Redesigning a demand planning process in an integrated planning environment of an internal end-to-end supply chain

Martens, L.J.M.

Award date:
2014
Redesigning a demand planning process in an integrated planning environment of an internal end-to-end supply chain

by
L.J.M. Martens

BSc Industrial Engineering and Management
Student identity number 0827437

in partial fulfilment of the requirements for the degree of

Master of Science
in Operations Management and Logistics

University supervisors
Prof. dr. A.G. de Kok, Eindhoven University of Technology, OPAC
Dr. K. H. van Donselaar, Eindhoven University of Technology, OPAC

Company supervisor
A. Sasso, Hilti AG, Head of Materials Management MO/Region
Subject headings: supply chain integration, demand planning, hierarchical demand planning, demand forecasting, hierarchical forecasting, top-down forecasting, forecast evaluation, forecast error measures
Abstract

As part of an overall project to integrate the internal end-to-end supply chain, Hilti AG aims to redesign the demand planning decision function in order to improve its contribution to the overall supply chain performance. In this project, scientific recommendations are provided to guide this process of redesigning.

In order to provide demand forecast information for the tactical and operational level of the planning hierarchy, the functionality of Hierarchical Forecasting (HF) has been examined. It has been concluded however that scientific literature does not provide any general guidelines on the definition of optimal forecast levels using HF. A quantitative analysis based on sales data of Hilti AG indicated that it seems beneficial to know the optimal forecasting level per cluster in advance. Hilti AG is therefore recommended to experiment with software assistance to perform data-analysis in order to define the optimal forecasting level per cluster in advance. Since the optimal level of forecasting showed to be potentially unstable over time, this experimentation should be done with caution.

From a process perspective, Hilti AG is suggested to review their statistical and consensus forecasting activities based on the idealized process design, provided in this project. Furthermore, it is recommended to centralize the responsibility of aligning demand and supply plans.

A combination of the forecast accuracy metrics \( wMAPE \) and \( MPE \) has been suggested to properly provide feedback on demand planning performance and support resource allocation decisions. In addition, guidelines have been provided on the methodology of measurement and target setting.
Acknowledgements

The completion of this master thesis marks the end of my master Operations Management and Logistics at the University of Technology Eindhoven. Especially in a broad area as Operations Management and Logistics, graduating is not a one-man's project. I would like to dedicate some lines to showing my gratitude to all that have helped me in completing this project.

In the first place, I would like to thank Ton de Kok for guiding me during the last 1,5 years of my master. I still remember our first meeting in which we discussed the possibilities of mentoring my graduation. In this talk, I asked him about his motivation for being active in the area of operations management. The same enthusiasm and passion as he showed in providing his answer have been a valuable source of inspiration in conducting my master thesis project.

I am honestly charmed by the way Hilti welcomes and integrates new employees in the community and additionally provides them with all means to achieve valuable results. I think everybody is jointly doing an amazing job in creating a warm but challenging working environment. To all I came across during my research I hereby would like to show my gratitude.

I am grateful to Roeland Baaijens and Rüdiger Kübler for providing me the opportunity to participate and contribute to a highly interesting and challenging real-life project. I owe special thanks to Alessandro Sasso. His way of guiding me the past months has enabled me, challenged me and motivated me to fulfill my potential. I admire the way he manages to deal with the responsibility of operationally steering the HIP project.

I will happily remember my time operating in the GLMM-team thanks to the interesting conversations and great support of Bernd Wohlgenannt, Monica Benko, Alessandra Ruzzi and Erwin Schütz.

A simple ‘thank you’ would not suffice to show my appreciation for the (mental) support and love of my family, and in particular my parents. The past few years they so often enabled and supported me in every possible way to explore ‘the world’. I am blessed to always be able to rely on them.

Lars Martens
August 2014
Management summary

Hilti is currently integrating its internal end-to-end supply chain with the objective of improving the overall performance with respect to inventory levels, operational costs and customer service. In this report, scientifically grounded guidelines are provided on redesigning the demand planning decision function in order to improve its contribution on the tactical and operational level of the planning hierarchy to the overall supply chain performance. Through reducing demand uncertainty by providing accurate demand forecast information at different levels of aggregation, the overall supply chain performance can be increased.

According to the methodology of Interactive Planning (Ackoff, 1974), this project identified the gaps between the current and ideal state of demand planning. Subsequently, design solutions have been provided. We will highlight those findings per design question.

How to embed demand planning in an integrated planning concept?

Broft (2014) and Kreuwels (2014) assessed that the current planning concept at Hilti lacks a tactical level at which supply and demand should be centrally aligned for a horizon of 15-18 months at an aggregate level. Demand forecasts are currently used directly as an input for an item-specific net requirements calculation, based upon which planning activities at the operational level are executed. The works of Broft (2014) and Kreuwels (2014) provide recommendations on the setup of a new integrated planning concept, which includes the tactical level in the planning hierarchy.

In this project, it has been assessed how the planning activities at both the tactical and operational level should be supported by demand forecast information. In table 0.1, the forecast requirements per phase of the supply chain are concisely presented for both planning levels. Plans on the tactical and operational level are suggested to be updated respectively monthly and weekly.

Table 0.1 – Demand forecast requirements per hierarchical level

<table>
<thead>
<tr>
<th>Hierarchical level</th>
<th>Type of planning decision</th>
<th>Level of forecasts required</th>
<th>Horizon of forecasts required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tactical level</td>
<td>Procurement and Production</td>
<td>Agg. prod. /month</td>
<td>≥ 12 months</td>
</tr>
<tr>
<td></td>
<td>Distribution</td>
<td>Total volume/</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>location/month</td>
<td></td>
</tr>
<tr>
<td>Tactical parameter - SS</td>
<td>Item/(location)/month</td>
<td>3 months</td>
<td></td>
</tr>
<tr>
<td>Operational level</td>
<td>Procurement and Production</td>
<td>Item/week</td>
<td>Longest cum. lead time</td>
</tr>
<tr>
<td></td>
<td>Distribution</td>
<td>Item/location/week</td>
<td></td>
</tr>
</tbody>
</table>

How to provide demand forecast information at different levels of aggregation?

Table 0.1 showed the need for demand forecast information at various levels of aggregation. Literature suggests the use of Hierarchical Demand Planning (HDP), which involves Hierarchical Forecasting (HF) to provide forecast information at different levels of aggregation. HF comprises the application of statistical forecast techniques at pre-defined levels of aggregation and the use of various (dis)aggregation methods to obtain information at other levels. Scientific literature distinguishes the Bottom-Up approach (applying statistical forecast techniques on the low level and consolidating the forecasts to obtain information at higher levels) and the Top-Down approach (applying statistical forecast techniques on the high level and disaggregating forecasts according to
pre-defined rules to obtain information at lower levels). Since literature does not provide any general guidelines for the selection process of the type of approach to be used, no scientifically grounded recommendations can be provided on the selection process of the level of statistical forecasting. The lack of knowledge on the potential accuracy damage of selecting the incorrect forecast approach impedes us from conducting a cost-benefit analysis.

A quantitative analysis based on sales data of Hilti, provided the following insights on the debate between the Top-Down approach and direct disaggregate forecasts to optimize forecast accuracy at the disaggregate level:

Hilti

- *It seems beneficial to be aware of the optimal HF-approach per cluster in advance* – given the specifications of the test, overall forecast accuracy improvement of using the optimal approach per cluster is 6.3%-8.9% compared to using the same HF-approach for all clusters
- *The optimal HF-approach per cluster could be unstable over time* – given the specifications of the test, the optimal HF-approach changed for 46% of the clusters from 2012 to 2013

**Experiment with software assistance to select an optimal forecast level per cluster in advance, but do this with caution since the optimal level could be unstable over time!**

Research Field on Hierarchical Forecasting

- *Given the use of the specifications of the analysis, the TD-approach slightly outperformed the direct disaggregate approach in 2013 while the performances were fairly equal in 2012*
- *In line with existing theory, the TD-approach seems to be more effective for clusters with positively correlated subaggregate variables compared to clusters with uncorrelated subaggregate variables.*

**How to setup the demand planning process to efficiently generate accurate demand plans?**

Statistical forecasting is currently a mixed responsibility of Materials Management in the logistic regions and in the corporate headquarters. The gap-analysis showed that the corporate process is ineffective due to the facts that it is based on replenishment history, history correction is poorly executed, events are improperly integrated, forecasts review is poorly executed and no accuracy measurement is performed. In addition, the process is said to be inefficient due to its too high frequency of execution. Since the main arguments in favor of the corporate forecasting process do not hold either, it is recommended to assign the full responsibility of forecasting items replenished by the logistic regions to the regions themselves. Items for non-integrated markets which are directly replenished by HAG are still forecasted by HAG MM.

Next to the shift of forecast responsibility, the following recommendations are provided:

1. *Review the statistical and consensus forecasting processes based on the idealized design with respect to sales outlier identification, forecast model and parameter selection/modification and the setup of consensus meetings*
2. *Safety stock and replenishment methods (‘MRP-types’) should be regarded as tactical parameters which are updated less frequent (e.g. quarterly) than what is currently done*
3. *A Supply Chain Specialist Team (SCST) should be installed to centralize the responsibility of the alignment of demand and supply*
4. *Forecast accuracy should be measured for all individual activities that influence the height of demand forecasts to objectively identify the causes of inaccuracy. Measurement is recommended to be done with the wMAPE-metric, as explained in the next paragraph.*
Demand planning should be controlled in order to:

1. Improve the performance of demand planning
2. Quantify the impacts of demand uncertainty to support supply planning decisions

The information feedback loops corresponding to both objectives are depicted in figure 0.1.

**Objective 1**

- Use the wMAPE-metric (weighted mean absolute percentage forecast error)
- Apply the metric at the basic level at which forecasts are provided to subsequent planning activities
- Define target groups based on the coefficient of variance and volume of demand
- Proportionally average APE-values based on the importance of products
- Set different accuracy targets per market due to different lag-values

**Objective 2**

- Use the MPE-metric (mean percentage forecast error) to support decisions on resource allocation
- Apply the metric at the level at which main decisions on resource allocation are taken and track values over time to support the decisions

**Recommendations**

For the further improvement of demand planning, Hilti is recommended to focus on:

1. The acceptance of main stakeholders on the proposed redesign
2. The potential benefits of software assistance in determining the best level of forecasting
3. The potential benefits of an automatic forecast model selection function
4. The use of forecast accuracy information to support decisions on inventory investments

The research field of Hierarchical Forecasting is recommended to further examine:

1. The effectiveness of TD-approach and the direct disaggregate approach under various forecast models and performance measures
2. The influence of homogeneity and volatility of subaggregate variables on the effectiveness of HF-approaches

**How to setup a control-mechanism for demand planning?**
# Table of Contents

## Chapter 1  Introduction
1.1 Project approach 1  
1.2 Company description 1  
1.3 Reason for action 2  
1.4 Stakeholder analysis 2  
1.5 Methodology 3  

## Chapter 2  AS-IS Analysis
2.1 Introduction 5  
2.2 Supply, product and market characteristics 5  
2.3 The demand planning process 7  
2.4 The use of the output of demand planning 12  
2.5 Conclusion 13  

## Chapter 3  Idealized design
3.1 Introduction 15  
3.2 Demand Planning Framework 15  
3.3 Idealized design - Workflow 15  
3.4 Idealized design - Structure 16  
3.5 Idealized design - Process 18  
3.6 Idealized design - Control 21  

## Chapter 4  Gap-analysis
4.1 Introduction 25  
4.2 Gap-analysis - Workflow 25  
4.3 Gap-analysis - Structure 25  
4.4 Gap-analysis - Process 26  
4.5 Gap-analysis - Control 28  
4.6 Problem definition 30  

Chapter 5  TO-BE design
  5.1  Introduction 33
  5.2  Workflow 33
  5.3  Structure 33
  5.4  Process & Control 45
  5.5  Control 49

Chapter 6  Recommendations on implementation 55
  6.1  Overall recommendations on implementation 55
  6.2  Recommendations regarding demand planning structure 55
  6.3  Recommendations regarding demand planning process 56
  6.4  Recommendations regarding demand planning control 56

Chapter 7  Conclusions and Recommendations 57
  7.1  Answering the design questions 57
  7.2  Recommendations 59

References 61

Appendix A 63

Appendix B 66

Appendix C 68

Appendix D 69
## List of Figures

**Figure 0.1** – Feedback loops control objectives  

**Figure 1.1** – Overview main types of stakeholders  
**Figure 1.2** – Project structure

**Figure 2.1** – Simplified Supply Chain Hilti AG (excluding spare parts)  
**Figure 2.2** – Product hierarchy  
**Figure 2.3** – Market hierarchy  
**Figure 2.4** – AS-IS monthly demand planning process  
**Figure 2.5** – Idef0-element

**Figure 3.1** – DP Framework  
**Figure 3.2** – HF-debates in literature (a circle represents forecast variable i)  
**Figure 3.3** – Idealized design monthly demand planning process  
**Figure 3.4** – Demand Planning Control

**Figure 4.1** – Cause-and-effect diagram demand planning

**Figure 5.1** – Structure matrix  
**Figure 5.2** – Use of sample data  
**Figure 5.3** – Distribution of clusters based on correlation and average volatility, 2013  
**Figure 5.4** – Share of TD2-clusters along correlation, 2013  
**Figure 5.5** – Share of TD2-clusters along average CoV, 2013

**Figure A.1** – Stakeholder analysis phase 2

**Figure B.1** – Item classification  
**Figure B.2** – Demand segmentation matrix

**Figure C.1** – Ideal design Hierarchical Planning Framework by Broft (2014) and Kreuwels (2014)
Figure D.1 – Distribution of clusters based on correlation and average volatility, 2012  
Figure D.2 – Share of TD2-clusters per range of correlation, 2012  
Figure D.3 – Share of TD2-clusters per range of average CoV, 2012  
Figure D.4 – Metric analysis control objective 1
**List of Tables**

**Table 0.1** – Demand forecast requirements per hierarchical level  

**Table 2.1** – Overview forecast workload  
**Table 2.2** – Overview unofficial forecast accuracy metrics  
**Table 2.3** – Overview of operational decision functions that are based on NRC

**Table 3.1** - Overview forecast requirements for planning decisions  
**Table 3.2** – Overview arguments HF-approaches  
**Table 3.3** – Forecast accuracy metric characteristics  
**Table 3.4** – Soft-characteristics metric objective 1  
**Table 3.5** – Soft-characteristics metric objective 2 (resource allocation)

**Table 4.1** – Overview HF-incentives at Hilti  
**Table 4.2** – Overview gaps in statistical forecasting  
**Table 4.3** – Gap-analysis metrics objective 1

**Table 5.1** – Model identification rules (adopted from Gardner and McKenzie, 1988)  
**Table 5.2** – Smoothing constant settings for grid search  
**Table 5.3** – Overall results 2013  
**Table 5.4** – Combined influence correlation and volatility, 2013  
**Table 5.5** – Overall results 2012  
**Table 5.6** – Stability of preferred forecast approach from 2012 to 2013  
**Table 5.7** – Scenario analysis 2013  
**Table 5.8** – Example of functionality of forecast accuracy metric objective 2

**Table 6.1** – Frequently overheard assumptions at Hilti

**Table A.1** – Stakeholder analysis phase 1

**Table B.1** – List of available statistical forecasting techniques

**Table D.1** – Combined influence correlation and volatility (2012)
Table D.2 – Influence volatility on performance of approaches for uncorrelated clusters given $P_x, P_y \geq 0.3$

Table D.3 – Influence volatility on performance of approaches for uncorrelated clusters given $P_x$ or $P_y < 0.3$
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>APA</td>
<td>Absolute Percentage Accuracy</td>
</tr>
<tr>
<td>APE</td>
<td>Absolute Percentage Error</td>
</tr>
<tr>
<td>APO</td>
<td>Advanced Planner and Optimizer (planning software in SAP)</td>
</tr>
<tr>
<td>BA</td>
<td>Business Area</td>
</tr>
<tr>
<td>BOM</td>
<td>Bill of Materials</td>
</tr>
<tr>
<td>BU</td>
<td>Business Unit</td>
</tr>
<tr>
<td>BU-approach</td>
<td>Bottom-Up approach</td>
</tr>
<tr>
<td>BW</td>
<td>Business Warehouse</td>
</tr>
<tr>
<td>CF</td>
<td>Consensus Forecast(ing)</td>
</tr>
<tr>
<td>CoV</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td>CW</td>
<td>Central Warehouse</td>
</tr>
<tr>
<td>DC</td>
<td>Distribution Center</td>
</tr>
<tr>
<td>DFU</td>
<td>Demand Forecast Unit</td>
</tr>
<tr>
<td>DP</td>
<td>Demand Planning</td>
</tr>
<tr>
<td>GLM</td>
<td>Global Logistics Materials Management</td>
</tr>
<tr>
<td>GLMM</td>
<td>Global Logistics Materials Management Market/Regions</td>
</tr>
<tr>
<td>HAG</td>
<td>Hilti AG (used to indicate headquarters)</td>
</tr>
<tr>
<td>HC</td>
<td>Hilti Center</td>
</tr>
<tr>
<td>HDP</td>
<td>Hierarchical Demand Planning</td>
</tr>
<tr>
<td>HF</td>
<td>Hierarchical Forecasting</td>
</tr>
<tr>
<td>HIPP</td>
<td>Hilti Integrated Planning Project</td>
</tr>
<tr>
<td>HNA</td>
<td>Hilti North America</td>
</tr>
<tr>
<td>INP</td>
<td>Introduction New Product</td>
</tr>
<tr>
<td>JF</td>
<td>Judgmental Forecast(ing)</td>
</tr>
<tr>
<td>LEC</td>
<td>Logistics Europe Central</td>
</tr>
<tr>
<td>LESE</td>
<td>Logistics Europe South East</td>
</tr>
<tr>
<td>MAD</td>
<td>Mean Absolute Deviation</td>
</tr>
<tr>
<td>(w)MAPA</td>
<td>(weighted) Mean Absolute Percentage Accuracy</td>
</tr>
<tr>
<td>(w)MAPE</td>
<td>(weighted) Mean Absolute Percentage Error</td>
</tr>
<tr>
<td>MASE</td>
<td>Mean Absolute Scaled Error</td>
</tr>
<tr>
<td>MM</td>
<td>Materials Management</td>
</tr>
<tr>
<td>MO</td>
<td>Market Organization</td>
</tr>
<tr>
<td>MPE</td>
<td>Mean Percentage Error</td>
</tr>
<tr>
<td>MPS</td>
<td>Master Production Schedule</td>
</tr>
<tr>
<td>MRP</td>
<td>Material Requirements Planning</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Squared Error</td>
</tr>
<tr>
<td>NDC</td>
<td>National Distribution Center</td>
</tr>
<tr>
<td>NRC</td>
<td>Net Requirements Calculation</td>
</tr>
<tr>
<td>RDC</td>
<td>Regional Distribution Center</td>
</tr>
<tr>
<td>RF</td>
<td>Rolling Forecasts (sales)</td>
</tr>
<tr>
<td>ROP</td>
<td>Reorder Point</td>
</tr>
<tr>
<td>SCOP</td>
<td>Supply Chain Operations Planning</td>
</tr>
<tr>
<td>SCST</td>
<td>Supply Chain Specialist Team</td>
</tr>
<tr>
<td>SF</td>
<td>Statistical Forecast(ing)</td>
</tr>
<tr>
<td>SFI</td>
<td>Sales Forecasting Integration</td>
</tr>
<tr>
<td>SP</td>
<td>Sales Planning</td>
</tr>
<tr>
<td>TD-approach</td>
<td>Top-Down approach</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

In this section we will discuss the project approach and provide a concise introduction of the company at which the project is executed. In addition, we systematically select and discuss the methodology according to which the project is structured.

1.1 Project approach

Operations management is about designing, implementing and controlling supply chain operations processes in manufacturing and service industries. Bertrand and Fransoo (2002) explain that scientific management can be considered as the root of the development of (quantitative) operations management. Scientific management came up in the beginning of the 20th century and proclaimed a systematic working method to enhance the performance of operational processes. This involved the application of analytic techniques which revealed and quantified the aspects that affect this performance. Due to fact that scientific management did not produce any generic scientific knowledge, it cannot be regarded as a science itself. However, its approach still reminds us of the final objective of the scientific research area (operations management) it has resulted in: the performance improvement of operational processes.

This graduation project will be executed in line with the original essence of scientific management in the sense that the primary objective is to analyze a real-life operational process and redesign it in order to improve its performance. The focus lies more on the application of available scientific models and theories on operational processes rather than to produce generic knowledge. However, in situations where scientific models or theories are missing or insufficiently support certain design decisions, quantitative and qualitative analyses are conducted in order to scientifically ground design decisions and contribute to the generic scientific knowledge available.

In this chapter we will start with the introduction of the company Hilti AG at which this project is executed. Second, the reason for action is discussed which provides further insights regarding the system under consideration. Subsequently, a general stakeholder analysis is conducted in order to properly select the methodology for the execution of the project.

1.2 Company description

<table>
<thead>
<tr>
<th>Company name</th>
<th>Hilti AG (Hilti Aktiengesellschaft, Hilti Group)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Headquarters</td>
<td>Schaan, Liechtenstein</td>
</tr>
<tr>
<td>Industry</td>
<td>Manufacturing – Professional construction tools/systems</td>
</tr>
<tr>
<td>CEO</td>
<td>Dr. Christoph Loos</td>
</tr>
<tr>
<td>Net sales (2013)</td>
<td>4340 CHF million</td>
</tr>
<tr>
<td>Net income (2013)</td>
<td>304 CHF million</td>
</tr>
<tr>
<td>Nr. of employees</td>
<td>21,456 (31st December 2013)</td>
</tr>
</tbody>
</table>

In 1941, the brothers Martin and Eugene Hilti set up a mechanical workshop where they manufactured construction tools on a small scale. Starting as a machine shop in a garage in Schaan, their enterprise gradually expanded internationally promoting their own product designs. Nowadays, this company is known as Hilti AG, providing full system solutions for the professional construction industry. Hilti’s characterizing red-and-black colored assortment includes amongst others drilling and demolition products, screw fastening systems, anchoring systems and fire stop systems. Although the majority of their products is still being sold in (Western-) Europe, sales are
expanding globally. This is mainly visible in the American (21%) and Asian (12%) markets. In addition to their tangible output, Hilti provides consultancy services for major construction projects all over the world. Current research and development practices aim at improving construction industry product standards in order to make buildings more resistant to extreme environmental conditions.

Hilti operates in the spirit of ‘passionately creating enthusiastic customers and building a better future’. In March 2014, Hilti presented the ‘Champion 2020 Strategy’. The strategic objective of the new Champion 2020 formula is ‘sustainable value creation through market leadership and differentiation’. In this, market leadership refers to the company’s relative market share in key markets. Differentiation reflects the intention to (exclusively) focus on those products which are of high value to their customers and those markets in which Hilti regards it likely to achieve and sustain leadership positions.

1.3 Reason for action

In the beginning of 2013, Global Logistics Materials Management (GLM) launched the Hilti Integrated Planning Project (HIPP) in order to create an integrated planning concept for their end-to-end internal supply chain. The final objective is to reduce the high levels of stock and operational costs and to improve customer service. Initial analyses, conducted in the fall of 2013 by a team of students of the University of Eindhoven, indicated the lack of a tactical level in the operational planning activities (Broft, 2014; Kreuwels, 2014). In scientific literature, this level is suggested to contain aggregated planning activities that focus on a mid-term horizon of 6 to 24 months. Broft (2014) and Kreuwels (2014) showed that demand forecasts are currently, without any tactical intervention, directly used as an input to calculate the item-specific net requirements. Recommendations provided by Broft (2014) and Kreuwels (2014) focused on the integration of a tactical level in the hierarchical planning structure at which Supply Chain Based Sales and Operations Planning is performed. Four follow-up projects have been assessed to structurally redesign the planning concept at Hilti:

1. The Workflow-project aims to integrate and synchronize all demand, production and distribution planning activities in the end-to-end supply chain.
2. The Sales Planning-project intends to redesign the demand planning process as such that it ensures the efficient generation of accurate demand plans to be used in an integrated planning process.
3. The Production Planning-project should result in the implementation of the process that centralizes production planning activities on the tactical and operational level of the planning hierarchy.
4. The Organization-workstream should install a central team which is responsible for the alignment between demand, production and distribution planning.

This master thesis project contributes to the Sales planning-workstream by providing a scientific perspective on the redesign of the demand planning process. The initial research proposal (Martens, 2014a) confirmed the need of redesigning this process through extensive research in order to increase the overall supply chain performance in an integrated planning environment. The master thesis project is conducted at the Global Logistics Materials Management Market/Regions department (GLMM), located in Nendeln, Liechtenstein.

1.4 Stakeholder analysis

Jackson (2003) argues that the selection of an appropriate project approach depends on the complexity of the problem and the characteristics of the organizational context. Organizational
contexts can be labeled as either unitary, pluralist or coercive, referring to the alignment of interests of the most important stakeholders. It is of paramount importance to be aware of the problem context, since a successful implementation of projects is to great extent dependent on the compliance of the main stakeholders (as shown by Nutt, 2002).

In order to classify the organizational context for the current project, a concise stakeholder analysis is conducted. In broad terms, a stakeholder is defined as any individual or group that can affect or be affected by the outcomes of the project.

In the first phase of the analysis, stakeholders are identified and their interests are revealed. An overview of the stakeholders is presented in figure 1.1, the detailed results of the first phase are listed in table A1 in Appendix A. Eindhoven University of Technology is included as a stakeholder since it poses a number of requirements in order for the project to satisfy the conditions of a master thesis project for the degree of MSc Operations Management and Logistics.

![Figure 1.1 – Overview main types of stakeholders](image)

The goal of the second phase of the analysis is to categorize the stakeholders based on their power and level of interest in the project. In this, we regard power as the extent to which stakeholders are able to persuade or demand others into following certain courses of action. For the categorization we have made use of the power/level of interest matrix as provided by Gardner, Rachlin, and Sweeny (1986). The results are presented in figure A1 in Appendix A.

The results of the second phase show that all stakeholders have different interests. However, the interests of the highly interested and powerful stakeholders are mostly not coercive. The only situation of opposing interests is observed between the desire of Materials Management and Marketing regionally and globally to have flexibility in the input of customer demand and the request of production plants to have stable production plans. The early awareness of the existence of coercive interests eases the definition and integration of the correct trade-off between these conflicting objectives in the design process.

### 1.5 Methodology

In the previous paragraph we have noticed that the project involves a variety of stakeholders with different interests. Since these interests are mostly not coercive of nature, we classify the problem context as ’pluralist’. We label the problem itself as ’complex’, due the wide product portfolio of Hilti...
and the large variety of existing procedures. In situations of complex problems in a pluralist context, Jackson (2003) recommends the use of either *Strategic Assumption Surfacing and Testing*, *Interactive Planning* or *Soft Systems Methodology*.

The structure of this project (figure 1.2) is based upon *Interactive planning*. This methodology is invented and elaborated in detail by Ackoff (1974, 1981). *Interactive planning* is a methodology that focusses on tackling the complexity of problems while acknowledging their societal context. It is based on the principles of stakeholder participation, sustainability of research design and holistic problem solving. This methodology is regarded as appropriate since it does approach the research design from a holistic perspective and constructively involves important stakeholders. Additionally, the fact that Hilti is already familiar with the research structure through the initial research by Broft (2014) and Kreuwels (2014) contributes to the ease of use of the methodology.

![Figure 1.2 - Project structure](image)

For the redesign of a process, it is of paramount importance to be aware of the current state of the process. In the AS-IS analysis it is therefore aimed to obtain a clear understanding of the way demand planning is currently performed at Hilti. The second phase entails the creation of the ideal state of the process as if it were designed from scratch. The idealized design will be predominantly based on existing scientific models and theories in the area of demand planning. Subsequently, an analysis of the gaps between this idealized design and the AS-IS state of the process is conducted. Based on the gaps identified, design objectives and questions are formulated and the scope of the project is defined. In the TO-BE design, it is aimed to (partially) close the gaps between the desired and current state of the process. The last phase will provide concise recommendations on the implementation of the TO-BE design.
Chapter 2

AS-IS Analysis

By means of conducting an AS-IS analysis, it is intended to obtain an accurate description of the current state of a well-defined factual (sub-)system. A proper understanding of this current state supports the identification of issues impeding ideal performance and the creation of suitable solution designs.

2.1 Introduction

The ‘Reason for Action’ in the previous chapter revealed the sub-system under consideration: the demand planning process at Hilti. We define demand planning as the entire process of the generation of demand forecasts and the creation of demand plans based on these forecasts. Demand forecasts support in reducing demand uncertainty. Reduction of this demand uncertainty contributes to lowering inventory levels, operational costs and increasing customer service levels. In make-to-stock environments, demand forecasts are typically used as production triggers.

In this chapter we intend to provide an accurate description of the factual state of demand planning at the tactical and operational level of the planning hierarchy. The term ‘factual’ refers to the co-existence of two realities: a formal reality and an informal reality. The formal reality of a process is the set of activities in the process as it is officially prescribed to be executed. The informal reality refers to the set of all unofficial regulations and practices. With the aim of describing the factual state of the process, we intend to fully capture the informal reality and only integrate that part of the formal reality which overlaps with the informal reality.

In order to understand the way demand planning is currently performed at Hilti, we will start with describing the product and market characteristics that define the demand planning context. Subsequently, we start unravelling the separate demand planning operations in further detail.

2.2 Supply, product and market characteristics

Demand planning is a fairly easy practice for the supply of a single item, for a single customer given a certain period of time. The complexity of demand planning rises along with the size of the product portfolio and customer base. To obtain an understanding of the complexity of the demand planning process at Hilti, we will discuss the product and market characteristics that influence the execution of demand planning. We will however start with describing the simplified supply chain of Hilti to gain insights on its structure and the various organizational entities involved in the supply of products.

![Figure 2.1 – Simplified Supply Chain Hilti AG (excluding spare parts)](image-url)
Hilti supplies her products through an internal end-to-end supply chain (figure 2.1). This implies that it owns and controls the majority of operations in the supply chain. External parties exclusively contribute in the supply of components/final products and their distribution.

Globally, Hilti controls 8 production plants which are responsible for the in-house production. Five of these plants are located in Europe, two in China and one in Mexico. Three of the eight plants cover the production of Power Tools, the remaining plants are responsible for the supply of Consumables.

Next to the in-house production, Hilti cooperates with approximately 900 allied suppliers which cover the total out-house production. Small material flows and materials flows that need to be shipped to locations with a long lead time are replenished from the production plants or allied suppliers via HAG Warehouses to central warehouses (CW), national (NDC) or regional distribution centers (RDC) in the logistic regions. All three HAG consolidation points are located in Central Europe. For large material flows and material flows to a short lead time location, the local warehouses can be replenished directly by the production plants or allied suppliers. The logistic regions replenish retail through either a single-tier or a two-tier distribution structure.

Hilti reaches its customers via either Hilti Centers (HC) or via specialized Hilti Vans which provide additional service in the form of visiting the customer and providing instructions on the usage of the products. In case of large sales orders, customers can also be served directly by RDCs.

**Product characteristics** - The described supply chain distributes a large variety of products with different production and sales characteristics to customers all over the world. In order to get an understanding of the complexity this causes, we discuss the various product and customer types at Hilti.

The formal product hierarchy of Hilti is depicted in figure 2.2. Products are divided over three different Business Areas (BA): BA Electric Tools & Accessories (ET&A), BA Fastening and Protection Systems (F&P) and BA Energy and Industry (E&I). The final amount of different items (active sold items excluding spare parts) is approximately 31,000. Only 7% of these unique end-items generates 80% of the turnover.

All these products have different lead times, sales volumes and storage, handling and transportation requirements. To be able to manage all these different product appropriately, the products are segmented based on their sales value and customer order frequency. Products with a label of T, A, B, C or D respectively compose 50%, 30%, 15%, 4% and 1% of total sales. Item ranked as X, Y or Z are respectively fast-moving, sporadic and non-moving. The classification matrix is presented in figure B.1, Appendix B.

The described classification serves amongst others in prioritizing products for resource allocation decisions in demand forecasting. Generally, the T,A,B,C,D-label of a product is regarded to reflect the importance of the product for the company. The segmentation matrix supports as well in setting forecast accuracy targets for a cluster of items.

In order to correctly apply replenishment methods and make stock positioning decisions, items are clustered based on the coefficient of variance of the order sizes and the order frequency (figure B.2 in Appendix B.). Based on this segmentation, products are either labeled as CG1 (normal demand), CG2 (variable demand) or CG3 (sporadic demand).
**Market characteristics** - In figure 2.3 the market hierarchy is depicted. The global market is divided into logistic and sales regions. Logistic regions reflect the division of the global market from a supply perspective: every region has its own CWs/NDCs/RDCs and is responsible for the replenishment of the Hilti Centers.

Sales regions reflect the division of the global market from a marketing/sales perspective. Financial performance indicators are calculated per sales region. The definition of logistic regions and sales regions mostly overlaps, but does not have to be equal necessarily.

Every logistic region consists of one or multiple demand groups. A demand group is either equal to a country-specific market or represents a set of those markets. Market organizations (MOs) are country-specific departments which are responsible for the retail within their sales area. Countries with a market organization are referred to as integrated, while countries without an MO and without fixed distribution channels are called non-integrated. In the latter category, sales volumes are relatively low. Products are sold via agents and dealers.

### 2.3 The demand planning process

The previous section has provided us with an idea of the size and complexity of the context of demand planning. In this section we will discuss the activities in the specific phases of the demand planning process, the departments involved in this process and the way these interact.

Demand planning provides demand forecasts related to a product, time and customer. Literature generally distinguishes three types of processes to provide those demand forecasts:

1. **Statistical forecasting** (SF) – Demand forecasts are acquired through the application of statistical forecast techniques on historical sales data.
2. **Judgmental forecasting** (JF) – Human beings generate demand forecasts based on their experience and in-depth knowledge of the market.
3. **Consensus forecasting** (CF) – This process aligns the statistical and judgmental forecasts to create one single demand forecast value to be used for further planning operations.

The aforementioned types of forecasting processes are in Hilti executed on both a global and a regional level. Figure 2.4 provides a process overview of the generation and alignment of the statistical and judgmental forecasts on both levels. The diagram is based on the findings of analyses of the forecasting processes regionally in LEC (Logistics Europe Central), HNA (Hilti North America), LESE (Logistics Europe South-East) and globally in BUs PT&A (Power Tools & Accessories) and DF (Direct Fastening). This subset of organizations is argued to be representative for the entire set of forecasting organizations since it includes regions representing long and short lead time markets, big and small markets (related to demand group) and BUs representing the products of all BAs. It is recommended to read to content of the explanation box on page 9 carefully to properly interpret the figure 2.4.
Conceptual Demand Planning Process // AS-IS

Marketing - HQ

MM - HQ

Week 4

Week 2

Week 2

Week 4

Corporate Forecast Review

APO DP Run

Final HAG SF

Hist replenishment data

AS - IS monthly demand planning process

Figure 2.4 – AS-IS monthly demand planning process
REDESIGNING A DEMAND PLANNING PROCESS IN AN INTEGRATED PLANNING ENVIRONMENT

Legend figure 2.4

Process modeling language - The main process modeling language used for constructing the diagram in figure 2.4 is *idef0*. The most important element of this language is the process-symbol depicted in figure 2.5. Demand planning concerns many data transformation activities. In the diagram we will use the input arrow to indicate a dataset that is different each time the process is executed. We will use the control arrow when we want to represent data upon which the process execution is based.

IT landscape - At Hilti, final forecast data is entered into the Advanced Planner and Optimizer (APO), add-in of SAP (ERP-system). The PP/DS function in APO translates the final forecasts into requirements according to MRP-logic. Once the planning is confirmed, actual transactions are executed in the R3-system within SAP. From here, the sales information and master data are stored in the Business Warehouse (BW).

In the following paragraphs we will describe how the organizations execute the different process phases in further detail. We will start with discussing the regional demand planning process (swimminglanes 3 and 4 in figure 2.4), after which the global process will be described (swimminglanes 1 and 2 in figure 2.4)

2.3.1 The demand planning process at regional level

Statistical forecasting (lane 4, figure 2.4) - Materials Management (MM) in the logistic regions generates statistical forecasts on a monthly basis for products with 'normal' demand (high volume and order frequency) and a long lead time from the HAG warehouses. These forecasts are based on the sales history and have a horizon of 18 months. Forecasts are provided per Demand Forecast Unit (DFU). A DFU is at Hilti defined as an item, related to the local warehouse it is distributed from and the demand group it is destined for.

In the initial process phase (History Adjustment) non-recurrent outliers in the historical demand patterns are corrected. In general, this is done for all T- and A-items individually. The B-, C- and D-items are reviewed through alert management. Two alerts are used which display whether the forecast of the previous month was either too high or too low compared to the actual sales.

In the first whole weekend of the month an APO DP run calculates statistical forecasts through a predefined statistical forecast technique per active DFU. Phase-in and phase-out items are mainly forecasted manually. For phase-in products, INP-plans (Introduction New Product) are created which contain the forecasted figures per item per demand group for the first year in monthly buckets. After the INP-plan of a product has ended, Materials Managers select an appropriate statistical forecast technique and parameters based on the course of the available sales history and their knowledge and experience. The list of available forecast techniques is presented in Appendix B (table B.1).

After the APO DP run, Materials Managers review the quality of generated statistical forecasts preventively. Again, all T- and A-items are in general reviewed individually and alerts are used to review the B-, C- and D-items. Based on experience and knowledge, a Materials Manager may decide to change to forecast technique or its parameters, to manually adjust the forecast or to change the MRP-type which leads to a change of the forecast responsibility. Likewise the phase of history adjustment, the amount of DFUs is too high for Materials Managers to review all T- and A-items and alerts for B-, C- and D-items. Worldwide, Hilti performs more than 160,000 forecasts. An overview of the workload per analyzed organization is presented in table 2.1.
In April 2014, HNA has started testing with statistically forecasting consumables on a product family-level (per location per month). Forecasts are disaggregated to a DFU-level based on the proportions of separate low level forecasts. Tools are still forecasted on a DFU-level, since the demand patterns within the families are not regarded to be similar (which is regarded to be the case for consumables).

In order to reduce the forecast workload, LEC started to jointly forecast the demand groups MO DE and MO NL from June onwards. LESE performs only one forecast per item for the products replenished by the warehouse in Sofia (Bulgaria). This warehouse replenishes the majority of the countries in Eastern Europe.

**Judgmental forecasting (lane 3, figure 2.4)** - The sales planning is a regional judgmental forecast provided by the regional Marketing department. Although the term sales planning (SP) is not commonly used within Hilti, this term will be used in this document to refer to these forecasts. SP is not created at a fixed point in time nor with a constant frequency. Data are provided in sales value and mostly on a product family level.

**Consensus forecasting** - In the third week of the month, the judgmental forecasts and the statistical forecasts are aligned. This is done in Sales Forecasting Integration Meetings (SFI). These meetings are generally attended by Product and Materials Managers. SFIs are used to discuss INPs, Sales Growth (Sales Planning), Event Management, Supply Issues and Phase-out Management.

For the alignment of the statistical forecasts and the sales planning, the statistical forecasts need to be expressed in sales value and aggregated to the level of the sales planning. Based on possible differences, it may be decided to manually adjust the aggregated statistical forecasts. The extent to which Marketing is involved in the disaggregation of these aggregated forecasts to DFU-level varies among regions. The frequency of SFI-meetings and the extent to which the judgmental forecasts are challenged by MM differ per region as well.

Next to the alignment on promotions planning in the SFI-meetings, promotion forecasts can be communicated and integrated into the statistical forecasts every day of the month.

**Safety stock update and Master data maintenance** - Automatically generated safety stock update proposals are used to adjust statistically calculated safety stocks in APO. In the adjustment activity in the last week of the month, manually set safety stocks are reviewed as well.

At the end of the month, master data (lead time information, obsolescence classes and MRP-types) are maintained. The ‘MRP-type’ of an item indicates whether an item is forecasted and where it is forecasted, the way the item is replenished and how the height of the safety stock is determined. Although master data maintenance is mentioned as being a monthly process, the MRP-type of a product can be changed weekly.
**Forecast accuracy measurement** – On a monthly basis, Global Controlling generates a forecast performance report which contains the forecast performance figures for every region. Calculations of these performance figures are based on the following definition of the official KPI for Forecast Quality:

\[
  \text{if } (F_{i,t-2,t} < D_{i,t}), \text{ then } FQ_{i,t} = \frac{F_{i,t-2,t}}{D_{i,t}} \times 100\% , \text{ otherwise } FQ_{i,t} = \frac{D_{i,t}}{F_{i,t-2,t}} \times 100\% \quad (2.1)
\]

where:
- \( F_{i,t-x,t} \) Forecast of DFU \( i \) made at time \( t-x \), for time \( t \)
- \( D_{i,t} \) Actual demand of DFU \( i \) at time \( t \)

Formula 2.1 shows that forecast quality depends on the ratio of the actual demand of a specific month and demand forecasts for that month, which are generated two months prior to that month (i.e. the lag is two months). The official targets per item type set the lower bounds of forecast quality for T- and A-items at 70%, for B-items at 65% and for the average of all items at 60%.

Out of discontent with the formal accuracy measurement methodology, several regions have created their own accuracy measurement methodologies. An overview of the metrics in the analyzed regions is provided in table 2.2. It should be noted that the metric definition of HNA MM is based on demand and forecast value instead of demand and forecast volume.

**Table 2.2 – Overview unofficial forecast accuracy metrics**

<table>
<thead>
<tr>
<th>Forecast Org.</th>
<th>HNA MM</th>
<th>LEC MM</th>
<th>LESE MM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Metric name</strong></td>
<td>FA-monetized</td>
<td>FQ-aggregated</td>
<td>Official FQ</td>
</tr>
<tr>
<td><strong>Metric Definition</strong></td>
<td>( \left( 1 - \frac{\sum_{i=1}^{N} \sum_{t=1}^{T}</td>
<td>F_{i,t-2,t} - D_{i,t}</td>
<td>}{\sum_{i=1}^{N} \sum_{t=1}^{T} F_{i,t-2,t}} \right) \times 100% )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level</th>
<th>DFU/month</th>
<th>Item/location/month</th>
<th>DFU/month</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Based on the metrics in table 2.2, the performance of demand planners is discussed. Regionally set performance targets may differ from the official targets.

In addition to the aforementioned official and unofficial forecast performance metrics, the Mean Absolute Deviation (MAD) is measured over the previous 3 or 12 months for the calculation-method of several safety stock types. The MAD is however not used for any sort of forecast performance review.

**2.3.2 The demand planning process at global level**

**Statistical forecasting (lane 1, figure 2.4)** - HAG MM (MM in Headquarters) performs statistical forecasts per BU for:

1. *ROP-items* - items replenished via reorder point-method with a short lead time to the RDC
2. *Sporadic items* - items with low sales volume and order frequency
3. *Demand in non-integrated markets*

Different than the statistical forecasts by the logistic regions, the forecasts by Global Headquarters are based on the replenishment history (from HAG warehouse to RDC or non-integrated market). Forecasts are generated on a weekly basis, in monthly buckets with a horizon of 18 months. These are on DFU-level, where the location represents a HAG warehouse and the demand group is HAG.
Due to the continuously changing history (as a result of weekly changing MRP-types), it is difficult to structurally perform history adjustment and to measure forecast accuracy. Outliers in the history are intended to be eliminated by immediately correcting sales history when promotions occur. This can be done by the Materials Managers in the logistic regions or by HAG MM in case Regional MM properly forwards event information. Every Monday, the quality of the newly generated statistical forecasts is preventively checked through alert-management.

**Judgmental forecasting (lane 2, figure 2.4)** - Global Marketing provides budgetary sales forecasts per BU, which we will refer to as sales rolling forecasts (RF). Based upon the strategic policy and the knowledge and experience of the Global Product Managers, these rolling forecasts are created three times a year: in February (for the current year), in April (for the current and next year) and in October (for the next year). Data are expressed in sales value, in monthly buckets and mostly on a product family-level.

In the third week, a demand preview is created for allied suppliers to provide them with insights on the forecasted demand for the upcoming 12 months. In the same week, HAG MM of each BU creates a sales report in which the rolling forecasts are compared with the sum of the year-to-date sales and the statistical forecasts (sum of MO and HAG SF) for the rest of the year. In Global SFI-meetings it is analyzed whether the statistical forecasts deviate from the rolling forecasts and which forecast needs to be adjusted if there is a difference. In case the regional statistical forecasts need to be adjusted, Marketing HQ will contact Marketing in the different regions to agree upon the adjustment of the statistical forecasts. When Marketing Region agrees with MM Region upon the height of the forecast change, the statistical forecasts are manually adapted.

### 2.4 The use of the output of demand planning

The Net Requirements Calculation (NRC) is run according to MRP-logic in APO on a daily basis with a horizon of 5 months and on a weekly basis with a horizon of 18 months. The requirements vary frequently since they are based on the forecasts in APO which are changed from unrevised statistical forecasts into the revised statistical forecasts and lastly the consensus forecasts during one month. In order to derive the net requirements in daily buckets, the monthly forecasts in APO are equally divided over circa 20 working days in a month. The theses of Broft (2014) and Kreuwels (2014) indicate a lack of a tactical level in the hierarchical planning structure. This implies that currently there are no (or to very low extent) centralized planning functions that align demand and supply on an aggregated level. The net requirements are currently without any intervention of central planning used directly as an input for operational decision functions. In order to understand the current use of demand forecasts for decision functions on the operational level, the processes in Plant 4 (Thüringen, Austria) and Plant 6 (Kaufering, Germany) have been analyzed. An overview of the results is presented in table 2.3.

**Table 2.3 – Overview of operational decision functions that are based on NRC**

<table>
<thead>
<tr>
<th>Plant</th>
<th>Decision function</th>
<th>Length plans</th>
<th>Level of information</th>
</tr>
</thead>
<tbody>
<tr>
<td>P4</td>
<td>Procurement</td>
<td>No plan</td>
<td>Item/day</td>
</tr>
<tr>
<td></td>
<td>Personnel planning</td>
<td>No plan</td>
<td>Item/week</td>
</tr>
<tr>
<td></td>
<td>Production Planning and Scheduling</td>
<td>2 weeks</td>
<td>Item/week</td>
</tr>
<tr>
<td>P6</td>
<td>Procurement</td>
<td>No plan</td>
<td>Item/day</td>
</tr>
<tr>
<td></td>
<td>Personnel planning</td>
<td>1 week</td>
<td>Item/week</td>
</tr>
<tr>
<td></td>
<td>Production Planning and Scheduling</td>
<td>1 week</td>
<td>Item/week</td>
</tr>
</tbody>
</table>
Based on the analysis, it is concluded that there is no real alignment on demand and supply planning on the operational level. Procurement orders are directly released based on the information of the NRC (after explosion of the BOM). Suppliers are however informed in advance on the expected procurement quantities through demand previews.

P4 is rather flexible in hiring personnel: quantities can be adapted daily. Therefore, personnel planning is not or to low extent performed. P6 is restricted by the German labor law and is therefore forced to plan personnel one week in advance.

Additionally, consensus forecasts in APO are used for determining the safety stock levels of the production plants, the HAG warehouses and the warehouses in the logistic regions. The safety stock updates are done on a monthly basis. Several installed safety stock calculation types are based on the average daily forecasted demand for the next 90 days (forecast horizon of 3 months).

2.5 Conclusion

The process of demand planning should provide accurate demand plans that support in reducing demand uncertainty. Demand uncertainty negatively affects inventory levels, operational costs and customer service. The AS-IS analysis showed that demand planning at Hilti is performed at both the global and regional level. In general, forecasts are obtained through generating separate statistical and judgmental forecasts and subsequently aligning these via consensus forecasting.

Among the regional demand planning processes, various differences have been observed. The differences are related to the way the processes are executed and controlled. In all organizations, the statistical forecasting workload is perceived as high due to the high number of DFUs. Several regions are experimenting with statistically forecasting at a higher level, assuming it will decrease this workload.

The output of demand planning is directly, without any tactical intervention, used to support planning activities on the operational level.
Idealized design

The Idealized design is the ideal setup of a process as if it were constructed from scratch without any restrictions. It specifies the objectives and ideals, which provide the main direction on where to go. In this project, scientific literature is the basis for the construction of the idealized design.

3.1 Introduction

The literature review of Martens (2014b) provides a detailed overview of the existing literature on demand planning in an integrated planning environment. For the construction of the idealized design, we will only include the most relevant findings.

In order to create a valuable idealized design, three criteria should be taken into account (Jackson, 2003). The design must be:

1. Technologically feasible - The design should be able to work with currently available technological resources or it should be based on likely technological developments.
2. Operationally viable - The design should function properly if it would be implemented in the current environment
3. Capable of being continuously improved - Instead of creating a fixed design, the intention should be to construct an ideal-seeking system that adapts to a changing environment.

We will start with providing a demand planning framework, according to which we will provide further insights on the idealized design of demand planning at Hilti.

3.2 Demand Planning Framework

Kilger and Wagner (2008) provide a framework for the setup of demand planning. This framework consists of three main components: structure, process and control. The aspect of 'structure' entails the hierarchical structuring of the input and output of demand planning in the dimensions of product, customer and time. The process-component refers to the phases of the demand planning process itself. The component of 'control' relates to the definition of basic accuracy metrics and methodologies. The three components are related and should be aligned to obtain a high demand planning performance. In the upcoming paragraphs, the idealized design is discussed according to this framework. We will however start with describing how demand planning is ideally embedded in an integrated planning environment.

3.3 Idealized design - Workflow

This section will provide information on how to ideally embed the decision function of demand planning in a future integrated planning concept at Hilti. The fundamentals for this new integrated planning concept have been provided by Broft (2014) and Kreuwels (2014). Their recommended design of a hierarchical planning framework for Hilti is presented in figure C.1 (Appendix C). In this concept, demand forecasts support in reducing the demand uncertainty, which negatively affects supply chain performance. An overview of the demand forecast requirements per decision function on each level is presented in table 3.1.

Figure 3.1 – DP Framework
Table 3.1 – Overview forecast requirements for planning decisions

<table>
<thead>
<tr>
<th>Hierarchical level</th>
<th>Planning decisions</th>
<th>Level of forecasts required</th>
<th>Horizon of forecasts required</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tactical level</strong></td>
<td>Personnel Planning</td>
<td>Agg. product/month</td>
<td>≥ 12 months</td>
</tr>
<tr>
<td></td>
<td>Aggregate Material Requirements Planning</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Aggregate Production Planning</td>
<td>Agg. product/month</td>
<td>≥ 12 months</td>
</tr>
<tr>
<td></td>
<td>Capacity planning</td>
<td>Agg. product/month</td>
<td>≥ 12 months</td>
</tr>
<tr>
<td></td>
<td>Distribution planning</td>
<td>Total volume/location/month</td>
<td>≥ 12 months</td>
</tr>
<tr>
<td></td>
<td>Tactical parameter - SS</td>
<td>Item/(location)/cum. lead time</td>
<td>3-6 months</td>
</tr>
<tr>
<td><strong>Operational level</strong></td>
<td>Personnel planning</td>
<td>Item/week</td>
<td>Longest cum. lead time</td>
</tr>
<tr>
<td></td>
<td>Material Requirements Planning</td>
<td>Item/week</td>
<td>Longest cum. lead time</td>
</tr>
<tr>
<td></td>
<td>Master Production Schedule</td>
<td>Item/week</td>
<td>Longest cum. lead time</td>
</tr>
<tr>
<td></td>
<td>Transportation scheduling</td>
<td>Item/location/week</td>
<td>Longest cum. lead time</td>
</tr>
</tbody>
</table>

**Tactical level** - On this mid-term level, planning decisions are taken based on aggregate demand forecasts. The term ‘aggregate’ refers to the fact that planning decisions are based on consolidated data in the product, customer and/or time dimension(s). Silver et al. (1998) explain that the horizon of tactical plans often ranges between six months to two years. Reuter and Rohde (2004) state that an aggregate plan should consider at least one seasonal cycle in order to balance all demand peaks. The horizon is thus defined as minimally 12 months (preferably 15-18 months). Plans on the tactical level are suggested to be updated monthly. The tactical parameter setting function originates from the work of De Kok and Fransoo (2003). This function is installed to coordinate the safety stock, lead time and workload parameters of the supply chain. Broft (2014) and Kreuwels (2014) suggest updating these parameters on a quarterly basis.

**Operational level** - The Supply Chain Operations Planning function on the short-term level centralizes the control of the release of materials and resources for the production in the transformational units (TU). Based on short-term demand forecasts, a Master Production Schedule is created. This is an item-specific production plan, often expressed in weekly or even daily time buckets. Silver et al. (1998) state that the minimum planning horizon of the master production schedule should be as long as the longest of the cumulative lead times of the items in the master schedule. Plans on the operational level are suggested to be updated weekly.

In order to prevent the instability of plans, Silver et al. (1998) explain the use of time fences. These time fences mark the period in which forecasts are not integrated in the master schedule. The demand time fence (DTF) bounds the period that is specified by the transportation lead time. The planning time fence (PTF) is dependent on the production (and procurement) lead time and the DTF. By not integrating forecasts for t=0 for [-PTF, 0], one freezes a part of the master schedule.

3.4 **Idealized design - Structure**

In the previous paragraph it is mentioned that demand forecasts are required on various levels of aggregation in the dimensions of products, markets and time. In this paragraph we will discuss the most appropriate approaches to obtain this information on different hierarchical levels.

In order to satisfy the desire and need of stakeholders to make planning decisions on different levels of aggregation, Hax and Meal (1975) suggest the application of the centralized forecast approach of Hierarchical Demand Planning (HDP). HDP structures demand planning
information hierarchically in the dimensions of product, geographical location and time. HDP-systems make use of hierarchical forecasting (HF) methodologies in order to create statistical forecasts. Hierarchical forecasting entails the application of mathematical models on (a) predefined level(s) of aggregation. The forecast information for other levels of aggregation is subsequently derived through various (dis)aggregation methods.

In literature, two main HF-methodologies can be distinguished. These are the bottom-up approach (BU-approach) and the top-down approach (TD-approach). The BU-approach applies statistical forecast techniques on the lowest level of aggregation and consolidates the resulting forecasts to obtain figures on a higher level of aggregation. The TD-approach statistically forecasts aggregated time series and uses proration methods in order to provide data on the disaggregate levels. The literature on HF is divided into two debates (visually presented in figure 3.2):

Debate 1: BU-approach versus the direct aggregate approach in order to optimize accuracy at the aggregate level.

Debate 2: TD-approach versus the direct disaggregate approach in order to optimize forecast accuracy at the disaggregate level.

Intuitive arguments in favor of the BU/direct disaggregate approach and TD/direct aggregate approach are presented in table 3.2. Empirical results in scientific literature on both debates are concisely discussed subsequently.

Table 3.2 – Overview arguments HF-approaches

<table>
<thead>
<tr>
<th>Advantage(s)</th>
<th>TD or direct aggregate approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The approach has the potential to properly capture the demand patterns on the disaggregate level</td>
<td>• Through aggregation one may cancel out the random distortions in individual time series</td>
</tr>
<tr>
<td>• Highly volatile series are difficult to forecast</td>
<td>• Aggregation of time series may discard valuable pattern information of disaggregate non-stationary time series, which may make aggregate time series too complex to model and forecast</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Disadvantage(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The disaggregation process often introduces bias on the disaggregate level (only for TD-approach)</td>
</tr>
</tbody>
</table>
REDESIGNING A DEMAND PLANNING PROCESS IN AN INTEGRATED PLANNING ENVIRONMENT

**Literature on debate 1** - Results of Dunn, Williams and DeChaine (1976) and Gordon, Morris and Dangerfield (1997) rule in favor of the use of the BU-approach in order to obtain accurate forecasts at the aggregate level. Fliedner (1999) shows that a direct aggregate forecast however may outperform the BU-approach when the individual time series are highly negatively or positively correlated. Widiarta, Viswanathan and Pipili (2009) confirm that the aspect of correlation affects the accuracy of the direct aggregate forecasts, but argue that the effectiveness is also related to the autocorrelation of the individual time series.

**Literature on debate 2** - The results of Dangerfield and Morris (1992) show that the highest forecast accuracy at the disaggregate level is achieved by direct disaggregate forecasts. Under the assumption of stationary demand, Zotteri and Kalchschmidt (2007) argue however that for high forecast accuracy at the disaggregate level, the TD-approach is more appropriate than the direct disaggregate approach in situations where demand has a high degree of variability and patterns are homogeneous.

Disaggregation in the TD-approach is done by calculating an item’s share of the total forecast through certain proration methods. Results of an empirical analysis of the effectiveness of 21 proration methods by Gross and Sohl (1990) indicate that disaggregation rules which are based on the simple average of an item’s share over a certain period of historical sales tend to yield the highest forecast accuracy at the disaggregate level.

**Conclusion** - Martens (2014b) concludes that there is no general consensus in scientific literature on the conditions that make one type of forecasting approach outperform the other in both debates. Almost 60 years of scientific research has so far not resulted (yet) in specific guidelines on when to apply what type of HF-approach. We can only conclude that the effectiveness of approaches is situational. Based on an extensive literature review on HF, Ouwehand (2006) concludes that results of the majority of papers tend to prefer a direct disaggregate forecast over a TD-approach, and a BU-approach over a direct aggregate forecast. As long as the specific criteria that determine the best forecast approach have not been assessed and the risk of choosing one approach over the other is not quantified, no specific recommendations can be provided for the decision making on the application of HF-approaches.

### 3.5 Idealized design - Process

The previous paragraph explained how to obtain demand forecast information at various levels, based on the forecast information at a pre-defined level. In this section, the ideal setup of the process to obtain the forecast information at that level is discussed. Based on the frameworks of Silver et al. (1998) and Kilger and Wagner (2008), an idealized conceptual process design is composed. This process is depicted in figure 3.3. The timeline reflects only a suggested division of activities over time, more important is the sequence of the activities. Safety Stock (and MRP-types) is regarded as tactical parameters. Broft (2014) and Kreuwels (2014) suggest updating these parameters on a quarterly basis.

**Statistical forecasting (lane 1, figure 3.3)** - In cleaning the historical data, demand planners have to adjust for all non-recurrent events that resulted in an extreme deviation of sales values. These can be promotions, strikes, environmental conditions etcetera. Moore and McCabe (1989) suggest using the modified z-score method to identify outliers. Only in case the demand planner is completely aware of the underlying cause of the outlier, the outlier could be replaced by a representative value (for instance according to the method of four period moving average). Otherwise, we suggest replacing the value of the outlier with the sales value corresponding to the nearest z-score value.
As a result of unsatisfactory forecast accuracy results of previously performed forecasts, statistical models and their parameters may need to be updated. A tracking signal (explained in section 3.5.1) could be useful to identify demand patterns for which the forecast model or parameters need to be updated.

Voudouris, Owusu, Dorne and Lesaint (2008) list several important criteria which should be satisfied for a proper execution of statistical forecasting:

1. **Statistical models can only be applied when there is enough historical data available to generate valid predictions** - Assuming very small random variation, Hyndman and Kostenko (2007) suggest a theoretical minimum of 3 observations for trend models and 17 monthly observations for the Holt-Winters model (capturing trend and seasonality).

2. **Patterns in time series should be properly detected in order to select appropriate forecast techniques** - Depending on the expertise of demand planners in identifying patterns in time series, the selection of statistical models could benefit from the use of IT-supported automatic model selection. When the demand planner has specific domain knowledge, it could be better to select the forecasting technique manually. Lawrence, Goodwin, O’Connor and Onkal (2006, p. 499) define domain knowledge as ‘..any information relevant to the forecasting task other than the time series’.

Directly after statistical forecasting, demand planners can adjust the statistical forecast when there is important domain knowledge available, not included by the statistical model. This revision should be done according to the **revised extrapolation forecasts** – procedure as described by Armstrong and Collopy (1998). Statistical forecast are only allowed to be manually adjusted when pre-defined triggers (promotions etcetera) indicate the domain knowledge.

**Judgmental forecasting (lane 2, figure 3.3)** - Next to the statistical forecasts, other departments (marketing/sales) can generate full judgmental forecasts. These forecasts are based on in-depth market knowledge and historical sales data. In order to enhance the forecast accuracy and reduce presence of bias, the methodology of *structured analogies* could be used (Green & Armstrong, 2007). This methodology prescribes to identify similar situations in the past which may provide useful information regarding the future course of sales.

**Consensus forecasting** - In order to align the judgmental and statistical forecasts, monthly consensus meetings should be organized for the teams that generate those forecasts (Fosnaught, 1999). In case the different types of forecasts differ in height, the teams should have a discussion about the underlying assumptions of the forecast. The overall intention of the consensus meetings is to align on a final set of assumptions (Drumm, 1993). When this alignment is achieved, final forecast numbers can easily be derived.

To facilitate better and shorter discussions in the consensus meetings, the different teams could briefly update each other in advance on the assumptions they have made to generate their forecasts (Drumm, 1993). Based on this information, the teams could create an own opinion on these assumptions prior to the consensus meeting.

**Creation of demand plans (lane 3, figure 3.3)** – Broft (2014) and Kreuwels (2014) provide a conceptual idealized design for alignment of the planning processes on the tactical level. Based on the mid-term unconstrained demand forecasts (consensus forecasts), a preliminary delivery plan is created by a Supply Chain Specialist Team (SCST) in collaboration with Demand Management and Marketing in the logistic regions. Whereas Broft (2014) and Kreuwels (2014) regard statistical demand forecasting and sales planning (judgmental forecasting) as two different decision functions, our idealized design prescribes to combine both activities into one demand planning decision function. By making demand planning one responsibility of multiple teams with different objectives, one is more likely to control the bias in forecasting.
Figure 3.3 – Idealized design monthly demand planning process
The preliminary distribution plan serves as input for the creation of the preliminary production plan, which is done by SCST in collaboration with the production plant(s). Both plans are subsequently used to agree upon a Supply Chain based Sales and Operations Plan. At this moment, (allied) suppliers can be informed regarding the expected orders by providing them with demand previews.

At the operational level, short-term forecasts are used for the creation of the master production schedule. An order acceptance-function is installed to decide which short-term forecasts are integrated in the actual master planning schedule. Broft (2014) and Kreuwels (2014) stress the use of engagement rules to support this decision making process.

3.6 Idealized design - Control

In scientific literature, forecast accuracy metrics are widely discussed. Due to the fact that there is no metric that is universally best (Silver et al., 1998), this debate remains ongoing. However, literature on the use of these metrics is relatively sparse. In order to construct the idealized control mechanism for demand planning in an integrated planning environment, we will first identify the main motivations for controlling demand planning. In figure 3.4 the ideal forecast accuracy feedback loops are depicted. On the one hand, forecast accuracy information is used in demand planning activities. On the other hand, the information serves as an input for supply planning decisions.

Based on figure 3.4, we formulate the following objectives of demand planning control:

**Objective 1**  *Improving the demand planning performance* - Forecast accuracy is a useful performance indicator, based on which erroneous practices in demand planning can be identified and resolved.

**Objective 2**  *Quantifying the (financial) impacts of demand uncertainty to support supply planning decisions* – The inability to precisely forecast the actual demand of future periods causes demand uncertainty. Since planning decisions on the tactical level and operational level are based on mid-term and short-term forecasts, they incorporate this demand uncertainty. Metrics servicing this objective should quantify the demand uncertainty to create an awareness of its (financial) impacts and support decision making in supply planning activities.
Metrics in literature - The majority of the forecast accuracy metrics discussed in literature, is based on the following definition of a forecast error:

$$\varepsilon_{i,t} = D_{i,t} - F_{i,t-x,t} \quad (3.1)$$

, where:

- $D_{i,t}$ the actual demand of SKU $i$ at time $t$
- $F_{i,t-x,t}$ the forecasted demand of SKU $i$ made at time $t-x$ for time $t$

In equation 3.1, the variable $x$ represents the forecast lag. The forecast lag corresponds to the time period between the moment forecasts are integrated in the master production schedule and the moment when the actual customer demand occurs. Due to the different replenishment lead times of markets, each market has an own value for the forecast lag.

The characteristics of a forecast accuracy metric typify its functionality. Those characteristics can be divided into hard-characteristics (which each metric should satisfy in order to provide effective feedback) and soft-characteristics (which may differ per control-objective). An overview of the characteristics is provided in table 3.3.

Table 3.3 – Forecast accuracy metric characteristics

<table>
<thead>
<tr>
<th>Type of characteristic</th>
<th>Characteristic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soft</td>
<td>Summarisable</td>
<td>In order to provide a summarized forecast accuracy value for a cluster of forecasts, the metric should display an absolute error value. (Kilger &amp; Wagner, 2008). This impedes over-forecast and under-forecast errors from levelling out.</td>
</tr>
<tr>
<td></td>
<td>Scale-independent</td>
<td>In order to fairly compare forecast accuracy values of different time series, the accuracy value should be independent of the scale (volume) of time series (Hoover, 2009).</td>
</tr>
<tr>
<td></td>
<td>Intuitively understandable</td>
<td>In order for practitioners to properly decide on follow-up actions, the result of a metric should be easily interpretable (Hoover, 2009).</td>
</tr>
<tr>
<td>Hard</td>
<td>All input data available</td>
<td>For a metric to effectively display forecast accuracy, all input data which are required for the computation of metric values should be available for all instances of the demand planning structure (Kilger &amp; Wagner, 2008).</td>
</tr>
<tr>
<td></td>
<td>Applicable to all time series</td>
<td>A metric should provide representative accuracy results for both intermittent and normal demand patterns.</td>
</tr>
<tr>
<td></td>
<td>Stakeholders’ agreement</td>
<td>All stakeholders should agree (after education and training) to the use of metrics for a control mechanism to be implemented successfully (Kilger &amp; Wagner, 2008).</td>
</tr>
</tbody>
</table>

For each of the two objectives of a control mechanism, we will specify which soft-characteristics the forecast accuracy metrics should have. In addition, further recommendations on the use of metrics per objective are provided.

3.6.1 Objective 1 - Metric requirements and use

The forecast accuracy metric for objective 1 should be applied on the basic level of forecast information, which is provided to supply planning decision functions. The soft-requirements for the metric are presented in table 3.4.
Table 3.4 – Soft-characteristics metric objective 1

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Required?</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Summarisable</strong></td>
<td>✓</td>
<td>Due to the wide product portfolio of Hilti, it is desired to review accuracy performance and set targets for a cluster of products.</td>
</tr>
<tr>
<td><strong>Scale-independent</strong></td>
<td>✓</td>
<td>In order to check the forecast improvement over time or compare the performance between demand planners or other companies, the metric should be scale-independent.</td>
</tr>
<tr>
<td><strong>Intuitively understandable</strong></td>
<td>✓</td>
<td>In order to correctly identify the causes of inaccuracy and properly decide on corrective actions, the practitioner should perfectly understand accuracy results.</td>
</tr>
</tbody>
</table>

**Target setting** - Forecast accuracy targets should depend on the ability to accurately predict the future course of sales. This ability related to the concept of controllability, as described by Van Wanrooij (2012). Van Wanrooij (2012) explains that controllability is linked to demand variability, for which the coefficient of variance of order sizes (CoV) is said to be a proper indicator. This parameter is a measure to express the degree of noise in data. Order frequency can indicate a cause of this variability. The two indicators are similar to the criteria underlying the demand pattern classification of Boylan, Syntetos and Karakostas (2008), which are frequently used in literature to indicate the ‘forecastability’ of demand patterns. In general, ‘forecastability’ decreases with the coefficient of variance of the order sizes and increases with the frequency of ordering. It should be noted that the forecast lag and level of aggregation at which forecast accuracy is measured also influence forecast accuracy values. Since the forecast lag influences ‘forecastability’, forecast accuracy targets should be set differently per market.

**Process control** - With regards to the process, forecast accuracy should be measured for all individual forecasting activities that directly affect the forecasted figures (Kilger & Wagner, 2008; Silver et al., 1998). Relating to the idealized process diagram in figure 3.3, this implies that measurements should take place after statistical forecasting, judgmental revision of the statistical forecasts, the judgmental forecasts and the total consensus forecasts. Tracking accuracy values after each of those process steps enables to accurately identify the problems that negatively affect the forecast performance.

**Tracking signal** – In addition to a standard forecast accuracy metric (for target setting etcetera), Harrison and Davies (1964) suggest the use of a cumulative sum tracking signal in order to detect when a forecasting model is no longer adequate. This tracking signal monitors the bias in a forecasting process. The tracking signal is defined as:

\[ U_{i,t} = \frac{C_{i,t}}{MAD_t} \]  

(3.2)

where:

\[ C_{i,t} = \epsilon_{i,t} + C_{i,t-1} \]

\[ MAD_t = \frac{1}{T} \sum_{x=0}^{T} |\epsilon_{i,t-x}| \]

If the forecasting process shows no bias, \( U_{i,t} \) should fluctuate around 0. A positive value indicates a bias to underforecast, while a negative value suggests a bias to overforecast. Forecast models and/or parameters should be modified once the tracking signal exceeds a predefined control limit. In general, it is suggested to use a value of ± 4 as a boundary.
3.6.2 Objective 2 - Metric requirements and use

To mitigate the negative effects of demand uncertainty on customer service levels, one can adjust the parameters of lead time, inventory and capacity. In order to quantify the amount of inventory and capacity that is needed to protect oneself against demand uncertainty, a (set of) metric(s) is needed that displays the amount of excess or shortage capacity/inventory in the past. We will discuss the requirements and use of metrics supporting both resource allocation and inventory investment decisions.

**Resource allocation decisions** - Forecast accuracy metrics to support decisions on resource allocation should be applied at the level of aggregation at which those decisions are made. It is suggested to track accuracy values over time to verify the consistency of forecast errors. Table 3.5 lists the metric requirements to support decisions on resource allocation.

**Table 3.5 – Soft-characteristics metric objective 2 (resource allocation)**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Required?</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summarisable</td>
<td>🔴</td>
<td>In order to decide upon resource allocations, information on the direction of forecast errors is required. In addition, over-forecast and under-forecast errors are allowed to level out.</td>
</tr>
<tr>
<td>Scale-independent</td>
<td>✓</td>
<td>In order to relate the metric results easily to the financial impacts, the metric should be scale-independent.</td>
</tr>
<tr>
<td>Intuitively understandable</td>
<td>✓</td>
<td>In order to properly decide on corrective actions, the practitioner should perfectly understand accuracy results.</td>
</tr>
</tbody>
</table>

**Inventory investment decisions** - Metrics to support inventory investment decisions are predominantly used to calculate the height of safety stocks. The requirements of a metric depend on the type of safety stock calculation method used. In general, the height of safety stocks is amongst others based on the cumulative forecast error over the replenishment lead time. This error clearly reveals the impact of inaccurate forecasts on the inventory hedge which is installed to maintain desired customer service levels in situations of (high) demand uncertainty.
Chapter 4

Gap-analysis

In the Gap-analysis, the differences between this ideal setup and current state of the process are identified. These gaps impede the process from achieving the ideal performance. Based on the identified gaps, design objectives and questions are formulated.

4.1 Introduction

The existence of gaps between the idealized design and the AS-IS state of demand planning results in a sub-optimal contribution of demand planning to the overall supply chain performance. In this section, we aim to identify those gaps by comparing the ideal state (chapter 3) with the current state of the process (chapter 2). The structure of this gap-analysis is similar to the setup of the idealized design.

4.2 Gap-analysis - Workflow

An extensive gap-analysis on the workflow is provided by Broft (2014) and Kreuwels (2014). They distinguish between gaps on the tactical level and gaps on the operational level.

**Gaps on tactical level** - In general, Broft (2014) and Kreuwels (2014) state that a tactical level in the planning hierarchy of Hilti is missing. This implies that there is no structural alignment on supply and demand on an aggregated level. The lack of aggregate materials requirements planning, capacity planning, production planning and distribution planning may eventually lead to plants having to postpone or even deny production orders due to the unavailability of material or resources. In addition, safety stock is updated too frequently which may yield high demand volatility upstream (bullwhip effect).

**Gaps on operational level** - Kreuwels (2014) remarks that there is no plant that works with a master production schedule in the entire GM (Global Manufacturing) division. A master production schedule would enable production plants to smoothen their production workload to greater extent, thereby reducing the operational and capital costs. Short-term forecasts are used only to low extent for the purposes of procurement, personnel planning and production planning and scheduling.

**Consequences** - As the key alignment-activities on the tactical and operational level are missing, the low level net requirements (which are based short-term forecasts) are directly converted into production orders. Due to the facts that forecasts are revised and adjusted on a daily basis and the net requirements are calculated with a similar frequency, production plants receive highly unstable inputs. This makes it impossible for those plants to create stable production plans, which will have negative impacts on operational costs, inventory levels and customer service.

4.3 Gap-analysis - Structure

In the AS-IS analysis, it is mentioned that forecasts are in principle performed at a DFU-level per month. However, several separate departments apply a certain HF-approach assuming this will reduce the forecast workload. Table 4.1 provides an overview of these incentives. To locate the incentives in one of the scientific debates on literature (column 4), we define the item-location-market organization per month as the disaggregate level and any higher level as the aggregate level.
Given the knowledge provided by the idealized design, we discuss the effectiveness of these approaches individually.

In HNA, the TD-approach is used (at product family level) for obtaining high forecast accuracy at a disaggregate level for all consumables. Scientific literature has not concluded (yet) whether a TD-approach may increase or decrease forecast accuracy at the disaggregate level. Since the results of Dangerfield and Morris (1992) for instance indicate that forecast accuracy could be harmed by the TD-approach, we classify the incentive of HNA as a possible gap.

In LEC, LESE and the BUs, it is intended to achieve high forecast accuracy at aggregate level through direct aggregate forecasting. Literature on the debate between direct aggregate forecasting and the BU-approach does not provide any guidelines for the selection of an approach to optimize forecast accuracy at the aggregate level. Since the BU-approach could outperform the direct aggregate approach (as shown by Dunn et al. (1976) and Gordon et al. (1997)), forecasting through the direct aggregate approach is labeled as a possible gap.

In general, we should remark that forecasting on the default DFU-level could be classified as a possible gap as well. Since scientific literature does not provide any general guidelines, forecasting on a DFU-level could cause a sub-optimal demand planning performance.

Further data-analysis should reveal the actual impacts of the different forecast approaches used on forecast accuracy.

4.4 Gap-analysis - Process

Figures 2.4 and 3.3 show respectively the AS-IS and Idealized forecasting process. By comparing both diagrams, gaps are identified.

4.4.1 Gaps in statistical forecasting

Table 4.2 provides an overview of the gaps and shows in addition whether the gaps negatively affect the efficiency of the process or the accuracy of forecasts.

<table>
<thead>
<tr>
<th>Type</th>
<th>Gap</th>
<th>Explanation</th>
<th>Consequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of corporate</td>
<td>Weekly execution of corporate forecasting</td>
<td>Purely from process-point-of-view, it is inefficient to perform weekly statistical forecasts when forecasts are created in monthly buckets.</td>
<td><strong>Inefficiency</strong></td>
</tr>
<tr>
<td>forecasting process</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>History adjustment</td>
<td>No strong focus</td>
<td>There is no strong focus on history adjustment in all regions and no structured correction in the corporate process.</td>
<td><strong>Inaccuracy</strong></td>
</tr>
<tr>
<td></td>
<td>Ineffective alerts</td>
<td>The alerts used for history correction do not identify outliers in historical sales, but reveal situations of low forecast accuracy.</td>
<td><strong>Inaccuracy</strong></td>
</tr>
</tbody>
</table>
Corporate forecasting is based on replenishment history | Replenishment history (instead of actual customer sales) is less suitable to use as a basis to generate statistical forecasts. These data represent the net requirements (instead of actual sales) and are influenced by for instance lot sizing rules. As a result, patterns in these time series are unreliable and more difficult to detect.  

Modification of forecast models and parameters | The preventive alerts currently used to modify forecast models do not properly identify when a forecast model needs to be changed. The continuous use of inappropriate forecast models could result in inaccurate forecasts. In addition, the modification is supposed to be done prior to the statistical forecast run in APO.  

Forecast model selection | Nowadays, forecast models are selected at the end of INPs, when only 12 monthly observations are available. Unless there are historical data available of a predecessor of the new product which are expected to be similar to the sales pattern of the new product, a (seasonal) statistical forecast technique cannot be correctly selected after 1 year of sales data (minimally 17 monthly observations are required).  

Lack of knowledge | Nowadays, demand planners select a forecasts technique based on their experience (and trial and error). Since the demand planners do not necessarily have a statistical background, selected models may not properly capture patterns in time series. The selection procedure may benefit from a more structured approach and/or further integration of an automatic model selection functionality.  

Integration of events | Officially, events affecting corporately planned items should be communicated well in advance according to formal guidelines. In practice, this is often not done.  

Logistic Regions versus Sales Regions | The definition of logistic regions does not always overlap with the one of sales regions. In these situations, Marketing and Sales belong to a different region than Materials Management. Due to this difference in regions, promotions and other events are sometimes poorly communicated.  

**Safety Stock and MRP-type update frequency** – Safety stock and MRP-types are characterized as tactical parameters which should be updated quarterly. This is currently done on respectively a monthly and weekly basis. The high updating frequency may cause high demand volatility upstream.
4.4.2 Gaps in judgmental forecasting

Structured procedures as the Delphi-method and the use of structured analogies may support the performance of judgmental forecasting. Currently, Marketing provides judgmental forecasts based on insights of Marketing at MO-level, historical data and their own experience. Although the application of one of the more structured approaches is likely to result in higher forecast accuracy, it could make the process more time-consuming. In order to assess the value of the judgmental forecasts, they should be documented officially. This is currently not (always) done.

4.4.3 Gaps in consensus forecasting

With regards to consensus forecasting, four gaps are observed:

1. **Frequency of consensus meetings** - Not all markets have monthly SFI-meetings. It is however suggested to do this on a monthly basis, to properly integrate the judgmental forecasts in the statistical forecasts in order to ensure more accurate tactical and operational demand plans.

2. **Focus of discussions in consensus meetings** - MM and Marketing should discuss the assumptions based on which Marketing generated the judgmental forecasts in order to reduce the bias. Currently, judgmental forecasts are sometimes directly integrated in the statistical forecasts without discussion (for instance in LEC).

3. **Aggregation/disaggregation** – In order to fairly compare the statistical and judgmental forecasts in the consensus meetings, the statistical forecasts need to be aggregated and converted into sales value. After integration of the judgmental forecasts, numbers should be disaggregated to a DFU-level. It is unclear to what extent the conversion into sales value is done properly. The disaggregation to DFU-level is mainly based on experience. To increase the performance of this process, Marketing should be involved to greater extent in case Product Managers have more detailed information. Proration rules, as discussed by Gross and Sohl (1990), may support this process as well.

4. **Pre-consensus meeting** - SFI-meetings could be more efficient by sharing the assumptions underlying the forecasts in advance. This is currently not (or to low extent) done.

4.4.4 Gaps in creation of demand plans

Nowadays, unconstrained consensus forecasts are used directly to calculate the net requirements. Either the net requirements calculation or rolling forecasts are subsequently directly used as an input for procurement, production planning and distribution planning. Due to the lack of a tactical level and the alignment of supply and demand at the operational level, inventory costs will be high and customer service sub-optimal as a result of the facts that:

- Production may have to postpone or even deny orders as a result of capacity constraints
- Production may have to postpone or even deny orders as a result of material unavailability, since procurement and production may be misaligned
- Inventory levels may be too high due to the possible misalignment of procurement, production and distribution (e.g. production and distribution capacity constraints impede from processing all procured or produced components/items)

4.5 Gap-analysis - Control

The description of the ideal control mechanism for demand planning in an integrated planning environment defines the metric requirements and use per control objective. In this part of the gap-analysis, we will concisely identify the gaps per objective of demand planning control.
4.5.1 Gaps in control objective 1

Use of metrics - The idealized design states that a forecast accuracy metric for objective 1 should be summarisable, scale-independent and intuitively understandable. In addition, the metric should be applicable to all types of time series, should have the stakeholder's agreement and all required input data should be available. In table 4.3 it is assessed whether the currently used official and unofficial forecast accuracy metrics at Hilti meet those requirements.

Table 4.3 – Gap-analysis metrics objective 1

<table>
<thead>
<tr>
<th>Metric</th>
<th>Official</th>
<th>HNA</th>
<th>LEC</th>
<th>LESE</th>
<th>BUs</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Level</em></td>
<td>Official FQ</td>
<td>FA-monetized</td>
<td>FQ-aggregated</td>
<td>Official FQ</td>
<td>-</td>
</tr>
<tr>
<td><em>Lag</em></td>
<td>DFU</td>
<td>DFU</td>
<td>Item/location</td>
<td>DFU</td>
<td>-</td>
</tr>
<tr>
<td><em>Horizon</em></td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td><em>DFU</em></td>
<td>6 months</td>
<td>12 months</td>
<td>6 months</td>
<td>6 months</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Official</th>
<th>HNA</th>
<th>LEC</th>
<th>LESE</th>
<th>BUs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summarisable</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Scale-independent</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Understandable</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Stakeholders’ agreement</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Availability of input data</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Applicable to all types of time series</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

The results in table 4.3 reveal that all existing accuracy metrics are not intuitively understandable. For HNA, this is due to the fact that the metric measures the absolute percentage error based on the forecasted value instead of the actual demand value. Mainly in situations of a higher forecast bias, a percentage error scaled over the forecast will yield results which are more difficult to understand than the results of a metric based on the actual demand value. In addition, the metric proportionally aggregates based on the actual value of products. This makes it more difficult for demand planners to identify the cause of poor performance, since it may not be immediately obvious which products they should focus on in order to achieve their accuracy targets.

The official FQ-metric is hard to understand, due to the fact that the denominator may switch. In addition, it incorrectly uses a lag of 2 months for all markets. The FQ-aggregated metric, used in LEC, is applied at an item-location level. This level does however not correspond to the basic level of aggregation at which forecasts are provided to subsequent planning stages (DFU-level). As a consequence, the results of this metric do not properly support decision making on corrective actions to improve the performance of demand planning.

Although the official FQ-measure is formally used for all markets, the existence of various self-created metrics indicates that FQ is not generally accepted by the markets. If we regard all MM-teams and Marketing teams in all markets, BUs and production plants to be stakeholders, we argue that the locally self-created metrics do not have an overall stakeholders’ agreement either. These metrics are only used locally.

Table 3.2 demonstrates once more that forecast accuracy is currently not measured for the corporate forecasts in the BUs.
Target setting - Currently, forecasts accuracy targets are set based on the TABCD-label of products. This label indicates the relative importance of the product in terms of sales value. The idealized design suggests however that forecast accuracy targets should be related to the ‘controllability’ of items. Controllability is indicated by the coefficient of variance of order sizes. By not taking this criterion into account when setting forecast accuracy targets, one risks to define targets which are unachievable or too easy to achieve.

Process control - At the moment, forecast accuracy is measured once per month based on the values of demand forecasts after the second weekend of the month. To accurately track the root causes of forecast inaccuracy, the idealized design suggests to measure forecast accuracy for each individual forecasting activity in which forecasts are generated or adjusted. Current forecast accuracy measurements do not enable a proper identification of the cause of forecast inaccuracy.

Tracking signal - Next to the accuracy metrics used globally and in the regions, no accuracy measurement is performed to display the systematic bias of forecasts. The idealized design suggests that a tracking signal, indicating the systematic bias, is useful for the identification of inadequate forecast models.

4.5.2 Gaps in control objective 2

Nowadays, demand forecasts are not or to low extent used for resource allocation decisions on the tactical and operational level. In addition, there is no forecast accuracy information provided to support these decisions. As a result, resource allocation decisions are more likely to be taken incorrectly and result in excess or shortage planning.

At the moment, the mean absolute deviation (MAD) metric is used to support two safety stock calculation methods (SB5 and SB8). For SB5, the mean absolute deviation of the monthly consumption in the last 12 months is required. For SB8, the average positive forecast error in the past 3 months is used as an input. Since the safety stock calculation is out of the scope of this project, we will not assess the quality of the methods used.

4.6 Problem definition

The gap-analysis revealed the main gaps between the desired state of demand planning and the way the process is currently being executed at Hilti. An overview of the gaps is presented in the cause-and-effect diagram in figure 4.1. We conclude in general that all aspects that deviate from the ideal state of demand planning in the structure-, process- and control-components have a negative impact on the effectiveness and/or efficiency of demand planning. The effectiveness is related to the ability to provide accurate demand plans. Efficiency corresponds to the amount of resources that is required to generate the outputs. The gaps in the category ‘workflow’ impede the optimal use of the outputs of demand planning in an integrated planning environment. The complete set of gaps causes demand planning to sub-optimally contribute to the overall supply chain performance. This may result in too high inventory levels, too high operational costs and poor customer service.

In the next phase of the project, a TO-BE design is created with the intention of partially closing the gaps between the idealized design and the current state of the process. Due to time restrictions, it has been decided to tackle only the red-colored gaps on a conceptual level (figure 4.1). In addition, the actual implementation of the TO-BE design is regarded to be out of scope. Although the safety stock calculation process is taken into account in the redesign of demand planning, the actual methodologies used to compute the height of safety stock are excluded from the scope as well.
REDESIGNING A DEMAND PLANNING PROCESS IN AN INTEGRATED PLANNING ENVIRONMENT

Figure 4.1 – Cause-and-effect diagram demand planning

Based on the gaps, the following design objective is defined:

Providing scientifically grounded guidelines on redesigning the demand planning decision function on the tactical and operational level in an integrated planning environment with the objective of improving the contribution to the overall performance of the internal end-to-end supply chain

First, the scientifically grounded guidelines should be related to the redesign of the demand planning process itself, aiming to increase its effectiveness and efficiency. Second, the guidelines should concern the alignment of demand planning with subsequent decision functions in a future integrated planning process. As indicated by the gaps in the workflow-component, the majority of these decisions functions is currently not in place. We will therefore make use of the recommended operations planning concept of Broft (2014) and Kreuwels (2014) as a reference model.

Related to the design objectives, we state the following design questions which correspond to the four different sections of the gap-analysis:

1. In what way should demand planning be aligned with subsequent decisions functions on the tactical and operational level in order to improve the overall supply chain performance?
2. How should the demand planning process make use of Hierarchical Forecasting approaches in order to create an optimal balance between forecasting accuracy and process efficiency?
3. How should demand planning process be redesigned from a process-perspective in order to efficiently create accurate demand plans?
4. In what way should the control-mechanism be redesigned in order to continuously improve demand planning performance and quantify demand uncertainty with the intention to support planning decisions in further stages of integrated planning?

As the demand planning framework in figure 3.1 suggests, the guidelines on the redesign of the individual component of demand planning should be aligned in order to improve demand planning performance as a whole.
Chapter 5

TO-BE design

In order to (partially) close the gaps between the idealized design and the AS-IS state of the process, a TO-BE design is created. Scientific insights and creativity are used to provide appropriate means of bringing the current demand planning process towards the idealized state.

5.1 Introduction

As formulated in the previous chapter, the objective of this project is to provide scientifically grounded guidelines on the setup of demand planning in an integrated planning environment. Guidelines are most effective when the practitioner understands the mechanisms and motivations underlying these guidelines. In this chapter we will therefore highly focus on the aspects that need to be considered when redesigning demand planning instead of simply providing ‘the answer’.

The guidelines presented in this chapter aim at partially closing the gaps between the desirable future state of demand planning and the current state. The TO-BE design phase is typically repeated several times to continuously improve the process. The structure of the TO-BE design presented in this chapter is similar to the setup of the gap-analysis.

5.2 Workflow

This project aims to redesign the demand planning process in an integrated planning environment. Since an integrated planning concept is currently not yet in place, we will elaborate on the TO-BE design of an integrated planning concept at Hilti, provided by Broft (2014) and Kreuwels (2014). We use the overview of demand forecast requirements in table 3.1 as a reference. Since the calculation of safety stock is out of the scope of this project, we will elaborate on the current demand forecast requirements for safety stock calculation: item- (and location) specific forecasts per month with a horizon of 3 months.

5.3 Structure

Table 3.1 lists all levels of aggregation at which demand forecast are required. In this section, it is discussed how these forecasts should be obtained.

Currently, the majority of demand forecasts is generated at DFU-level per month. The overview of literature on hierarchical forecasting in the idealized design showed that currently no general guidelines are available for decision making on the level of aggregation at which forecasts should be generated. Any design decisions that would suggest any different level of statistical forecast generation would therefore not be scientifically grounded. As a consequence, no scientifically grounded recommendations can be provided for the determination of the optimal forecast level. In addition, there is no general consensus in literature on the extent to which a change in aggregation level could harm accuracy. This impedes us from conducting a cost-benefit analysis for the change of the forecast level.

When forecast information is provided at one level (currently DFU-level per month) to subsequent decisions functions forecasts, this information should be (dis)aggregated to satisfy the need of forecast information at different levels. For the tactical level, forecasts should be simply consolidated to the levels presented in table 3.1. For the operational level, forecasts should be aggregated to the item (and location) level. However, the forecast information should be disaggregated into weekly buckets. This could be done through continuing with the policy of disaggregating into equal shares.
5.3.1 Quantitative analysis

In the remainder of this section, we aim to contribute to available scientific knowledge by conducting an exploratory quantitative analysis on the performance of the TD-approach and the direct disaggregate approach with the objective of optimizing forecast accuracy at the disaggregate level. The objective of this analysis is twofold. On the one hand, it is aimed to assess whether the effectiveness of the approaches is related to the characteristics of a cluster of time series. On the other hand, it is intended to verify whether the stability of the effectiveness of the approaches over time. For the first objective, we define two main criteria which may influence the performance of forecast approaches:

1. **Homogeneity of sub-aggregate variables** – In accordance with the arguments of Ouwehand (2006) and the research of Zotteri and Kalchschmidt (2007) and Dangerfield and Morris (1992), we argue that the effectiveness of the TD-approach depends on the similarity of the subaggregate time series. Especially for non-stationary time series, essential information of the patterns in these time series might be lost when time series are aggregated (Ouwehand, 2006). By clustering homogeneous time series, one prevents the aggregate process from becoming too complex to model and forecast.

2. **Volatility of sub-aggregate time series** – Pooling time series could be useful, since the aggregated time series may show less variation as random fluctuations cancel out. Duncan et al. (2001) state as a basic principle that the improvement in forecast accuracy of pooling methods increases with the volatility of time series.

Previous articles have examined the individual influences of homogeneity (Dangerfield & Morris, 1992) and volatility (Duncan et al., 2001). However, no research has been found that looked at the joint influence on the effectiveness of the approaches. As long as no specific guidelines for the selection of a forecast approach are provided, it may be useful to be aware of the stability of the effectiveness of forecast approaches over time.

**Hypotheses** - Assuming that homogeneity will decrease the complexity of modeling and forecasting the aggregate time series, we expect the effectiveness of the derived forecasts to increase with the homogeneity. Similarly, we expect the derived forecast approach to be more effective in cases of volatile subaggregate time series, since the effect of cancelling out random variations when consolidating data has more influence. In order to analyze the combined effects of the criteria, we test the accuracy of derived forecasts in six situations which are reflected in the matrix of figure 5.1. In section 5.3.2, we will explain that homogeneity is measured by the correlation between time series. Segments (1) and (2) contain clusters with uncorrelated time series, segments (3) and (4) cover clusters with positively correlated time series and in segments (5) and (6) clusters with negatively correlated time series are located. For the segments with low volatility, we expect that positive correlation will increase the effectiveness of the TD-approach. The share of TD-clusters is thus expected to be higher in segment (3) than in segments (1) and (5). For uncorrelated clusters, we expect the effectiveness of the TD-approach to increase with the volatility of time series (share of TD-clusters is higher in 2 than in 1).

Regarding the stability of the effectiveness of approaches, we expect that the preference of forecast approach remains rather stable over time.
5.3.2 Data and methodology

For the quantitative analysis, we have made use of 48 months of historical sales data of 873 items in the logistic region of Hilti North America. This sample includes items with normal and variable demand (for definition see figure B.2, Appendix B) and spare parts with a reported sales value in all 48 months. Items with sporadic demand are excluded from the sample.

All 873 items are distributed through the warehouses in Tulsa and Dayton. Hence, each item is split over 2 DFUs (one for Tulsa and one related to Dayton). We define the aggregate level as the consolidation of the two DFUs (aggregation in the location dimension). The separate time series of the item per location represent the disaggregate level. In total, we thus have 873 clusters and 1746 subaggregate variables, where each cluster is composed of two subaggregate variables.

The motivation for selecting this sample of data is twofold. On the one hand, employees at Hilti indicated that a substantial amount of time series of the same items in different locations is expected to be homogeneous to great extent. A uniform distribution of clusters along the aspect of homogeneity would enable to draw statistically significant conclusions. On the other hand, the current aggregation method ensures stable cluster sizes. The impact of homogeneity and volatility on the performance of HF-approaches for clusters which contain more subaggregate variables may be different. Hence, a stable cluster size thus ensures a fair analysis of the influence of criteria.

In this analysis, the forecast accuracy of the derived and direct forecast approach has been measured over the last 12 months of the sample data. We define this part of the sample as the out-of-sample data (months 1 to 12). The in-sample data (months 35 to 0) are used for forecast model selection, parameter estimation, initialization of forecast models and the determination of homogeneity and volatility of the time series. The use of the sample is presented in figure 5.2.

**Figure 5.2 – Use of sample data**

*History correction* - In the first phase of the experiment, the historical sales data are adjusted through the modified z-score method. This method is preferred over the normal z-score method, since it is more robust to outliers. The modified z-score is calculated according to equation 5.1.

\[
    z_i = \frac{D_{i,t} - \bar{D}_i}{\text{Median of absolute deviation}_i}
\]

(5.1)

where

- \( D_{i,t} \) is the actual demand of item \( i \) at time \( t \)
- \( \bar{D}_i \) is the median of the in- and out-of-sample sales data of item \( i \)

and the median of absolute deviation is obtained as

\[
    \text{Median of absolute deviation}_i = \text{median}(\{|D_{i,t} - \bar{D}_i|\})
\]

(5.2)

All observations showing a z-score outside the bandwidth of [-3.5; 3.5] are labeled as outlier. These thresholds are chosen based on the work of Iglewicz and Hoaglin (1993). The sales value of outliers are replaced with the value corresponding to the nearest z-score threshold.
Model selection - In order to generate the demand forecasts, we use the exponential smoothing technique as discussed by Silver et al. (1998). In our analysis, six different models have been used. For non-seasonal demand patterns, we distinguish between patterns with a constant level, patterns with a damped trend and patterns with a linear trend. According to Gardner and McKenzie (1985), a model with damped trend may be preferred over a model with a linear trend in situations where the time series are noisy or the trend is erratic. For seasonal demand, we integrate seasonal influences in the same three models as used for non-seasonal demand. The overview of the resulting total of six forecast models is provided in the third column of table 5.1. The exact functionality of the models is explained in explanation box D.2 in Appendix D.

We adopt the model identification procedure of Gardner and McKenzie (1988) for the selection of the appropriate model per time series. This procedure examines six variances for each in-sample adjusted historical sales data (table 4.2). The forecast technique corresponding to the lowest variance is recommended to be selected. Similar to our analysis, Gardner and McKenzie (1988) assume multiplicative seasonality in case a seasonal model is indicated. They argue this is a safe assumption since Gardner and McKenzie (1989) have shown that the difference in forecast accuracy of additive and multiplicative seasonal exponential smoothing models is little.

Table 5.1 – Model identification rules (adopted from Gardner and McKenzie, 1988)

<table>
<thead>
<tr>
<th>Case</th>
<th>Series yielding minimum variance</th>
<th>Model selected</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>(D_t)</td>
<td>Constant level</td>
</tr>
<tr>
<td>B</td>
<td>((1-B) D_t)</td>
<td>Damped trend</td>
</tr>
<tr>
<td>C</td>
<td>((1-B)^2 D_t)</td>
<td>Linear trend</td>
</tr>
<tr>
<td>D</td>
<td>((1-B^p) D_t)</td>
<td>Seasonal, constant level</td>
</tr>
<tr>
<td>E</td>
<td>((1-B)(1-B^p) D_t)</td>
<td>Seasonal, damped trend</td>
</tr>
<tr>
<td>F</td>
<td>((1-B)^2(1-B^p) D_t)</td>
<td>Seasonal, linear trend</td>
</tr>
</tbody>
</table>

The computation of the six different variances in column two of table 5.1 is based in the backshift operator \(B\). This operator reflects the proportion of value of the time series at time \(t\) that represents the height of the value of the time series at time \(t-1\). The backshift operator is computed through formula 5.3. \(B^p\) is defined as the backshift operator over the number of \(p\) periods in a season (equation 5.4). At Hilti, products with a seasonal demand pattern show generally a seasonal period of 1 year (\(p = 12\) months).

\[
B = \frac{D_t}{D_{t-1}} \quad (5.3) \quad B^p = \frac{D_t}{D_{t-p}} \quad (5.4)
\]

where
\(D_t\) the adjusted demand at time \(t\)
\(p\) periods per seasonal cycle

Parameter optimization - Via a grid search, the smoothing constants and dampening parameters are found, which result in an optimal fit of the model on the data of months -23 to 0. The ranges in which the optimal values for the smoothing parameters are searched, are presented in table 5.2.

Table 5.2 – Smoothing constant settings for grid search

<table>
<thead>
<tr>
<th>Constant</th>
<th>(\alpha)</th>
<th>(\gamma_{nw})</th>
<th>(\Phi)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range</td>
<td>[0.01;0.30]</td>
<td>[0.05;0.50]</td>
<td>[0.05;0.95]</td>
</tr>
</tbody>
</table>
The range of $\alpha$ is similar to the range recommended by Silver et al. (1998). Based on the $\alpha$-value, $\alpha_{HW}$ and $\beta_{HW}$ are determined according to formulas 5.5 and 5.6.

$$
\alpha_{HW} = [1 - (1 - \alpha)^2]
$$

(5.5)  

$$
\beta_{HW} = \frac{\alpha^2}{[1 - (1 - \alpha)^2]}
$$

(5.6)

The optimal value of $\gamma_{HW}$ is only determined for seasonal models. For constant models, the value of $\varphi$ automatically equals 0, while this parameter has a fixed value of 1 for models with a linear trend.

In the analysis, we have forecasted the future demand of one period ahead. The initialization and parameter updating formulas for every model and are elaborated in explanation box D1 in Appendix D. For the initialization of the simple exponential smoothing model (model 1) the average of the first 12 months of data is taken.

The initialization for the models which incorporate seasonality is based on the deseasonalized and adjusted time series. The initialization and deseasonalization is done based on the initial seasonal indices, which are computed and normalized based on the centered 12-period moving average of all in-sample data. During the forecasting process, the seasonality factors are updated and renormalized every period. With the intention to estimate smoothing constants, all other models are initialized based on the first 12 months of the in-sample data. With the objective of forecasting the out-of-sample data, the last five models are initialized based on all in-sample data.

**Performance measure** - In order to assess the optimal value of the smoothing constants and measure the forecast performance of the out-of-sample data, the Mean Absolute Deviation (MAD) metric is used (equation 5.7). We argue that the MAD is most appropriate for the purpose of our analysis, since it is summarisable (required since we are interested in the summarized value of the forecast performance of the subaggregate variables in a cluster over the out-of-sample data). It is decided not to make use of a MAPE-metric, since this metric has difficulty in dealing with time periods in which the demand is zero or close to zero.

$$
MAD = \frac{1}{T} \sum_{t=1}^{T} |\epsilon_{i,t}|
$$

(5.7)

**Disaggregation in TD-approach** - Disaggregation in the TD-approach is done based on periodically updated proportional factors. These factors range from 0 to 1 and represent the proportion of the consolidated forecast that represents the derived forecasts on the low level. Based on the results of Gross and Sohl (1990), the analysis includes two scenarios for disaggregation. In the first scenario (TD-1), the periodical proportional factors are obtained through formula 5.8. For the second scenario (TD-2), formula 5.9 is used

$$
P_{i,t} = \left(\frac{\sum_{x=1}^{T} d_{i,t-x}}{D_{t-x}}\right) / T
$$

(5.8)  

$$
P_{i,t} = \left(\frac{\sum_{x=1}^{T} d_{i,t-x}}{T} / \left(\frac{\sum_{x=1}^{T} D_{t-x}}{T}\right)\right)
$$

(5.9)

,where

- $d_{i,t}$ the adjusted historical demand of product $i$ at time $t$
- $D_t$ the adjusted historical consolidated demand of all subaggregate variables
- $T$ the total number of time periods used to assess the proportional factors. The value of $T$ is set to 12 months to incorporate the full seasonal cycle.

37
Homogeneity measure - In order to assess the homogeneity of a cluster, we adopt the measure of correlation as used by Widiarta et al. (2009) and Dangerfield and Morris (1992). Correlation is represented by Pearson’s $\rho$ and is determined by formula 5.10.

$$\rho_{xy} = \frac{\sum_{t=-35}^{T} (x_t - \bar{x})(y_t - \bar{y})}{\sigma_x \cdot \sigma_y}$$ \hspace{1cm} (5.10)

, where

- $x_t$ the amount of adjusted demand in time bucket $t$ for subaggregate variable $x$
- $\bar{x}$ the in-sample mean of the adjusted demand subaggregate variable $x$
- $T$ the amount of observations of a subaggregate variable
- $\sigma_x$ the sample standard deviation of the adjusted demand of variable $x$

Correlation basically reflects the extent to which the course of one variable is dependent on the course of the other variable(s). The range of Pearson’s $\rho$ lies between -1 and 1, where -1 reflects highly negatively correlated time series and 1 stands for a highly positive correlation. The correlation of a cluster is determined based on the in-sample sales data. Formula 5.10 is only suitable to calculate the correlation of two variables. For future tests with bigger cluster sizes, we suggest computing the average of the individual multiple correlation coefficients in these situations. The algebraic instruction for this method is presented in explanation box D.2 (Appendix D).

Volatility measure - The volatility of time series is measured by the coefficient of variation (CoV) of detrended and deseasonalized data, as suggested by Duncan et al. (2001). The CoV is defined as the ratio between the standard deviation and the mean of the in-sample data. We assume the presence of a trend and/or seasonality only when the model identification procedure yields a model with a trend and/or seasonality.

Test of stability - In order to assess the stability of the preference of a HF-approach per cluster, we repeat the analysis for the year 2012 (months -11 to 0). By comparing the results of 2012 and 2013 it is intended to provide conclusions on the stability of the preference of a HF-approach. For the analysis of the data of 2012, correlation and average volatility is still determined based on all in-sample data. The forecast model is identified based on these data as well. The optimal smoothing constants are estimated through a grid search over data of 2011.

5.3.3 Analysis of results

Overall results - In table 5.3, the overall results of the analysis for the out-of-sample data (2013) are presented. In columns (4) and (5) the share of clusters is expressed for which a TD-approach with respectively proration rule 1 and proration rule 2 is preferred. It is observed that both TD-approaches are favored over the direct disaggregate approach. Among the two TD-approaches, TD-2 yields the best forecast accuracy in most clusters (59,9%). In columns (5) and (6) the average difference of the average cluster MAD of the TD-approaches compared to the average cluster MAD of the direct approach is presented. The last two columns show the average cluster performance of both approaches (direct and TD-2) compared to the ideal situation in which the optimal HF-approach would have been known in advance. Overall, the TD-2 approach shows to be less deviating from the ideal situation than the direct approach. For the four different BA-labels (Accessories, Consumables, Spares and Tools) all clusters are on average predicted best through the TD-2 approach. For T- and D-items the TD-1 approach appeared to be effective as well.

Table 5.3 shows that TD-2 approach outperforms the TD-1 approach in most situations. In order to ensure a clear presentation of the upcoming results, we will therefore provide merely the results of the TD approach based on proration rule 2.
Table 5.3 – Overall results 2013

<table>
<thead>
<tr>
<th>Group</th>
<th>Label</th>
<th>N (% of total sample)</th>
<th>%TD1</th>
<th>% TD2</th>
<th>Av. cluster MAD - TD1 comp. to direct</th>
<th>Av. cluster MAD - TD2 comp. to direct</th>
<th>Av. MAD Direct comp. to ideal</th>
<th>Av. MAD TD2 comp. to ideal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>Total</td>
<td>873 (100%)</td>
<td>58,1%</td>
<td>59,9%</td>
<td>-1,6%</td>
<td>-2,0%</td>
<td>+8,9%</td>
<td>+5,1%</td>
</tr>
<tr>
<td>BA</td>
<td>Acc.</td>
<td>146 (16,7%)</td>
<td>63,7%</td>
<td>67,1%</td>
<td>-2,5%</td>
<td>-2,6%</td>
<td>+9,2%</td>
<td>+5,1%</td>
</tr>
<tr>
<td></td>
<td>Con.</td>
<td>645 (73,9%)</td>
<td>56,1%</td>
<td>57,4%</td>
<td>-1,3%</td>
<td>-1,7%</td>
<td>+8,8%</td>
<td>+5,3%</td>
</tr>
<tr>
<td></td>
<td>Spares</td>
<td>73 (8,4%)</td>
<td>63,0%</td>
<td>65,8%</td>
<td>-3,1%</td>
<td>-3,7%</td>
<td>+8,6%</td>
<td>+3,4%</td>
</tr>
<tr>
<td></td>
<td>Tools</td>
<td>9 (1,0%)</td>
<td>55,6%</td>
<td>66,7%</td>
<td>-3,4%</td>
<td>-4,1%</td>
<td>+9,4%</td>
<td>+3,2%</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
<td>51 (5,8%)</td>
<td>62,7%</td>
<td>60,8%</td>
<td>-4,1%</td>
<td>-4,5%</td>
<td>+9,8%</td>
<td>+3,2%</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>244 (27,9%)</td>
<td>61,1%</td>
<td>61,5%</td>
<td>-2,8%</td>
<td>-3,2%</td>
<td>+9,3%</td>
<td>+4,2%</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>394 (45,1%)</td>
<td>56,3%</td>
<td>59,1%</td>
<td>-0,5%</td>
<td>-1,0%</td>
<td>+8,7%</td>
<td>+6,0%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>160 (18,3%)</td>
<td>55,0%</td>
<td>57,5%</td>
<td>-1,6%</td>
<td>-1,9%</td>
<td>+8,0%</td>
<td>+4,7%</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>24 (2,7%)</td>
<td>62,5%</td>
<td>66,7%</td>
<td>-3,4%</td>
<td>-3,4%</td>
<td>+10,9%</td>
<td>+5,1%</td>
</tr>
</tbody>
</table>

**Objective 1 – Criteria that determine effectiveness of approaches** – Figure 5.3 shows the distribution of the clusters based on the correlation and the average volatility of the subaggregate variables. Each dot represents one cluster of which the color indicates whether the cluster is in 2013 on average forecasted best by the TD-2 approach or the direct disaggregate forecast approach. The majority of the clusters has a correlation in the range of [-0,4;0,6] and an average coefficient of variance lower than 1.

![Distribution of clusters based on correlation and average volatility (2013)](image)

**Figure 5.3 – Distribution of clusters based on correlation and average volatility, 2013**
In figure 5.4, the share of TD2-clusters is presented per correlation-range. In the ranges of high negative correlation and high positive correlation, the share of TD2-clusters is larger. However, it should be noted that the sample size in these ranges is very low. In all other ranges, the share of TD2-clusters is relatively similar. In order to check the individual influence of the average volatility of subaggregate variables on the effectiveness of forecast approaches, the share of TD2-clusters is presented per range of average coefficient of variance as well (figure 5.5). Although there are a few clusters with an average volatility outside the ranges presented in figure 5.5, these ranges are not integrated since the amount of clusters they include is very low. Figure 5.5 shows that the share of TD2-clusters in the range of low volatility [0,0;0.2] is relatively high (71%). The ranges of higher volatility show a slightly lower share of TD2-clusters, although in each range the amount of TD-2-clusters exceeds the number of clusters for which a direct disaggregate approach is preferred.

To analyze the joint influence of the aspects of homogeneity and volatility on the effectiveness of the HF-approaches, the share of TD2-clusters is computed for six different segments. The results of this analysis are presented in table 5.4. Since the amount of clusters in segments (4), (5) and (6) are very low, no statistical significant conclusions can be based upon the results of these segments. The results of the first three segments show that the share of TD-clusters in is higher when the subaggregate variables are positively correlated and the average volatility is low compared to the clusters which are uncorrelated and have a low average volatility. The share of uncorrelated TD2-clusters with low average volatility is higher than the share of uncorrelated TD2-clusters with high volatility.

**Table 5.4 – Combined influence correlation and volatility, 2013**

<table>
<thead>
<tr>
<th>Segment</th>
<th>Correlation (r)</th>
<th>Average volatility (CoV)</th>
<th>N</th>
<th>% TD2-clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Uncorrelated (-0.5 &lt; r &lt; 0.5)</td>
<td>Low (CoV &lt; 0.5)</td>
<td>576</td>
<td>62%</td>
</tr>
<tr>
<td>2</td>
<td>Uncorrelated (-0.5 &lt; r &lt; 0.5)</td>
<td>High (CoV ≥ 0.5)</td>
<td>230</td>
<td>51%</td>
</tr>
<tr>
<td>3</td>
<td>Pos. correlated (r ≥ 0.5)</td>
<td>Low (CoV &lt; 0.5)</td>
<td>49</td>
<td>73%</td>
</tr>
<tr>
<td>4</td>
<td>Pos. correlated (r ≥ 0.5)</td>
<td>High (CoV ≥ 0.5)</td>
<td>16</td>
<td>63%</td>
</tr>
<tr>
<td>5</td>
<td>Neg. correlated (r ≤ -0.5)</td>
<td>Low (CoV &lt; 0.5)</td>
<td>2</td>
<td>50%</td>
</tr>
<tr>
<td>6</td>
<td>Neg. correlated (r ≤ -0.5)</td>
<td>High (CoV ≥ 0.5)</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>
Objective 2 – Stability of approach preference – In order to assess the stability of the preference of forecast approach over time, the quantitative analysis has been repeated for the year 2012. The results of this analysis are presented in table 5.5 and in Appendix D (figure D.1, D.2, D.3 and table D.1). Table 5.5 reveals that for the year 2012 the TD-approach is still preferred for the slight majority of clusters. The best type of TD-approach differs per group of products. We remark that the lowest average MAD for clusters is on average obtained through the direct disaggregate forecasting approach. Interestingly, the direct approach is also preferred for T-items, which are regarded to be most important for Hilti. Figure D.2 shows that the share of TD2-clusters gradually grows with the correlation. However, in the ranges of high positive correlation, the sample size is too low to support any conclusions. We observe in figure D.3 that the share of TD2-clusters is the highest for low-volatility and high-volatility ranges. No specific link between volatility and the effectiveness of the TD2-approach can however be assessed. The data in table D.1 display that the share of TD2-clusters remains relatively high in segment (3). Similar to 2013, segment (1) has a higher share of TD2-clusters than segment (2).

Table 5.5 – Overall results 2012

<table>
<thead>
<tr>
<th>Group</th>
<th>Label</th>
<th>N (% of total sample)</th>
<th>%TD1</th>
<th>% TD2</th>
<th>Av. cluster</th>
<th>Av. cluster</th>
<th>Av. MAD</th>
<th>Av. MAD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MAD - TD1 comp. to direct</td>
<td>MAD - TD2 comp. to direct</td>
<td>Direct comp. to ideal</td>
<td>TD2 comp. to ideal</td>
</tr>
<tr>
<td>Total</td>
<td>Total</td>
<td>873 (100%)</td>
<td>52,8%</td>
<td>52,7%</td>
<td>+0,8%</td>
<td>+0,4%</td>
<td>7,3%</td>
<td>6,3%</td>
</tr>
<tr>
<td>BA</td>
<td>Acc.</td>
<td>146 (16,7%)</td>
<td>54,8%</td>
<td>53,4%</td>
<td>-0,0%</td>
<td>-0,1%</td>
<td>8,0%</td>
<td>6,2%</td>
</tr>
<tr>
<td></td>
<td>Con.</td>
<td>645 (73,9%)</td>
<td>51,8%</td>
<td>52,2%</td>
<td>+0,8%</td>
<td>+0,3%</td>
<td>7,3%</td>
<td>6,1%</td>
</tr>
<tr>
<td></td>
<td>Spares</td>
<td>73 (8,4%)</td>
<td>53,4%</td>
<td>50,7%</td>
<td>+2,4%</td>
<td>+3,4%</td>
<td>6,1%</td>
<td>8,4%</td>
</tr>
<tr>
<td></td>
<td>Tools</td>
<td>9 (1,0%)</td>
<td>88,9%</td>
<td>88,9%</td>
<td>-4,5%</td>
<td>-4,3%</td>
<td>6,9%</td>
<td>1,9%</td>
</tr>
<tr>
<td>TABC</td>
<td>T</td>
<td>51 (5,8%)</td>
<td>47,1%</td>
<td>49,0%</td>
<td>+5,4%</td>
<td>+5,4%</td>
<td>6,0%</td>
<td>10,2%</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>244 (27,9%)</td>
<td>60,2%</td>
<td>59,8%</td>
<td>-2,6%</td>
<td>-2,9%</td>
<td>9,9%</td>
<td>4,9%</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>394 (45,1%)</td>
<td>45,1%</td>
<td>47,0%</td>
<td>+2,7%</td>
<td>+2,3%</td>
<td>5,6%</td>
<td>6,8%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>160 (18,3%)</td>
<td>56,3%</td>
<td>54,4%</td>
<td>-0,7%</td>
<td>-0,8%</td>
<td>8,0%</td>
<td>5,4%</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>24 (2,7%)</td>
<td>62,5%</td>
<td>66,7%</td>
<td>+2,1%</td>
<td>+2,3%</td>
<td>7,4%</td>
<td>8,7%</td>
</tr>
</tbody>
</table>

The results of a comparison between the preference of forecast approach per cluster in 2012 and 2013 are provided in table 5.6. The results show that from 2012 to 2013, the preferred forecast approach has changed for 46% of all clusters. Of all 460 TD2-clusters in 2012, 37% changed to a Direct-cluster. Conversely, 56% of all 413 Direct-clusters in 2012 is best forecasted by the TD2-approach in 2013. The share of clusters of which the preferred forecast approach shifts is relatively equal among all segments (not including segment 5 and 6 since the amount of clusters in this segment is very low).

Table 5.6 – Stability of preferred forecast approach from 2012 to 2013

<table>
<thead>
<tr>
<th>Segment</th>
<th>N</th>
<th>Total changed</th>
<th># TD2-clusters 2012</th>
<th>2012-2013 From TD2 to Direct</th>
<th># Direct-clusters 2012</th>
<th>2012-2013 From Direct to TD2</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>873</td>
<td>46%</td>
<td>460</td>
<td>37%</td>
<td>413</td>
<td>56%</td>
</tr>
<tr>
<td>1</td>
<td>576</td>
<td>45%</td>
<td>298</td>
<td>34%</td>
<td>278</td>
<td>58%</td>
</tr>
<tr>
<td>2</td>
<td>230</td>
<td>47%</td>
<td>112</td>
<td>46%</td>
<td>118</td>
<td>47%</td>
</tr>
<tr>
<td>3</td>
<td>49</td>
<td>43%</td>
<td>37</td>
<td>30%</td>
<td>12</td>
<td>83%</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>38%</td>
<td>12</td>
<td>33%</td>
<td>4</td>
<td>50%</td>
</tr>
</tbody>
</table>
Now we have analyzed that the preference of forecast approach remains stable from 2012 to 2013 for only 54% of the clusters, we intend to simulate the consequences of three different scenarios. These scenarios reflect different decision alternatives at the end of 2012 to forecast 2013:

- **Scenario 1**: All clusters are forecasted by the direct disaggregate approach.
- **Scenario 2**: All clusters are forecasted by the TD2-approach.
- **Scenario 3**: The forecast approach per cluster is equal to the optimal approach in 2012 (similar to the functionality of a software application that selects a forecast approach corresponding to the optimal forecast approach for a predefined historical sales period).

The results of this scenario analysis are presented in table 5.7. Scenarios 2 and 3 outperform scenario 1 in general and for every product group. For the majority of groups, scenario 2 shows to be best. The data in the last three columns show the average of the average MAD per cluster of the three scenarios compared to the average of the average MAD per clusters of the ideal situation (if the optimal HF-approach per cluster would have been used).

### Table 5.7 – Scenario analysis 2013

<table>
<thead>
<tr>
<th>Group</th>
<th>Label</th>
<th>N (% of total sample)</th>
<th>Av. MAD Scen. 2 comp. to Scen. 1</th>
<th>Av. MAD Scen. 3 comp. to Scen. 1</th>
<th>Av. MAD Scen. 1 comp. to ideal</th>
<th>Av. MAD Scen. 2 comp. to ideal</th>
<th>Av. MAD Scen. 3 comp. to ideal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>Total</td>
<td>873 (100%)</td>
<td>-2.0%</td>
<td>-1.9%</td>
<td>+8.9%</td>
<td>+5.1%</td>
<td>+5.9%</td>
</tr>
<tr>
<td>BA</td>
<td>Acc.</td>
<td>146 (16.7%)</td>
<td>-2.6%</td>
<td>-2.5%</td>
<td>+9.2%</td>
<td>+5.1%</td>
<td>+5.9%</td>
</tr>
<tr>
<td></td>
<td>Con.</td>
<td>645 (73.9%)</td>
<td>-1.7%</td>
<td>-1.6%</td>
<td>+8.8%</td>
<td>+5.3%</td>
<td>+6.1%</td>
</tr>
<tr>
<td></td>
<td>Spares</td>
<td>73 (8.4%)</td>
<td>-3.7%</td>
<td>-2.9%</td>
<td>+8.6%</td>
<td>+3.4%</td>
<td>+5.1%</td>
</tr>
<tr>
<td></td>
<td>Tools</td>
<td>9 (1.0%)</td>
<td>-4.1%</td>
<td>-4.9%</td>
<td>+9.4%</td>
<td>+3.2%</td>
<td>+2.4%</td>
</tr>
<tr>
<td>TABCD</td>
<td>T</td>
<td>51 (5.8%)</td>
<td>-4.5%</td>
<td>-2.7%</td>
<td>+9.8%</td>
<td>+3.2%</td>
<td>+6.4%</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>244 (27.9%)</td>
<td>-3.2%</td>
<td>-2.8%</td>
<td>+9.3%</td>
<td>+4.2%</td>
<td>+5.0%</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>394 (45.1%)</td>
<td>-1.0%</td>
<td>-0.9%</td>
<td>+8.7%</td>
<td>+6.0%</td>
<td>+7.0%</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>160 (18.3%)</td>
<td>-1.9%</td>
<td>-2.0%</td>
<td>+8.0%</td>
<td>+4.7%</td>
<td>+5.1%</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>24 (2.7%)</td>
<td>-3.4%</td>
<td>-6.1%</td>
<td>+10.9%</td>
<td>+5.1%</td>
<td>+2.8%</td>
</tr>
</tbody>
</table>

### 5.3.4 Discussion

**Overall results** - The results of the analysis apply to clusters of two items and are related to the specific set of forecasting techniques and performance measures used. Although the analysis of Dangerfield and Morris (1992) was similar in the sense that they made use of clusters of two items and the Winters’ exponential smoothing model, their results indicated that the direct disaggregate approach was favored for the majority of forecasts (74%). The fact that our results deviate may be the result of the difference in performance measures and amount of historical data used to estimate smoothing constants. Dangerfield and Morris (1992) used the relative error metric MAPE instead of MAD. They showed additionally that the use of MSE would have resulted in a share of 34% for which the direct disaggregate approach is favored. A theory to explain this difference might be that the direct disaggregate approach results in relatively more underforecast errors than the TD-approach. The MAPE is more forgiving in these situations than a scale-dependent error (such as MAD or MSE), since the error is divided by the high demand value. Next to the difference in performance measure, Dangerfield and Morris (1992) made use of more historical data to optimize smoothing constants. Dangerfield and Morris (1992) show that a worse fit of models slightly favors the TD-approach.
Criteria that determine effectiveness approaches - The clusters used in this analysis showed not to be uniformly distributed along the parameters of correlation and average volatility. Since some segments contained a very low amount of clusters, no statistically valid conclusions can be drawn on the individual and joint influence of the parameters on the performance of the forecast approaches. The high share of TD-clusters in segments (3) and (4) in both 2012 and 2013 however suggest that the TD-approach is more effective positively correlated clusters. This is in line with the findings of Dangerfield and Morris (1992). We remark that correlation might not be the best measure to quantify the homogeneity of the underlying pattern of a time series, since it is influenced by noise. Indicators of homogeneity which are less affected by noise may better explain the performance of the TD-approach.

The average volatility of a cluster shows not to have a specific influence on the performance of the approaches. The fact that no specific effect is found may be the related to the measure that is used in this analysis to indicate the volatility of time series. The average coefficient of variance of deseasonalized and detrended time series may not be appropriate since it is not related to the proportion of the coefficient of variance of the subaggregate variables nor to the item's share of sales in the cluster. An additional analysis for uncorrelated clusters in which we differentiated between clusters of which the subaggregate variables are both lowly volatile, both highly volatile or one lowly and one highly volatile still showed no specific relation between volatility and the performance of HF-approaches either. The results are presented for the situations in which the amount of sales of subaggregate variables is fairly equal (table D.2, Appendix D) and unequal (table D.3, Appendix D). The lack of a relation between volatility and the performance of HF-approaches, could be the result of the fact that the sample does not include sporadic demand; Van Wanrooij (2012) shows that sporadic demand reflects more than 50% of the highly volatile items (CoV>1.5).

Stability - The analysis indicated that the optimal forecast approach changed for 46% of all clusters from 2012 to 2013. Since only one periodic change is considered, results cannot be generalized. In order to obtain a better indication of the stability, more periodic changes should be analyzed.

The share of clusters of which the optimal forecast approach changes from 2012 to 2013 is roughly equal per segment. The results also indicate that the difference in share of TD-clusters remains relatively equal throughout 2012-2013. Again, we remark that the single measurement of stability does not provide enough evidence to ground any valid general conclusions.

5.3.5 Conclusions and recommendations

In this analysis it has been aimed to obtain insights on the effect of time series characteristics on the effectiveness of HF-approaches. In addition, it has been analyzed to what extent the preference for an approach is stable over time. We will formulate concise conclusions and recommendations for the research field of Hierarchical Forecasting and Hilti AG.

Research field of Hierarchical Forecasting

1. Given the specifications of the analysis, the TD-approach slightly outperformed the direct disaggregate approach in 2013 while the performances were fairly equal in 2012 – Since the results showed to be (highly) dependent on the specifications of the test, the conclusion has a low degree of generalizability.

Extend the external validity of the results by:
- Enhancing cluster sizes
- Testing with different forecast techniques
- Testing with different performance measures (e.g. percentage error metrics, relative error metrics)
2. For both 2012 and 2013, the TD-approach seemed to be more effective for positively correlated clusters than uncorrelated clusters – Due to a limited sample size, this statement is however not statistically valid.

Further examine the effect of homogeneity on the performance of HF-approaches by:
- Testing with a sample that is uniformly distributed along the correlation of clusters
- Testing with different measures of homogeneity that are less affected by noise

3. Volatility seemed not to have influence on the performance of HF-approaches for uncorrelated clusters with an average CoV of deseasonalized and detrended time series between 0 and 1 – the fact that volatility did not seem to improve the relative performance of the TD-approach may be due to the fact that the sample only included time series with a relatively low volatility (CoV<1,5).

Further examine the effect of volatility on the performance of HF-approaches by:
- Including sporadic demand in the sample, as this contains a high share of highly volatile time series

Hilti AG

1. Due to the fact that there are no general criteria/guidelines that assess the optimal forecast level, only data-analysis can indicate the optimality of forecast levels - It is confirmed once more that the preference for an HF-approach is related to the forecast models and performance measures used and cannot (yet) be determined by the characteristics of time series. Data-analysis is therefore required to assess the optimal forecast level for a certain period of time.

2. It seems beneficial to be aware of the optimal forecast level per cluster in advance - given the specifications of the test, the overall forecast accuracy improvement of using the optimal approach per cluster is 6.3%-8.9% compared to using one and the same HF-approach for all clusters

3. The preference of the forecast level per cluster could be unstable over time - a comparison of the results of 2012 and 2013 showed that, given the specifications of the test, the optimal HF-approach changed for 46% of the items. As a consequence, the results of using software assistance showed to be overall less accurate than the best single approach for all items, but still better than the worst single approach. Software assistance could therefore be beneficial, but does not have to result in the highest accuracy.

Experiment with software assistance for executing data-analysis in order to define the optimal forecast level per cluster in advance!
Warning: Do this with caution, since the optimal level may differ over time!
5.4 Process & Control

Based on the AS-IS process state, the idealized process design and the identified gaps, a conceptual TO-BE process design has been created (figure 5.6). The following design decisions have been made in order to partially close the gaps:

1. Assign the full responsibility for forecasting for integrated markets to logistic regions
2. Review the statistical forecasting processes based on the idealized design
3. Review the consensus forecasting processes based on the idealized design
4. Safety stock and replenishment methods (‘MRP-types’) should be regarded as tactical parameters which are updated less frequent (e.g. quarterly) than what is currently done
5. Centralize the responsibility of the alignment of demand and supply
6. Forecast accuracy should be measured for all individual forecasting activities

In the following paragraphs, we will structurally provide arguments for the six design decisions.

5.4.1 Assign responsibility for forecasting for integrated markets to logistic regions

The first design decision assigns the full responsibility of forecasting all items, replenished in integrated markets, to MM in the logistic regions. This implies that HAG MM will not have the (partial) responsibility anymore of providing forecasts for ROP-items and sporadic items for the integrated markets. HAG MM will however still be responsible for the forecasting process for non-integrated markets that are served directly from HAG warehouses.

Motivation for design decision - The motivations for this design decision originate from seven identified gaps in the corporate forecasting process:

1. The weekly execution of the process is inefficient
2. The forecasts are based on replenishment history, which is not appropriate for statistical forecasting as it is influenced by lot sizes etcetera.
3. Domain knowledge may not always be properly integrated, due to sometimes poor communication between the logistic regions and headquarters

Due to the continuously changing history as a result of changing MRP-types:

4. There is no structural history correction
5. It is difficult to properly identify patterns in historical data and select forecast models
6. It is complex to preventively review the quality of forecasts
7. There is no forecast accuracy measurement. This impedes the improvement of the process.

Shifting the responsibility of the forecast generation for all products in a logistic region to the region itself will resolve all the aforementioned problems.

Disproving arguments in favor of current setup - Now we have provided an overview of the discrepancies of the current corporate forecasting process, the motivations for the setup of the AS-IS process are analyzed. Nowadays, there exist three main arguments in favor of the (weekly) execution of forecasting in the headquarters:

Argument 1: Aggregated forecasts (as executed by HAG MM) for sporadic items (and ROP-items) are assumed to result in higher forecast accuracy on the aggregate level.

Argument 2: The desire to be able to weekly update MRP-types creates the need for corporate forecasting to generate forecasts on a weekly basis.

Argument 3: Non-integrated markets do not perform any statistical forecasts themselves. Forecasting by HAG MM, based on the replenishment history, is said to be a proper solution for this problem.
**REDESIGNING A DEMAND PLANNING PROCESS IN AN INTEGRATED PLANNING ENVIRONMENT**

**Argument 1** is based on the expectancy of the achievement of higher forecast accuracy when sporadic items of different regions are forecasted on an aggregate level. In the idealized design (section 3.3), it has been argued that scientific literature has not provided any proof of the effectiveness of direct aggregate forecasts over bottom-up forecasts for the forecast generation on the aggregate level. The results of Dunn et al. (1976) and Gordon et al. (1997) for instance disprove argument 1. Although there is no consistent scientific evidence to disprove argument 1, there is no consistent scientific evidence in favor of argument 1 either.

Currently, logistic regions can change MRP-types on a weekly basis. In order to have up-to-date corporate forecasts that comply with this frequency, the corporate forecasts are generated weekly as well (Argument 2). We argue that the MRP-type of items should be regarded as a tactical parameter, of which should be updated monthly at most (for instance for INP and phase-out products). For all active products, we suggest to update MRP-types together with other tactical parameters on a quarterly basis (as recommended by Broft (2014) and Kreuwels (2014)).

Aggregated forecasts for non-integrated markets are currently performed by HAG MM based on the replenishment history (argument 3). Since the non-integrated markets cannot provide statistical forecasts themselves, the TO-BE design suggests maintaining the forecasting process for these markets in headquarters. Ideally, this is done based on the actual sales history. However, in reality it may be complex to obtain this information from agents and dealers in non-integrated markets. The only feasible solution would in this case be to forecast based on the replenishment history. One should bear in mind that this could lead to higher demand variability through the occurrence of the bullwhip effect. Since there is no consistent evidence claiming that Bottom-Up forecast approach outperforms the direct aggregate forecast approach, it is suggested to keep forecasting the items for non-integrated markets at an aggregate (global) level.

**Consequences** - Due to the numerous gaps in the corporate forecasting process and the fact that the two main arguments in favor of corporately forecasting items for integrated markets do not hold, we argue that the design decision to assign the full responsibility of forecasting all items, replenished by logistic regions, to the regions themselves. The MRP-type will therefore not indicate forecast responsibility anymore. However, MRP-types still indicate the way an item is replenished and thus has an impact on the stock positioning, the safety stock determination and replenishment method. This implies that for instance the reorder point replenishment method can still be maintained in the suggested redesign.

Although the responsibility of forecasting shifts, HAG MM still has a role in the alignment of the consensus forecast and the (financial) sales rolling forecast at a global level. The current functionality of the sales report should still be maintained in order to analyze per BU to what extent the (financial) sales rolling forecast and the consensus forecast are in line. For the generation of rolling forecasts, it is recommended to make to greater extent use of regional market information instead of purely relying on the insights of Marketing HQ on the global market.

**5.4.2 Review statistical forecasting based on idealized design**

**History correction** - For the history correction activity, a new set of alerts should be installed to identify outliers in the historical demand pattern. In this, we define an outlier as any sales value that significantly deviates from other sales values in the time series. Only in case the demand planner is completely aware of the cause of the outlier and the fact that it is non-recurrent, the value could be replace by for instance the four period moving average. If not, it is suggested to replace the value with the threshold value that is used for the identification of outliers.

**Modification forecast models and parameters** - Prior to the APO DP run, statistical forecast models and their parameters should be modified based on the forecast accuracy of the statistical
forecasts (without any integration of promotional or judgmental forecasts). The model modification process could be supported by the use of alerts, which indicate high forecast inaccuracy. In addition, a tracking signal (equation 3.2) should be used which indicates the forecast bias over time.

Selection of forecast models - Statistical forecast techniques can only provide accurate results when they are applied on sufficient historical sales data. After an INP, only 12 months of data is available. This is insufficient to properly capture seasonal patterns. It is therefore recommended to make optimal use of the demand pattern of an item's predecessor or to identify expected similar patterns of other products (use of structured analogies) to properly select forecast models.

The selection of forecast models is currently done based on the knowledge and experience of demand planners. Since human beings often tend to respond to randomness in data as if it is a pattern, the selection process is prone to errors. Data-analysis should assess whether the automatic model selection functionality within APO is better capable of selecting forecast models.

When reviewing statistical forecasts, demand planners should be aware that the forecasts are used as mid-term and short-term forecasts. Demand planners should therefore not only aim to achieve high forecast accuracy on the short-term, but should focus as well on the course of the forecasts on the mid-term (at least a rolling horizon 12 months). This stresses the need of additional preventive alerts that analyze the course of forecasted demand on the mid-term.

Integration of events - After the statistical forecast run, promotional forecasts can be communicated with the demand planners on a daily basis. Statistical forecasts can be manual adapted based upon these projected promotional sales. In case the promotional forecasts concern (a) time bucket(s) within the horizon of the master production schedule, the demand planners should align with the Supply Chain Specialist Team (SCST) on the supply potential.

5.4.3 Review consensus forecasting based on idealized design

Sales planning - On a monthly basis, Regional Marketing should generate an official sales plan based on the in-depth market knowledge and the alignment with Global Marketing. Ideally, the forecasts are expressed in sales quantities for a rolling horizon of 18 months. In practice, Regional Marketing is only able to provide forecasts expressed in sales value. Whereas these projections are currently provided until the end of the current year, it should be aimed to provide forecasts for a rolling horizon of at least 15 months (corresponding to the horizon of the tactical level).

(Pre-)Consensus meetings - In the pre-SFI meeting, Regional Marketing is supposed to share the assumptions which they have made to create their judgmental forecasts. This meeting does not have to be an official gathering of people: the main purpose is to share the assumptions underlying the judgmental forecasts prior to the SFI-meeting in order to increase the quality and efficiency of the discussions in the SFI-meeting. The pre-SFI could also result in initial forecast alignment proposals already.

SFI-meetings should be organized on a monthly basis. All teams that generate or influence forecasts to a certain extent need to be involved (Materials Management and Marketing). Teams should aim to discuss the assumptions underlying the height of the forecasts. Once an agreement on these assumptions is reached, it is a relatively easy task to align on the height of the consensus forecast values. To focus of the SFI-meetings should be on a rolling horizon of 15-18 months to provide good input for the aggregate planning activities on the tactical level (and operational level). Statistical forecasts (plus promotions) need to be aggregated to the levels at which sales plan projections are provided to properly align both types of forecasts. Additionally, the statistical forecasts (plus promotions) need to be expressed in sales value. It is of paramount importance that this conversion is executed correctly.
**Figure 5.6 – TO-BE monthly demand planning process**
The results of the SFI-meetings are the unconstrained consensus forecasts expressed in sales value at various levels of aggregation in monthly buckets. The results need to be converted and disaggregated to provide demand forecast information at the level required for supply planning. In order to do this correctly, Regional Marketing could be involved to greater extent in case product managers have more specific knowledge. If not, the disaggregation can be structured by doing it based on an item’s average share in the consolidated sales over a certain historical period. This period should be at least 12 months to incorporate seasonal influences.

5.4.4 Lower the updating frequency of Safety stock and MRP-types

Safety stock values and MRP-types are regarded as tactical parameters. Broft (2014) and Kreuwels (2014) suggest updating tactical parameters on a quarterly basis. Exceptional cases are items which are phased-in or out. For these items, a monthly update of the parameters may be required. Figure 5.6 shows that the responsibility of the update of these parameters is split between Regional MM and HAG MM. A more centralized alternative is provided by Broft (2014) and Kreuwels (2014), which recommend the tactical parameters to be updated by BU-Management.

5.4.5 Centralize the responsibility for alignment demand and supply

Broft (2014) and Kreuwels (2014) recommend installing a Supply Chain Specialist Team (SCST) which operates on the tactical and operational level with the objective of aligning supply and demand. The SCST is responsible for creating the preliminary delivery and production plans at the tactical level, which serve as an input for the Supply Chain based Sales and Operations Plan. In addition, SCST is responsible for the alignment of supply and demand at the operational level. Different than the design of Broft (2014) and Kreuwels (2014), our recommendation is to base the supply plans on the unconstrained consensus forecasts, which are the result from the alignment of the statistical forecasts and sales planning. By making demand planning (statistical forecasting and sales planning) one responsibility of multiple teams with different objectives, one is more likely to control the bias in forecasting.

Instead of HAG MM, the SCST should now be responsible for sending the demand previews to (allied) suppliers.

5.4.6 Measure accuracy for all individual forecasting activities

In order to identify the causes of forecast inaccuracy in the demand planning process, it is of paramount importance to objectively measure forecast accuracy of all individual activities that generate or in a certain way influence the height of forecasts. In figure 5.6, the moments at which forecast accuracy needs to be measures are marked by the *-symbol. In order to measure forecast accuracy on various moment, the system should support the storage of forecast information at all those moments in time. In calculating the forecast accuracy metric, the lag of each market should be taken into account.

5.5 Control

In this section we will discuss design decisions for the redesign of the demand planning control mechanism. We will start with rephrasing the different objectives of demand planning control and the specific requirements they set for forecast accuracy metrics. Based on these requirements, a set of metrics is proposed that effectively services these objectives.
5.5.1 Basics of a control mechanism

In order to structurally (re)design the control mechanism, we start with rephrasing the objectives that are serviced through measuring the accuracy of demand forecasts in an integrated planning environment. Two key objectives are:

Objective 1: Improving the demand planning performance - The idealized design indicated that the results of the metric(s) serving this objective should be summarisable, intuitively understandable and scale-independent.

Objective 2: Quantifying the (financial) impacts of demand uncertainty to support supply planning decisions - The idealized design indicates that metric(s) for resource allocation decisions should be non-absolute, intuitively interpretable and scale-independent.

The gap-analysis revealed that Hilti currently uses metrics for control-objective 1 that does not satisfy all requirements, while there is no metric at all supporting resource allocation decisions (objective 2). In order to partially resolve the identified gaps, we will analyze which set of metrics is most appropriate to service both objectives of demand planning control. Per objective, we will discuss the arguments for the selection of the metric, how to use it and how to deal with its disadvantages.

Any forecast accuracy metric should measure accuracy based on the actual demand of a certain month and the forecasts generated \( x \) months ahead of that month. We refer to the value of \( x \) as the forecast lag. The idealized design suggests that length of this forecast lag depends on the replenishment lead time.

5.5.2 Metric selection for control objective 1

Analysis - For objective 1, we need an accuracy metric that is summarisable (absolute/ squared), scale independent and intuitively interpretable. Based on the literature review of Martens (2014b), three potential metrics have been selected. The MAPE (Mean Absolute Percentage Error) displays the average absolute forecast error as a percentage of the actual demand. The MASE (Mean Absolute Scaled Error) yields the (average) absolute forecast error, scaled by the in-sample mean absolute error of the naïve forecast method. The MAD/Mean (Mean Absolute Deviation /Mean demand) calculates the (average) absolute forecast error and scales it over the in-sample mean demand. Hyndman (2006) provides a more detailed description of the functionality of the three metrics. The results of an analysis on the characteristics of the three metrics are presented in figure D.4 (Appendix D). This table also provides a more detailed definition of each metric.

Conclusion - Based on the results of the analysis, we select the MAPE as the most appropriate metric to service objective 1. Although the MASE performs well on most criteria, the difficulty to interpret the result of the metric makes it inappropriate. The MAD/Mean metric basically has the same functionality as MAPE. However, the fact that the absolute forecast error is scaled over the mean in-sample demand makes the results unreliable for non-stationary demand patterns.

The MAPE represents the mean of absolute forecast errors as a percentage of the actual demand at time \( t \). The definition of the absolute percentage error is provided by equation 5.11. The advantages of this metric are the facts that it is intuitively understandable and that it facilitates performance comparison of different forecasts. This is not only useful to compare the performance improvement of a demand planner, but also between demand planners or departments. Since the metric is most commonly used in industry, it also highly appropriate for benchmarking the performance against the forecast performances of other companies.
In addition, the metric facilitates performance review on a higher level of aggregation since its outputs are summarisable.

### 5.5.3 Use of metric for control objective 1

The results of the metric for the first objective should be provided to all organizations that generate or influence demand forecasts. In order to properly identify the causes of forecast inaccuracy, forecast accuracy should be measured for every forecasting activity that influences the height of demand forecasts. The moments at which forecasted values have to be stored are identified in section 5.4.6. In order to provide specific feedback on the forecast performance, the absolute percentage error should always be computed at the basic level at which demand information is provided to subsequent supply planning stages.

In order to use the MAPE-metric to review performance and set accuracy targets at a high level of aggregation, the following recommendations are provided:

1. Create target-groups based on the characteristics of demand patterns that influence the potential height of forecast accuracy
2. Proportionally aggregate individual APE-values within target-groups based on the importance of items for the company
3. Assess the height of targets per target-group based on the forecast lag and data-analysis

We will discuss these three instructions in further detail.

**1. Create target-groups based on the characteristics of demand patterns that influence the height of the potential forecast accuracy** - In the idealized design it is suggested that predictability of demand patterns depends on the controllability of items. This is indicated by the coefficient of variance of the order sizes. The potential height of forecast accuracy is lower for high coefficients of variance. The scale of CoV is already existing in Hilti (distinction between U, V and W items as depicted in figure B.2, Appendix B). As a consequence, it is recommended to:

**Set targets differently for U-, V- and W-items**

The MAPE-metric has difficulty in dealing with demand patterns that contain time buckets in which the actual demand is close or is equal to zero. In period where the actual demand is zero, MAPE will yield an error. We suggest the frequently used method of artificially replacing the demand values of zero by a value of one in order to overcome this problem. In situations where the actual demand is close to zero, MAPE will show inflated results. One can deal with the inflated MAPE-results in situations with an actual demand close to zero by defining special targets for items of which the demand pattern includes those situations (intermittent demand). One should be aware that the inflated accuracy results of the intermittent group cannot be fairly compared with the accuracy results for normal demand patterns. As a result, it is recommended to:

**Set targets based on the sales volume of items and make a clear distinction between targets for normal and intermittent demand**
REDESIGNING A DEMAND PLANNING PROCESS IN AN INTEGRATED PLANNING ENVIRONMENT

2. Proportionally aggregate individual APE-values within target-groups based on the importance of items for the company - While the normal MAPE simply averages multiple APE-values of separate forecasts, literature (for instance Kilger and Wagner, 2008) argues that the weighted MAPE (wMAPE) may better reflect the actual situation of the supply chain. The wMAPE (equation 5.12) proportionally averages separate APE-values assuming that the APE-value of a certain demand pattern is more important than another. Importance can be related to for instance sales volume, sales value and type of customer. This degree of importance is included through the use of weighting factors.

\[
wMAPE = \sum_{i=1}^{N} \sum_{t=1}^{T} \frac{|\varepsilon_{i,t}|}{D_{i,t}} \times 100\%
\]

, where

\[
w_i = \frac{x_i}{\sum_{i=1}^{N} x_i}
\]

In the previous paragraph, we recommended to set forecast accuracy targets amongst others based on sales volume. Since the sales quantities of all items in target-groups thus are roughly similar, item-importance can be incorporated by proportionally averaging APE-values based on sales value. Since it is easier to allocate resources and define the way of working based on a discrete scale than on a continuous scale of importance, it is recommended to make use of discrete weighting factors. This could be realized by assigning values for \(x_i\) (equation 5.12) per T,A,B,C,D-label, reflecting the importance of product-type. Should it be anyway desired to base the \(wMAPE\) on a continuous scale of sales value or sales volume, it is recommended to use equation 5.13.

\[
wMAPE = \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} |\varepsilon_{i,t}|}{\sum_{i=1}^{N} \sum_{t=1}^{T} D_{i,t}} \times 100\% \tag{5.13}
\]

Proportionally average APE-values based on a discrete scale of item-importance (e.g. sales value)!

3. Assess the height of targets per group based on the forecast lag and data-analysis - Academic literature does not provide any models to properly define the height of forecast accuracy targets. It is therefore suggested to analyze the current forecast accuracy values per target group and set challenging targets accordingly. In general, predictability of a demand pattern decreases for higher values of the coefficient of variance and lag of forecasting. Since the lag of forecasting is related to the lead times of markets, it is recommended to set forecast accuracy targets differently per market.

The wMAPE-metric provides feedback on the performance of demand planners in terms of the absolute percentage forecast error. In case it is desired to receive this feedback in terms of forecast accuracy instead of forecast error (provided by wMAPE), the wMAPA can be used (Kilger & Wagner, 2008). This variation of the wMAPE is only different in the sense that it transforms the absolute percentage error into absolute percentage accuracy according to equation 5.14. Whereas the results of the wMAPA may be better interpretable, they are less representative since all absolute errors exceeding 100% are presented as 0% accuracy.

\[
APA_{i,t} = \max\{100\% - APE_{i,t}; 0\%\} \tag{5.14}
\]

Set forecast accuracy targets differently per market and assess the height based on data-analysis!
5.5.4 Metric selection and use for control objective 2

Metrics for objective 2 should support resource allocation decisions and inventory investment decisions on the tactical and operational level. Since safety stock calculation is out of the scope of this project, we only discuss the metric selection and use for resource allocation decisions.

**Resource allocation decisions** – To support decisions on resource allocation, a metric is required to be non-absolute and intuitively understandable. To easily relate forecast errors to the resource consumption in terms of hours of labor or financial value, we argued that scale-dependency is useful as well. Similar to the metric selection process for the first objective, percentage error metrics or scale independent metrics are said to best comply with these requirements. For the metric selection for resource allocation decisions, the non-absolute versions are regarded. Out of these, we select the MPE-metric (Mean Percentage Error) as being the most appropriate since its results are intuitively understandable. The results of the metric are obtained through equation 5.15.

$$MPE_I = \frac{1}{T} \sum_{t=1}^{T} \frac{\varepsilon_{I,t}}{F_{I,t}} \times 100\%$$  \hspace{1cm} (5.15)

where:

- $I$ is a set of products \(\{i=1,\ldots,N\}\) based on which resource allocation decisions are taken.

Capacity decisions are typically made based on demand forecasts. It is for this reason that the forecast error in equation 5.15 is divided by the forecast value. The results indicate the size of shortage or excess capacity over a certain time period. By averaging the results over a time period $T$, one captures the bias (consistent under- or overcapacity over time). An example of the use of the MPE-metric is provided in explanation box 5.1.

In order to make properly use of the MPE-metric in capacity decision making, the metric should be applied on the levels of aggregation as suggested in table 3.1. The metric can be applied on the tactical and operational level in order to analyze the quality of capacity planning in the past. Results of the MPE-metric should be provided to the organizations that are responsible for resource allocation decisions. It should be noted that results can never be summarized into one value!

**Use the MPE-metric to support resource allocation decisions and apply the metric on the level of aggregation at which those decisions are taken!**

**Explanation box 5.1 - Example of use of MPE-metric**

Consider the set of items in table 5.8. Assume that for this set of items a production or distribution capacity planning decision needs to be made. Based on the forecasts, it is decided to attribute capacity for 100 products. The actual demand in the end turns out to be 80 products. The MPE-value displays 20% overcapacity. In case the MPE-metric shows consistent overcapacity over time, capacity planners might consider adjusting the planned capacity based on this error.

**Table 5.8 – Example of functionality of forecast accuracy metric objective 2**

<table>
<thead>
<tr>
<th>Item</th>
<th>$D_t$</th>
<th>$F_t$</th>
<th>$\varepsilon_t$</th>
<th>MPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>30</td>
<td>-20</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>15</td>
<td>+5</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>50</td>
<td>-10</td>
<td>-20%</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>5</td>
<td>+5</td>
<td></td>
</tr>
<tr>
<td>Sum</td>
<td>80</td>
<td>100</td>
<td>-20</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 6

Recommendations on implementation

In order to close the gap between the TO-BE design and the actual execution, concise recommendations on the implementation are provided.

6.1 Overall recommendations on implementation

In chapter 1, this project has been classified as the challenge of tackling a complex problem in a pluralistic organizational context. Jackson (2003) states that the measures of success for this type of projects are:

- **Effectiveness** - the extent to which achievements are in line with the objectives
- **Elegance** - the extent to which stakeholders are enthusiastic about the reform and agree with about the outcomes of the project

Due to the fact that the available time was limited, the progress of the project so far is mainly related to the first measure. Since the aspect of 'elegance' is however at least as important for a successful implementation, Hilti is recommended to pay further attention to achieving the acceptance of the main stakeholders. In order to ensure this acceptance, the stakeholder analysis performed in chapter 1 should be elaborated in further detail. This analysis would enable to create appropriate strategic plans in order to convince opposing stakeholders with high power for certain design decision or integrate correct trade-offs of opposing interests in the redesign. This stakeholder analysis should be executed per component of demand planning in three phases:

1. **The identification of stakeholders and their interests**
2. **The categorization of stakeholders based on their interest in the project and level of power**
3. **The creation of a strategic plan to**
   a. Approach different types of stakeholders during the further redesign and implementation
   b. Align opposing interests of stakeholders with high power or integrate correct trade-offs of opposing interests in the redesign.

In the remainder of this chapter we will provide concise recommendations on the implementation of design decisions per component of demand planning.

6.2 Recommendations regarding demand planning structure

**Education** – Although this project does not provide any specific guidelines on determining the optimal level of forecasting, the presented information is valuable in the sense that it provides clarity on available scientific knowledge on hierarchical forecasting. This information is recommended to be spread across different departments in order to prevent the arousal of incentives for forecasting at higher/lower levels which are based on erroneous assumptions. In table 6.1 the most frequently overheard assumptions are listed and related (scientific) arguments and evidence is provided for their validation.

**IT-investment and training** – The potential benefits of software assistance on the selection of optimal forecast levels per cluster should be tested through a pilot in a logistic region. This requires investment in a software application and its assistance as well as investment in the training of demand planners to properly make use of the software application.
Table 6.1 – Frequently overheard assumptions at Hilti

<table>
<thead>
<tr>
<th>Assumption</th>
<th>Evidence and arguments</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘Forecasting at a higher level increases forecast accuracy’</td>
<td>Although forecasting at a higher level may cancel out random distortions in low level sales data, valuable pattern information of the low level data is lost. This could make the aggregated data difficult to forecast. Empirical research showed that there is no consistent evidence in favor of forecasting at a higher level to increase accuracy at that high level or at the low level.</td>
</tr>
<tr>
<td>‘Forecasting at a higher level is more efficient’</td>
<td>Statistical forecasting at a higher level than the level of forecast information provided to supply planning yields the same amount of forecasted values as directly forecasting at this low level. The TD-approach however only applies one forecast technique at the high level instead of the use of multiple techniques at the low level. The amount of work of the initial selection and modification of forecast techniques would therefore be reduced by using the TD-approach. The TD-approach does however not increase the efficiency of the activity of reviewing forecast accuracy and preventive alerts, since forecasts accuracy (and preventive alerts) still needs to be reviewed at the basic level provided to supply planning (low level).</td>
</tr>
</tbody>
</table>

6.3 Recommendations regarding demand planning process

Resource planning – The suggested process design decisions in section 5.4.1 influence the workload of the regional forecasting process. A resource plan needs to be generated in order to express the impacts of the design decision on the resources of personnel, money, equipment and data & knowledge.

Education and Training – In order to improve current statistical and consensus forecasting activities based on the idealized design, demand planners should be provided with educational and training-modules which focus on history correction using new alerts, forecast model modification through the use of a tracking signal, the focus in consensus meetings and create awareness of the use of forecasts for further supply planning decisions.

IT-support – The IT-system should support the creation of stable distribution and production plans through lowering the frequency of the net requirements calculation. Demand planners should be supported to communicate and update of forecasts within the horizon of the MPS with the SCST.

6.4 Recommendations regarding demand planning control

Education and Training – Stakeholders’ agreements is a hard requirement of forecast accuracy metrics, which should be satisfied in order to ensure the effectiveness of forecast accuracy metrics. This agreement may come after training and education on demand planning. The more effective use of metrics stimulates the improvement of the demand planning performance.

IT support – IT should facilitate the proper calculation of accuracy values incorporating the appropriate metrics, lags, target groups and ways of aggregation. In addition, the system should store forecasted values and measure accuracy values for each individual forecasting activity.

Use of labels – In order to ease use the use of the wMAPE- and MPE-metrics proposed in this project, the two metrics could be labeled differently. These labels should be formulated as such that they indicate directly what objective the metric services and how the results should be interpreted.
Chapter 7

Conclusions and Recommendations

The project is concluded by answering the defined design questions. In addition, recommendations are provided on future directions of improvement.

In the transition towards integrated operations planning, Hilti aims to redesign and align the main decision functions in a hierarchical planning structure in order to improve the overall performance of the internal end-to-end supply chain. In this master thesis, scientifically grounded guidelines have been provided on the redesign of the demand planning decision function on the tactical and operational level of the planning hierarchy. The project predominantly focused on the application of available theories and models in scientific literature on demand planning. In situations where these theories and models appeared to be deficient or not existing at all, quantitative and qualitative analyses have been conducted in order to ground design decisions and contribute to the research area.

7.1 Answering the design questions

We will provide concise conclusions by answering the design questions, underlying the TO-BE design of demand planning at Hilti.

1. In what way should demand planning be aligned with subsequent decisions functions on the tactical and operational level in order to improve the overall supply chain performance?

On the tactical level of the planning hierarchy, demand planning should provide mid-term demand forecasts with a horizon of at least 12 months (preferably 15-18 months). Since the plans on this tactical level are recommended to be updated monthly, demand information needs to be provided on a monthly basis. The forecasts are required on an aggregated product level in monthly buckets. The exact levels that are required depend on the decision function the forecasts serve. In accordance with the recommendations of Broft (2014) and Kreuwels (2014), the tactical parameters are suggested to be updated on a quarterly basis.

On the operational level, plans are updated on a weekly basis. Short-term forecasts are required to be item-specific and in weekly buckets. For transportation scheduling, the demand information needs to be location-specific as well. The horizon of the short-term forecasts should be as long as the longest cumulative lead time of items.

2. How should the demand planning process make use of Hierarchical Forecasting approaches in order to create an optimal balance between forecasting accuracy and process efficiency?

In order to provide all required demand information at different levels of aggregation with different horizons, scientific literature suggests the use of Hierarchical Demand Planning (HDP). HDP involves the concept of Hierarchical Forecasting (HF), which entails the application of statistical forecast techniques on pre-defined levels of aggregation and the use of various approaches to obtain demand information at other levels. In literature, there is no general consensus on the type of approach (TD or BU) that results in the highest forecast accuracy at either the aggregate or disaggregate level. There is no general consensus on the criteria that influence the performance of both approaches either. Since literature does not provide any general guidelines on the implementation of HF and the potential damage of selecting the
incorrect approach is not quantified, no scientifically grounded recommendations can be provided for the determination of the optimal forecast level.

In order to examine the influence of time series’ characteristics on the performance of HF-approaches and to assess the stability of the preference of HF-approach over time, a quantitative analysis has been conducted. The results indicated that it seems beneficial to be aware of the optimal forecast level per cluster in advance and that the preference of forecast level for a cluster could be unstable over time. Hilti is therefore recommended to experiment with software assistance in defining the optimal forecast level per cluster in advance with caution. In addition, the analysis showed that the correlation of subaggregate variables may influence the effectiveness of the TD-approach.

3. **How should demand planning process be redesigned from a process-perspective in order to efficiently create accurate demand plans?**

Due to the existence of various gaps in the corporate forecasting process, it is suggested to assign the full responsibility of forecasting ROP-items and sporadic items of integrated markets to the logistic regions which replenish those items. HAG MM will however maintain the responsibility of forecasting the items of non-integrated markets which are directly replenished by the HAG warehouses. A second recommendation is to review the statistical and consensus forecasting processes based on the idealized design. In monthly SFI-meetings, the underlying assumptions of the judgmental forecasts (forecasts by regional Marketing) should be discussed. The aligned consensus forecasts are subsequently used by a newly installed Supply Chain Specialist Team (SCST), which is responsible for the alignment between demand and supply.

Safety stock and MRP-types should be regarded as tactical parameters, which are suggested to be updated less frequent than what is currently done (e.g. quarterly).

In order to properly identify the causes of forecast inaccuracy, forecast accuracy should be measured objectively for all individual forecasting activities.

4. **In what way should the control-mechanism be redesigned in order to continuously improve demand planning performance and quantify demand uncertainty with the intention to support planning decisions in further stages of integrated planning?**

When redesigning the control-mechanism, it is of paramount importance to be aware of the two objectives it serves. These objectives each set different requirements for accuracy results.

In order to provide feedback on forecasts performance most effectively, a forecast accuracy metric is required which is summarisable, intuitively understandable and scale-independent. A qualitative analysis showed the wMAPE (weighted mean average percentage error) to be the most appropriate. The metric needs to be applied on the basic level of demand forecast information which is provided to subsequent supply planning activities. Forecast accuracy target-groups need to be defined based on the aspects that influence the potential height of accuracy: coefficient of variance of order sizes and sales volume. APE-values should to be proportionally averaged based on the importance of items (e.g. value). The height of targets need to be assessed through data-analysis and depend on the forecast lag. This implies that targets need to be set differently per logistic region.

In order to support resource allocation decisions, a forecast accuracy metric is required that is non-absolute, intuitively understandable and scale-independent. The MPE (mean percentage error) metric that computes the forecast error as a percentage of the forecasted value is suggested to be used. This metric should be applied on the level of aggregation at which resource allocation decisions are taken. By tracking accuracy values over time, consistent over- or undercapacity can be exposed.
7.2 Recommendations

7.2.1 Recommendations for Hilti

According to the methodology of Ackoff (1974; 1981), the TO-BE design phase is typically executed various times in order to continuously improve the performance of demand planning. In the future executions of the TO-BE design phase, it is suggested to pay attention to the following topics:

- **The acceptance of the process redesign by main stakeholders** - Although this issue has been stressed in chapter 6 of this document, we address this topic once more since it is highly important for the successful implementation of the current and future redesigns of demand planning. The high number of stakeholders and variety of interests increase the complexity of the design and implementation of a process redesign.

- **The potential benefits of software applications that iteratively calculate the expected optimal forecast level** – The results of the quantitative analysis in section 5.3 showed that it seems beneficial to be aware of the optimal forecast level per cluster. Since scientific literature does not provide any general guidelines (yet) on the decision making process, it could be beneficial to experiment with software applications that select forecast level based on the optimal forecast approach over a historical period. These applications do not necessarily have to result in the best accuracy performance, but may prevent achieving bad forecast performance through incorrectly selecting one single approach. It is however recommended to experiment with caution, since the results of the quantitative analysis indicated that the preference of forecast level could be unstable over time.

- **The potential benefits of the use of the automatic model selection option** – In the AS-IS analysis, it has been assessed that demand planners select forecast models based on their experience and knowledge. Since this judgmental selection is prone to errors, we recommend to further experiment with the use of the automatic model selection option in APO. The automatic model selection recommends forecasts models that show the highest accuracy over a certain historical period. Although we are not familiar with the exact functionality of the selection procedure, we warn for the risk of ‘over-fitting’. The model that results in the highest accuracy given a certain historical period could have obtained this result by best complying with the random noise of data instead of with the underlying pattern of a time series. Hence, this model does not have to provide the best future forecasts. By defining a relatively long ex-post horizon and not iteratively calculate the optimal model too frequent, one could reduce the negative effects of ‘over-fitting’.

- **The use of forecast accuracy results for the support of inventory investment decisions** – Since safety stock calculation is out of the scope of this project, it has not been assessed whether the provision of forecast accuracy results for this purpose best facilitates the calculation of optimal safety stocks. It should be further examined in what way forecast accuracy results should support the calculation of the optimal inventory hedge against demand uncertainty given the supply chain planning logic at Hilti. From a managerial perspective, it could also be useful to install a forecast accuracy metric that supports inventory investment decisions on an aggregated level.
7.2.2 Recommendations for the research field

Next to recommendations for Hilti, the analyses conducted in this project provide the following further recommendations for the research field of Hierarchical Forecasting:

- **Extend the external validity of the results of the quantitative analysis by testing with larger cluster sizes, different forecast models and various performance measures** – Given the specifications of the analysis, the TD-approach slightly outperformed the direct disaggregate approach for one year, while the performances were fairly equal for another year. However, a comparison with other researches showed that results are influenced by the specifications of the test. The generalizability of the conclusion should therefore be increased by varying those specifications.

- **Further examine the effect of homogeneity on the performance of HF-approaches by using uniformly distributed samples and testing with different measures of homogeneity** – Initial results showed that a positive correlation of subaggregate variables could increase the effectiveness of the TD-approach. More positively and negatively correlated cluster data are however required to draw a general conclusion.

- **Further examine the effect of volatility on the performance of HF-approaches by including sporadic/intermittent demand in the sample** – Sporadic/intermittent demand is generally highly volatile. By including this demand in the sample, a better distribution of data along the measure of volatility is obtained. This may provide more insights on the influence of volatility on the effectiveness of HF-approaches.
References


REDESIGNING A DEMAND PLANNING PROCESS IN AN INTEGRATED PLANNING ENVIRONMENT


## Appendix A

### Table A.1 – Stakeholder analysis phase 1

<table>
<thead>
<tr>
<th>Stakeholder</th>
<th>Interest</th>
</tr>
</thead>
</table>
| Hilti AG Executive Board, shareholders          | ▪ Sustainably improving (corporate values of):  
▪ Customer focus – high customer service level  
▪ People development  
▪ Productivity  
▪ Operational expenses  
▪ The reform should be in accordance with the corporate responsibility                                                                                                                                 |
| Global Logistics Materials Management (GLM) including Head of Global Logistics | ▪ Implementing one general and efficient demand planning process which is applicable on a multitude of products in wide range of markets  
▪ Launching one general demand planning process and workflow to better align the practices of both the Distribution side (MOs) as the Replenishment side of the network (production plants, allied suppliers, normal suppliers).  
▪ Implementing one standard KPI for forecast accuracy that accurately captures the quality and effectiveness of the forecasting process. A standard scientifically supported KPI could better facilitate benchmarking practices.  
▪ The reform should result in structured communication and enhanced transparency  
▪ Reducing operational and inventory holding costs as a result of higher forecast accuracy and better horizontal integration.  
▪ Improving sales performance by increasing the product availability and customer service levels  
▪ The new demand planning process should assure a fit for growth                                                                                                                                 |
| Process Competence Center (PCC)                 | ▪ Implementing one general and efficient forecasting procedure which is applicable on a multitude of products in wide range of markets.  
▪ Continuing with, implementing (or outsourcing) one general and efficient forecasting software application which can easily be integrated in the current IT-systems landscape and does not require much maintenance. This software application should also be attractive from a financial point of view.  
▪ Selecting a forecasting software application that is sustainable under the current market growth expectations in order to prevent a relatively rapid software change |
<table>
<thead>
<tr>
<th>Stakeholder</th>
<th>Interest</th>
</tr>
</thead>
</table>
| **Materials Management – Markets/Headquarters** | Operating according a simple and efficient forecasting procedure and workflow that  
- maximally reduces the workload of the demand planner. In this, workload is related to both the amount of work to perform one forecast (at any aggregation level) as the number of forecasts to be executed.  
- provides the highest possible forecast accuracy to optimize the combination of stock out and inventory holding costs.  
- allows for flexibly adapting to customer demand changes |
| **Marketing - Markets/Headquarters** | Operating according a simple and efficient forecasting procedure and workflow that  
- allows for flexibly adapting to customer demand changes  
- results in high customer service levels |
| **Final customer/ consumer** | The process redesign should result in an:  
- equal or reduced, but stabilized product lead time  
- equal or improved customer service level  
- equal or improved product quality  
- equal or reduced product cost |
| **Supplier - internal (Production plants) - external (Raw material and allied suppliers)** |  
- Low volatility in product demand – a stable production planning for a fixed period of time  
- Order confirmation as early as possible, leaving more time to adapt to this demand |
| **Academic Supervisors / Eindhoven University of Technology** |  
- Guiding a project to be successfully completed in order to increase the likelihood of obtaining (quantitatively and/or qualitatively) more graduation projects at Hilti AG. In this, a successful completion implies a project that is executed to the full satisfaction of Hilti AG and meets the standards of proper academic research design.  
- Gaining insights and inspiration for further academic research  
- The scope of the project should be defined as such that the project can be completed within a time frame of 21 weeks (excluding the preparation phase). |
### Figure A.1 – Stakeholder analysis phase 2

<table>
<thead>
<tr>
<th>Level of interest</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Power</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>II</td>
<td>III</td>
</tr>
</tbody>
</table>
|                   |   Final customers  
|                   |   Hilti AG Executive Board |   GLM  
|                   |   MM – Headquarters & Global 
|                   |   Marketing – Headquarters & Global 
|                   |   PCC  
|                   |   Internal suppliers  
|                   |   Academic supervisors |
| Low               | I   | IV   |
|                   |   External suppliers |
Appendix B

Figure B.1 – Item classification

Figure B.2 – Demand segmentation matrix
### Table B.1 – List of available statistical forecasting techniques

<table>
<thead>
<tr>
<th>Forecasting technique</th>
<th>Model name</th>
<th>Forecast strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Series Based</td>
<td>Constant Model</td>
<td>1st order Exponential Smoothing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Automatic Alpha Optimization (1st order)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Moving Average</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weighted Moving Average</td>
</tr>
<tr>
<td></td>
<td>Trend Model</td>
<td>1st Order Exponential Smoothing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2nd Order Exponential Smoothing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Automatic Alpha Optimization (2nd Order)</td>
</tr>
<tr>
<td></td>
<td>Seasonal Model</td>
<td>Seasonal Model based on Winter’s Method</td>
</tr>
<tr>
<td></td>
<td>Seasonal + Linear Regression</td>
<td>Seasonal Linear Regression</td>
</tr>
<tr>
<td></td>
<td>Median Method</td>
<td>Empirical Method</td>
</tr>
<tr>
<td></td>
<td>Seasonal Trend Model</td>
<td>Seasonal Linear Regression</td>
</tr>
<tr>
<td></td>
<td>Automatic model selection</td>
<td>Test for Constant, Trend, Seasonal, Seasonal Trend</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Test for Trend</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Test for Season</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Test for Trend and Season</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Seasonal Model and test for Trend</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trend Model and test for Season</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trend pattern with an additional test for seasonal pattern</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Test for Constant, Trend, Seasonal, Seasonal Trend</td>
</tr>
<tr>
<td></td>
<td>History</td>
<td>Copy History</td>
</tr>
<tr>
<td></td>
<td>Manual Forecasting</td>
<td>Based on planner's best judgment</td>
</tr>
<tr>
<td></td>
<td>Croston Method</td>
<td>Croston Method</td>
</tr>
<tr>
<td></td>
<td>Causal</td>
<td>Linear Regression</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple Linear Regression</td>
</tr>
</tbody>
</table>
Appendix C

Figure C.1 – Ideal design Hierarchical Planning Framework by Broft (2014) and Kreuwels (2014)
Appendix D

Explanation Box D1 – Initialization and updating procedures forecasting models

As explained in section 5.3.1, the initialization and parameter updating procedures of the models used in the analysis are adopted from the work of Silver et al. (1998). The procedures have been adjusted to the situation of forecasting the demand of one period ahead. We will firstly describe the procedures for the non-seasonal models, after which the procedures for the seasonal models will be discussed.

Non-seasonal models

Model with constant level (Simple exponential smoothing)

The forecast for period \( t+1 \) made in period \( t \) is defined as

\[
F_{t,t+1} = \hat{a}_t
\]

where

\[
\hat{a}_t = \hat{a}_{t-1} + \alpha (D_t - \hat{a}_{t-1})
\]

This updating formula is initialized by the determination of \( \hat{a}_0 \), which is equal to the average demand in the last 12 periods.

Models with damped or linear trend (Second order exponential smoothing)

The forecast for period \( t+1 \) made in period \( t \) is defined as

\[
F_{t,t+1} = \hat{a}_t + \phi \hat{b}_t
\]

where

\[
\hat{a}_t = \alpha_{HW} D_t + (1 - \alpha_{HW})(\hat{a}_{t-1} + \phi \hat{b}_{t-1})
\]

\[
\hat{b}_t = \beta_{HW}(\hat{a}_t - \hat{a}_{t-1}) + (1 - \beta_{HW})\phi \hat{b}_{t-1}
\]

These updating formulas are initialized by the determination of \( \hat{a}_0 \) and \( \hat{b}_0 \), which are defined as

\[
\hat{a}_0 = \frac{6}{n(n+1)} \sum_{t} tD_t + \frac{2(2n - 1)}{n(n+1)} \sum_{t} D_t
\]

\[
\hat{b}_0 = \frac{12}{n(n^2 - 1)} \sum_{t} tD_t + \frac{6}{n(n+1)} \sum_{t} D_t
\]

where

\( n \) the number of historical time periods used
Explanation Box D1 (cont’d) – Initialization and updating procedures forecasting models

**Seasonal models**
*Model with constant level (φ=0), model with damped trend (φ=[0.05;0.95]), model with linear trend (φ=1)*

The forecast for period $t+1$ made in period $t$ is defined as

$$F_{t,t+1} = (\hat{a}_t + \phi \hat{b}_t)\hat{S}_{t+1-p}$$

where

$$\hat{a}_t = \alpha_{HW}(D_t/\hat{S}_{t-p}) + (1 - \alpha_{HW})(\hat{a}_{t-1} + \phi \hat{b}_{t-1})$$

$$\hat{b}_t = \beta_{HW}(\hat{a}_t - \hat{a}_{t-1}) + (1 - \beta_{HW})\phi \hat{b}_{t-1}$$

$$\hat{S}_t = \gamma_{HW}(D_t/\hat{a}_t) + (1 - \gamma_{HW})\hat{S}_{t-p}$$

During the updating of the seasonal factors $\hat{S}_t$, the factors are renormalized over the last $p$ months of data. In this, $p$ refers to the number of periods in one seasonal cycle. In the analysis, the value of 12 months is used. This updating formulas are initialized by the determination of $\hat{a}_0$ and $\hat{b}_0$, which are equal to the initialization formulas of the non-seasonal models. The equations should however be applied on the deseasonalized time series. This is done by correcting this time series by the initial seasonal factors, which are computed through the centered 12-period moving average and subsequently normalized. This procedure is described in further detail by Silver et al. (1998).
For the computation of the correlation of clusters containing more than two subaggregate variables, formula 5.10 is not applicable. Instead, in section 5.3.1 it is suggested to take the average of the multiple correlation coefficients of all subaggregate variables.

For a cluster with three variables, the multiple correlation coefficient for each variable is obtained as:

\[
R_{y;x_1,x_2} = \frac{\sqrt{\rho_{y,x_1}^2 + \rho_{y,x_2}^2 - 2\rho_{y,x_1}\rho_{y,x_2}\rho_{x_1,x_2}}}{\sqrt{1 - \rho_{x_1,x_2}^2}}
\]

For larger clusters, matrices need to be used for the computations (Abdi, 2007). For this, we start with creating matrix X based on all dependent variables \((j=1,\ldots,J)\). Additionally, we define the independent variable as column vector y.

\[
X = \begin{bmatrix}
1 & x_{1,1} & \cdots & x_{1,J} \\
\vdots & \vdots & \ddots & \vdots \\
1 & x_{N,1} & \cdots & x_{N,J}
\end{bmatrix} \quad y = \begin{bmatrix}
y_1 \\
\vdots \\
y_N
\end{bmatrix}
\]

We define b as:

\[
b = (X^TX)^{-1}X^Ty
\]

The regression sum of squares is obtained as:

\[
SS_{\text{regression}} = b^TX^Ty - \frac{1}{N}(1^T y)^2
\]

And the total sum of squares (regression and error sum of squares) is:

\[
SS_{\text{total}} = y^Ty - \frac{1}{N}(1^T y)^2
\]

The multiple correlation coefficient is then obtained as:

\[
R_{y;x_1,\ldots,x_J} = \sqrt{\frac{SS_{\text{regression}}}{SS_{\text{total}}}}
\]
Figure D.1 – Distribution of clusters based on correlation and average volatility, 2012

Figure D.2 – Share of TD2-clusters per range of correlation, 2012
Figure D.3 – Share of TD2-clusters per range of average CoV, 2012

Table D.1 – Combined influence correlation and volatility (2012)

<table>
<thead>
<tr>
<th>Segment</th>
<th>Correlation (r)</th>
<th>Average volatility (CoV)</th>
<th>N</th>
<th>% TD2-clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Uncorrelated (-0.5 &lt; r &lt; 0.5)</td>
<td>Low (CoV &lt; 0.5)</td>
<td>576</td>
<td>52%</td>
</tr>
<tr>
<td>2</td>
<td>Uncorrelated (-0.5 &lt; r &lt; 0.5)</td>
<td>High (CoV ≥ 0.5)</td>
<td>230</td>
<td>49%</td>
</tr>
<tr>
<td>3</td>
<td>Pos. correlated (r ≥ 0.5)</td>
<td>Low (CoV &lt; 0.5)</td>
<td>49</td>
<td>76%</td>
</tr>
<tr>
<td>4</td>
<td>Pos. correlated (r ≥ 0.5)</td>
<td>High (CoV ≥ 0.5)</td>
<td>16</td>
<td>75%</td>
</tr>
<tr>
<td>5</td>
<td>Neg. correlated (r ≤ -0.5)</td>
<td>Low (CoV &lt; 0.5)</td>
<td>2</td>
<td>0%</td>
</tr>
<tr>
<td>6</td>
<td>Neg. correlated (r ≤ -0.5)</td>
<td>High (CoV ≥ 0.5)</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

Table D.2 – Influence volatility on performance of approaches for uncorrelated clusters given $P_x, P_y ≥ 0.3$

<table>
<thead>
<tr>
<th>Segment</th>
<th>Coefficient of variance subagr. var. $x$ and $y$</th>
<th>N - 2013</th>
<th>% TD2-clusters 2013</th>
<th>N - 2012</th>
<th>% TD2-clusters 2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$0 &lt; x &lt; 0.5 ; 0 &lt; y &lt; 0.5$</td>
<td>371</td>
<td>64%</td>
<td>373</td>
<td>53%</td>
</tr>
<tr>
<td>2</td>
<td>$0 &lt; x &lt; 0.5 ; 0.5 &lt; y &lt; 1$</td>
<td>156</td>
<td>58%</td>
<td>152</td>
<td>45%</td>
</tr>
<tr>
<td>3</td>
<td>$0.5 &lt; x &lt; 1 ; 0.5 &lt; y &lt; 1$</td>
<td>85</td>
<td>54%</td>
<td>82</td>
<td>49%</td>
</tr>
</tbody>
</table>

Table D.3 – Influence volatility on performance of approaches for uncorrelated clusters given $P_x$ or $P_y < 0.3$

<table>
<thead>
<tr>
<th>Segment</th>
<th>Coefficient of variance subagr. var. $x$ and $y$</th>
<th>N - 2013</th>
<th>% TD2-clusters 2013</th>
<th>N - 2012</th>
<th>% TD2-clusters 2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$0 &lt; x &lt; 0.5 ; 0 &lt; y &lt; 0.5$</td>
<td>71</td>
<td>56%</td>
<td>69</td>
<td>58%</td>
</tr>
<tr>
<td>2</td>
<td>$0 &lt; x &lt; 0.5 ; 0.5 &lt; y &lt; 1$</td>
<td>65</td>
<td>48%</td>
<td>69</td>
<td>51%</td>
</tr>
<tr>
<td>3</td>
<td>$0.5 &lt; x &lt; 1 ; 0.5 &lt; y &lt; 1$</td>
<td>32</td>
<td>47%</td>
<td>35</td>
<td>49%</td>
</tr>
</tbody>
</table>
### Metric analysis control objective 1

<table>
<thead>
<tr>
<th>Metric</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong></td>
<td>$\frac{</td>
<td>\varepsilon_{i,t}</td>
<td>}{D_{i,t}} \times 100%$</td>
</tr>
<tr>
<td><strong>Criterion</strong></td>
<td><strong>Summarisable</strong></td>
<td><strong>Scale-independent</strong></td>
<td><strong>Stakeholders’ agreement</strong></td>
</tr>
<tr>
<td><strong>Understandable</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Availability of input data</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Applicable to all types of time series</strong></td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
</tr>
</tbody>
</table>

*For the sake of simplicity, the definition only represents the forecast error of a single item at a single point in time.*

**Figure D.4** – Metric analysis control objective 1