Gradient descent optimization algorithm (Yamashita et al., 2018). 24

15 A common CNN architecture form in which convolutional layers are continuously stacked between ReLus, then passed through a pooling layer, and then passed through one or more fully connected layers (O'Shea and Nash, 2015). 25

16 Architecture of Feedforward Neural Network (FNN) and Recurrent Neural Network (RNN) 26

17 Visualization of the differences in cell structure in an RNN vs. LSTM network (Olah, 2022). 27

18 PRISMA flowchart illustrating the process of the literature selection 31

19 Flowchart describing data preparation method used for LC HMN data 33

20 The architecture for performing data analysis. 34

21 The architecture for the discrete event simulation. 35

22 Overview of the research flow 37

23 Year-wise publication of articles 38

24 Distribution of machine learning and statistical approach in articles in the publication year 39

25 Distribution of machine learning, statistical and the two combined 40

26 Bullwhip effect 41

27 Categorized demand observations based on the Syntetos method (Chuang et al., 2021) . . . 42

28 An example of decomposition of a time series. 44

29 An example of ensemble learning 47

30 Control model describing the AS-IS situation of Logistics Center Helse Midt-Norge 49

31 AS-IS order policy at LC HMN 51

32 Illustration of the three echelon supply chain. 53

33 Data filtering overview 59

34 Elbow method. 60

35 A clustering of all the materials, based on CV-value and Inventory Turnover Ratio. 60

36 A demand categorization of all the materials. 61

37 Lead time distribution within demand categories 63

38 Histogram of the clusters and demand category. 64

39 A demand categorization for cluster zero in Figure 38 65

40 Decomposition of three selected representative materials for each category: Erratic (4001095), Lumpy (4012198), and Smooth (4003841). 67

41 Auto-correlation of three selected representative materials for each category: Erratic (4001095), Lumpy (4012198), and Smooth (4003841). 69

42 Augmented Dickey-Fuller test for the three selected representative materials for each category: Erratic (4001095), Lumpy (4012198), and Smooth (4003841). 71

43 Sparsity analysis of the selected materials 72

44 Data point analysis of the selected materials 73

45 Simplified textual description of simulation algorithm 76

46 Scope of dynamic reorder point calculation for each forecast interval 77

47 The demand from 30/08/2021 to 26/08/2022 (260 business days) for the three selected representative materials for each category: Erratic (4001095), Lumpy (4012198), and Smooth (4003841). 80

48 AS-IS reorder point policy for the three selected representative materials for each category: Erratic (4001095), Lumpy (4012198), and Smooth (4003841). 82

49 Cumulative stock and demand levels for the three selected representative materials for each category: Erratic (4001095), Lumpy (4012198), and Smooth (4003841). 84

50 Average service level based on AS-IS fixed reorder point 85

51 Average inventory level AS-IS reorder point policy 86

52 Simplified illustration of optimization search algorithm 88

53 Distribution of optimal forecast interval and window size combination after optimization. . 89

54 Average service level result of optimal aggregation type 90

55 Percentage change in average inventory by strategy. 91

56 Holding cost analysis 93

57 Safety factor and the corresponding service level for Proposed Fixed ROP. 94

58 The visualization shows the effect of safety factors on service levels for Basic Dynamic ROP across all forecast intervals. 95

59 Distributions of window size for each "locked" forecasting interval (Basic Dynamic ROP) . 96

60 Forecast interval impact on service level 97

61 Aggregation impact on average inventory 98

62 Aggregation impact on holding cost for Basic Dynamic ROP 99

63 Predictions for erratic representative material (4001095) 101

64 Predictions for lumpy representative material (4012198) 102

65 Predictions for smooth representative material (4003841) 104

66 Most accurate model: LSTM - erratic representative material (4001095). 106

67 SMA (1W-10) model - erratic representative material (4001095). 107

68 Most accurate model: Holt-Winters Exponential Smoothing - lumpy representative material (4012198). 108

69 SMA (1W-10) model - lumpy representative material (4012198). 109

70 Most accurate model: SARIMAX - smooth representative material (4003841) 110

71 SMA (1W-10) model - smooth representative material (4003841) 110

72 Table of service level and safety factors (Arnold, 2017) 152

List of Tables

1 General Formulas in Inventory Management 6

2 Thesis Structure 7

3 Formulas of activation functions visualized in Figure 13 23

4 Explanation of notations in Figure 17 27

5 The stepwise development of search words. 29

6 Explanation of attributes of the data set. 55

7 The percentage of missing data. 56

8 The resulting representation of zero-values for the different forecast intervals in Figure 43 . 72

9 The resulting loss of data points for the different forecast intervals in Figure 44 73

10 Stepwise dynamic reorder point calculation 78

11 Inputs to the simulation of the selected materials, excluding demand 81

12 Holistic view of change in service Level and average inventory. 91

13 Strategy transition: from under 97% → over 97% 92

14 Strategy transition: from over 97% → under 97% 92

15 Replenishment policy impact on holding cost [NOK]. 93

16 Holistic view of change in service level and holding cost. 94

17 From under 97% → over 97% 98

18 From over 97% → under 97% 99

19 Forecast interval impact on holding cost [NOK]. 99

20 Forecast performance results for material 4001095 102

21 Forecast performance results for material 4012198 103

22 Forecast performance results for material 4003841 104

23 Simulation results based on demand forecasting model for erratic representative material (4001095). 105

24 Simulation results based on demand forecasting model for lumpy representative material (4012198). 107

25 Simulation results based on demand forecasting model for smooth representative material (4003841). 109

26 Research questions and their respective sections for discussion 111

27 SLR Result Table A 154

28 SLR Result Table B 155

29 SLR Result Table C 156

30 SLR Result Table D 157

31 SLR Result Table E 158

32 SLR Result Table F 159

List of Acronyms

AI Artificial Intelligence. 20, 21

ANN Artificial Neural Network. 22–25

ARIMA Autoregressive Integrated Moving Average. 22, 45

CNN Convolutional Neural Network. 23–25, 47

DDLT Demand During Lead Time. 12

DES Discrete Event Simulation. 35

DL Deep Learning. 20

EOQ Economic Order Quantity. 45

ERP Enterprise Resource Planning. 33, 59

FNN Feedforward Neural Network. 23, 25

KPI Key Performance Indicator. 116, 136

LC HMN Logistics Center Helse Midt-Norge. i, ii, vi, vii, 2–7, 31–33, 48–54, 58, 59, 65, 72, 76, 79–84, 98, 114–116, 119, 121, 122, 128, 130, 134, 135, 139

LSTM Long Short-Term Memory. 26, 27, 45, 47, 100, 128–133

MAE Mean Average Error. 18, 100, 101

MASE Mean Absolute Scaled Error. 18, 100, 101, 103–110, 129–133

ML Machine Learning. 20, 21, 39, 40

MRP Material Resource Planning. 59

NN Neural Network. 45, 46

OUL Order-Up-To. 10, 11, 40, 41, 44

PRISMA Preferred Reporting Items for Systematic Reviews and Meta-Analyses. 30

RMSE Root Mean Squared Error. 17, 18, 100, 101

RNN Recurrent Neural Network. 23, 25, 26, 45, 46

ROP Reorder Point. viii, 11, 51, 52, 76–78, 85–87, 90–95, 97–99, 105, 118–122, 124, 126, 127, 129–131

SAP System Analysis Program Development. 32, 33

SARIMAX Seasonal Autoregressive Integrated Moving Average with Exogenous variables. 22, 100, 128–133

SKU Stock Keeping Unit. ii, 12, 52, 56, 114

SL Service Level. 118, 119, 129

SLR Systematic Literature Review. 28, 29

SS Safety Stock. 12

SVM Support Vector Machines. 45–47

1 Introduction

This section contextualizes the research and presents the rationale and purpose. Section 1.1 outlines the context and rationale of the research. Section 1.2 describes the literature gap found through the systematic literature review. The literature gap serves as the basis for the present study. Section 1.3 provides a problem analysis that includes the research objective and research questions. Section 1.4 outlines the scope of the research. Lastly, Section 1.5 describes the structure of the thesis.

1.1 Background and Motivation

The Norwegian sector of health and long-term care (HC) is substantial. It employed 310,000 man-years in 2017, representing 13% of total employment (Leknes et al., 2019). The demand for HC is affected by the size and composition of the population, particularly the age and location. Norway's population is becoming more centralized, and the average lifespan is increasing (SSB, 2022). Consumption of HC resources increases significantly with age. As a result, central locations in Norway are expected to experience the highest increase in health care employment (Leknes et al., 2019).

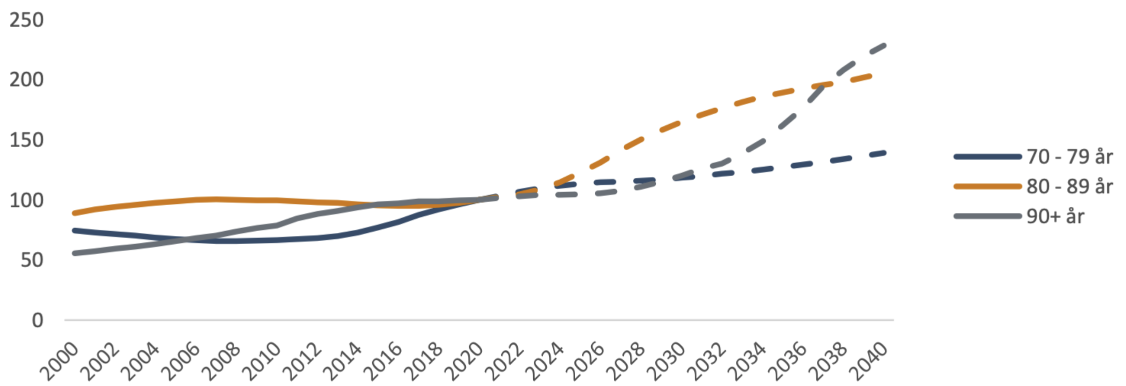

Figure 1: The elderly development index in Norway from 2000 to 2020 and the projected future development to 2040. The year 2020 equals an index of 100 (Mellbye and Gierløff, 2018).

The demographic changes characterized by a constantly growing and aging population is the main driver of the increasing demand for health services. As seen in Figure 1, in the last 20 years, the number of people older than 70 years has increased by 30%. The peak is not yet expected to be reached and is predicted to increase by more than 60% over the next 20 years. Growth is especially strong for age groups 80 – 89 years and older than 90 years (Mellbye and Gierløff, 2018).

The expenditures of elderly people are expected to experience a significant increase in the future, attributed both to a general increase in specific expenditures for the elderly and to the increase in expenditures for decedents (Gregersen, 2014). The older population leads to a smaller percentage of the population working and paying taxes, thus expenses for pensions and health care services are going to increase. At the same time, the oil and gas industry is not expected to be as an important driver of economic growth and productivity as it has been in the past (Mellbye and Gierløff, 2018). The government pension fund (Oljefondet) will not continue to grow as quickly (Meld.St, 2017). With more elderly and lower oil income, the effective use of resources in the public sector becomes even more important in the future (Mellbye and Gierløff, 2018).

In 2018 the real expenditure, adjusted for inflation, of Norwegian long-term care and health (HC) per inhabitant was 7.9 times the level in the year 1970. Out of these expenditures, 85% were tax-financed, and 17.5% of total public expenditure was devoted to HC. In 2017 the public "ensure-for-responsibility" seizure accounted for 14% of all man-years in the Norwegian economy (Holmøy, 2020).

According to the estimation by Mellbye and Gierløff (2018), Norwegian municipalities have the potential to save over 100 billion NOK through digital transformation between 2018 and 2027, with nearly half of the savings expected to be achieved within the health and social care sector. The results are based on a number of reports and analyses from Norway and abroad and the potential within several technological solutions, such as sensors, algorithms, machine learning, and data analysis.

The Norwegian government is encouraging the utilization of data opportunities to enhance the country's GDP and improve the efficiency of the public sector. It is crucial that Norway makes better use of data to successfully transition to a more sustainable society and economy (Meld.St, 2021). There is a consensus across political parties that the digital transformation of the public sector has great potential as a vital tool to address current challenges related to productivity, transformation, and efficiency in society. By embracing digital transformation, the public sector can achieve significant improvements in productivity, optimizing the utilization of public resources (Skodbo, 2017).

Increased use of big data will help streamline municipal tasks and change the way employees work (Mellbye and Gierløff, 2018). Municipalities are now prioritizing the implementation and exploitation of digitization opportunities. Eight out of ten municipal executive directors, IT directors, and municipal chiefs believe that machine learning will be important to improve work processes in the future (Mamre et al., n.d.). However, it is worth noting that nearly 40% of municipalities are yet to initiate machine learning in their IT projects. Notably, within the public sector, a substantial 89% confirms improved process quality through IT investments, with an additional 81% reporting efficiency gains (Skodbo, 2017). Consequently, there arises a critical necessity to conduct research focusing on data analytics and machine learning within the Norwegian health sector.

By analyzing data and information from Logistics Center Helse Midt-Norge (LC HMN), it is possible to conduct a more in-depth analysis and contribute to the discussion of the application of machine learning in the Norwegian healthcare system. The municipal sector depends on the success of digital transformation, as the trend in the next 10 years is towards higher demand for care services and lower income growth (Mellbye and Gierløff, 2018). Eivind Moen, Chief Operations Officer for St. Olav's Operational Service - Logistics and Supply, said in November 2022 (Midt-Norge-RHF, 2022):

> *"It is expected that patient treatment in Helse Midt-Norge will increase in the next eight to ten years. Logistics must not be an Achilles heel but must keep up. We need to be able to increase volume without necessarily increasing costs"*

LC HMN is responsible for the storage, distribution, and management of non-pharmaceutical medical equipment for Helse Midt-Norge. With a large number of goods stored and distributed from the logistics center, storage efficiency is an important part of increasing volume without necessarily increasing costs. When analyzing large amounts of data, it is possible to identify patterns and trends that can be used to improve inventory management and decision-making. Historical data can contribute to the efficiency of inventory levels and order frequency (Omsorgsdepartementet, 2022). This forms the basis for the theoretical motivation for this project.

1.2 Literature Gap

Existing research has primarily focused on the utilization of statistical methods and machine learning techniques for forecasting purposes. These studies have explored various applications of forecasting, such as sales prediction (R. Snyder, 2002), lead time demand distribution anticipation (Willemain et al., 2004), and demand forecasting for safety stock determination (Harvey and Ralph D. Snyder, 1990). Within the relevant articles, the primary objective has been to forecast demand in order to determine either the safety stock or the quantity of the orders.

However, the existing literature lacks sufficient research on the implementation of dynamic reorder points within a continuous replenishment policy and the potential benefits that can be derived from incorporating machine learning techniques for demand prediction. Furthermore, there is a notable gap in research when it comes to considering this combination in conjunction with service-level constraints.

1.3 Research Objective and Questions

As stated in Section 1.1, an analysis of data and information from Logistics Center Helse Midt-Norge (LC HMN) is an important contribution to the discussion of the application of machine learning in the Norwegian healthcare system. As visualized in Figure 2, this thesis aims to fill the literature gap stated in Section 1.2, and provide a contribution to the scientific literature on the intersection of machine learning, inventory management, and demand forecasting.

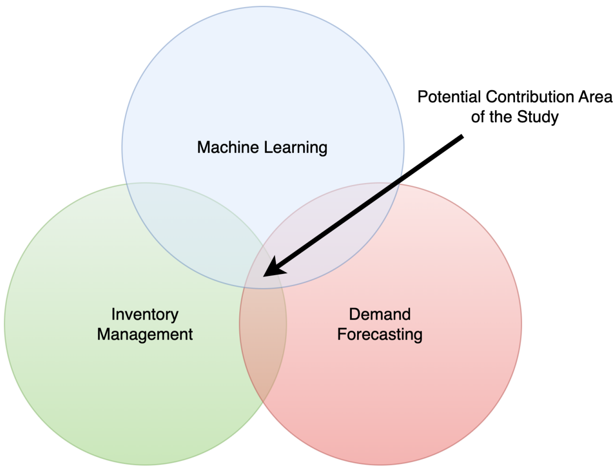

Figure 2: Potential contribution area illustrated

The objective of this research involves a contribution to the existing scientific knowledge in the intersection field represented in Figure 2, throughout an in-depth analysis of Logistics Center Helse Midt-Norge, focusing on dynamic reorder point policy. More specifically, the *objective* is to:

"Make a contribution to the current literature by investigating the effects of implementing demand forecasting and a dynamic reorder point policy for Logistics Center Helse Midt-Norge"

To achieve the objective and fulfill the purpose of the thesis, a series of research questions will be addressed and answered throughout the study. These research questions were formulated through an iterative process of research and discussion. The present study is led by the following research questions:

RQ1: What is the state-of-the-art within demand forecasting for inventory management?

This research question aims to explore the established use cases of time series-based forecasting for inventory management. Through conducting a systematic review of the literature, a comprehensive mapping of the various demand forecasting methods can be achieved. The systematic review of the literature will be presented in Section 4, supported by the theory presented in Section 2. The research question will be discussed in Section 6.1.

RQ2: How can the inventory be classified?

In the second research question, the focus will be on investigating relevant methods and measures that can be utilized for classifying the inventory of Logistics Center Helse Midt-Norge. The research question will be answered by performing a data analysis, and exploring findings from the systematic literature review (Section 4). The second research question will be discussed in Section 6.2.

RQ3: How can the AS-IS fixed reorder point be improved through a dynamic reorder point?

This research question aims to investigate the impact of transitioning from the current fixed reorder point to a dynamic reorder point. The research will employ a simulation method and rely on basic forecasting techniques to support the analysis. The results will be discussed in Section 6.4.

RQ4: What is the impact of implementing advanced forecasting methods for the dynamic reorder point?

The fourth and final research question explores the effects of implementing advanced forecasting methods on dynamic reorder points using the simulation method. More specifically, it builds upon the third research question by investigating the impact of machine learning and more advanced forecasting techniques on the simulation, comparing them to the basic techniques examined in the third research question. The findings are discussed in Section 6.6.

1.4 Research Scope

Logistics Center Helse Midt-Norge (LC HMN) provides non-pharmaceutical medical equipment to Helse Midt-Norge. As visualized in Figure 3, Helse Midt-Norge includes the four healthcare organizations: Helse Nord-Trøndelag Levanger, St. Olavs Hospital, Helse Møre og Romsdal Molde, and Helse Møre og Romsdal Ålesund. In terms of health care, Helse Midt-Norge has the responsibility for 732 000 Norwegian citizens (Asperud, 2020).

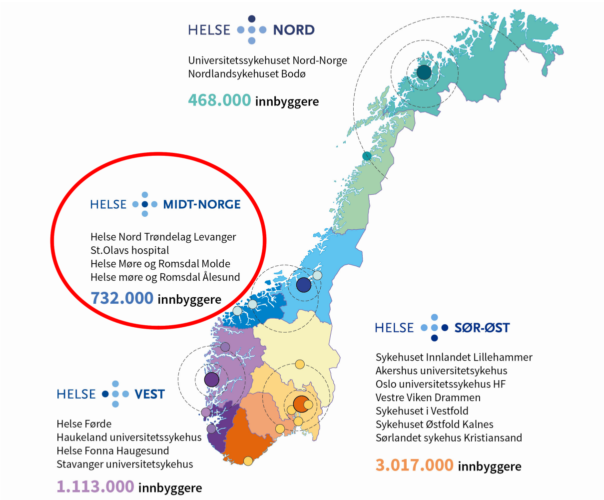

Figure 3: Established cooperation structures between municipalities and health care organizations in Norway (Asperud, 2020)

Given that St. Olav's Hospital is the largest customer of LC HMN, the flow of goods between these two parties is likely to be representative of the general flow of goods within LC HMN. For this reason, this thesis will focus on inventory management exclusively related to outgoing goods delivery to St. Olav's Hospital, which includes: St. Olavs Department Øya, St. Olavs Department Orkanger, St. Olavs Department Røros and satellite units.

Name	Formula	Parameters	Dependent on Demand
Safety Stock	SS $= \sigma \cdot SF$	σ - Standard Deviation SF - Safety Factor	True
Reorder Point	ROP $= DDLT + SS$	DDLT - Demand During Lead Time SS - Safety Stock	True
Target (Maximum) Inventory Level	T $= D \cdot (R+L) + SS$	D - Demand per unit of time L - Lead-time Duration R - Review-period Duration SS - Safety Stock	True
Order Quantity	Q $= T - I$	Q - Order Quantity I - Inventory on Hand	True
Economic Order Quantity	EOQ $= \sqrt{\frac{2 \cdot A \cdot S}{i \cdot c}}$	A - Annual Demand S - Ordering Cost i - Carrying Cost c - Unit Cost	True

Table 1: General Formulas in Inventory Management

Table 1 demonstrates that all the mentioned formulas are directly or indirectly dependent on demand. For this reason, this thesis will examine how demand as a parameter affects inventory management. The flow of outgoing goods for Logistics Center Helse Midt-Norge is a key factor that affects inventory levels. Therefore, the scope of this thesis will include the demand for these goods. For data analysis, a representative sample of goods will be analyzed in order to draw conclusions about demand patterns and forecasting challenges for the company as a whole. The sample size will be determined on the basis of statistical considerations and the specific research goals of the study.

The investigation will involve evaluation of the effects of incorporating insights from demand analysis, by simulating the warehouse environment. Specifically, the focus will be on experimenting with the reorder point. However, due to the time constraints of this master's thesis, the research scope will exclude the examination of the potential impact on order size resulting from demand insights.

In summary, the scope of this thesis will be limited to inventory management, specifically examining reorder points at the case company LC HMN. The study will primarily concentrate on the demand for outgoing goods delivery to St. Olav's Hospital.

1.5 Thesis Structure

Table 2: Thesis Structure

Section 1 Introduction	The introduction describes the background and purpose of the project, research objectives, questions and scope, and the structure of the report.
Section 2 Theoretical Background	The theory section introduces the theoretical background of the research context and discusses different theoretical perspectives defined in the relevant literature. It provides a theoretical overview.
Section 3 Methodology	The methodology section outlines the rationale for selecting the research approach, as well as providing details on the process of conducting the systematic literature review, empirical case study, data analysis, and simulation.
Section 4 Systematic Literature Review	In this section, the results and findings of the systematic literature review are presented. The systematic literature review is concerned with reviewing recently published literature in the field of inventory management, demand forecasting, and machine learning. Relevant articles are utilized for answering the research objective and research questions.
Section 5 Empirical Case Study	This section presents the results of an empirical case study, including an extensive description of how LC HMN operates, a data analysis, and a multi-scenario analysis.
Section 6 Discussion	The discussion section of this thesis will analyze and interpret the findings obtained from both the systematic literature review and the empirical case study. These findings will be discussed in the context of the research questions.
Section 7 Conclusion	The conclusion summarizes the results, present a step-wise guideline, highlights research contributions, and present the limitations and potential further work.

2 Theoretical Background

This section presents a summary of the established theoretical frameworks found in the existing literature, which serve as the foundation for the research context. These frameworks will contribute to the interpretation of the findings obtained from the systematic literature review and the empirical case study. Theoretical perspectives related to inventory management, demand forecasting, and artificial intelligence will be presented.

2.1 Inventory Management

Inventory management is the process of managing the supply of materials to ensure that the right amount of stock is available at the right time (Stevenson, 2021). It involves tracking the movement of goods from the point of origin to the point of consumption, as well as managing the storage and distribution of goods. Effective inventory management is essential to ensure that the right amount of stock is available to meet customer demand and to avoid overstocking or understocking (Arnold, 2017).

This process involves several key considerations, such as deciding what and when to order. While the quantity of items to be ordered is an important aspect of inventory management, the order quantity is excluded from the scope of this thesis, as outlined in Section 1.4. Timing plays a critical role in inventory replenishment to prevent stockouts and minimize the potential for excess inventory, necessitating the consideration of trade-offs between factors such as holding costs and service level (Stevenson, 2021). Therefore, this section presents the relevant costs and terms associated with inventory management.

2.1.1 Relevant Costs in Inventory Management

In addition to the initial purchase price, the acquisition of materials encompasses several direct expenses that necessitate careful consideration. This subsection specifically delves into two primary costs within this domain: holding costs and ordering costs.

Holding Costs

For inventory management decisions, there are several costs to take into consideration. The price paid for a purchased item, known as the item cost, includes the cost of the item itself and any additional direct costs incurred in bringing the item into the facility. The item price is necessary to calculate the holding costs. The holding cost increases in direct proportion to inventories (Arnold, 2017). Holding costs can be divided into three types of cost: capital cost, storage cost, and risk costs. Capital expenses can be affected by interest rates, company credit standing, and possible investment opportunities (Stevenson, 2021). The location and type of storage needed will affect the storage cost. Risk cost is linked to the risk of carrying a type of product that can obtain a loss of product value due to technologically aging or perishable goods that may expire.

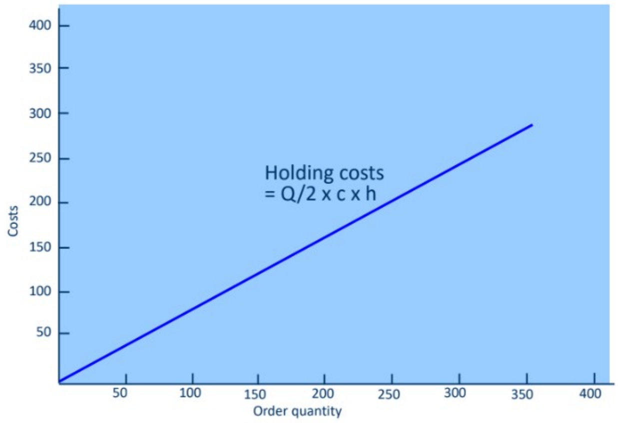

Figure 4: Graphical representation of the holding cost.

The total holding costs can be expressed as shown in Figure 4, where Q equals the order quantity in units, c equals the unit purchase cost, and h equals the holding cost per year as a fraction of product cost (S. Chopra, 2016).

Ordering Costs

Ordering costs refer to the expenses associated with the process of placing and fulfilling an order. Ordering costs are often fixed expenditures, independent of the number of units ordered (Stevenson, 2021). Ordering costs encompass various expenses involved in the preparation, follow-up, expedition, receipt, approval of payment, and accounting charges associated with order acceptance and invoice payment (Arnold, 2017).

As the number of orders placed by a company increases, the accumulated ordering expenses also rise. However, by utilizing order releases within the framework of blanket orders, which are placed in large quantities and cover extended time periods, it is possible to reduce the overall order cost.

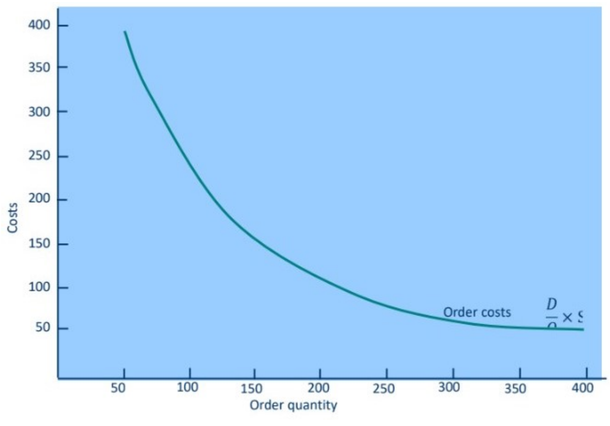

Figure 5: Graphical representation of the order cost.

The ordering cost is expressed as shown in Figure 5, where D is the annual demand, Q is the given lot size, and S is the order cost for each order placed. A warehouse can withstand a higher total ordering cost if it leads to a decrease in total holding costs. When a company only purchases goods as needed, it results in more frequent orders but reduces the inventory held. To achieve the right balance between order quantities and minimize overall costs, a company must carefully monitor its ordering costs and holding costs (Stevenson, 2021).

Stockout, Safety Stock & Service Level

A stockout is possible if demand during the lead time is higher than anticipated. Due to backorder fees, lost sales, and perhaps even lost customers, a stockout could be costly. Carrying extra inventory to protect against situations where demand during lead time is higher than anticipated can help to reduce stockouts (Stevenson, 2021).

The purpose of safety stock is to mitigate uncertainty in supply and demand and to avoid the possibility of stockout. Uncertainty may occur in two ways: quantity uncertainty and timing uncertainty. Safety stock is an additional amount of stock carried and is generally the most commonly used buffer (Arnold, 2017). The level of safety stock is determined by the variability of demand, the frequency of reordering, the lead time, and the desired service level (Arnold, 2017). The calculation of safety stock (SS) is expressed as:

$$\text{SS} = \sigma \times \text{SF} \tag{1}$$

Where σ represents the standard deviation of demand within a time period, and the safety factor (SF) represents the number of standard deviations needed for the desired service level. The connection between safety factor and service level can be seen in Appendix A.

The service level of a warehouse is expressed as the proportion of orders fulfilled without a stockout. A service level of 95% implies that a stockout is possible only during the time interval between a customer's order and the warehouse's replenishment and that the warehouse is able to supply the customer's order 95% of the time (Arnold, 2017). The service level is calculated as follows:

$$\text{Service Level} = \frac{n_{\text{orders}} - n_{\text{stockouts}}}{n_{\text{orders}}} \tag{2}$$

2.1.2 Replenishment Policies

This section provides an overview of two replenishment policies in inventory management: the periodic review policy and the continuous review policy. These policies serve as approaches for maintaining inventory levels and ensuring efficient replenishment processes.

Periodic Review Policy

In a periodic review policy, the inventory level is evaluated at predetermined intervals and the Order-Up-To (OUL) is determined as the sum of the existing inventory and the size of the replenishment lot. An order is initiated when the sum of the current inventory level and the lot size reaches the order-up-to level (OUL) (S. Chopra, 2016). The period of time between orders is known as the review interval. Each order's size may vary depending on the demand encountered between orders and the inventory level at the time of ordering.

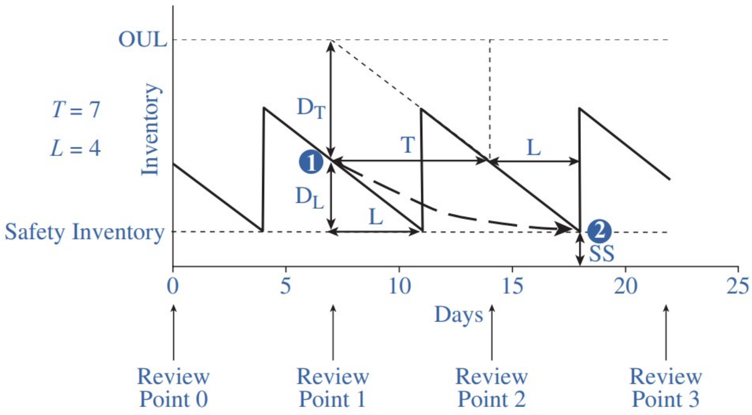

Figure 6: Example of Periodic Review Policy (S. Chopra, 2016).

Figure 6 shows how an inventory with a periodic review policy would operate. A fixed review interval is sat, in this example, seven days, and the quantity being ordered equals the Order-Up-To (OUL) minus the quantity on stock.

Continuous Review Policy

In an inventory management system utilizing a continuous review policy, technology is utilized to constantly monitor inventory levels. This enables orders to be triggered when the inventory falls below a predetermined reorder point (Stevenson, 2021). A continuous review policy, often referred to as Reorder Point (ROP) policy, should take into account the uncertainty of demand during the lead time period.

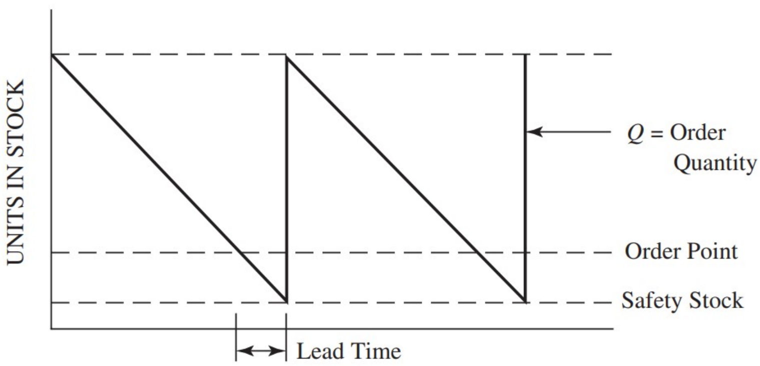

Figure 7: Example of Continuous Review Policy (Arnold, 2017).

Inventory replenishment must consider reordering points and safety stock levels to guard against stock out. Mean lead time demand should be taken into account when determining the necessary quantity of product to be stored (John J. Bartholdi and Hackman, 2019). Incorporating lead time variability into the fundamental inventory model can improve its accuracy. By combining safety stock levels and lead time demand, an reorder point, also referred to as ROP, can be determined to mitigate the risk of stockouts and minimize the impact of lead time fluctuations (Arnold, 2017). The reorder point is defined as:

$$\text{Reorder Point} = \text{DDLT} + \text{SS} \tag{3}$$

Where Demand During Lead Time (DDLT) is the anticipated demand for a Stock Keeping Unit during the lead time required to replenish the inventory. As explained earlier, the Safety Stock (SS) is the additional quantity of stock held to mitigate uncertainties in demand and lead time. As for order quantity, in continuous review policies, the order size is kept fixed between replenishment (S. Chopra, 2016).

2.2 Time-Series Forecasting

In this section, relevant statistical time series forecasting methods are presented. Time-series forecasting involves analyzing historical data to predict future values based on patterns and trends within the data.

A time series comprises sequentially arranged raw data points, such as monthly sales for a specific product spanning multiple years. Statistical methods are applicable when sufficient data spanning several years exist for a product, and when there are clear and relatively consistent relationships and patterns present (Chambers et al., 1971). Time-series analysis enables the detection and elucidation of various phenomena, including:

- Seasonality-related patterns or systematic variations within a data series.
- Repetitive patterns occurring at regular intervals of two or more years.
- Analysis of data trends.
- Examination of growth rates associated with these trends.

Wold's Theory of Decomposition posits that a time series can be deconstructed into four distinct components: trend (T), seasonal (S), cyclical (C), and residual (I). This theory suggests that by breaking down a time series into its constituent components, one can gain a deeper understanding of the underlying patterns and trends (Chase, 2016, Treyer, 2010). Based on the characteristics of these components, diverse demand patterns can be generated (Treyer, 2010). Figure 8 illustrates the combination of a trend and a seasonal component with the corresponding demand patterns.

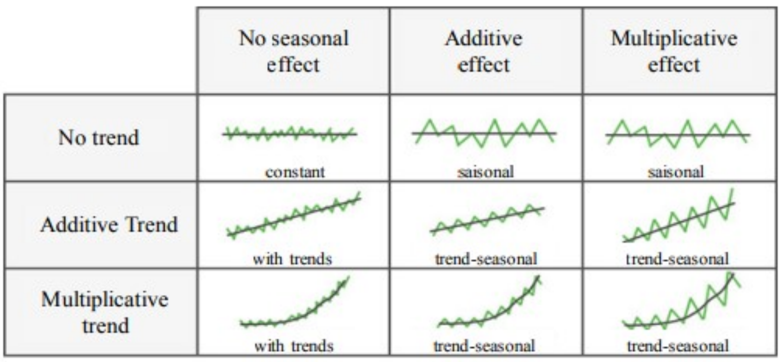

Figure 8: Demand patterns resulting from trend and seasonal components (Moroff et al., 2021b).

Time-series analysis models aim to identify recurring patterns within historical observations, enabling the forecasting of future events. This analysis is valuable to businesses in making informed decisions regarding product sales, inventory management, staffing requirements, and marketing strategies (Shumway, 2017). Time-series forecasting approaches can generally be categorized into two main methodologies: machine learning methods and traditional statistical methods. These methodologies diverge in their approach to forecasting, with machine learning methods employing algorithms to detect patterns within the data, while traditional statistical methods rely on mathematical models for prediction (S. Zhang et al., 2017).

Typical Use-cases for Time Series

Time series are commonly used to address various challenges, such as business forecasting, understanding past behavior, and evaluating current achievements.

- **Business forecasting:** Common scenarios where predictive analytics is applied involves a retailer seeking to anticipate future sales volume and a financial investor predicting the future trajectory of a particular stock.
- **Understanding past behavior:** By analyzing historical time series data, a supplier can gain insight into sales patterns over specific months or periods of the year, including fluctuations and trends. This examination enables the supplier to develop a deeper understanding of seasonal variations within the market.
- **Evaluate current accomplishments:** Utilizing time series analysis, one can make predictions regarding future events and subsequently evaluate performance by comparing actual outcomes with the earlier predictions. This allows for an evaluation of progress and performance during the designated period.

A trend in a time series refers to a noticeable pattern in the data that shows how it changes over time. Typically, a trend involves either an increase or decrease in the values of the series and lasts for a short period before fading away. In Figure 9, you can see six different types of trends visually represented.

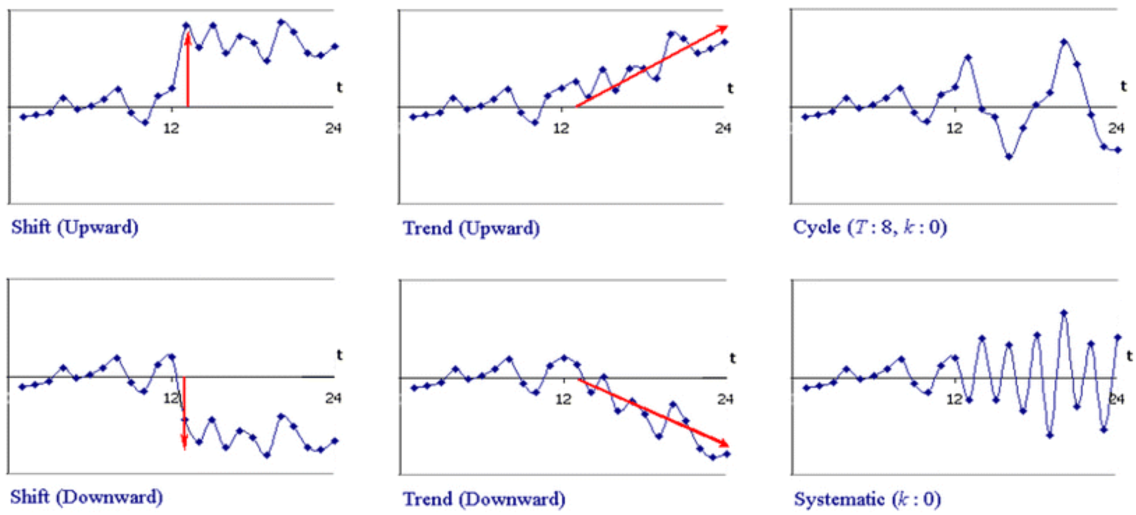

Figure 9: Common trends of time series

Time-series forecasting encompasses univariate and multivariate methodologies. Univariate time-series forecasting leverages a singular attribute from the temporal sequence as an input to predict the subsequent output. On the contrary, multivariate time series forecasting incorporates a collection of attributes for each temporal point to generate a single output for the subsequent time step (Vercellis, 2011).

Simple Moving Average

The Simple Moving Average (SMA) is a method that calculates the average value of a variable over a specified period of time. It is called "simple" because it assigns equal weight to all observations within the specified period (Wheelwright et al., 1998).

Let X_t be a time series variable at time t, and let n be the number of observations used for calculating the SMA. The formula for calculating the SMA at time t is as follows:

$$SMA_t = \frac{X_{t-n+1} + X_{t-n+2} + \ldots + X_{t-1} + X_t}{n}$$

Here, $X_{t-n+1}, X_{t-n+2}, \ldots, X_{t-1}, X_t$ are the n observations used to calculate the SMA at time t.

The SMA method is simple to understand and easy to calculate. It provides a quick estimate of the current trend in the time series. However, it has some limitations. SMA does not consider the weightage of different observations, and it may not capture sudden changes or outliers in the data. Additionally, since SMA assigns equal weight to all observations, it may not be suitable for time series data with varying trends or seasonality.

Holt-Winters Exponential Smoothing

The Holt-Winters exponential smoothing model is a forecasting technique used to analyze and predict time series data. It is an extension of simple exponential smoothing that takes into account both trend and seasonality in the data (Vercellis, 2011). The model was proposed by Charles Holt and Peter Winters in the 1960s. The Holt-Winters model is based on three smoothing components (Wheelwright et al., 1998):

- Level equation: Represents the average value of the series over time. It is denoted by the symbol l.

$$l(t) = \alpha \cdot (y(t) - s(t-L)) + (1-\alpha) \cdot (l(t-1) + b(t-1))$$

- Trend equation: Accounts for the increasing or decreasing pattern in the data. It is denoted by the symbol b.

$$b(t) = \beta \cdot (l(t) - l(t-1)) + (1-\beta) \cdot b(t-1)$$

- Seasonality equation: Captures the repeating patterns or cycles in the data. It is denoted by the symbol s.

$$s(t) = \gamma \cdot (y(t) - l(t)) + (1-\gamma) \cdot s(t-L)$$

Where:

- $y(t)$ represents the actual value at time t.
- L denotes the length of the seasonal period (e.g., 12 for monthly data with annual seasonality).
- α, β, and γ are smoothing parameters that control the weights given to the new observations versus the existing components. They should be chosen or estimated to optimize forecasting accuracy.

The model uses three equations to update these components at each time step. Once the components are updated, they are used to make forecasts for future time periods by projecting the level, trend, and seasonality forward. The Holt-Winters model has different variations, including additive and multiplicative models, depending on whether the seasonal component is added or multiplied by the other components (Wheelwright et al., 1998). The choice between the two depends on the characteristics of the data and the underlying patterns.

2.3 Descriptive Statistics

In this section, relevant descriptive statistics are presented, including univariate analysis, multivariate analysis, and forecasting performance metrics. Univariate analysis focuses on single variables, while multivariate analysis explores relationships between multiple variables.

2.3.1 Univariate Analysis

Univariate analysis is a statistical approach employed for utilized to assess the relationship between an individual independent variable and a dependent variable. This method entails investigating the distributions, measures of central tendency, and dispersion of a singular variable in isolation (Vercellis, 2011).

Arithmetic Mean ($\bar{X}$)

One commonly employed method for quantifying central tendencies in a dataset is the arithmetic mean, also known as the mean or average. The arithmetic mean is determined by summing all the values within a sample and dividing the sum by the total number of values in the sample. It is considered the most appropriate measure of central tendency when the data follow a normal distribution, although it can be influenced significantly by the presence of large outliers (Vercellis, 2011). The arithmetic mean can be represented mathematically as follows:

$$\mathrm{AM}(X) = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{4}$$

Median

The median is defined as the central value within a given sample and is particularly useful when the data exhibit skewness or contain outliers (Miller, 1993).

$$M = \begin{cases} a\left[\frac{N}{2}\right] & \text{if } N \text{ is even} \\ \frac{a\left[\frac{N-1}{2}\right]+a\left[\frac{N+1}{2}\right]}{2} & \text{if } N \text{ is odd} \end{cases} \tag{5}$$

The median is a more robust estimator than the mean, as shown in Equation 4 and Equation 5. The median refers to the statistical measure that identifies the central value within a sorted dataset, while the mean encompasses all values in its calculation. When the data set contains significant outliers, the median is preferred to use because it is not affected by extreme values. This attribute makes the median a more resilient statistic (Vercellis, 2011).

Variance (σ^2) & Standard Deviation (σ)

Variance and standard deviation are statistical indicators that quantify the proximity of individual data points to the mean of a given dataset. In datasets characterized by a limited dispersion, data points tend to cluster closely around the mean, leading to lower values for both variance and standard deviation. Conversely, datasets with a wider range of values that differ from the mean exhibit higher values for variance and standard deviation. Consequently, when all values within a data set are identical, both variance and standard deviation assume a value of zero (P. Kaur et al., 2018). Below are the mathematical representation of the two:

$$\text{Var}(X) = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2 \tag{6}$$

$$\text{SD}(X) = \sqrt{\text{Var}(X)} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2} \tag{7}$$

2.3.2 Bivariate & Multivariate Analysis

Bivariate and multivariate analyses are separate methodologies employed in data analysis to examine the associations among variables. Bivariate analysis focuses on exploring the relationship between two variables, whereas multivariate analysis investigates the relationships involving three or more variables. Both approaches serve the purpose of unveiling underlying patterns and trends within the data, which may not be evident in its raw form.

Covariance

The computation of covariance involves determining the average of the products derived from the differences between the values of two variables and their respective means. This computation yields a singular numerical value capable of spanning from positive to negative values. A positive covariance indicates a positive association between the variables, implying that an increase in one variable is likely to correspond with an increase in the other. Conversely, a negative covariance signifies a negative association, suggesting that an increase in one variable tends to coincide with a decrease in the other (Vercellis, 2011).

$$\text{Cov}(X, Y) = \frac{\sum_{i=1}^{N} (X_i - \bar{X})(Y_i - \bar{Y})}{N - 1} \tag{8}$$

Correlation

Correlation provides a quantitative measure of the linear relationship between two random variables X and Y within a given dataset. The resulting coefficient resides within the open interval $(-1, 1)$, where values of 1 or -1 indicate a strong positive or negative correlation, respectively, between the variables, while a value of 0 signifies no correlation (Vercellis, 2011). The correlation coefficient, denoted as $\text{Cor}(X, Y)$, is defined as:

$$\text{Cor}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \quad \text{if } \sigma_X \sigma_Y > 0, \tag{9}$$

where $\text{cov}(X, Y)$ represents the covariance, and σ_X and σ_Y denote the standard deviation of X and Y, respectively. It should be noted that $\text{Cor}(X, Y)$ is only applicable when both standard deviations are finite and positive.

2.3.3 Forecasting Performance Metrics

This section presents relevant performance metrics to measure the performance of forecasting models.

RMSE

Root Mean Squared Error (RMSE) serves as a widely utilized performance metric for quantifying the disparity between predicted and actual values in regression problems. The process of computing RMSE can be outlined into three straightforward steps (Willmott and Matsuura, 2005).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_{i,pred} - y_{i,true})^2} \tag{10}$$

Where:

- n is the total number of observations
- $y_{i,pred}$ is the predicted value for observation i
- $y_{i,true}$ is the true (actual) value for observation i

Firstly, the "total square error" is computed by summing the squared errors of each individual observation. This method incorporates a weighting scheme that emphasizes the contribution of larger errors by assigning them greater importance due to their squared magnitude. Consequently, if the total error is concentrated within a decreasing number of increasingly significant individual errors, the total square error will consequently increase. Subsequently, the total square error is divided by the number of observations

(n), yielding the mean square error (MSE). Lastly, the RMSE is derived by taking the square root of the MSE.

MAE

Mean Average Error (MAE) serves as a performance metric that quantifies the average magnitude of errors in model predictions, without considering their direction. Unlike RMSE, which assigns greater weight to larger errors (Hyndman and Koehler, 2006), MAE treats all errors equally, providing a clearer indication of the model's accuracy in terms of the absolute difference between predicted and actual values (Chai and Draxler, 2014).

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_{i,pred} - y_{i,true}| \tag{11}$$

Where:

- n is the total number of observations
- $y_{i,pred}$ is the predicted value for observation i
- $y_{i,true}$ is the true (actual) value for observation i

MASE

Mean Absolute Scaled Error (MASE) proves valuable in comparing the efficacy of diverse forecasting methods, as it facilitates the standardization of error metrics across varying datasets and time series (Hyndman and Koehler, 2006). MASE accomplishes this by comparing the absolute errors of a forecast with those of a naive forecast, which represents a straightforward forecasting approach assuming that future values of the time series will match the most recent observed value.

$$\text{MASE} = \frac{\frac{1}{T}\sum_{t=1}^{T} |y_t - \hat{y}t|}{\frac{1}{T-1}\sum t = 2^T |y_t - y_{t-1}|} \tag{12}$$

Where:

- y_t represents the actual value of the time series at time t
- $\hat{y}_t$ represents the forecasted value at time t
- T represents the total number of time periods

By contrasting the two error sets, the MASE provides a measurement of forecast accuracy that is normalized in relation to the errors of the naive forecast. This normalization facilitates a more straightforward interpretation and comparison of forecast accuracy, ensuring consistency across distinct time series and forecasting methodologies.

The ranges of the MASE statistic can be explained as follows:

- MASE = 1: The evaluated forecasting method is as good as the Naïve method.
- MASE < 1: The forecasting method utilized is better than the Naïve method. The smaller the MASE, the better the forecasting method is relative to the Naïve method.
- MASE > 1: The forecasting method performs worse than the Naïve method. There's no point in using the forecasting method.

2.4 Data Preprocessing

Data preprocessing is a crucial step within the data analysis pipeline. This step encompasses the vital activities of data cleaning, data transformation, and data reduction, all of which are undertaken to prepare the dataset for subsequent analyses.

Data Cleaning

Data cleaning is a fundamental procedure in data science, aiming to minimize the presence of noise within the data and enhance the accuracy and reliability of the dataset (Brownlee, 2022). It is the process of refining the data into a usable format by locating missing data, correcting incorrect values, removing duplicate entries, and standardizing the data, thus ensuring its suitability for analysis (Sattler and Schallehn, 2001).

Data Transformation

Data transformation plays a crucial role in data preparation as it aims to enhance manageability by reducing the complexity of the data. This process enables the identification of patterns and trends within the data with greater ease, thereby enhancing the accuracy of the dataset (Chandrasekar et al., 2017).

Name	
John	Tel:7188751243
	Fax:7188751200
Mike	Tel:7186359762
	Fax:7186359700
Frank	Tel:5198780763
	Fax:5198780700
Julie	Tel:5176543809
	Fax:5176543800

Name	Tel	Fax
John	718-875-1243	718-875-1200
Mike	718-635-9762	718-635-9700
Frank	519-878-0763	519-878-0700
Julie	517-654-3809	517-654-3800

Figure 10: Syntatic transformation, (Chu and Ilyas, 2019)

Chu and Ilyas (2019) described the underlying method of syntactic transformation as a central part of the data preparation. Syntactic transformation entails the manipulation of data values to adhere to a specific format or predefined set of rules, such as a database schema (Abdallah et al., 2017). Figure 10 presents a demonstrative instance of syntactic transformation, wherein the data undergo processing to enhance manageability and facilitate subsequent analysis.

Data Reduction

Data reduction constitutes a critical phase within the data preprocessing workflow, with the aim of reducing the volume of data while preserving its integrity. By employing data reduction techniques, the amount of data to be stored and analyzed is reduced, leading to improved efficiency and accuracy in data processing. The main objectives of data reduction encompass the reduction of storage space, processing time, and the complexity of data analysis. Various techniques can be employed to achieve data reduction, including data compression, selection of attribute subsets, aggregation, and data transformation (Namey et al., 2008).

The selection of attributes in subsets involves identifying a subset of attributes in the dataset that are most relevant to the analysis. Aggregation consolidates similar records or values into a single entry. Lastly, data transformation techniques facilitate the conversion of data from one form to another, such as from numerical to categorical or from unstructured to structured representations (Hall and Holmes, 2003).

2.5 Artificial Intelligence

This section provides an overview of the hierarchy of artificial intelligence, machine learning, and deep learning. Additionally, it explores concepts of time series forecasting, presenting relevant models and techniques.

Artificial Intelligence is a complex concept that has been difficult to define precisely. Alan Turing proposed the *Turing Test* in his seminal work *Computing machinery and intelligence* (Turing, 2009), which suggests that a machine may be regarded as intelligent if it is indistinguishable from a human in conversation by a neutral observer. Artificial Intelligence refers to a machine's capacity to communicate, think, and act autonomously in both familiar and unfamiliar environments in a manner comparable to that of a human (Du-Harpur et al., 2020). AI is a broad field of research that encompasses Machine Learning and Deep Learning. Figure 11 demonstrates that Deep Learning is an integral part of Machine Learning. The three technologies, namely Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL), are observed to be subsets of each other.

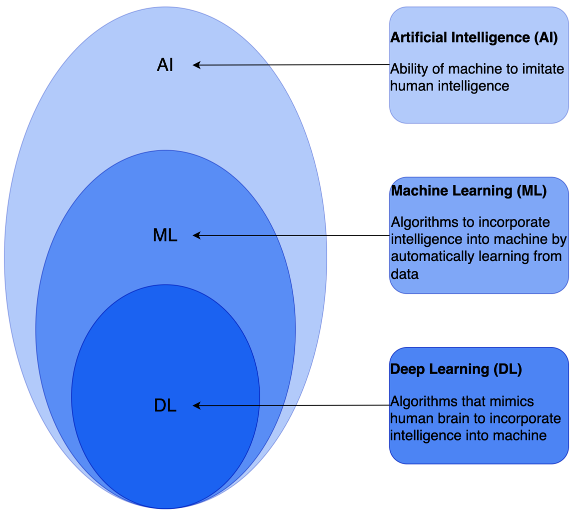

Figure 11: Deep learning is a subset of machine learning, which in turn is a subset of artificial learning.

2.5.1 Machine Learning

Machine Learning is a subset of Artificial Intelligence (Figure 11) that includes methods to allow machines to learn from data without being explicitly programmed (Samuel, 1988). By employing algorithms and statistical models, computers can be trained to analyze data and identify underlying patterns within it. This enables them to perform specific tasks without explicit guidance from human operators. In Machine Learning, the algorithm can be trained to improve its accuracy by striving to reduce the error and increase the probability of its predictions being correct (Jakhar and I. Kaur, 2020).

Supervised & Unsupervised Learning

In the field of machine learning, supervised learning refers to tasks where the objective is to discover an optimal function that accurately maps a given set of inputs (e.g., images) to their corresponding correct outputs (labeled data). This process relies on the availability of a predetermined pair of training datasets. Conversely, unsupervised learning focuses on the identification of latent patterns within data, such as clusters or groups, in the absence of prior knowledge or explicit labeling (Vercellis, 2011).

Overfitting

Overfitting is a frequently encountered phenomenon in machine learning, characterized by a situation in which a model demonstrates satisfactory performance when evaluated on training data but fails to generalize effectively when applied to data from different sources. The primary cause of overfitting is often attributed to the complexity of the model itself. A highly complex model tends to learn complex functions that may not be present in other datasets. Consequently, when confronted with new data, the excessively complex model becomes more prone to overfitting, thereby hindering its ability to generalize accurately (Hawkins, 2004).

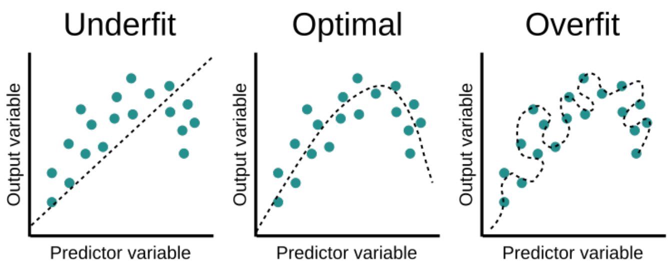

Figure 12: Under- and overfitting examples

Regression Tree

Regression trees are a supervised machine learning technique used to predict numerical outcomes based on a set of features (Nasteski, 2017). This method of data analysis involves partitioning the data into smaller subsets and using these subsets to build a model that captures the relationships between the features and the target variable (Elith et al., 2008). This model can be used to identify linear and non-linear relationships in the data.

SARIMAX

Seasonal Autoregressive Integrated Moving Average with Exogenous variables (SARIMAX) is a time series forecasting model that builds upon the popular Autoregressive Integrated Moving Average (ARIMA) model. The main difference between SARIMAX and ARIMA is that SARIMAX includes additional variables, known as exogenous variables, that can help to explain the behavior of the time series (Vagropoulos et al., 2016). Exogenous variables are variables that are not part of the time series being forecasted but are believed to influence it.

The SARIMAX model is typically represented as SARIMAX(p, d, q) (P, D, Q)s, where:

- p is the order of the autoregressive term (AR)
- d is the order of differencing required to make the time series stationary (I)
- q is the order of the moving average term (MA)
- P is the seasonal order of the autoregressive term (SAR)
- D is the seasonal order of differencing required to make the time series stationary (SI)
- Q is the seasonal order of the moving average term (SMA)
- s is the number of time periods in a season

The SARIMAX model is fitted to the time series data using maximum likelihood estimation. Once the model is fitted, it can be used to make forecasts by extrapolating the trend and seasonal components of the time series, as well as incorporating the effects of any exogenous variables (Williams and Hoel, 2003).

K-Means Clustering

K-Means Clustering is an unsupervised machine learning technique utilized for the purpose of categorizing data points into a predetermined number (K) of distinct clusters (Morissette and Chartier, 2013). The algorithm operates by initially assigning each data point to one of the K clusters and subsequently adjusting the centroid position of each cluster iteratively. The objective is to minimize the distance between the data points and their respective cluster centroids. This iterative process continues until a local optimum is reached, where the distance between the data points and centroids is minimized to the greatest extent possible (Vercellis, 2011).

2.5.2 Deep Learning

Deep learning is a subset of machine learning (Figure 11) that includes computational models and algorithms that mimic the architecture of organic neural networks in the brain.

Artificial Neural Network

An Artificial Neural Network, in contrast to conventional machine learning models, exhibits the capability to capture and leverage non-linear associations between input and output variables (Yang and X.-J. Wang, 2020). This means that Artificial Neural Network's can learn complex patterns in data that cannot be identified by traditional models (Sarker, 2021). They are used in a wide range of applications, including computer vision, natural language processing, and time series forecasting (Du and Sun, 2006; Baroni, 2020;

Khashei and Bijari, 2010). Feedforward Neural Network, Convolutional Neural Network and Recurrent Neural Network are all variations of Artificial Neural Network that have been modified to address specific problem domains and exhibit improved performance.

Non-linearity

Non-linearity plays a crucial role in the functioning of Artificial Neural Network, serving to modulate or limit the output produced by linear activation functions (Albawi et al., 2017). Sigmoid and Tanh functions are commonly used for the non-linear transformation step of ANN (J. Wang et al., 2016). However, as the complexity of the neural network architecture increases, the gradient signal may decrease, leading to the "vanishing gradient" problem (Krizhevsky et al., 2017). To address this issue, the Rectified Linear Unit (ReLU) activation function offers a constant gradient for positive input values Figure 12. Additionally, ReLU and Softplus, both unsaturated activation functions, provide positive inputs with a constant gradient. Comparative studies have shown that deep ANN using ReLU activation functions outperform their counterparts using Tanh units (Krizhevsky et al., 2017).

Figure 13: Graphical representation of four selected activation Functions

Name	Sigmoid	Softplus	ReLU	Tanh
Formula	$\frac{1}{1+e^{-x}}$	$ln(1+e^{x})$	$max(0,x)$	$\frac{2}{1+e^{-2x}}-1$

Table 3: Formulas of activation functions visualized in Figure 13

Backpropagation

Backpropagation is a fundamental procedure for calculating the gradient of an objective function with respect to the weights in a multilayer network. It applies the chain rule to determine the gradient by working backward from the output to the input of each module. By iteratively propagating gradients through the network, starting from the top and moving down to the bottom, backpropagation allows for the computation of gradients with respect to the weights. This technique is vital in training neural networks, as it allows weight adjustments to minimize the objective function and improve performance (LeCun et al., 2015).

Gradient Decent

Gradient descent is an optimization algorithm that is used to minimize a function iteratively. In the context of neural networks, it is commonly used to update the network's weights during the backpropagation process (LeCun et al., 2015). The gradient of the loss function signifies the path of greatest ascent, requiring the update of all parameters in the opposite direction of the gradient, facilitated by an adaptable step size established by the learning rate (Yamashita et al., 2018).

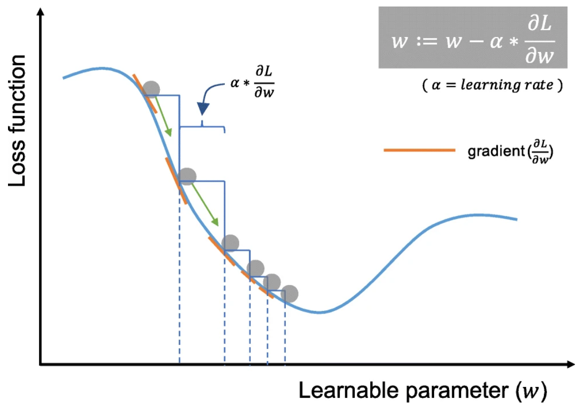

Figure 14: Gradient descent optimization algorithm (Yamashita et al., 2018).

Training ANNs has posed challenges due to the issue of backpropagated gradients that either amplify or diminish with each time step. Consequently, over multiple time steps, these gradients tend to either explode or vanish (LeCun et al., 2015).

Convolutional Neural Network

A type of ANN referred to as Convolutional Neural Network employs layers that apply filters to specific attributes within localized regions of an image. The CNN architecture typically comprises convolutional layers, pooling layers, and fully connected layers. Convolutional layers function as input filters, generating "output images" based on the applied filters. The pixels within the output image are obtained through a linear combination of neighboring pixels situated at corresponding positions within the output image. The size of the filter determines the extent of the neighborhood considered. Convolutional layers can also decrease spatial information by eliminating padding. Pooling layers are utilized to reduce the image dimensions while preserving the most informative data. The fully connected layers, also known as dense layers, are responsible for classification tasks. The underlying concept is that convolutional and pooling layers extract relevant features from the image, which are then utilized by the fully connected layer to classify the obtained features (Alzubaidi et al., 2021).

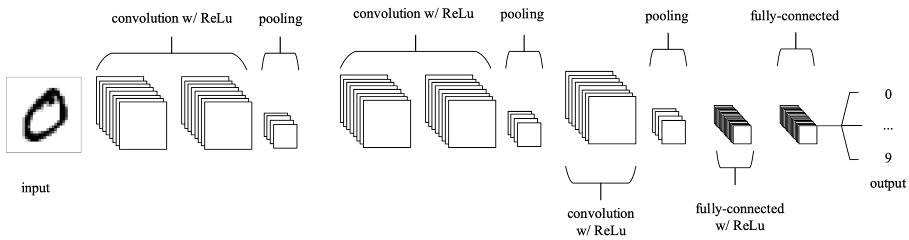

Figure 15: A common CNN architecture form in which convolutional layers are continuously stacked between ReLus, then passed through a pooling layer, and then passed through one or more fully connected layers (O'Shea and Nash, 2015).

CNNs can be utilized for time series forecasting by treating the time series as a one-dimensional sequence and applying the convolutional operation across the temporal dimension (Zhao et al., 2017). This allows the CNN to automatically learn and extract relevant features from the time series data.

Feed Forward Neural Network

A FNN represents a prominent variant of Artificial Neural Network extensively employed in diverse machine learning tasks, including pattern recognition and classification. Its architecture comprises multiple interconnected layers of neurons, with the initial layer receiving input data and the final layer generating the output. The intermediate layers, referred to as hidden layers, play a crucial role in modifying the data as it passes through the network, in one direction from the input layer to the output layer (Benardos and Vosniakos, 2007).

Recurrent Neural Network

A Recurrent Neural Network (RNN) represents a deep learning model widely utilized for the analysis of sequential data. It finds application in various domains, such as language translation, music generation, time-series forecasting, and financial prediction (Vathsala and Holi, 2020; Boulanger-Lewandowski et al., 2012; Qin et al., 2017; Cao et al., 2019).

RNN build upon the fundamental principles of learning employed by FNN, including backpropagation, for adjusting internal weights. However, as depicted in Figure 16, RNN possess a distinctive characteristic not found in FNN: the ability to retain a memory of prior inputs and outputs. This memory enables RNNs to recognize patterns over extended sequences of data. The preservation of this memory is achieved through recurrent connections, which establish connections between nodes in the network that are reused when new inputs are processed. Consequently, the network can "remember" past information and utilize it for making enhanced predictions or decisions (Tu et al., 2019). Notably, Recurrent Neural Networks incorporate an additional layer that facilitates interconnection among inputs, establishing relationships between each input and its preceding counterparts.

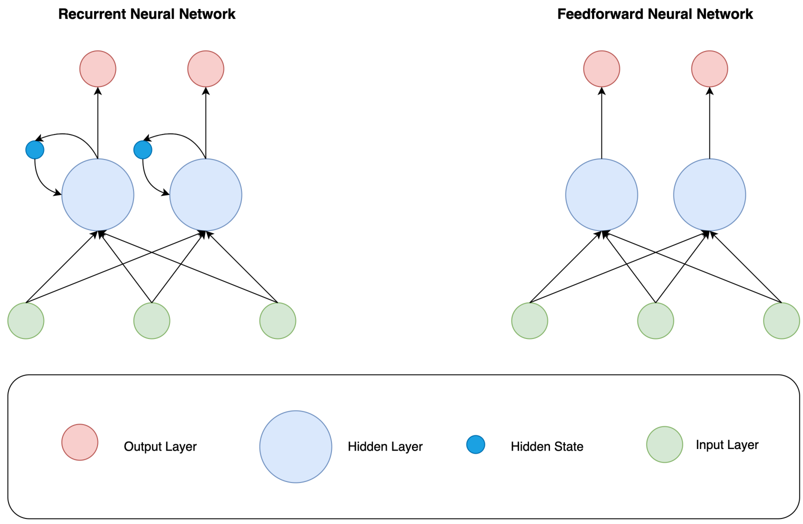

Figure 16: Architecture of Feedforward Neural Network (FNN) and Recurrent Neural Network (RNN)

Long Short-Term Memory

The Long Short-Term Memory (LSTM)) is a variant of RNN that is designed to address the problem of vanishing and exploding gradients, described in Section 2.5, which can occur when training RNNs for long sequences (Hochreiter, 1998). Figure 17 visualizes the RNN and LSTM cell, which will be described in detail.

The LSTM architecture is made up of four interconnected layers (DiPietro and Hager, 2020). The input vector across the cell is used to generate the cell state, which is not subject to any activation functions and does not contain any cell keys. This allows information to pass through without interference. To decide which information is to be removed from the cell state, a sigmoid layer (forget gate) is used to compute a number between 0 and 1 for each number in the cell state. If the value is 1, it is kept, while if it is 0, it is forgotten. The input gate and the tanh gate layers are then used to determine what information is stored in the cell state. The input gate layer determines which values will be updated, while the tanh layer calculates the new candidate value vector to add to the state. These two parts are combined to make a state update, which determines the range of candidate values depending on the time the model chooses to update each value and the number of times each value is updated (Siami-Namini et al., 2019).

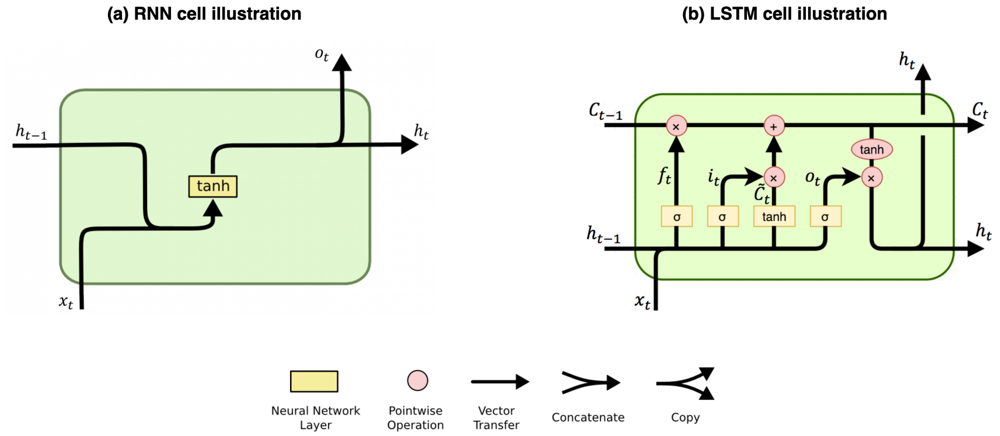

Figure 17: Visualization of the differences in cell structure in an RNN vs. LSTM network (Olah, 2022).

The decision regarding the generated output relies on the passing of the cell state through a sigmoid layer and a tanh layer, respectively. The results of each layer are then multiplied to restrict the number of outputs. Although different versions of this method exist, they share the concept of using the cell state to determine which parts to output. The initial state of the LSTM cell can be set to 0; however, for optimal accuracy, it is best to treat it as a learned parameter. This can be done by starting with random guesses, then using backpropagation to update the prediction errors back to the initial state values, and updating with gradient descent. Once this is done, the parameter should be kept as a learned parameter (Olah, 2022).

Explana-tion	Input Vector	Hidden Layer Vectors	Output Vector	Bias Vector	Parameter Matrices	Parameter Matrix	Activation Functions
Notation	x_t	h_t, C_t	o_t	b	U, W	V	$\sigma, tanh$

Table 4: Explanation of notations in Figure 17

3 Methodology

This section provides an outline of the methodological approach adopted for this thesis. It covers a justification for selecting the specific research methods used to address the research questions. The methods presented include a systematic literature review, an empirical case study, data analysis, and simulation. Lastly, an overview of the research is presented, outlining the overall structure and flow of the thesis based on the research questions and the corresponding methods used.

3.1 Systematic Literature Review

This study applies a Systematic Literature Review (SLR) approach to answering the first research question presented in Section 1.3. The method was utilized in the fall of 2022 as a part of a preliminary study prior to this thesis. The systematic literature review includes a theoretical understanding of how demand forecasting has been used in connection with inventory control, while also identifying research gaps in the domain.

Narrative, or traditional, reviews are widely criticized for being a singular, descriptive report of a writer's contributions to the field, often selected to be included in the field. The implicit bias of the researcher is often chosen (Tranfield et al., 2003). Systematic reviews differ from standard narrative reviews in that they use a procedure that is repeatable, scientific, and transparent. The goal is to reduce bias through extensive searches in the literature of published and unpublished research, as well as by providing an audit record of the opinions, methodology, and conclusions (Tranfield et al., 2003).

Systematic reviews of the literature are a means of providing an objective theoretical evaluation of a specific topic. A systematic review is an effective way to provide an objective theoretical analysis (Hopayian, 2001). Hence, this type of review makes it easier to identify, evaluate, and interpret research in a particular area, by first examining existing concepts, practices, and theories, and then summarizing the state of reproducible research in a specific area. Therefore, this type of review facilitates the identification, evaluation, and interpretation of studies in a given area, first examining existing concepts, practices, and theories and then summarizing the state of reproducible research in a particular area (Rowley and Slack, 2004; Seuring and Muller, 2008).

Literature Identification and Selection

The relevant literature for the SLR was identified using two databases, Scopus and WebOfScience, respectively. Four different experiments were conducted, as indicated in Table 5. Several subthemes emerged during the search process. Some keywords are listed below:

- Time series
- Demand forecasting
- Lot sizing
- Reorder point
- Economic order quantity
- Lead time
- Stock level

After experimenting we ended up with the following used search terms: ("time series" OR "timeseries") AND ("demand" OR "order") AND ("forecast" OR "forecasting") AND ("frequency" OR "lot sizing" OR "reorder" OR "lead time" OR "leadtime" OR "eoq" OR "economic order quantity") AND ("supply" OR "inventory" OR "stock"). The search was limited to articles written in the English language, resulting in a total of 103 articles identified from the two databases, 60 and 43 articles from Scopus and Web of Science. Table 5 shows a systematic search conducted in the databases Scopus and Web of Science, where the results of the searches are presented in a step-by-step manner.

Search	Searchwords	Scopus Results	Web of Science Results	Total
1	("time series") AND ("demand" OR "forecasting") AND ("reorder" OR "lead time" OR "leadtime" OR "eoq")	487	238	725
2	("time series" OR "timeseries") AND ("demand") AND ("forecasting") AND ("dynamic" OR "reorder" OR "lead time" OR "leadtime" OR "eoq" OR "economic order quantity") AND ("supply" OR "inventory" OR "stock")	156	102	257
3	("time series" OR "timeseries") AND ("demand" OR "order") AND ("forecast" OR "forecasting") AND ("frequency" OR "lot sizing" OR "reorder" OR "lead time" OR "leadtime" OR "eoq" OR "economic order quantity") AND ("supply" OR "inventory" OR "stock")	137	98	235
4	("time series" OR "timeseries") AND ("demand" OR "order") AND ("forecast" OR "forecasting") AND ("frequency" OR "lot sizing" OR "reorder" OR "lead time" OR "leadtime" OR "eoq" OR "economic order quantity") AND ("supply" OR "inventory" OR "stock")	60	43	103

Table 5: The stepwise development of search words.

The 103 resulting articles were exported to the EndNote library, where a duplication removal was performed. This resulted in 92 unique articles. Data from each article were extracted and put into an excel spreadsheet for systematic reasons. The spreadsheet consisted of one row for each of the 92 articles. For each row, information such as reference link, authors, title, year of publication were stored. A systematic screening process was conducted by pre-defining questions to separate relevant articles from irrelevant ones. This process included a screening of every abstract and title.

- Is the article related to forecasting?
- Could the objective of the article be relevant in inventory management?
- Does the article and the author seem trustworthy?

The screening resulted in the removal of 53 articles, leaving 60 articles for further analysis. The next step in the literature selection was a full reading (eligibility) of the remaining 60 articles. The following questions were used as inclusion criteria:

- Within the scope of this thesis?
- Is the article of sufficient quality to be included in the systematic literature review?
- Is the article in the subject area?

Consequently, 43 papers were removed. After the assessment, a total of 17 articles were selected for inclusion in the SLR.

With 17 articles remaining, the research topics could be said to have limited available data. As the topic of review has not been widely studied before, forward and backward snowballing was included to add additional potentially relevant articles that were not identified in the initial search. In general, snowballing is a data collection method that uses existing data sources to find additional data sources. Snowballing allows researchers to identify sources of information that they may not have been able to find through keyword searches, enabling them to access a wider range of literature. This means that the review can be more comprehensive and the quality of the results is improved (Wohlin, 2014). The snowballing process involved the search for additional references through the reference lists of articles already included, resulting in 21 additional articles included for further analysis. The resulting total number of articles after snowballing is therefore 38 papers. In Figure 18, a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart summarizes the literature discovery and selection procedure.

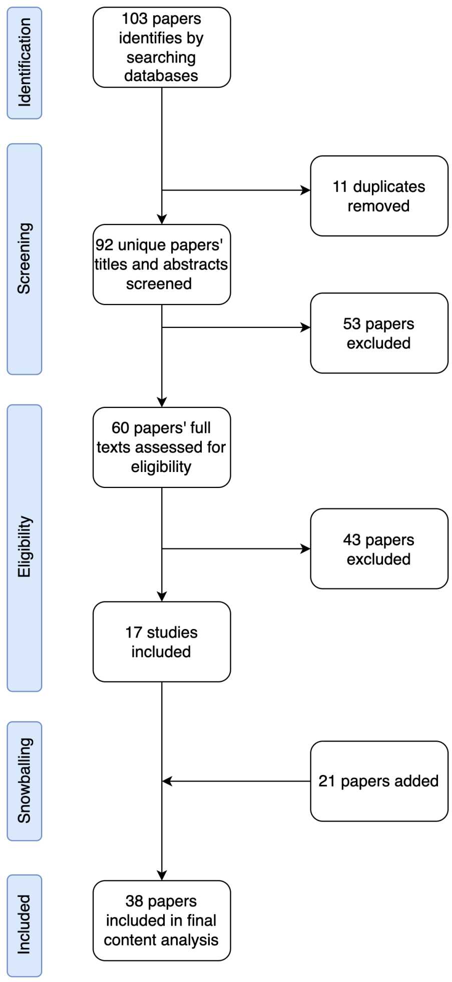

Figure 18: PRISMA flowchart illustrating the process of the literature selection

3.2 Empirical Case Study

This method was used in answering research questions two, three, and four. Gerring (2004) describes a *case study* as an intensive study of a single unit with the aim of generalizing across a larger set of units. For the purposes of this project, the unit is Logistics Center Helse Midt-Norge (LC HMN). Moreover, Flyvbjerg (2011) argues that case studies cannot provide reliable information on the broader class, but may be useful in the early stages of research because they provide hypotheses that can be systematically tested in more cases.

Baxter, Jack et al. (2008) argues that, following the establishment of research questions best suited for a qualitative case study and the selection of the case and its boundaries, the study should be executed. The choice of a particular type of case study design is based on the general purpose of the study. The most suitable case study type is important to define whether you need to describe the case, investigate the case, or compare between cases (Baxter, Jack et al., 2008). Yin (2009) states that descriptive case studies are used to describe the intervention or phenomenon and the real world context in which it occurred. This makes the case study approach applicable to the research activities in LC HMN related to this project.

The case study of LC HMN was conducted with a study of second-hand data. This technique tries to answer questions about the "what," "how", or "why" phenomenon, rather than "how many", or "how much," which are answered by quantitative methods. When the goal is to understand how a community or individuals within it perceive a specific issue, qualitative approaches are frequently applicable (Baruch, 1999).

A variety of different approaches were utilized to gather data and gain insight into the case company. Data were collected through meetings with various personnel from the case company, both at NTNU Trondheim, the case company's office in Heimdal, and remotely. The preparations for the meetings were made by email, specifying the topics of discussion. Relevant questions were prepared in advance of each meeting to initiate an open discussion on the predetermined topic. Primarily, the section leader or SAP engineer attended the meetings on behalf of LC HMN. Follow-up discussions were conducted with LC HMN representatives to guarantee the accuracy of the collected information.

Dodgson (2019) states that all qualitative research is contextual; it occurs within a specific time and place between two or more people. In qualitative research, the researcher's identity is assumed to influence the findings of the study; objectivity is not present. Indeed, often said: "The researcher is the researcher's research instrument" (Dodgson, 2019). Reflexivity can serve as a quality indicator to be incorporated into the process section. Tjora (2021) claims the analysis should be done in collaboration with multiple people. This is how collective reflexivity is created. Christoffersen and Johannessen (2012) claims that researchers often have a preconception about the topic being studied and that this has an impact on how data are interpreted and analyzed. Furthermore, preconceptions influence the information to be highlighted. Two researchers participating in this study will independently analyze the interview results. This dynamic can lead to improved reflexivity when compared to a researcher working alone, as opposing preconceptions can be present in the analysis. However, it should be noted that the presence of shared preconceptions can lead to a degree of bias, thus potentially decreasing the quality of the results.

3.3 Data Analysis

This section provides an overview of the data analysis process conducted in this study, which consists of two main parts: data preparation and quantitative analysis. The data analysis method is utilized for answering the second research question, as a part of the empirical case study. The section highlights the steps involved in data collection and data preprocessing.

3.3.1 Data Preparation for Case Company Data

As described in Section 2.4, data preparation is an important data analysis process and takes place at an early stage. The method starts with data collection and is followed by three major steps: data cleaning, data transformation, and data reduction. This study adapts the described data method as a preparation for the quantitative analysis described in Section 5.

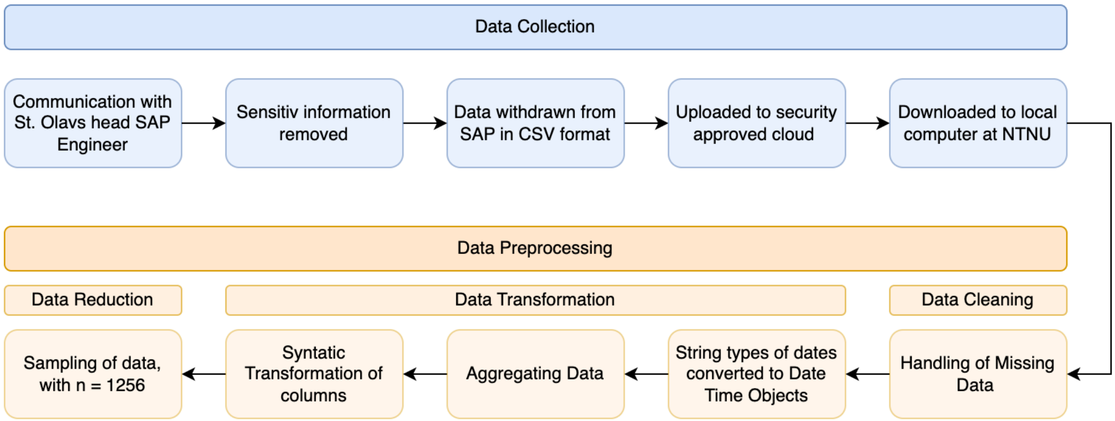

Figure 19: Flowchart describing data preparation method used for LC HMN data

As visualized in Figure 19, the preparation of data involves a series of steps, which in turn can be divided into two main parts: data collection and data preprocessing.

Data Collection

The data collection for this thesis was collected from the Logistics Center Helse Midt-Norge and exported by their integrated ERP system, called SAP. The withdrawal of data type and quantity was determined following a series of meetings with LC HMN, including the SAP Head Engineer. The amount of data to be withdrawn was determined based on what was feasible. Prior to the withdrawal, the data were anonymized to ensure security. A cloud service, which had been pre-approved for security, was utilized to upload and transfer data before downloading it onto a local computer.

Data Preprocessing

The data preprocessing was done using Python, with the use of pandas software library (Pandas, 2020). The first step of data preprocessing was to clean the data. This involved handling missing data, where columns with a high prevalence of NaN values were removed. The next step was data transformation, which was a quite complex part of the preprocessing. All data types were checked and converted, if necessary. Originally, all dates were stored as strings; these were converted to a date-time object. In addition, aggregation was performed based on weekly demand, before the data were restructured based on a syntactic transformation. Lastly, a sample set of the data was made, containing 1256 unique materials.

3.3.2 Quantitative Data Analysis

Quantitative data analysis is a method used in the field of statistics to analyze and interpret numerical data. This type of analysis typically involves the use of mathematical and statistical techniques to identify patterns, relationships, and trends in data sets and to draw conclusions and make predictions based on that information.

Programming Language

Several techniques and theories were used for data analysis. To analyze the data provided, the authors of this study used Python programming language. Python is an interpreted high-level general-purpose computer programming language developed by Guido Van Rossum in the late 1980s (Virtanen et al., 2020). Python is a suitable programming language for data analysis for a number of reasons. First, it is a high-level interpreted language that makes it easy to write, debug, and maintain. It also has a large and active community of users, which has resulted in the development of a wealth of high quality open source libraries and frameworks for data analysis, machine learning, and scientific computing (Fabian Pedregosa et al., 2011). Through quantitative analysis of the data set, Python has been a valuable tool to produce quality results more effectively.

Data Structure

To store and process the data, the *Pandas* Python tool was used. Pandas provide high-level, easy-to-use data structures and data manipulation tools to work with numerical, tabular, and time-series data (McKinney, 2010). Some of the key features of Pandas include its DataFrame and Series data structures, which are designed to store and manipulate tabular data in a way that is similar to a spreadsheet. For the data analysis performed for this study, large Excel files were exported from SAP and then read as a DataFrame structure for easy handling.

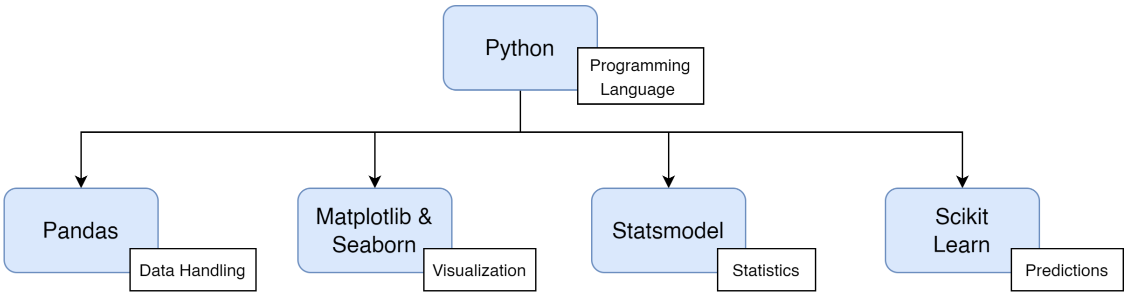

Figure 20: The architecture for performing data analysis.

Visualization

For presenting visuals, the Python packages *Matplotlib* and *Seaborn* were used. The Python package Matplotlib is a powerful tool for creating visualizations of data. For this study, visualizations such as line plots, bar charts, and histograms are utilized. It can also be used to customize the appearance of plots, such as by adding titles, labels, and legends, and applying different styles and color schemes (Hunter, 2007). The Python package Seaborn is a library for creating statistical visualizations. It is built on top of the popular plotting library matplotlib, and provides a high-level interface for drawing attractive and informative statistical graphics (Waskom, 2021). For the quantitative data analysis, Seaborn is applied as a scatterplot and histogram.

Statistic Models

The python package *Statsmodels* was used for statistical analysis in Python. The key benefit of statsmodels is its extensive range of statistical tests and regression models (Seabold and Perktold, 2010). For data analysis for this study, statistical analysis such as seasonal-trend decomposition, autocorrelation, and augmented Dickey-Fuller was applied. With its simple and user-friendly interface, statsmodels contributed to performing complex statistical analyzes and accessing the results in a convenient way.

3.4 Simulation

This method was utilized in order to answer the third and fourth research questions. There are two primary methods for conducting supply chain analysis: analytical and simulation-based. The analytical approach involves deriving optimal solutions through calculations, but it heavily relies on assumptions and available formulas. On the other hand, simulation-based analyses can effectively capture the intricacies of complex systems, ensuring that any modifications to the input are connected to a corresponding set of outputs (Holden, 2017).

To be able to visualize and compare the performance of various replenishment policies, simulation was utilized as a valuable tool. Simulation allows testing of any aspect of a desired change without incurring significant costs to implement the change (Banks, 1999). As both inbound and outbound logistics can be represented as a two-step process, it was preferred to code the simulation in the programming language Python. There are various useful tool kits for simulation, including the utilized Python package Simpy (2023), which is a process-based discrete-event simulation framework based on Python.

In the context of inventory simulation, Discrete Event Simulation (DES) encompasses the discrete variables involved, such as order points and inventory levels. These variables are typically expressed as whole numbers or integers, representing distinct events or states within the simulation model. The process-oriented nature of DES signifies that each activity in the simulation is modeled as a process. There are multiple "application-specific threads" and a "general thread" responsible for managing the event set (Matloff, 2008).

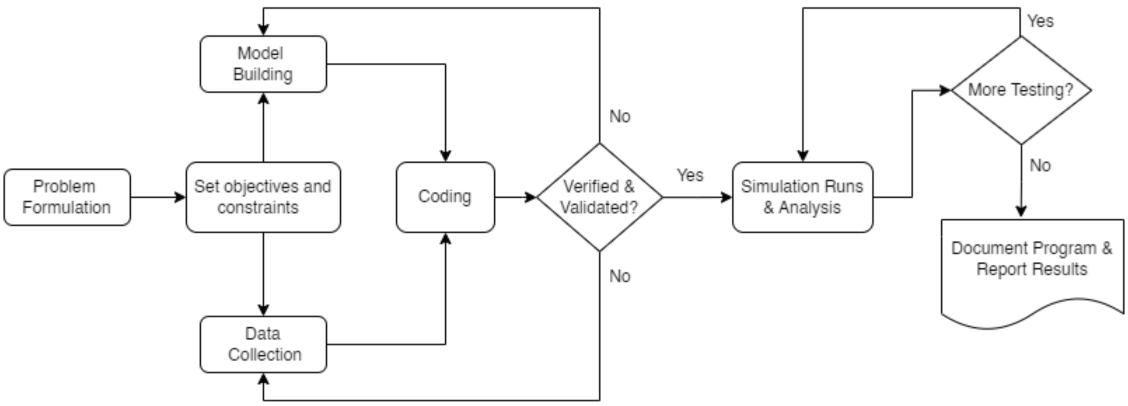

Figure 21: The architecture for the discrete event simulation.

In Figure 21, the architecture for the procedure used to construct the simulation model can be seen. Several assumptions were made in order to represent the inventory management system. One disadvantage of building such a simulation model is the complexity of accurately representing the real system, including factors that are not captured in the data, such as human influence (Banks, 1999).

Following extensive testing and evaluation of both the data and the simulation model, it will be possible to monitor the critical performance metrics to evaluate the effectiveness of a replenishment policy. The resulting inventory levels for a given demand can be used to calculate measures such as the number of orders for the given period of time, the average inventory level, the unfilled demand, and the service level for the given replenishment policy.

3.5 Research Overview

The research overview, depicted in Figure 22, outlines the flow of the thesis. It commences by emphasizing significant motivating factors discussed in Section 1.1. These motivations subsequently give rise to four interrelated research questions (RQs) formulated through iterative research and discussion.

Different methodologies were employed to address these research questions. The first research question involved a systematic literature review that explored demand forecasting techniques for inventory management. The second, third, and fourth research questions employed specific approaches as part of a comprehensive empirical case study method. The second research question used data analysis through clustering and categorization techniques to gain insights into demand patterns and identify materials for improvement. Research questions three and four were centered around conducting multi-scenario analysis utilizing simulation as the chosen method. The figure presents five strategies for comparison, including the existing policy, potential improvements, and various forecasting models. Outcome measures such as inventory level, service level, holding cost, and forecasting accuracy were considered.

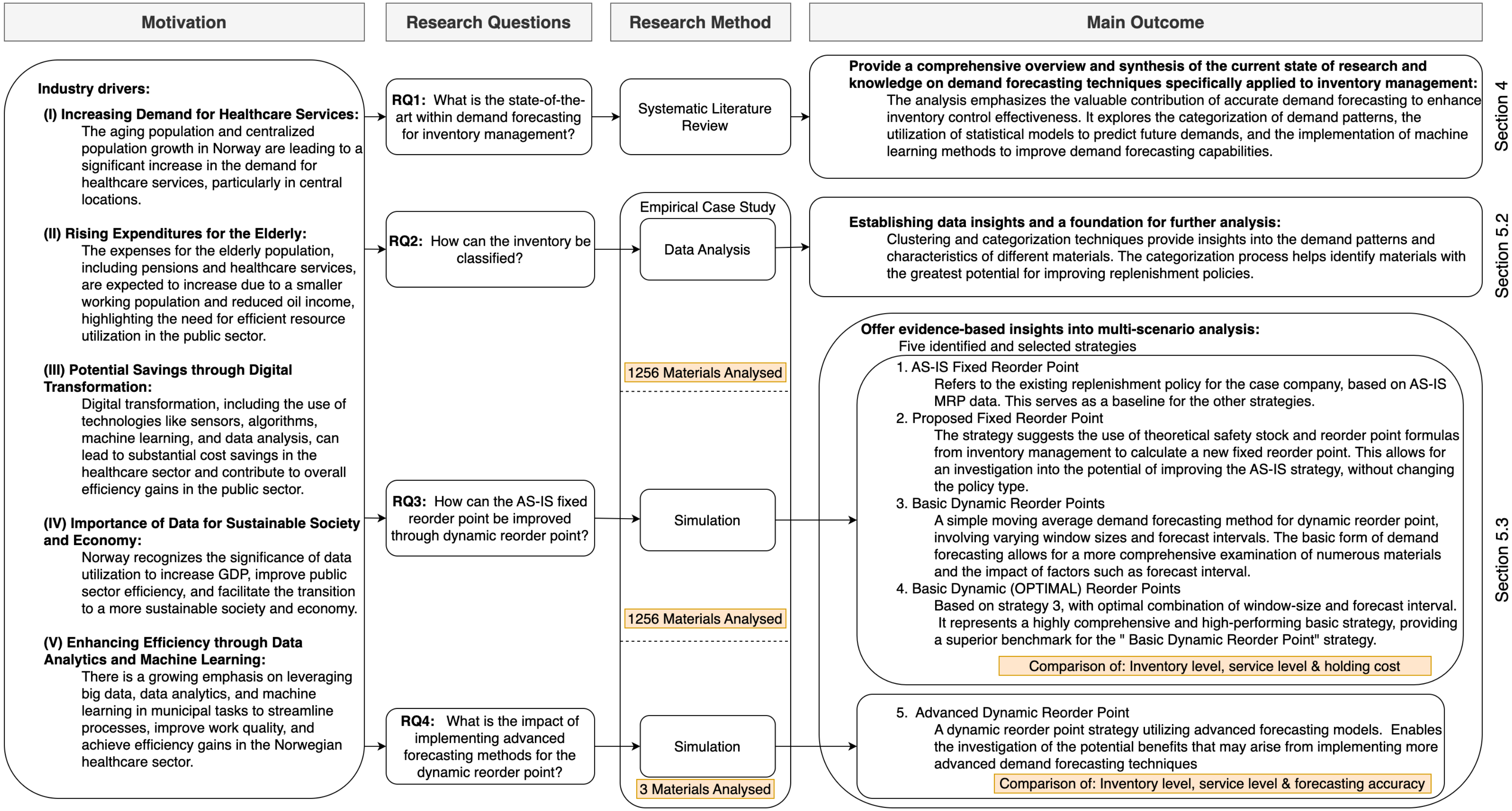

Figure 22: Overview of the research flow

4 Systematic Literature Review

This section reveals the results of the systematic literature review. The review was performed in the fall of 2022 as a part of a preliminary study prior to this thesis. Section 4.1 provides a descriptive analysis of the field under investigation, and the resulting articles from the systematic literature review are presented in six tables (Appendix B). In Section 4.2 the results of the content analysis are presented, utilized with a deductive approach, where the categories were taken from existing theories and defined before analyzing the articles.

4.1 Descriptive Analysis

To obtain a general understanding of the analyzed literature, a primary mapping was conducted based on a yearly publication analysis, see Figure 23. This approach allowed insight into the distribution and evolution of research on this topic over time. The graph will show a steady increase in the number of articles over the course of about 30 years.

The number of published articles on demand-based inventory management, as seen in Figure 23, has shown an increasing trend in recent years, indicating a growing interest in the topic within the academic community. This trend is particularly evident from 2015 to 2017, during which a high level of activity was observed in this field. However, prior to 2006, few significant contributions to this area of research were made.

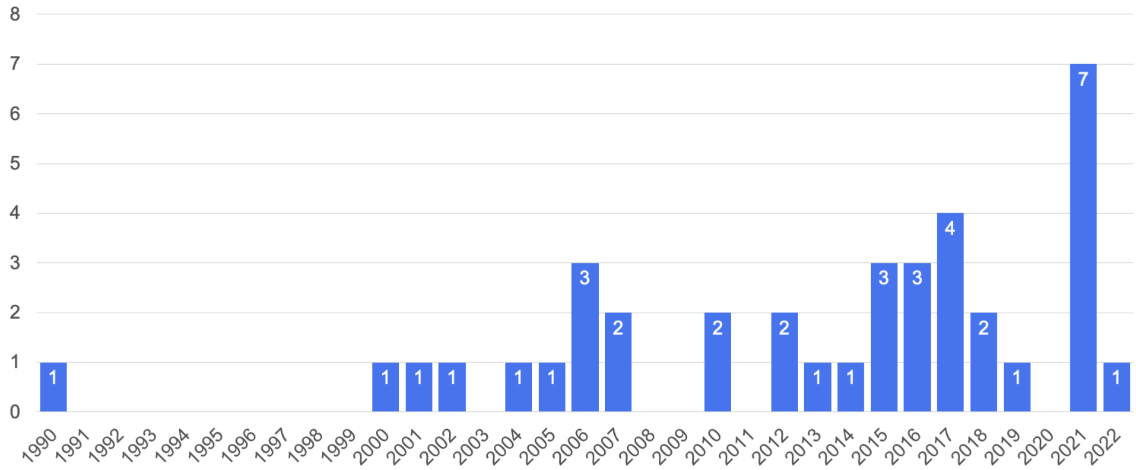

Figure 23: Year-wise publication of articles

As seen in Figure 23 there are many publications in 2021. This is due to the snowballing that was carried out and which was included in the SLR to add more relevant articles. Figure 24 shows a more detailed distribution with respect to the different groups. The three different groups are the following;

- Machine Learning - Represents the articles using the machine learning methods exclusively
- Statistical - Represents the articles using the statistical methods exclusively
- Machine Learning ∩ Statistical - Represents the articles using both machine learning and statistical methods

Figure 24 shows that the articles based on an Machine Learning (ML) method are represented from 2012 and later. The articles based on both statistical and Machine Learning (ML) methods have a more spread distribution, going back to 2001. The articles based on statistical methods are spread throughout the whole scale. For all three groups, there is an increase over the years, but it seems like Machine Learning is becoming more and more relevant, taking over for statistical methods.

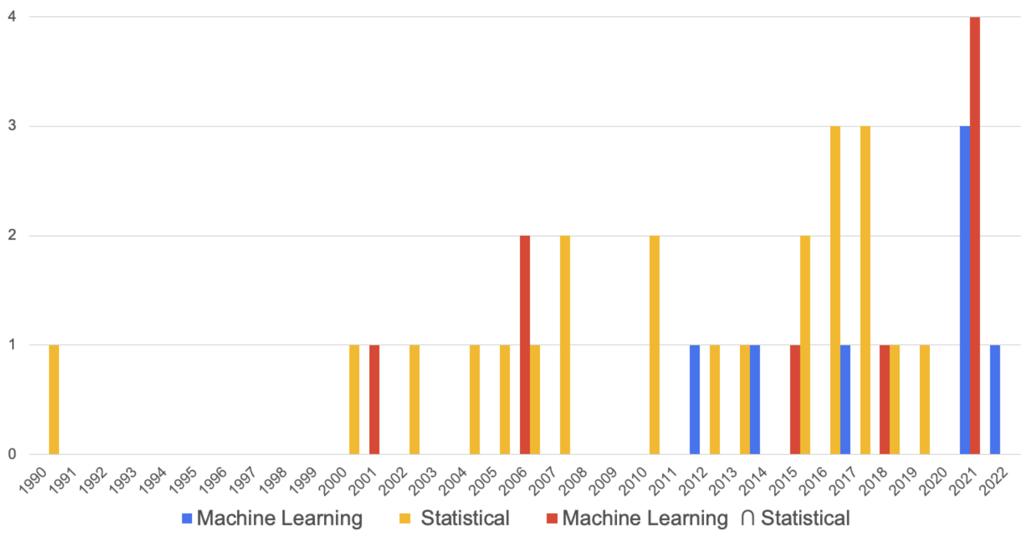

Figure 24: Distribution of machine learning and statistical approach in articles in the publication year

The cake diagram shown in Figure 25 is a visual representation of the proportions of different groups of methods used within the selected articles. The groups referred to are Machine Learning (ML) and statistical methods, along with their intersection. The diagram shows that “Machine Learning" makes up 18% of the total, while “statistical" makes up 58% of the total. The diagram also shows that there is an intersection between the two categories, which represents 24%.

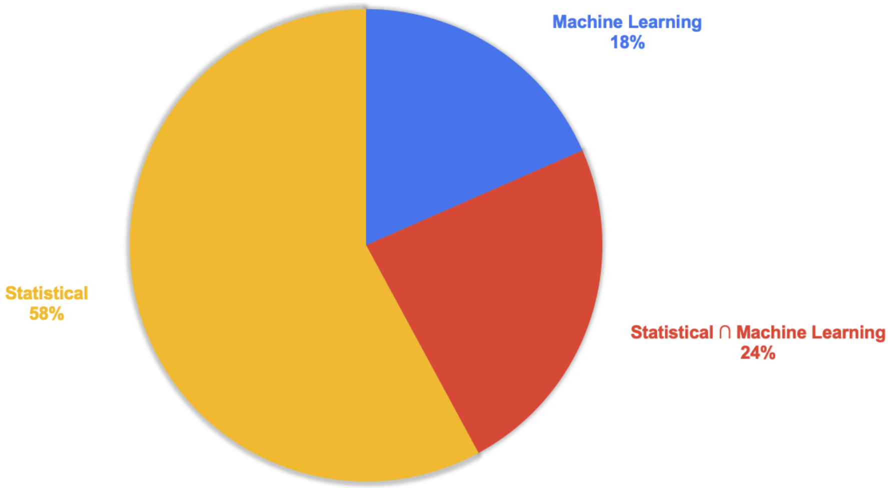

Figure 25: Distribution of machine learning, statistical and the two combined

Figure 25 illustrates a cake diagram that provides a visual representation of the proportions of different methods within the topic. It suggests that while a significant portion of Machine Learning research involves statistical methods, statistical is the most used method in the selected articles.

4.2 Content Analysis

This section presents the findings of the content analysis conducted as part of the systematic literature review. The articles categorized and reviewed can be seen in Appendix B. The analysis involved a thorough examination and evaluation of the articles included in the review. Through content analysis, these articles were systematically categorized and analyzed to enhance our understanding of the research area. This process involves exploring various aspects of the articles, such as research objectives, methodologies employed, key findings, and theoretical frameworks used.

4.2.1 Inventory Control & Policies

Determining appropriate procurement lot sizes and buffer inventory levels under uncertain market conditions while guaranteeing a particular degree of leanness and service level is one of the most fundamental problems in warehousing and manufacturing companies (Piasecki, 2009). In terms of leanness, a previous study by Hofer et al. (2012) found that increasing inventory leanness via underlying lean initiatives can result in improved profitability due to lower operational costs. The results of the report demonstrated the negative correlation between stock levels and the company's financial performance.

Replenishment policies can either follow a periodic order policy, or restocking can take place when needed, also called a continuous review policy (Section 2.1.2). The first is characterized by having a specific cycle time for each raw material bought from each supplier (Santos et al., 2022). According to Bhagwat and Sharma (2007), the overall order cycle time might affect supply chain response time and thereby directly influence customer satisfaction. The other restocking strategy, also known as the Order-Up-To (OUL) policy, is characterized by periodic stock level reviews, with the quantity ordered equal to the goal level minus the reorder point. It is seen as a riskier strategy that is more exposed to stock shortages. Clausen and H. Li (2022) considers an OUL policy and suggests a dynamic inventory model with applications of big data and machine learning.

While several approaches may be used for optimal replenishment, various sectors frequently demand various replenishment approaches. Santos et al. (2022) states that large product portfolios in highly dynamic situations may require flexible order cycles or even non-cyclical restocking practices for businesses to be adaptable to changing situations.

The authors of S. Chopra (2016) state that periodic review policies require more safety stock than continuous review policies for the same level of product availability. In addition, whatever forecasting method is used, Order-Up-To policy will always result in the bullwhip effect (Dejonckheere et al., 2003). In Order-Up-To policy, the bullwhip phenomenon is unavoidable when forecasting is necessary; it is the price to pay to forecast unstable demand and to detect trends.

Figure 26: Bullwhip effect

The bullwhip effect, where order fluctuations worsen as they go up the supply chain from retailers to wholesalers to manufacturers to suppliers, is one result of a lack of supply chain coordination (S. Chopra, 2016). The supply chain's input on demand is distorted by the bullwhip effect, with each stage's estimate of demand being different. Most importantly, smoothing replenishment rules have been recognized as the most powerful approach to avoid the bullwhip effect (Dejonckheere et al., 2003). However, various studies have found that mitigating the bullwhip effect can increase inventory volatility, resulting in low service levels (Costantino et al., 2016).

L. Wang and H. Chen (2022) purposed an algorithm that calculated the optimal OUL for every item for a given time. They used a periodic review policy to replenish multiple items at the same time. The inventory status of each item was evaluated once per time unit and the inventory replenishment of each item is regulated by the item-dependent policy. They also created a policy that combined a continuous and periodic review strategy. The requirement to trigger the replenishment of this strategy was that the aggregated total demand for all products from the previous order reaches a given number of units or a given number of time units have passed.

Due to the challenges in predicting stockouts, incorporating service-level constraints are frequently utilized for inventory models (Tsai and Liu, 2015). Özsen and Thonemann (2015) took into account an expediting policy with service-level requirements and created a successful linear optimization technique to determine the ideal parameters of the policy.

4.2.2 Demand Forecasting

According to Dolgui and Prodhon (2007), inventory control still presents numerous untapped opportunities, particularly when demand and time variation are co-evaluated. Accurate forecasting of consumer demands and procurement lot-sizing plays a crucial role in addressing complex industrial-level stock management challenges (Dolgui, Louly et al., 2005). The characteristics of demand data significantly influence the choice of an appropriate forecasting approach (Vercellis, 2011). This section presents the key findings from the systematic literature review in the domain of demand forecasting.

Classification of Demand

In most occurrences of an inventory, there will be hundreds or thousands of different products. Wenhan Fu (2018) experienced that if an inventory holds several products, it is not practical to conduct deep research to fit the demand pattern for each product, which causes a lot of computational costs that are not efficient. Therefore, some categorization methods have been built to group products and invented to identify suitable forecasting methods for classification results. Syntetos, Boylan et al. (2005) used mathematical proof to quantify the classification matrix. They proposed two cut-off values to calculate different demand patterns. The demand intervals are determined by the inter-demand interval ($p = 1.32$), and the other coefficient is the demand variation ($CV^2 = 0.49$). A high value of p indicates a low frequency of demand and a high value of CV^2 indicates a high volatility of demand.

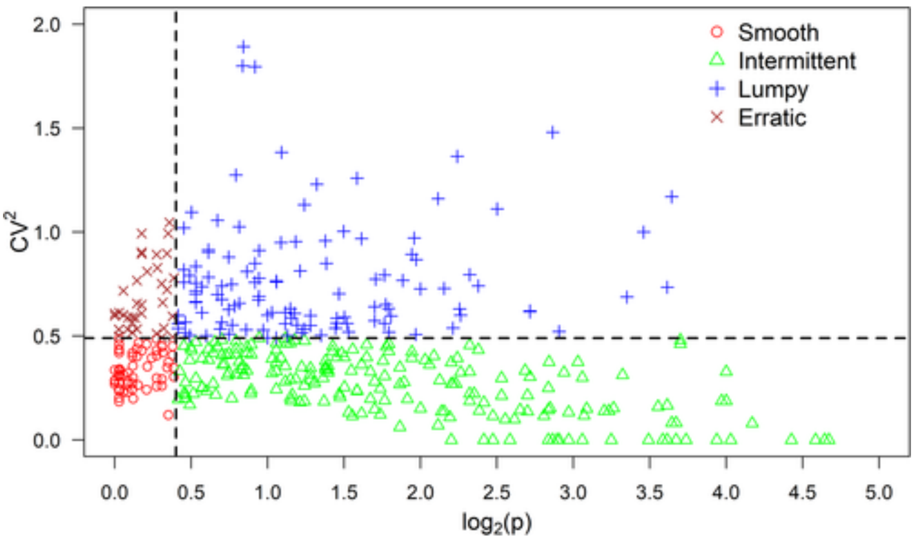

Figure 27: Categorized demand observations based on the Syntetos method (Chuang et al., 2021)

Using these values, the demand pattern of a product can be classified into four types of demand; smooth, intermittent, erratic, and lumpy. An example of this is Chuang et al. (2021), which classified its demand series based on the metrics proposed by Syntetos, Boylan et al. (2005). Chuang et al. (2021) categorized

426 weekly demand series into the four mentioned categories of demand. Figure 27 shows the distribution of demand types, and clearly one can observe the highly skewed demand distributions that will be difficult to predict. This distribution violates the assumption in standard inventory formulas, which assumes that lead time demand is represented by a normal distribution (Van der Auweraer et al., 2019).

Applications of Statistical Models

It is important to be aware of the impact of assumptions that are made when performing research. Pollack-Johnson et al. (1990) states that it is important to test the assumptions taken, reconcile them with known theory (e.g. laws of supply and demand), and avoid biases due to one's affiliation. The study carried out by Barrow and Kourentzes (2016) investigated whether a weighted combination of forecasts led to an improvement in the distribution of forecast error with respect to the properties of normality and unbiasedness. The extent of impact the distribution had on the calculation of the safety stock was investigated by comparing the one-step ahead average safety stock with a combination of different statistical forecast methods. The models that were combined were exponential smoothing, AR, ARIMA, Theta method, and MAPA. By utilizing the empirical distribution of the forecast error, Barrow and Kourentzes (2016) were able to overcome the limitations of the theoretical inventory calculations, which fail to account for aberrations from normality and any covariance between forecast errors of cumulative demand over lead time.

According to R. Snyder (2002) who considered different ways to forecast sales of slow- and fast-moving car parts, when transactions are small, the discrete nature of demand can become important. The use of a discrete probability distribution defined across the entire number, as well as exponential smoothing updates to its mean, is problematic, as the associated simulated data may exhibit weird behavior (Grunwald et al., 1997). Thus, a skewed distribution may require proper modeling of demand data.

Harvey and Ralph D. Snyder (1990) investigated non-stationary models for exponential smoothing and derived the connected formulas for the variance of lead time demand under different conditions. Exponential smoothing is often applied for inventory forecasting and the method usually relies on formulas such as the variance of demand given the lead time. Since the variance depends on the assumption of a stable demand process, it is inconsistent in applications such as exponential smoothing. Harvey and Ralph D. Snyder (1990) holds the view that exponential smoothing models have a tendency to underestimate the required safety stock, and therefore targeted service levels can be at risk.

Kourentzes et al. (2020) claims that reducing the bias of the forecast is more important than accuracy since forecasts and variance for inventory management are used to make decisions regarding reordering and safety stock. A prediction that is highly accurate in-sample but skewed out-of-sample might have a negative impact on inventory performance. On the other hand, even if the forecast function has missing terms, a less accurate forecast that remains reasonably unbiased in the out-of-sample might be preferred. Similarly, because it does not minimize the approximation error, this strategy is less likely to overfit, even when unneeded terms are included in the prediction function.

A way to better understand the distribution of the time-series data is to apply the seasonal-trend decomposition using STL to decompose the time series into three components; seasonality, trends, and residuals. The trend component is a long-term pattern that indicates the time series' growth or decreases during the observed period. Seasonality is a pattern of change that repeats itself across time at particularly regular intervals smaller than a year, such as weekly, monthly, or quarterly. N. Li et al. (2021) states that the main advantage of performing an STL is the flexibility to handle different kinds of seasonality, strong estimates of trends and seasonal components that are not affected by outliers, the ability to break down time series with missing values and perform fast calculations. The rate of seasonal change and the smoothness of trend cycles can be controlled by two key parameters; the trend cycle window and the seasonal window.

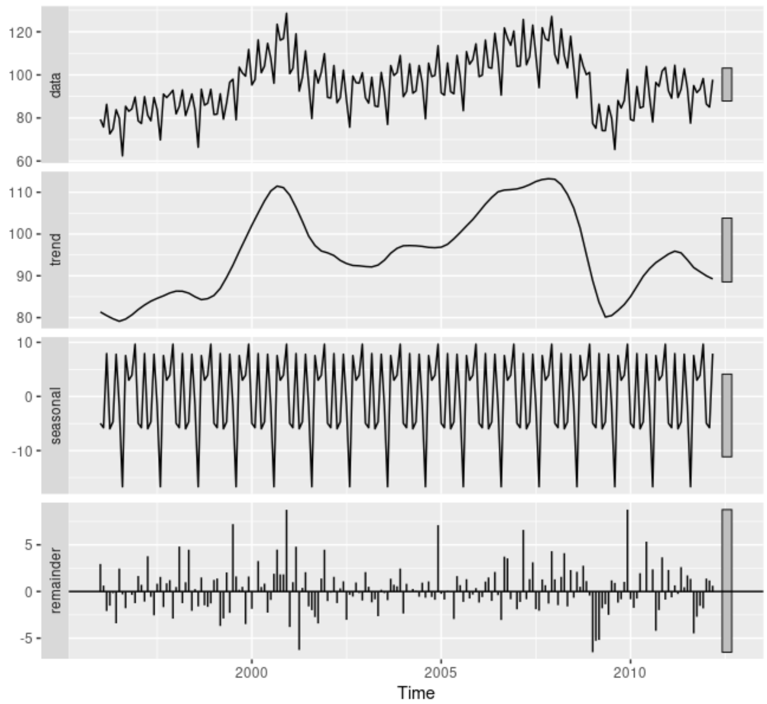

Figure 28: An example of decomposition of a time series.

On the other hand, STL has one critical important drawback: STL depends on its own history. STL is not capable of capturing structural changes, such as nonlinear patterns, that are associated with explanatory variables.

Willemain et al. (2004) developed a bootstrapping forecast model that did not require independent and normal distributed demand. Bootstrapping is a procedure that re-samples the data into many simulated samples. Willemain et al. (2004) compared the results of bootstrapping with exponential smoothing and the Cronston method. When the aim was to anticipate the complete distribution of lead-time demand, the study found that the Cronston method did not give an overall benefit over exponential smoothing. However, for short lead periods, the bootstrap clearly outperformed exponential smoothing. However, Willemain et al. (2004) has, on the other hand, not taken into consideration that the bootstrapping method resamples from the initial sample, thereby outliers may skew the estimates from the resamples.

X. Zhang (2007) considered a Order-Up-To (OUL) level policy and examined the problem of inventory control when the demand variance reveals temporal heteroskedasticity. They decomposed the forecast error variance for lead-time demand into two additive components; forecasting-error variance under homogeneous demand and forecast error variance due to temporal heteroskedasticity. X. Zhang (2007) stated that the operations management community has paid little attention to variability at higher demand moments. In fact, X. Zhang (2007)'s study shows that if a service-level approach is adopted for managing inventory, the actual service level can deviate from the desired one if the volatility in the demand variability is not recognized. Furthermore, ignoring temporal heteroskedasticity can increase inventory costs by up to 30% when demand autocorrelation is highly positive.

Tasdemir and Hiziroglu (2019) performed an experiment to optimize the raw material inventory management system of a manufacturer of oriented strand board (OSB). OSB does fluctuate in demand as a seasonal product, driven primarily by the construction sector. Tasdemir and Hiziroglu, 2019 made use of the Winters Seasonal Multiplicative Forecasting Model (Winters, 1960) and Regression Based Forecast Modelling (Hyndman and Athanasopoulos, 2018) as a consequence of missing data for specific periods. Tasdemir and Hiziroglu (2019) proposed a fixed-period quantity dynamic lot size method that resulted in an 80% reduction in annual ordering cost, a 62% reduction in safety stock holding cost, and a 57% reduction in the cost of the raw material inventory management system.

Applications of Machine Learning for Demand Forecasting

The goal of the upcoming fifth version of an industrial revolution is to be more human-centric or to put more emphasis on the interaction between humans and machines, both physically (e.g. robots) and computationally (e.g. artificial intelligence-based decision support) (Breque et al., 2021). In addition, two more key pillars of this industrial revolution are industrial resilience and environmental responsibility. These pillars are frequently exploited by computational strategies. Forecasting is a central objective of statistical modeling and has been widely used in a variety of domains. It can mitigate the uncertainty of the future through the use of existing data to project uncertain future (Karimnezhad and Moradi, 2016), and by exploiting the predictions companies can help themselves choose the most sustainable option (Santos et al., 2022). Alicke (2022) states that the implementation of AI-enabled supply chain management has allowed early adopters to improve inventory levels by 35%.

Sillanpää and Liesiö (2018) found that by modeling consumer demand with distributions for replenishment orders in retail, the accuracy of replenishment order forecasts improved significantly and could result in substantial cost savings. Similarly, Kim and Jeong (2018) considered a mass-producing factory that had an excessive amount of material. The factory had products that did not respond to demand, resulting in high inventory maintenance costs. Kim and Jeong (2018) proposed an ARIMA model that predicted future demand in the temporal variability or seasonal element and created an EOQ-based demand forecasting model.

An approach to capture the dynamic behaviors of the underlying processes when analyzing time series is applying deep learning methods. Pacella and Papadia (2021) did overcome the statistical complexities by making use of the Long Short-Term Memory (LSTM) network for demand forecasting in supply chain management. Their results indicated that an LSTM effectively models the non-linearity of time series, and surpassed the traditional linear forecasting method in performance.

There are many different deep learning algorithms to make use of for forecasting demand. Carbonneau et al. (2008) applied tools such as Neural Network (NN), Recurrent Neural Network (RNN), and Support Vector Machines (SVM). The performance of these methods was compared to standard baseline approaches such as naive forecasting, moving averages, linear regression, and time series models. For their analysis, they used two sources of skewed demand data. The first comes from a simulation of the extended supply chain, while the second comes from the estimated value of the new orders received.

The research of Carbonneau et al. (2008) suggested that RNN and SVM were the most accurate forecasting techniques compared to moving average and Naive Bayes, which were among the worst performers. However, statistical analysis of the findings revealed that there were no statistically significant differences in forecast accuracy between RNN, SVM, NN, and MLR. Hence, for both datasets, the increases in forecast accuracy of RNN and SVM over MLR were minor. Furthermore, on both datasets, MLR outperformed neural networks. The result can be explained by the problem of the neural network overfitting. The RNN performed better due to its ability to detect temporal patterns. SVM performed better on training sets that did not extend to the testing set. This demonstrates the limits of the SVM in reaching genuine generalization.

Moroff et al. (2021a) selected six types of models to forecast demand. In terms of forecasting errors, two models from the fields of statistics, machine learning, and deep learning were studied. Through their study, Moroff et al. (2021a) showed that the different models have various qualities based on the demand pattern. The deep learning model Multiplayer Perceptron (MLP) was the best performing model, yet the statistical models like Holt-Winters were reasonable for specific types of demand. They experienced that the potential of machine learning and deep learning methods can be increased by including data that highly correlate with the demand pattern (e.g. weather forecasts).

Hybrid Models

One approach to leveraging the strengths of both deep learning methods and high-performing statistical methods is to combine the two as a hybrid model. Arunraj and Ahrens (2015) states that there is no universal forecasting model that can be applied to all kinds of problems. For example, it is unrealistic to expect a model that forecasts the price of a product with greater accuracy to also predict demand for the same product. As a result, forecast accuracy can only be enhanced by integrating and using two or more models with varied capabilities rather than a single unique model with restricted capability.

Integrating and combining several machine learning models is called ensemble learning. Chuang et al. (2021) studied a semiconductor producer and showed that one aggregated model outperformed many of the individual time series models. Chuang et al. (2021) combined ensemble learning with cross-validation, which is a useful technique to assess how the analysis generalizes to the given data set. Applying cross-validation in any form will mitigate the chance of over-fitting (Makridakis et al., 2020). The combination of ensemble learning and cross-validation is potentially the new standard within machine learning (Bojer and Meldgaard, 2021).

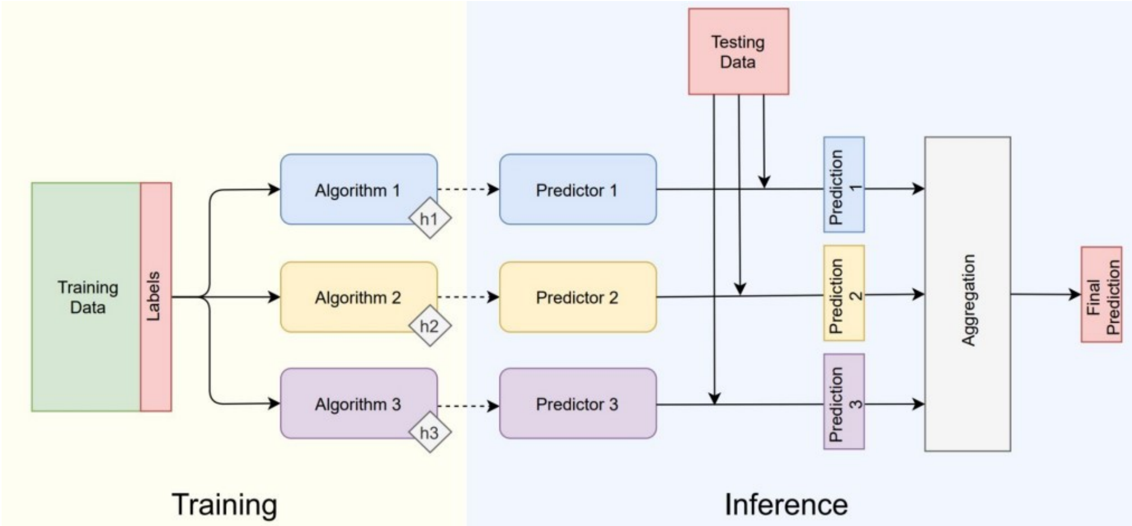

Figure 29: An example of ensemble learning

The study conducted by N. Li et al. (2021) utilized the statistical method STL and extracted the linear and seasonal components, and then combined this with a machine learning model called *eXtreme Gradient Boosting (XGBoost)*. This type of hybrid forecasting model handles changes in trends and seasonality, non-linear patterns, and correlations among predictors. N. Li et al. (2021) had great success in applying the hybrid model for the ordering strategy, reducing the order frequency by 60% and reducing the level of inventory by 40%, with a low risk of shortage. Similarly, Hua and B. Zhang (2006) proposed a hybrid model consisting of two machine learning methods, *logistic regression* and SVM, to predict the intermittent demand for car parts. The model worked in such a way that the occurrence of non-zero demand was first predicted and then the lead time demand was forecasted. The hybrid model outperformed traditional statistical methods such as exponential smoothing, Cronston's method, and bootstrapping method for all lead times.

These results are similar to those reported by Tang and Ge (2021), who proposed a hybrid model consisting of Convolutional Neural Network (CNN) and an Long Short-Term Memory (LSTM). Tang and Ge (2021) stated that the CNN has technical defects in terms of gradient disappearance and gradient explosion. The disadvantages of the CNN model are complemented by the LSTM model, which can handle the problem of gradient disappearance. In contrast to studies conducted by Hua and B. Zhang (2006), Chuang et al. (2021) and N. Li et al. (2021), Tang and Ge (2021) discovered that including additional variables such as transit warehouse inventory and material attributes resulted in better forecasts than the model based on a single variable. These discoveries are similar to those reported by Moroff et al. (2021a).

5 Empirical Case Study: Logistics Center Helse Midt-Norge

This section provides an account of the findings obtained from the empirical case study conducted. Initially, the characteristics of the case company are presented, encompassing an introduction to the company itself and a detailed examination of the current state (referred to as the AS-IS situation). Subsequently, the results of the data analysis conducted on the case-company data are presented, which includes data collection, preparation, and analysis process. Finally, the outcomes of a multi-scenario analysis are presented. The results presented were derived from a comprehensive analysis of the collected data and a Python-based simulation model. The simulation was executed extensively, with a total of 62 800 runs. The identified and selected strategies examined are visually represented in Figure 3.5.

5.1 Characteristics of The Case Company

Logistics Center Helse Midt-Norge (LC HMN) is organized under Helse Midt-Norge (HMN) and is responsible for the storage and supply of all non-medical consumables to hospitals, medical treatment centers, and doctors' offices in the Midt-Norge region. Examples of goods in the warehouse are masks, gloves, toilet paper, etc. These are non-perishable goods, where the goods have a very long shelf life (more than two years). The rotation of products and the current turnover rate of approximately 6 months ensure that the expiry date of the goods will never be a concern, as they are managed according to a first-in, first-out (FIFO) policy.

Located in Heimdal, a municipality in Trøndelag, Logistics Center Helse Midt-Norge serves as a dual-purpose facility, operating as both a warehouse and a cross-docking station. Throughout the day, goods are delivered from suppliers to the warehouse, where they are sorted into designated sections based on criteria such as hygiene. Four times a day, vehicles depart from the warehouse to transport goods to St. Olav's Hospital, which is owned by Helse Midt-Norge, a healthcare organization under the ownership of the Norwegian state. St. Olav's Hospital serves regional and national roles.

Helse Midt-Norge owns both the hospitals and the warehouse, establishing a shared objective of effective hospital operations. Given the hospital setting, the definition of successful operations encompasses specific requirements. The warehouse's primary focus is not solely on maximizing profit but prioritizing safety and service at the highest level. Emphasis is placed on delivering the required goods to the hospitals and their sub-departments promptly.

5.1.1 Control Model

The model visualized in Figure 30 shows how warehouse procurement planning has a large effect on both the product flow and the information flow of the warehouse. For this reason, the model is of great relevance for inventory management at LC HMN. The scope of this thesis is marked in red, clarifying what part of warehouse management is included in the scope of the topic. Due to the stated connection between procurement planning and the rest of warehouse flow, the entire model is relevant to the thesis.

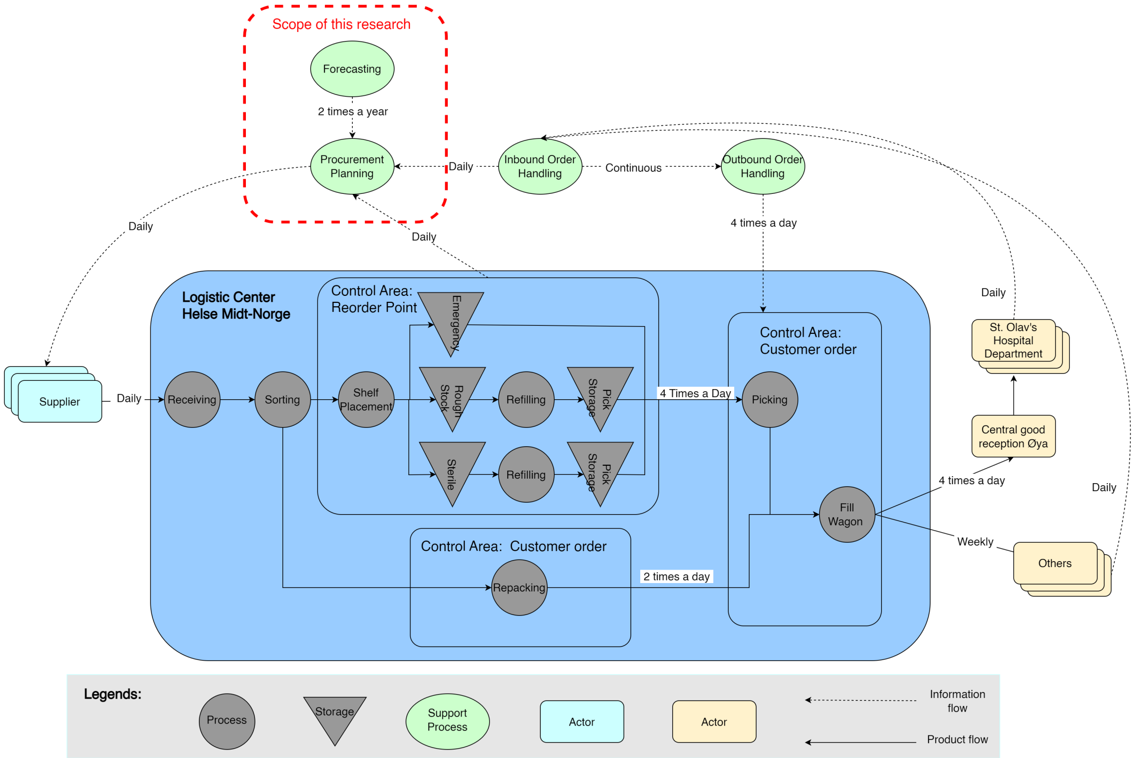

Figure 30: Control model describing the AS-IS situation of Logistics Center Helse Midt-Norge

The model flow consists of two components: a product flow component and an information flow component. The product flow is represented by solid lines, while the information flow is represented by dashed lines. Following, important aspects of the two components will be analyzed.

5.1.2 Product Flow

Within a warehouse setting, the term "product flow" pertains to the progression of goods and products, commencing from their arrival at the warehouse and concluding with their departure for delivery to customers (Arnold, 2017). The process of product flow in a warehouse is essential for effectively managing goods and maintaining sufficient inventory levels to meet customer demand. In this section, this study will examine the various stages of product flow within the LC HMN, including the receipt of goods, storage, organization, and picking and packing for shipment. Understanding the key elements of product flow in the warehouse is essential to identify potential areas for improvement within our scope of inventory management.

Supplier

It is important to note that the inbound lead time specified in the supplier agreements should be realistic and accurately reflect the actual lead time. For LC HMN, it appears that the stated lead time of three days is not consistent with the actual delivery time, which can range from a few days to several months. This inconsistency can cause confusion and potentially lead to problems in meeting deadlines and fulfilling orders. In general, the warehouse receives goods daily, up to many times a day.

Storage

There are three underlying layers of storage; emergency, rough and sterile. The goods are evenly moved from the sterile and rough storage to picking storage, from where picking takes place. Items in the emergency warehouse are not moved to the picking area but are picked directly if necessary. Not all items are stored in the emergency area. Selection varies depending on factors such as season. The safety stock is included in the rough storage and sterile storage.

Cross-Docking

In addition to the regular warehouse, LC HMN is also equipped with a cross-docking system where delivered goods are repacked onto different wagons. There is a unique cart for each department at St. Olav's Hospital. There is no storage of goods in this part of the warehouse. The handling of the goods during cross-docking is continuous, as the goods are received throughout the day. The fully repackaged wagons are collected and sent twice a day, with departure at 10:00 and 14:00.

Shipping

Within the shipping and coordination department, it is ensured that each pre-picked wagon is sent with the correct car. To St. Olav's departments, there are four departures each day, five days a week:

- 08:00 AM
- 10:00 AM
- 12:00 PM
- 14:00 PM

Delivery to St. Olav's departments goes through a central good reception at the Department of Øya before being further distributed to the departments. The delivery for the actors marked as "others" in the AS-IS control model, most deliveries are sent weekly by a courier.

5.1.3 Information Flow

In a warehouse, information flow refers to the movement of data and information related to the warehouse's operations and inventory (Arnold, 2017). This includes information about incoming and outgoing goods, inventory levels, and other critical aspects of the warehouse's function. Effective information flow is essential to ensure that the warehouse is able to operate effectively and meet the needs of customers and other stakeholders. In this section, the various sources and types of information that are relevant to LC HMN will be explored, as how this information is collected, processed, and disseminated throughout LC HMN. The understanding of the key elements of information flow in the warehouse is elementary in identifying potential areas for improvement within the thesis scope.

Inbound Order Handling

Incoming orders from the departments of St. Olav's Hospital (outbound orders) are printed on paper. The following requirements are made from the warehouse to the departments of St. Olav's Hospital in order to be able to deliver at the agreed time:

- Order no later than 2 days before shipment departure if shipment leaves at 08:00 am from the warehouse.
- Order the day before to get delivery with the shipment leaving at 10:00 am or later.

Outbound Order Handling

In order to handle outgoing goods delivery, goods are taken from both cross-docking and normal picking. As mentioned above, there are four departures every day, Monday through Friday. Cross-docking goods are shipped with delivery at 10:00 or 14:00. Items picked are sent according to the delivery request received in the order.

Procurement Planning

LC HMN utilizes a continuous reorder point policy. Every night SAP ERP (system applications and products, enterprise resource planning) generates new purchase orders for suppliers. Orders are automatically generated for items where the inventory level has dropped below the item's reorder point.

The orders are then manually transferred by the order manager. Automated orders are double-checked against the inventory level. Here, additional orders are often added as needed. Potential remaining order quantities are also taken into account. Additionally, there is made an effort to consider the ordering of complete pallets, although it is not the top priority. The procurement planning will be further analyzed in Section 5.1.4

5.1.4 Order Policy

To ensure that the inventory of LC HMN is sufficient, SAP checks every night for potential new orders. The system is based on the continuous reorder points system, in which an order is placed if the inventory level of an item is below the Reorder Point (ROP), see Figure 31.

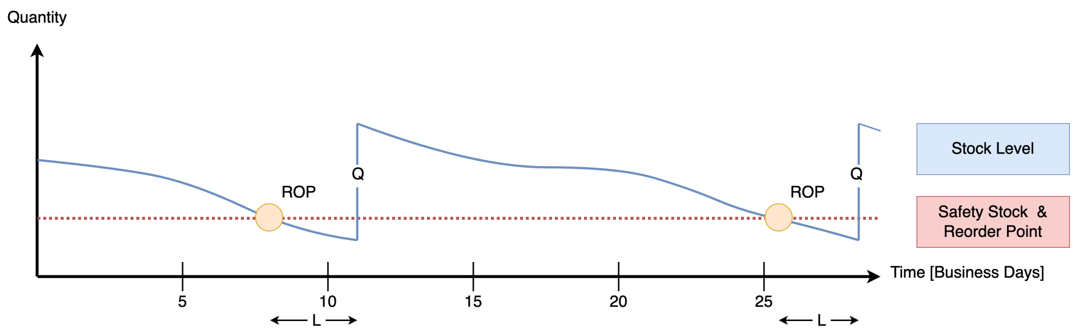

Figure 31: AS-IS order policy at LC HMN

For LC HMN, the ROP is set to correspond to the safety stock. Indicating that in periods, the inventory level may drop below the safety stock level (Figure 31).

Each individual material has a fixed order size (Q) that remains constant. This fixed order size is determined once per year, taking into account the demand from the previous year. Modifications can be made to the specific order size during the ordering process, driven by economic considerations. For instance, if there is an opportunity to fill a pallet or if it is advantageous to make minor modifications, these adjustments can be implemented. It's worth noting that there are no quantity discounts offered for orders, and as a result, the order quantity is not influenced by this factor. It is important to mention that the ordering cost is not taken into account during the determination of the fixed order size (Q) for each individual material.

The lead time (L) is not constant but varies with each order and for each individual material. For most materials, the typical lead time is stated to be between 3 and 10 days.

Service Level

LC HMN operates with a service level of minimum 97%. This means that LC HMN aims to have 97% of all orders fulfilled within the expected delivery time or less.

Substitute Products and Unfulfilled Demand

It is common for the product range to be periodically supplemented with new products or new suppliers of similar products. These products serve as substitutes for existing products in the range and should be taken into account when calculating the inventory level for LC HMN, as it is a crucial factor in the inventory management process. If the demand is higher than the inventory level, LC HMN will deliver substitute products.

5.1.5 Holding Cost

As described in Section 2.1.1, the holding cost consists of capital cost, storage cost, and risk costs, and increases with inventory levels. Through communication with the contact person at the case company LC HMN, a holding cost relative to the value of the SKUs could not be stated.

Azzi et al. (2014) conducted a multi-case study to gain insights into how industrial managers currently calculate the holding cost parameter and to assess the impact of manual versus automated warehousing systems on defining the structure of inventory costs. Their results stated that inventory holding cost parameters range between a minimum of 21.9% and a maximum of 32.9% of the inventory value on hand. For warehouse types with similar levels of automation as LC HMN, the traditional rule of thumb stating that holding costs amount to 25% of the product value is a reasonable assumption. Thus, a holding cost of 25% is assumed for LC HMN in this thesis.

5.1.6 The Supply Chain

As visualized in Figure 32, the LC HMN is part of a three (four)-level supply chain within the scope derived in Section 1.4. LC HMN has categorized suppliers into "major" and "minor" suppliers. The major suppliers deliver goods directly to the logistics center using their own trailers. The minor suppliers use a third party, Posten, to deliver goods to the logistics center.

From the logistics center, goods are sent with their own courier to the central warehouse of St. Olav's Department Øya, which further distributes to the departments at St. Olav's Hospital Department Øya. For the remaining departments, Orkanger, Satelite Units, and Røros, orders are sent directly from the logistics center. Of the total number of deliveries, 80.03% goes to the Øya department, 8.91% to Orkanger, 6.99% goes to satellite units, and 0.47% goes to Røros. The remaining 3.61% goes directly to different doctor offices and private individuals.

In the near future, LC HMN will deliver to hospitals in Levanger and Namsos in Helse Nord-Trøndelag, as well as Ålesund, Volda, Kristiansund, and Molde in Helse Møre og Romsdal. When LC HMN delivers to all these healthcare facilities, they will be responsible for ensuring that the healthcare facilities have the goods they need. As mentioned in Section 1, these healthcare associations are responsible for approximately 732 000 citizens of the Norwegian population.

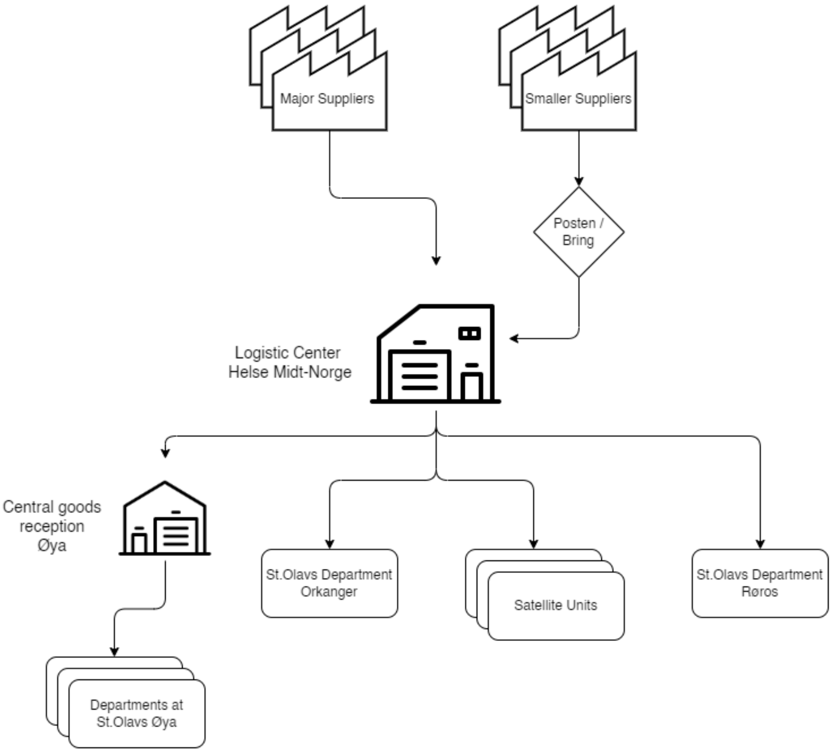

Figure 32: Illustration of the three echelon supply chain.

Order requests from different departments are sent directly to LC HMN. There is a dedicated team that checks the buffer stocks at the various departments as well as the central warehouse of St. Olav's Department Øya, and manually submits orders. The order quantity is determined by this team based on experience and judgment.

5.2 Data Analysis

In this section, the results from the performed data analysis is presented. This includes collection of the data, preprocessing, categorization, time series analysis and aggregation analysis.

5.2.1 Data Collection

For this thesis, three files representing the inbound logistics, purchasing data, and outbound logistics were extracted from LC HMN's SAP system by their SAP engineer.

Inbound logistics refers to the transportation and management of goods and materials that are received from suppliers. Purchasing data represents details such as the supplier, material description, quantity purchased, price, delivery schedule, and payment terms. Outbound logistics refers to all the orders performed by the different hospitals in the region of LC HMN. The data sources are stored as CSV files (comma-separated values), with a total size of 147 MB, containing data registered from the years 2017 to 2022.

The three data files were combined into a single table, making it easier to access all the data. As Table 6 illustrates, the combined data table consists of 25 different columns. The data table contains a variety of different types of data, which can be divided into two main categories, numerical and textual. The numerical data consist of integers, floats, and other numerical values, while the textual data consist of text strings.

Original Attribute	English Named Attribute	Explanation	Format
Innkjøpsordre	Purchase order	Delivery ID unique pr delivery	Float
Posisjon	Position	Line number of the purchase order	Float
Material	Material	Item ID unique pr item	Float
Batch	Batch	Batch from production	Integer
Materialkorttekst	Material short text	Description of item, included size	String
Fabrikk	Factory	Factory ID unique pr factory	Integer
Lager	Storage	Storage ID. All are the same and represent Logistic Center HMN	String
Bevegelsestype	Movement type	Type of movement of the material	Integer
Bevegelsestype: Tekst	Movement type: Text	Extra description of movement type	String
Materialdokument	Material document	Identifier for underlying documents related to the withdrawal of goods	Integer
Posisjon mat.dok.	Position material document	Position of the product in the material document	Integer
Konteringsdato	Posting date	The date on which funds are taken or added to a checking account	String
Kvantum i RKE	Quantum in RCU	Quantum in the Regional coordinating unit	Integer
Reg.kvantumsenhet	Registered unit of measure	Unit of measure	String
Basiskvantumsen-het	Base quantum unit	Unit of measure	String
Kvantum Inngående	Quantum Inbound	Quantity of items inbound	Integer
Kvantum Utgående	Quantum Outbound	Quantity of items outbound	Integer
Beløp lok. val.	Amount local currency	NOK per material unit	Float
Nettopris	Unit price	Price per material unit	Float
Valuta	Currency	The currency of the "Nettopris" column	String
Bestillingsdato	Date of order	DD/MM/YYYY	String
Leveringsdato	Date of delivery	DD/MM/YYYY	String
Registreringsdato	Date of registration	DD/MM/YYYY	String
Leverandør	Supplier	Supplier ID. Unique per supplier	Integer
Bilagstopptekst	Attachment header	Name of the supplier	String

Table 6: Explanation of attributes of the data set.

In the data table, numerical columns such as *Material*, *Quantum Inbound*, *Quantum Outbound*, and *Unit Price* are considered the most important numerical data for this case study. The numerical float *Material* is a unique identification for the type of material. The numerical integer *Quantum Inbound* represents the amount of a given material that is received from the supplier at a given time, and *Quantum Outbound* represents the quantum shipped from the warehouse to a given hospital at a given time, while the decimal number (float) *Unit Price* is the price of the material unit.

Textual data can provide more detailed descriptions of the data and can give context to the numbers. In this data set, textual data such as labels and dates assist describing the numerical data and provide additional information about the subject of the data set. For this data set, textual columns such as

Registered quantity unit, *Material short text*, *Date of Order*, *Date of delivery*, and *Date of Registration* are considered the most important features for analysis.

It is important to note that *Material* is the most important attribute. This attribute represents a unique code for each distinct Stock Keeping Unit (SKU).Hence, in this thesis, the term *material* specifically refers to a Stock Keeping Unit (SKU).

5.2.2 Data Preprocessing

This section highlights the essential techniques employed as part of the necessary preprocessing of the collected data. It encompasses data cleaning, engineering, and filtering.

5.2.2.1 Data Cleaning & Engineering

The initial stage of data cleaning involves performing a fundamental exploration of attributes to gain insight into the data. As mentioned in Table 6, the given dataset contains 25 columns.

Removing Irrelevant Attributes

The first step was to verify the presence of missing values, commonly referred to as "Not a Number" (NaN). Table 7 shows the attributes that have missing values. The attributes *Batch*, *Leverandør* and *Bilagstopptekst* have a share of missing values close to 100% due to the removal of confidential information described in Section 5.2.1. While the attribute *Innkjøpsordre* has a share of missing values of 2,3 %. These attributes were removed, as they will presumably not have an important contribution to upcoming the time-series analysis. Attributes such as *Posisjon*, *Fabrikk*, *Lager*, *Bevegelsestype*, *Bevegelsestype:Tekst*, *Materialdokument*, *Posisjon mat.dok.*, *Konteringsdato*, *Kvantum i RKE*, and *Reg.kvantumsenhet* were also dropped, simply due to not being relevant for the data analysis.

Attributes	Innkjøpsordre	Batch	Leverandør	Bilagstopptekst
Percentage of NaNs	2,3 %	97,2 %	100 %	99,9 %

Table 7: The percentage of missing data.

Ensuring The Right Data Types

Any inconsistent formats should be corrected to ensure that the data is consistent across multiple systems and possible to utilize for further analysis. For example, columns that describe dates should be standardized to a format that can be handled by the software used to analyze. The attributes *Bestillingsdato*, *Leveringsdato*, and *Registreringsdato* describe the date of order, the date of order arrival, and the date of shipment respectively. These three attributes are of data type *String*. Converting these attributes to date-time objects makes it possible to index, measure, and record changes in data points over time, including seasonality and trends. Date-time objects are also used to set the boundaries of time-series analysis, such as the start and end of the analysis, and determine the frequency of data points.

Aggregation of Data to Daily Level

Initially, the data points are at an hourly level. To simplify analysis and predictions, the data were aggregated on a daily level and grouped accordingly to the type of material. This higher forecast interval ensures that the analysis and predictions are more simplified. Furthermore, the reduction of data points to a higher forecast interval helps to smooth out the curve of the time series. Additionally, the reduction of the dataset's size makes it more manageable and easier to work with. This also allows for easier identification of trends and patterns in the data, enabling more informed decision-making.

Lead Time Calculation

Feature engineering was performed to calculate the lead time for each order. This was achieved by extracting the business days between the "Bestillingsdato" and "Leveringsdato" columns. Extracting only the business days will be more representative than including weekends and vacations since the logistics are on hold under weekends and vacations. The resulting lead time metric provided insights into the reliability and consistency of the inbound logistics process. To further analyze the data, the average lead time for each material was calculated.

Dataframe

After calculating the lead time metrics, a data frame was created for future use. The data frame consisted of one line per material, with several columns including the lead time mean, as well as other relevant information such as *Korttekst*, *Lagerkvantumsenhet*, *Nettopris*, and *Valuta*. This data frame provided a comprehensive overview of the inbound logistics process for each material and could be used for further analysis and insights into the replenishment process.

Removing Outliers

Outliers are data points that are significantly different from the other data points in a dataset. It is essential to remove outliers from a dataset to obtain accurate and reliable results. The SciPy Python package, as seen in Figure 20 under the Methodology chapter, was utilized to calculate the outliers for both the outbound logistics (*Kvantum ut*) and inbound logistics (*Kvantum inn*) data points in the dataset. The z-score is a measure of how many standard deviations a data point is away from the mean of the data set. If a data point has a z-score larger than 4, it is considered an extreme outlier.

Removing outliers with a z-score larger than 4 is important because these values can have a significant impact on statistical analysis and machine learning algorithms. These values can skew the mean and standard deviation of the dataset, leading to inaccurate results. In some cases, extreme outliers can also cause a model to overfit, meaning it performs well on the training data but poorly on new data. By removing extreme outliers, the dataset is more representative.

Average Inventory Level

Calculating and knowing the average inventory level for all materials in an inventory is crucial for several reasons. Knowing the average inventory level helps in reducing costs. Maintaining excessive inventory levels can tie up a significant amount of capital, and the cost of storing excess inventory can quickly add up.

The inventory level for all materials at all given times was not retrievable from the LC HMN database, but they conducted inventory counts three to four times a year. The inventory level was calculated by starting with a manual inventory count and taking the difference between cumulative demand and cumulative purchases. This allowed them to see the inventory level for all materials for a year ahead. Then the average for all these inventory levels was calculated, so that the average inventory level for all materials for the given time period could be seen.

Unit Price Calculation

The unit cost for each material was calculated by fetching the total tied-up capital for each material, divided by the total quantity for each material. The unit price of each material can help manage the inventory more effectively by tracking the most expensive materials and improving the ordering policies accordingly.

Coefficient of Variation

The coefficient of variation provides a way of understanding the variability of demand relative to the average level of demand. This measure was calculated by first calculating the standard deviation and average values of demand for each material, then dividing the standard deviation by the average value. By comparing the coefficient of variation for different materials, items that are highly volatile can be identified. Thus, the measure is useful to better understand and manage the demand variability.

Inventory Turnover Ratio

The inventory turnover ratio is a measure of how quickly a company is able to sell and replace its inventory. The inventory turnover ratio calculated involves dividing the total demand for each material by the average inventory level. The average inventory level is calculated as described under paragraph *Average Inventory Level*.

Average Demand Interval

The average demand interval is a measure that reflects the average time between two consecutive demands of a material. This measure is calculated by taking the total number of periods, given in days, weeks, or months, and diving that number by the total number of periods with demand.

5.2.2.2 Data Filtering

Figure 33 visually represents the sequential steps involved in filtering the preprocessed data, ultimately leading to a subset of data, stated suitable for analysis. The diagram underscores the importance of data quality and completeness, as well as the significance of data cleaning and preprocessing in obtaining a reliable and representative dataset for the case study data analysis.

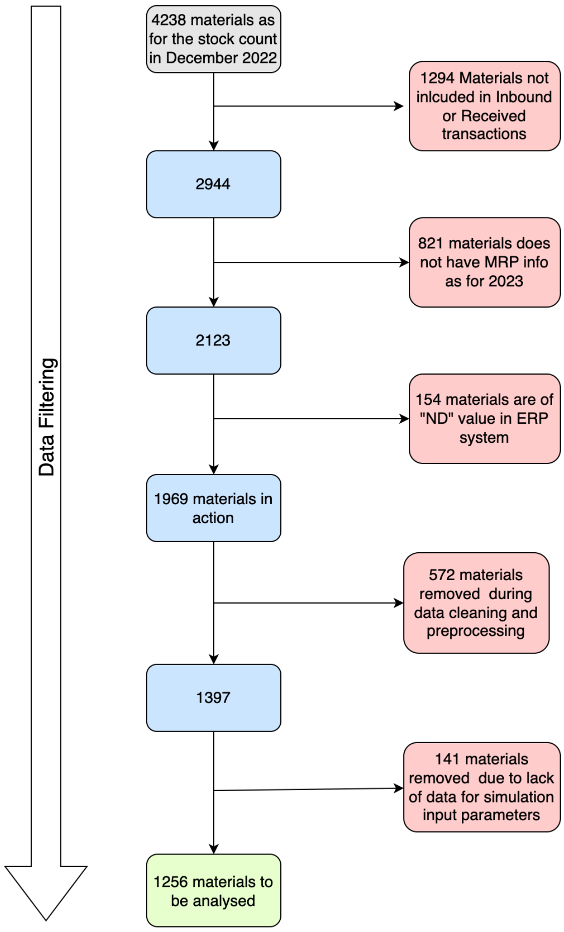

Figure 33: Data filtering overview

Starting with a raw dataset updated in December 2022, a filtering step is applied to remove materials that are not included in inbound or received transactions. This initial removal reduces the dataset by 1294 materials. Following the removal of materials not included in transactions, another filtering step is conducted to identify materials lacking crucial Material Resource Planning (MRP) information for the year 2023. Subsequently, the dataset is carefully examined to identify materials labeled as "ND" (No Data) in the ERP system. The next step involves data cleaning and preprocessing, described in Section 5.2.2.1, where materials are filtered out based on criteria such as data inconsistencies or incomplete records. Before proceeding with the final analysis, a filtering step is conducted to exclude materials lacking sufficient data for simulation input parameters (Section 5.2.2.2). The resulting subset of 1256 materials represents the final dataset that is considered suitable for detailed analysis.

5.2.3 Material Categorization

Analyzing all the materials that LC HMN holds is crucial for identifying the materials that have the greatest potential for improving operational policies and reducing holding costs. In this section, a scientific approach to material analysis involves using statistical and mathematical tools to quantify and interpret the data was performed.

5.2.3.1 K-Means Clustering

In order to classify all the materials in terms of demand and holding inventory, a graph (Figure 35) was generated to visualize the relationship between the coefficient of variation (CV), which is a measure of demand variability, and the inventory turnover ratio, which reflects the number of times an inventory is sold out within a time period.

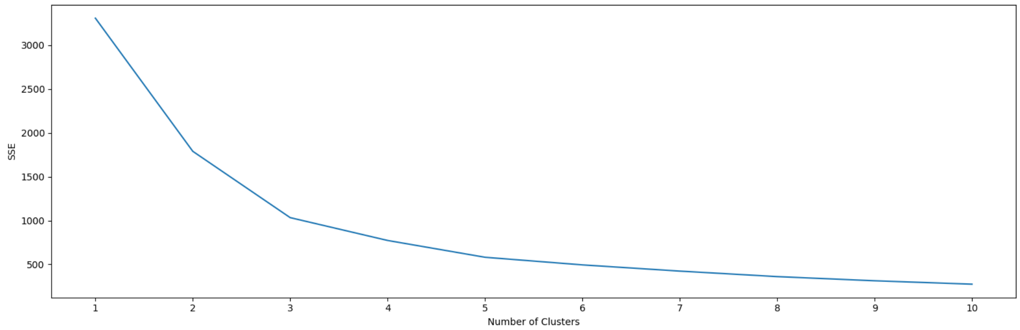

Figure 34: Elbow method.

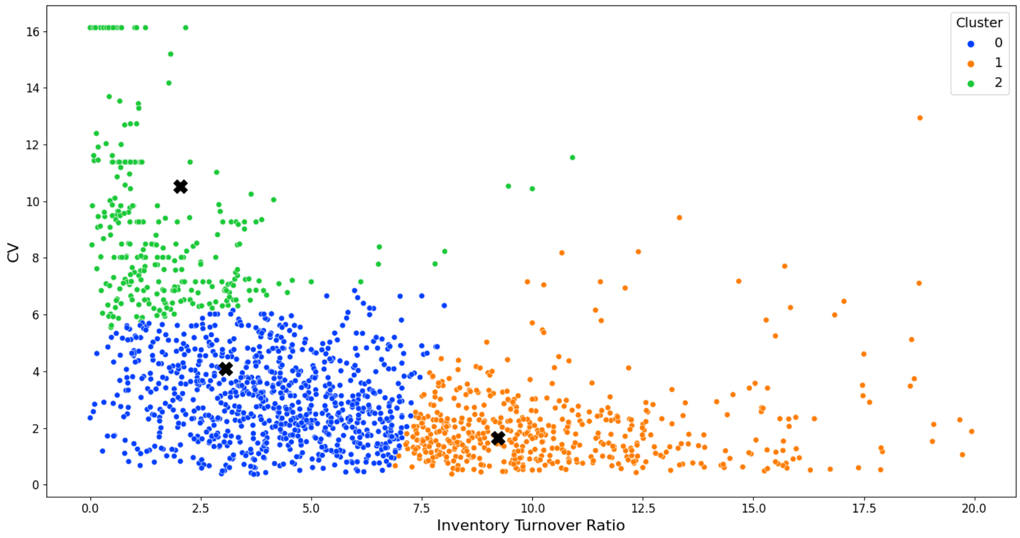

Figure 35: A clustering of all the materials, based on CV-value and Inventory Turnover Ratio.

As seen above, Figure 35 shows the clustering of all 1256 materials included in the subset remaining after the data filtering process in Section 5.2.2.2. All the materials are divided into three clusters, defined through the use of *Elbow Method* by the KMeans algorithm, as seen in Figure 34 (F. Pedregosa et al., 2011). A higher coefficient of variation (CV) means a higher level of dispersion around the mean, which means higher volatility in demand in this case. A CV value larger than 2 is considered high, meaning that the standard deviation is twice as significant as the demand. A high value of inventory turnover ratio will prove a relatively low average inventory level, compared to the demand (Arnold, 2017).

Each individual material in Figure 35 is assigned to a cluster. Cluster group number one contains materials with a relatively high inventory turnover ratio ($ITR > 7$), which means that the materials are often selling, indicating a high demand. The coefficient of variation (CV) for cluster one is also relatively low ($CV <$

3), meaning that the standard deviation of demand is less or equal to third times the average demand.

Materials belonging to cluster group zero represent materials with a relatively lower inventory turnover ratio. This cluster group account for approximately 50% of the different material types and have slightly slower demand than the previously mentioned cluster groups 1. The final cluster group, number 2, contains materials with a high ratio between standard deviation and average demand. This cluster group includes material types that are rarely ordered or substitute items that have been ordered if the originally preferred item was unavailable.

5.2.3.2 Syntetos Method

The use of the coefficient of variation and inventory turnover ratio as variables provides some insight into the variability and efficiency of inventory management, but may not capture the full complexity of demand patterns or the relationship between demand and inventory. Thus a second type of classification was performed, to potentially provide a more nuanced and informative view of the data.

As discovered in the systematic literature review in Section 4.2, Syntetos, Boylan et al. (2005) proposed two cut-off values to calculate different demand patterns. The demand intervals are determined by the average demand interval *($ADI = 1.32$)*, and the other coefficient is the demand variation *($CV^2 = 0.49$)*. A high value of ADI indicates a low frequency of demand and a high value of CV^2 indicates a high volatility of demand. Squaring the coefficient of variation can help emphasize the impact of high variability, while the average demand interval can capture the frequency and regularity of demand.

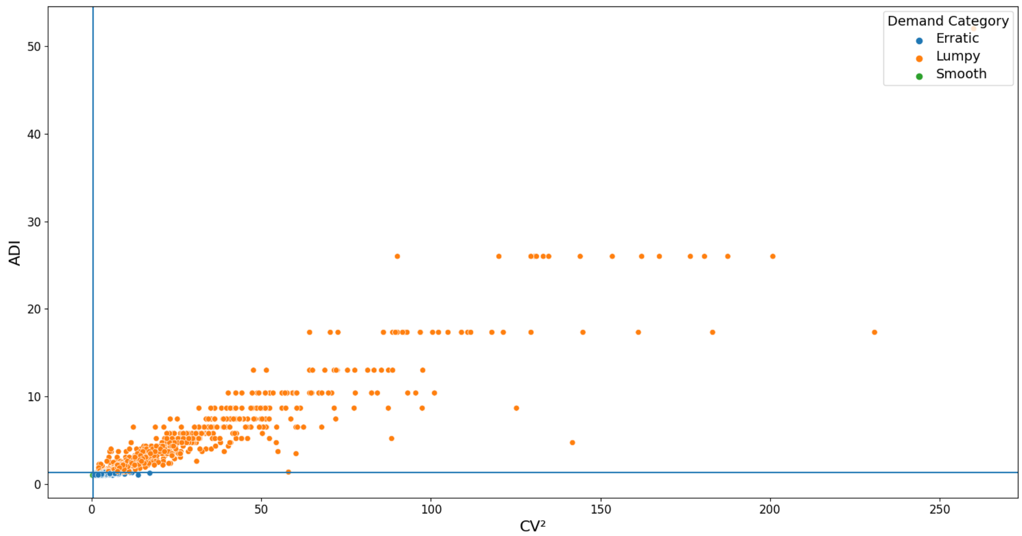

Figure 36: A demand categorization of all the materials.

Figure 36 shows the ADI - CV^2 plot for all the 1256 materials. The materials are divided into four categories, based on the two cut-off values. These categories are erratic, lumpy, smooth, and intermittent, derived by Syntetos, Boylan et al. (2005).

- Smooth Demand (ADI $<$ 1.32 and CV^2 $<$ 0.49): The demand is consistent in quantity and time intervals between occurrences of demand
- Intermittent Demand (ADI $>=$ 1.32 and CV^2 $<$ 0.49): The demand is consistent in quantity, but variation in time intervals between occurrences of demand
- Erratic Demand (ADI $<$ 1.32 and CV^2 $>=$ 0.49): The demand has high variation in quantity, but consistent in time intervals between occurrences of demand
- Lumpy Demand (ADI $>=$ 1.32 and CV^2 $>=$ 0.49): The demand has high variation in quantity and time intervals between occurrences of demand

Among the selection of materials, there are 37.7% of the items in the erratic category, 57.2% of the items in the lumpy category, 5.1% of the items in the smooth category, and zero items in the intermittent category. There is a presence of materials with extreme values. This makes the plot to be rather unrepresentative, as the large scale of the axes distorts the visual display of the data.

Lead Time Distribution

An analysis of the distribution of the average lead time for all materials was conducted. Sorted by the demand categories presented earlier in this section, the lead time distribution in days is relatively similar for each of the demand categories presented in Figure 37.

As can be observed, the majority of materials have an average lead time of between two to five days, regardless of the demand categories. Thus the graphs presented in Figure 37 provide evidence that the average lead time in days for all materials has no clear impact on the categorization of the materials based on demand.

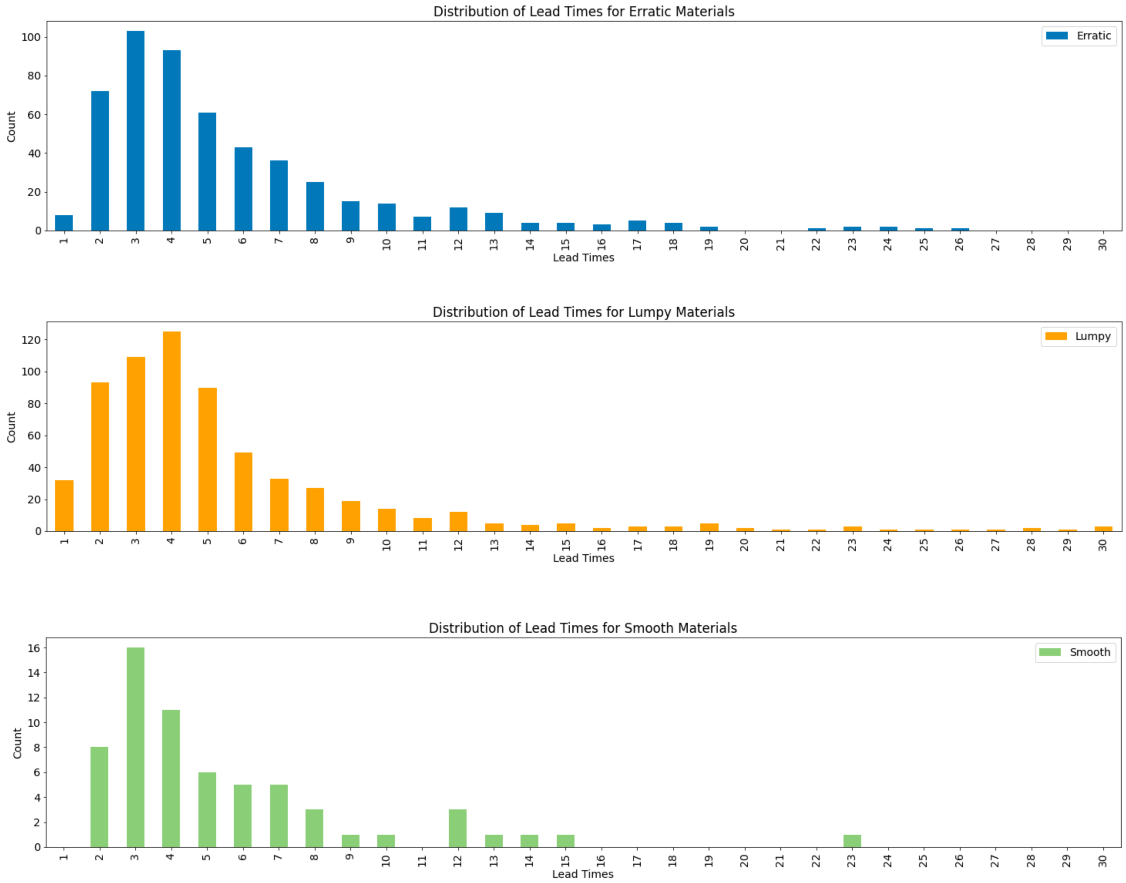

Figure 37: Lead time distribution within demand categories

5.2.3.3 Combining K-Means Clustering and Syntetos Method

Looking closer at the clustering performed in Figure 35, and combining those three clusters with the classification of demand performed in Figure 36, Figure 38 is created to see the distribution of the demand categories within the three clusters. As one can observe, cluster number zero contains the majority of the materials, measuring 61.5% in total.

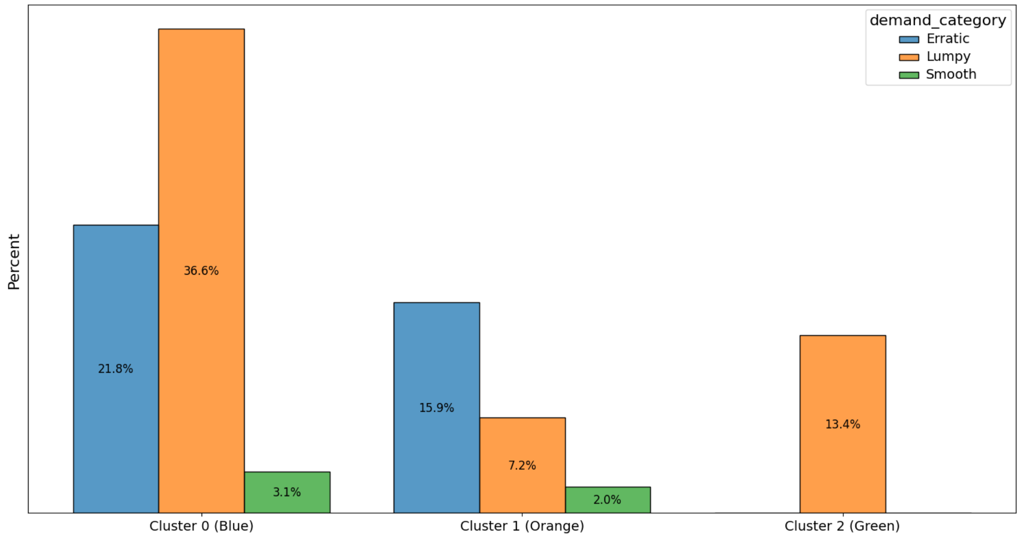

Figure 38: Histogram of the clusters and demand category.

The demand categories within cluster zero appear to be skewed distributed, where 36.6% of the materials within this cluster have a *lumpy* type of demand, and 21.8% has a *erratic* type of demand, while the remaining 3.1% is a *smooth* type of demand. This suggests that materials within this cluster may exhibit a diverse range of demand patterns, which may require customized forecasting and inventory management strategies. Similar to cluster zero, cluster 1 has a skewed distribution with 15.9% erratic demand, 7.2% with lumpy demand, and 2% with smooth demand. In contrast, cluster two exhibits a single type of demand.

As discussed in Section 5.2.3.1, cluster zero has the highest potential in replenishment policy due to its slow-moving inventory and high demand variability, making it challenging to predict demand and plan inventory levels accurately. Cluster one has a relatively lower CV value and a higher inventory turnover ratio, indicating that demand is consistent and inventory levels are not too high throughout the period. Therefore, the potential for this cluster is lower than for cluster zero. For cluster two, the CV value is so high that predicting demand becomes very difficult to accomplish.

A Closer Look At Cluster Zero

Considering that 61.5% of the materials fall under this cluster, characterized by excessive inventory and a considerable variation in demand, it presents the most significant opportunity for improvement for LC HMN.

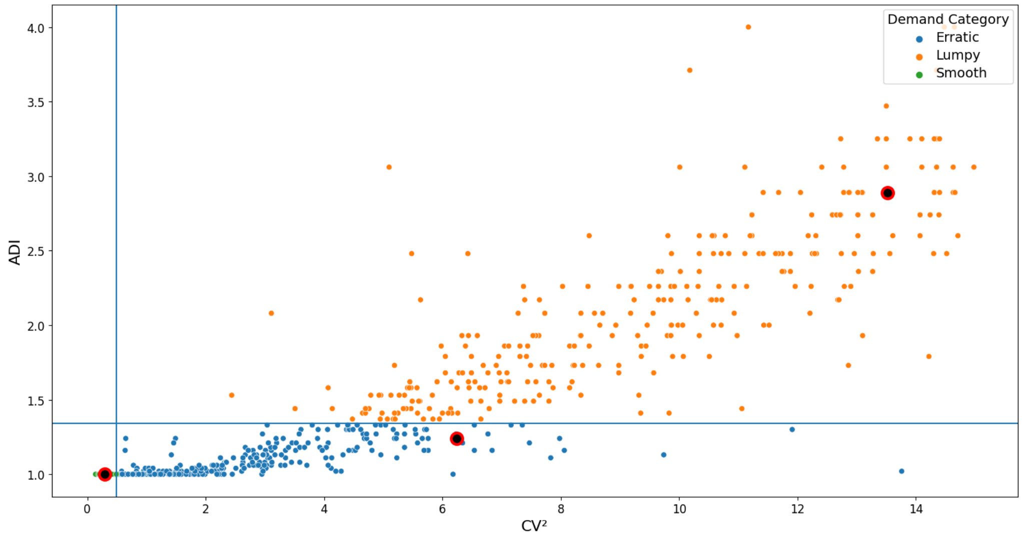

Figure 39: A demand categorization for cluster zero in Figure 38

The ADI - CV^2 plot in Figure 39 provides insights into the demand behavior for the materials within cluster zero. A single material from each category was selected to represent their respective demand category. The three selected materials from cluster zero exhibit different demand patterns, with material 4003841 having smooth demand, material 4001095 having erratic demand, and material 4012198 having lumpy demand. These materials are denoted as red dots in Figure 39, and were chosen as representatives of their respective demand categories.

5.2.4 Time Series Analysis

This section will cover the execution of three types of time series analysis: seasonal-trend decomposition, autocorrelation analysis, and augmented Dickey-Fuller test. Seasonal-trend decomposition will be used to decompose the time series data into seasonal, trend, and irregular components. Autocorrelation analysis will be used to measure the degree of correlation between a time series and its lagged values. The augmented Dickey-Fuller test will be used to determine whether a time series is stationary or not, which is a crucial assumption in many time series models.

5.2.4.1 Decomposition of the Time-Series

Performing a seasonal trend decomposition of time series data is beneficial to better understand the different components of time series (Cleveland et al., 1990). Seasonal-trend decomposition works by breaking the data into three components: the trend component, the seasonal component, and the remainder of the disturbance component. The trend component represents the underlying long-term direction of the data;

the seasonal component captures the seasonal fluctuations; and the remaining component represents all the remaining fluctuations in the data that do not fit into either the trend or the seasonal components, which could be important factors, such as economic cycles or unusual events.

Figure 40 shows a seasonal trend decomposition of three different materials, each representing a demand category from Figure 39. From the first plot for each material, called *Observed*, one can examine that the three materials have completely different demand patterns. Material 4001095 has a pattern with regular occurrences of demand in time, but the quantity of demand may still vary widely, with some periods experiencing high demand and others experiencing lower levels.

For material 4012198, the observed demand is characterized by a large variation in time, thus infrequent orders, but the order quantity is relatively large and consistent. In other words, instead of customers placing orders at a steady pace, they place them sporadically in large quantities, resulting in "lumps" of demand. From the observed pattern of the last material, 4003841, one can monitor a pattern with demand every day throughout the year, but the demand varies to an extent. This pattern provides more data points, making it easier to predict the such type of demand.

Looking at the trend component (the second plot for each material), materials 4001095 and 4012198 have a trend line that is similar to some extent. The trend line has a step-wise line of trend, indicating a type of trend in the data that is characterized by sudden and significant shifts in the values of the data over time. In other words, instead of exhibiting a smooth or gradual change, the data changes abruptly from one level to another, creating a "step" pattern. The discrete steps reflect the infrequent demand that these three materials do represent.

For the third material (4003841), the line of trend is easier to analyze. The trend line remains almost flat until April 2022, when there is a slight increase, followed by a sharp decline towards May. During the months leading up to June, there is a notable increase in the material, followed by a subsequent decline and stabilization towards the end of the year.

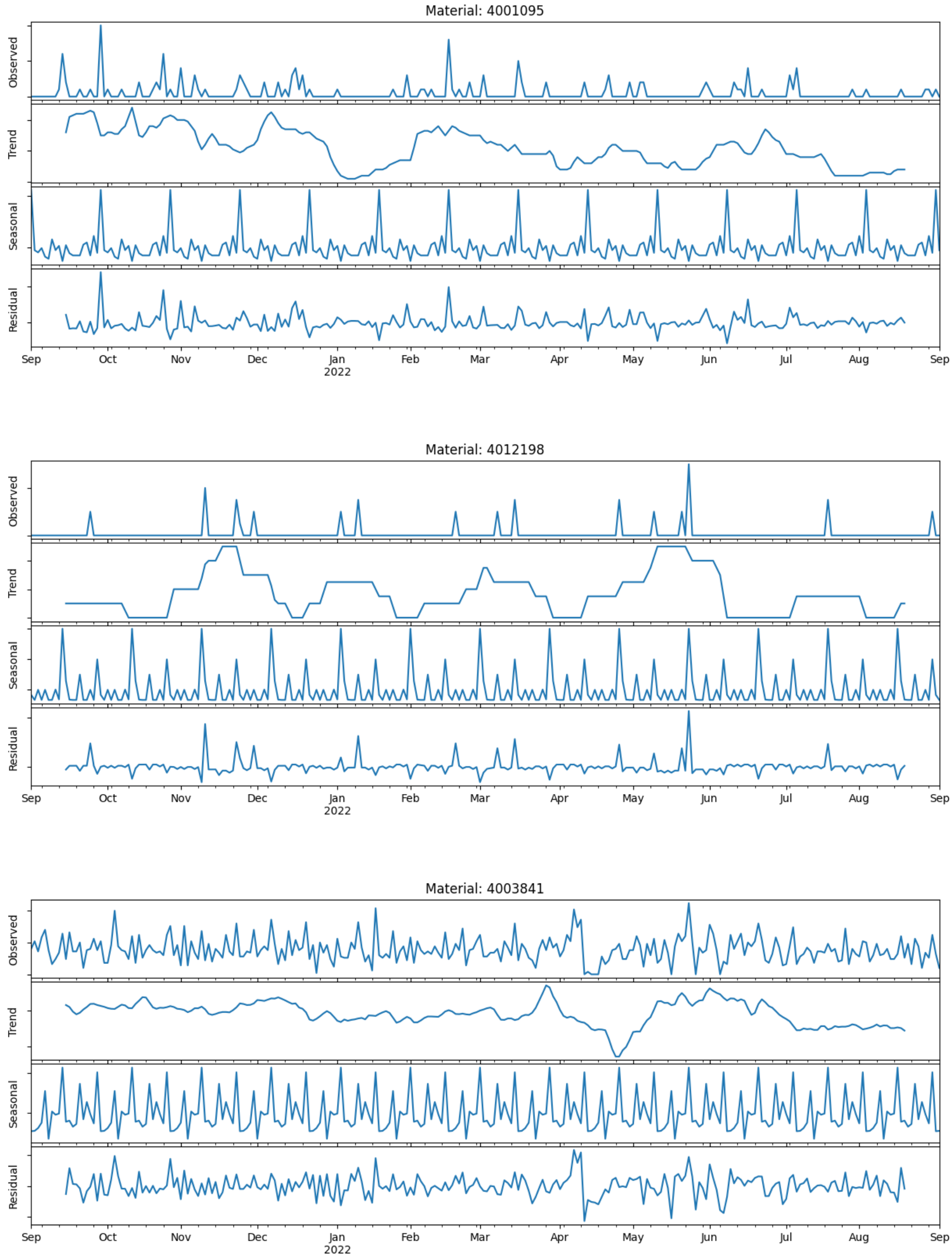

Figure 40: Decomposition of three selected representative materials for each category: Erratic (4001095), Lumpy (4012198), and Smooth (4003841).

The seasonal component called *Seasonal* is represented as the third graph for each of the three selected materials in Figure 40. When comparing the three materials, one can observe that all materials have a seasonal cycling variation that occurs every month. Thus, all three materials have a monthly repeating cycle in the time span between September 2021 and September 2022. The residual component is plotted in the fourth graph. The residuals, also called noise, refers to random fluctuations in time series that do not follow a systematic pattern. Noise can be caused by random events or measurement errors.

5.2.4.2 Auto-Correlation

Auto-Correlation is useful to see how time series develop. Past observations of a given time series, often called lags, are compared with the current value of the time series, which can help in identifying patterns and determining the order of an autoregressive model (Bollerslev, 1986).

The autocorrelation analysis of material 4001095 indicates an initial sharp decline in autocorrelation values close to zero, followed by relatively minor fluctuations in autocorrelation over the entire time period, occurring on both the negative and positive sides. Material 4001095 has no correlations above 0.25 or below -0.25. This indicates that there is no significant linear relationship between the lags for material 4001095.

The plot for material number 4012198 is somewhat similar to material 4001095, besides the majority of values are primarily on the negative side. Material 4012198 differs from 4001095 with some values above 0.25, meaning that it exists proof of a positive correlation between the lags for material 4012198.

The analysis of material number 4003841 reveals a significant observation regarding the autocorrelation behavior. The autocorrelation function exhibits a gradual, linear increase over the considered period, and the graph appears to have a slightly concave shape. The material has several values of auto-correlations close to 2.5, indicating that there is a statistically significant positive correlation between the lags for material 4003841.

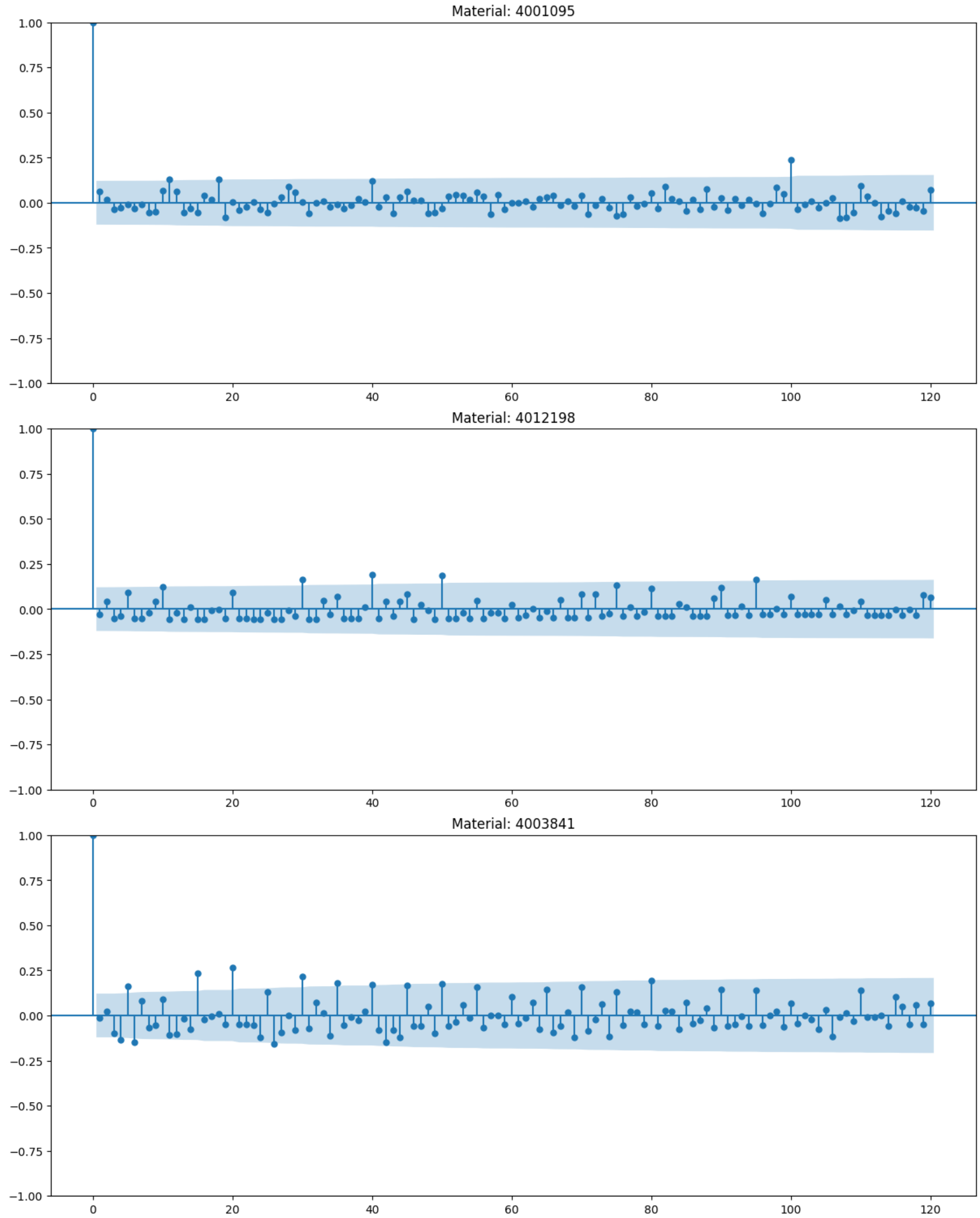

Figure 41: Auto-correlation of three selected representative materials for each category: Erratic (4001095), Lumpy (4012198), and Smooth (4003841).

5.2.4.3 Stationary Test

To verify whether the time series data can be utilized for machine learning demand forecasting, an Augmented Dickey-Fuller test (ADF) (Dickey and Fuller, 1979) was performed. The ADF test is a statistical test that is commonly used to determine whether a time series is stationary. The stationary state is an important assumption for many time series analysis methods, so it is often useful to perform the ADF test to check whether this assumption holds for a given time series.

In Figure 42, the Dickey-Fuller test is performed for the three materials, each representing a demand category. For each material, the original time series is plotted, in addition to the rolling mean and standard deviation, where the rolling window is set to 30 days. In each plot the test statistic, p-value, number of lags used, number of observations used, and the 1%, 5%, and 10% critical values are present. The test statistic measures the extent to which a unit root is present. A p-value that results from the augmented Dickey-Fuller test below a threshold suggests rejecting the null hypothesis, which implies that the observation is stationary and does not have a time-dependent structure (Brownlee, 2021). Otherwise, a p-value above the threshold indicates that the test failed to reject the null hypothesis, and thus it is non-stationary.

The number of lags is the number of time intervals that are included in the test. This number is chosen based on the autocorrelation function and is used to control the potential correlation between successive observations in the time series. The number of observations utilized represents the number of data points in the time series. The critical values (1%, 5%, and 10%) represent the threshold that the test statistic must exceed in order to reject the null hypothesis. The hypothesis test is set as follows:

- Null Hypothesis (p-value > 0.05): It suggests that the time series has a unit root, meaning it is non-stationary. It has some time-dependent structure.
- Alternate Hypothesis ($p <= 0.05$): It suggests that the time series does not have a unit root, which means it is stationary. It does not have a time-dependent structure.

The material 4001095 exhibits a p-value of 0.0, which indicates that the null hypothesis is rejected in favor of the alternate hypothesis. The test statistic of -15.694, which is substantially less than the critical values of 1%, 5%, and 10%, implies that the time series of material 4001095 is stationary and lacks any temporal structure.

For the second material (4012198), the observed p-value of 0.0 implies that the test fails to reject the null hypothesis and thus accepts the alternative hypothesis. The time series of material 4012198 has no time-dependent structure and thus is stationary. Examining the test statistic of -16.4621, which is smaller than all the critical values, implies that the test has strongly rejected the null hypothesis at all the significance levels. This indicates that the test results are highly significant and provide strong evidence against the null hypothesis.

Material 4003841 has a p-value of 0.0003, which is smaller than the selected p-value for the null hypothesis. Thus the null hypothesis is rejected and the alternative hypothesis is accepted. The test statistic of -4.3635 is marginally smaller than the 1% critical value, but significantly smaller than the 5% and 10% critical values. This means that the test has rejected the null hypothesis at all significance levels.

Thus all three selected materials do not have a unit root, and their statistical properties remain constant over time. The time series of the three selected materials is thus stationary and can be utilized for demand prediction methods such as statistical methods and machine learning methods.

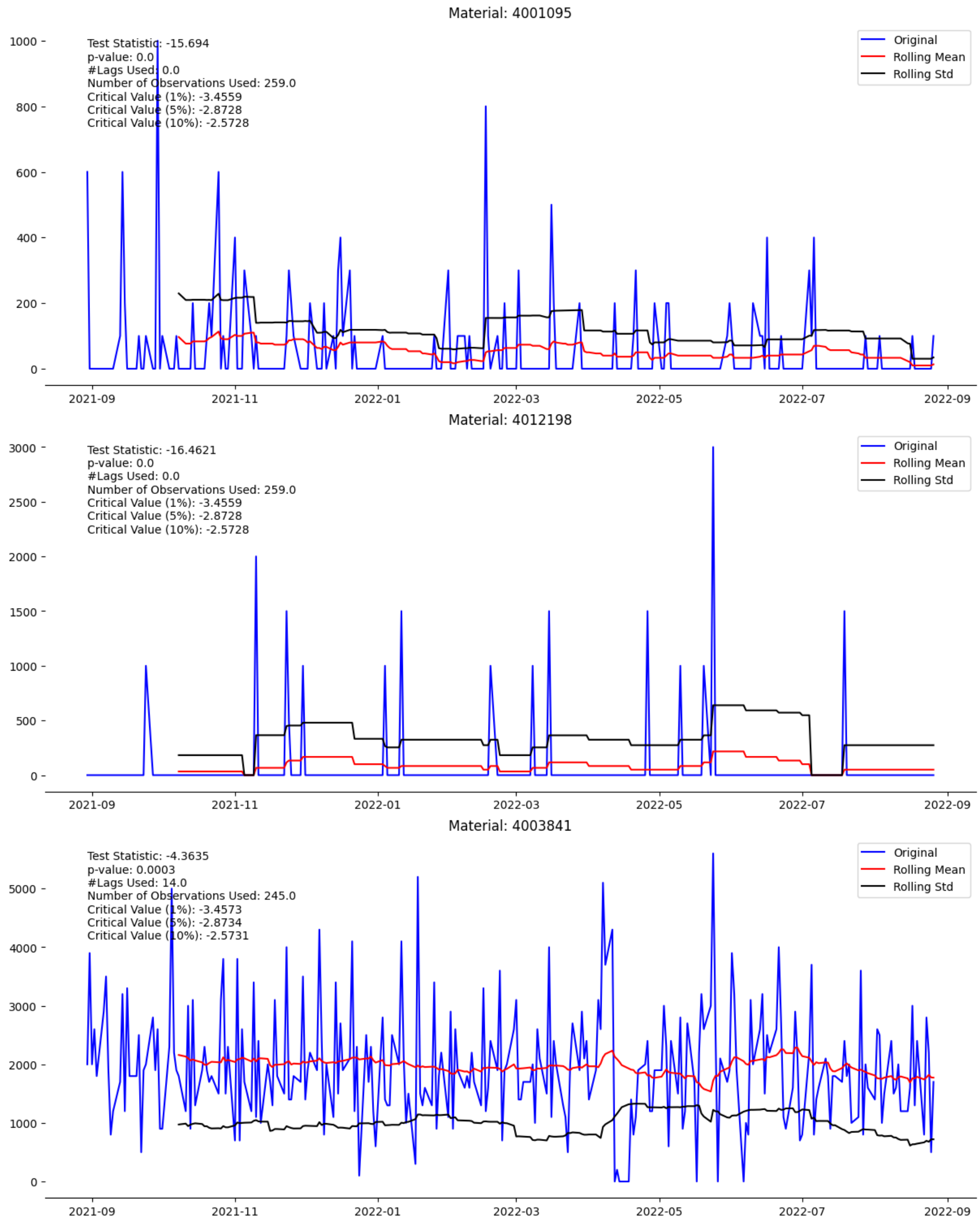

Figure 42: Augmented Dickey-Fuller test for the three selected representative materials for each category: Erratic (4001095), Lumpy (4012198), and Smooth (4003841).

5.2.4.4 Aggregation Analysis

Sparse data is a common occurrence in many real-world problems. There are multiple possible causes for sparsity (T. Chen and Guestrin, 2016). It could be the presence of missing values, frequent zero entries in the statistics, or artifacts of feature engineering.

The observed data is to some extent sparse due to the absent demand for certain dates with zero values. The number of zero values also defines the average demand interval (Section 5.2.2.1), determining which demand category a given item belongs to. If the data is aggregated to a higher level than the base level (which is daily for the preprocessed data of LC HMN), such as a weekly level, it will naturally reduce the number of null values.

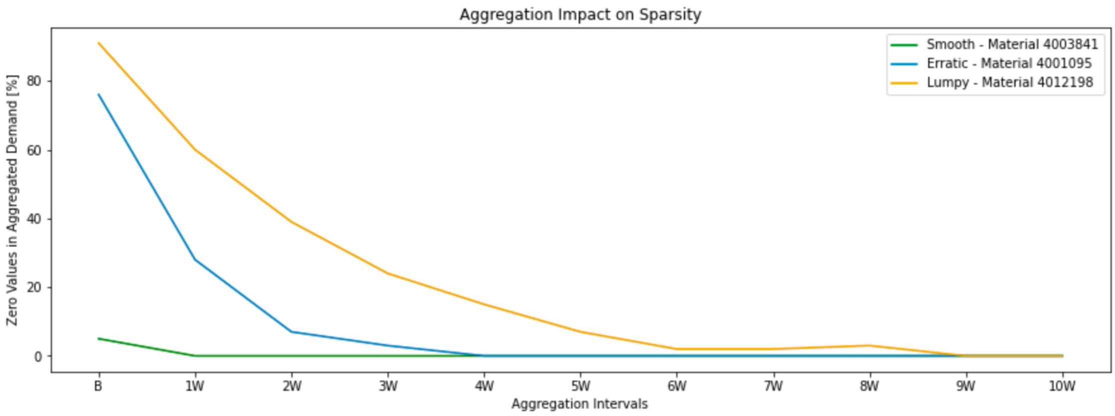

Figure 43: Sparsity analysis of the selected materials

Forecast Interval	Smooth - Material 4003841	Erratic - Material 4001095	Lumpy - Material 4012198
B	5 %	76 %	91 %
1W	0 %	28 %	60 %
2W	0 %	7 %	39 %
3W	0 %	3 %	24 %
4W	0 %	0 %	15 %
5W	0 %	0 %	7 %
6W	0 %	0 %	2 %
7W	0 %	0 %	2 %
8W	0 %	0 %	3 %
9W	0 %	0 %	0 %
10W	0 %	0 %	0 %

Table 8: The resulting representation of zero-values for the different forecast intervals in Figure 43

Figure 43 shows the reduction of zero values when aggregating the time series to a certain level of time. The x-axis spans from business days (noted as "B") to 10 weeks time ("10W"). For material 4003841, which is a smooth type of demand, the percentage of zero values in the aggregated data drops from 5% to 0% when going from business days to a weekly aggregation. This reflects that material 4003841 exhibits a relatively short average demand interval, with demand occurring frequently throughout the representative year.

Material 4001095 has a more significant drop of zero values when aggregating data. When aggregating from business days to one week, the drop in zero values goes from 76% to 28%, and when aggregating from one week to two weeks, the drop in zero values is from 28% to only 7%. The third material, 4012198, has a more gradual slope of development than the latter material. When aggregating the data from business days to weekly, material 4012198 has a reduction in zero values from 91% to 60%. From weekly to every other second week, the reduction is from 60% to 39%.

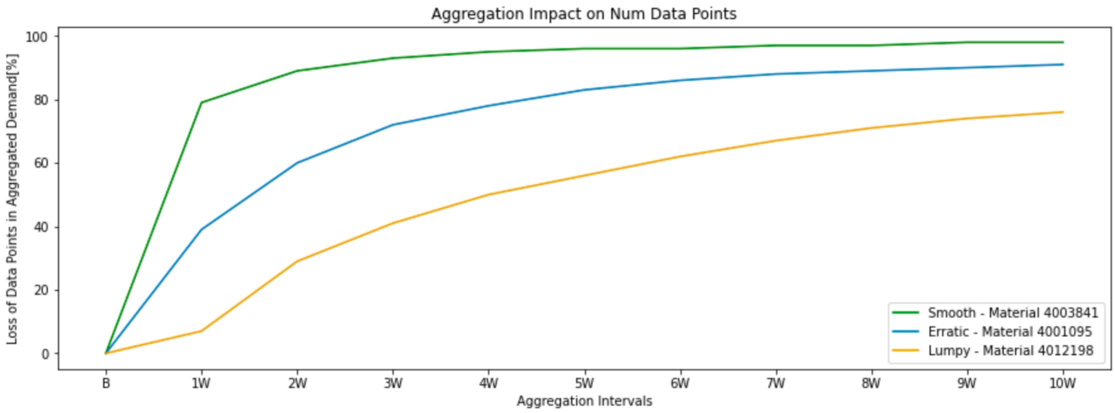

Figure 44: Data point analysis of the selected materials

Forecast Interval	Smooth - Material 4003841	Erratic - Material 4001095	Lumpy - Material 4012198
B	0 %	0 %	0 %
1W	79 %	39 %	7 %
2W	89 %	60 %	29 %
3W	93 %	72 %	41 %
4W	95 %	78 %	50 %
5W	96 %	83 %	56 %
6W	96 %	86 %	62 %
7W	97 %	88 %	67 %
8W	97 %	89 %	71 %
9W	98 %	90 %	74 %
10W	98 %	91 %	76 %

Table 9: The resulting loss of data points for the different forecast intervals in Figure 44

A downside of aggregating to a higher level (e.g. weekly) is that the number of data points is reduced. Although small sample sizes are often prevalent, limited data can pose a challenge in recognizing patterns (Raudys, Jain et al., 1991) and making accurate predictions. Therefore, a trade-off between reducing null values and having enough data points must be made.

Figure 44 shows the loss of data points when aggregating data. Material 4003841 has a drastic reduction of data points of 80% when aggregating to weekly time intervals, showing the downside of aggregating data. Material 4001095 has a reduction of 39%, while material 4012198 has a reduction of 7%. The forecast interval determines the minimum forecast interval. The answer to which forecast interval is the most optimal, thus the forecast interval, will be explored and described for all the materials later in Figure 5.3.3.

5.3 Multi-Scenario Analysis

In this section, the results from a performed multi-scenario analysis are presented. The following five identified and selected strategies will be presented and analyzed.

- AS-IS Fixed Reorder Point
- Proposed Fixed Reorder Point
- Basic Dynamic Reorder Point
- Basic Dynamic (OPTIMAL) Reorder Point
- Advanced Dynamic Reorder Point

As depicted in Figure 3.5, the strategies are compared through a multi-scenario analysis. Firstly, the Python-based simulation model is introduced, including the description of the logic behind dynamic reorder points. Next, the AS-IS Fixed Reorder Point is presented, providing a detailed analysis of the three selected representative materials, in addition to a more general overview of all 1256 materials. Moving on, the Proposed Fixed Reorder Point, Basic Dynamic Reorder Point, and Basic Dynamic (OPTIMAL) Reorder Point strategies are examined. An analysis of all materials is conducted, comparing these three strategies based on inventory level, service level, and holding cost. Lastly, the Advanced Dynamic Reorder Point strategy is investigated. This analysis focuses on the three selected materials and compares this strategy against the other four mentioned strategies, considering inventory level, service level, and forecasting accuracy.

5.3.1 Python-based Simulation Model

In this section, the simulation model programmed in Python specifically for this thesis will be presented. The simulation model is used in order to analyze and compare the five identified and selected strategies described in Figure 3.5.

In order to simulate different order policies, a number of assumptions and parameters need to be determined. The model used to simulate different policies is a time-slicing, deterministic discrete simulation model. A time-slicing deterministic discrete simulation model is a type of simulation model that breaks down a system into discrete time slices but does not introduce random variability or uncertainty into the model (Banks, 1999).

The simulation model is designed to simulate the behavior of an inventory system, taking into account factors such as demand, lead time, reorder points, and order size. The SimPy library (Simpy, 2023) is used to create a discrete-event simulation, which allows the modeling of complex real-world systems in a controlled and repeatable manner.

Characteristics of the Simulation Model

The following bullet points outline the main characteristics and aspects of the simulation model, shedding light on its design and functionality. The simulation model incorporates several key features and considerations to accurately represent the inventory management system under investigation.

- Discrete-event simulation: The model uses the SimPy library to create a discrete-event simulation, where events occur at discrete points in time.
- The time slice is in business days: There are no gaps between working days in the simulation model.
- The deterministic model assumes that all inputs to the model are known with certainty. In other words, there is no randomness or variability in the simulation.
- Maximum one replenishment order issued at the time.
- Unfulfilled demand monitoring: The model continuously monitors whether the demand is met or not.
- Lead time: The model includes lead time, which is the time between placing an order and receiving it.
- Reorder points: The model uses reorder points to trigger an order when inventory levels fall below a certain threshold.
- Fixed order size: The model uses a fixed order size, where the same quantity is ordered every time an order is placed.
- Demand: The model incorporates demand, which is pre-defined.
- Inventory level monitoring: The model continuously monitors inventory levels and triggers an order if inventory falls below a reorder point.
- Order placement and arrival: The model includes processes for placing and receiving orders, which involve lead time and a fixed order size.
- Performance metrics: The model does not explicitly calculate performance metrics such as service level or inventory cost, but they can be derived from the simulated data.
- Initial Stock: The initial stock is calculated using a custom formula.

5.3.1.1 Simulation Logic

The logic of the simulation model is visually represented in Figure 45. The logic is repeated for each day, with the number of loops in the simulation determined by the number of business days with demand.

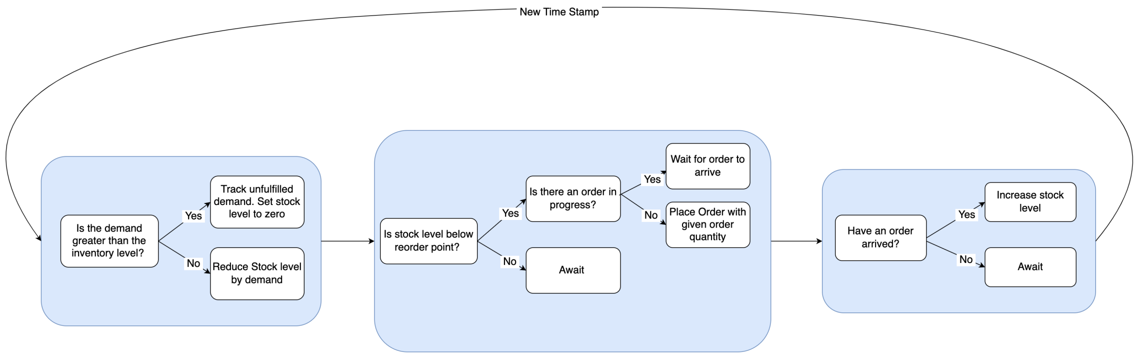

Figure 45: Simplified textual description of simulation algorithm

Initial Inventory Level

This section describes the way the initial inventory level is calculated for the simulation of the Proposed Fixed ROP and the dynamic ROP strategies. As Table 11 shows, with today's replenishment policy, the fixed reorder points and order size are substantially large compared to the initial stock, indicating that LC HMN seeks to maintain a high level of inventory to ensure that they do not run out of stock.

Using the existing inventory level (AS-IS) as the initial inventory level, a conceptual replenishment policy will be influenced by the existing inventory level (AS-IS). Therefore, it becomes essential to calculate an initial inventory level to simulate conceptual strategies, ensuring its relativity and enabling comparisons with both the AS-IS replenishment policy and other conceptual replenishment policies.

The calculation for determining the initial inventory level involves two components: the initial Reorder Point (ROP) and the fixed order size. The initial stock is obtained by adding the initial ROP to half of the fixed order size.

$$\text{Initial Stock} = \text{ROP}_{\text{Initial}} + \frac{\text{Fixed Order Size}}{2}$$

The initial ROP represents the reorder point value for the first day of the simulation, which can be either fixed or dynamic. The fixed order size, determined by LC HMN, corresponds to the specific material being simulated. The initial inventory level serves as the starting point for the simulation and is calculated in the manner described to make the initial inventory level relative to the replenishment policy used in the specific simulation. This facilitates the analysis and simulation of various inventory management policies for each individual sampled material.

Calculation of Dynamic Reorder Point

The dynamicReorder Point (ROP) plays a crucial role in the simulation and testing of various inventory management strategies. It is designed to adapt to changing demand patterns and ensure sufficient inventory levels, based on theoretical formulas inherent to inventory management.

Figure 46 depicts the parameters involved in the calculation of the dynamic reorder point for a specific forecast interval, denoted as the period "0". The horizontal lines shown (light blue, orange, and green) in the figure represent the demand forecasts. These forecasts are obtained using a forecasting model that relies on historical demand data prior to the specified time period. The demand forecast for the upcoming period is a single numerical value that represents the total forecasted demand within the forecast interval of n days. The total forecasted demand is leveled for the n days within the forecast interval. This results in the forecasted demand remaining constant for the n-days within each forecast interval, regardless of the specific forecasting model employed.

The figure also includes dotted lines that represent the daily deviations between the observed demand (represented as a curved blue line) and the forecasted demand. These deviations serve to highlight the variations or discrepancies between the actual demand experienced and the forecasted values.

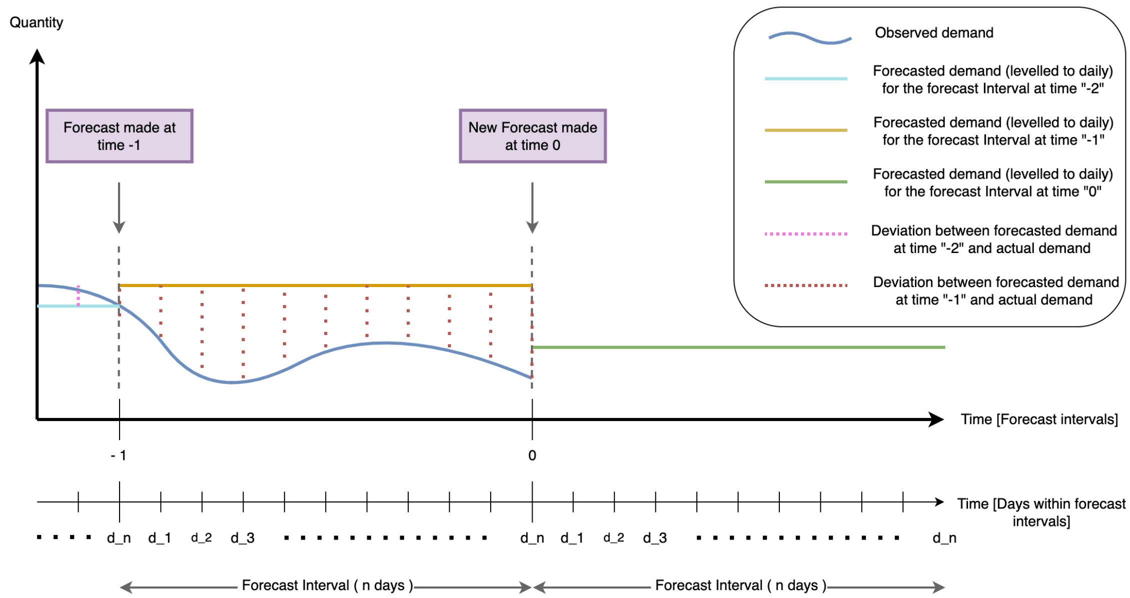

Figure 46: Scope of dynamic reorder point calculation for each forecast interval

Overall, Figure 46 provides a visual aid for understanding the dynamic reorder point calculation process, specifically in relation to the forecasted demand and its deviations. Further on, Table 10 explains the stepwise calculation of the dynamic reorder point.

Step	Step Description	Formula	Formula Description
1	Calculation of forecast interval sigma (σ_{FI}) based on the error (deviation) between the earlier forecasted demand and the actual demand	$\sigma_{FI} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(DailyDeviation)^2}$	"n" is the number of days in the forecast interval. "DailyDeviation" is the deviation between forecasted demand at time "-1" and actual demand at the n^{th} day, see Figure 46.
2	Calculation of lead time sigma (σ_{LTI}) based on an adjustment made to σ_{FI}, to compensate for differences between lead time interval (LTI) and forecast interval (FI).	$\sigma_{LTI} = \sigma_{FI} \times \sqrt{\frac{LTI}{FI}}$	"FI" is the number of days in the forecast interval. "LTI" is the number of days in the lead time for the given material. The resulting σ_{LTI} is the adjusted forecasting error sigma.
3	Calculation of safety stock (SS)	$SS = SafetyFactor \times \sigma_{LTI}$	"$SafetyFactor$" is the number of standard deviations provided as safety stock, determined by Appendix A
4	Calculation of demand during lead time (DDLT)	$DDLT = ForecastedDemand_{Time0} \times LTI$	$ForecastedDemand_{Time0}$ is the forecasted demand at time 0 for the upcoming n forecast interval days, leveled to daily demand.
5	Calculation reorder point (ROP)	$ROP = DDLT + SS$	This ROP is constant for the upcoming forecast interval.

Table 10: Stepwise dynamic reorder point calculation

Table 10 outlines the stepwise dynamic reorder point calculation process. It presents a description of each step, along with the corresponding formulas and their explanations. Step 1 of the calculation involves determining the forecast interval sigma (σ_{FI}). This value is calculated based on the deviations (errors) between the earlier forecasted demand and the observed demand. The formula for σ_{FI} is given as the square root of the average of the squared daily deviations, divided by the number of days in the forecast interval (n), see Equation 7 in the theory section. The "Daily Deviation" represents the deviation between the forecasted demand at forecast interval "-1" and the actual demand on the n^{th} day, as illustrated in Figure 46.

Step 2 focuses on the calculation of the lead time sigma (σ_{LTI}). It involves adjusting σ_{FI} to compensate for differences between the lead time interval (LTI) and the forecast interval (FI). The formula for σ_{LTI} is obtained by multiplying σ_{FI} by the square root of the ratio between LTI and FI. Here, "FI" represents the number of days in the forecast interval, while "LTI" represents the number of days in the lead time for the given material. The resulting σ_{LTI} represents the adjusted forecasting error sigma.

Step 3 focuses on the calculation of the safety stock (SS). The formula for SS involves multiplying the "Safety Factor" (number of standard deviations) by σ_{LTI} (Arnold, 2017). The value of the Safety Factor is determined by Appendix A.

Step 4 involves calculating the demand during the lead time (DDLT). The formula for DDLT is obtained by multiplying the forecasted demand (leveled to a daily demand) at time *0* for the upcoming n forecast interval days by the lead time interval (LTI).

Finally, in Step 5 the Reorder Point (ROP) is calculated. The formula for Reorder Point (ROP) involves adding DDLT and SS, see Equation 3 in the theory section. The resulting ROP remains constant for the upcoming number of days in the forecast interval.

Determination of Key Performance Indicators

As discovered in Section 5.1.4, the service level is highly important for LC HMN, having a lower service level constraint of 97%. Therefore, the service level is included as a key performance indicator in the simulation result.

The average inventory level is an important metric used to evaluate the performance of a simulation. It serves as the second key performance indicator (KPI). This metric measures the quantity of inventory held during the simulation period for the simulated material and is directly linked to the calculation of holding cost.

Limitations of the Simulation Model

To ensure accurate result interpretation and informed decision-making, it is crucial to recognize the limitations of the simulation model. The following are the acknowledged limitations of the constructed simulation model:

- Order Size: The simulation model assumes a fixed order size.
- Demand patterns: The model assumes a predefined demand pattern.
- Fixed lead time: The model assumes a fixed lead time for orders.
- Limited performance metrics: The model focuses on a limited set of performance metrics, such as service level and average inventory level.
- Simplified supply chain: The model assumes a single-level supply chain, which may not accurately reflect the complexities of real-world supply chains.
- The simulation does not handle unfulfilled demand: In the event that the demand exceeds the inventory level for a specific time period, the unmet demand (represented by the negative difference between the demand and the inventory level) will be tracked, but the demand will not be met.

5.3.2 AS-IS Fixed Reorder Point

In this section, a detailed analysis of the AS-IS Fixed Reorder Point is presented for the three representative materials selected from Figure 39. Subsequently, the analysis is extended to encompass all 1256 materials. The first presented detailed analysis is meant to give a deeper understanding of what the analysis includes. By expanding the examination to include the entire range of materials, a broader perspective is captured.

5.3.2.1 Analysis of The Three Materials

The preprocessed actual demand of the selected materials in the time range of 30/08/2021 to 26/08/2022 was passed as input to the simulation model. The distribution of the preprocessed demand in the given time range is shown in Figure 47. The time range equals 260 business days, which is visualized on the x-axis.

The date range differs from the range used in Section 5.2.3, since the dates prior to 30/08/2021 will be utilized as a training set for predictive algorithms, and the dates post 30/08/2021 will be predicted as input for conceptual replenishment policies before the AS-IS replenishment policies will be compared with the conceptual replenishment policies.

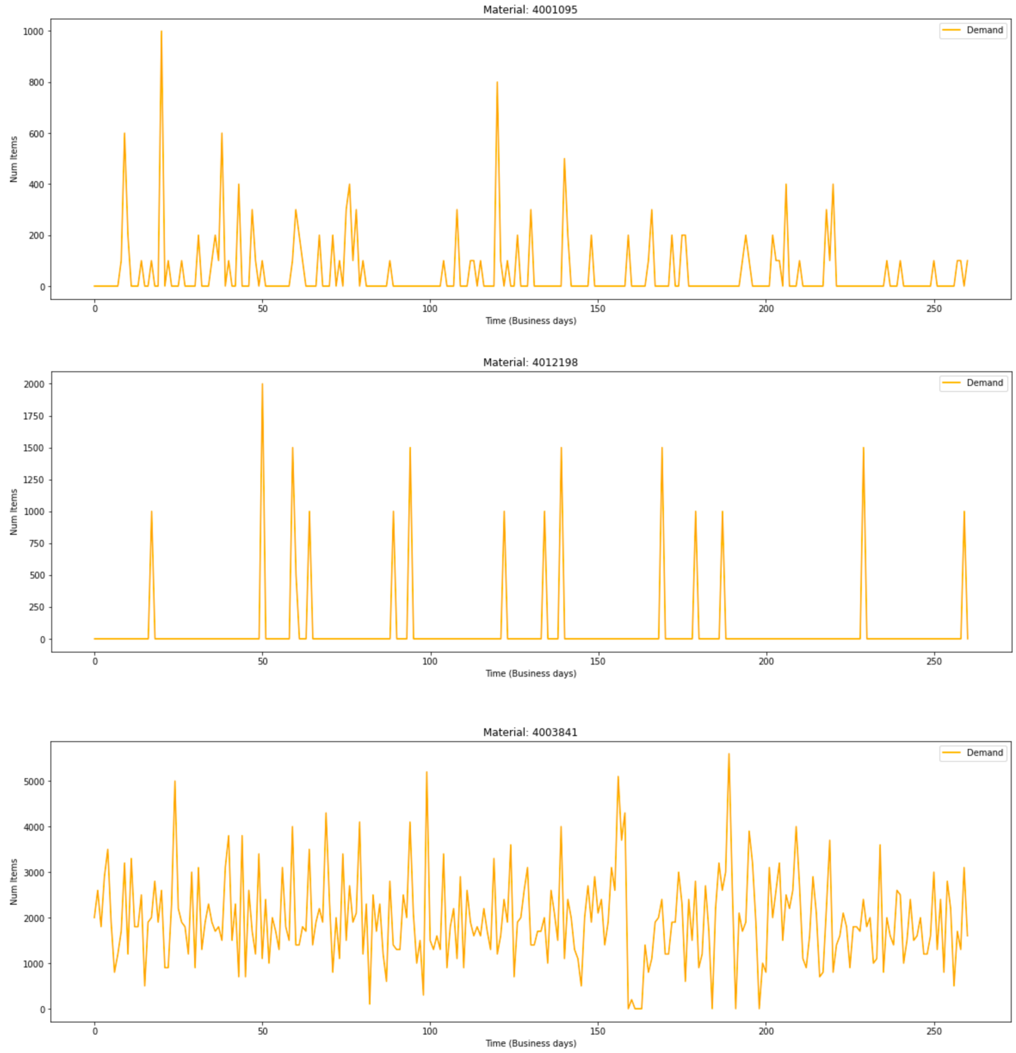

Figure 47: The demand from 30/08/2021 to 26/08/2022 (260 business days) for the three selected representative materials for each category: Erratic (4001095), Lumpy (4012198), and Smooth (4003841).

Despite the time interval used for simulation is different from the time interval used for data analysis (Section 5.2.3), the three materials are assumed to have a similar demand pattern for the two periods, therefore still belong to the same demand categories assigned and explained in Section 5.2.3.2.

In Table 11, a series of other parameters used in the simulation is presented. The initial stock, reorder point and order size has been determined through emails and meetings with the head SAP Engineer at LC HMN. The lead time is, as described in Section 5.2.2.1, calculated based on the historical orderings and inbound reception of goods.

Material	Initial Stock	Lead Time	Reorder Point(s)	Order Size
4001095	5 000 items	4 business days	3 000 items (fixed)	4 000 items (fixed)
4012198	12 000 items	4 business days	8 000 items (fixed)	8 000 items (fixed)
4003841	144 400 items	12 business days	130 000 items (fixed)	28 800 items (fixed)

Table 11: Inputs to the simulation of the selected materials, excluding demand

Replenishment Policy Result

In Figure 48, the AS-IS Fixed Reorder Point of LC HMN for the selected materials is simulated. The average inventory level for the materials is represented as the blue line. The gradual descent of the inventory level is caused by demand, and the shift in the vertical axis represents the arrival of an order to the warehouse.

As stated in Section 5.3.2, the reorder point is fixed, and is visualized as a red horizontal dotted line. Once the inventory level drops below the reorder point, an order is placed to replenish the inventory. The demand for the given material is represented by the orange line. The variation in demand directly reflects the variation in inventory level.

The light green dotted line represents the average inventory level for the simulated 261 business days. The green whole line at the bottom of the plot represents the unfulfilled demand. For the selected materials, this is equal to zero throughout the whole simulation, which indicated a service level of 100% for the time period.

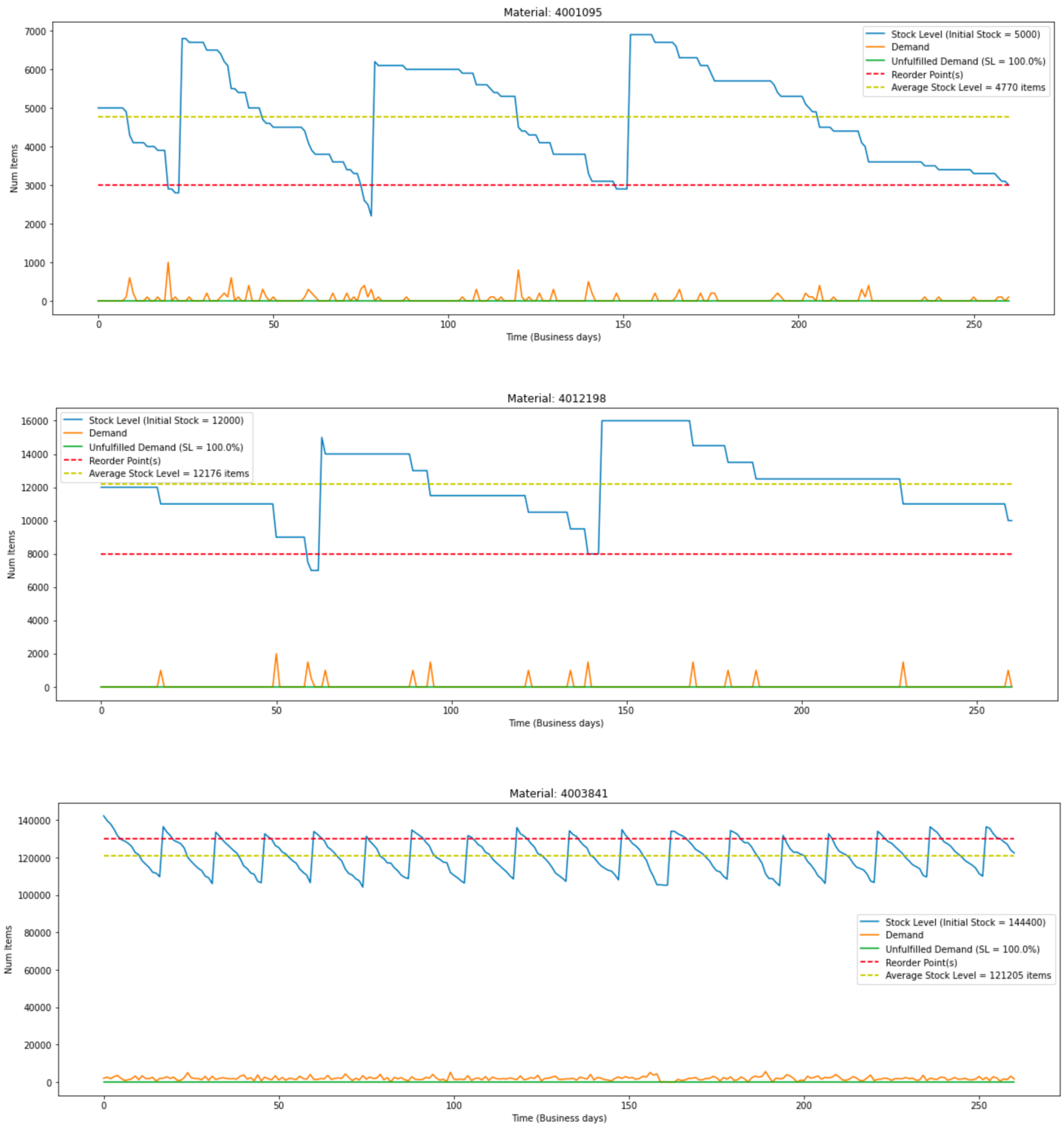

Figure 48: AS-IS reorder point policy for the three selected representative materials for each category: Erratic (4001095), Lumpy (4012198), and Smooth (4003841).

For material 4001095, there have been made three replenishments during the 261 business days, triggered by the inventory level crossing the reorder point. The calculated initial stock is 5000 units. As discovered under Section 5.2.3, material 4001095 has a demand pattern that is categorized as *erratic*, meaning that the demand interval and demand quantity vary with time, which results in a high level of uncertainty about how much inventory is needed to be kept in stock. LC HMN needs to maintain a service level close to 100%, and therefore, they mitigate against this difficult demand pattern by maintaining a high inventory level.

Material 4012198 exhibits a comparable inventory-level pattern to material 4001095. The calculated initial stock is 12 000 units. As discussed in Section 5.2.3, the demand pattern of material 4012198 was classified as "lumpy," indicating that there is variation in the demand interval but not significant variation in the demand quantity. Thus LC HMN holds a fairly high inventory level for material 4012198 to protect against demand uncertainty.

The last material, 4003841, has a quite different stock-level pattern compared to the two latter materials. The calculated initial stock is 144 400 units. Under Section 5.2.3, the demand pattern of material 4003841 was categorized as *smooth*. This implies that the demand pattern of material 4003841 is characterized by lower fluctuations in demand quantity and a stable demand interval.

Cumulative Stock and Demand Level Results

In Figure 49 the cumulative order arrival and the cumulative demand are visualized. The cumulative order arrival is visualized as a blue line. The vertical shifts in the blue line represent the arrival of materials from the suppliers to LC HMN. The cumulative demand is represented as the orange line, showing the total demand for the material through the time span. The green area between the cumulative demand and the cumulative order arrival represents the inventory level. For each business day, the inventory level is shown as the distance between the cumulative order arrival line, and the cumulative demand line. The more linear the cumulative demand is, the more *smooth* the demand for the material is. As one can observe from Figure 49, material 4003841 which is categorized as a *smooth*, has a relatively more linear line of cumulative demand than material 4012198, which is categorized with a *lumpy* demand pattern.

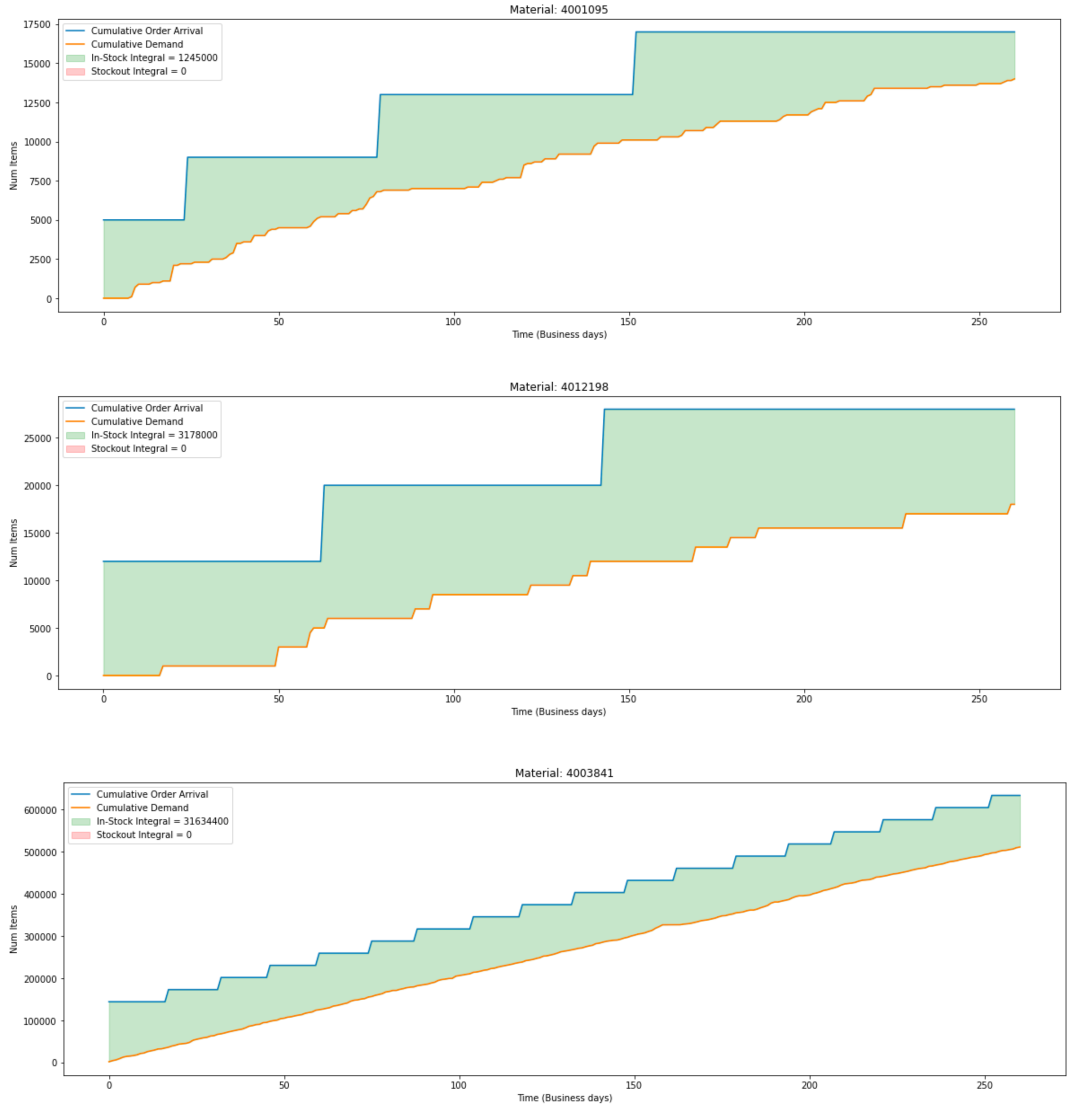

Figure 49: Cumulative stock and demand levels for the three selected representative materials for each category: Erratic (4001095), Lumpy (4012198), and Smooth (4003841).

The intersection of the blue line and the orange line in Figure 49 indicates the point at which the total cumulative demand matches the cumulative order arrival. At this point, the inventory level of the material will be zero, and the LC HMN will need to place additional orders in advance to meet ongoing demand. Conversely, if the cumulative order arrival line is above the cumulative demand line, it suggests that there is excess inventory, which could lead to overstocking and increased carrying costs. On the other hand, if the cumulative demand line is above the cumulative order arrival line, it indicates a potential shortage of the material, which could lead to lost deliveries.

In that case, the two graphed lines have crossed, and the integral of the lost deliveries will be visualized with a red color, also noted as "Stockout Integral" in the graphs for all materials. Fortunately, with today's replenishment policy, all three materials have excessive inventory for all times through the given time span, resulting in no stock-outs and a service level of 100%.

5.3.2.2 Analysis of All Materials

In this section, the analysis of the AS-IS Fixed Reorder Point is extended to encompass all 1256 materials. By expanding the examination to include the entire range of materials, a broader perspective is captured.

Service Level

Figure 50 displays the AS-IS fixed ROP service level for four different categories: Total, Erratic, Lumpy, and Smooth, calculated through simulation, where "total" represents the three others combined. The service level percentage is shown on the y-axis and ranges from 92 to 100%. The x-axis displays the four categories. Each bar represents the service level percentage for each category, and the exact percentage is labeled on top of each bar. The plot compares the service levels between the four categories.

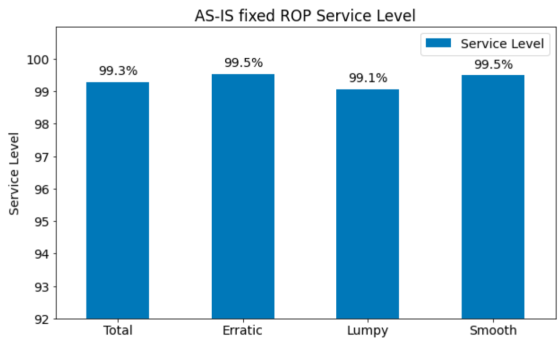

Figure 50: Average service level based on AS-IS fixed reorder point

The overall service level for all materials is 99.3%. When considering specific material categories, both Erratic and Smooth materials exhibit slightly higher service levels compared to Lumpy materials, which have a service level of 99.1%.

Average Inventory Level

Figure 51 displays the number of materials in different average inventory level buckets for the three demand categories. The x-axis shows the different average inventory level buckets, which are labeled as ranges of values, and the y-axis shows the number of materials in each bucket. The data is visualized using a bar chart, with each category having a different color.

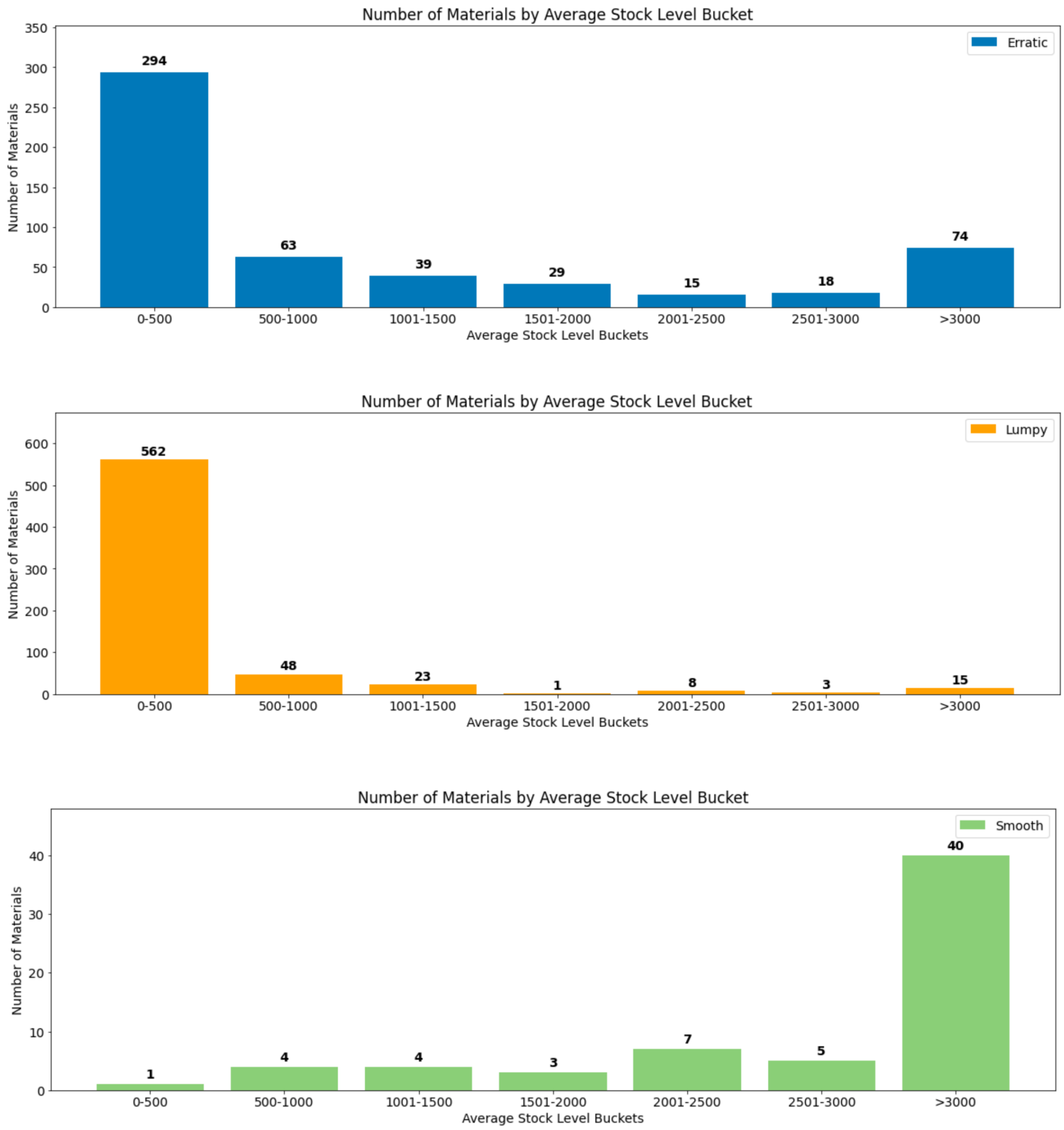

Figure 51: Average inventory level AS-IS reorder point policy

In the erratic category, the majority of materials are concentrated in the 0-500 average inventory level range, with some materials distributed across other buckets. For the lumpy category, approximately 85% of materials are found in the 0-500 bucket, showing a high concentration in this range. Lastly, the smooth category stands out with the highest number of materials (40) in the > 3000 bucket, indicating a larger inventory level.

5.3.3 Proposed Fixed ROP & Basic Dynamic (OPTIMAL) ROP

In this section, the simulation outcome for two of the five strategies from Figure 3.5 is presented: Proposed Fixed ROP and Basic Dynamic (OPTIMAL) ROP. A description of the Basic Dynamic ROP strategy is included in this section to give an understanding of the relationship between the strategies. For the Basic Dynamic ROP strategy, the analysis itself will be presented later in Section 5.3.5

The Proposed Fixed ROP strategy suggests the use of theoretical safety stock and reorder point formulas from inventory management to calculate a new fixed reorder point. This allows for an investigation into the potential of improving the AS-IS strategy, without changing the policy type.

Basic Dynamic ROP utilizes a simple moving average demand forecasting method for dynamic reorder points, involving varying window sizes and forecast intervals. The basic form of demand forecasting allows for a more comprehensive examination of numerous materials and the impact of factors such as forecast interval. The Basic Dynamic (OPTIMAL) ROP strategy is based on "Basic Dynamic ROP", with the optimal combination of window size and forecast interval. It represents a highly comprehensive and high-performing basic strategy, providing a superior benchmark for the "Basic Dynamic Reorder Point" strategy.

The Basic Dynamic ROP and Basic Dynamic (OPTIMAL) ROP are based on the dynamic reorder point calculation as derived in Section 5.3.1.1. SMA was chosen due to the fact of the simplicity and popularity of the method (Ali and Boylan, 2012), thus the strategies utilizing the SMA forecasting method are referred to as *"Basic"*.

Safety Factor

As described in Table 10, the safety factor has to be determined in order to calculate the dynamic reorder point. The case company has set a service level constraint of 97%. To determine the safety factor for a desired service level of 97%, the normal curve can be used (Arnold, 2017). The safety factor, or z-value, indicates the number of standard deviations from the mean in a standard normal distribution that corresponds to a specific service level (Radasanu et al., 2016). A safety factor of 1.88, which corresponds to 97% service level (Appendix A), is used as as input parameter in the reorder point calculations of the simulation model.

Search Space & Simulation Runs

The Basic Dynamic ROP and Basic Dynamic (OPTIMAL) ROP strategies involved the following search space:

- Forecast intervals: Business days (B), weekly (1W), every second week (2W), every third week (3W), and every fourth week (4W)
- Window Sizes: 1-10 [Forecast intervals]

Where forecast interval refers to the time between each forecast made, and the window size refers to the number of forecast intervals of historical demand data taken into consideration when predicting the demand for the upcoming period.

As described in Figure 3.5, the Basic Dynamic Reorder Point strategy involves a variety of window size and forecast interval parameters. In this thesis, the Basic Dynamic Reorder Point strategy will involve all parameter combinations in the presented search space. This means 50 (5 forecast intervals x 10 window sizes) versions of the Basic Dynamic Reorder Point strategy are analyzed. Again, this means 50 unique Basic Dynamic Reorder Point strategies will be tested for all 1256 materials. This equals 62 800 runs of the simulation model.

The Basic Dynamic (OPTIMAL) reorder point strategy is based on one single optimal combination of the 50 parameters in the given search space. This means only one version of the Basic Dynamic (OPTIMAL) strategy is simulated for each of the 1256 materials.

Optimization Technique Used

For each of the 1256 materials, the optimal combination of parameters in the search space was needed to be found in order to analyse the Basic Dynamic (OPTIMAL) reorder point strategy. For small search spaces, a brute force optimization technique can be appropriate (García and Mena, 2013). A brute force optimization technique was used in this case, with the objective of minimizing average inventory. Due to the service level constraint found in Section 5.1.4, the constraint of 97% service level was included in the optimization objective and logic.

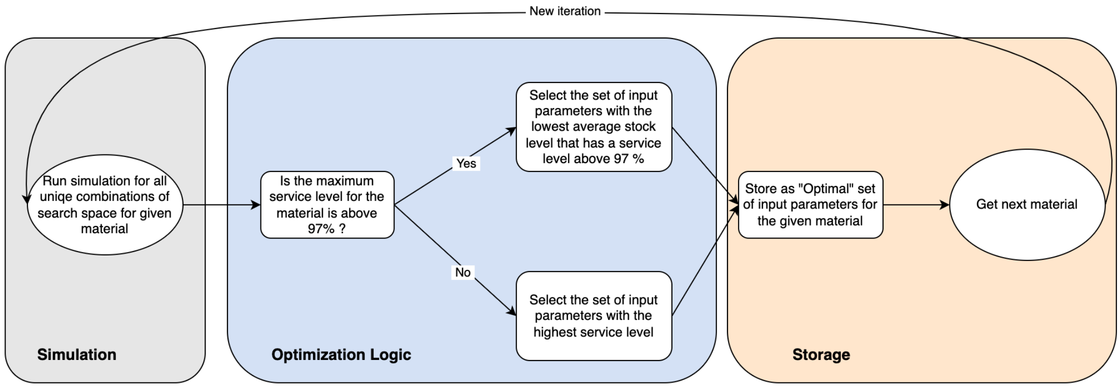

Figure 52: Simplified illustration of optimization search algorithm

Figure 52 illustrates the logic of the optimization objective, showing that the main objective is to minimize the average inventory level for each material, also taking into account the service level constraint of the case company.

Distribution of Forecast Interval & Window Size Combination After Optimization

As mentioned and analyzed earlier under Section 5.2.4.4, the level of aggregation of the time series data is connected to the forecast interval and has a great impact on how many data points that is available for predictive models. In order to determine the best-performing forecast interval, a search was conducted for the forecast interval and the optimal number of time units of window size to use as the data basis for the predictive model. Window size is defined as the number of past observations used as a basis for a prediction (Wheelwright et al., 1998).

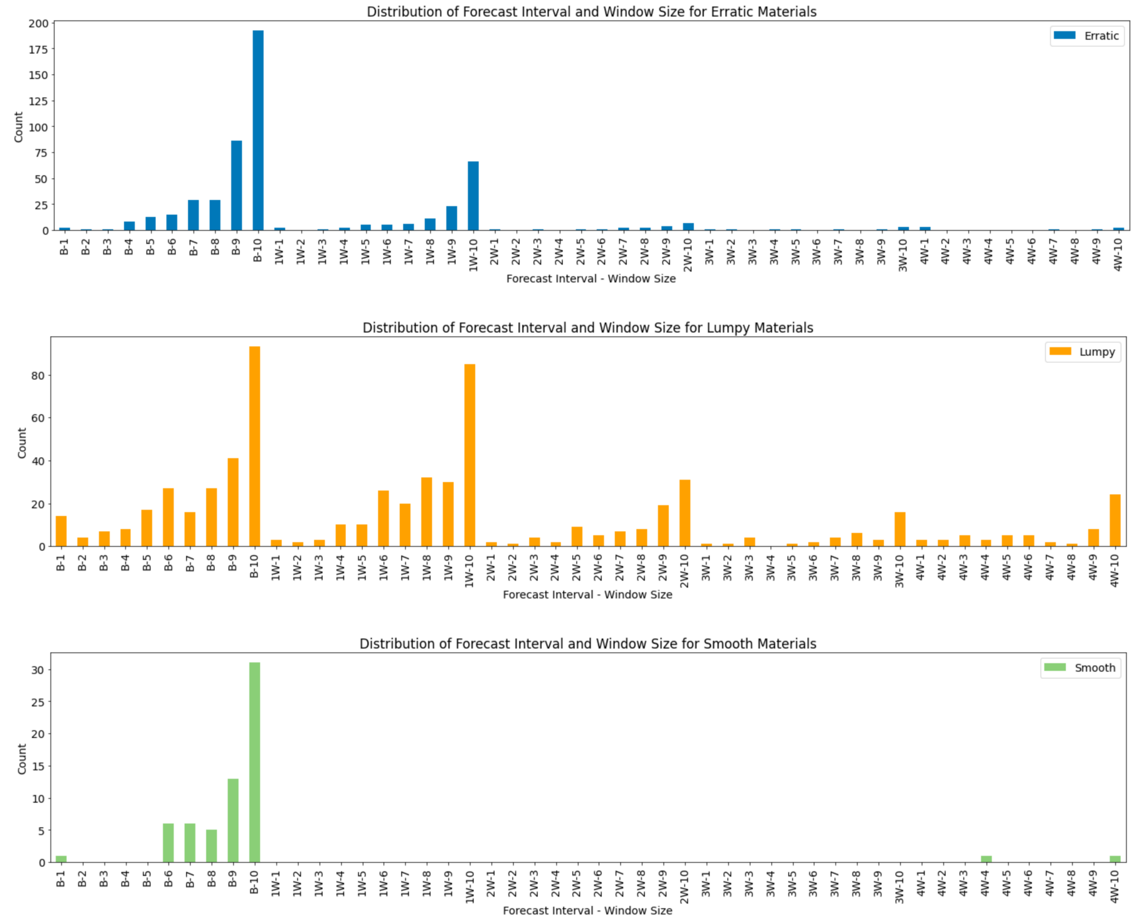

Figure 53: Distribution of optimal forecast interval and window size combination after optimization.

As mentioned earlier in this section, a search space of four types of forecast intervals ("B" for business days, "1W" for weekly, "2W" for every second week, etc.) and 10 time units of window size results in 50 permutations. The figure displayed in Figure 53 illustrates the distribution of number of materials based on their best-performing parameters for forecast interval and window size. Each y-axis value corresponds to the number of materials having the given x-axis value (combination of forecast interval and window size) as their optimal parameter combination.

For the erratic materials displayed in the blue graph, the majority of materials select the daily forecast interval ("B") as the most optimal and seek the longest window size to determine the predicted day. No higher forecast interval than weekly ("1W") is significantly popular for erratic materials. For the lumpy category displayed in the orange graph, one can observe a repeating increase in window size for all forecast intervals. In other words, regardless of the forecast interval, the majority of materials within the lumpy category seek the largest window size. The forecast interval that is most frequently chosen is the daily interval ("B"), although the weekly interval ("1W") is also quite commonly selected. Lastly, for the demand category *Smooth*, no higher forecast interval than daily ("B") is significantly represented, thus making forecast interval at a daily level the most relevant. In terms of window size, the majority of materials within the *smooth* category seek the largest window size.

Service Level Statistics

Figure 54 shows the average service level for the different replenishment policies, divided into the three demand categories. The AS-IS fixed ROP strategy with fixed ROP is visualized in blue bars, the Proposed Fixed ROP is visualized in orange bars, and the Basic Dynamic (OPTIMAL) ROP is visualized in green bars. As mentioned earlier in the paragraph *Search Space & Simulation Runs*, the Basic Dynamic (OPTIMAL) ROP solution consists of parameters such as forecast interval ("B" for daily, "1W" for weekly, "2W" for bi-weekly, etc.) and window size (number of time units used as a basis for prediction) optimized for each individual material.

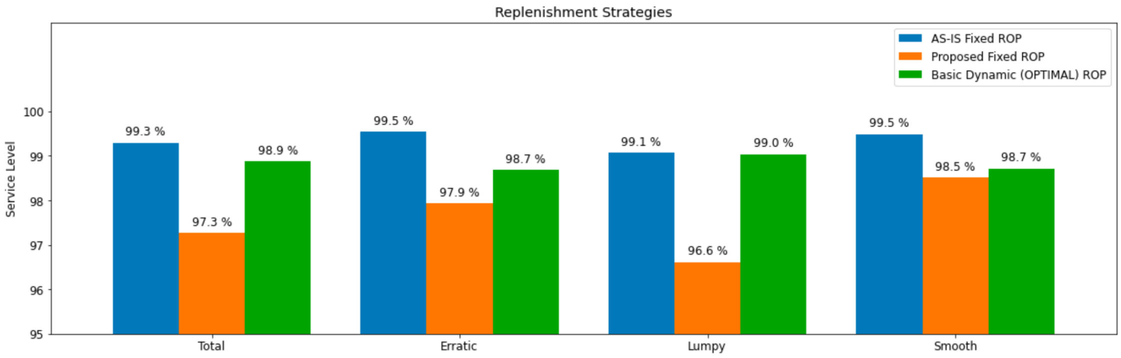

Figure 54: Average service level result of optimal aggregation type

Analyzing one demand category at a time, the performance of all three replenishment policies for the *smooth* category do perform well with all service levels above the 97% threshold. For the *lumpy* demand category, the AS-IS Fixed ROP strategy and the Basic Dynamic (OPTIMAL) ROP is superior to the Proposed Fixed ROP strategy, with 99.1% and 99% compared to 96.6%, which is below the threshold of 97%. For the *erratic* demand category, all the replenishment policies are above the threshold of 97%.

Average Inventory Transition

Figure 55 shows the resulting change in average inventory for all demand categories when transitioning from one replenishment policy to another replenishment policy. In the bar chart, there are two types of transitions. The one marked as a dark green represents the percentage change in average inventory when converting from the AS-IS Fixed ROP to the Proposed Fixed ROP. The light green marked bars represent the percentage change in average inventory when converting from the AS-IS Fixed ROP to the Basic Dynamic (OPTIMAL) ROP.

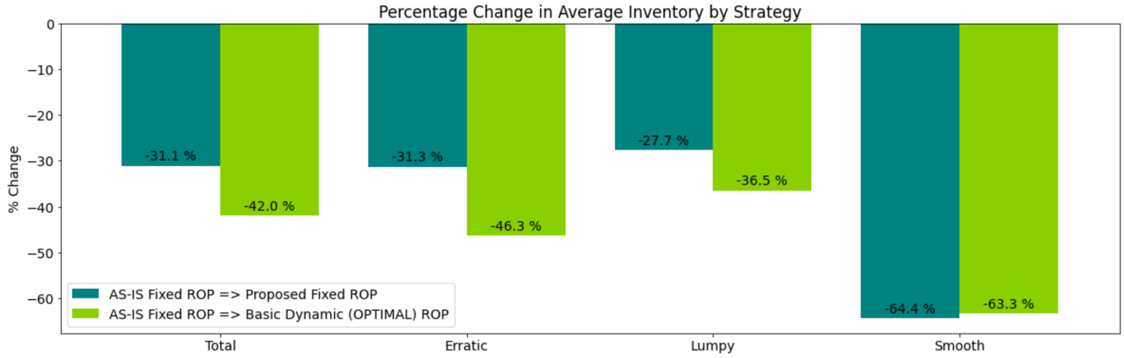

Figure 55: Percentage change in average inventory by strategy.

One can clearly observe from Figure 55 that going from the AS-IS fixed ROP to the Basic Dynamic (OPTIMAL) ROP is reducing the average inventory by the greatest amount for the demand categories *erratic* and *lumpy*. In contrast, the greatest decrease in average inventory for the *smooth* demand category occurs when transitioning from the AS-IS fixed ROP to the Proposed Fixed ROP.

Holistic View of The Change in Service Level and Average Inventory

The resulting change in average inventory must be seen in combination with the change in service level. Thus the results shown in Figure 54 and Figure 55 are combined into Table 12. Analyzing the results in Table 12 by demand category, starting with the erratic category, the greatest reduction in average inventory level will occur by switching to a Basic Dynamic (OPTIMAL) ROP. This will result in a reduction of −46.3% in the average inventory level, and consequently, a 98.7% service level.

For the lumpy demand category, the greatest reduction will also occur when switching to a Basic Dynamic (OPTIMAL) ROP. This will result in a reduction in the average inventory level of −36.5% and a service level of 99.0%. In contrast to the two previously mentioned demand categories, for the smooth demand category, the highest reduction in average inventory level will result from choosing the Proposed Fixed ROP. This will result in a reduction of −64.4% in average inventory level and a service level of 98.5%, compared to a −63.3% reduction in average inventory level and resulting service level of 98.7% with a Basic Dynamic (OPTIMAL) ROP. The difference in average inventory reduction between Proposed Fixed ROP and Basic Dynamic (OPTIMAL) ROP is minor, with a −1,1% difference.

Policy Transition	**Demand Category**	**% Change in Average Inventory**	**Percentage Point Change in SL**	**Resulting SL**
AS-IS Fixed ROP → Proposed Fixed ROP	Erratic	−31.3%	−1.6%	97.9%
	Lumpy	−27.7%	−2.5%	96.6%
	Smooth	−64.4%	−1%	98.5%
	Total	−31.1%	−2%	97.3%
AS-IS Fixed ROP → Basic Dynamic (OPTIMAL) ROP	Erratic	−46.3%	−0.8%	98.7%
	Lumpy	−36.5%	−0.1%	99.0%
	Smooth	−63.3%	−0.8%	98.7%
	Total	−42%	−0.4%	98.9%

Table 12: Holistic view of change in service Level and average inventory.

Materials That Surpasses The Service Level Constraint

Table 13 shows the number of materials that go from under 97% to above 97% when converting from the AS-IS Fixed ROP strategy to either a Proposed Fixed ROP or a Proposed Dynamic ROP (OPTIMAL) strategy.

- AS-IS Fixed ROP → Proposed Fixed ROP: This transition resulted in a total of six materials that surpassed the service level constraint of 97%. All of these six materials belong to the *lumpy* demand category.
- AS-IS Fixed ROP → Basic Dynamic (OPTIMAL) ROP: This transition resulted in 41 materials that surpassed the service level constraint. Nine of these were of the *erratic* category, 32 were of the *lumpy* category, and zero materials was of the *smooth* category.
- Proposed Fixed ROP → Basic Dynamic (OPTIMAL) ROP: This transition led to a total of 260 materials that surpassed the service level constraint of 97%. Splitted into demand categories, 90 of them were *erratic*, 166 were *lumpy*, and four materials were *smooth*.

Policy Transition	Total	Erratic	Lumpy	Smooth
AS-IS Fixed ROP → Proposed Fixed ROP	6	0	6	0
AS-IS Fixed ROP → Basic Dynamic (OPTIMAL) ROP	41	9	32	0
Proposed Fixed ROP → Basic Dynamic (OPTIMAL) ROP	260	90	166	4

Table 13: Strategy transition: from under 97% → over 97%

Materials That Drops Below The Service Level Constraint

Table 14 presents the quantities of materials that transition from having a service level over 97% to below 97% when converting from the AS-IS Fixed ROP strategy to either a Proposed Fixed ROP or a Basic Dynamic (OPTIMAL) ROP strategy.

- AS-IS Fixed ROP → Proposed Fixed ROP: This strategy transition resulted in a total of 229 materials that ended up with a service level below the constraint of 97%. Amongst these 229 materials, 83 of them belong to the *erratic* category, 142 belong to the *lumpy* category, and four materials belong to the *smooth* category.
- AS-IS Fixed ROP → Basic Dynamic (OPTIMAL) ROP: This transition resulted in two materials that went below the service level constraint. Both instances are of the *lumpy* category.
- Proposed Fixed ROP → Basic Dynamic (OPTIMAL) ROP: This transition had only one instance, and belongs to the *lumpy* category.

Policy Transition	Total	Erratic	Lumpy	Smooth
AS-IS Fixed ROP → Proposed Fixed ROP	229	83	142	4
AS-IS Fixed ROP → Basic Dynamic (OPTIMAL) ROP	2	0	2	0
Proposed Fixed ROP → Basic Dynamic (OPTIMAL) ROP	1	0	1	0

Table 14: Strategy transition: from over 97% → under 97%

Holding Costs Statistics

Figure 56 presents a graphical representation of the analysis conducted on holding costs. The figure showcases a bar chart with three bars representing different st: AS-IS Fixed ROP, Proposed Fixed ROP, and Basic Dynamic (OPTIMAL) ROP. The x-axis of the chart represents the different categories of inventory, namely *erratic*, *lumpy*, and *smooth*. These categories are based on the demand patterns of the inventory items. The y-axis denotes the holding costs in NOK (Norwegian Kroner). The heights of the bars indicate the magnitude of the holding costs for each category in the respective strategies.

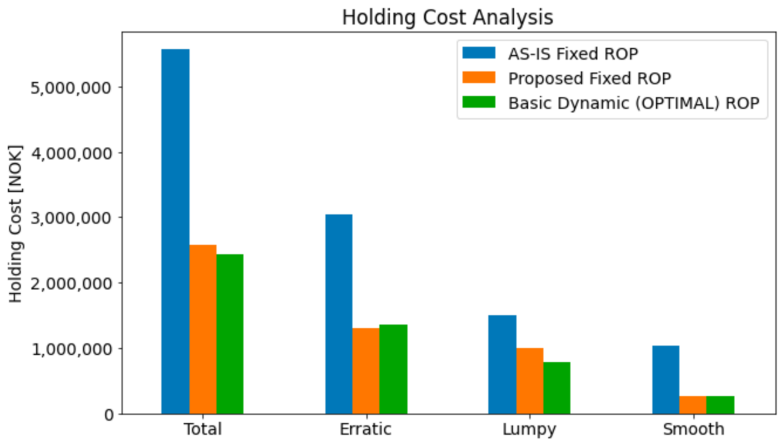

Figure 56: Holding cost analysis

The first strategy, AS-IS Fixed ROP, is represented by the tallest bars for all, indicating the highest holding costs across all demand categories. This strategy suggests that the current fixed operational strategy results in substantial holding costs of, in total, 5 565 119 NOK.

The second strategy, Proposed Fixed ROP, exhibits lower holding costs compared to the first strategy. The bars representing this strategy are shorter, indicating a reduction in holding costs for all three categories. This suggests that implementing a fixed operational strategy based on the Proposed Fixed ROP approach has yielded cost savings. In total, this strategy results in a holding cost of 2 568 329 NOK.

The third strategy, Basic Dynamic (OPTIMAL) ROP, is in total depicted by the shortest bar, with a total holding cost of 2 486 232 NOK, indicating the lowest holding costs among the three strategies. For the different categories, this strategy indicates a lower holding cost for the lumpy materials, and a higher holding cost for the erratic and smooth, compared to the Proposed Fixed ROP strategy.

In total, Figure 56 is emphasizing the cost-saving potential of implementing an optimized dynamic operational strategy. The results suggest that moving from the current fixed operational strategy to a more dynamic and optimized approach can lead to significant reductions in holding costs for the company.

Description	AS-IS Fixed ROP	Proposed Fixed ROP	Basic Dynamic (OPTIMAL) ROP
Total	5 565 119	2 568 329	2 429 790
Erratic	3 040 175	1 310 617	1 364 921
Lumpy	1 497 127	991 860	793 142
Smooth	1 027 817	265 851	271 725

Table 15: Replenishment policy impact on holding cost [NOK].

The table labeled "Replenishment Policy Impact on Holding Cost" (Table 15) provides the numerical data that corresponds to the previously presented Figure 56. It presents the holding costs in NOK (Norwegian Kroner) for three different operational strategies: AS-IS Fixed ROP, Proposed Fixed ROP, and Basic Dynamic (OPTIMAL) ROP, and the given categories.

Policy Transition	**Demand Category**	**% Change in Holding Cost**	**Percentage Point Change in SL**	**Resulting SL**
AS-IS Fixed ROP → Proposed Fixed ROP	Erratic	−56.9%	−1.6%	97.9%
	Lumpy	−33.7%	−2.5%	96.6%
	Smooth	−74.1%	−1%	98.5%
	Total	−53.8%	−2%	97.3%
AS-IS Fixed ROP → Basic Dynamic (OPTIMAL) ROP	Erratic	−55.1%	−0.8%	98.7%
	Lumpy	−47%	−0.1%	99.0%
	Smooth	−73.6%	−0.8%	98.7%
	Total	−56.3%	−0.4%	98.9%

Table 16: Holistic view of change in service level and holding cost.

5.3.4 Safety Factor Impact

As stated in Section 5.3.3, the safety factor used in the reorder point calculations in the simulation is 1.88. This section investigates the impact of changing the safety factor to corresponding higher service levels. The safety factors of 1.88, 2.05, and 2.33 correspond to 97%, 98%, and 99% service levels respectively, which can be seen in Appendix A.

Impact on Service Level - Fixed ROP

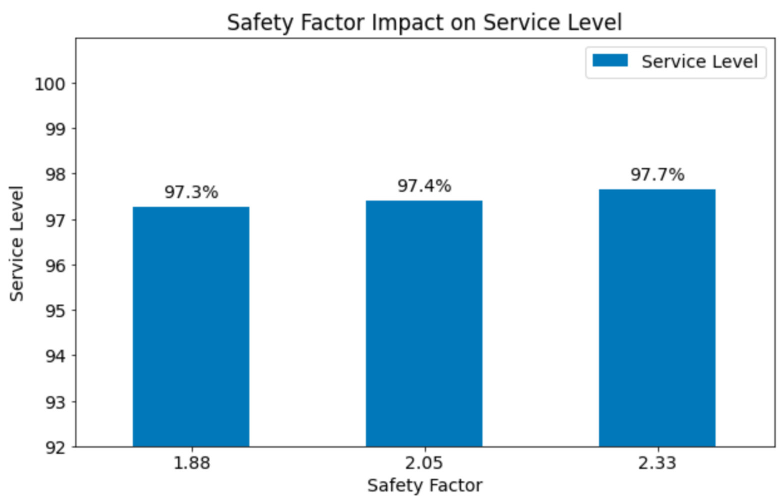

Figure 57: Safety factor and the corresponding service level for Proposed Fixed ROP.

The visualization presented in Figure 57 shows the effect of safety factors on service levels for Proposed Fixed ROP. The analysis is based on all the 1256 materials. The results show that, on average for all materials, a safety factor of 1.88 will result in a service level of 97.3%. For a safety factor of 2.05, it leads to an average service level of 97.4%. With a safety factor of 2.33, the service level increases to 97.7%. It can be observed that the service level increases with the safety factor, and all values are above the requirement of 97%.

Impact on Service Level - Basic Dynamic ROP

The analysis conducted in Figure 58, shows the safety factor impact on Basic Dynamic ROP for all 1256 forecast intervals. It can be observed that for all forecast intervals and safety factors, the corresponding service level is above the requirement of 97%. Additionally, the service level exhibits a linear increase as the forecast interval increases.

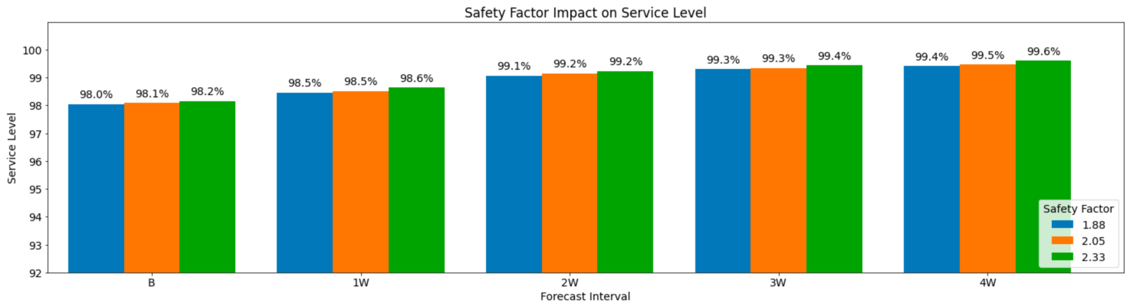

Figure 58: The visualization shows the effect of safety factors on service levels for Basic Dynamic ROP across all forecast intervals.

5.3.5 Forecasting Interval Impact on Basic Dynamic ROP

This section aims to explore the extent to which the optimal solution can be relaxed and examine the possibility of generalizing the forecast interval. As described in Section 5.3.3, the optimal solution involves setting specific forecast intervals and window sizes for each individual material, which leads to complexity and potentially results in high costs. The different materials were forced to use B, 1W, 2W, 3W, and 4W, and thus it was possible to see how they performed compared to the optimal solution and if a given forecast interval applied to all materials can perform equally well or close to the fully customized and optimal solution.

Distributions of Window Size

Figure 59 shows the distribution of the selected window size for all materials, for each "locked" forecast interval. "locked" meaning the Basic Dynamic ROP is locked to using the given forecast interval, but chooses the optimal window-size. Note that the scale of the y-axis is different between the plots. Starting with the *erratic* demand category, a noticeable observation is that the distribution of window sizes exhibits an exponential increase towards the largest window size of 10 time units, which applies to all levels of forecast interval for the *erratic* materials.

For the *lumpy* demand category, the distribution of window sizes exhibits a more linear increase towards the largest window size of 10 time units. The distribution of window sizes is more evenly compared to the *erratic* category. In contrast, for the *smooth* demand category, the distribution of window size is highly skewed, with window sizes of six and higher the most represented. Moreover, the majority of materials within the *smooth* demand category select the 10 time units of window size. In total, looking at all demand categories, the pattern of distribution is similar for all the levels of window size. This suggests that for all demand categories, the majority of materials will tend to opt for the largest window size, regardless of the forecast interval utilized.

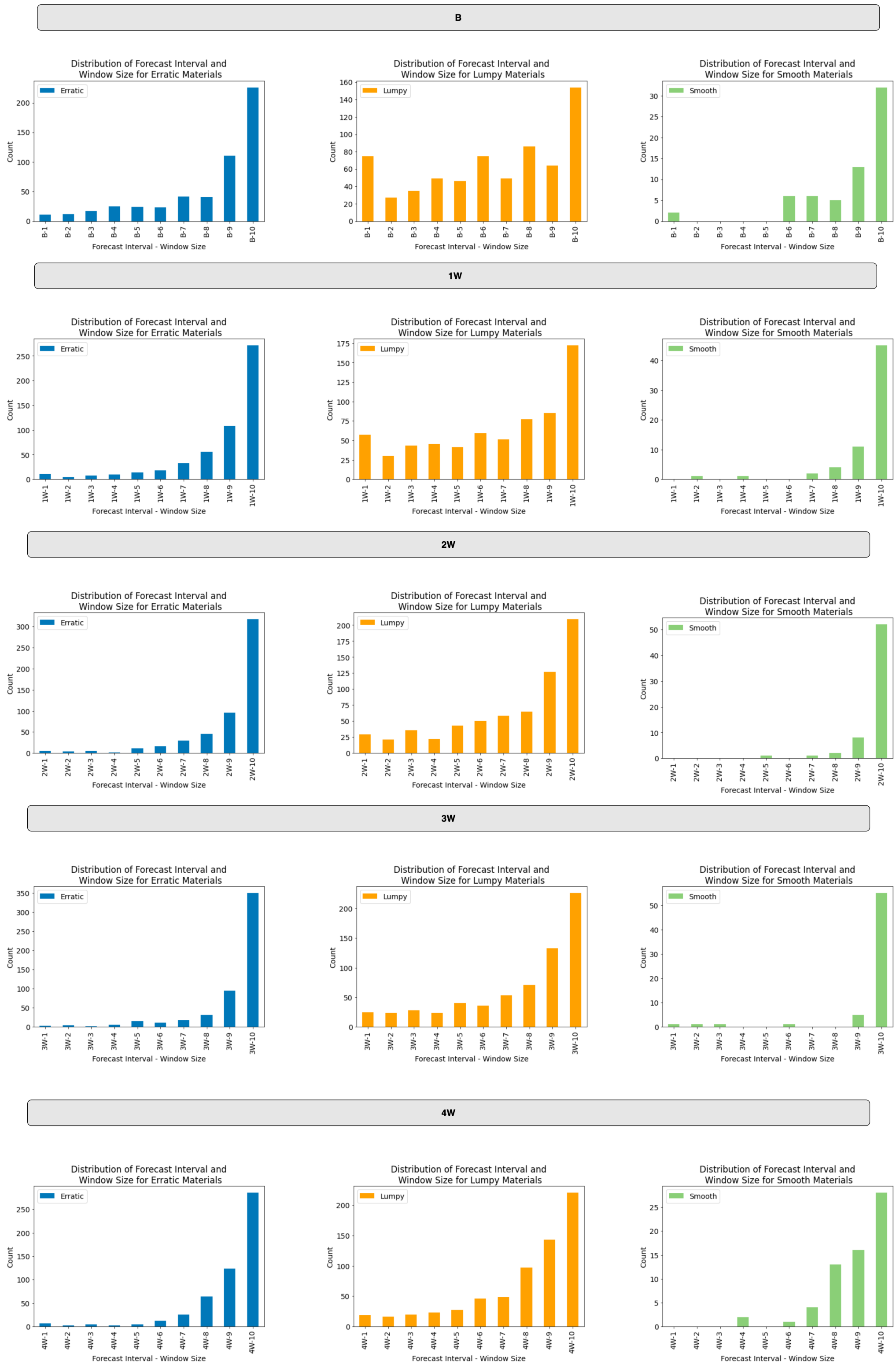

Figure 59: Distributions of window size for each "locked" forecasting interval (Basic Dynamic ROP)

Service Level

The impact of the forecast interval on the simulated inventory and its resulting service level is illustrated in Figure 60. The grouped bar plot shows the different resulting service levels, divided into the demand categories, and in total as seen to the far left.

It is evident that there exists a positive correlation between the forecast interval and the resulting service level. This phenomenon is consistent across all demand categories and therefore applies to all materials in total. Comparing the different forecast intervals with the fully customized and optimal solution ("OPTIMAL"), the weekly forecast interval ("1W") results in a better service level for demand categories *erratic* and *smooth*, but not *lumpy*. In total, the "OPTIMAL" solution, visualized in a brown-colored bar, results in a 0.4% higher service level than the weekly forecast interval ("1W"), as visualized as a blue-colored bar.

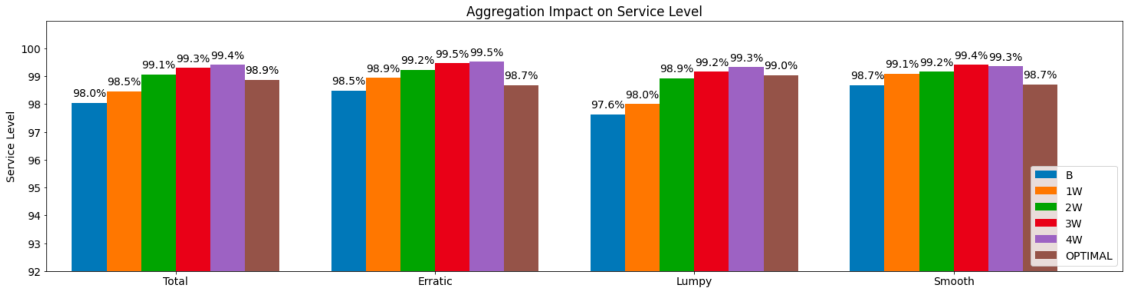

Figure 60: Forecast interval impact on service level

Average Inventory

The impact of the forecast interval on the simulated inventory and its resulting percentage change in average inventory compared to the AS-IS strategy with fixed ROP is illustrated in Figure 61. The grouped bar plot shows the resulting change, divided into the demand categories, and summarized to a total as seen to the far left.

The percentage change in average inventory level must be analyzed in combination with the achieved service level as visualized in Figure 60. Examining the performance of different forecast intervals for the three demand categories, and disregarding the optimal solution ("OPTIMAL"), the daily forecast interval, denoted as "B", outperforms the higher levels of forecast intervals in terms of reducing the average inventory for all three categories of demand (*erratic, lumpy, smooth*).

However, as can be seen from Figure 60, daily forecast interval performs the worst in terms of service level performance for all categories of demand. Here, the daily forecast interval results in a service level of 98.7% for the *smooth* category, 97.6% for the *lumpy* category, and 98.5% in the *erratic* category. Comparing daily forecast interval with weekly forecast interval, a weekly forecast interval increases the service level by 0.5 percentage points up to 98.5% in total, while reducing the average inventory by 36.6%. Thus, the reduction in the average inventory level for the weekly forecast interval is very close to the optimal solution, while maintaining a service level of 98.5% overall.

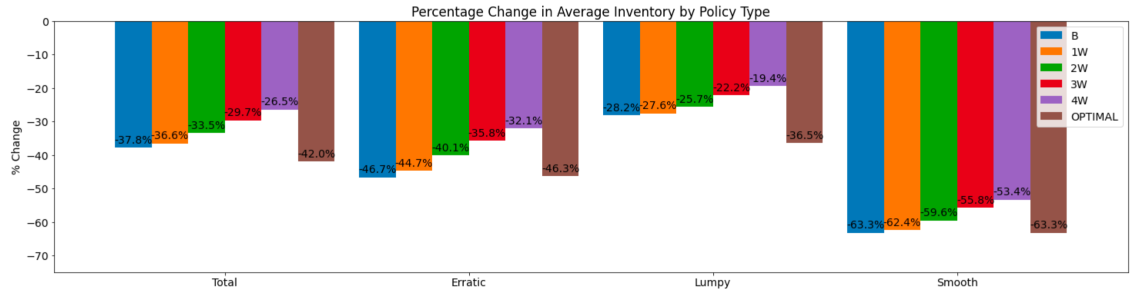

Figure 61: Aggregation impact on average inventory

Materials That Surpasses The Service Level Constraint

As described earlier under section 5.1.4, LC HMN operates with a service level constraint of 97%. Table 17 shows the distribution of materials that have gone from under 97% in service level to equal or above 97% service level when going from AS-IS fixed ROP to dynamic ROP. Materials going from under 97% in service level to above 97% in service level is seen as a positive impact on the inventory performance. One can observe that the "4W" (i.e. every fourth week) will make 36 materials go from under 97% service level to above, where 30 materials of those belong to the *lumpy* type of demand. The daily forecast interval ("B") is the worst-performing among them all, with 25 materials in total going from under 97% to above 97%.

When looking at the demand categories in Table 17, it is clear that the majority of materials in the *lumpy* category achieve over 97% service level for the most materials, regardless of the forecast interval. The smooth category has the fewest movements, with zero materials for all the forecast intervals.

Forecast Interval	Total	Erratic	Lumpy	Smooth
B	25	4	21	0
1W	27	5	22	0
2W	30	7	23	0
3W	33	7	26	0
4W	36	6	30	0
OPTIMAL	41	9	32	0

Table 17: From under 97% → over 97%

Materials That Drops Below The Service Level Constraint

Table 17 shows the distribution of materials that goes from over 97% service level to below 97% when going from the AS-IS fixed ROP to dynamic ROP with the tested forecast intervals. In this case, the "OPTIMAL" forecast interval, where each material has the optimal forecast interval and window size, performs the best with only two materials going from over 97% service level to below. Besides the "OPTIMAL" forecast interval, the "4W" forecast interval is the best performing with only six materials that move under the limit of 97%. The worst performing forecast interval is the daily forecast interval ("B"), where 44 materials go from above 97% to under 97%.

In terms of demand categories, the *lumpy* demand category is the one with the most materials that go from over 97% service level to below 97% service level regardless of forecast interval level. In contrast, the *smooth* category is the category with the fewest materials that goes from over 97% service level to below 97%.

Forecast Interval	Total	Erratic	Lumpy	Smooth
B	44	2	41	1
1W	33	4	28	1
2W	14	2	11	1
3W	9	0	9	0
4W	6	0	6	0
OPTIMAL	2	0	2	0

Table 18: From over 97% → under 97%

Holding Cost Statistics

The plot titled "Aggregation Impact on Holding Cost" (Figure 62) illustrates the impact of different aggregations on holding costs for the Basic Dynamic ROP. The plot is presented as a bar chart, with each bar representing a specific forecast interval: "B," "1W," "2W," "3W," "4W," and "OPTIMAL." The x-axis of the chart displays the categories of inventory, namely "Erratic," "Lumpy," and "Smooth." The y-axis represents the holding costs in NOK (Norwegian Kroner). The heights of the bars indicate the magnitude of the holding costs for each inventory category under the corresponding forecast interval.

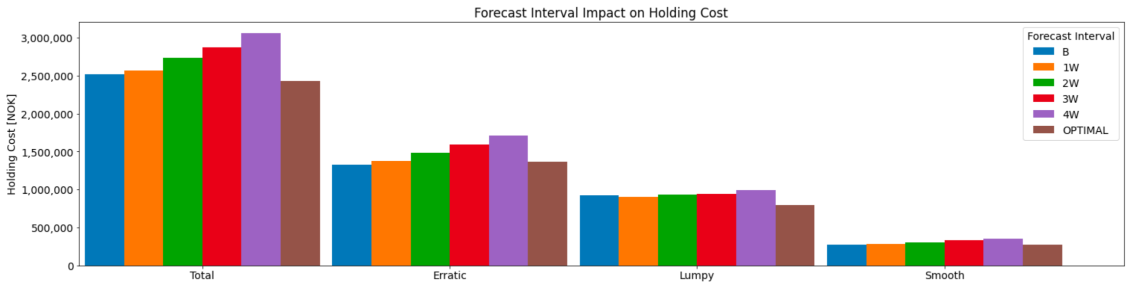

Figure 62: Aggregation impact on holding cost for Basic Dynamic ROP

For the total analysis, the bars corresponding to the "B", "1W" and "OPTIMAL" approaches are the lowest, indicating the lowest holding costs among all the aggregations.

Description	B	1W	2W	3W	4W	OPTIMAL
Total	2,515,767	2,566,095	2,731,473	2,874,717	3,059,023	2,429,790
Erratic	1,324,185	1,377,541	1,488,888	1,596,022	1,714,509	1,364,921
Lumpy	919,964	901,140	934,408	944,904	990,928	793,142
Smooth	271,616	287,414	308,175	333,791	353,585	271,725

Table 19: Forecast interval impact on holding cost [NOK].

The table labeled "Forecast Interval on Holding Cost" (Table 19) presents the numerical data that corresponds to the previously presented Figure 62. It provides a detailed breakdown of the holding costs in NOK (Norwegian Kroner) for different forecasting interval approaches across three inventory categories: "Erratic," "Lumpy," and "Smooth."

5.3.6 Advanced Dynamic Reorder Point

In this section, the fifth strategy presented in Figure 3.5 is analyzed, this being the "Advanced Dynamic Reorder Point". Advanced Dynamic Reorder Point is a dynamic reorder point strategy utilizing advanced forecasting models. It enables the investigation of the potential benefits that may arise from implementing more advanced demand forecasting techniques. The strategy is analyzed of each of the three selected representative materials for each category found in Figure 39. Firstly, the advanced demand forecasting (used of the Advanced Dynamic Reorder Point strategy) results are presented. Further on, the simulation outcome from each analyzed advanced demand forecasting technique is presented.

5.3.6.1 Advanced Demand Forecasting

Several methods from various disciplines, including statistical models, machine learning models, and deep learning models, were tested to predict the three time series that represent different demand categories. The time series data is aggregated to weekly data, so for example, 365 daily data points will result in 52 data points of weekly data. All models predicted iteratively One-step-ahead (one aggregated week) ahead in time for a whole year.

The models were selected based on their different characteristics. The simplest method, the Naive method, assumes that the predicted interval is the same as the last observed interval (Hyndman and Koehler, 2006). The naive model will be used as a benchmark for comparison with the other models. The second model is Holt-Winters (Section 2.2), a version of exponential smoothing with trend and seasonality (added in an additive way). Holt-Winters is a model from the statistical domain.

The third model, SARIMAX (Section 2.5.1), is a more complex method belonging to the domain of machine learning, that uses a combination of auto-regression, differencing, and moving average to model the time series. It also differs from the previously mentioned models in that it can include exogenous variables. The fourth and final model, LSTM (Section 2.5.2), is designed to handle time series data with long-term dependencies. LSTM models use a memory cell to remember past observations and a set of gates to control the flow of information into and out of the cell. LSTM is a type of neural network, making it a part of the deep learning domain.

All four models represent a domain of time series prediction. By exploring and understanding the strengths and weaknesses of each model applied to the case study's data, a selection of the most appropriate method for a specific time series can be performed. To compare the performance of the selected models, the three metrics RMSE, MAE, and MASE (Section 2.3.3) were chosen to measure the accuracy of the models.

Erratic Representative Material - 4001095

Figure 63 shows the observed line of demand and the predicted line of demand for the four models applied for material 4001095. The material is characterized as erratic, indicating that it has high variation in quantity, but is consistent in time intervals between occurrences of demand. This can also be seen in the observed demand (plotted as the blue line). The line has high peaks with demand over 1000 units at periods, but it also has periods with demand close to zero, making it hard to predict.

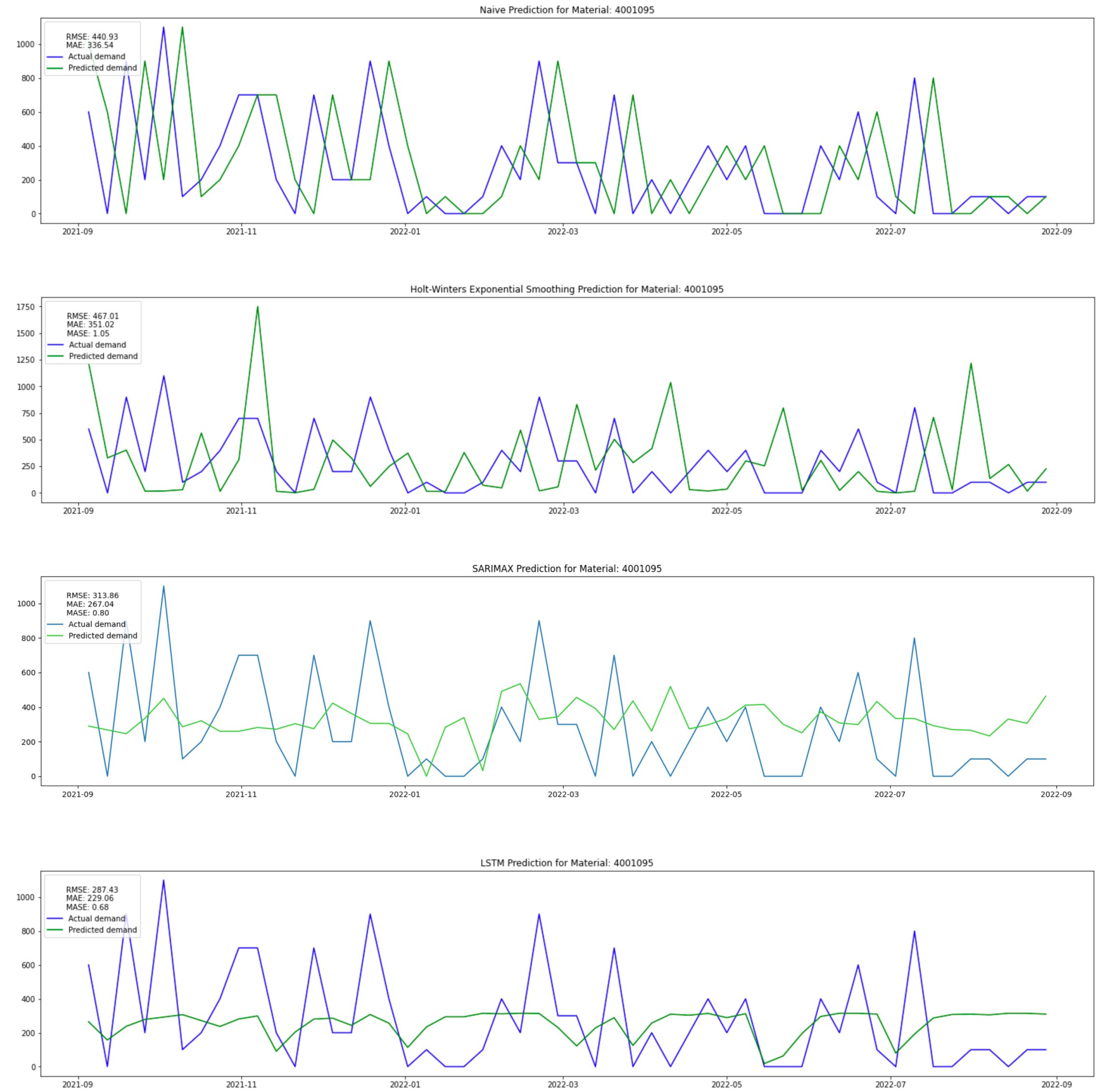

Figure 63: Predictions for erratic representative material (4001095)

Looking closer at the performance of the different models, the three metrics RMSE, MAE, and MASE must be considered together. Looking at the MASE metric, the Holt-Winters does perform worse than the Naive method, making it irrelevant for further use. The best-performing models are SARIMAX and LSTM, where LSTM is slightly superior to the SARIMAX model, with a MASE score of 0.68 (as seen in Table 20). The LSTM model is the best-performing model because it strikes a straight line, with some exceptions. The prediction thus becomes a kind of horizontal straight line, which aligns itself roughly with the average of the high and low observed values.

Model	RMSE	MAE	MASE
Naive	440.93	336.54	1.00
Holt-Winters	467.1	351.02	1.05
SARIMAX	313.86	267.04	0.80
LSTM	287.43	229.06	0.68

Table 20: Forecast performance results for material 4001095

Lumpy Representative Material - 4012198

Figure 64 shows the observed and predicted demand for the *lumpy* material. As derived under Section 5.2.3, a lumpy material has high variation in quantity and intervals between occurrences of demand. This can be observed from Figure 64, as the observed quantity varies with 1000 units, and the time interval between occurrences of demand varies from weeks to several months.

Figure 64: Predictions for lumpy representative material (4012198)

The lumpy representative material has a high representation of non-existent demand, meaning that the time series data contains a large number of zeros. With that in mind, the performance of the selected models is contrasting. The SARIMAX and the LSTM have a MASE score of 1.03, meaning that they perform slightly worse than a naive prediction. The Holt-Winters Exponential Smoothing is the best performing among the models, with a MASE of 0.77.

Model	RMSE	MAE	MASE
Naive	832.05	500.00	1.00
Holt-Winters	679.63	394.58	0.77
SARIMAX	631.96	525.85	1.03
LSTM	604.08	526.32	1.03

Table 21: Forecast performance results for material 4012198

Smooth Representative Material - 4003841

The third and last material, 4003831, represents the *smooth* demand category. Materials in this category have a low variance in demand size and a consistent time interval between occurrences of demand. It can be observed that the median demand for this material is around 11,000 units, with maximum demand reaching up to 16,000 and minimum demand dropping down to 5,000 units. However, the latter values are individual occurrences and not something that repeats throughout the given year being observed.

Figure 65: Predictions for smooth representative material (4003841)

Materials that are categorized as *smooth* have more consistent demand, and therefore more data points from the data set utilized. This provides a better basis for detecting any seasonal variations and trends, which can be important inputs for a prediction model. When interpreting the results in Table 25, none of the selected models perform exceptionally well, but SARIMAX is the best-performing model with a MASE score of 0.8, meaning it is 20% better than a naive forecast.

Model	RMSE	MAE	MASE
Naive	2671.57	1723.08	1.00
Holt-Winters	2368.98	1715.15	0.98
SARIMAX	1966.23	1398.77	0.80
LSTM	2253.72	1701.19	0.97

Table 22: Forecast performance results for material 4003841

5.3.6.2 Simulation Results

In this section, the impact of advanced forecasting methods on the dynamic ROP strategy will be explored, focusing on weekly aggregated data. The comparison will be made against the Basic Dynamic ROP strategy with one-week forecasting intervals. The specific version of the Basic Dynamic ROP strategy is referred to as "SMA (1W-10)".

Erratic Representative Material - 4001095

Table 23 shows the simulation results of various demand forecasting models for material 4001095. The models are evaluated based on three performance metrics: mean absolute scaled error (MASE), service level (SL), and average inventory.

MASE measures the accuracy of the models in predicting demand, while SL indicates the percentage of time that the inventory is able to meet the demand as a result of the simulation. Average inventory describes the resulting average inventory as a result of the simulation.

The models compared in the table include Naive, Holt-Winters, SARIMAX, LSTM, and SMA (1W-10). The Naive model is used as a baseline, and its performance is shown to have a MASE of 1.00, an SL of 100%, and an average inventory of 3,189. The other models are then evaluated based on how much they can improve upon the Naive model's performance.

Model Classification	Model	MASE	SL	Average Inventory
Baseline model	Naive	1.00	100 %	3 189
Advanced Models	Holt-Winters	1.05	100 %	2 851
	SARIMAX	0.80	99.82 %	2 645
	LSTM	0.68	100 %	2 414
Basic Model	SMA (1W-10)	1.02	100 %	2 223

Table 23: Simulation results based on demand forecasting model for erratic representative material (4001095).

The most precise advanced demand prediction model is shown to be LSTM, with a MASE score of 0.68. This results in the simulation results of service level at 100 % and average inventory at 2414. The SMA model scores worse in MASE score, although has the lowest simulated average inventory of 2 223.

Figure 66 visualizes the simulation of the most accurate prediction model for material 4001095 based on the MASE score. The in-stock integral is 627 620 and the stockout integral is 0, indicating no stockouts. The calculated initial stock is 3407.

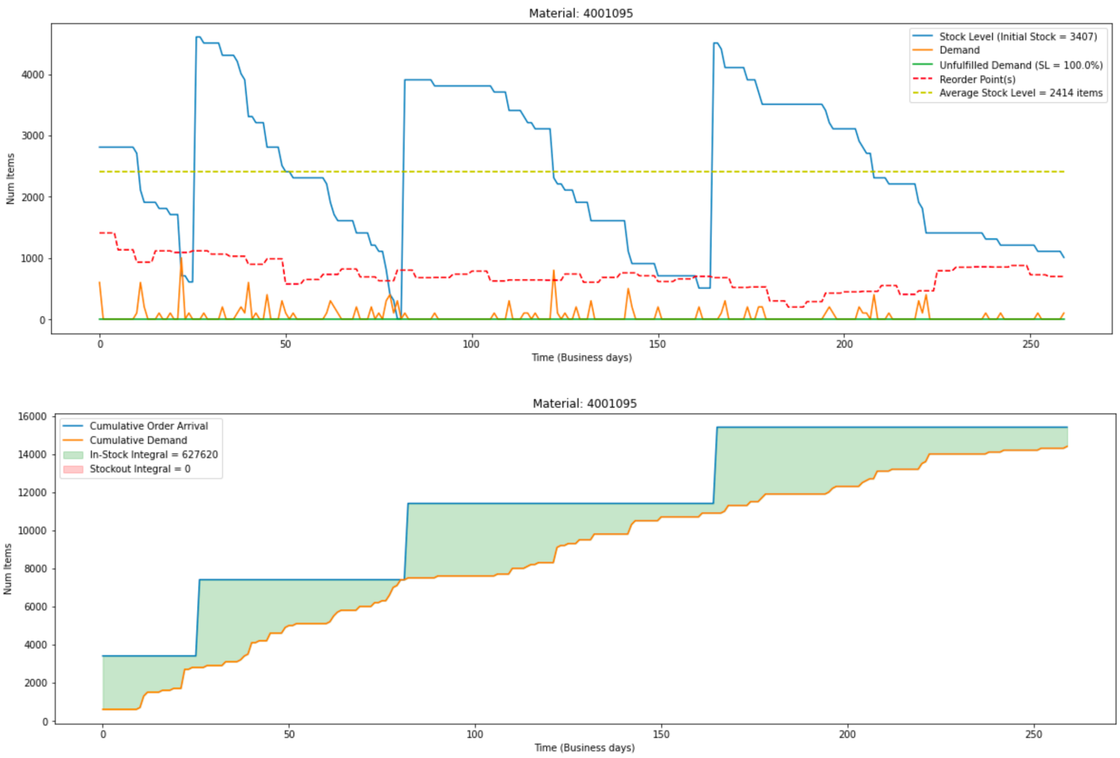

Figure 66: Most accurate model: LSTM - erratic representative material (4001095).

Figure 67 visualizes the simulation of the optimal SMA model for material 4001095 based on the MASE score. The in-stock integral is 578 020 and the stockout integral is 0, indicating no stockouts. The calculated initial stock is 2847.

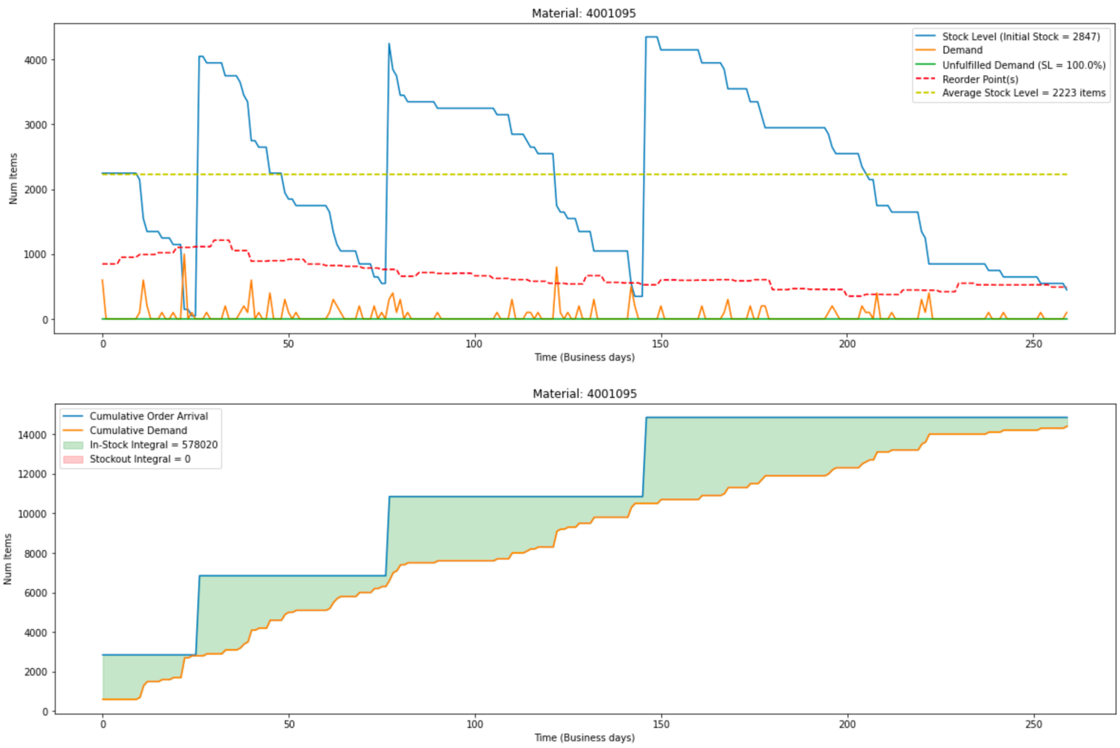

Figure 67: SMA (1W-10) model - erratic representative material (4001095).

Lumpy Representative Material - 4012198

In Table 24 the most precise advanced demand prediction model is shown to be Holt-Winters, with a MASE score of 0.77. This results in the simulation results of service level at 96.78 % and average inventory at 5 066. The SMA model scores worse in MASE score, although has the lowest simulated average inventory of 5 014.

Model Classification	Model	MASE	SL	Average Inventory
Base model	Naive	1.00	100 %	5 997
Advanced Models	Holt-Winters	0.77	96.78 %	5 066
	SARIMAX	1.03	100 %	5 198
	LSTM	1.03	100 %	5 019
Basic Model	SMA (1W-10)	0.95	100 %	5 014

Table 24: Simulation results based on demand forecasting model for lumpy representative material (4012198).

Figure 68 visualizes the simulation of the most accurate prediction model for material 4012198 based on the MASE score. The in-stock integral is 1 210 212 and the stockout integral is 1 692. The calculated initial stock is 4452.

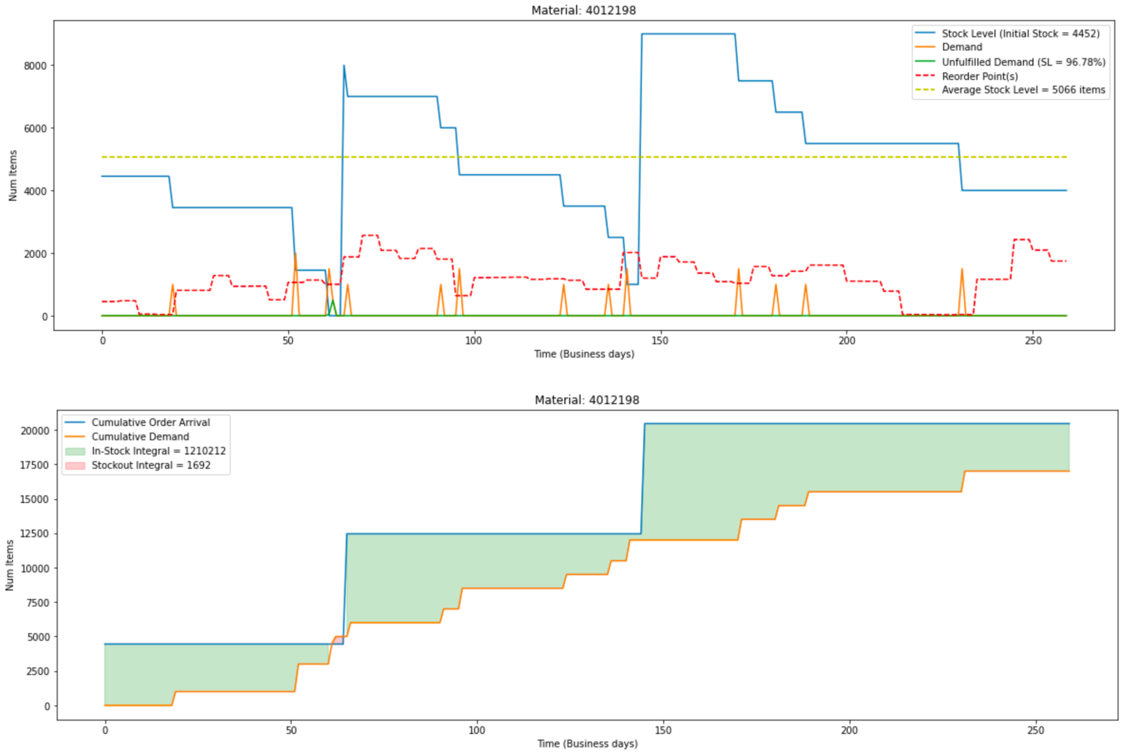

Figure 68: Most accurate model: Holt-Winters Exponential Smoothing - lumpy representative material (4012198).

Figure 69 visualizes the simulation of the optimal SMA model for material 4012198 based on the MASE score. The in-stock integral is 1 303 680 and the stockout integral is 0, indicating no stockouts. The calculated initial stock is 5618.

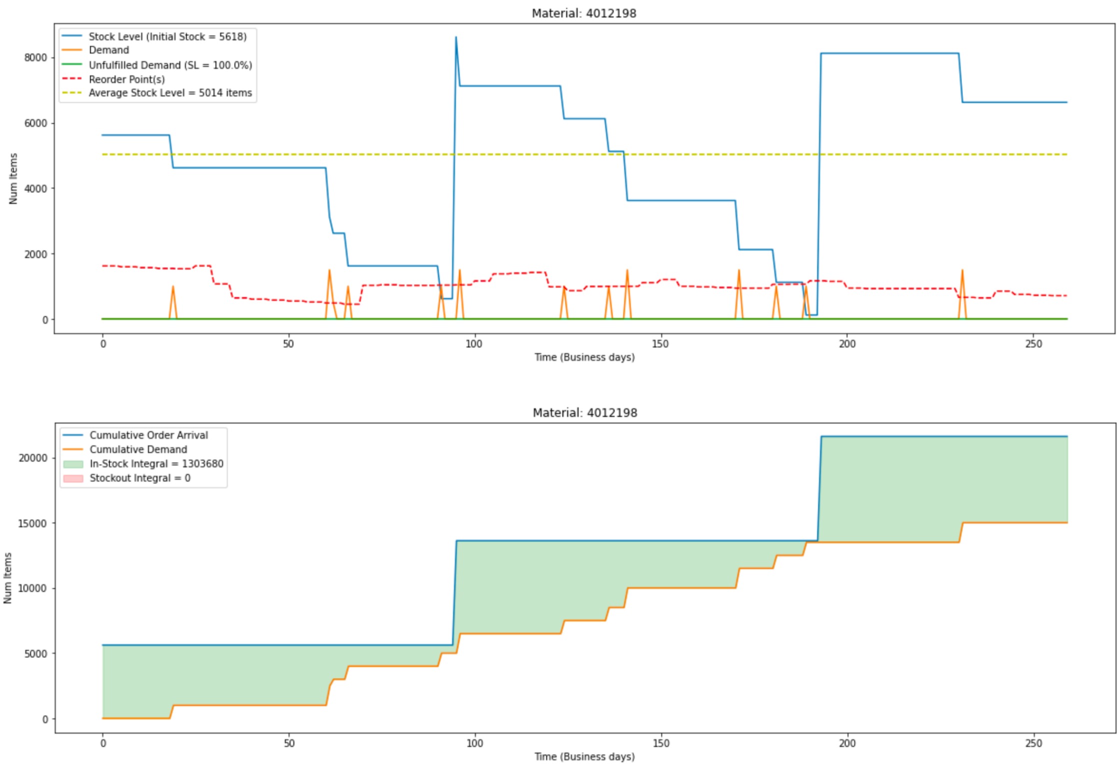

Figure 69: SMA (1W-10) model - lumpy representative material (4012198).

Smooth Representative Material - 4003841

In Table 25 the most precise advanced demand prediction model is shown to be SARIMAX, with a MASE score of 0.80. This results in the simulation results of service level at 99.81 % and average inventory at 23 952. The SMA model scores higher in MASE score with 0.73 and has a lower average inventory of 18 998.

Model Classification	Model	MASE	SL	Average Inventory
Baseline Model	Naive	1.00	100 %	24 527
Advanced Models	Holt-Winters	0.98	100 %	27 058
	SARIMAX	0.80	99.81 %	23 952
	LSTM	0.97	100 %	23 147
Basic Model	SMA (1W-10)	0.73	99.35 %	18 998

Table 25: Simulation results based on demand forecasting model for smooth representative material (4003841).

Figure 70 visualizes the simulation of the most accurate prediction model for material 4003841 based on the MASE score. The in-stock integral is 6 081 932 and the stockout integral is 952. The calculated initial stock is 23 952.

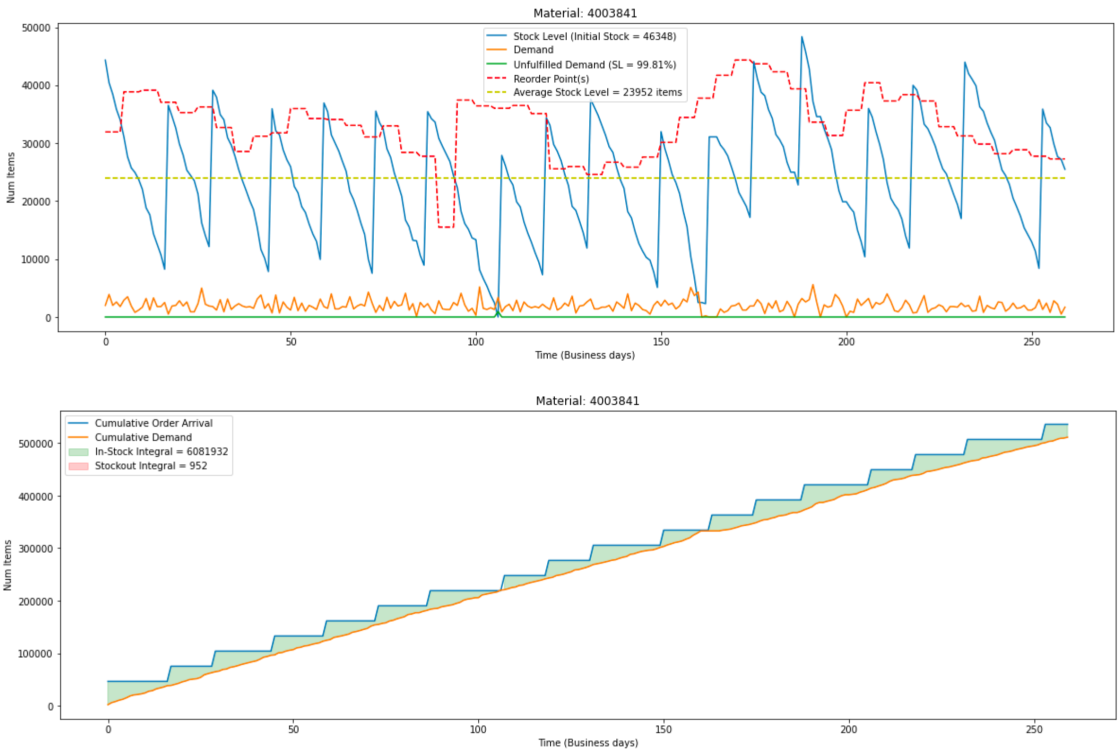

Figure 70: Most accurate model: SARIMAX - smooth representative material (4003841)

Figure 71 visualizes the simulation of the optimal SMA model for material 4003841 based on the MASE score. The in-stock integral is 4 617 403 and the stockout integral is 9 563. The calculated initial stock is 41679.

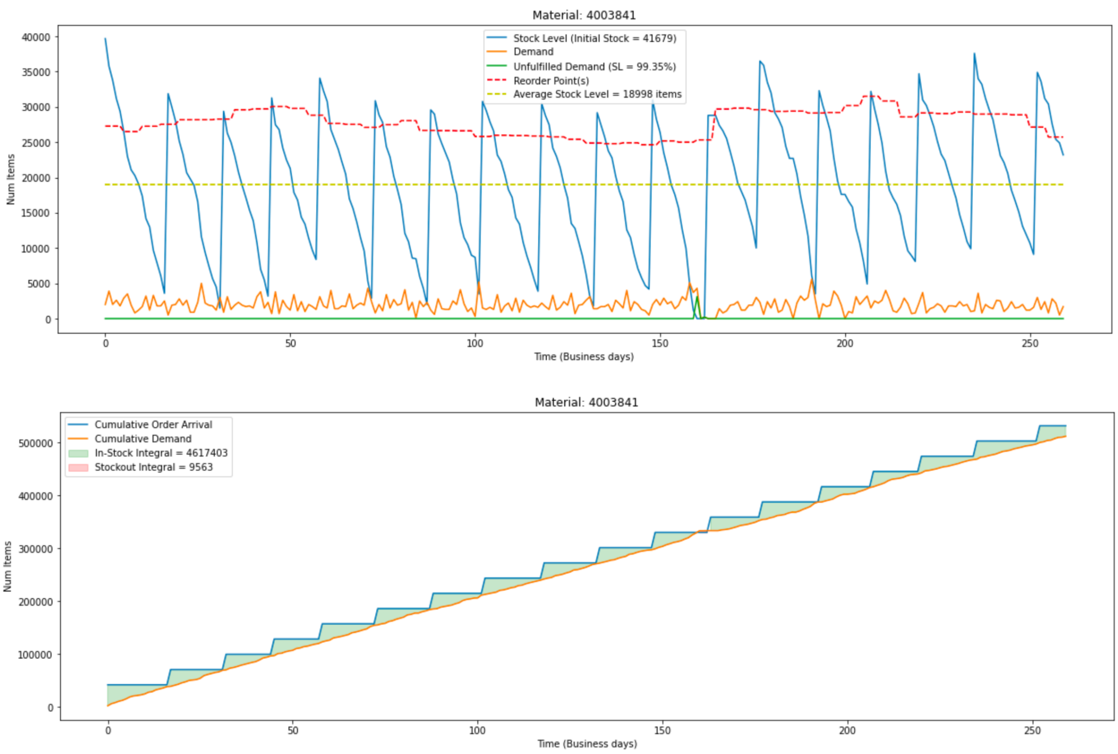

Figure 71: SMA (1W-10) model - smooth representative material (4003841)

6 Discussion

The following section analyzes and interprets the key findings of the study, specifically addressing the research questions outlined in Section 1.3. It involves a detailed examination of the results, establishing connections with the existing literature, and exploring the implications of the findings. For the sake of clarity, this section is divided into six subsections, each dedicated to one or more research questions. Table 26 illustrates the link between the research questions and the subsections. Section 6.3 addresses the simulation model, which served as the basis for the findings related to both research question three and research question four.

Research Questions	**Discussed in Section**
RQ1: What is the state-of-the-art within demand forecasting for inventory management?	6.1
RQ2: How can the inventory be classified?	6.2
RQ3: How can the AS-IS fixed reorder point be improved through dynamic reorder point?	6.3 6.4 6.5
RQ4: What is the impact of implementing advanced forecasting methods for the dynamic reorder point?	6.3 6.6

Table 26: Research questions and their respective sections for discussion

Within each subsection, the relevant results are briefly summarized before engaging in a comprehensive discussion that emphasizes their connection to the research questions. Furthermore, unexpected or surprising discoveries are identified and the theoretical and practical implications of the results are explored. Throughout the discussion, a critical evaluation is conducted, focusing specifically on the limitations and weaknesses of the study in relation to the research questions.

6.1 The State-of-The-Art in Demand Forecasting for Inventory Management

The key to improving inventory performance is to use forecasts as a better anchor point to improve stock control (Bacchetti and Saccani, 2012). In the systematic review of the literature conducted in Section 3.1, it was discovered that Barrow and Kourentzes (2016) employed five different weighted statistical models to forecast the safety stock. The study conducted by R. Snyder (2002) explored various approaches to forecast sales of slow- and fast-moving car parts. Harvey and Ralph D. Snyder (1990) delved into the investigation of exponential smoothing for non-stationary time series in order to forecast safety stock. Willemain et al. (2004) successfully developed a bootstrapping forecasting model, enabling accurate prediction of the entire distribution of lead-time demand. Additionally, Tasdemir and Hiziroglu (2019) utilized two forecasting models to address missing data in specific periods and subsequently forecasted lot sizes, ultimately proposing a fixed period quantity dynamic lot size method.

The literature review findings (Section 4.2) emphasize the importance of considering the application area and data characteristics when choosing a forecasting method (Moroff et al., 2021a). It is clear that demand forecasting plays a crucial role in determining the safety stock and other formulas presented in Table 1. The content analysis performed in Section 4.2, classified demand forecasting methods into two primary categories: statistical methods and machine learning methods.

As observed in Figure 25 in Section 4.1, 58% of the articles reviewed in Section 4.2 utilize a statistical method, while the remaining 42% use machine learning methods or a combination of both. On the other hand, one can observe from Figure 24 that there is an increased number of articles that apply only the machine learning method or combine it with statistical methods. On the contrary, there is a decline in the number of articles exclusively utilizing statistical methods. It is important to note that the resulting descriptive statistics may be biased, as the articles were selected based on the criteria defined by the authors in Section 3.1. Thus, the observations should not be generalized to all publicly published articles within this study research area. However, the growing trend of machine learning model applications can be attributed to their ability to learn from data and make predictions without relying on assumptions about the underlying data structure (Karimnezhad and Moradi, 2016)

The utilization of machine learning has been shown to improve demand forecasting by enabling algorithms to learn from new data and adapt accordingly. This approach considers a broader range of factors and allows nuanced predictions, leading to a significant improvement in the accuracy of demand forecasts (Moroff et al., 2021a; Tang and Ge, 2021). Machine learning algorithms can handle large amounts of data more efficiently than statistical methods, which can be particularly useful for time series data with a large number of observations. Their capacity to capture complex patterns, including nonlinear relationships, surpasses that of statistical methods, making them superior in generating accurate predictions (N. Li et al., 2021; Pacella and Papadia, 2021).

One of the key advantages of employing machine learning in demand forecasting is the ability to incorporate diverse data sources (Moroff et al., 2021a). Unlike traditional methods that rely on limited datasets, such as outbound logistics, machine learning methods can leverage additional data such as demographics, clinical information, and economic indicators. By considering a broader spectrum of influencing factors, machine learning models provide a comprehensive view of demand drivers, leading to more precise forecasts (Tang and Ge, 2021)

Statistical models are based on the assumption that data have a certain structure, such as linearity or normal distribution, which may not always be the case (R. Snyder, 2002; Harvey and Ralph D. Snyder, 1990). This can lead to inaccurate predictions if the data do not conform to the assumed structure. To manage the impact of the data distribution on demand forecast, a weighted combination of different statistical models can exploit the empirical distribution of forecast errors (Barrow and Kourentzes, 2016), thus exceeding the statistical assumptions that the data are normally distributed.

In the context of forecasting time series data, it is important to consider that machine learning methods may not always be the optimal choice. While these methods have gained popularity for their ability to handle complex patterns and make accurate predictions, they have certain drawbacks that need to be addressed. Machine learning methods can be challenging to understand and interpret due to their complexity and can be prone to overfitting if not tuned properly (Carbonneau et al., 2008). Overfitting refers to the scenario in which a model has attained such a deep understanding of the patterns within the training data that it struggles to apply that knowledge to unseen data (Section 2.5.1). To avoid overfitting, it is important to use techniques such as cross-validation, which is a useful technique to assess how the model generalizes to the given data set (Chuang et al., 2021).

It can be difficult to pre-establish the amount of data needed for a predictive task. The amount of data required for machine learning tasks depends on factors such as the complexity of the problem and the complexity of the algorithm (Cerqueira et al., 2019). In general, a larger sample size tends to be advantageous, as it offers a more comprehensive and diverse range of patterns and relationships for the machine learning model to learn from. A larger dataset can help mitigate the risk of overfitting. However, the relationship between sample size and prediction accuracy is not always linear. The objective is to strike a balance between capturing crucial information for model training and ensuring computational feasibility. This is achieved by selecting a relatively small, yet statistically significant sample size for the training set, typically a few thousand observations (Vercellis, 2011).

Thus, the choice of forecasting method is strongly determined by the sample size and the characteristics of the time series. The characteristics of a time series can be defined as whether the time series is stationary, skewed in distribution (Syntetos, Boylan et al., 2005; Carbonneau et al., 2008; Kourentzes et al., 2020), the degree of autocorrelation (X. Zhang, 2007), and the decomposed components (trend, seasonality, and residuals) (N. Li et al., 2021). However, it is important to note that even with a large amount of data, there is no guarantee that a machine learning method will outperform a statistical method in terms of forecasting accuracy (Cerqueira et al., 2019). This argument aligns with the *No Free Lunch* Theorem (Wolpert, 1996), which asserts that there is no "one-size-fits-all" learning algorithm that is universally optimal for every situation. Consequently, a careful evaluation of different methods is essential to identify the most effective approach.

6.2 Material Categorization Methods

During the case study (Section 5), it was discovered that LC HMN maintains an inventory of several thousand materials. As revealed in the systematic literature review (Section 4), it is not practical to conduct in-depth research to fit the demand pattern for each individual material when dealing with a large number of materials (Wenhan Fu, 2018). This would result in significant computational costs, which is economically inefficient. Therefore, in Section 5.2.3, various methods and variables were explored to classify the SKUs of LC HMN.

6.2.1 An Algorithmic Approach

An algorithmic approach was used to evaluate the demand for a given sample of materials (as seen in Figure 33) using KMeans classification (Section 5.2.3.1. The analysis incorporated the utilization of two key variables: the coefficient of variation (CV), calculated as the ratio of the standard deviation of demand to the average demand, and the inventory turnover ratio. The application of KMeans clustering resulted in the identification of three distinct clusters, as illustrated in Figure 35. The clusters revealed that the inventory of LC HMN encompasses a wide range of demand patterns for different materials. Some materials exhibit a high frequency of orders, indicating a consistently high demand, while others demonstrate infrequent orders with low demand throughout the analyzed year.

Notably, the findings from the clustering analysis highlight that more than 50% of the items in the inventory have an inventory turnover ratio below seven. This ratio signifies that the inventory is sold out seven times or fewer within the analyzed year, indicating a relatively slow-moving inventory for a significant portion of the items. This finding suggests that there is potential for more frequent replenishment by Logistics Center Helse Midt-Norge, assuming the lead time for the materials allows such an approach. This observation highlights the possibility of optimizing inventory management by adopting a more responsive and proactive replenishment policy.

Furthermore, the clustering of materials revealed considerable demand variation among product groups, particularly for Cluster 2 (Figure 35). The demand pattern exhibited by materials in Cluster 2 indicated sparse demand, with some years experiencing no orders at all. The sparsity observed can be attributed to the presence of substitute materials. This observation aligns with the earlier case study of Logistics Center Helse Midt-Norge (Section 5.1), which identified a significant number of substitute materials in the inventory due to logistical and production issues with suppliers in recent years. Consequently, these items are unlikely to be part of the regular material assortment.

On the other hand, Cluster 1 represented a group with a relatively higher inventory turnover ratio and low variation in demand. However, it is important to note that this group exhibits a higher turnover ratio. Thus, it is not necessarily given that this group has significant potential to increase its already high inventory turnover ratio. The relatively stable and predictable demand pattern suggests that the inventory turnover ratio for this group may have reached its maximum potential, and additional attempts at improvement may have limited returns.

In general, the algorithmic approach employed in this study provided valuable insights into the demand characteristics of materials within Logistics Center Helse Midt-Norge. The clustering analysis highlighted the potential for more frequent replenishment for a substantial portion of the inventory and revealed variations in demand patterns and the presence of substitute materials. These findings can guide decision-making in enhanced inventory management at Logistics Center Helse Midt-Norge.

6.2.2 A More Nuanced Categorization

Besides the algorithmic approach used for material categorization, an additional method was tested that accounted for the temporal variation in demand occurrences. During the systematic literature review conducted in Section 4, an insightful method developed by Syntetos, Boylan et al. (2005) was discovered. In addition to capturing the variation in demand, the average demand interval (ADI) was also measured and a classification matrix was plotted utilizing these two variables (Figure 39). The findings revealed that the inventory of LC HMN encompassed three of the four demand categories (*lumpy, erratic, and smooth*) described by Syntetos, Boylan et al. (2005), indicating the efficacy of the method in capturing variations in demand magnitude and the time intervals between occurrences of demand.

This finding aligns with the research of Dolgui, Louly et al. (2005), who emphasized the existence of numerous untapped opportunities in inventory control. They highlighted that by considering both demand and time variation in a co-evaluated manner, practical solutions can be achieved for complex inventory management. Moreover, X. Zhang (2007)'s study further supports the utilization of such an approach, demonstrating that failure to recognize the volatility in demand variability can lead to deviations from the desired service level when adopting a service-level approach in inventory management. Therefore, accurately measuring and understanding the extent of demand variation becomes crucial, particularly given the substantial number of materials with high demand variation present in the inventory of LC HMN.

It is important to note that LC HMN operates under the constraint of maintaining a minimum service level of 97%. This underscores the importance of precisely assessing and managing demand variability to maintain the desired service level, while also enhancing inventory control by enabling targeted replenishment policies for the different demand categories. The feasibility of this demand classification method will be discussed later in the context of simulation. Categorizing all materials makes it easier to apply predictive algorithms to each specific material.

To analyze the distribution of average lead time within each demand category, a distribution graph was generated, providing a visual representation of the materials in each category. As depicted in Figure 37, the lead time exhibits a relatively even distribution across all three types of demand: *erratic*, *lumpy*, and *smooth*. This suggests that lead time does not significantly impact the variation in demand or the intervals between demand occurrences. As mentioned earlier in this section, more than 50% of the items from LC HMN have an inventory turnover ratio lower than seven. Keeping this in mind, it should be noted that the typical lead time for materials (as observed in Figure 37) falls within the range of two to eight days. The relatively "low" average lead time implies the potential for a higher inventory turnover ratio, and consequently the possibility of more frequent replenishment. However, the reason for a low inventory turnover compared to the average lead time could be attributed to the fact that the demand for the majority of items that Logistics Center Helse Midt-Norge holds exhibits significant variability in demand.

It is important to note that the method developed by Syntetos, Boylan et al. (2005) overlooks crucial factors such as product price, lead time variations for different materials, and information about suppliers. Unfortunately, the supplier data for all the materials were categorized as sensitive information and were unavailable to the authors of this study. It would be beneficial to further categorize the materials based on suppliers to explore potential benefits, thus highlighting a limitation of the dataset and the employed classification method. Additionally, the method does not consider any temporal dependencies, such as trend and seasonality, when applying the classification. Regarding material types, LC HMN holds a certain amount of substitute goods (as discussed in Section 5), which are intended to replace out-of-stock materials from the preferred supplier. However, these materials have not been mapped or categorized as the same material.

6.3 The Python-Based Simulation Model

The motivation behind the use of a simulation model was to compare and evaluate different replenishment policies. The simulation model was implemented in Python with the SimPy library (Section 5.3.1). It captures the dynamic nature of inventory levels, allowing analysis of Key Performance Indicator (KPI)s such as average inventory and service level. Using simulations, different scenarios and conditions can be assessed, adjusting parameters to explore inventory management strategies and their impact. The model provides a comprehensive view of the inventory management process, tracking events like order placement, arrival, and fulfillment. This data enables making evidence-based decisions, comparing different approaches, and quantitatively evaluating their effectiveness in improving average inventory levels while maintaining high service levels.

Scalability is one of the advantages of utilizing a simulation model implemented in Python with the SimPy library. This means that the simulation can be easily applied to a wide range of materials and various parameters, making it possible to simulate different scenarios and conditions. In addition, the scalability of the simulation model facilitates the analysis of large datasets and the generation of reliable statistical results. It can handle a significant number of simulation runs, increasing the robustness and accuracy of the evaluation. With a scalable simulation, it becomes feasible to generate sufficient data points and make evidence-based decisions regarding inventory management.

Furthermore, the simulation model can incorporate various forecast methods, enhancing its applicability. Forecasting plays a crucial role in inventory management (Bacchetti and Saccani, 2012), as accurate predictions of demand can significantly impact inventory levels and service levels. By allowing the flexibility to choose any forecasting method for the simulation, it becomes possible to explore its effectiveness and compare its impact on the calculation of reorder points and the resulting average inventory levels.

As revealed in the case study (Section 5), LC HMN follows a fixed order size approach, although the order size can be modified occasionally. These modifications are made based on factors such as the purpose of filling the pallets or making minor adjustments if necessary (Section 5.1.4). However, in the simulation model, the order size remains fixed and no changes can be made to the fixed order size during the simulation.

With a fixed order size, it becomes challenging to respond quickly to changes in customer demand, supplier lead times, or unforeseen disruptions in the supply chain. This lack of responsiveness can lead to missed sales opportunities or excess inventory that become outdated.

Another limitation is the assumption of a fixed lead time for orders. In reality, lead times can vary due to factors such as supplier performance and shipping delays. This variability is not taken into account in the simulation model. Furthermore, the simulation model assumes a single-level supply chain, which does not accurately represent the complexities of real-world supply chains. In practice, supply chains can involve multiple levels of suppliers and distributors, each with its own lead times and inventory policies. Additionally, the model structure is relatively simple and does not incorporate nuances of inventory management such as inventory obsolescence and expiration or product substitution.

It is also important to note that the simulation model does not handle unfulfilled demand. If the demand exceeds the inventory level in a given time slice, the unfulfilled demand will be tracked, but the demand will not be met. This limitation can affect the accuracy of the results, as unfulfilled demand can have significant implications for inventory management decisions.

A further limitation of the presented simulation model is the utilization of a custom initial stock formula, where the initial stock is calculated as the sum of the reorder point at the beginning of the simulation period and half of the fixed order size (Section 5.3.1.1). A drawback of this approach is that the simulation does not account for the actual initial stock levels of each simulated material at the beginning of the period. Consequently, the model fails to capture the potential impact of varying initial stock levels on subsequent inventory management strategies. However, this simplification offers an advantage by facilitating the simulation and analysis of different inventory management strategies for each material, thereby enhancing the ease of experimentation and evaluation.

Despite these limitations, the simulation model implemented in Python with the SimPy library provides a valuable tool to compare and evaluate different replenishment policies in inventory management. It captures the dynamic nature of inventory levels and allows the analysis of key performance indicators such as average inventory and service level.

6.4 Improving the AS-IS Fixed Reorder Point

This section examines the effectiveness of transitioning from the existing AS-IS Fixed ROP strategy to two alternative strategies: the Proposed Fixed ROP and the Basic Dynamic (OPTIMAL) ROP. The results presented in Section 5.3.3 shed light on the impact of these systems on inventory management and service levels.

The AS-IS Situation

The simulation results, based on the AS-IS data of the case company, provide valuable insights into the inventory management practices for the 1256 sampled materials. The AS-IS fixed ROP replenishment strategy for three of the sampled materials, one from each demand category (Figure 39), is visualized in Section 5.3.2.1. After simulating the AS-IS situation for the three materials, several key observations can be made. First, inventory levels exhibited a consistent pattern of being significantly higher than the corresponding demand, as indicated by the blue line in Figure 48, compared to the orange demand line. This finding challenges the author's initial expectation that the inventory levels would closely align with the demand patterns. Furthermore, the cumulative stock and demand levels depicted in Figure 49 reveal a substantial green area, representing the integral of the stock level surplus. In a perfect scenario, this green area would be minimal, indicating a close match between stock levels and demand. However, the considerable surplus of inventory reflected in Figure 49 suggests a discrepancy in the demand and inbound ordering of goods.

This interpretation aligns with previous research on inventory management, which emphasizes the importance of aligning inventory levels with actual demand to optimize operational efficiency (Gao et al., 2022). Further analysis of various strategies can have a significant impact on existing inefficiencies, as the simulation outcomes expose weaknesses in the current inventory management approach.

In terms of service level for the AS-IS situation, the resulting service levels for the various demand categories vary from 99.1% to 99.5% (Figure 50). This shows that in the simulated warehouse, all categories of materials exhibit a service level above the constraint of 97%.

The Proposed Fixed ROP & The Basic Dynamic (OPTIMAL) ROP

By analyzing the simulation results presented in Table 12, one can observe the influence of two strategy transitions on the average inventory and service level (SL). The first strategy transition is from the AS-IS Fixed ROP strategy to the Proposed Fixed ROP strategy, while the second transition is from the AS-IS Fixed ROP to Basic Dynamic (OPTIMAL) ROP. The change in average inventory and SL are presented for the different demand categories: *Erratic*, *Lumpy*, and *Smooth*, as well as all materials in total, expressed as *Total*.

Taking into account the service level constraint of 97% for the case company, the simulation results can be analyzed accordingly. For the transition to the Proposed Fixed ROP, the results indicate a decrease in the average inventory ranging from 27.7% to 64.4% across the demand categories. However, the SL decreases by 1.0% to 2.5% compared to the AS-IS system. With the decline in SL, the resulting SL values range from 96.6% to 98.5%, which indicates a drop below the company's 97% SL constraint for a significant amount of materials. According to the findings presented in Table 14, a total of 229 materials are observed to have transitioned from having an initial service level above 97% to falling below the service level constraint of 97%. These negative movements primarily affect the *erratic* and *lumpy* materials, with 83 and 142 materials, respectively, moving below the service level constraint. However, when considering the *smooth* materials, only 4 materials have moved under the service level constraint, which can be considered a negligible number in this context. This observation is significant due to the higher variability in demand for *erratic* and *lumpy* materials, as stated in Section 5.2.3.2. Consequently, it is also logical to expect that these materials will exhibit greater sensitivity with the change in fixed ROP value. In total, the Proposed Fixed ROP shows potential to reduce average inventory levels, but does not maintain an acceptable service level.

In contrast, the transition to Basic Dynamic (OPTIMAL) ROP demonstrates more substantial improvements. The results of Basic Dynamic (OPTIMAL) ROP calculated using the SMA forecasting model for demand forecasting indicate significant improvements in average inventory levels without a significant drop in service level performance. Basic Dynamic (OPTIMAL) ROP utilizes the SMA model with the optimal combination of window size and forecast interval parameters. As stated in Section 5.3.3, the optimization process involved a brute-force search of 50 unique parameter combinations for each of the 1256 materials analyzed.

The findings (as seen in Table 12) show that in different demand categories, implementing the Basic Dynamic (OPTIMAL) ROP strategy leads to substantial reductions in average inventory levels. Specifically, the average inventory reduction ranges from 36.5% to 63.3% across the demand categories, indicating a better inventory management efficiency. The percentage point change in SL is smaller compared to the transition to the Proposed Fixed ROP strategy, ranging from 0.1% to 0.8%. Consequently, the resulting SL values for the Basic Dynamic (OPTIMAL) ROP range from 98.7% to 99.0% for the demand categories, with an overall SL of 98.9%. This is supported by the results in Table 14, stating that only two SKUs have moved from having an AS-IS service level above 97% to a service level less than the service level constraint of 97%. These results suggest that the Basic Dynamic (OPTIMAL) ROP achieves superior inventory reduction while still meeting the 97% service level constraint. Not only is the service level constraint met, but as stated in Table 13, 41 materials have moved from having an AS-IS service level of less than 97%, to a service level higher than the LC HMN service level constraint. These materials represent the demand categories *erratic* and *lumpy*, and as mentioned earlier in the discussion, these categories are sensitive to changes in the calculated ROP.

When transitioning from Proposed Fixed ROP to Basic Dynamic (OPTIMAL) ROP (Table 13), a significant advantage of the Basic Dynamic (OPTIMAL) ROP strategy over the Proposed Fixed ROP strategy is indicated by the fact that 90 *erratic* materials and 166 *lumpy* materials move from below to above the service level constraint of 97%. Only four *smooth* materials cross the boundary from below to above 97% when transitioning from Proposed Fixed ROP to Basic Dynamic (OPTIMAL) ROP. Taking into account the minimal difference in average inventory reduction between the AS-IS Fixed ROP strategy and both the Proposed Fixed ROP and Basic Dynamic (OPTIMAL) ROP strategies (Table 12), it is evident that there are negligible differences between the two suggested policies for *smooth* materials. Thus, there is strong evidence to suggest that it is not necessary, from a cost and time perspective, to implement the Basic Dynamic (OPTIMAL) ROP for materials within the *smooth* demand category.

In summary, the simulation results provide valuable insight into the impact of transitioning to different inventory management strategies. Both the Proposed Fixed ROP and the Basic Dynamic (OPTIMAL) ROP show considerable reductions in average inventory compared to the AS-IS Fixed ROP. The Proposed Fixed ROP exhibits a considerable decrease in service level. The Basic Dynamic (OPTIMAL) ROP, on the other hand, showcases substantial inventory reduction with minimal impact on the service level. Taking into account the service level constraint and the need for inventory optimization derived in the discussion of the AS-IS situation (Section 6.4), the case company should carefully evaluate the benefits and trade-offs associated with implementing either of the proposed inventory management policies.

Analysis of Optimal Combinations of Forecast Intervals and Window Sizes for Different Material Types

The distribution of optimal combinations of the forecast interval and window size for the 1256 materials analyzed is presented in Figure 53. This analysis provides valuable information on the patterns observed among different categories of materials.

The distribution of optimal combinations for *erratic* materials exhibits a significant variation. The preferred forecast interval for these materials is predominantly business days (B) and one week (1W). Interestingly, these materials tend to require larger window sizes, with 10 being the most common. This supports the description of *erratic* materials derived in Section 5.2.3.2, that *erratic* materials are characterized by high variation in quantity, but consistent in time intervals between occurrences of demand, requiring relatively larger window sizes for accurate forecasting.

In contrast, the *lumpy* materials demonstrate a more dispersed distribution. The forecast intervals of "B" and "1W" exhibit a more even distribution across various window sizes, including the smaller ones, despite the dominance of window size 10. However, for forecast intervals of "2W", "3W" and "4W", the larger window sizes appear to be more dominant. This indicates that there is no clear determination of the best forecast interval for this category and that *lumpy* materials require different forecast intervals and window sizes to effectively capture the underlying fluctuations. This aligns with the description of *lumpy* materials in Section 5.2.3.2, where they are described as difficult to predict.

Smooth materials, on the other hand, predominantly favor a forecast interval of "B". Among the available window sizes, 6 to 10 stand out as the most dominant choice. Among these, window size 10 appears to be the most frequently selected. This suggests that *smooth* materials exhibit more stable and predictable behaviors, where shorter forecast intervals and relatively high window sizes are sufficient for accurate forecasting.

The correlation between material types and optimal combinations suggests that certain patterns and relationships exist. For example, *erratic* materials tend to require shorter forecast intervals and larger window sizes, possibly due to their unpredictable nature. On the other hand, *lumpy* materials have a more diverse distribution, indicating the need for flexibility in choosing forecast intervals and window sizes based on the specific characteristics of the demand pattern. This highlights the importance of considering material-specific factors when selecting forecast intervals and window sizes.

Furthermore, the findings contribute to the existing knowledge by showcasing the relationships and dependencies between the forecast interval, the window size, and the characteristics of the materials. It becomes evident that for the *lumpy* materials, there is no one-size-fits-all approach to forecasting. Instead, a nuanced understanding of demand patterns and characteristics is necessary. On the other hand, for *erratic* or *smooth* materials, the findings suggest that a uniform forecast interval may be applicable, as indicated by the more concentrated distributions around shorter forecast intervals.

Examining the Potential Savings in Holding Cost of The Analyzed ROP Strategies

The comparative analysis of holding costs for different replenishment strategies, as seen in Section 5.3.3, reveals significant cost-saving potential by implementing the Proposed Fixed ROP or Basic Dynamic (OPTIMAL) ROP strategy. The results show that the AS-IS fixed ROP strategy incurs the highest holding costs, while the Proposed Fixed ROP strategy and the dynamic ROP (OPTIMAL) strategy result in a holding cost reduction of 53.8% and 56.3% respectively.

The results in Figure 56 and Table 15 must be considered in the context of the resulting service levels (Figure 54). These results are combined in Table 16. The findings show that when considering individual demand categories separately, distinct variations arise in terms of which replenishment strategies yield the lowest holding costs.

Starting with the *erratic* category, the Proposed Fixed ROP strategy yields the lowest holding costs of 1 310 617 NOK, with a 56.9% reduction compared to the AS-IS Fixed ROP strategy. However, when using the Proposed Fixed ROP strategy for *erratic* materials, the service level would be 0.8% lower compared to the Dynamic ROP(OPTIMAL) strategy.

For *lumpy* materials, the Basic Dynamic (OPTIMAL) ROP strategy performs significantly better, with a 47% lower holding cost compared to the AS-IS Fixed ROP strategy. The resulting service level for *lumpy* materials using the Basic Dynamic (OPTIMAL) ROP strategy would be 99% as opposed to 96.6% with the Proposed Fixed ROP strategy, which falls below the 97% constraint. Thus, the Proposed Fixed ROP does not fulfill the requirements of LC HMN for *lumpy* materials.

For *smooth* materials, the Proposed Fixed ROP strategy demonstrates better performance compared to the Basic Dynamic (OPTIMAL) ROP approach, resulting in a 74.1% lower holding cost compared to the AS-IS Fixed ROP strategy. The service level for Proposed Fixed ROP would be 98.5% compared to 98.7% for Basic Dynamic (OPTIMAL) ROP. These findings support what was discussed in Section 6.6.3, that for *smooth* materials, the Dynamic Reorder Point (ROP) values exhibit a close to straight line similar to a fixed ROP strategy. This implies that when weighing complexity against performance, the Proposed Fixed ROP strategy would be the preferred replenishment strategy for *smooth* materials, as the Basic Dynamic (OPTIMAL) ROP requires additional computation time.

Lastly, when examining the holding cost for the three demand categories in Table 19 and combining it with the distribution of average demand buckets in Figure 51, some interesting insights can be derived. It should be noted that most materials in the *erratic* and *lumpy* demand categories are concentrated within the 0-500 bucket, which also happens to be the bucket with the highest holding cost for both categories, regardless of the forecast interval (Table 19). On the contrary, the majority of materials in the *smooth* demand category are found within the average inventory level bucket of >3000 (Figure 51). This observation suggests that the demand categories with the most unpredictable demand patterns, namely *lumpy* and *erratic*, are associated with the most expensive materials held by LC HMN.

Thus, the implications of the analysis suggest that implementing a dynamic operational strategy can contribute to improved cost efficiency and profitability for the company. However, a dynamic strategy may not be the best performing for all materials. The demand categorization of the materials reveals that a dynamic strategy may not necessarily be the optimal choice for all three represented categories of demand, and the preferred strategy depends on the specific category being considered. When considering the demand patterns of inventory items and adjusting replenishment strategies accordingly, holding costs can be significantly reduced. This optimization can lead to better resource allocation and improve overall supply chain performance.

6.5 Exploring the Impact of Forecast Intervals on Basic Dynamic ROP Strategy

In Section 6.4, the analysis of the Basic Dynamic (OPTIMAL) ROP strategy was discussed. This strategy was optimized at a material level, as explained in Section 5.3.3. The analysis focused on each material's utilization of the best combination of both forecast interval and window size for the Simple Moving Average (SMA) model used for demand forecasting. This section examines the extent to which the optimal solution can be relaxed, in contrast to the analysis conducted on the optimal solution. To accomplish this, the simulation of materials was carried out using each of the five forecast intervals sequentially, while optimizing only the window size for each specific material.

The results presented in Section 5.3.5 were obtained by simulating all distinct combinations of the 1256 materials. The analysis utilizes the simple moving average (SMA) demand forecasting technique, considering a range of window sizes from 1 to 10. The forecast intervals examined include business days (B), one week (1W), two weeks (2W), three weeks (3W), and four weeks (4W). Simulations were conducted to gain valuable insight into the trade-off between service levels and average inventory levels, particularly in light of forecast intervals.

The examination of the service level results reveals a consistent improvement as forecast intervals increase. In the case of the "Total" category in Figure 60, service levels for the forecast intervals B, 1W, 2W, 3W and 4W are observed to be 98.0%, 98.5%, 99.1%, 99.3%, and 99.4% respectively. These findings demonstrate that longer forecast intervals contribute to higher service levels. It is important to note that all forecast intervals exceed the service level constraint of 97%, set by LC HMN.

Upon analyzing the average inventory results, it is evident that the implementation of the Dynamic ROP strategy has generally led to reductions in average inventory levels across various forecast intervals and demand categories. As shown in Figure 61, the average inventory levels for forecast intervals B, 1W, 2W, 3W, and 4W in the "Total" category have been reduced by 37.8%, 36.6%, 33.5%, 29.7%, and 26.5% respectively. These reductions signify improved efficiency and cost savings for the case company. This finding highlights the negative correlation (Section 2.3.2) between service levels and average inventory reduction, which can be seen in Figure 60 and Figure 61. Although longer forecast intervals enhance service levels, they may also lead to less of a reduction in average inventory levels, potentially incurring higher holding costs for the company.

As presented in Figure 60, the relationship between service level and forecast interval is examined. As discussed earlier, there is a positive correlation between the two, indicating that as the forecast interval increases, the service level also increases. The plot illustrates this relationship, showing that as the forecast interval extends, the service level tends to rise.

Upon analyzing the various demand categories in light of the forecast intervals, it becomes evident that the *lumpy* category shows the sharpest rise in service level, advancing from 97.6% to 99.3%. Subsequently, the *erratic* category demonstrates the second-highest increase, while the *smooth* category exhibits a more gradual incline. These observations align with the classification of demand (Section 5.2.3.2), which identified different demand categories based on variations in demand patterns and events (Figure 36). The *lumpy* category encompasses materials with the greatest fluctuations in demand and occurrences of events, which explains its significant improvement in service level. The steep increase in service level for the *lumpy* demand category is likely associated with increased uncertainty related to longer forecast intervals, resulting in higher simulated reorder points.

Comparing the different forecasting strategies, it is observed that the "OPTIMAL" solution achieves a service level of 98.9%, outperforming strategies "B" (98.0%) and "1W" (98.5%). However, the "OPTIMAL" solution falls behind the performance of strategies "2W," "3W," and "4W". It is important to consider that the discrepancy between the service levels of the different forecast intervals is minimal, with the worst-case scenario being a 0.9% difference compared to the more complex "OPTIMAL" solution. From a statistical point of view, this difference is practically negligible.

The findings presented in Figure 61 highlight the differences between the OPTIMAL forecast interval and the other intervals for various categories of material. In particular, the *smooth* materials category exhibits the least difference in the average inventory reduction between OPTIMAL and the other forecast intervals, with a 10.1% difference between the 4W and the OPTIMAL. This suggests that the choice of forecast interval has a relatively minor impact on inventory reduction for *smooth* materials.

However, for both the *smooth* and *erratic* material categories, the "B" and "1W" forecast intervals perform nearly as well as the OPTIMAL interval. This indicates that, in terms of reducing average inventory, "B" and "1W" can be considered viable alternatives to OPTIMAL for these categories. On the other hand, when it comes to the *lumpy* material category, OPTIMAL outperforms all forecast intervals by a substantial margin. The average inventory reduction achieved by OPTIMAL is 36.5%, compared to the range of 28.2% to 19.4% for the forecast intervals. This result suggests that the *lumpy* category is particularly sensitive to the choice of forecast interval and that OPTIMAL proves to be significantly more effective in this case.

In general, OPTIMAL shows the highest reduction in average inventory among all forecast intervals, achieving a remarkable -42.0%. However, it is worth noting that the lower forecast intervals, such as "B" and "1W," are not far behind in terms of inventory reduction. These intervals achieve reductions of -37.8% and -36.6%, respectively. Although OPTIMAL provides the most significant improvement, the relatively close performance of "B" and "1W" suggests that they can be considered reasonable alternatives, especially considering their comparable inventory reduction levels.

Table 17 presents the results of the forecast intervals on the movement of materials, focusing on the number of materials that transition from below the service level constraint to above it, which is set at 97%. Upon examining the data in Table 17, it is evident that the forecast intervals have an impact on the movement of materials. The range of materials moved varies between 25 and 36 when considering different forecast intervals. It is important to note that the optimal forecast interval results in the movement of 41 materials, which is the highest number among all alternatives. This suggests that the optimal forecast interval has the best potential to satisfy the service level constraint.

When analyzing the specific material categories, it is observed that the *smooth* materials do not exhibit any movement across the service level constraint for any of the forecast interval options. In contrast, *erratic* materials demonstrate a relatively smaller range of materials moved, ranging from 4 to 9. This implies that forecast intervals have a limited impact on the movement of *erratic* materials, suggesting that they may exhibit more consistent demand patterns or have relatively stable inventory levels, making them less sensitive to changes in forecast intervals.

On the other hand, *lumpy* materials exhibit a wider range of materials moved, ranging from 21 to 32. This indicates that forecast intervals have a more significant influence on the movement of *lumpy* materials. These materials have more sporadic demand patterns, making them more sensitive to changes in forecast intervals. Therefore, adjusting the forecast intervals for *lumpy* materials could potentially lead to a more accurate forecast and improved inventory management.

Table 18 presents the results of the analysis on the movement of materials from over to under the service level constraint of 97% for different forecast intervals. The findings reveal a negative correlation between the forecast interval and the number of materials moved from over to under the service level constraint. As the forecast interval increases, indicating a longer-term forecast, the number of materials that are transferred decreases. Examining the subcategories of *erratic*, *lumpy*, and *smooth*, it is observed that the majority of materials moved from over to under the service level constraint are classified as *lumpy*, regardless of the forecast interval. However, as the forecast interval lengthens, the number of *lumpy* materials transitioning from over the service level constraint to under decreases, suggesting that a longer-term forecast provides better insights into managing such *lumpy* materials. In contrast, the *erratic* and *smooth* materials show minimal movement from over to under the service level constraint, regardless of the forecast interval. Conversely, *smooth* and *erratic* materials, are less impacted by the specific forecast intervals.

The findings contribute to a clearer understanding of the relationship between forecast intervals, service levels, and average inventory in the dynamic ROP strategy. The results of the study provide empirical evidence that longer forecast intervals positively influence service levels within the dynamic ROP strategy. However, longer forecast intervals may result in less reduction in average inventory levels, potentially leading to higher holding costs. The section also highlighted the significance of demand categories, with the *lumpy* category showing the highest increase in service level, followed by the *erratic* category. The optimal solution, named "OPTIMAL," outperformed shorter forecast intervals in terms of service level and average inventory reduction, but the differences were relatively minor. For *smooth* and *erratic* materials, forecast intervals "B" and "1W" were viable alternatives to OPTIMAL, while OPTIMAL performed significantly better for *lumpy* materials. The analysis of material movement, which involved investigating the number of materials that move below or above the service level limitation, revealed that longer forecast intervals have a more noticeable effect on *lumpy* materials. In contrast, *erratic* materials showed limited sensitivity to changes in forecast intervals. Furthermore, the findings suggested a negative correlation between forecast interval length and the number of materials moving from over to under the service level constraint, with longer-term forecasts providing better insights for managing *lumpy* materials.

The Relationship Between Forecast Intervals and Window Sizes for Demand Categories

The results visualized in Figure 59 show the distribution of window sizes for analyzed forecast intervals across different demand categories. The key findings reveal a positive correlation between the forecast interval and window size for all three demand categories. It was observed that as the forecast interval increased, the distribution of window sizes also increased. Among the demand categories, *smooth* materials exhibited the most rapid increase in window size, followed by *erratic* materials with a relatively fast ascent. Lumpy materials displayed a significantly slower change in the distribution compared to *erratic* and *smooth* materials, but there was still a positive correlation with an increased forecast interval. An interesting observation was that for a forecast interval of 4 weeks, there was a shift in the correlation, where smaller window sizes became more frequent. This trend was consistent across all material categories.

The results indicate that as the forecast interval lengthens, there is a tendency for larger window sizes to be associated with the analyzed forecast intervals. This finding suggests that a longer forecast interval requires a broader window of historical data to accurately predict demand patterns.

The transition of window size distributions towards higher values occurs most rapidly for *smooth* materials, somewhat slower for *erratic* materials, and slowest for *lumpy* materials. There is a correlation between the speed of transition of window size distributions towards higher values and the complexity of the demand category. As derived in Section 5.2.3, *smooth* materials are the easiest to predict, *erratic* materials are more challenging, and *lumpy* materials are the most difficult to predict. This finding implies that the relationship between the optimal window size and forecast interval is affected by the complexity level of the demand pattern.

Additionally, the shift in correlation observed at a 4-week forecast interval, where smaller window sizes become more frequent, suggests that the forecast interval is sufficiently large for nearby historical data to have a greater impact. One possible explanation for this phenomenon is that longer forecast periods introduce higher levels of uncertainty, which justifies excluding data from the distant past. By focusing on more recent data within a narrower time frame (utilizing smaller window sizes), the forecast can be more applicable and reliable. According to this observation, when employing forecast intervals exceeding three weeks, the importance of recent data increases.

The presented analysis on the distribution of window sizes for forecast intervals has certain limitations that should be taken into consideration. One important limitation is the given search space of window sizes and forecast intervals stated in Section 5.3.3. Since the analysis did not consider forecast intervals beyond 4 weeks, it is important to note that the shift in correlation observed at the 4-week forecast interval, where smaller window sizes became more frequent, may not hold true for even longer forecast intervals.

Holding Cost - Forecast Interval

In order to examine the effect on holding cost when relaxing the optimal parameters of window size and forecast interval for the Basic Dynamic ROP, the simulation results from Section 5.3.5 were analyzed. The results displayed in Figure 60 demonstrate that all forecast intervals produce a service level above the defined 97% constraint for the *erratic* materials. The associated holding cost, however, presented in Table 19, indicates that for *erratic* materials, the Basic Dynamic ROP strategy with a daily "B" or one-week ("1W") forecast intervals performs closest to the OPTIMAL solution.

The Basic Dynamic ROP utilizing daily forecasts ("B") results in a holding cost of 1 324 185 NOK, which is slightly better than the holding cost of 1 364 921 for the Basic Dynamic (OPTIMAL) ROP strategy. The Basic Dynamic (OPTIMAL) ROP strategy yields a slightly higher service level, with an improvement of 0.2 percentage points.

In contrast, the resulting service level for the Basic Dynamic ROP (1W) is 98.9%, which is 0.2 percentage points higher than the Basic Dynamic (OPTIMAL) ROP. However, this improvement comes at a cost. The holding cost for the Basic Dynamic ROP with a "1W" forecast interval amounts to 1 377 541 NOK, which is slightly higher (0.9%) compared to the Basic Dynamic (OPTIMAL) ROP strategy. Thus utilizing either daily ("B") or weekly ("1W)" forecast intervals are preferred for *erratic* materials.

For the *lumpy* category in Table 19, the Dynamic ROP utilizing the "1W" forecast interval is the one performing closest to the one with "OPTIMAL" parameters, resulting in a holding cost of 901 140 NOK versus the holding cost of 793 142 NOK for the Dynamic ROP (OPTIMAL) strategy. In contrast to the *erratic* materials, the service level for *lumpy* materials decreases from 99% to 98% when utilizing a one-week ("1W") forecast interval instead of the "OPTIMAL" parameters. Therefore, considering both holding cost and service level, the best-performing solution for the *lumpy* materials is the Basic Dynamic (OPTIMAL) ROP strategy.

In contrast to the *erratic* and *lumpy* materials, the Dynamic ROP strategy using days ("B") as the forecast interval performs best in terms of holding cost for *smooth* materials. The Dynamic ROP strategy with the "B" forecast interval yields a 0.04% lower holding cost compared to the Dynamic ROP (Optimal) strategy. Analyzing the resulting service levels (Figure 60), the "Optimal" parameters and the daily forecast interval ("B") result in an equal service level of 98.7%. Surprisingly, the service level increases to 99.1% when the weekly forecast interval ("1W") is used.

6.6 The Effectiveness of Advanced Forecasting Models for The Dynamic ROP

In this section, the findings from Section 5.3.6 are examined to assess the potential impact of advanced forecasting models on inventory performance. As mentioned earlier, this analysis was only conducted on three materials, each representing one of the demand categories *smooth*, *erratic*, and *lumpy*.

6.6.1 Data Aggregation Analysis: Balancing Null Values and The Number of Data Points

The implications of the analysis on data aggregation (Section 5.2.4.4) and the selection of one-week aggregation for advanced forecasting methods (Section 5.3.6.1) in the dynamic ROP strategy have important consequences for theory and practice. The aggregation analysis results support the notion that aggregating the time series data to a higher level, such as weekly aggregation, can effectively reduce the number of null values in sparse datasets. Reducing null values through aggregation improves the robustness of forecasting models and enhances their ability to capture demand patterns.

Furthermore, the analysis reveals that one-week aggregation strikes a balance between reducing null values and maintaining an adequate number of data points. Increasing the aggregation level higher than "1W" results in a drastic reduction in data points, while one-week aggregation "1W" yields a significant reduction of zero values compared to daily ("B") data. This finding supports the idea that one-week aggregation provides a suitable compromise, allowing for accurate demand analysis and prediction while preserving an acceptable sample size for pattern recognition. Aggregation is directly linked to the forecast interval, as the level of aggregation determines the shortest possible forecast interval.

6.6.2 Performance Assessment of Advanced Forecasting Models for Different Demand Pattern Categories

The demand forecasting analysis conducted in Section 5.3.6.1 offers insights into the effectiveness of various models when applied to three distinct demand categories: *erratic*, *lumpy*, and *smooth* materials. This research focuses on three specific materials, namely 4001095, 4012198, and 4003841, which have been classified as *erratic*, *lumpy*, and *smooth*, respectively (Figure 39). The forecasting time span encompassed 52 weeks, allowing for a comprehensive assessment of long-term trends and patterns. The forecasting performed was one-step-ahead forecasting, where one data point (aggregated week) was forecasted at the time, indicating weekly forecasts. Several models were tested during this analysis, each contributing to the understanding of demand forecasting dynamics and their implications for inventory management and planning.

For the *erratic* representative material (4001095), it is evident that predicting demand accurately is a challenging task due to the high variation in demand quantity and inconsistent intervals between occurrences of demand. Among the four models evaluated (Table 20), the Holt-Winters model performed worse than the Naive method, indicating that it is not suitable for this type of demand. On the other hand, both the SARIMAX and LSTM models showed promising results, with the LSTM model slightly outperforming SARIMAX. This finding suggests that deep learning approaches, such as LSTM, can effectively capture the complex patterns and dynamics present in *erratic* demand. This result aligns with the findings of Pacella and Papadia (2021), which experienced that LSTM surpassed the traditional linear forecasting methods in performance. However, it is important to note that the LSTM model faced the issue of the vanishing gradient problem (Section 2.5.2) due to limited data points, which may have affected its performance. Therefore, further investigation and experimentation with larger datasets could potentially improve the performance of deep learning models to forecast *erratic* demand.

Turning to the *lumpy* representative material (4012198), which is characterized by high variation in quantity and intervals between demand occurrences. Due to the uncertainties, the *lumpy* category can pose great challenges for time-series forecasting (Jiang et al., 2017). The results indicate that the Holt-Winters model performed the best among the selected models (Table 21). This finding aligns with previous research suggesting that exponential smoothing methods, like Holt-Winters, are well-suited for handling *lumpy* demand and can provide accurate forecasts (Kiefer et al., 2021). On the other hand, both the SARIMAX and LSTM models performed worse than the Naive method, likely due to the high representation of non-existent demand (zeros) in the time series. It is worth noting that Holt-Winters and SARIMAX models have an advantage in handling zero values compared to LSTM, which requires preprocessing techniques like imputation (Che et al., 2018). These findings emphasize the importance of considering the specific characteristics of demand patterns when selecting an appropriate forecasting model.

For the *smooth* representative material (4003841), which exhibits low variance in demand and consistent intervals between demand occurrences, the SARIMAX model demonstrated the best performance among the evaluated models (Table 22). This result suggests that SARIMAX's ability to incorporate seasonality, trends, and exogenous variables, such as week number and month, can improve the accuracy of demand forecasts for *smooth* materials. These results are similar to those reported by Tang and Ge (2021) and Moroff et al. (2021a), which experienced that additional variables included resulted in better forecasts for inventory management. Although the Holt-Winters model also performed well, it showed a slightly lower performance compared to SARIMAX. The LSTM model performed similarly to Holt-Winters, indicating that deep learning approaches may not provide significant advantages for forecasting smooth demand when compared to statistical models.

In summary, this research provides new insights into the performance of different forecasting models for the specific demand categories at LC HMN. The findings support existing literature that suggests the suitability of certain models for different demand patterns. The results highlight the strengths and weaknesses of each model and provide guidance for selecting the most appropriate method for LC HMN, based on the characteristics of the time series data.

From a practical perspective, the findings have implications for inventory management. Accurate demand forecasting is crucial for optimizing inventory levels, reducing costs, and meeting customer demands. By understanding the performance of different models for different demand categories, practitioners can make informed decisions regarding the selection and implementation of forecasting methods. For instance, for *erratic* demand materials, deep learning models like LSTM can be considered, while for *lumpy* demand materials, Holt-Winters or other exponential smoothing methods may be more appropriate. For *smooth* demand materials, incorporating seasonality and trends through models like SARIMAX can lead to improved forecasting accuracy.

6.6.3 Comparative Analysis of Forecasting Techniques for Dynamic ROP strategy on Three Demand Categories

Section 5.3.6.2 investigated the application of advanced forecasting techniques on three distinct materials, which corresponded to the three demand categories illustrated in Figure 36. The materials under examination include Erratic Representative Material (4001095), Lumpy Representative Material (4012198), and Smooth Representative Material (4003841). To evaluate the performance of these techniques, the simulation results were analyzed using three key metrics: the Mean Absolute Scaled Error (Mean Absolute Scaled Error (MASE)), Service Level (Service Level (SL)), and average inventory.

Erratic Demand

The simulation results in Table 23 demonstrate the effectiveness of different forecasting models for Erratic Representative Material (4001095). The Naive model serves as the baseline, with a MASE score of 1.00, a service level of 100%, and an average inventory of 3189. The advanced models, including Holt-Winters, SARIMAX, LSTM, and SMA (1W-10), were compared against the Naive model. Among these models, the LSTM model achieved the lowest MASE score of 0.68, indicating its superior accuracy in predicting demand. The LSTM model also achieved a service level of 100% and an average inventory of 2414 in the simulation. On the other hand, the SMA model had a higher MASE score of 1.02 but achieved the lowest simulated average inventory of 2223.

Several factors can explain the lower average inventory with the SMA (1W-10) method compared to the LSTM method, despite handling *erratic* demand. Erratic demand, as derived in Section 5.2.3.2, is characterized by a high variation in demand quantity and consistent time intervals between the appearance of demand. Thus, in general, *erratic* demand patterns make it challenging to accurately forecast future demand (Jiang et al., 2017). Firstly, it is important to consider the sparsity of the data. As depicted in Figure 43 and Table 8, the proportion of zero values for the *erratic* representative material 4001095 decreases from approximately 76% to 28% when aggregating from business days "B" to weekly level "1W". Furthermore, Figure 44 demonstrates a 40% reduction in data points after aggregation to the weekly level "1W". Although the use of weekly intervals for forecasting and reducing the number of zero values, material 4001095, also known as the *erratic* material, still exhibits a relatively sparse distribution with approximately 28% zero values.

The SMA model, as shown in Figure 67, might struggle to capture the complexities of such sparse demand patterns, which could lead to sub-optimal inventory management. On the other hand, the LSTM model (as seen in Figure 66), a type of recurrent neural network specifically designed to handle sequential data (Section 2.5.2), achieves a lower MASE score, indicating its superior accuracy in demand forecasting.

Examining the visualizations of the simulated inventory levels based on the LSTM predictions further supports its effectiveness. Figure 66 illustrates a more fluctuating reorder point (ROP), thus more dynamic, suggesting that the LSTM model adapts to the *erratic* demand patterns by adjusting inventory levels accordingly. This dynamic ROP enables the system to respond quickly to sudden changes in demand, ensuring a better balance between stock availability and cost. In contrast, the SMA shown in Figure 67 presents a less fluctuating ROP, indicating that the SMA (1W-10) model is closer to a fixed ROP for the simulated period.

The performance of the forecasting models should be seen in relation to the resulting average inventory level and service level. Although LSTM has the lowest MASE score and a service level of 100%, it has a higher resulting average inventory level compared to SMA(1W-10).

As referenced in Section 5.3.1.1, the dynamic reorder point (ROP) is based on the forecasted demand and the error between the last forecast and the last observed demand. If there is a forecast high demand, a substantial error in the previous demand forecast, or a combination of both, the reorder point will be set at a higher level. Thus, the LSTM demand forecast has a large error, a large forecasted demand, or both, leading to the periodically high ROPs. This can also be observed in Figure 66, where the calculated Reorder Point (ROP)s adjusts significantly higher ahead of high spikes in demand. These results show that the use of the model with the highest forecasting accuracy may not necessarily result in the best-performing inventory management in terms of reducing inventory levels.

Lumpy Demand

For *lumpy* Representative Material (4012198), as shown in Table 24, the advanced forecasting models considered were Holt-Winters, SARIMAX, LSTM, and SMA with a one-week forecast interval and a window size of 10. The Naive model was again used as the baseline, with a MASE score of 1.00, a SL of 100%, and an average inventory of 5997. In this case, Holt-Winters demonstrated the best performance among advanced models, with a MASE score of 0.77, a SL of 96.78%, and an average inventory of 5066. The SMA model achieved the lowest average inventory of 5014, despite having a higher MASE score of 0.95. This highlights the trade-off between forecast accuracy and inventory levels. While advanced models may outperform the Naive model in terms of forecast accuracy (MASE score), they may not necessarily result in better simulation outcomes in terms of service level and average inventory.

When comparing the average inventory levels between the SMA (1W-10) and Holt-Winters methods for *lumpy* representative material (4012198), it is important to note that the SMA model achieved a lower average inventory of 5014 units, while Holt-Winters resulted in an average inventory of 5066 units. This finding raises some considerations given the nature of the analyzed and predicted demand, which is *lumpy*. Lumpy demand implies that demand has a high variation in quantity and time intervals between occurrences of demand (Section 5.2.3.2), making it the most challenging category to predict. It is important to recognize that the Holt-Winters method achieved a service level (SL) of 96.78%, which is just below the service level constraint set by the case company at 97%. Although the difference is relatively small, it suggests that the Holt-Winters model may not fully meet the desired service level constraint set by LC HMN. This could be a concern for the company's operations, as a higher service level is usually preferred.

Furthermore, it is worth mentioning that the data used for analysis contained a significant amount of sparsity, including many NaN values or zero values. As visualized in Figure 43, the percentage of zero values was reduced from 91% to 60% when aggregating to a weekly level, and the number of data points was only reduced with 7% (Figure 44). Thus, a significant number of zero values remained in the dataset.

This sparsity may have presented challenges to advanced forecasting models, potentially impacting their ability to accurately capture the underlying demand patterns. In such cases, simpler models like the SMA with a one-week forecast interval and window size of 10 could have handled sparsity better, resulting in a more stable and smooth line of Reorder Point (Figure 69) compared to the fluctuating Reorder Point observed in the Holt-Winters simulation (Figure 68).

Despite SARIMAX, LSTM, and SMA(1W-10) yielding a MASE score close to 1 (Table 24), all three simulated algorithms achieved a service level of 100% and maintained a similar average inventory level close to 5000 units. It demonstrates that the three algorithms exhibited comparable performance for the *lumpy* material 4012198, with none showing superiority over the others.

Taking these factors into account, it becomes evident that the Holt-Winters alternative has certain limitations when applied to *lumpy* demand scenarios. The resulting suboptimal service level, its sensitivity to sparsity in the data, and the fluctuating nature of the ROP all contribute to its less favorable performance compared to the SMA method. Although the Holt-Winters method excels in terms of forecast accuracy (MASE score), it does not yield better simulation outcomes in terms of service level, inventory levels, and thus cost-efficiency. The SMA method could be a more suitable choice for managing *lumpy* demand, providing a smoother inventory replenishment process and potentially mitigating the challenges posed by irregular demand patterns.

Smooth Demand

The performance of the forecasting techniques for Smooth Representative Material (4003841) is presented in Table 25. Advanced forecasting models considered were Holt-Winters, SARIMAX, LSTM, and SMA with a one-week forecast interval and a window size of 10. The Naive model served as the baseline, with a MASE score of 1.00, an SL of 100%, and an average inventory of 24 527. SARIMAX performed the best among advanced models, with a MASE score of 0.80, a SL of 99.81%, and an average inventory of 23 952. Surprisingly, SMA (1W-10) achieved the lowest MASE score of 0.73 and the lowest average inventory of 18 998.

The comparison between the SMA (1W-10) and SARIMAX methods for forecasting Smooth Representative Material (4003841) reveals interesting insights. Despite both methods achieving service levels above the acceptable threshold of 97%, there are noteworthy differences to consider. Unlike the other materials analyzed, the data for *smooth* Material is not sparse, implying consistency in quantity and time intervals between occurrences of demand.

As shown in Figure 43, material 4003841 transitions from 5% null values to 0% when aggregating from a daily level ("B") to a weekly level ("1W"). However, the number of data points is reduced by 79% when transitioning from a daily ("B") to a weekly ("1W") level, as shown in Table 9. Although the data is not sparse, the reduction of data points makes it challenging for more advanced models to draw conclusions, as they often perform better with larger available datasets (Y. Zhang and Ling, 2018).

Upon examining Figure 70, which visualizes the simulated inventory levels based on the SARIMAX prediction, it becomes evident that the dynamic reorder point (ROP) exhibits significant fluctuations. This fluctuating ROP might be unnecessary given the *smooth* nature of the demand. On the contrary, Figure

68, representing the simulated inventory levels based on the SMA prediction, displays a more stable and smoother dynamic ROP. This observation suggests that the SMA model provides a more appropriate and well-aligned ROP for managing the inventory of the *smooth* material.

The SMA method, with a MASE score of 0.73, outperforms SARIMAX (MASE score of 0.80) in terms of forecast accuracy. Moreover, it achieves a lower average inventory of 18 998 units compared to SARIMAX's average inventory of 23 952 units. This reduction in average inventory amounts to approximately 20.1% lower inventory when using the SMA approach.

The lower average inventory obtained with the SMA model shows its effectiveness when utilized for *smooth* materials. When exploring the simulation results of the SMA for the *smooth* material in Figure 71, the line of ROPs is not far from striking a fixed line of ROP for the simulated period. It indicates that implementing an alternative approach to the dynamic ROP strategy, such as a fixed ROP strategy, could yield similar or even better performance. This can be supported by the results in Section 5.3.3, which showed that for *smooth* materials, employing the "Proposed Fixed ROP" would reduce the average inventory by utterly 1.1% compared with the "Basic Dynamic (OPTIMAL) ROP" strategy. It is important to note that the difference could potentially be larger, in favor of the "Proposed Fixed ROP" strategy, when not utilizing "OPTIMAL" parameters for the Dynamic ROP strategy.

In conclusion, the SMA method demonstrates several advantages over SARIMAX when utilized for *smooth* Representative Material (4003841). It produces a more stable and smoother dynamic ROP, indicates lower average inventory levels, and exhibits similar high service levels. These positive outcomes suggest that exploring alternative approaches, such as a fixed ROP strategy inspired by SMA, could lead to improved inventory management and operational efficiency.

Examining the Linkage

Analyzing the correlation between simulation results and the MASE score, it is observed that a lower MASE score indicates a more accurate forecasting model. However, the relationship between the MASE score and the service level or average inventory is not straightforward. Although the advanced models generally outperformed the Naive model in terms of accuracy and average inventory, there were variations among the advanced models themselves. For example, in the case of Erratic Representative Material (4001095), the LSTM model had the lowest MASE score and achieved a service level of 100%, while the SMA model had a higher MASE score, but resulted in the lowest average inventory. In contrast, for Smooth Representative Material (4003841), the SMA model had a lower MASE score and achieved a relatively lower average inventory compared to advanced models. Thus, better forecasting accuracy does not imply better inventory management in terms of simulation results.

The complexity of the models analyzed for different materials plays a crucial role in their performance. In some cases, LSTM, a deep learning model considered more advanced than SARIMAX and Holt-Winters, may perform worse than simpler models. For the *erratic* material 4001095, the LSTM model exhibits better accuracy in demand prediction but is outperformed by the simpler SMA model in terms of average inventory. Similarly, for the *lumpy* material 4012198, the Holt-Winters model outweighs advanced models in capturing seasonality and trend, although LSTM and SARIMAX underperform in terms of MASE score. In the case of the *smooth* material 4003841, SMA(1W-10) performs better than SARIMAX and LSTM in terms of the MASE score, indicating that the simpler model can adequately capture the demand patterns. The choice of the appropriate model should consider the material's characteristics, the complexity of demand patterns, and the trade-off between accuracy, simplicity, interpretability, and inventory performance.

From these findings, it is evident that there is no direct correlation between the MASE score and the simulation results in terms of service level and average inventory. The advanced forecasting models, such as LSTM and SARIMAX, which are more complex and utilize advanced techniques such as deep learning, may not consistently outperform simpler models like SMA in the simulation. While the advanced models show potential to improve forecast accuracy, their superiority in simulation outcomes, such as achieving higher service levels or lower average inventory, is not guaranteed. These results challenge the assumption that better accuracy in demand forecasting leads directly to better operational performance.

7 Conclusion

The motivation for this thesis concerns the increasing demand for healthcare services in Norway due to demographic changes, including a growing and older population. As the government expects higher demand for care services and lower income growth, there is a need for efficient resource utilization in the public sector, particularly in the health sector (Section 1.1). By leveraging data analytics and machine learning, there is potential to improve inventory management and decision-making.

The existing literature lacks research on applying dynamic reorder points in a continuous replenishment policy and the potential advantages of integrating machine learning methods for demand prediction. Furthermore, there is a significant research gap in exploring this combination along with service-level restrictions. Thus, the primary objective of this study was to:

"Make a contribution to the current literature by investigating the effects of implementing demand forecasting and a dynamic reorder point policy for Logistics Center Helse Midt-Norge"

The objective was achieved by addressing four specific research questions. The selected research questions, the methods used to answer them, the key findings, and the corresponding contributions are presented as follows:

RQ1: "What is the state-of-the-art within demand forecasting for inventory management?"

The first research question focused on established methods and application areas for demand forecasting in inventory management. By conducting a systematic literature review, the study emphasized the importance of considering the intended application area and the characteristics of the data when selecting a demand forecasting method. The reviewed articles highlighted the significance of both statistical and machine learning-based approaches in determining safety stock and other relevant inventory management formulas.

Although machine learning methods are capable of handling fluctuating demand curves with seasonality and trends, it is important to note that a large amount of data does not guarantee superior forecasting accuracy compared to statistical methods. The suitability of a forecasting method depends on factors such as data characteristics, the number of available data sources, and the sample size. Moreover, conducting comprehensive research on demand patterns for each product in large inventories can be impractical due to significant computational costs. Therefore, it is advisable to explore the variables that can classify the SKUs in the inventory before applying any forecasting model. In addition, incorporating external data, such as demographic or economic information, has the potential to enhance forecast accuracy. Therefore, as discussed in Section 6.1, the most appropriate forecast method depends on the data characteristics, the number of available data sources and the sample size.

RQ2: "How can the inventory be classified?"

The systematic literature review (SLR), used to address RQ1, highlighted the importance of assessing demand variability and suggested that categorization items could be advantageous. The second research question aimed to explore the potential this categorization of the inventory of LC HMN. By investigating the possible categorization of materials within the inventory, this study aimed to uncover potential benefits and insights that could enhance the efficiency and effectiveness of LC HMN's operations.

Initially, the study used an algorithmic approach to categorize materials based on two variables: coefficient of variation and inventory turnover ratio. The clustering analysis revealed the presence of diverse demand patterns and highlighted the potential for more frequent replenishment of a significant portion of the inventory. Additionally, the study also examined the effectiveness of a technique devised by Syntetos, Boylan et al. (2005), discovered during a systematic literature review. The technique considered both the demand size and time intervals between demand occurrences. This proved effective in capturing demand fluctuations and divided the materials of LC HMN into the three categories *lumpy*, *erratic*, and *smooth*. The three material categories formed the basis for further analysis linked to research questions three and four.

RQ3: "How can the AS-IS fixed reorder point be improved through a dynamic reorder point?"

In order to answer the third and fourth research questions, a Python-based simulation model was developed, making it possible to simulate conceptual strategies and make evidence-based decisions. Four reorder point strategies were identified and selected for simulation related to the third research question. The AS-IS reorder point strategy was used as a baseline for comparison, while the two proposed dynamic strategies utilizing simple moving average (SMA) forecasting, namely Basic Dynamic ROP, and Basic Dynamic (OPTIMAL) ROP, were conceptual strategies intended to challenge the AS-IS method. Lastly, the Proposed Fixed ROP strategy served as a simple alternative to the more demanding dynamic strategies.

The simulation results demonstrate that both the Basic Dynamic and Basic Dynamic (OPTIMAL) strategies outperform the Proposed Fixed ROP strategy by significantly reducing average inventory while minimizing the impact on service level reduction. Specifically, the OPTIMAL proposal achieves an impressive 42% inventory reduction while maintaining a service level of 98.9%. The results emphasize the significance of the forecast interval in the Basic Dynamic approach, as the Basic Dynamic ROP approach closely approximates the optimal solution for specific forecast intervals. Within the various demand categories, there are significant variations in the results, with the dynamic solution demonstrating superior performance for *erratic* and *lumpy* materials compared to *smooth* ones. Taking into account all demand categories collectively, the greatest reduction in holding costs for all materials is achieved by implementing the Basic Dynamic (OPTIMAL) ROP strategy, resulting in a substantial decrease of 56.3%.

In general, the simulation results for both the Basic Dynamic ROP and the Basic Dynamic (OPTIMAL) ROP indicate favorable results for Logistics Center Helse Midt-Norge. The findings suggest a decrease in average inventory levels and holding costs while still achieving satisfactory service levels. This demonstrates an improvement on the AS-IS fixed reorder point strategy.

RQ4: "What is the impact of implementing advanced forecasting methods for the dynamic reorder point?"

To address research question four, the advanced dynamic ROP strategy was introduced. Four different versions of this strategy were simulated for the three selected representative materials: *erratic* (4001095), *lumpy* (4012198), and *smooth* (4003841). The four versions differed in the forecasting model, employing Naive, Holt-Winters, SARIMAX, and LSTM. Apart from the Naive model, these models are considered significantly more complex than the simple moving average used to answer research question three.

For *erratic*, *lumpy*, and *smooth* materials, the most accurate forecasting models were LSTM, Holt-Winters, and Sarimax, respectively. Despite the more precise forecasting compared to SMA, the simulation results show that the Basic Dynamic ROP strategy outperforms the Advanced Dynamic ROP strategy in terms of average inventory and service level for the three selected materials. As discussed in Section **??**, this may be caused by the limited data available. Although Basic Dynamic ROP outperforms Advanced Dynamic ROP in terms of inventory performance, the difference is minimal for *lumpy* and *erratic* categories.

In total, considering the amount of historical data available, the simulation of the three materials shows that implementing advanced forecasting methods does not have a positive impact on the dynamic reorder point.

7.1 Implementation Guidelines for Dynamic Reorder Point Based on Demand Forecasting

This section provides a comprehensive guide for practitioners on the implementation of dynamic reorder points in a continuous replenishment policy system that is based on demand forecasting for stock-keeping units (SKUs). The guidelines presented are derived from extensive research conducted through a systematic literature review and an empirical case study. By following these guidelines, practitioners can effectively leverage demand forecasting to improve their inventory management processes.

Step 1: Identity Characteristics and Key Performance Indicators

In order to successfully implement a dynamic reorder point policy, a thorough understanding of the inventory operations is crucial. This can be achieved by establishing a control model that outlines the flow of products and information within the system. Additionally, gaining a clear understanding of the supply chain, including the roles of different entities and their interactions, is vital.

Identifying potential requirements and constraints is essential. Specifically, it is important to determine if there are any specific service level constraints that should be considered during the implementation of the dynamic reorder point policy. To compare the dynamic reorder point policy with the existing policy, it is necessary to identify relevant Key Performance Indicator (KPI). Service level and average inventory level are two proposed KPIs in relation to the dynamic reorder point implementation.

Step 2: Data Collection

The second step focuses on data collection. This step involves collecting the necessary demand data for all SKUs within the inventory. It is crucial to obtain a comprehensive understanding of the data by examining and explaining all relevant attributes.

During the data collection process, it is essential to identify which attributes can potentially correlate with demand patterns. These attributes may include factors such as currency, supplier, and unit price. By recognizing such attributes, practitioners can develop more accurate demand forecasts and make informed decisions. Additionally, it is crucial to collect information on the lead time for each SKU. If lead time data are not available, historical records of inbound order placement and order arrival need to be obtained in order to calculate the historical lead times. Understanding the lead time is vital to determine the reorder point, as it helps to account for the time required to replenish stock and prevent stockouts.

Step 3: Data Preprocessing

After collecting the data, the next step is preprocessing. This is essential for preparing the collected data for further analysis and forecasting. This step encompasses standard procedures such as data cleaning, filtering, and feature engineering.

When implementing a dynamic reorder points policy, a categorization of the data is found to be essential. The Syntetos method has been shown to categorize the demand patterns of SKUs effectively and is therefore strongly recommended. This method classifies demand patterns into categories such as *intermittent*, *erratic*, *lumpy*, and *smooth*. Categorizing data based on these patterns helps to understand the nature of demand variability across different SKUs. This understanding enables practitioners to select appropriate forecasting models and design effective reorder point policies that align with the specific demand characteristics of each SKU.

Step 4: Determine The Order Policy to Implement for each category

In this step, the selection of the appropriate order policy for each category is performed. The significance of accurate and reliable data cannot be overstated when establishing a precise ordering policy for each SKU. The following models are recommended for each demand category, allowing the practitioner to assess the most suitable model considering factors such as data quantity, data quality, and the trade-off between implementation cost and corresponding utility:

- *Erratic* category: Dynamic ROP utilizing SMA or Dynamic ROP utilizing LSTM (Long Short-Term Memory)
- *Lumpy* category: Dynamic ROP utilizing SMA or Dynamic ROP utilizing Holt-Winters Exponential Smoothing
- *Smooth* category: Dynamic ROP utilizing SMA or Fixed ROP

Performing time-series analysis, which includes decomposition, autocorrelation, and stationarity tests, is recommended for implementing order policies that utilize advanced forecasting techniques (LSTM and Holt-Winters) for dynamic reorder point (ROP). These analyses provide valuable insights into the underlying patterns, trends, and stationarity properties of the data, allowing the determination of appropriate parameters for advanced models.

For both *erratic* and *lumpy* SKUs, SMA or an advanced model is proposed. When choosing which of the two to utilize for the specific dynamic reorder point policy, several factors should be considered. SMA requires minimal customization and configuration while being computationally lightweight, which makes the forecasting method suitable to apply for a large number of distinct SKUs. On the other hand, advanced forecasting methods need to be individually tuned for each SKU, increasing the overall implementation time and effort. For *smooth* SKUs, it is not advisable to use advanced forecasting models. Instead, it is recommended to consider a fixed reorder point as a viable alternative to SMA.

Step 5: Determine Forecast Interval

The fifth step involves determining the suitable forecast interval for demand forecasting. It is recommended to customize the choice of forecast interval for each demand category, as determined in the third step.

- In the case of the *erratic* category, shorter forecast intervals are proposed, such as daily or weekly. Employing fixed forecast intervals for all SKUs displaying an *erratic* demand pattern is feasible.
- For the *lumpy* category, it is recommended to determine the forecast interval individually for each distinct SKU. This approach takes into account the specific demand behavior of the SKU, leading to a more customized forecasting strategy.
- Regarding the *smooth* category, the findings indicate that shorter forecast intervals, such as daily or weekly, generally produce better outcomes. However, the influence of forecast interval length on this category is not as evident, allowing for the consideration of longer intervals.

In general, implementing dynamic reorder points benefits from shorter forecast intervals, as they tend to provide more accurate results. However, it is important to consider that shorter intervals require more computational resources, which may result in increased costs. Therefore, a careful evaluation and trade-off must be made beyond the suggestions provided. It is also crucial to consider the impact on service levels when determining the forecast interval.

Step 6: Calculation of Dynamic Reorder Point

In Step 6, the logic of the dynamic reorder point calculation should be determined. This calculation needs to be performed for each SKU at each forecast interval period and is therefore computationally heavy. To facilitate this process, it is recommended to implement a custom algorithm, such as one developed in Python.

Developing a self-produced algorithm allows for flexibility and customization in calculating the reorder points based on the specific requirements and characteristics of the inventory system. This approach enables practitioners to incorporate their own business rules and considerations into the algorithm, leading to more tailored and accurate reorder point calculations.

To guide the development of the algorithm, Figure 46 and the corresponding formulas in Table 10 can be utilized as references. These visual representations and formulas provide valuable insights into the factors and variables that should be considered when calculating the dynamic reorder points.

Step 7: Test and evaluate on historical data

Before the real-time implementation of dynamic reorder points, it is essential to conduct tests and evaluations using historical data. This step allows for the identification of potential errors or issues in the implementation process. By testing the dynamic reorder point calculations on historical data, practitioners can gain confidence in the accuracy and effectiveness of the approach before applying it in real-time inventory management.

Step 8: Gradually Implementing in Continuous Replenishment Policy System

After successfully testing and evaluating the dynamic reorder point calculations on historical data, the next step is to gradually implement the approach in the continuous replenishment policy. It is advisable to start with a small sample size during the initial phase of real-time implementation.

Before increasing the sample size, it is recommended to conduct a thorough analysis to ensure that the results align with the desired outcomes. This analysis should consider the KPIs determined in Step 1. By gradually implementing the dynamic reorder point approach and closely monitoring the outcomes, practitioners can ensure a smooth transition and improved inventory management practices. This stepwise approach allows for continuous improvement and adjustment based on the observed performance, enabling practitioners to fine-tune the implementation and achieve the desired inventory management outcomes.

7.2 Contribution

This research makes a significant contribution to the existing literature by exploring the synergies between three distinct fields: machine learning, inventory management, and demand forecasting. The study focused on the practical implications for inventory management at the case company, LC HMN.

By employing simulation-based multi-scenario analysis, this study contributes with evidence-based results, demonstrating the clear advantages and feasibility of implementing a dynamic reorder point policy. The results highlight the positive outcomes of categorizing materials and emphasize the importance of data analysis in achieving efficient inventory management. Furthermore, this study makes a valuable contribution by examining the impact of employing machine learning and other advanced demand forecasting techniques. The research underscores the importance of not undervaluing the use of simpler statistical models, as they can also yield favorable results.

Therefore, this study contributes to the existing scientific literature by laying the foundations for further analysis of the effectiveness of dynamic reorder points in inventory management. It provides a detailed investigation into the advantages and viability of employing data analytics within a particular case company, serving as the basis for future studies examining the wider implementation of data analytics in the healthcare sector of Norway. Consequently, the investigation has successfully achieved its objective and has comprehensively addressed and answered the related research questions.

7.3 Limitations and Further Work

A limitation of this study is that the data set did not take into account the impact of the COVID-19 pandemic on the demand for healthcare services and supplies. As described in Section 5.2.1, the data set included data in the time range 2020-2022 and therefore does not reflect any changes in demand patterns that may have occurred as a result of the pandemic. This means that the results of the analysis may not be fully representative of the current demand for healthcare products and services and may not accurately reflect the potential benefits and challenges of using dynamic reorder points and demand forecasting in the current context.

Another limitation of the study is the lack of consideration and mapping of substitute products. This exclusion could result in an incomplete assessment of the availability and usage of alternative options, which could lead to stockouts in real-world scenarios.

The limited amount of data provided by the case company presents a significant limitation when it comes to advanced demand forecasting. Without sufficient data points, accurate predictions become challenging to achieve. Limited data restrict the ability to identify patterns, trends, and correlations that are essential for robust forecasting models. The findings suggest that this limitation specifically affected the fourth research question.

Further research should explore the possibility of including more data in the analysis. Since historical demand data is limited, it would be interesting to access consumption data directly from hospitals, which are one step further in the supply chain. These additional data could be used to evaluate the potential benefits of implementing a Vendor-Managed Inventory (VMI) system (Waller et al., 1999), where the logistics center takes a proactive approach by initiating deliveries to hospitals based on real-time consumption data.

As mentioned in the scope of the thesis in Section 1.4, the variability in the order size has been excluded from the study. Further research should explore the effect of including variations in order size.

References

Abdallah, Zahraa Said, Lan Du and Geoffrey I Webb (2017). *Data Preparation.*

Albawi, Saad, Tareq Abed Mohammed and Saad Al-Zawi (2017). 'Understanding of a convolutional neural network'. In: *2017 international conference on engineering and technology (ICET).* Ieee, pp. 1–6.

Ali, Mohammad M and John E Boylan (2012). 'On the effect of non-optimal forecasting methods on supply chain downstream demand'. In: *IMA Journal of Management Mathematics* 23.1, pp. 81–98.

Alicke, Knut (2022). *Succeeding in the AI supply-chain revolution.* URL: https://www.mckinsey.com/industries/metals-and-mining/our-insights/succeeding-in-the-ai-supply-chain-revolution (visited on 4th Oct. 2022).

Altunkaynak, Abdüsselam (2014). 'Predicting water level fluctuations in Lake Michigan-Huron using wavelet-expert system methods'. In: *Water resources management* 28.8, pp. 2293–2314.

Alzubaidi, Laith et al. (2021). 'Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions'. In: *Journal of big Data* 8.1, pp. 1–74.

Arnold, J.R. Tony (2017). *Introduction to Materials Management.* Pearson Education, Inc.

Arunraj, Nari Sivanandam and Diane Ahrens (2015). 'A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting'. In: *International Journal of Production Economics.*

Asperud, Brynhild (2020). 'Utfordringer knyttet til covid-19-testing og løsningsforslag fra NITO'. In.

Azzi, Anna et al. (2014). 'Inventory holding costs measurement: a multi-case study'. In: *The International Journal of Logistics Management.*

Bacchetti, Andrea and Nicola Saccani (2012). 'Spare parts classification and demand forecasting for stock control: Investigating the gap between research and practice'. In: *Omega* 40.6, pp. 722–737.

Banks, Jerry (1999). 'Introduction to simulation'. In: *Proceedings of the 31st conference on Winter simulation: Simulation—a bridge to the future-Volume 1*, pp. 7–13.

Baroni, Marco (2020). 'Linguistic generalization and compositionality in modern artificial neural networks'. In: *Philosophical Transactions of the Royal Society B* 375.1791, p. 20190307.

Barrow, Devon K. and Nikolaos Kourentzes (2016). 'Distributions of forecasting errors of forecast combinations: Implications for inventory management'. In: *International Journal of Production Economics.*

Baruch, Yehuda (1999). 'Response rate in academic studies-A comparative analysis'. In: *Human relations* 52.4, pp. 421–438.

Baxter, Pamela, Susan Jack et al. (2008). 'Qualitative case study methodology: Study design and implementation for novice researchers'. In: *The qualitative report* 13.4, pp. 544–559.

Benardos, PG and G-C Vosniakos (2007). 'Optimizing feedforward artificial neural network architecture'. In: *Engineering applications of artificial intelligence* 20.3, pp. 365–382.

Beutel, Anna-Lena and Stefan Minner (2012). 'Safety stock planning under causal demand forecasting'. In: *International Journal of Production Economics* 140.2, pp. 637–645.

Bhagwat, Rajat and Milind Kumar Sharma (2007). 'Performance measurement of supply chain management: A balanced scorecard approach'. In: *Computers & industrial engineering.*

Bojer, Casper Solheim and Jens Peder Meldgaard (2021). 'Kaggle forecasting competitions: An overlooked learning opportunity'. In: *International Journal of Forecasting.*

Bollerslev, Tim (1986). 'Generalized autoregressive conditional heteroskedasticity'. In: *Journal of econometrics* 31.3, pp. 307–327.

Boulanger-Lewandowski, Nicolas, Yoshua Bengio and Pascal Vincent (2012). 'Modeling temporal dependencies in high-dimensional sequences: Application to polyphonic music generation and transcription'. In: *arXiv preprint arXiv:1206.6392.*

Brahimi, Nadjib et al. (2017). 'Single-item dynamic lot-sizing problems: An updated survey'. In: *European Journal of Operational Research* 263.3, pp. 838–863.

Breque, Maija, Lars De Nul and Athanasios Petridis (2021). 'Industry 5.0: towards a sustainable, human-centric and resilient European industry'. In: *Luxembourg, LU: European Commission, Directorate-General for Research and Innovation.*

Brownlee, Jason (2021). *Ensemble learning algorithms with Python: Make better predictions with bagging, boosting, and stacking.* Machine Learning Mastery.

— (2022). *Data preparation for machine learning.*

Cao, Jian, Zhi Li and Jian Li (2019). 'Financial time series forecasting model based on CEEMDAN and LSTM'. In: *Physica A: Statistical mechanics and its applications* 519, pp. 127–139.

Carbonneau, Real, Kevin Laframboise and Rustam Vahidov (2008). 'Application of machine learning techniques for supply chain demand forecasting'. In: *European Journal of Operational Research.*

Cerqueira, Vitor, Luis Torgo and Carlos Soares (2019). 'Machine learning vs statistical methods for time series forecasting: Size matters'. In: *arXiv preprint arXiv:1909.13316.*

Chai, Tianfeng and Roland R Draxler (2014). 'Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature'. In: *Geoscientific model development* 7.3, pp. 1247–1250.

Chambers, John C, Satinder K Mullick and Donald D Smith (1971). *How to choose the right forecasting technique.* Harvard University, Graduate School of Business Administration Cambridge, MA ...

Chandrasekar, Priyanga et al. (2017). 'Improving the prediction accuracy of decision tree mining with data preprocessing'. In: *2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC).* Vol. 2. IEEE, pp. 481–484.

Chase, Charles W (2016). *Next generation demand management: People, process, analytics, and technology.* John Wiley & Sons.

Che, Zhengping et al. (2018). ‘Recurrent neural networks for multivariate time series with missing values’. In: *Scientific reports*.

Chen, Tianqi and Carlos Guestrin (2016). ‘XGBoost: A Scalable Tree Boosting System’. In: New York, NY, USA: Association for Computing Machinery. URL: https://doi.org/10.1145/2939672.2939785.

Christoffersen, Line and Asbjørn Johannessen (2012). *Forskningsmetode for lærerutdanningene*. Abstrakt.

Chu, Xu and Ihab F Ilyas (2019). *Data Cleaning*. ACM/Association for Computing Machinery.

Chuang, Howard Hao-Chun, Yen-Chun Chou and Rogelio Oliva (2021). ‘Cross-item learning for volatile demand forecasting: An intervention with predictive analytics’. In: *Journal of Operations Management* 67.7, pp. 828–852.

Clausen, Johan Bjerre Bach and Hongyan Li (2022). ‘Big Data Driven Order-up-to Level Model: Application of Machine Learning’. In: *Comput. Oper. Res.*

Cleveland, Robert B et al. (1990). ‘STL: A seasonal-trend decomposition’. In: *J. Off. Stat* 6.1, pp. 3–73.

Costantino, Francesco et al. (2016). ‘Smoothing inventory decision rules in seasonal supply chains’. In: *Expert Systems with Applications* 44, pp. 304–319.

Dejonckheere, Jeroen et al. (2003). ‘Measuring and avoiding the bullwhip effect: A control theoretic approach’. In: *European journal of operational research* 147.3, pp. 567–590.

Dickey, David A and Wayne A Fuller (1979). ‘Distribution of the estimators for autoregressive time series with a unit root’. In: *Journal of the American statistical association*.

DiPietro, Robert and Gregory D Hager (2020). ‘Deep learning: RNNs and LSTM’. In: *Handbook of medical image computing and computer assisted intervention*. Elsevier, pp. 503–519.

Dodgson, Joan E (2019). ‘Reflexivity in qualitative research’. In: *Journal of Human Lactation* 35.2, pp. 220–222.

Dolgui, Alexandre, Aly Ould Louly and Caroline Prodhon (2005). ‘A survey on supply planning under uncertainties in MRP environments’. In: *IFAC Proceedings Volumes* 38.1, pp. 1–12.

Dolgui, Alexandre and Caroline Prodhon (2007). ‘Supply planning under uncertainties in MRP environments: A state of the art’. In: *Annual reviews in control* 31.2, pp. 269–279.

Du, Cheng-Jin and Da-Wen Sun (2006). ‘Learning techniques used in computer vision for food quality evaluation: a review’. In: *Journal of food engineering* 72.1, pp. 39–55.

Elith, Jane, John R Leathwick and Trevor Hastie (2008). ‘A working guide to boosted regression trees’. In: *Journal of animal ecology* 77.4, pp. 802–813.

Flyvbjerg, Bent (2011). ‘Case study’. In: *The Sage handbook of qualitative research* 4, pp. 301–316.

Fu, Wenhan, Chen-Fu Chien and Zih-Hao Lin (2018). ‘A hybrid forecasting framework with neural network and time-series method for intermittent demand in semiconductor supply chain’. In: *IFIP International Conference on Advances in Production Management Systems*. Springer, pp. 65–72.

Gao, Dandan et al. (2022). 'Inventory Management'. In: *Enterprises' Green Growth Model and Value Chain Reconstruction: Theory and Method*. Springer, pp. 251–269.

García, Juan Andrés Martín and Antonio José Gil Mena (2013). 'Optimal distributed generation location and size using a modified teaching–learning based optimization algorithm'. In: *International journal of electrical power & energy systems* 50, pp. 65–75.

Gardner Jr, Everette S (2006). 'Exponential smoothing: The state of the art—Part II'. In: *International journal of forecasting* 22.4, pp. 637–666.

Gerring, John (2004). 'What is a case study and what is it good for?' In: *American political science review* 98.2, pp. 341–354.

Gilbert, Kenneth (2005). 'An ARIMA supply chain model'. In: *Management Science* 51.2, pp. 305–310.

Glock, Christoph H, Eric H Grosse and Jörg M Ries (2014). 'The lot sizing problem: A tertiary study'. In: *International Journal of Production Economics* 155, pp. 39–51.

Goltsos, Thanos E et al. (2021). 'Inventory–forecasting: Mind the gap'. In: *European Journal of Operational Research*.

Gregersen, Fredrik Alexander (2014). 'The impact of ageing on health care expenditures: a study of steepening'. eng. In: *The European journal of health economics* 15.9, pp. 979–989. ISSN: 1618-7598.

Grunwald, Gary K, Kais Hamza and Rob J Hyndman (1997). 'Some properties and generalizations of non-negative Bayesian time series models'. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*.

Hall, Mark Andrew and Geoffrey Holmes (2003). 'Benchmarking attribute selection techniques for discrete class data mining'. In: *IEEE Transactions on Knowledge and Data engineering* 15.6, pp. 1437–1447.

Du-Harpur, X et al. (2020). 'What is AI? Applications of artificial intelligence to dermatology'. In: *British Journal of Dermatology* 183.3, pp. 423–430.

Harvey, Andrew and Ralph D. Snyder (1990). 'Structural time series models in inventory control'. In: *International Journal of Forecasting*.

Hawkins, Douglas M (2004). 'The problem of overfitting'. In: *Journal of chemical information and computer sciences* 44.1, pp. 1–12.

Hill, Arthur V, Weiyong Zhang and Gerald F Burch (2015). 'Forecasting the forecastability quotient for inventory management'. In: *International Journal of Forecasting* 31.3, pp. 651–663.

Hochreiter, Sepp (1998). 'The vanishing gradient problem during learning recurrent neural nets and problem solutions'. In: *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 6.02, pp. 107–116.

Hofer, Christian, Cuneyt Eroglu and Adriana Rossiter Hofer (2012). 'The effect of lean production on financial performance: The mediating role of inventory leanness'. In: *International Journal of Production Economics* 138.2, pp. 242–253.

Holden, Lauren (2017). 'Inventory Optimization Using a SimPy Simulation Model'. PhD thesis. East Tennessee State University.

Holmøy, Erling. (2020). 'Skatteregningen for helse-og omsorgsutgifter mot 2060'. In.

Hopayian, Kevork (2001). 'The need for caution in interpreting high quality systematic reviews'. In: *Bmj* 323.7314, pp. 681–684.

Hua, Zhongsheng and Bin Zhang (2006). 'A hybrid support vector machines and logistic regression approach for forecasting intermittent demand of spare parts'. In: *Applied Mathematics and Computation*.

Hunter, John D. (2007). 'Matplotlib: A 2D Graphics Environment'. In: *Computing in Science & Engineering*.

Hyndman, Rob J and George Athanasopoulos (2018). *Forecasting: principles and practice*. OTexts.

Hyndman, Rob J and Anne B Koehler (2006). 'Another look at measures of forecast accuracy'. In: *International journal of forecasting* 22.4.

JAIN, ASHU and LINDELL E ORMSBEE (2001). 'A decision support system for drought characterization and management'. In: *Civil Engineering Systems* 18.2, pp. 105–140.

Jakhar, Deepack and Ishmeet Kaur (2020). 'Artificial intelligence, machine learning and deep learning: definitions and differences'. In: *Clinical and experimental dermatology* 45.1, pp. 131–132.

Jans, Raf and Zeger Degraeve (2008). 'Modeling industrial lot sizing problems: a review'. In: *International Journal of Production Research* 46.6, pp. 1619–1643.

Jiang, P et al. (2017). 'An adaptive data-driven model for erratic demand forecasting'. In: *IFAC-PapersOnLine* 50.1, pp. 9020–9025.

John J. Bartholdi, III and Steven T. Hackman (2019). 'Warehouse & Distribution Science'. In: Georgia Institute of Technology, p. 296.

Karimnezhad, Ali and Fahimeh Moradi (2016). 'Bayes, E-Bayes and robust Bayes prediction of a future observation under precautionary prediction loss functions with applications'. In: *Applied Mathematical Modelling* 40.15-16, pp. 7051–7061.

Kaur, Parampreet, Jill Stoltzfus, Vikas Yellapu et al. (2018). 'Descriptive statistics'. In: *International Journal of Academic Medicine* 4.1, p. 60.

Khashei, Mehdi and Mehdi Bijari (2010). 'An artificial neural network (p, d, q) model for timeseries forecasting'. In: *Expert Systems with applications* 37.1, pp. 479–489.

Kiefer, Daniel et al. (2021). 'Demand forecasting intermittent and lumpy time series: Comparing statistical, machine learning and deep learning methods'. In: vol. 2020-January. Cited by: 7, pp. 1425–1434. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85104188541&partnerID=40&md5=429cf3c921c51421ea631ce2d15ad8f2.

Kim, Jeong-A and Jongpil Jeong (2018). 'Simulation Evaluation for Efficient Inventory Management Based on Demand Forecast'. In: *Computational Science and Its Applications – ICCSA 2018*. Ed. by Osvaldo Gervasi et al. Springer International Publishing.

Kourentzes, Nikolaos, Juan R. Trapero and Devon K. Barrow (2020). ‘Optimising forecasting models for inventory planning’. In: *International Journal of Production Economics.*

Kozik, Piotr and Jarosław Sęp (2012). ‘Aircraft engine overhaul demand forecasting using ANN’. In: *Management and Production Engineering Review.*

Krizhevsky, Alex, Ilya Sutskever and Geoffrey E Hinton (2017). ‘Imagenet classification with deep convolutional neural networks’. In: *Communications of the ACM* 60.6, pp. 84–90.

LeCun, Yann, Yoshua Bengio and Geoffrey Hinton (2015). ‘Deep learning’. In: *nature* 521.7553, pp. 436–444.

Leknes, Stefan et al. (2019). ‘Regionale framskrivinger av etterspørsel etter helse-og omsorgstjenester, 2017-2035’. In.

Li, Na et al. (2021). ‘A decision integration strategy for short-term demand forecasting and ordering for red blood cell components’. In: *Operations Research for Health Care.*

Liao, T Warren and Pei-Chann Chang (2010). ‘Impacts of forecast, inventory policy, and lead time on supply chain inventory—a numerical study’. In: *International Journal of Production Economics* 128.2, pp. 527–537.

Makridakis, Spyros, Evangelos Spiliotis and Vassilios Assimakopoulos (2020). ‘The M4 Competition: 100,000 time series and 61 forecasting methods’. In: *International Journal of Forecasting.*

Mamre, M. et al. (n.d.). ‘Kommunalundersøkelsen 2017 - Status for digitalisering i norske kommuner’. In: *Evry* ().

Matloff, Norm (2008). ‘Introduction to discrete-event simulation and the simpy language’. In: *Davis, CA. Dept of Computer Science. University of California at Davis. Retrieved on August* 2.2009.

McKinney, Wes (2010). ‘Data Structures for Statistical Computing in Python’. In: *Proceedings of the 9th Python in Science Conference.*

Meld.St (2017). *Perspektivmeldingen 2017.*

— (2021). *Data som ressurs.*

Mellbye, Christian S. and Caroline W. Gierløff (2018). ‘Smarte kommuner - Hva er gevinstpotensialet?’ In: *Menon Economics.*

Midt-Norge-RHF, Helse (2022). ‘Her er Helse Midt-Norges nye logistikksenter’. In.

Miller, James N (1993). ‘Tutorial review—Outliers in experimental data and their treatment’. In: *Analyst* 118.5, pp. 455–461.

Morissette, Laurence and Sylvain Chartier (2013). ‘The k-means clustering technique: General considerations and implementation in Mathematica’. In: *Tutorials in Quantitative Methods for Psychology* 9.1, pp. 15–24.

Moroff, Nikolas Ulrich, Ersin Kurt and Josef Kamphues (2021a). ‘Machine Learning and Statistics: A Study for assessing innovative Demand Forecasting Models’. In: *Procedia Computer Science.*

— (2021b). ‘Machine Learning and statistics: A Study for assessing innovative demand forecasting models’. In: *Procedia Computer Science* 180, pp. 40–49.

Namey, Emily et al. (2008). 'Data reduction techniques for large qualitative data sets'. In: *Handbook for team-based qualitative research* 2.1, pp. 137–161.

Nasteski, Vladimir (2017). 'An overview of the supervised machine learning methods'. In: *Horizons. b* 4, pp. 51–62.

O'Shea, Keiron and Ryan Nash (2015). 'An introduction to convolutional neural networks'. In: *arXiv preprint arXiv:1511.08458*.

Olah, Christopher (2022). *Understanding LSTM Networks*. URL: https://colah.github.io/posts/2015-08-Understanding-LSTMs/index.html (visited on 28th Oct. 2022).

Omsorgsdepartementet, Helse- og (2022). *Nasjonal helse- og sykehusplan 2020-2023*. URL: https://www.regjeringen.no/contentassets/e353a5d022d84deabd969a5fe043783e/no/pdfs/i-1194_b_kortversjon_nasjonal_helse.pdf (visited on 4th Oct. 2022).

Özsen, Raik and Ulrich W Thonemann (2015). 'Determining optimal parameters for expediting policies'. In: *Manufacturing & Service Operations Management*.

Pacella, Massimo and Gabriele Papadia (2021). 'Evaluation of deep learning with long short-term memory networks for time series forecasting in supply chain management'. In: *Procedia CIRP*.

Pandas (Feb. 2020). *pandas-dev/pandas: Pandas*. Version latest. DOI: 10.5281/zenodo.3509134. URL: https://doi.org/10.5281/zenodo.3509134.

Pedregosa, F. et al. (2011). 'Scikit-learn: Machine Learning in Python'. In: *Journal of Machine Learning Research* 12, pp. 2825–2830.

Pedregosa, Fabian et al. (2011). 'Scikit-learn: Machine learning in Python'. In: *the Journal of machine Learning research*.

Perera, Kushan C et al. (2016). 'Ensemble forecasting of short-term system scale irrigation demands using real-time flow data and numerical weather predictions'. In: *Water Resources Research* 52.6, pp. 4801–4822.

Piasecki, David J (2009). *Inventory management explained: a focus on forecasting, lot sizing, safety stock, and ordering systems*. Ops Publishing.

Pollack-Johnson, Bruce et al. (1990). 'Predicting doctorate production in the U.S.A.: Some lessons for long-range forecasters'. In: *International Journal of Forecasting*.

Qin, Yao et al. (2017). 'A dual-stage attention-based recurrent neural network for time series prediction'. In: *arXiv preprint arXiv:1704.02971*.

Radasanu, Alin Constantin et al. (2016). 'Inventory management, service level and safety stock'. In: *Journal of Public Administration, Finance and Law* 09, pp. 145–153.

Raudys, Sarunas J, Anil K Jain et al. (1991). 'Small sample size effects in statistical pattern recognition: Recommendations for practitioners'. In: *IEEE Transactions on pattern analysis and machine intelligence* 13.3, pp. 252–264.

Rowley, Jennifer and Frances Slack (2004). 'Conducting a literature review'. In: *Management research news*.

S. Chopra P. Meindl, D.V. Kalra (2016). *Supply Chain Management: Strategy, Planning and Operation*. Pearson India Education Services.

Samuel, Arthur L (1988). 'Some studies in machine learning using the game of checkers. II—recent progress'. In: *Computer Games I*, pp. 366–400.

Santos, João A.M. et al. (2022). 'Many-objective optimization of a three-echelon supply chain: A case study in the pharmaceutical industry'. In: *Computers & Industrial Engineering*.

Sarker, Iqbal H (2021). 'Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions'. In: *SN Computer Science* 2.6, p. 420.

Sattler, K-U and Eike Schallehn (2001). 'A data preparation framework based on a multidatabase language'. In: *Proceedings 2001 International Database Engineering and Applications Symposium*. IEEE, pp. 219–228.

Seabold, Skipper and Josef Perktold (2010). 'Statsmodels: Econometric and statistical modeling with python'. In: *9th Python in Science Conference*.

See, Chuen-Teck and Melvyn Sim (2010). 'Robust approximation to multiperiod inventory management'. In: *Operations research* 58.3, pp. 583–594.

Seuring, Stefan and Martin Muller (2008). 'From a literature review to a conceptual framework for sustainable supply chain management'. In: *Journal of cleaner production* 16.15, pp. 1699–1710.

Shumway, Robert H. (2017). *Time Series Analysis and Its Applications*. Springer International.

Siami-Namini, Sima, Neda Tavakoli and Akbar Siami Namin (2019). 'The performance of LSTM and BiLSTM in forecasting time series'. In: *2019 IEEE International Conference on Big Data (Big Data)*. IEEE, pp. 3285–3292.

Sillanpää, Ville and Juuso Liesiö (2018). 'Forecasting replenishment orders in retail: value of modelling low and intermittent consumer demand with distributions'. In: *International Journal of Production Research* 56.12, pp. 4168–4185.

Simpy (2023). *Simpy Docs*. URL: https://simpy.readthedocs.io/ (visited on 9th Mar. 2023).

Skodbo, Morten (2017). 'IT i praksis2017'. In.

Snyder, Ralph (2002). 'Forecasting sales of slow and fast moving inventories'. In: *European Journal of Operational Research*.

Snyder, Ralph D, J Keith Ord and Adrian Beaumont (2012). 'Forecasting the intermittent demand for slow-moving inventories: A modelling approach'. In: *International Journal of Forecasting* 28.2, pp. 485–496.

Spedding, TA and KK Chan (2000). 'Forecasting demand and inventory management using Bayesian time series'. In: *Integrated Manufacturing Systems*.

SSB (2022). *Befolkning*. URL: https://www.ssb.no/befolkning/folketall/statistikk/befolkning?fane=tabell%5C&sort=nummer%5C&tabell=212146 (visited on 11th Dec. 2022).

Stevenson, William J (2021). *Operations management*. McGraw-Hill Education New York.

Syntetos, Aris A, M Zied Babai and Everette S Gardner Jr (2015). 'Forecasting intermittent inventory demands: simple parametric methods vs. bootstrapping'. In: *Journal of Business Research* 68.8, pp. 1746–1752.

Syntetos, Aris A, John E Boylan and JD Croston (2005). 'On the categorization of demand patterns'. In: *Journal of the operational research society*.

Tang, Zheng and Yuemeng Ge (2021). ‘CNN model optimization and intelligent balance model for material demand forecast’. In: *International Journal of System Assurance Engineering and Management*, pp. 1–9.

Tasdemir, Cagatay and Salim Hiziroglu (2019). ‘Achieving cost efficiency through increased inventory leanness: Evidences from oriented strand board (OSB) industry’. In: *International Journal of Production Economics*.

Tjora, A (2021). *Kvalitative forskningsmetoder i praksis*. Gyldendal Akademisk.

Tranfield, David, David Denyer and Palminder Smart (2003). ‘Towards a methodology for developing evidence-informed management knowledge by means of systematic review’. In: *British journal of management* 14.3, pp. 207–222.

Treyer, Oscar AG (2010). *Business Forecasting: Anwendungsorientierte Theorie quantitativer Prognoseverfahren*. 3365. Haupt.

Tsai, Shing Chih and Chung Hung Liu (2015). ‘A simulation-based decision support system for a multi-echelon inventory problem with service level constraints’. In: *Computers & Operations Research*.

Tu, Zhuozhuo, Fengxiang He and Dacheng Tao (2019). ‘Understanding generalization in recurrent neural networks’. In: *International Conference on Learning Representations*.

Turing, Alan M (2009). ‘Computing machinery and intelligence’. In: *Parsing the turing test*. Springer, pp. 23–65.

Vagropoulos, Stylianos I. et al. (2016). ‘Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting’. In: *2016 IEEE International Energy Conference (ENERGYCON)*.

Van der Auweraer, Sarah, Robert N Boute and Aris A Syntetos (2019). ‘Forecasting spare part demand with installed base information: A review’. In: *International Journal of Forecasting*.

Vathsala, MK and Ganga Holi (2020). ‘RNN based machine translation and transliteration for Twitter data’. In: *International Journal of Speech Technology* 23.3, pp. 499–504.

Vercellis, Carlo (2011). *Business intelligence: data mining and optimization for decision making*. John Wiley & Sons.

Virtanen, Pauli et al. (2020). ‘SciPy 1.0: fundamental algorithms for scientific computing in Python’. In: *Nature Methods*.

Waller, Matt, M Eric Johnson and Tom Davis (1999). ‘Vendor-managed inventory in the retail supply chain’. In: *Journal of business logistics* 20.1, p. 183.

Wang, Chien-Chih, Chun-Hua Chien and Amy JC Trappey (2021). ‘On the application of ARIMA and LSTM to predict order demand based on short lead time and on-time delivery requirements’. In: *Processes* 9.7, p. 1157.

Wang, Jinjiang et al. (2016). ‘A multi-scale convolution neural network for featureless fault diagnosis’. In: *2016 International Symposium on Flexible Automation (ISFA)*. IEEE, pp. 65–70.

Wang, Lei and Haoxun Chen (2022). ‘Optimization of a stochastic joint replenishment inventory system with service level constraints’. In: *Computers & Operations Research* 148, p. 106001.

Waskom, Michael L (2021). 'Seaborn: statistical data visualization'. In: *Journal of Open Source Software*.

Wenhan Fu, Chen-Fu Chien & Zih-Hao Lin (2018). 'A Hybrid Forecasting Framework with Neural Network and Time-Series Method for Intermittent Demand in Semiconductor Supply Chain'. In: *IFIP Advances in Information and Communication Technology*.

Wheelwright, Steven, Spyros Makridakis and Rob J Hyndman (1998). *Forecasting: methods and applications*. John Wiley & Sons.

Willemain, Thomas R., Charles N. Smart and Henry F. Schwarz (2004). 'A new approach to forecasting intermittent demand for service parts inventories'. In: *International Journal of Forecasting*.

Williams, Billy M and Lester A Hoel (2003). 'Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results'. In: *Journal of transportation engineering*.

Willmott, Cort J and Kenji Matsuura (2005). 'Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance'. In: *Climate research* 30.1, pp. 79–82.

Winters, Peter R (1960). 'Forecasting sales by exponentially weighted moving averages'. In: *Management science*.

Wohlin, Claes (2014). 'Guidelines for snowballing in systematic literature studies and a replication in software engineering'. In: *Proceedings of the 18th international conference on evaluation and assessment in software engineering*, pp. 1–10.

Wolpert, David H (1996). 'The lack of a priori distinctions between learning algorithms'. In: *Neural computation* 8.7, pp. 1341–1390.

Yamashita, Rikiya et al. (2018). 'Convolutional neural networks: an overview and application in radiology'. In: *Insights into imaging* 9.4, pp. 611–629.

Yang, Guangyu Robert and Xiao-Jing Wang (2020). 'Artificial neural networks for neuroscientists: a primer'. In: *Neuron* 107.6, pp. 1048–1070.

Yin, Robert K (2009). *Case study research: Design and methods*. Vol. 5. sage.

Zahraie, Banafsheh, Mohsen Nasseri and Fariborz Nematizadeh (2017). 'Exploring spatiotemporal meteorological correlations for basin scale meteorological drought forecasting using data mining methods'. In: *Arabian Journal of Geosciences* 10.19, pp. 1–15.

Zarandi, MH Fazel, Morteza Pourakbar and IB Turksen (2006). 'An intelligent agent-based system for reduction of bullwhip effect in supply chains'. In: *2006 IEEE International Conference on Fuzzy Systems*. IEEE, pp. 663–670.

Zhang, Senlin et al. (2017). 'Data-based line trip fault prediction in power systems using LSTM networks and SVM'. In: *Ieee Access* 6, pp. 7675–7686.

Zhang, Xiaolong (2007). 'Inventory control under temporal demand heteroscedasticity'. In: *European Journal of Operational Research*.

Zhang, Ying and Chen Ling (2018). 'A strategy to apply machine learning to small datasets in materials science'. In: *Npj Computational Materials* 4.1, p. 25.

Zhao, Bendong et al. (2017). 'Convolutional neural networks for time series classification'. In: *Journal of Systems Engineering and Electronics* 28.1, pp. 162–169.

Zohra Benhamida, Fatima et al. (2021). 'Demand Forecasting Tool For Inventory Control Smart Systems'. In: *Journal of Communications Software and Systems*.

Appendix

A Service Level

Service Level (%)	Safety Factor
50	0.00
75	0.67
80	0.84
85	1.04
90	1.28
94	1.56
95	1.65
96	1.75
97	1.88
98	2.05
99	2.33
99.86	3.00
99.99	4.00

Figure 72: Table of service level and safety factors (Arnold, 2017)

B SLR Result Tables

Table 27: Results of SLR

Reference	Objective	Method		Application Area			
		Machine Learning	**Statistical**	**Supply Chain**	**Manufacturing**	**Operation Management**	**Inventory Controll**
Harvey and Ralph D. Snyder, 1990	Highlight the potential role of exponential smoothing models in inventory control		Exponential smoothing, exponentially weighted moving average				x
Spedding and Chan, 2000	Comparing different models for demand forecasting for inventory management		Bayesian dynamic linear, ARIMA				x
JAIN and ORMSBEE, 2001	Short-term water demand forecasting	ANN	Auto-Regressive, Moving Average Regression			x	
R. Snyder, 2002	Forecasting sales of slow and fast moving car parts		Exponential smoothing, Bernoulli & ARIMA (Cronston method)				x
Willemain et al., 2004	Address the problem of forecasting intermittent (or irregular) demand		Exponential smoothing, Cronston's method, variant of bootstrap		x		
Gilbert, 2005	Model demand in supply chain		ARIMA	x			
Gardner Jr, 2006	The state of the art in exponential smoothing		Exponential Smoothing, ARIMA				x

Table 28: Results of SLR

Reference	Objective	Method		Application Area			
		Machine Learning	**Statistical**	**Supply Chain**	**Manufacturing**	**Operation Management**	**Inventory Controll**
Hua and B. Zhang, 2006	Hybrid forecasting model, forecasting intermittent demand	SVM, Logistic Regression	For comparing models: exponential smoothing, Cronston's method				x
Zarandi et al., 2006	Reduce bullwhip effect through predicting fuzzy demand	Neural Network (for pattern classification)	ARIMA	x			
X. Zhang, 2007	The effect of variance heterogeneity on the performance of inventory control		ARIMA				x
Jans and Degraeve, 2008	Dynamic Lot-sizing		Dantzig–Wolfe decomposition, Lagrange relaxation				x
Liao and Chang, 2010	Impact of forecasting: numerical study		ACOR (comb. of exp. smoothing, Holt Method, Dholt, Pegel Method and DPegel)	x			x
See and Sim, 2010	Optimize a multiperiod inventory control problem under ambiguous demand		ARMA				x

Table 29: Results of SLR

Reference	**Objective**	**Method**		**Application Area**			
		Machine Learning	**Statistical**	**Supply Chain**	**Manufacturing**	**Operation Management**	**Inventory Controll**
Ralph D Snyder et al., 2012	Examination of various different approaches to demand forecasting for products with intermittent demand		Exponential smoothing, Cronston method, Harvey-Fernandes method				x
Kozik and Sęp, 2012	Spare part demand prediction	Multilayer perceptron Artificial Neural Network (MLP ANN)					x
Glock et al., 2014	Systematic literature review		Harris' model				x
Altunkaynak, 2014	Predicting water level fluctations	ANN, MLP		x			
Arunraj and Ahrens, 2015	Predicting food sales	SARIMA, SARIMAX	Multiple Linear Regression, Quantile Regression				x
Hill et al., 2015	Predict the forecastability quotient Q		Single Exponential Smoothing, Double Exponential Smoothing				x
Syntetos, Babai et al., 2015	Forecast inventory demand, bootstrapping vs. exponential smoothing		Single Exponential Smoothing, Crostons method, Syntetos-Boylan				x

Table 30: Results of SLR

Reference	Objective	Method		Application Area			
		Machine Learning	**Statistical**	**Supply Chain**	**Manufacturing**	**Operation Management**	**Inventory Controll**
Barrow and Kourentzes, 2016	Investigate the effect that forecast combinations have on the shape of the forecast error distributions		Exponential Smoothing, AR, ARIMA, Theta Method, MAPA				x
Perera et al., 2016	Forecast weather predictions with ensemble forecasting		ARIMAX			x	
Zahraie et al., 2017	Forecasting meteorological drought	SVM, GMDH				x	
Brahimi et al., 2017	Single-Item Lot-Sizing Problem (SILSP), dynamic demand		Lagrangian relaxation, Dynamicx Programming				x
Sillanpää and Liesiö, 2018	Examine the added value of modelling consumer demand with distributions		point-estimate model, probabilistic model	x			x
Beutel and Minner, 2012	Compare methods for inventory levels and safety stock		least squares regression, Linear Programming model				x

Table 31: Results of SLR

Reference	Objective	Method		Application Area			
		Machine Learning	**Statistical**	**Supply Chain**	**Manufacturing**	**Operation Management**	**Inventory Controll**
Fu et al., 2018	Hybrid demand forecasting approach to capture the intermittent demand pattern of semiconductor products	RNN (LSTM)	Syntetos-Boylan Approx.				x
Kim and Jeong, 2018	Predict demand for a manufacturing factory		ARIMA				x
Tasdemir and Hiziroglu, 2019	Increase inventory leanness thorugh forecasting demand		Winters Seasonal Multiplicative Forecasting Model, Regression Based Forecast Modelling				x
Tang and Ge, 2021	Improve the accuracy of inventory demand forecast	CNN, CNN-LSTM (Hybrid)			x		x
Chuang et al., 2021	Demand forecasting	Multilayer Perception (MLP)	Conston's method, Exponential Smoothing, Moving Average, Syntetos-Boylan Approx., Teuner-Syntetos-Babai method, ARIMA, Error Trend Seasonality (ETS)				x

Table 32: Results of SLR

Reference	Objective	Method		Application Area			
		Machine Learning	**Statistical**	**Supply Chain**	**Manufacturing**	**Operation Management**	**Inventory Controll**
C.-C. Wang et al., 2021	Demand forecasting	LSTM	ARIMA				x
Pacella and Papadia, 2021	Customer demand forecasting technique comparison	Forward LSTM, Bidirectional LSTM		x			
Zohra Benhamida et al., 2021	Demand forecasting tool for inventory control	Feed Forward Neural Network	Theta, Arima				x
Goltsos et al., 2021	Systematic Literature Review	Machine Learning (ML)					x
Clausen and H. Li, 2022	Optimization	Artificial Neural Network (ANN)					x
N. Li et al., 2021	Hybrid demand forecasting for blood demand	LSTM, XGBoost	Linear Regression, STL				x

NTNU

Norwegian University of
Science and Technology

ST. OLAVS HOSPITAL
UNIVERSITETSSYKEHUSET I TRONDHEIM