CHALMERS
UNIVERSITY OF TECHNOLOGY

Systematic Lead Time Analysis

Master of Science Thesis
in the Program Production Engineering

ARNÒR JÒNSSON

VICTOR SVENSSON

Department of Technology Management and Economics
Division of Supply and Operations Management
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2016
Report No. E 2016:026

MASTER'S THESIS E2016:026

Systematic Lead Time Analysis

ARNÓR JÓNSSON

VICTOR SVENSON

Tutor, Chalmers: Peter Almström
Tutor, Organisation: Leif Pehrsson

Department of Technology Management and Economics
Division of Supply and Operations Management
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2016

Systematic Lead Time Analysis

ARNÓR JÓNSSON
VICTOR SVENSON

Master's Thesis E2016:026

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

Chalmers Reproservice
Gothenburg, Sweden 2016

Abstract

Current trends of mass customization, decreased product life-cycles and innovation increase the need for short factory lead time to respond swiftly to changing markets. The thesis constitutes a case study, where lead time was studied in an engine factory owned by a European car manufacturer. Executives have set a target to decrease the factory lead time by 70 %.

The thesis aimed to measure and analyse current lead time, and to provide recommendations to increase efficiency for lead time data collection, processing, distribution and analysis in the future. A process mapping was performed to understand the manufacturing system and contextual issues.

ID tags are scanned at each stage in production to recognize variants and instruct equipment and workers. Lead times were measured using time stamps from those ID scans, which were stored in manufacturing execution systems. Issues were encountered with old- and non-existing systems, where samples of lead time data had to be taken.

JMP statistical software and control charts were used to analyse the lead time data, searching for root-causes to exceptional variation and long lead times. All lead times were modified by multiplying them with an un-disclosed scaling factor in order to accommodate confidentiality. Relations in the data remain intact, but the actual lead times are not revealed. From the distribution of total lead time in the results, it was shown that 91 % of the time was spent in buffers and storages.

Lead time was highly affected by the weekday. Engines started close to the weekend got trapped, and were often finished after the weekend. The data was separated and two data modifications were constructed in order to give added insight to other relations, trends and parameters that would otherwise remain hidden under the weekday variation.

Two ideas were presented to automate data collection and processing, which was extensively time consuming since it was done manually in the thesis. A software could be built in-house, or JMP could be used and connected to MES databases.

The thesis clearly demonstrates how statistics and control charts can be used by organisations to systematically work with lead time data. The analytical approach was powerful, since it assured objectivity by focusing on evidence-based and data driven improvements.

Keywords: *Lead time reduction, Input data management, Production development, Statistical process control chart, Factory lead time*

Acknowledgements

There are several people whom we would jointly like to thank for their aid and support during our thesis work. First of all, we would like to thank our Chalmers supervisor and examiner, Peter Almström, for his guidance, advice and feedback during the thesis work. We would also like to extend our gratitude to the organisations supervisor Leif Pehrson, who initiated the thesis and continuously guided and encouraged us.

In addition to our head supervisors we would like to thank Dani Benjamin at the organization for his extensive help with data gathering and interpretation. Dani's insight into various parts of the factory under investigation was of great value, and running our ideas by him often made us realize things we missed or could improve. Thereby, Dani's input always made significant contributions to increase the quality of the thesis.

We would also like to thank Jerry Axén and Per Thim at the organisation for their excellent support and guidance with various questions and issues along the way. There are many employees within the organisation who generously devoted valuable time to share their knowledge. Without mentioning names, we still want to pay gratitude to all those people in the organisation. The openness and the great team spirit to help each other became evident and supported the thesis.

We would also like to thank Peter Hammersberg, senior lecturer at Chalmers, for his extensive help with industrial statistics. Peter's enthusiasm and passion for his field of statistics was of great value and sparked newfound interest in the field.

Our opponent, Sofia Wikstrand, also deserves appreciation for helping us to improve our work.

Separately,

Arnór would like to thank his wife Thelma Anderson for her devoted love and support.

Victor would like to thank his girlfriend Elin Abrahamsson for her love and encouragement.

Abbreviations

The following contains an alphabetic list of abbreviations and acronyms used throughout the report.

AGV: Automated Guided Vehicle

ANOVA: Analysis of Variance (tool in Six Sigma)

BoM: Bill of Material

CAMP: Component Assembly Machining Production execution

CAMP_IM6: CAMP base assembly line

CAMP_MAC: CAMP Machining

CAMP_XM: CAMP final assembly

CNC: Computer Numerical Control

CoV: Coefficient of Variability

CRISP-DM: Cross-Industry Standard Process for Data Mining

DES: Discrete-Event Simulation

DMAIC: Define, Measure, Analyse, Improve and Control (for existing processes).

DMADV: Define, Measure, Analyse, Design, Verify (for new and developing processes).

DoE: Design of Experiments (tool in Six Sigma)

ERP: Enterprise Resource Planning

ESP: Encapsulated PostScript (file format)

FIFO: First In First Out

FMS: Flexible Manufacturing System

HTML: Hyper Text Markup Language (standard for Internet communication)

ID: Identification

IoT: Internet of Things

IoTSP: Internet of Things, Services and People

IQR: Inter-Quartile Range

JPEG: Joint Photographic Experts Group (file format for pictures)

KPI: Key Performance Indicator

KSMB: Kransystem för motorbuffert (i.e. crane system for engine buffer)

KS1S: Kransystem 1, S-fabrik (i.e. crane system 1, S-factory)

LNPL: Lower Natural Process Limit

Mac OS: apple Macintosh Operative System

MES: Manufacturing Execution System

NULL: Invalid, cancelled or revoked

OFAT: One Factor At a Time (Traditional experimental method)

RMS: Reconfigurable Manufacturing System

SPC: Statistical Process Control

SQL: Structured Query Language (standard for databases)

UNPL: Upper Natural Process Limit

VSM: Value Stream Mapping

WIP: Work In Process

Table of content

1 Introduction

The chapter includes the background, the purpose, research questions, delimitations, and deliverables.

1.1 Background

The background includes descriptions of industrial context and trends, the organization of interest, engine working principles and main components, the problem area and the lead time definition.

1.1.1 Industrial context

The heritage of today's production dates back to the first industrial revolution (1760–1830), which marks the transition from agriculture to an economy based on industrial activities (Bellgran & Säfsten, 2010). Initially, independent craftsmen sold their own labour, knowledge and parts. Tradesmen emerged as coordinators who handled procurement and transportation between craftsmen specialized in various process steps. Soon, production processes were centralized to designated factories. During the second industrial revolution (1870-1915), standardized and interchangeable parts revolutionized manufacturing.

Henry Ford (1863-1947) is often associated with the moving assembly line, which he observed within the meatpacking industry and adapted to his car production plant to speed up production (Bellgran & Säfsten, 2010). Through the use of standardized interchangeable parts and the moving assembly line, Ford's model T could be mass-produced to low prices, which enabled larger segments of the public to buy cars. Through high degree of vertical integration, low lead times, low inventories and cost reductions Ford achieved an extensive competitive advantage (Hopp & Spearman, 2008).

Unlike Ford (whose business was governed by production technology) General Motors later reshaped sales, marketing and production to be focused on market demand (Hopp & Spearman, 2008). Flexible mass production was emphasized, resulting in more affordable cars for different needs and purposes.

Scientific management, often also called Taylorism after its inventor Frederic W. Taylor (1856-1915), constitutes the next important milestone of industrial development (Hopp & Spearman, 2008). The core of this management system consists of enhancing productivity through work analysis and the development of work standards, which expressed the rate at which a worker could perform certain tasks using the best-established procedure.

Today, Lean production constitutes a widespread and important field of manufacturing practice (Bellgran & Säfsten, 2010). Lean revolves around the systematic elimination of waste to enhance customer orientation, with the ultimate goal of achieving broad excellence in terms of e.g. high quality, low cost and shortened lead times (Liker & Meier, 2006).

In the automotive industry, the current average life cycle for a car is six years, but that figure is progressively declining (Waldchen, 2014). New car models are introduced and withdrawn at a faster pace than ever before, which increases research and development costs and drives the evolution towards more cost efficient, responsive, and customer focused manufacturing systems. In highly competitive industries, only agile and fast suppliers will prosper (Ignizio, 2009). Mass customization, merges efficiency and low production costs with extensive possibilities for personalization and customization (Coletti & Aicher, 2011). The movement towards mass customization is also recognizable in the automotive industry.

New technologies and demands for more sustainable solutions are shaping todays economies with no exception of the automotive industry where four disruptive technology trends have emerged: autonomous driving, diverse mobility, connectivity and electrification (McKinsey, 2016). Electrified vehicles (hybrid, fuel cell, plug-in and battery electric) are becoming more viable and competitive, as battery technology advances in terms of making them lighter, cheaper and longer lasting. The speed of transformation over to electrified vehicles depends on the interaction between customer pull and regulatory push. By the year 2030 electrified vehicles could make up 10 to 50% of all new vehicle sales. However, hybrid solutions make up a large portion of electrified vehicles, making the combustion engine very relevant beyond 2030.

1.1.2 Organization of interest

With the industrial context covered, it is necessary to give a brief introduction to the organization of interest in which the thesis is performed, before moving on to the problem area. The organization, which requested to remain anonymous throughout the thesis, is a European car manufacturer that produces its own engines. The thesis was carried out at their engine manufacturing facility, which is separate from their car production. All lead times were modified by multiplying them with an un-disclosed scaling factor in order to accommodate confidentiality. Relations in the data remained intact, but the actual lead times are not revealed.

Figure 1-1 shows a simplified overview of the engine factory, covering its manufacturing flow and interactions with customers and suppliers. Raw material (which in some cases are semi-finished components already processed by the supplier) is delivered from suppliers and machined into finalized components; cylinder heads, cylinder blocks, crankshafts and camshafts. Additional suppliers provide input material, components and sub-assemblies at various stages in base and final assembly. Finalized engines are then tested and shipped to customers. Comprehensive descriptions of the manufacturing system are included in later chapters.

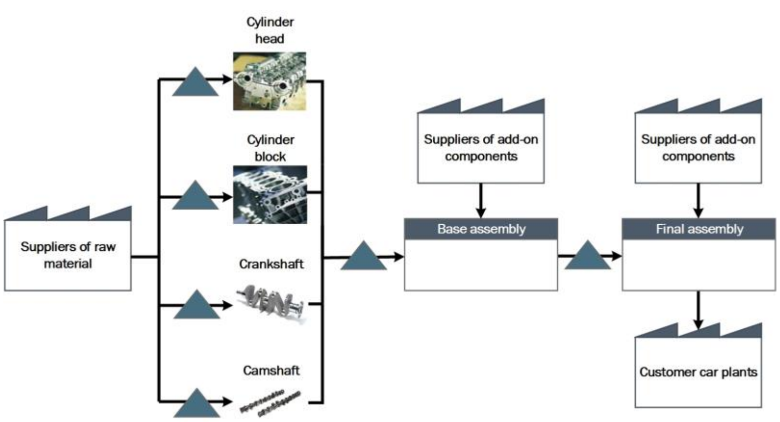


Figure 1-1 Simplified overview of engine production

Based on the future estimated development of the automotive industry the organization expects future product development trends to revolve around the following aspects:

- Increased environmental awareness
- Technological innovations
- Increased customer focus and responsiveness

These product development trends in turn increase the pressure on the organization to strengthen their manufacturing system abilities included in Table 1-1.

Table 1-1 The organizations thoughts on future development trends

Productivity	Ability to produce cost efficiently
Robust production flows	Ability to maintain production by effectively coping with system disturbances, such as quality problems, in-adequate line or sub-system balancing and breakdowns.
Flexibility	Ability to respond fast to short-term demand changes, but also to long-term changes, such as phasing in and out engine products and variants without extensive re-configurations.
Scalability	Ability to mix low- and high volume engine products and variants with regards to capacity.

1.1.3 Engine principles and components

Since the thesis focuses on engine production the purpose here is to provide an introduction to the working principles of combustion engines and their main components. Nikolaus August Otto invented the first petrol combustion engine in 1861, and therefore the engine type is often called “Otto engine” (Konrad, 2014). A patent for the first diesel engine was issued in 1892, but the engine was first used in regular cars in 1936. Today, both petrol and diesel engines are common.

Figure 1-2 shows the inner construction and main components of two four-cylinder engines, the main type of engines manufactured by the organization (a petrol version to the left and diesel version to the right). For both type of engines, the pistons in each cylinder is driven by the combustion of an air and fuel mixture (Konrad, 2014). For the diesel engine, the pressure and temperature in the combustion chamber (i.e. in the cylinder) are high enough for self-ignition of the air and fuel mixture. For the petrol engine, spark plugs are used to create a spark that drives the combustion of the air and fuel mixture. The pistons are linked to the crankshaft, which transform the linear movement of the pistons into rotary movement, which is then transferred to the wheels via the transmission.

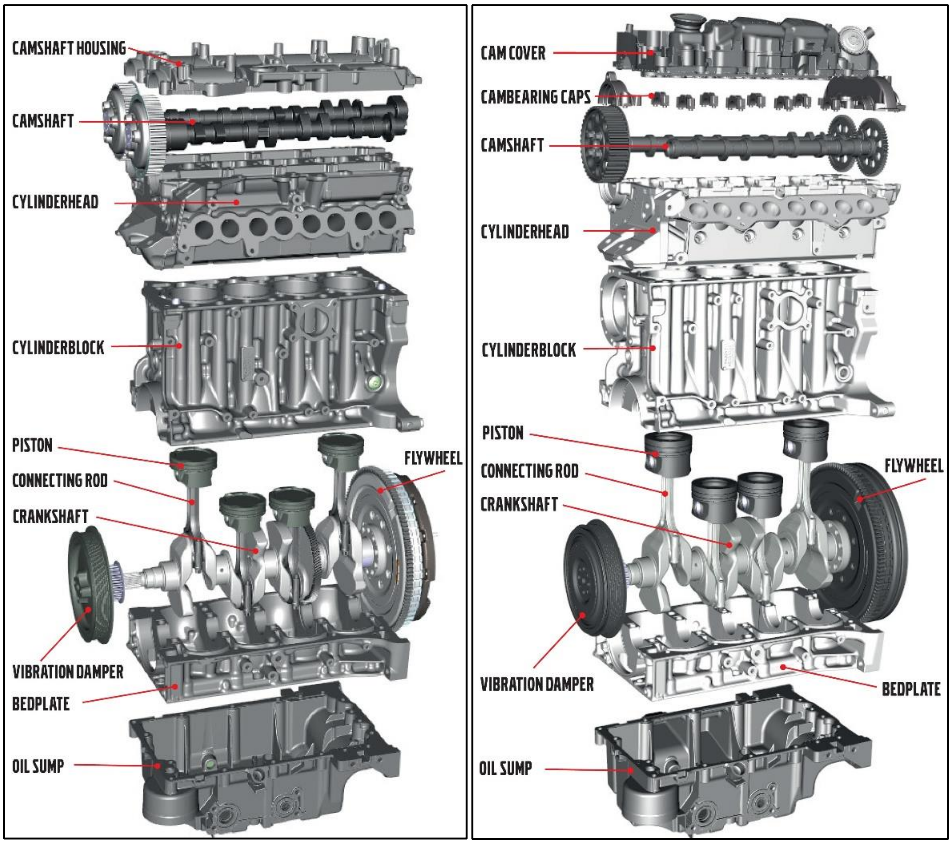


Figure 1-2 Exploded view of the organizations petrol- (left) and diesel (right) engines

Generally, car producers tend to concentrate on core competencies and key components, which has led to a drastic increase in outsourcing of component manufacturing to external suppliers over the last decades (Waldchen, 2014). However, engines and transmissions are strategically and technologically important, which means their development and production often is kept in-house. The organization follows this trend by having its own engine factory.

1.1.4 Problem area

When examining the historical development, one decisive factor for success becomes identifiable: a need for manufacturing speed that equals short factory and supply-chain lead times (Ignizio, 2009). Short lead times allowed Henry Ford to pay high wages, while still being cost efficient enough to crush competition. Lead time reduction is also fundamental in Lean production, where it is used jointly with waste elimination to enhance customer focus.

There are many potential benefits associated with lead time reduction, with the following being some important examples:

- Increased flexibility and responsiveness
- Decreased time to market
- Improved financial performance through decreased inventories and work in process
- Increased customer satisfaction, which supports growth in market shares.

Lead time reduction has been fundamental in manufacturing, and still requires much attention in the future development of factories. However, the incentives for achieving short lead times have changed over time. During Henry Ford's era, short lead time mainly facilitated high volume and low cost production. Although capacity and cost are still of significance, short lead times are nowadays needed for flexibility reasons. Today's trends of mass customization, decreasing product life-cycles and rapid technology development stress the need for short lead times to swiftly cope with frequent market and customer changes.

With respect to future growth and development trends, the organization foresees a need for increased customer focus and responsiveness through improved flexibility. In order to improve flexibility, top management has set a target to reduce the overall factory lead time with 70 %. This target shapes the focus on production development, and clearly indicates the necessity for new development projects focused on lead time reduction.

Lead time can be studied from different perspectives, with focus on specific lead time for machines and processes, production lines, the manufacturing system or the complete supply chain. For the thesis, it was decided to focus on the manufacturing system. The choice of perspective was governed by the managerial lead time target, as well as previous lead time studies in the organization. Lead time reduction projects have been conducted before, but they all focused mainly on lead time optimization for machines, processes and production lines. Consequently, there is a clear need for studies attaining a manufacturing system or complete supply chain perspective. Instead of investigating specific production lines, focus is on the connection of production lines, time spent in storages and buffers, and overall system performance. The benefits associated with such a perspective lies in the holistic picture it provides, which helps to avoid sub-optimizations. The drawbacks are the lack of detailed knowledge and the complexity of studying with the complete system.

1.1.5 Lead time definition

Lead time sounds quite intuitive, and most readers can probably agree that lead time is the time between start and finish of production. However, when does production start and stop in todays extended enterprises? Is it when pieces are loaded into the first machine, or is it when raw material is received from suppliers? From a broad perspective the time for product realization may be defined differently depending on the literature source and its focus. The organization also provides different records for the overall lead time in their factory, depending on the originating function or department providing the numbers. For the machining processes for instance that can depend on whether time spent in raw material storage is incorporated in the lead time or not.

The inconsistency with respect to lead time definition offers great confusion. Therefore, it is relevant to present a clear definition of the term, which is used throughout the thesis:

"Lead time is the total time from the arrival of raw material, through manufacturing, to the dispatch of finished products."

The particular definition of lead time chosen for the thesis allows for a holistic perspective that guides the research towards a focus on the complete internal manufacturing system.

1.2 Purpose

The purpose of the thesis is to determine whether lead times for the factories entire manufacturing system can be generated. Such unprecedented aggregated analysis would provide new and valuable insight into lead time reduction, to the means of increasing flexibility and lowering cost. By acquiring lead times for every stage of the manufacturing process cross-comparison of distribution and variation can be made between stages, revealing possible dependencies and constraining factors towards lowering lead times. Emphasis is put on how data extraction, processing and relaying data can be done in the most efficient way.

The purpose of the literature review is to provide theories and methods that are comparable to the way production is run in practice within the organization. The literature review also covers manufacturing paradigms and trends from a strategic perspective to enable analyses of possible impacts on lead time solutions.

1.3 Research questions

In order to guide the research, the following research questions were developed:

Question 1:

1. What is the current total lead time in the manufacturing system?
    - How is the lead time distributed within the system?
    - How does the lead time vary within the system?

Motivation:
In order to make strategic decisions that are governed or influenced by lead time, the current lead time must be known. It is important to be able to break current lead time down to specific parts of the system to analyse lead time distribution with regards to identifying improvement areas. Equally important is measuring lead time variation at different stages in the process to analyse how lead time fluctuates over time.

Question 2:

2. How can lead time data be continuously extracted, processed and distributed efficiently within the organization?

Motivation:
For the purpose of collecting lead time data at any given point, it is necessary to develop the means of continuously and effectively extracting and processing lead time data. It is vital to present lead time data in a clear and effective way that can be utilized for various departments from simulation to production planning. The process of data extraction also reveals which processes or buffers present a problem due to insufficient data.

Question 3:

3. What areas hold the biggest potential for lead time improvement and at what trade-off?

Motivation:
This question complements the findings from the first research question, imposing further studies for increased understanding. It triggers in-depth investigations of the true potential for improvement in areas with long and varying lead times. Due to the complexity of the complete manufacturing system, recommendations for lead time reduction are likely associated with certain trade-offs. The question therefore supports the development of such recommendations, but also incorporates the necessity for consciousness about the trade-offs.

1.4 Delimitations

To ensure adherence to the purpose, objectives and research questions, the following delimitations have been made:

- The thesis focuses on one of the organisations factories, meaning that proposes recommendations may not be directly transferable and applicable to other factories within the organisation.
- The thesis attains a system perspective, meaning that specific operations and production lines are not covered in detail. Instead, focus is on describing interactions and dependencies between operations and lines that jointly constitute the complete flow.
- The analysis will only focus on petrol engine lines since the diesel engine lines had recently been changed and are still in the ramp up face, resulting in insufficient data from such lines. (This will be further elaborated on in the result chapter)

2 Methodology

The chapter includes a description of the research method used, potential alternative analytical tools and methods, research triangulation and trustworthiness and credibility. The methodology chapter is intentionally brief and focused on the overall research method. For increased understanding and context further descriptions about detailed data collection, processing and analysis are to be found in the result chapter.

2.1 Research method

The thesis constituted a case study, since it focused on a specific manufacturing site within one organization (Bryman and Bell, 2011). The research consisted of a mix of quantitative methods utilizing numerical data extracted from production systems and qualitative approaches through process mappings, observations and interviews. Research questions were formulated in order to guide the research, especially with respect to the focus of the literature review, data collection and analysis.

Figure 2-1 illustrates the overall research approach used in the thesis. The theoretical foundation was based on a literature review, which mainly utilized scientific databases such as Summon, Proquest and Google Scholar. The literature review laid the theoretical foundation for data collection and analysis, enabling comparison between observed practices and scientific theories.

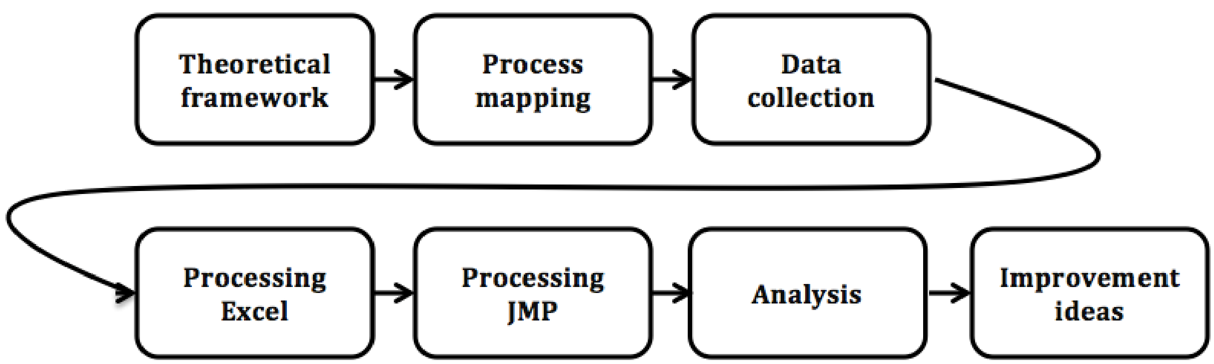


Figure 2-1 Research approach

A process mapping of the current state in the organization was conducted in accordance with the first stage of the value stream mapping (VSM) methodology. VSM is a Lean production tool for visualizing and analysing the current state and designing a future state for a product creation process (covering the whole flow from raw material to the products reaching the customers) (King and King, 2015). The process mapping was used to gain understanding and contextual awareness, which was proven valuable in later stages to put ideas in the right perspective.

For the complex manufacturing system investigated in the thesis, it was considered impossible to develop a realistic future state map solely based on the guidance from the VSM methodology. Although there are predefined questions available to govern the development of the future state map (Liker and Meier, 2006), they did not provide enough support. The predefined questions offered no guarantee for not overlooking crucial aspects, and it was therefore hard to assess the impact of certain improvement ideas. Therefore, the process mapping was used as a complement to the other research methods used.

For the purpose of measuring and analysing factory lead time, quantitative manufacturing data was extracted from the organizations internal manufacturing execution system (MES). For one area in the factory, where it was proven impossible to extract data in this way, a manual approach was developed for measuring lead times.

Several issues were encountered along the way, especially with respect to collection and processing of the quantitative data. To better understand, overcome and solve those issues, observations, discussions and semi-structured interviews with specialists from within the organization were used. Since the authors of the thesis lacked detailed knowledge and experience from specific areas in the organization, it was considered appropriate to utilize staff's experience.

Semi-structured interviews were chosen for their flexibility, allowing for deviations from predetermined questions to inquire to previously unidentified important aspects (Bryman and Bell, 2011). Three interviews were conducted with personnel within the organization to gain understanding and utilize competence needed to overcome problems within the specific field of each interviewee. Interviews were held with one person working at the IT department and two people working with logistics and material handling. Observations and discussions took place when visiting and examining various parts of the manufacturing system together with one of the supervisors.

The collected manufacturing data from the MES was processed in Excel and JMP statistical software, resulting in aggregated lead time measures for each stage of the manufacturing flow. JMP is a statistical software used for industrial statistics and exploratory data analysis, with focus on visualization through graphical representations (JMPa, 2016).

The analysis to identify areas of improvement and means of lowering the lead time was done using JMP. The lead time data was scrutinized through the use of various graphical representations. Initially, simple graphical representations were used to explore the data. In the next step, control charts were used to search for pattern, trends and root-causes for deviating and long lead times. The control chart is a specific type of graphical representation used in Six Sigma to systematically analyse complex manufacturing data, striving for consistent output to enhance process performance (Pyzdek, 2014; Wheeler, 2003). Additional manufacturing data was used to cross-compare with the lead time data. By exploring the data searching for patterns and deviations it was possible to identify potential improvement areas.

2.2 Alternative methods for the analysis

The lead time analysis was conducted in JMP, using various graphical representations and statistical tools. Control charts were used to a large extent, mainly because of its powerful systematic and powerful approach for data analysis. Simpler graphical representations were used as complements to better understand the data. Potentially, there are other methods available that could have been used for conducting similar analyses.

The complete VSM methodology could have been used, but as implied earlier it was regarded inappropriate due to the lack of guidelines. The processed data could also have been used as input to lead time optimization using simulation models. With respect to the limited timeframe and the authors personal interest and competence, this was not a viable option. JMP and control charts also have the advantage of focusing more on understanding and improving the quality of the lead time data, which was important for the organization since similar work has never been done before.

Beside control charts there were various other statistical tools for analysis available. In the thesis, control charts were used to examine exceptional variation, which could contain both outliers and relevant data. Outliers could also be detected using other graphical or numerical methods, such as histograms, scatter plots, Z-score or the inter-quartile range (IQR) method (Larose and Larose, 2015). An outlier is an abnormal data value, either an error or a planned or unplanned deviation from the norm. The decision on which method to use should mainly be governed by the distribution of the data set (Hammersberg, 2016).

The decision to use control charts was governed by the fact that control charts impose no assumption about the data (mainly no assumption about normal distribution). The box plot is another method, which also could have been used to identify outliers. Control charts were still chosen over box plots, due to its systematic approach for analysis.

Industrial statistics was discovered to be a vast field of research. Due to the authors of the thesis not being so experienced in statistics, it was difficult at first to research the whole field and assess the relevance of each area. Therefore, there was a risk of overlooking certain aspects that could be of interest for the thesis. In order to minimize that risk, three meetings were held with a senior lecturer at Chalmers currently doing research within industrial statistics and Six Sigma.

2.3 Research triangulation

Figure 2-2 shows the complementing mix of research methods used. The qualitative methods used where interviews and observations, while the data analysis was quantitative. Triangulation is believed to give greater adherence to the research questions compared to using a single method (Bryman and Bell, 2011).

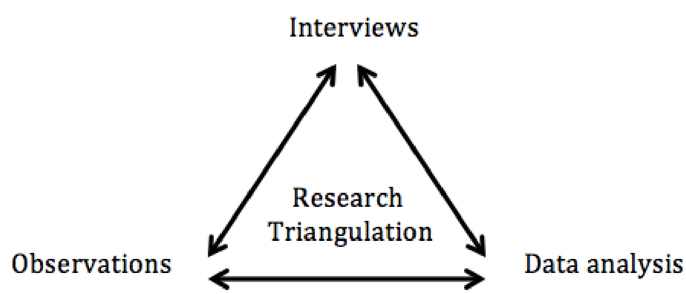


Figure 2-2 Research triangulation (Bryman and Bell, 2011)

Quantitative and qualitative methods are both associated with certain advantages and weaknesses, however using a combination of the two minimizes the impact of individual weaknesses and increases reliability (Bryman and Bell, 2011). Reliability refers to the ability to repeat the results of a study, while replication refers to the ability to replicate the method. The cross-comparison between quantitative and qualitative findings used in triangulation results in greater research validity. Validity refers to whether or not a measure really represents the concept that it is supposed to denote.

2.4 Trustworthiness and credibility

Respondent validations were utilized to ensure the trustworthiness of qualitative data. This meant sending summarized observation and interview data back to participants for validation in order to ensure credibility (Bryman and Bell, 2011). All qualitative data was compiled and analysed immediately after its extraction to ensure that data was correctly interpreted also adding to credibility.

2.5 Ethics

The ethical issues that can arise between researchers and research participants can be broken down into four areas, regarding: harm to participants; lack of informed consent; invasion of privacy; and whether deception is involved (Bryman and Bell, 2011). As previously stated three interviews were conducted with employees in the organization. The interviews all gave full consent and were fully disclosed of the intentions of the interviews and how the interview material would be utilized for the purpose of the thesis.

Other ethical concerns can for instance regard data security, data sharing, intended use of data and conflicts of interest amongst stakeholders (Bryman and Bell, 2011). Great care was taken with all data shared by the organization, which was only used for completion of the thesis and deleted afterwards. No ethical issues arose regarding conflicts of interest among shareholders as all shared the same vision to a common goal.

3 Theoretical framework

The theoretical framework provides insight to state-of-the-art research and manufacturing paradigms related to the focus area of the thesis. The chapter contains descriptions about Lean, factory design, factory physics, industrial statistics, six sigma, big data and manufacturing data management for improvement analysis.

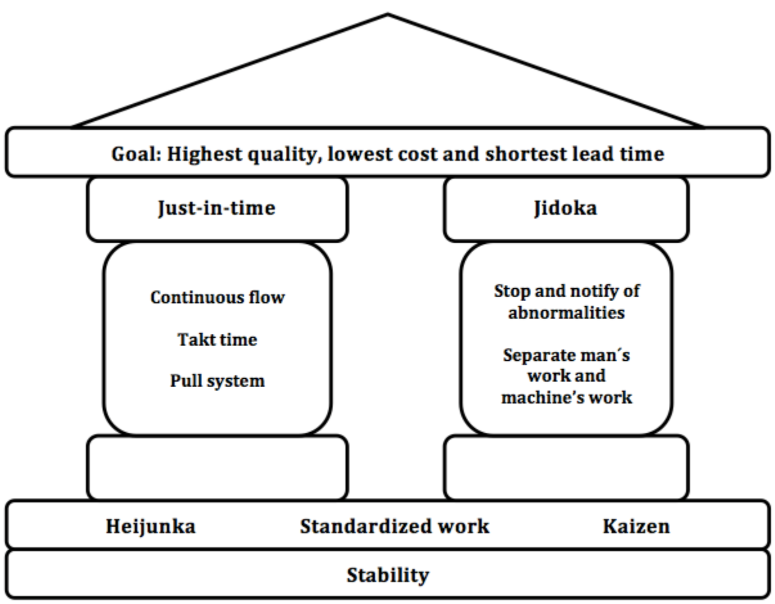
3.1 Lean Production

The Lean philosophy became highly popular in manufacturing industry in 1990 through the publication of the book "*The Machine that Changed the World"* (Liker and Meier, 2006). Lean thinking revolves around the implementation of improvement processes, which identify and eliminate inherent waste by examining root causes and proposing countermeasures. A crucial goal of waste reduction is to minimize lead time between customer order and delivery to enhance customer focus and -value.

Waste is categorized into eight major types, which all contribute to in-efficiency in the production system and thus decreases the customer value. The eight waste activities are; excess production, waiting time, transportation, overwork, inventory, movement, manufacturing of defect products and un-exploited competence. Excess inventory can be found in the form of raw material, work in progress (WIP) or finished goods all causing longer lead times.

The Lean philosophy revolves around long-term thinking and commitment to continuously improving and learning by analysing and understanding all processes in great detail (Liker and Meier, 2006). In order for a Lean organization to lower their lead times and enhance customer value, certain prerequisites must be fulfilled. These prerequisites and the mutual relation between them are often explained with the Lean house, which is shown in Figure 3-1.

Figure 3-1 Lean house (Liker and Meier, 2006)



The reference to a house is a pertinent one, since every part of the house builds on an equally important prerequisite (Liker and Meier, 2006). The foundation is comprised of levelling production (Heijunka) and standardizing work, which serves as the foundation for continuous improvement (Kaizen). The pillars consist of Just-in-time and Jidoka. Just-in-time is a term encompassing the reduction of lead times by pulling items through production based on customer demand instead of pushing items through production based on projected demand. Jidoka is a term for equipping every machine and empowering every worker to stop production at the first sign of an abnormal condition in order to resolve the issue so production does not have to be affected or stopped due to the same problem again. The objective of implementing the above-mentioned methodology is to increase quality while lowering cost and shortening lead times.

3.1.1 Value Stream Mapping

The primary goal of Lean is to reduce waste and improve material flow (King and King, 2015). In order to attain the holistic overview needed to understand where in the process waste exists, the VSM tool was invented. A value stream contains all the actions, both value and non-value adding, currently required to bring the product through main flow essential for its creation (Rother and Shook, 1999).

VSM visually represents the flow of material and information throughout the product realization process, providing a common ground for understanding (Liker and Meier, 2006). In addition, VSM identifies and visualizes the information trigger points for material flow. The tool holds several benefits for its user (King and King, 2015). It provides a detailed understanding of the current state, illustrating enablers and inhibitors of smooth flow, e.g. things that cause longer lead times and higher inventories. It provides an understanding of both separated operations and integrated processes, and aids in envisioning future Lean value streams.

VSM identifies the main forms of waste found in a process operation and accurately portrays the major effects such waste, providing insight into root cause analysis. The philosophy of VSM is to straighten out the overall flow of the value stream before diving deep into fixing individual processes (Liker and Meier, 2006). VSM should preferably be conducted with representatives from all process and functions building a strong cross-functional understanding of the entire overall process (King and King, 2015).

VSM revolves around the creation of a current- and a future state map (Liker and Meier, 2006). The purpose of the current state map is to understand the production flow and to provide a template for the development of an improved future state map. Future state questions have been developed to support the improvement thinking process when constructing the future state map. The current state also provides design information for application of Lean improvements like production levelling and pull replenishment systems (King and King, 2015).

3.1.2 Inventory

Inventory is considered to be an indicator of weakness in a process, commonly used to compensate for inflexibility, and a constant reminder that the process needs strengthening (Liker and Meier, 2006). Inventory can be split into different categories, e.g. raw material, procured components, finished goods and work in process, with each type of inventory typically being used to compensate for a specific weakness.

Inventory management is often faced with a paradox. The Lean philosophy clearly promotes elimination of all forms of waste, including excess inventory. However, prior to initial Lean improvements immature processes might be un-stable and reduction of inventory could jeopardize production performance. Therefore, it is highly important to find a suitable balance where sustainable inventory reductions (not too rabid) are made in harmony with process development. On the other hand, excess inventory hides problems, such as production imbalances, equipment downtime, long setup times and late deliveries from suppliers. Therefore, it is equally bad to reduce inventory too slowly.

3.1.3 Buffers

Buffers are important for the performance of production flows delivering products when the process cannot (Petersson et.al, 2010). Buffers can be categorized according to their function and one buffer may serve one or several functions. Firstly, disturbance buffers cover for unplanned disturbances such as breakdowns. Buffers for planned stops compensate for e.g. maintenance, vacations and installation of machines. Consumption buffers serve as necessary decoupling points between unsynchronized processes such as processes with different shift patterns. Lastly, sequence buffers may be needed to build or fix a sequence for the subsequent process where the sequence is of importance for the flow, e.g. to substitute for a product removed from the flow (e.g. due to quality problems).

The size of the disturbance- and sequence buffer should remain the same over time, while the size of the planned stop and consumption buffers should vary naturally over time. The disturbance buffer, together with the sequence buffer, provides the safety level or absolute minimum buffer level. The actual safety provided by the disturbance buffers size at any given time is dependent on the production takt time. The disturbance buffer needs to be increased following increased takt time to maintain a constant level of safety against stoppages in the flow. Characteristics of stops in terms of length, number and duration must be known in order to be able to dimension disturbance buffers. Dimensioning a buffer to cover any problem however is irrational. Instead judgment of expected stops and acceptable risk and cost should govern the bases of dimensioning.

3.2 Factory design

There are many interesting areas within the research field of manufacturing system design. In this chapter emphasis is on appropriate degrees of flexibility and adaptability. Flexibility may facilitate the ability to re-configure manufacturing systems to stay competitive in continuously evolving markets (Koren, 2006). The chapter starts by introducing Flexible- and Reconfigurable manufacturing systems, with focus on production lines, machines and tool flexibility. The concepts of agile- and transformable factories are then described from a more aggregated factory perspective.

3.2.1 Flexible- and Reconfigurable Manufacturing Systems

Historically, dedicated manufacturing lines enabled cost-efficient high volume production of products with relatively long life-cycles (Koren & Shpitalni, 2011). With increasing needs for customization and shortened product life-cycles, the need for flexibility and re-configurability have risen drastically since the 1980s. Flexible Manufacturing Systems (FMSs) was introduced in the 1980s to meet the increased need for flexibility. The intention was to enable production of many different variants in low volumes (Tolio, 2009). Flexibility was added to the manufacturing systems from the start to avoid extensive reconfigurations when introducing new products (Koren, 2006). However, it was proven difficult to achieve the right level of flexibility, which caused flexibility to remain un-used leading to high investment and running costs for FMSs (Tolio, 2009).

As a response to the drawbacks of excess flexibility in FMSs, the idea of reconfigurable manufacturing systems (RMSs) was introduced in the mid-90s (Koren, 2006). Instead of adding high degrees of flexibility when designing the manufacturing systems, RMSs are designed to be easily re-configurable when needed. RMSs are intended to provide cost-efficient and rapid re-configurations to cope with changes in customer demands. For instance new products or variants may be launched into existing manufacturing systems with minor system modifications. Typically, RMSs consist of sets of flexible machines and equipment, such as computer numerical control (CNC) machines, robot cells, measuring and inspection machines, tools etc. Re-configurability may be assessed from two perspectives; for individual machines or re-configuration with respect to the connection of machines forming the production line. In addition, re-configurability may also be discussed on a factory level, but is then often referred to as transformability (which is covered in the next section).

There are six parameters that quantify the re-configurability of RMSs: modularity, customization, ability to integrate new objects, scalability, ability for diagnosis and convertibility. Evaluation and validation of manufacturing system design and configurations (especially with regards to utilized flexibility) are often overlooked. In large companies, the evaluation might only be governed by following up on investment payback requirements (Tolio, 2009). Deciding appropriate degrees of flexibility often resides on manufacturing engineers, and long-term flexibility as a competitive weapon is not considered.

3.2.2 Transformable and smart factories

Companies must possess the ability to swiftly change the configuration of their factories to stay competitive on today's continuously evolving markets. There seem to be many alternative labels or "buzzwords", which tries to emphasise certain aspects of this ability.

According to Berger (2006) "agile production" was introduced in 1995 and describes manufacturing systems with parallel resources, which enables parallel item processing. However, Scheuermann, Verclas and Bruegges (2015) interpretation of agile production is different. The "agile factory" is about applying agile software engineering techniques in manufacturing. Agile software engineering means that feedback loops are created to allow the customer to influence the product during the realization phase, thus increasing the customization. As a demonstrating example, simple tracking devices were used to create a prototype of an agile factory, in which information feedback loops allowed the customer to effect the ordered product's configuration during production.

Instead of "agile factory", the term "Cyber-physical human system in manufacturing" is often used, since it also highlights three important design spaces: Cyber, physical and human space (Scheuermann, Verclas and Bruegge, 2015). Much research is currently being conducted on the subject of future smart factories. Besides "agile factories" there are many other similar terms used such as transformable factories, smart factories, ubiquitous factories, real-time factory, factory of things (Scheuermann, Verclas and Bruegge, 2015) and industrial Internet of Things, services and people (IoTSP) (ABB, 2016). One interesting aim of agile factories is to link virtual and physical development by making information accessible everywhere and anytime in the factories (Scheuermann, Verclas and Bruegge, 2015).

Re-configuring whole factories to raise competitiveness is resource demanding, and usually has to be done in the mid- to long-term perspective (Westkämper, 2006). By concentrating efforts on joint re-configurations of both production and logistics, the German processing industry was able to improve competitiveness in terms of throughput times and WIP.

3.3 Factory physics

Analysing manufacturing systems should be done in faceted ways, where e.g. popular Lean thinking is combined with scientific approaches, intuition and reasoning (Hopp & Spearman, 2008). The term "factory physics" may be used as a label for the scientific approach and understanding of manufacturing system behaviour. For instance, Lean clearly highlights the necessity of waste elimination to enhance factory performance, but fails to provide scientific guidelines to steer waste elimination activities.

Ignizio (2009) promotes a more scientific approach to manufacturing system analysis and defines a factory as: *"a nonlinear, dynamic, stochastic system with feedback".* The definition highlights the inherent variability and complexity of modern factories. By examining intrinsic relations between specific parameters it becomes possible to understand and predict manufacturing system behaviour. Human intuition is often useful here, since humans are good at understanding complex relations and for testing parameter changes to enhance performance (Hopp and Spearman, 2008).

However, Ignizio (2009) does not share this confidence in human intuition. With today's complex manufacturing systems, human intuition is often not enough to predict system behaviours.

Within the concept of "factory physics", some basic equations have been developed that helps engineers to understand the most fundamental behaviours of modern manufacturing systems. The first and most famous one is *"Little's law"* (see Equation 1), which quantifies the relation between lead time, WIP and throughput (Hopp and Spearman, 2008; Ignizio, 2009)

Equation 1: Little's law

$$Little's\ law: \quad WIP = Lead\ time * Throughput$$

Little's law can be proven mathematically, given certain assumptions (Hopp and Spearman, 2008). In real factory practice, Little's law has proven to provide adequate approximations of real behaviours (Ignizio, 2009). Assuming all parameters are measured in the same unit, Little's law is applicable for analysing single operations, production lines, manufacturing systems and complete factories (Hopp and Spearman, 2008).

In Equation 2, Little's law has been modified to accommodate discussion on lead time reduction (Hopp and Spearman, 2008). It becomes evident that lead time reduction can be achieved by reducing WIP or increasing the throughput rate.

Equation 2: Little's law modified to enhance lead time reduction

$$Lead\ time = \frac{WIP}{Throughut}$$

However, Equation 2 is deceptive, since the same absolute lead time (i.e. WIP/Throughput ratio) can be achieved by several combinations of absolute WIP and throughput values.

Little's law, as well as other "factory physics equations" (described further down) are based on several assumptions about manufacturing system conditions. The equations assume that every workstation connects only to one following process. All machines in a workstation are also assumed to have equal performance. By mathematically extending the equations, it becomes possible to describe complex manufacturing systems more accurately. Without any mathematical extensions, the equations are still able to approximate factory behaviours.

3.3.1 Variability

Everything in life is exposed to variations, and to some extent there is always a need to effectively understand and cope with variability (Hopp & Spearman, 2008). However, there is a great risk of variability being falsely accepted as a naturally inherent feature. In many cases, the reduction of variability might lead to faster and more resource efficient improvements in performance, compared to elevating capacity (Ignizio, 2009).

Variability is closely related to the areas of randomness and probability, and can be categorized either as random or controllable (Hopp & Spearman, 2008). Randomness is further classified as true or apparent, where apparent means that the viewer falsely interprets behaviour to be random while in reality it is not. The false interpretation is often caused by inadequate knowledge about how the manufacturing system actually works.

There are many potential sources of variability in manufacturing systems and they can be grouped as process- and flow related. Process variability refers to process-specific parameters such as setups, breakdowns, scrap rate and lead time. Flow variation concerns the general connected flow, where sources of variation may be related to planning and execution, production order release or strategies for material movement.

Analysing variability can be done using conventional statistical knowledge, such as calculating variance, median, and standard deviation. However, these measures are all expressed in absolute terms. The coefficient of variability (CoV) is a relative coefficient that enables comparison of variation across operations, production lines or manufacturing sub-systems (Hopp & Spearman, 2008; Ignizio, 2009). Equation 3 defines the CoV.

Equation 3: Coefficient of variability

$$Coeffiecient\ of\ variability\ (VoC) = \frac{\sigma}{\mu} = \frac{standard\ deviation}{mean}$$

As the mean value and standard deviation are expressed in the same unit, the CoV is a unit-less expression of variation (Ignizio, 2009).

There are different ways to reduce variability. *Variability pooling* is applicable for variability reduction for batching, queuing and managing buffers (Hopp & Spearman, 2008). Variability pooling tries to combine several sources of variability to minimize their impact. One example is the use of generic buffers instead of dedicated ones. Scheduling preventive maintenance as short and frequent activities instead of rare and long disruptions also helps to reduce factory variability (Ignizio, 2009).

Sometimes, production managers also tend to over-react by re-allocating resources to cope with un-balanced work loading. Such a managerial behaviour makes the manufacturing system un-stable, which increase variability and reduce performance.

3.3.2 Lead time

Manufacturing speed, expressed in terms of short lead times, has always been a profound measure of success (Ignizio, 2009). Short lead times was important for Henry Ford's mass production, and is still a key feature promoted within Lean to facilitate customer focus.

With Little's law as a base, one can approximate worst-case scenarios of WIP, throughput and lead times, which enables internal benchmarking (i.e. comparison between present state and the theoretically worst-case scenario) (Hopp & Spearman, 2008).

There is an explicit relation between factory loading and lead time, where lead time increases with increased workload (Ignizio, 2009). This relation is effectively illustrated using "factory performance curves", which are plots displaying factory workload on the x-axis and total lead time on the y-axis.

In order to create a factory performance curve, lead time data must be measured for different levels of system loading. In real production, such measurements are tedious and un-practical. A more suitable way to generate the curve is to run scenarios with different loading levels in a simulation model. The fact that lead time is dependent on factory loading is crucial to consider when evaluating factory performance. Comparing factory lead times between factories through benchmarking is only suitable if factory loading is considered.

3.4 Industrial statistics

Exploratory data analysis is the process of translating raw data into meaningful information, which captures, summarizes and conveys the most interesting characteristics of the data (Myatt, 2014). Myatt (2014) provides a framework for conducting exploratory data analysis and data mining projects. The framework comprises four main steps:

1. Define the problem and plan the study
2. Prepare the data (collect, characterize, clean and transform)
3. Select suitable method for analysis
4. Deploy, which involves summarizing and sharing study results with stakeholders.

Descriptive statistics summarizes the characteristics of data into measures of central tendency (e.g. median and mean) or measures of dispersion (e.g. range, maximum, minimum and standard deviation) (Duignan, 2016). Inferences should normally not be based solely on descriptive statistics measures, but instead complemented by exploratory graphics to aid in decision-making.

There is a wide range of powerful industrial statistical techniques and tools available, some of which have been recently adopted in the rapidly evolving field of big data mining and analytics (Myatt, 2014). Big data refers to datasets that are too vast to manage and analyse. Graphical representations are crucial to visualize statistical results in clear and interpretable ways, especially for decision-making purposes. The recommendation for the type of graphical representation tools to be used in specific situations depends primarily on the sample size and whether the data has been collected over time or not (Ryan, 2007). Assuming data was gathered over time, a rule of thumb is that the data should preferably also be plotted over time.

Steam-and-leaf displays, Digi-dot plots and Dot-plots are good for visualizing small datasets, but become blurry and hard to read for larger data sets. Histograms and boxplots are appropriate for initial analysis of large data sets. Histograms are bar charts that categorize data in classes and display the frequency of each class as the height of each bar. Working with large datasets sometimes requires sampling.

Statistical process control using control charts serves two purposes; data is properly displayed over time and it is possible to assess whether the process is statistically stable or not.

When working with industrial statistics, data distributions require much attention. There are many different statistical distributions, with some of the most fundamental ones being Normal, Binominal, Poisson, Geometric, Gamma, Exponential and Weibull.

The phrasing of a distribution being "normal" is an unfortunate one, since in practice there is no such thing as normal distribution (Ryan, 2007). However, normal distribution is often used to approximate the actual distribution of many random variables. Whenever calculating standard deviation the assumption is made that the underlying data is normally distributed as can be best describe through Figure 3-2. There μ denotes the mean value of the distribution and σ represents the standard variation.

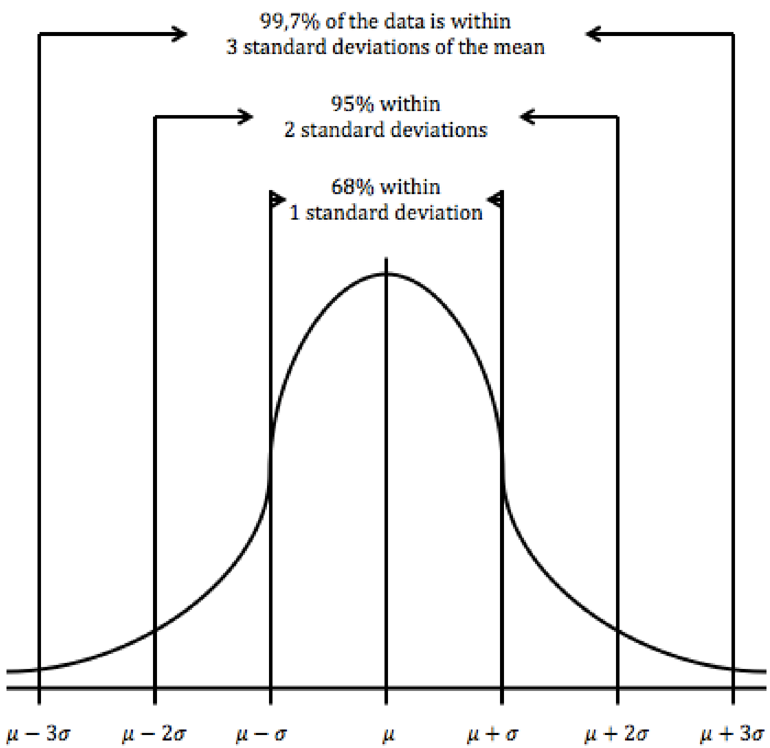


Figure 3-2 Normal distribution, mean and standard deviation (Ryan, 2007).

3.4.1 Detecting and eliminating outliers

Raw data, extracted from databases, is usually unprocessed making it incomplete and noisy (Larose and Larose, 2015). The data may e.g. be missing values, covering fields that are obsolete or redundant, or contain outliers.

Outliers are extreme values that lie an abnormal distance from remaining values, and can either represent an error in data entry or any planned or unplanned deviation from the norm such as preventive maintenance or breakdowns (Larose and Larose, 2015). Outliers can be identified through both graphical and numerical methods. One-dimensional histograms are commonly used when dealing with one variable, however two-dimensional scatter plots can be more helpful in revealing outliers when dealing with more than one variable.

The numerical Z-score method for outlier identification states that a data value is an outlier if it has a Z-score greater than 3. The Z-value test calculates the number of standard deviations by which the data varies from the mean by subtracting the mean from the data point and dividing that number with the standard deviation (see Equation 4):

Equation 4: Z-value normalization

$$Z = \frac{|x - \mu|}{\sigma}$$

$$, where\ x: observation\ value\ \mu: mean, \sigma: standard\ deviation$$

The method's shortcoming is that the incorporated mean and standard deviation are both sensitive to the presence of outliers, which creates a paradox. Presuming an outlier is added or removed from the dataset then both the standard deviation and mean will be affected by the presence of this new data value. Therefore using a method that is sensitive to the presence of outliers is not ideal. Therefore, a more robust statistical method for outlier detection has been developed, which is not as sensitive to the presence of the outliers themselves, namely the inter-quartile range (IQR) method.

Figure 3-3 illustrates the IQR method and exemplifies what regards as an outlier. The method divides the dataset into four quartiles, each containing 25% of the data. In other words, the first quarter contains 25 % of the data, the second quarter contains 50 % of the data, and etc. The IQR is a measure of variability and is calculated by subtracting Q1 from Q3.

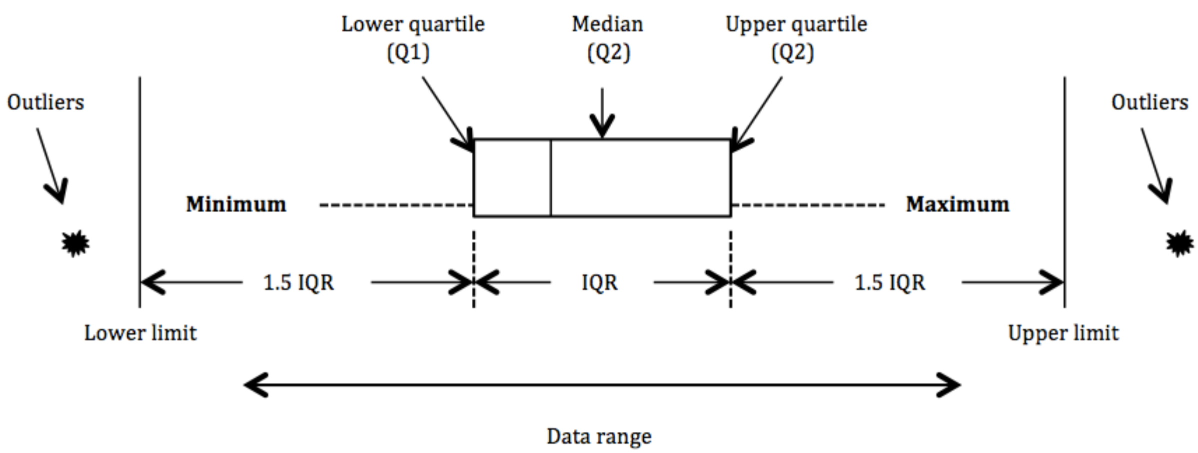


Figure 3-3 IQR outlier detection explained (Larose and Larose, 2015)

Assuming X is the data point value, an observation is detected as an outlier if one of the following two conditions holds:

$$\bar{X}\ located\ (1.5 * IQR)\ below\ Q_1$$
$$\bar{X}\ located\ (1.5 * IQR)\ above\ Q_3$$

Although the IQR method is more robust with respect to the presence of outliers, it still assumes normal distribution. Therefore, the distribution of the dataset must always be assessed and both methods used with caution.

3.4.2 JMP statistical software

Sound analysis of business data is crucial for strategic information-based decisions within improvement projects (JMPa, 2016). There are several software's available for industrial statistics with the most widespread tools being JMP Statistical Discovery, Minitab17, Microsoft Excel and R statistics (JMPa, 2016; Minitab, 2016; R Foundation, 2016). The software's all comprise similar versions of the most fundamental tools for statistical analysis such as basic statistics, regression analysis, statistical process control (SPC) charts and design of experiments (DoE).

The JMP statistical software intends to provide quick ways to "see your data", with the underlying philosophy of "one graph for every statistic and vice versa" (JMPa, 2016). JMP structures the information as data tables (similar to worksheets in Excel), where each column represents a variable and every row is an observation (JMPb, 2016). Thus, each data point consists of one row with information stored in one or multiple columns.

A powerful feature in JMP is that all points in plots are linked to the corresponding values in the data table (JMPb, 2016). For instance, abnormal data points may be identified and selected from the plots. The user then switches to the data table window and generates a new data table, which only contains the selected values. In this way, the abnormal values may be further examined separately.

Columns can be assigned properties, such as formulas or modelling types. The modelling types instruct JMP how to treat column values, and can be set either to continuous, ordinal or nominal. Continuous columns are continuous measurement values. Ordinal are values with mutual order, while nominal values are discrete without mutual ordering. Formulas can also be added to compute column values and the formulas can be based on calculations made on other column values.

The distribution platform and graph builder functions are appropriate for initial data analysis (JMPc, 2016; JMPd, 2016). The distribution platform uses a histogram to illustrate the distribution of all observations in the data table. The graph builder enables quick changes to the plot (e.g. replacing one variable with another) allowing the user to explore the data and discover interesting patterns and trends. These two visualization tools for initial exploration of the data are excellent because they impose no assumptions about the data, such as homogenous datasets (Hammersberg, 2016). Homogenous datasets are assumed to be normally distributed, without the presence of outliers. Generally, statistical analysis tools must be used with caution, since many of the tools are built from underlying assumptions of homogenous data.

Importing data to JMP can be done from Excel (copy/paste) (JMPb, 2016) There is also an add-in tool available to help advanced users to run Excel models in JMP. The Query builder connects JMP to SQL (Structured Query Language) databases, which enables importation of data directly from those databases to JMP. The user can specify SQL statements and add filters to customize data importing. Results can be shared as interactive HTML (standard for Internet communication) and Adobe Flash files to non-JMP users. Reports can also be formed into emails and sent to specified receivers.

3.5 Six Sigma

Six Sigma is a method that utilizes a set of techniques and tools for process improvement founded on variability reduction to ensure process enhancement and consistent output in manufacturing (Summers, 2010). The goal of Six Sigma is to achieve the highest quality of a manufacturing process by identifying causes of deviations or defects and removing them (Pyzdek, 2014). The use of the Greek letter sigma (σ) refers to statistics denoting variation from a standard.

A manufacturing process can be described with a sigma rating indicating its yield of defect-free opportunities. A Six Sigma process needs to produce 99.99966% defect-free opportunities, meaning that if an organization wants to achieve Six Sigma, they cannot produce more than 3,4 defects per million opportunities. Through constantly measuring and analysing defects, it is made possible to systematically eliminate them getting as close to perfection as possible.

The Six Sigma methodology follows one of two processes, DMAIC or DMADV, respectively (Pyzdek, 2014). DMAIC defines, measures, analyses, improves, and controls existing processes, which have fallen below specification and are looking for incremental improvement. DMADV defines, measures, analyses, designs, and verifies new and developing processes, which are striving for Six Sigma quality.

The benefits of Six Sigma include reduced- cost, waste and pollution (Summers, 2010). Additional benefits are better understanding of customer requirements and more customer satisfaction and last but not least shortened lead times. The Six Sigma toolbox comprises various statistical methods. For the purpose of the thesis, statistical process control and design of experiments, are described in following sub-chapters.

For its key implementation roles, Six Sigma has adopted a ranking system inspired by martial arts where colored belts are used (Pyzdek, 2014). Executive leaders are responsible for creating the vision and framing the direction of the development work, while Champions are responsible for the implementation across the whole organization. Master black belts serve as mentors or coaches, assisting the champions and guiding black- and green belts working in the organization. Black belts apply the Six Sigma methodology as specific projects, whereas green belts carry on the Six Sigma implementation along with their regular job responsibilities. In other words, master black belts and regular black belts work solely with Six Sigma, while green belts also have other responsibilities beyond Six Sigma implementation.

3.5.1 Statistical Process Control

Flowcharts, cause-and-effect diagrams (i.e. fishbone diagrams over cause-and effect relations), and bar charts are three examples of graphical representations, which in different ways help to visualize certain relations and aspects of datasets (Wheeler, 2003). For each kind of plot, data is displayed in a context where patterns and relations may be identified between different parts of the dataset. In some cases, these visual representations, combined with experience and knowledge of a process might constitute enough support process improvements. However, in many cases patterns, relations and cause-and-effect relationships are too complex for such ad-hoc analysis to be effective. For instance, cause-and-effect parameters might influence or intervene with each other. In many processes, one could also possibly have tens or hundreds of cause-and-effect relations. In such cases, analysis based on control charts has emerged as a powerful tool for systematically working with process improvements.

Six Sigma aims for highest quality through the identification and elimination of causes for process variations and defects (Pyzdek, 2014). Some people might relate the concept solely to reduction of scrap rates. However, processes and all kinds of data on key performance indicators (KPIs) have variation that limits process performance (Wheeler, 2003). Continuous monitoring and systematic reduction of variation helps to enhance process performance.

Figure 3-4 is an example of a control chart for an industrial process. Depending on the type of control chart, each point may represent either an individual observation, or a descriptive average for a sub-group of several observations (JMPe, 2015; Wheeler, 2003). The green line is the average for all data points, and constitutes a reference when searching for trends. The two red lines correspond to the upper- and lower natural process limits, which are used for distinguishing between routine- and exceptional variation (Wheeler, 2003). Routine variation always occurs and can be regarded as inherent noise. Exceptional variation should be interpreted as signals caused by special changes to the process, which should be further investigated.

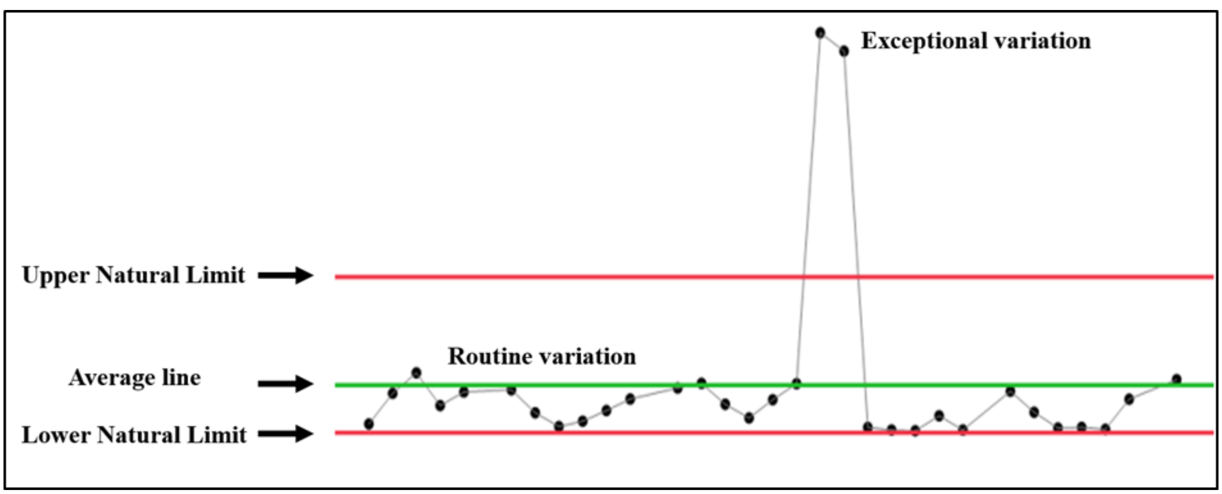


Figure 3-4 SPC average (X) chart example of an industrial process

A process that only contains routine variation is predictable, since it is possible to estimate future performance based on historical measures (Wheeler, 2003). Processes that contain both routine- and exceptional variation are unpredictable, since their future performance cannot be predicted.

The fundamental idea for enhancing process performance is to identify the parameters that cause exceptional variation to occur and then change parameter settings to decrease variation, striving for predictable process behaviour.

The control charts can be constructed based on different statistical measures. Figure 3-4 showed a so called "X chart", where each point in the plot represents a descriptive average for a sample group. The same kind of plot can be created based on individual observations, in which case each point represents the absolute value and the chart is called "individual chart".

In addition to the plots described above, the control chart view is often complemented by adding another plot that displays the process variation. The X chart is combined with a range (R) plot, while the individual chart is combined with a moving range (mR) plot (JMPe, 2015). The range is the difference between the maximum and minimum value in the sample group, while the moving range is the difference between two successive (individual) data points (JMPe, 2015; Wheeler, 2003). By convention, the range and moving range are defined as positive values since they are expressions of process variability.

Figure 3-5 is an example of a combined control chart, which comprises both an X- and an R- plot (JMPe, 2015). The top plot shows the average value over time, while the bottom plot shows the variation for each data point.

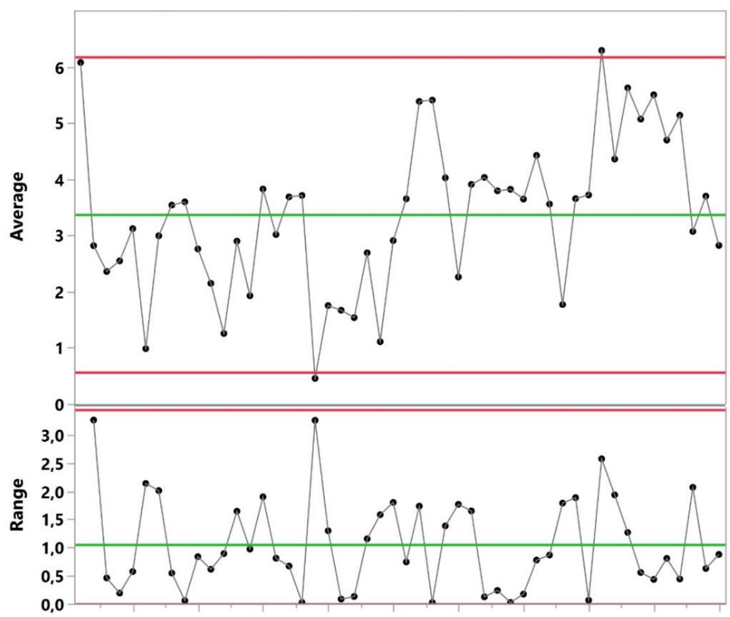


Figure 3-5 Statistical Process Control chart displaying average (X) and range (R) plots

All control charts have estimated natural process limits, which help distinguish between routine and exceptional variation (JMPe, 2015; Wheeler, 2003). Since the range and moving range are always positive, their lower natural limit cannot go below zero. The limits must be calculated with care. Too narrow limits make the control chart over-sensitive, while too wide limits increase the risk of missing out on exceptional variation.

There are several mathematical equations for calculating the lower- and upper natural process limits. Some simpler equations assume normal distribution, while the more advanced ones do not require assumptions about the distribution of data (Ryan, 2007). JMP uses advanced equations that allow control charts to be used without making assumptions about the distribution of the data (JMPe, 2015).

There are several rules, which can be utilized to systematically detect exceptional variation and trends. The simplest rule is to search for points outside the limits from either of the two complementing plot views. Other rules search for trends, for instance successive combinations of data points either above or below the average line.

3.5.2 Design of Experiments

Designed experiments are an important quality improvement tool utilizing analysis of variance (ANOVA) techniques to partition the variation in a response amongst other potential sources of variation (Pyzdek, 2014). The objective of experimenting with the complex interactions among parameters within a process is to gain the necessary insight to optimize the process (Summers, 2010). ANOVA is different from one factor at a time (OFAT) approaches traditionally used, holding all factors constant except for one (Pyzdek, 2014). One of the OFATs approach main drawback is that it is usually hard or impossible to hold all other variables constant, meaning there is no way to systematically account for experimental errors such as measurement variation. Another main drawback is that potential interaction causing synergy-effects between variables are not considered.

The DoE method on the other hand usually involves varying two or more variables simultaneously and through those means obtaining multiple measurements under the same experimental condition. Benefits of this method include the detection and measurement of interactions. Each value does the work of several values as the same observation can be used to estimate several different effects, and lastly experimental error can be quantified and used to determine the confidence of conclusions.

DoE experiments are usually done through either full factorial or fractional factorial design. (Summers, 2010) Full factorial design consists of all possible combinations of all possible discrete values of investigated factors. Fractional factorial designs only study a subset of possible combinations, which is less complicated and time consuming. Although not examining all possible combinations the fractional factorial way, when designed correctly, may still reveal the complex interactions between factors, including which factors hold more significance over others.

3.6 Big Data

As information technology constantly evolves, big data and the internet of things (IoT) have emerged as two important fields of research. IoT refers to the use of sophisticated devices, which are all connected to the Internet to enable smart communication and exchange of data (Farooq, et. al., 2015). The number of connected devices is expected to increase from approximately 14 billion in year 2014 to 50 billion in 2020. As the number of connected devises increases, large amounts of data must be analysed requiring analytic methods that are capable of handling massive datasets (Chen, et. al., 2015).

Big data can be defined as:

> "*Datasets whose size is beyond the ability of typical database software tools to capture, store, manage and analyse"* (Wang & Alexander, 2015).

There are also other alternative definitions. Big data is sometimes described in terms of 3 Vs, encompassing big data characteristics; volume, velocity and variety (Berman, 2013). Volume refers to the size of the dataset. Velocity means the data is rapidly changed and re-generated and variety means that different forms of data are mixed and combined (e.g. text, values, and images in structured or un-structured from different sources).

A common misinterpretation is classifying voluminous data, such as extensive spreadsheets from one data source, as big data. According to the definition all 3 Vs must be fulfilled for the data to be regarded as big. Sometimes, the 3 V definition is also extended into 6 Vs by adding value, variability and veracity (Wang & Alexander, 2015).

Despite the development of IoT and big data being closely interlinked (as the amount of data increases with the development of more connected devices) IoT is not a prerequisite for big data (Krumeich et al., 2014). Big data already applies in existing factories, which often lack extensive connectivity.

One example is medical centres, where data in the form of laboratory reports, drug orders and receipts, billing information, patient record history etc. is stored for various purposes such as supporting medical decisions and invoicing patients (Berman, 2013). The diverse data generated from different medical functions is voluminous, updated frequently and with additional data being continuously added. This vast pool of data qualifies as big data. Assuming the data would be compiled, it could potentially be utilized to enhance improvements in quality, time efficiency and cost.

Another concrete example derives from manufacturing industry, where big data analytics based on logistical models were utilized to improve production planning and logistical performance (Nywlt & Grigutsch, 2015). In a highly digitalized factory, lead time and order release data were measured, visualized and examined. The results from the study show high variability in lead time and order releases, which created unstable conditions that limited planning and logistical performance. Based on the insight about inherent process variability, actions were taken to create a more robust production environment. The increased awareness about variability in lead time and order releases simplified production planning.

3.6.1 Analytics and data mining

Big data analytics can be described as a process comprising four main steps: data mining, cleaning, conversion, and iterative data analysis (Berman, 2013). Data mining is usually combined with the creation of models to make estimated predictions about future scenarios.

Data mining and predictive analytics are defined as:

> *"The process of discovering useful patterns and trends in large data sets and extracting relevant information from large data sets in order to make predictions and estimates about future outcomes"* (Larose and Larose, 2015).

A cross-industry standard has been developed as systematic support for the data mining procedure. The standard is called Cross-industry standard process for data mining (CRISP-DM) and comprises six successive analytical steps:

1. Business and research understanding
2. Data understanding
3. Data preparation
4. Modelling
5. Evaluation
6. Deployment

3.7 Manufacturing data management for improvement analyses

There are several means available for analysing manufacturing systems to enhance performance with respect to KPIs. Factory physics points at the combination of Lean tools, insight and experience and scientific relations to analyse and improve production (Hopp and Spearman, 2008). The Lean theory promotes the use of VSM (Liker and Meier, 2006), while the use of discrete event simulation (DES) is powerful for complex systems (Ignizio, 2009). Although factory physics provides firm ground for analysis, it must often be complemented by simulation in today's complex factories, where dynamic and stochastic aspects must be considered.

However, irrespective of the type of analysis and optimization method used, the key to success and reliable results is access to correct basic input data (Jonsson and Mattsson, 2009). Item-, Bill-of-material (BoM)-, routing-, and work-centre data are main categories of basic data. This data is often fully or partially embedded in the company's enterprise resource planning (ERP) and MES. Item- and BoM data specifies product structure and components used. Routing data specifies how products and components are manufactured, the resources required and accumulated lead times. Work-centre data specifies available capacity for the various resources in the factory.

Efficient material- and production flows are dependent on a comprehensive and well-maintained database including all categories of basic data. Product development, order design, purchasing, production engineering and simulation teams are examples of company functions, which utilize basic data for various analyses and improvement projects (Jonsson and Mattsson, 2009; Robertson and Perera, 2001).

For example, production engineers can utilize lead times from the routing data to analyse set-up times, lead times and critical paths (Jonsson and Mattsson, 2009). The critical path is the longest accumulated lead time for procurement and manufacturing of an end product. Lead time analyses helps to identify items with the longest lead times. That information can then be used to make strategic priorities in production planning to control and reduce the lead times.

Input data management also plays a significant role in DES applications, where it is the single most time-consuming activity (Robertson and Perera, 2001). In some cases, as much as 40 % of the total time in simulation projects is spent on input data management (Skoogh, 2011). The main obstacles making input data management so time consuming is problem with finding the right data sources, scarce availability of data and limited possibilities to structure and process data in the simulation software.

The time consumption can be extensively reduced by optimizing and automating the process for data extraction, processing and transfer (Robertson and Perera, 2001). Figure 3-6 illustrates four methods for input data management to simulation models with various degrees of automation. For the first method, data collection, processing and distribution is done manually by the project team members. The data is incorporated into the simulation model. The second method separates the basic data from the model to increase flexibility. Basic data is still collected, processed and compiled manually to a structured spreadsheet, which is read automatically by the simulation model. The third method is automatic and the data is transferred from various ERP and MES databases via an intermediate database (acting as harmonizing interface) to the simulation model. The fourth method allows the simulation model to continuously communicate with the ERP and MES to assure access to the most recent data updates.

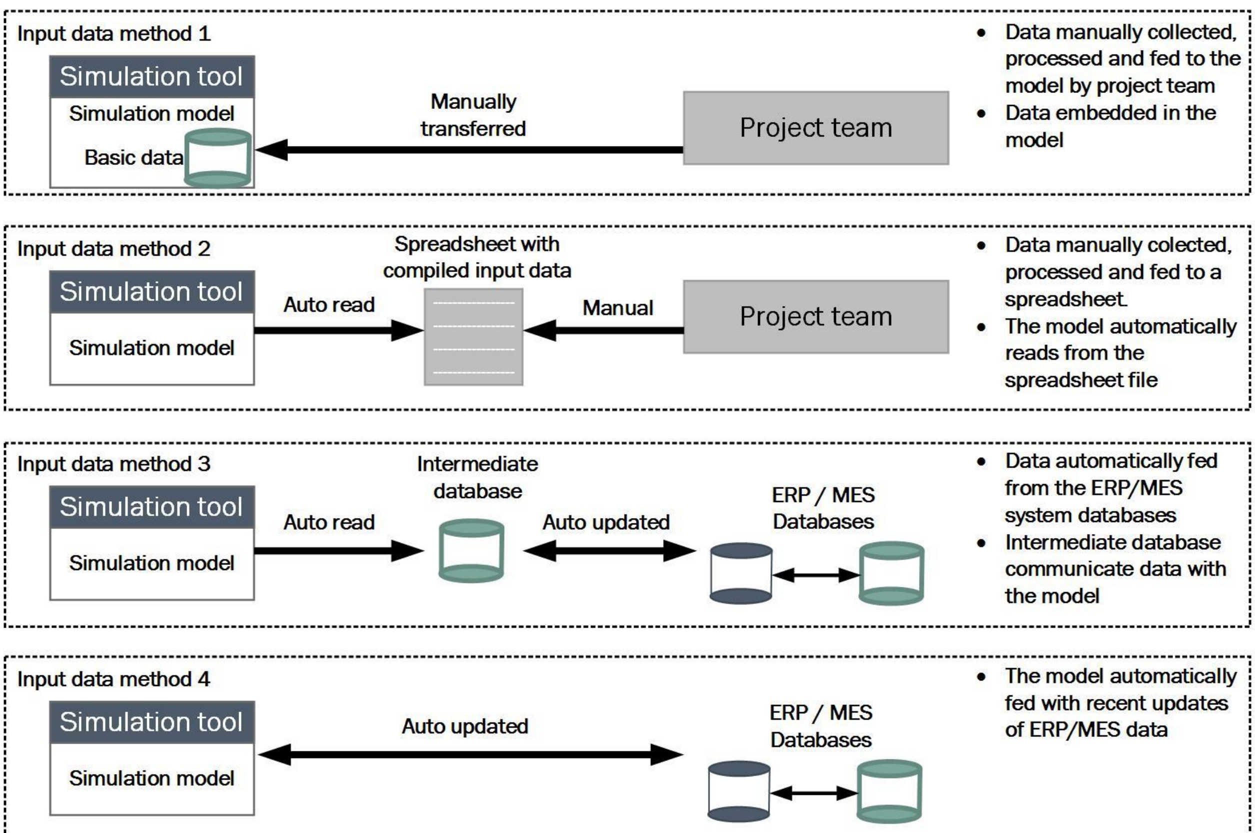


Figure 3-6 Methods for managing input data in simulation (Robertson and Perera, 2001)

Several ideas and case studies about simplifying input data management and connecting simulation models to ERP/MES can be found in literature. Three examples can be found through Randell and Blomsjö (2001), Skoogh (2011) and Cwikla (2014).

4 Results

This chapter includes results of the process mapping and lead time analysis.

4.1 Process mapping

The process mapping includes descriptions of the production system, development of critical path, approach for lead time measurements and description of production IT systems.

4.1.1 Production system

Figure 4-1 shows an overview of the factory's complete manufacturing system, spanning from raw material to finalized engines. The dashed line indicates the critical path used for investigating lead time in the thesis. The development and use of the critical path will be descried in detail in the next sub-chapter.

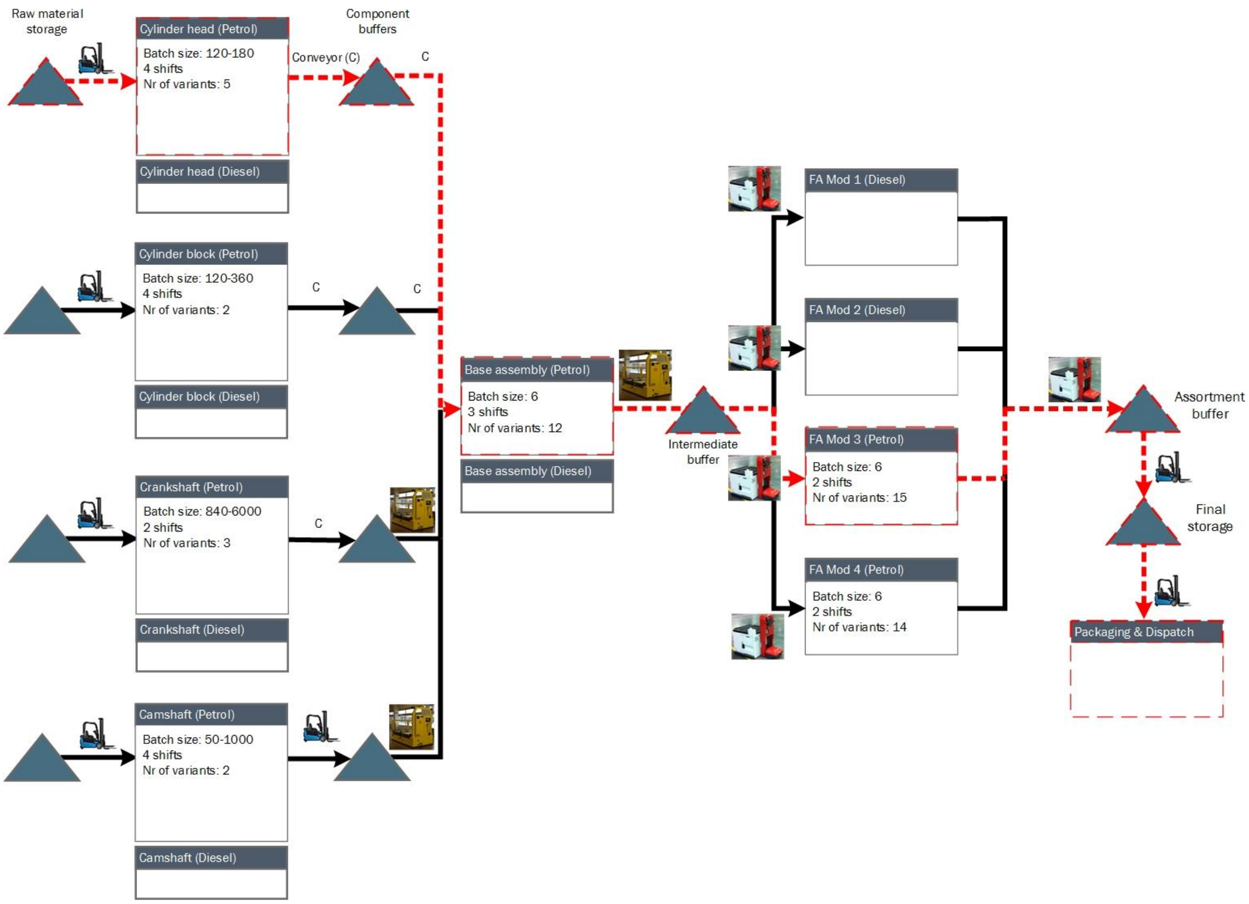


Figure 4-1 Production system overview

Four core-engine components (cylinder heads, cylinder blocks, crankshaft and camshafts), which are usually pre-processed by the supplier, are produced in-house through step-wise machining. Additional components are bought from suppliers and fed to production at various stages in the assembly processes. Throughout the factory, production of petrol and diesel variants are normally separated into dedicated production lines.

After machining, components are stored in dedicated buffers. The components are supplied to base assembly, either by dedicated conveyors (cylinder heads and cylinder blocks) or by large automated guided vehicles (AGVs) (crankshafts and camshafts). In base assembly internal engine components are assembled into so called base engines.

The intermediate buffer contains base engines in batches of six, which are stored according to the scheduled production sequence in the four final assembly modules. Engines are transferred from the intermediate buffer to each final assembly module by small AGVs, each holding one individual engine. In final assembly external components like turbos, plastic covers, tubes and wires are added.

Both kinds of AGVs are shared resources, meaning that the fleet of AGVs may be utilized differently depending on the workload of different parts of the factory. The assortment buffer and the final storage intend to merge the outfeed from the separate final assembly modules into the right delivery sequence and temporarily hold engines to re-create batches of six engines. Engines are taken from the final storage and loaded in outbound trucks in the packaging and dispatch area.

In Figure 4-1 batch sizes, work-shift patterns, number of variants and means of transportation has also been specified. Engines are delivered to the car plants on specialized transportation-racks, each capable of holding six engines of the same variant. The transportation racks were developed to optimize the usage of truck space and to ensure safe transportation.

Engine variants are clustered into planning groups based on assembly work content, which depends on the end product being configured as a low- or high performance type. Low performance means that the power output is lower. A high-performance engine is equipped with stronger parts and double-turbos, thus requiring longer assembly time. The specialized racks and planning groups mainly govern the batch size of six engines used in base and final assembly.

Regarding shifts, the 2 shifts refer to work being conducted on a day and an evening shift. For 3 shifts, a night shift is added (on top of day and evening). Where 4 shifts are run a weekend shift is added. The number of operators on each shift varies depending on the planned production volumes, which is expressed through the scheduled tact time. For instance, cylinder head machining currently run 4 shifts, with maximum tact (i.e. highest possible production volume) on days, nights and evenings. On weekends, the tact time is lowered, with fewer operators manning the line and subsequently smaller volumes produced. The shifts can be changed swiftly with respect to the customer demand.

The transportation means vary throughout the factory between regular forklifts, trolleys pulled by movers (i.e. Lean milk-round trains), dedicated conveyors, small-, and large AGVs.

Each type of buffer, whether holding machined components, base engines or finalized engines, has a target level that specifies the number of items that the buffer should contain. Buffer targets are normally defined in terms of production run-out time. The target level for the cylinder head buffer (i.e. between cylinder head machining and base assembly) is 36 hours. The target for the intermediate buffer (i.e. between base- and final assembly) is 3 hours of production. Using base assembly as an example, if production were to halt there, the final assembly modules should be able to consume the buffer and run according to plan for 3 hours before being starved of base engines. By expressing the buffer size in terms of run out time, the number of stored engines varies according to the tact time. With high tact time in final assembly, the number of base engines in the intermediate buffer is higher, compared to a situation with lower tact time. This way of managing buffers assures that the risk for material shortage remains constant over time. The same principles then apply for the cylinder head buffer.

The case of the assortment buffer is more unique whereas the buffer is intended to gather engines of the same variant to batches of six and build up the right sequence for delivery. The assortment buffer has the maximum capacity of 400 engines.

The maximum capacity of the raw material storage is difficult to define, since it depends on how material is stacked, how many variants are situated in the storage at each time and etc. However, the maximum capacity of the final storage is 4000 engines.

4.1.2 Development of the critical path for lead time measurements

From the production system overview (which was shown in Figure 4-1) and the complementing descriptions, it becomes evident that the production system is vast, advanced and complex. With respect to the limited timeframe for the, decisions were made to delimit the mapping and analysis of lead time. A “critical path”, which was highlighted in the production system overview (see Figure 4-1) by a dashed line, was defined and the measurement and analysis of lead time was focused along that path.

It was decided to focus solely on the production of petrol components and engine variants. This decision was governed by current conditions in base assembly. Previously, all engines were produced in one line, but dedicated lines have recently been built to separate petrol and diesel production. The newly built line produces diesel, while the old line produces petrol variants. The new diesel line is in a ramp-up phase with an abnormal level of disturbances. For the lead time analysis, it was considered more representative to study base assembly for petrol variants, since it was run under more normal and steady conditions. In order to make lead times comparable between processes, it was also decided to focus on petrol variants throughout the system.

For component production the lead time study was focused on cylinder heads due to lead time data being more easily accessible in that part of the system. As will be covered in detail in the later chapters (especially Production IT systems) different MES are used in various parts of the factory, which effect the possibilities of measuring lead time data. In some IT systems, basic production data can be connected to individual items while in other systems it is not possible. This is because some new versions of IT systems use ID tags on each item to control production.

The delimitations made to generate the critical path were governed by insight that was revealed during discussions with supervisors and staff in the process-mapping phase. Therefore, the delimitations were considered part of the process mapping results (rather than project-related delimitations to be included in the delimitation chapter).

4.1.3 Approach for lead time measurements

Figure 4-2 displays the critical path selected for the lead time study. The idea for measuring lead time is to collect “time stamps” from discrete points throughout the factory. A time stamp is the time registered and stored in various IT systems for certain activities in production. Since all cylinder heads and engines (in base- and final assembly) are equipped with ID tags, the idea was to compile time stamp data from a number of individual items for each part of the system. The compiled data would provide ground for statistical analysis.

Each time an item is moved to the next station in production, the ID tag on the item is scanned. The dots in Figure 4-2 represent the selected time stamps, one at the start and one at the end of each stage along the critical path. The difference between end- and start time yields the lead time for each stage. The sum of all successive lead times yields the accumulated total factory lead time for finalized engines.

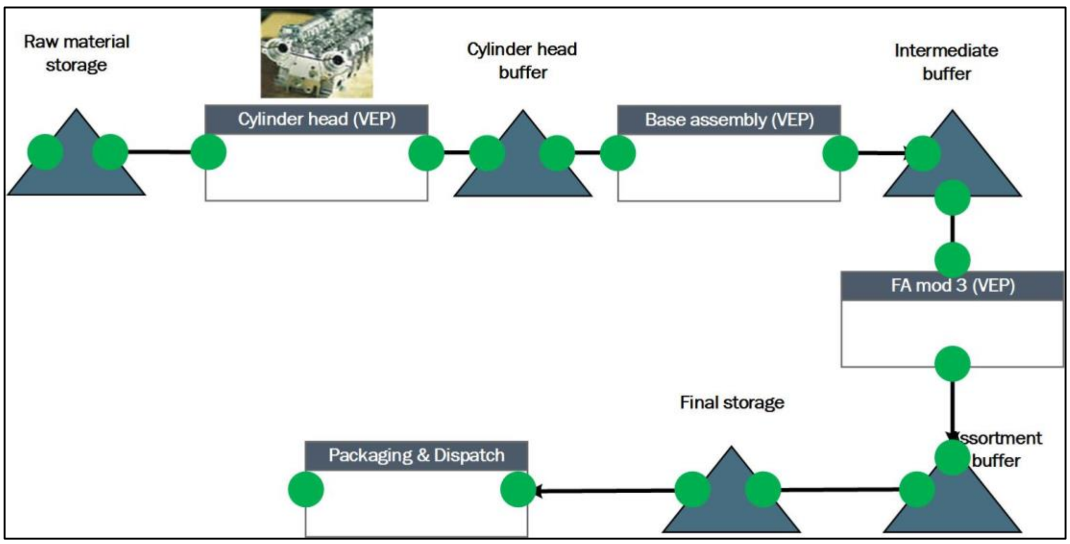


Figure 4-2 Production time stamps for lead time measurements

For cylinder head machining and base assembly, production is characterized by a high degree of automation. The first and last operations are robot cells, which load and unload material to and from the line. For these parts of the system, the time stamps were taken from the robot cells and correspond to the activity when the robot scans the ID tag.

The cylinder head-, intermediate-, assortment buffer and final storage are automatic storages controlled by production IT systems. Each buffer or storage has an “infeed” and an “outfeed” station, where pallets of material are dropped off or picked up by the conveyor or AGVs. The time stamp for each “drop off” and “pick up” activity is saved in various databases.

Final assembly and the packaging and dispatch area are both characterized by high degrees of manual work. In final assembly, AGVs flows through the assembly carrying individual engines, while humans are conducting the assembly work at various stations. The ID tag is scanned when the AGV carrier arrives and departures from the final assembly, which enables the time stamps to be taken. For the packaging and dispatch area, the starting time is taken from the outfeed from final storage. Information is also typed in and stored, when racks of engines are loaded into supplying trucks, making it possible to measure the end time for that stage.

For consistency reasons, time stamp data was selected from the same time period covering approximately four weeks of production. However, for some parts of the system only limited data outside of that period was available. Depending on the IT system, the time stamp data sometimes includes transportation time between process steps. In incidences where such times are incorporated they are difficult to separate from the lead times. Therefore, since those times are relatively short compared to overall lead times, they are not believed to have a significant effect on the results.

Several issues were encountered with regards to the extraction of these selected time stamps data and for the estimation of lead time. Since different issues were encountered for different parts of the system, they are described under each process-specific section in the Lead time analysis chapter. As implied earlier, these issues were related to the structure of the IT systems. Before describing the lead time results and the issues of lead time data extraction, it is therefore necessary to describe the structure of the IT systems along the critical path (which is done next).

4.1.4 Production IT systems

Figure 4-3 provides an overview of the different production IT systems used in the processes, buffers and storages along the critical path. The IT systems are used for production planning and manufacturing execution and control. Starting with the raw material storage, there is no IT system available, which made it impossible to utilize the IT system to extract lead time data from time stamps. Each separate IT system is linked to a dedicated database, where both historical and running data are stored.

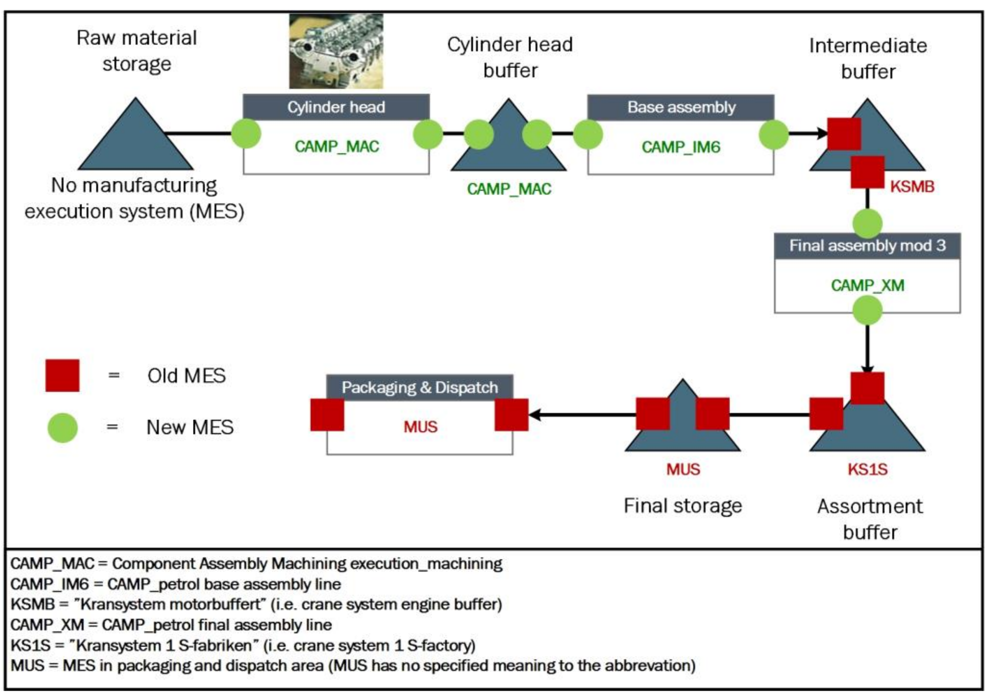


Figure 4-3 Production IT system overview

The CAMP_MAC, CAMP_IM6 and CAMP_XM (Component Assembly Machining Production execution) are the new systems, where lead time data was accessible. Limitations in storing capacity and user functionalities for the old systems complicated the extraction of time stamp data.

Although the three CAMP systems are similar, they are still organized as independent systems that need to be accessed separately. In the case of extracting time stamp data, one data inquiry (i.e. a request for data) was needed for each CAMP system. In total 7 data inquiries had to be created, each taking approximately 2 hours to construct and execute.

Since all IT systems are used in the everyday work for execution and control of the manufacturing process, the data inquiries had to be scheduled on nights or weekends or otherwise they would slow down the systems with the ultimate risk of stopping production.

Figure 4-4 illustrates the hierarchical structure of the IT systems and the links between different levels. All the systems described in the production IT system overview (Figure 4-3) are from the MES layer. The MES layer connects to the ERP and the shop floor execution systems, which jointly constitute the complete IT system. At the enterprise and resource planning level, a customer order list is generated based on actual demand. The customer order list is used to create the detailed production plan at MES level. The MESs communicate with the virtual device, which in turn specifies the activates to be performed by the production line equipment. MES data is temporarily stored internally and exported to an historical database. It would be possible to export data to a separate data warehouse that could be optimized for data mining and analytics, although this is not done today. For further details regarding the IT systems see **Appendix B**, which contains a summary of an interview held with one person working with MES at the IT department.

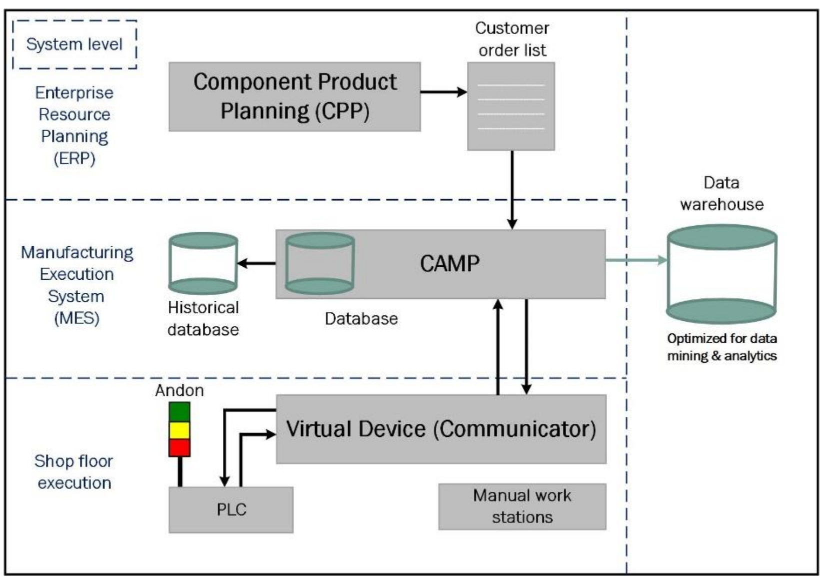


Figure 4-4 Detailed structure of the production IT systems

4.2 Lead time analysis

This chapter contains the results from gathering, analysing and visualizing lead time data. All stages from the selected critical path are covered in chronological order with process-specific results and issues being discussed under each section. For consistency reasons, data was selected from the same time period covering four weeks of production (from the 18th of January to the 22nd of February 2016).

4.2.1 Raw material storage

Figure 4-5 shows the process of raw material intake from external suppliers. Trucks with raw material park outside the factory gate. Goods documentation is submitted to the goods reception office at the gate, where data (advice note number, item number, and quantity) is entered manually into the logistics IT system. After the truck is released through the gate, a forklift driver assists with unloading of material to the outside storage. Material handlers working in the material storage collect material from the outside storage and place it in the inside storage area. Production material handlers collect material from the inside storage to feed the production lines.

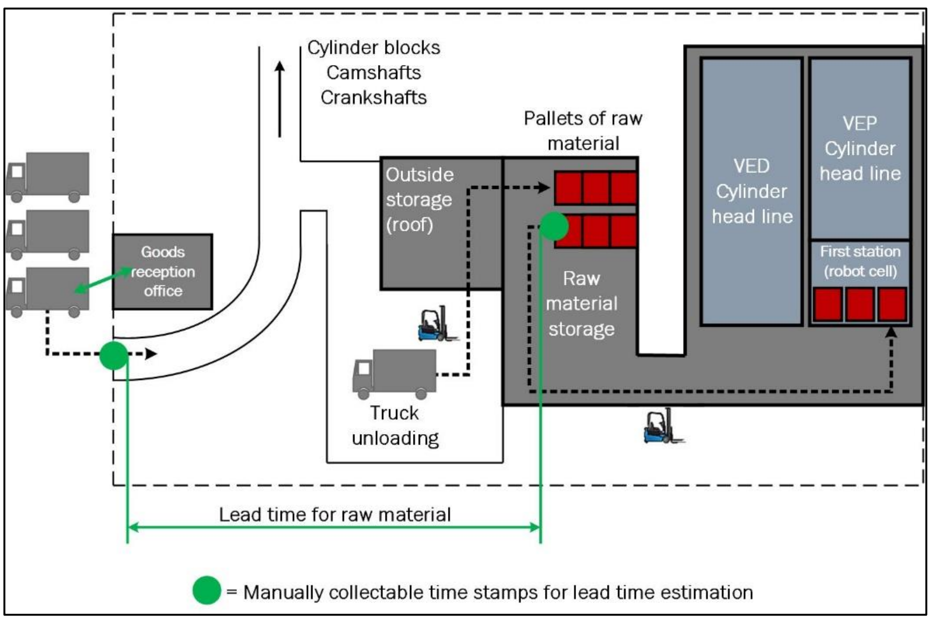


Figure 4-5 Raw material storage layout and manually collectable time stamps for lead time

All raw materials to be machined must acclimatize before being released into machining, due to measurement requirements (extremely low tolerances) as metal expands when heated and detracts when cooled. A specific target temperature has not yet been established (currently under investigation within the organisation), however all material must acclimatize for 24-48 hours in order to reach room temperature. There is no specified maximum storage capacity for the physical storage area. Nonetheless the material planning system uses a parameter of 48-hour maximum storage time for each item.

For additional details on the structure of the raw material storage see **Appendix C** and **Appendix D**, which contain interview summaries with employees working with production planning and material handling.

Since there was no MES in the raw material storage, it was impossible to use time stamp data as intended for measuring lead time. Instead, an alternative way of manually gathering data was needed. A process mapping was conducted to understand how the raw material storage worked, and to assess the possibilities for using documentation of activities to measure lead time. Figure 4-5 illustrate two discrete points, which could be used.

Throughout the process of raw material intake, cylinder heads are stored on the same pallet that they arrive on (in quantities of 30 pieces per pallet), which is labelled with advice note number, item number and quantity. When the pallets are fed to production, the labels are removed. By documenting the removal of the labels, the end time in raw material storage could be recognised. The information on the labels could then be used to trace the pallet back, through documentation in the IT logistic system, which gave the start time for when the particular pallet entered the factory gate. The procedure did not provide individual traceability, since serial numbers for each item are not available from the documentation at the goods reception office. However, it was possible to measure the lead time for each pallet in this way.

Documenting the removal of pallet labels was done by the material handler responsible for feeding production with material. Before discarding the label, the material handler documented the information needed together with current date and time. Information regarding 83 pallets was accumulated over a period of three days (From the 30th of March to the 1st of April 2016).

The lead time for each pallet was calculated by subtracting the arrival time at the goods reception gate from the time where the pallet was fed into the machining line. The average lead time for all pallets was then calculated, and the resulting lead time was 184.35 hours. It is of course relevant to relate the lead time value to other parts of the system, which is done in the later section of 4.2.10 Accumulated factory lead time.

Since lead time data for this stage was derived through manual collection of samples for a limited period of 3 days, the amount of data was considered to be too small for deeper statistical analysis. For next-coming parts where more data was available, the lead time analysis is more thorough, including graphs and detailed descriptions of data patterns and trends.

4.2.2 Cylinder head machining

Data was available through a new MES called CAMP_MAC and was received in the form of an Excel file. Data was available for the entire time period under investigation.

Substantial data processing was necessary in order to transform the data into a usable format, but the data processing itself will not be covered here in detail. Two step-lists were created, which are found in **Appendix E** and **Appendix F** to standardize and assist the data processing. One crucial aspect of the data processing was the removal of diesel item data, since the critical path focuses on petrol variants. Incomplete data rows (i.e. missing values) could also be identified, which were caused by engines being taken off the line midway through the process (e.g. due to quality issues detected during inspection). Those missing cell-values could easily be identified and the data rows removed.

Figure 4-6 shows a histogram with lead time up to 75 hours. Data points located beyond 75 hours are compiled in the bar furthest to the right. The histogram shows a large portion of data between 2-7 hours, but it also reveals that smaller portions of the data were scattered up to and over 75 hours. Peaks can be seen around 50-60 hours. The graph displays large variation, the source of which will be further elaborated on later in this chapter.

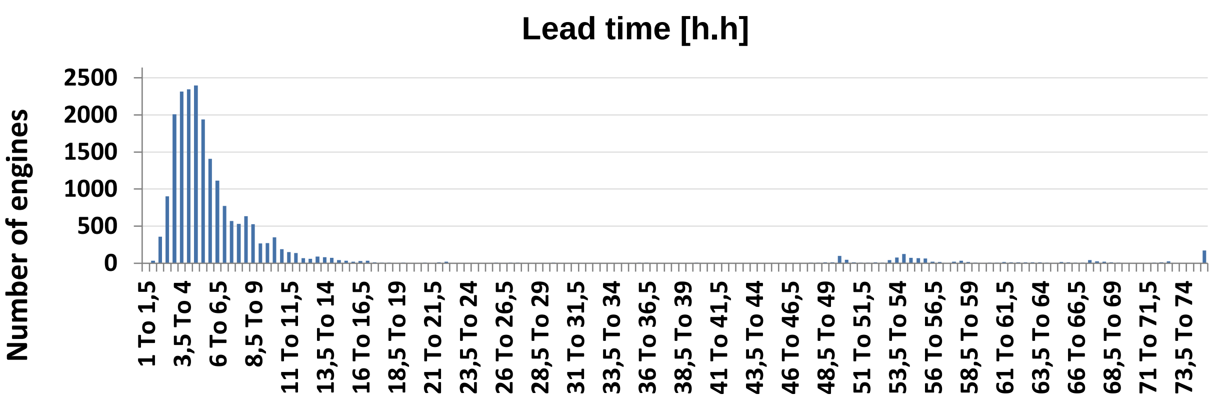


Figure 4-6 Cylinder head machining lead time up to 75 hours

Figure 4-7 shows another histogram with lead time. Here, the X-axis has been truncated at 20 hours to offer a closer look at the bulk of the data. It can be seen that the majority of the data resides within 2-7 hours. However, the truncated plot better reveals a data tail, slowly diminishing from 6 to 15 hours. Large variation exists around the bulk of the data, through what can be described as a tail. The source of which will be further elaborated on later in this chapter.

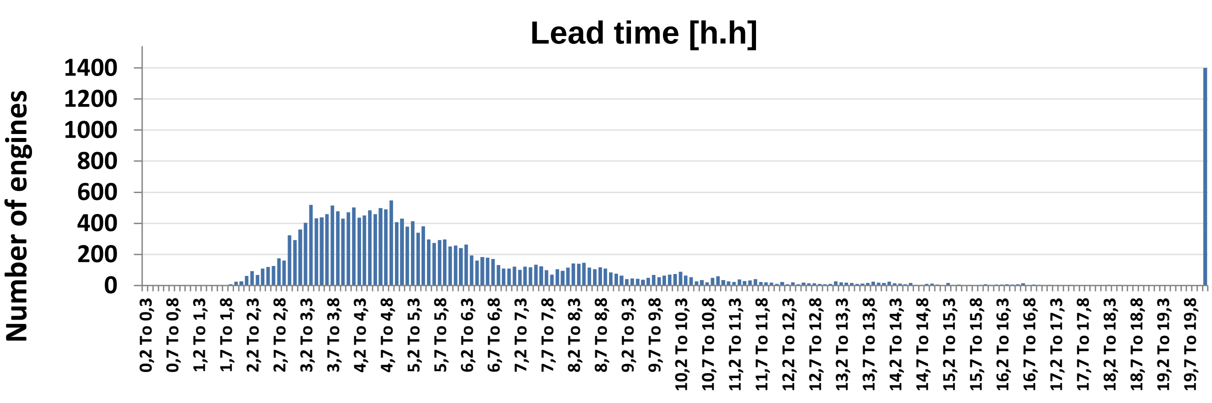


Figure 4-7 Cylinder head machining lead time up to 20 hours

Figure 4-8 illustrates a bar chart with the number of produced items per day corresponding to the height of each bar. Visualization of the data in this way helped to understand production scheduling and execution. No engines were completed on Saturdays, meaning that production was not run on any Saturday. A small number of engines were completed each Sunday, showing that production was run but at reduced tact time.

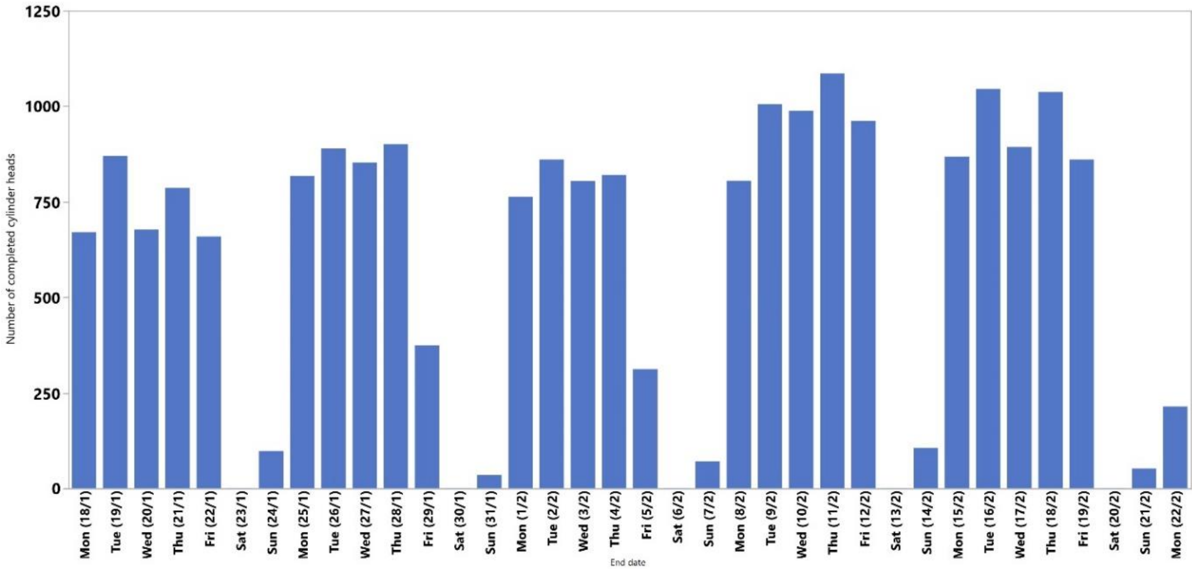


Figure 4-8 Number of produced cylinder head components each day

Figure 4-9 shows a control chart where each point in the top view represents lead time for individual cylinder heads. The bottom view shows the moving range, which is the difference between two successive points.

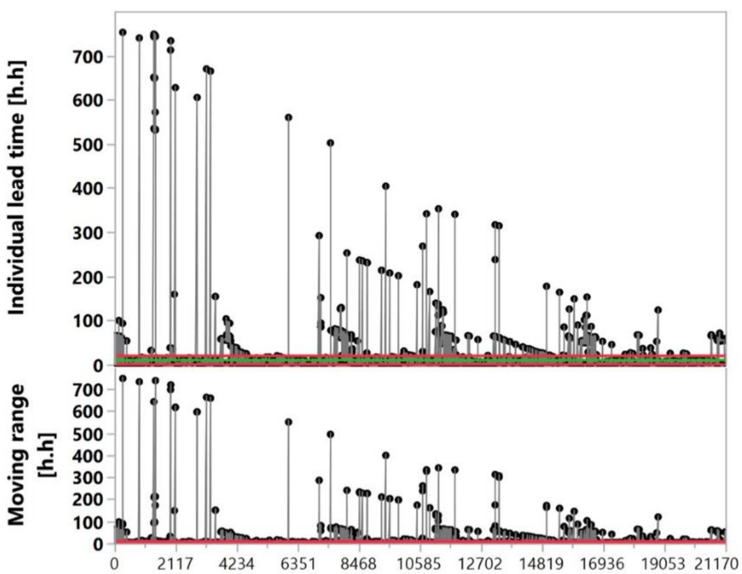


Figure 4-9 Control chart cylinder head machining based on individual lead time measures

Although the control chart is blurry due to the large amount of samples, it still provides relevant insight. Lead times seem longer in the beginning of the time period, with a clear declining trend towards the end. The horizontal line in the upper view indicate the average lead time for the complete time period. The two horizontal lines surrounding the horizontal average line are the natural process control limits. It can be seen that the process was unstable with respect to lead time, since numerous data points were outside the control limits.

The two alternative views normally provided in the control charts complement each other by highlighting different aspects. The top view shows the absolute values, while the bottom chart shows the variation.

Figure 4-10 illustrates another type of control chart with lead time, where data have been categorized into sub-groups. The points in the top view show the average lead time for a sub-group of cylinder heads that were all started to be produced on the same day. The bottom view shows the range, which is the difference between the longest and shortest lead time for the particular day. For instance, the difference between longest and shortest lead time of items started on the 18th of January was more than 700 hours. Some items were started and finished on the 18th of January, while others were not completed until the end of the period under investigation. There are various potential reasons that could have caused this extensive variation in lead time, for instance quality issues with raw material and breakdowns. However, without additional data the exact causes cannot be revealed.

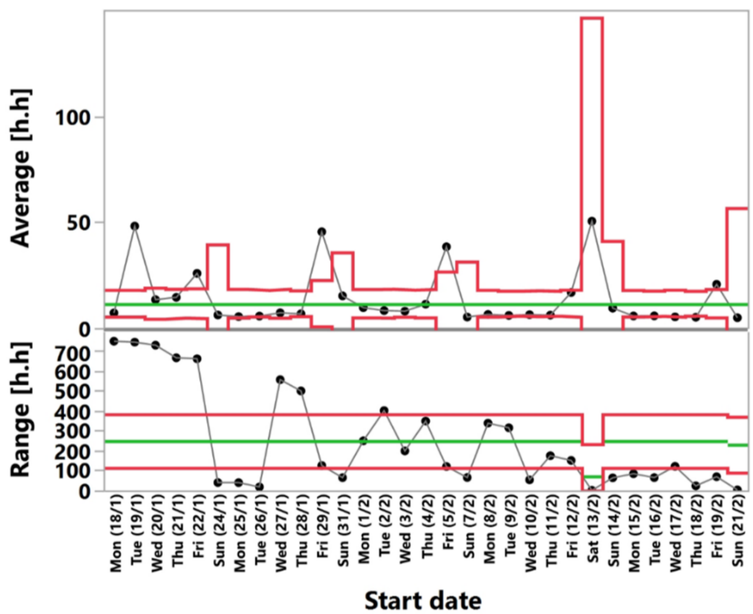


Figure 4-10 Control chart for cylinder head machining per start date

The natural process control limits are computed based on the sample size. For control charts with grouped data (like Figure 4-10), the control limits often vary throughout the period due to the fact that the number of items started in production differ. For example, in Figure 4-10 the control limits from Saturday the 13th were widespread, since there were fewer cylinder heads initiated in production compared to other days.

Sub-grouped control charts can occasionally show miss-leading patterns. Therefore, it is always crucial to assess the consequences of creating sub-groups. Figure 4-11 illustrates a control chart over lead time, but in this case the sub-grouping was based on end date (instead of start date). In the range view, exceptional variation could be found at the end of the period. This should be compared with Figure 4-10, where the range view showed exceptional variation in the beginning.

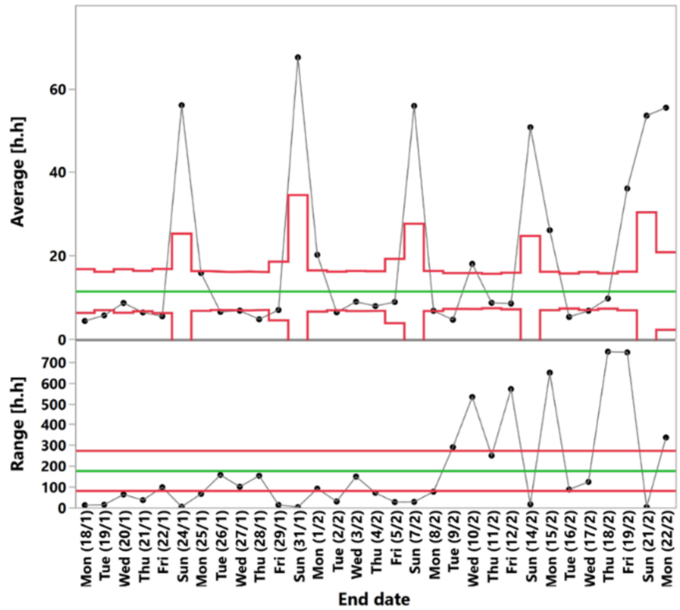


Figure 4-11 Miss-leading control chart for cylinder head machining per end date

The control chart based on start date (i.e. Figure 4-10) was easier to use for the analysis. By grouping on end date, the cylinder heads were displayed on finishing production dates, but the disturbances that caused the long and extensively varying lead times likely occurred prior to the end date. Thus, it was believed easier to track disturbances based on start dates.

The average lead time for this process stage was 11.48 hours. This may be related to the compiled lead time results along the whole critical path, which can be found in the later section of 4.2.10 Accumulated factory lead time.

Figure 4-12 illustrates the analytical procedure used to examine the control charts, which was done by using the JMP statistical software.

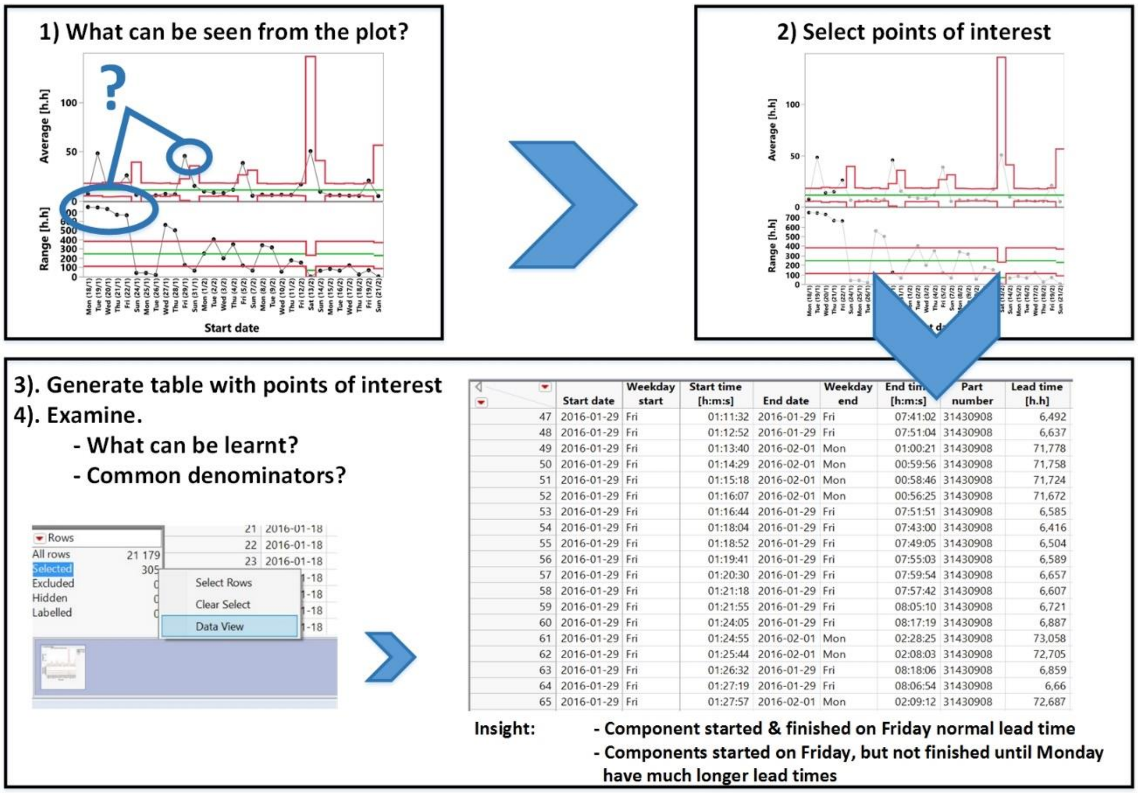


Figure 4-12 Analytical approach to assess data points of using JMP

In step 1, the control charts were closely examined, searching for patterns, trends and abnormal data points. Special attention was paid to exceptional variation (i.e. points outside the control limits).

In the JMP software, there are embedded links between the plot and the original data table, which were utilized to analyse abnormal data points. In step 2, one or several interesting data points were selected. Depending on the sample size, it was more or less appropriate to select several points simultaneously (if there was too much data, it was easier to examine point by point).

In step 3, a new data table was generated based on the points of interest. The bottom left picture shows a screenshot from JMP with one point from step 2 marked (the selected point was 29th of January). Through "data view", a new table was generated, which only contained the selected 305 rows of data.

In step 4, the new data table was examined in search for root-causes to the exceptional variation utilizing additional columns of data. The data table exemplified in Figure 4-12 provided some crucial insights. Since the control chart was sub-grouped on start date, all the data points represent cylinder heads initiated on the same day. From the "weekend end" column (which specified the week-day of the end date column), it was discovered that some cylinder heads were finished the same day, while others were completed on the next-coming Monday. It seemed like some cylinder heads were "trapped in the system" over the weekend, which caused the long lead times. The difference in lead time between cylinder heads completed either on Friday or Monday was large, which caused extensive variation in lead time.

The analytical approach described above was used to examine cylinder head machining, base- and final assembly. In all cases, the analysis pointed out the same issue of items being “trapped in the system” over weekends. In such cases production was not run at all or run with reduced tact time (due to fewer operators working the weekend shifts).

The next step of the analysis was to separate the data based on the insight about “trapped items”, then plot the separated data as new control charts to enable iterative analysis. The iterative analysis approach was thought of as “peeling the layers off an onion” to understand underlying parameters. By removing “trapped data” and making new plots, it became possible to detect new patterns, trends and abnormalities, which were no longer hidden by the impacts of the first discovered weekend-parameter.

Based on the idea of iterative analysis two scenarios called data modification were created. In data modification one, weekend data was removed and the separated data analysed. The second data modification was based on the findings from the first data modification, and will thus be described after the results from data modification one.

Data modification 1 (DM1):

Whenever production was stopped over weekends, some items that were started at the end of one week yet not finished until the beginning of the coming week. Assuming production would be run on full speed over weekends, this problem would not arise. However, when this occurs it skews (increases) the lead time average. Looking solely at the average lead time for the whole period of investigation would be decisive.

To increase the understanding about the reasons for lead time variations (i.e. what causes long lead time for certain time periods), this issue was first thought to be solved by filtering out “trapped items”. However, it proved difficult to construct such a filter appropriately. Firstly, that required either manual identification and removal of each “trapped item, or to filter out a selected time period. The method of filtering out a certain time period could be viable, assuming the time periods could be based on a fixed difference between shifts. The difference could be identified from the production schedules over shift patterns and tact times. However, when comparing schedules of production with the bar charts showing the number of items started each day, it was found that production was not executed according to plan.

Trials were made by filtering on start time, where data from a specific time on a Friday was removed. However, some items remained that were initiated on a Thursday and were not completed until after the following weekend. Instead, one could possibly remove all items completed in the beginning of Monday, but the problem was the same here. Some items that were started on Friday was not finished until Tuesday.

Due to the problems of filtering an alternative method was used, where the data was modified by excluding all lead time data exceeding 12 hours. The major issue with such an approach is where to set the limit. This was done based on the distribution of the data. The goal was to discard of peaks around 50-60 hours (as was shown in Figure 4-6) representing items “trapped” over weekends, while not cutting into the tail of the data (better shown in Figure 4-7). The benefit of this method is that the majority of “trapped” lead time data is discarded of. The drawback of this method, however, is that potential "true outliers" that could be of interest for the analysis, are discarded as well.

The resulting average lead time for DM1 was 5.33 hours. For compiled lead time results along the entire critical path, see the later section of 4.2.10 Accumulated factory lead time.

Figure 4-13 illustrates the control chart, based the separated data containing only lead times up to 12 hours.

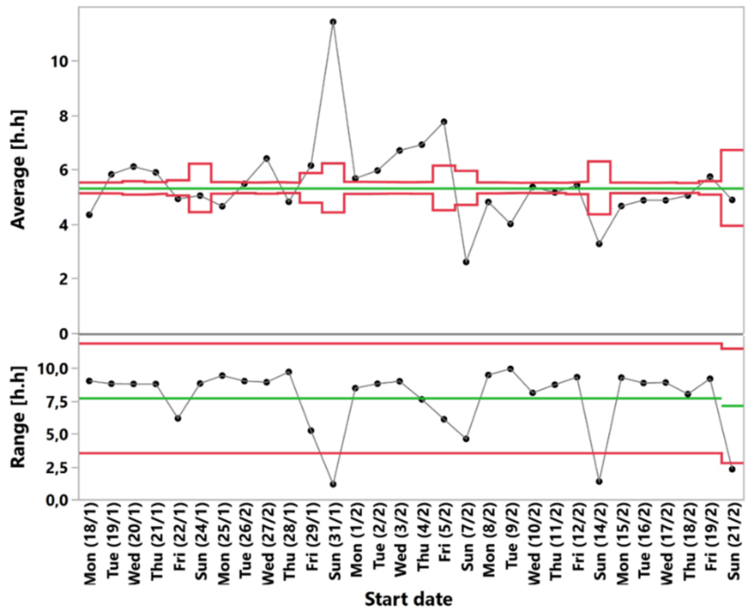


Figure 4-13 Control chart cylinder head DM1 based on average lead time per start date

Compared to the control chart of unmodified data (as seen in Figure 4-10) it can be observed that the creation of DM1 drastically lowered and narrowed the upper- and lower natural process limits for both the average and range views. Remnants of weekends might still be visible on the 31st of January when 43 engines were started between 22:00 and 00:00 on a Sunday and finished at approx. 10:00 on a Monday. From the data it seemed like production was stopped overnight resulting in the abnormally long lead times. Without confirmation such conclusions cannot be drawn as work could also have been carried out overnight at a reduced tact. However, the length of the time period (resembling a regular shift) points to halted work.

From the example above it can be seen that the 12-hour limit is no guarantee for removal of all "trapped items" as work can e.g. be halted over an 8-hour nightshift, not always an entire weekend. Potentially, another data modification could be done with a slightly lower limit to get rid of those long lead times. However, this was not considered to provide much new insight. Instead, it would be more interesting to reverse the situation and move swiftly over to assess the "ideal state", with the shortest lead times. Focusing on the shortest lead times was an attempt to better understand the prerequisites for achieving short lead times. Based on this reasoning, data modification 2 (DM2) was created to represent the ideal state. Another control chart was created based on "ideal state data", which was attained by separating data with a 4-hour cutting limit (instead of 12 hours as in DM1). The results of DM2 will now be described.

Data modification 2 (DM2):

As could be seen by looking at the distribution of the cylinder head data (in Figure 4-7) lead times are somewhat spread, with what can be described as a bumpy tail fading to the right (resulting in longer lead times). Instead of peeling of portions of the tail by creating new data modifications, it was decided to modify the data by trying to get as close to the ideal state as possible.

For consistency reasons the limit of 4 hours was used when developing DM2 for cylinder head, base- and final assembly. As with the 12-hour limit the, 4-hour limit was selected based on the distribution plot. The limit was also selected since it discarded all "trapped" items. As with DM1 modifying the data this way has the drawback of eliminating outliers from the relevant kept data.

Figure 4-14 illustrates the resulting control chart, which was created based on the separated data excluding all lead time data above 4 hours. Compared to the control chart in DM1 (Figure 4-13) the upper- and lower natural process limits for both the average and range have been further lowered and narrowed.

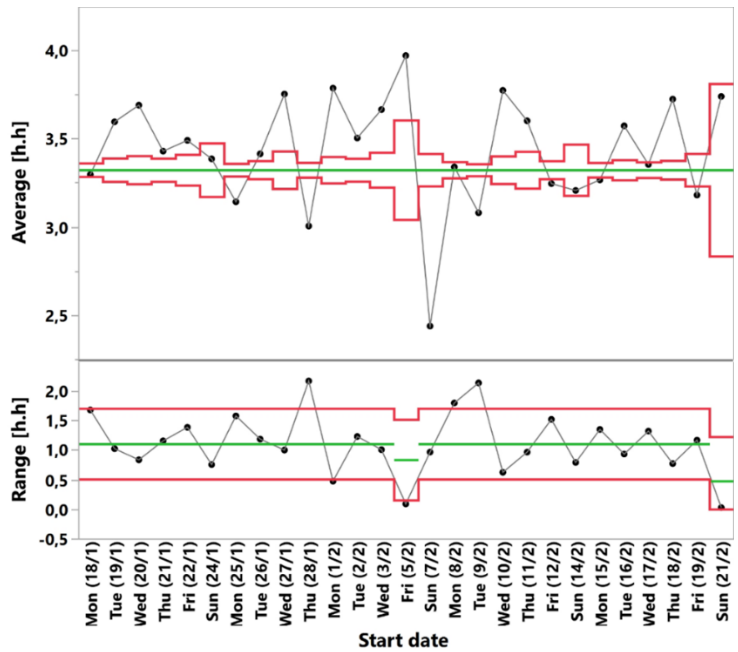


Figure 4-14 Control chart cylinder head DM2 based on average lead time per start date

Abnormally long lead times that can be contributed to "trapped" items and confirmed as such are not visible anymore. At this point, the next step was to examine the control chart in search for root-causes to the exceptional variation, trends and patterns. However, analysing the ideal stage was hard, since the analysis was dependent on access to additional data, which could be used to explain root-causes. Due to collection of data being difficult and time consuming, this data could not be extracted within the time frame of the thesis. Therefore, deeper analysis and cross comparing with other kinds of data, such as maintenance schedules, disturbances etc., had to be left out.

The resulting lead time for this process stage was 3.32 hours. For compiled lead time results, see the later section of 4.2.10 Accumulated factory lead time.

4.2.3 Cylinder head buffer

Data for the cylinder head buffer originates from the same MES as cylinder head machining. Therefore, descriptions of data extraction and processing from the cylinder head machining apply here, with no further explanations needed.

Since the cylinder head buffer is connected to cylinder head machining and base assembly, the lead time for the buffer is affected by the way production is managed in those stages. Therefore, it was regarded sufficient to analyse lead time in detail for cylinder head machining and base assembly. However, a histogram was still included since it provides a good overview of the lead time data.

Figure 4-15 shows a histogram of lead times up to 150 hours. The data shows variation throughout the spread of the data with large peaks around 20 and 35 hours. Several lower peaks are also visible around 50, 65, 80 and 95 hours.

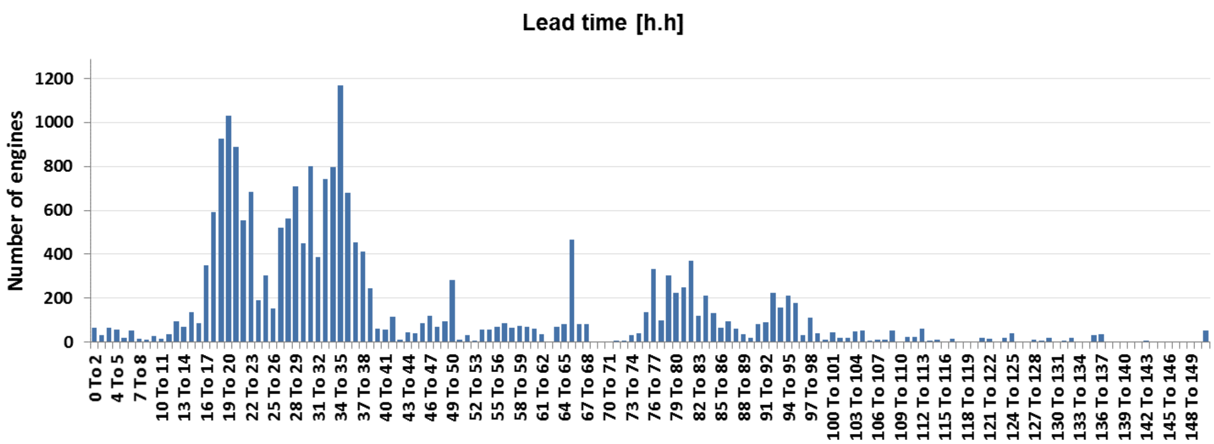


Figure 4-15 Histogram for cylinder head buffer data, up to 150 hours

The cylinder head buffer is target to hold 36 hours worth of material for the following process of base assembly. Thereby, the lead time target for the buffer is also 36 hours. Occasionally, the buffer might be strategically increased to cope with upcoming disturbances or imbalances between cylinder head machining and base assembly.

The resulting lead time for this process stage was 43.12 hours. For compiled lead time results, see the later section of 4.2.10 Accumulated factory lead time.

The histogram (see Figure 4-15) showed that the majority of the data was condensed within the target of 36 hours. However, a substantial amount of items stayed longer than the targeted 36 hours. The average lead time also exceeded the target by 7 hours.

Thus, it could be determined that the target of 36 hours of lead time in this buffer was not achieved during the time period.

4.2.4 Base assembly

Data for base assembly originates form a new MES called CAMP_IM6. However, the description of data extraction and processing from the cylinder head machining applies here, with no further explanations needed (due to CAMP data being structured in the same way).

Figure 4-16 shows a histogram with lead times up to 75 hours (all lead times beyond 75 hours are compiled in the bar furthest to the right). The data shows a large portion of data concentrated around 3 hours with a slight tail to the right fading out around 9 hours. The graph shows some data spread, with peaks around 18, 35, and 48 hours.

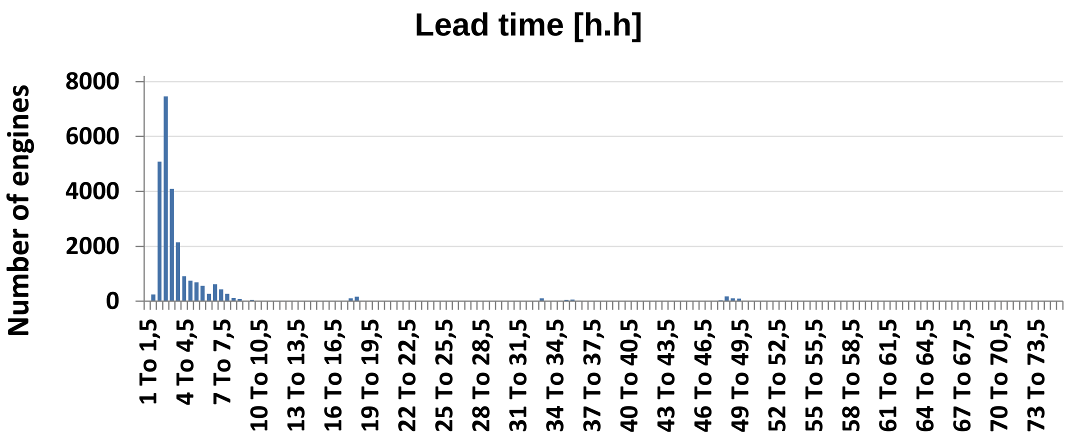


Figure 4-16 Histogram for cylinder head machining data, up to 75 hours

Figure 4-17 illustrates a control chart of lead times. Exceptional variation, with regards to both the average and range views, could be identified on several occasions (many of which occurs on Fridays or Saturdays).

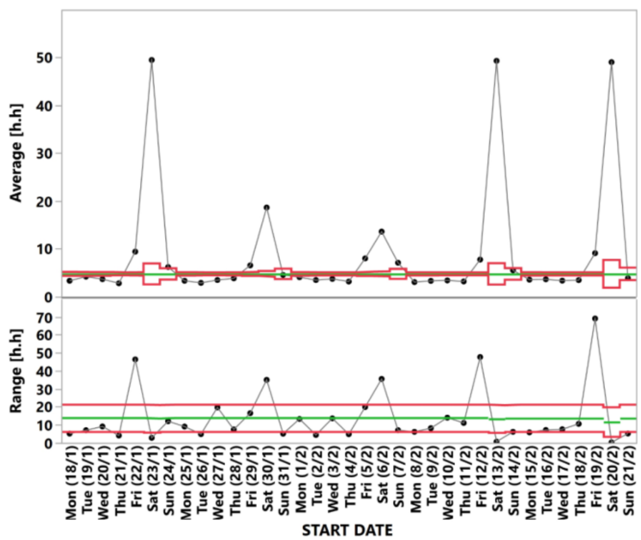


Figure 4-17 Control chart base assembly based on average lead time per end date

By looking at a bar chart graph displaying the number of completed engines each day (the same chart as used in cylinder head machining, see Figure 4-8), it became evident that production was run on both Saturdays and Sundays (it was decided not to include the bar chart here). The number of completed engines was substantially lower on weekends, which indicated that production was run on reduced tact. As in cylinder head machining items were "trapped" due to halted work or reduced takt over weekends causing the lead time to peak during those days.

The resulting lead time for this process stage was 4.76 hours. For compiled lead time results, see the later section of 4.2.10 Accumulated factory lead time.

DM1:

As previously described in the cylinder head machining chapter, DM1 was introduced as a method for discarding of items "trapped" over weekends in the lead time data (for more detailed descriptions of this see previous chapter 4.2.2 Cylinder head machining)

Figure 4-18 shows a control chart for DM1 based on average lead time per start date. Compared to the control chart of unmodified data (i.e. Figure 4-17), the upper and lower natural process limits have been substantially lowered and narrowed.

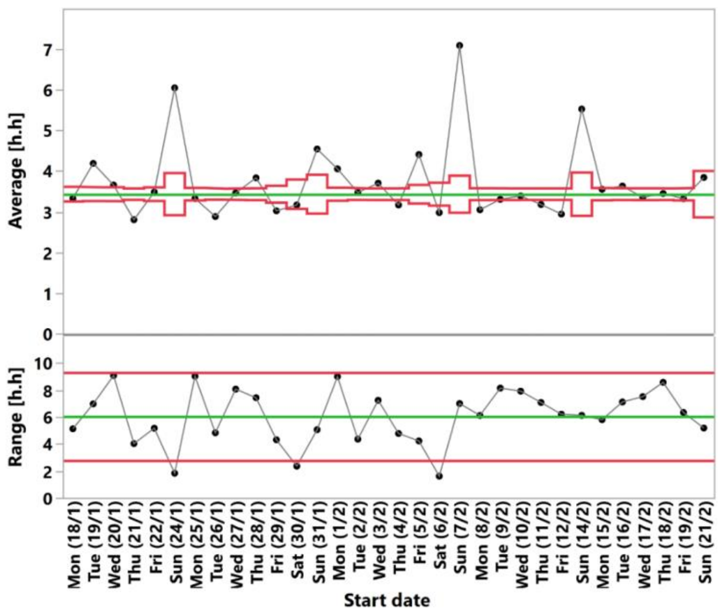


Figure 4-18 Control chart base assembly DM1 based on average lead time per start date

According to Figure 4-18, what can be construed as remnants of weekends is still visible on Sundays. By using the embedded links in JMP between the control chart and the underlying data table, abnormalities seen in the control chart could be examined. Some interesting aspects could be seen. For instance, on the 15th of February 75 engines were started between 22:00 and 00:00 on a Sunday and finished at between 02:00 and 07:00 on a Monday. It is difficult to distinguish between halted work and reduced tact as the cause for abnormally long lead times, needing conformation through additional data. However, due to the shortness between starting and finishing times the nightshift can be presumed to have been run on reduced tact.

The resulting lead time for this process stage was 3.43 hours. For compiled lead time results, see the later section of 4.2.10 Accumulated factory lead time.

DM2:

As previously described in the cylinder head machining chapter DM2 was introduced as a method for getting as close to the ideal state as possible. That was done by removing all lead time data above 4 hours thereby discarding of trapped items (for more info review the previous more detailed descriptions in 4.2.2 Cylinder head machining).

Figure 4-19 shows the control chart for DM2 based on average lead time per start date. Compared to the control chart for DM1 (Figure 4-18) the upper and lower process limits for both the average and range views have been further lowered and narrowed. In the ideal state, it could be observed that the process seemed relatively stable in the range view, but not in the average view. In the average view, exceptional variation was detected on many occasions. For instance, Sunday the 7th had substantially longer lead times than on regular weekdays. However, the range was smaller than usual, which meant all engines started on Sunday the 7th had long lead times. On Sundays operations were run with fewer employees resulting in longer tact time and more variation in production.

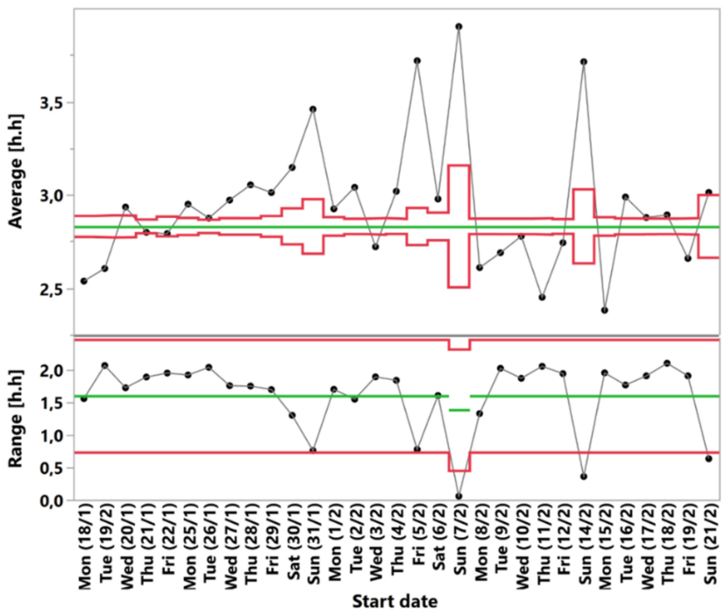


Figure 4-19 Control chart base assembly DM2 based on average lead time per start date

The resulting lead time for this process stage was 2.83 hours. For compiled lead time results, see the later section of 4.2.10 Accumulated factory lead time.

4.2.5 Intermediate buffer

Available data for the intermediate buffer originates from an old MES called KSMB (kransystem för motorbuffert), and was presented through Excel in a way making extraction of lead times very difficult. Figure 4-20 illustrates an example from the data received. Firstly, only 5000 lines of data can be saved and exported from the system at a time. Due to the limitation of 5000 exported lines the data has been cut off resulting in missing infeeds (inlag) at the beginning of the data table and missing outfeeds (dest) at the end of the data table. Secondly, data for in- and outfeed is displayed in separate lines. Lastly date and time is stored in a format unrecognized by Excel hindering the use of formulas to calculate lead time.

IDATE (ÅÅÅÅVVD)	ITIME (TTMMSSHH)	HISTORY_TYPE	SERIENR
2016115	4074531	INLAG	1588213
2016115	7345112	DEST	1588213
2016123	8124606	DEST	1588865
2016123	8132294	DEST	1588866

Figure 4-20 Excel data example from KSMB Manufacturing Execution System

Lead time extraction through the use of the KSMB MES data was therefore not feasible (without applying advanced programming to re-construct the data table). An alternative method was proposed where the outfeed from base assembly and the infeed from final assembly were used instead (with the intermediate buffer in-between). This made it possible to avoid working with the old data, since base- and final assembly use new MESs.

The selected method utilizing new system data was not without complications. The issue of having date and time data in a format recognizable by Excel was solved, since that was not a problem with data from the new MESs. Through this method, data for all of the period under investigation was made available. Otherwise the data had been limited to week 11 and 12 of 2016, which was the time period available in the old MESs.

However, a new problem was discovered since there was one input flow from base assembly to the intermediate buffer, but two output flows from the intermediate buffer to the petrol final assembly modules. Since there was only data for one of the two final assembly modules available, all in- and outfeed data rows could not be matched together (two-by-two). Also, in- and outfeed data was still stored on separate lines, which made it difficult to match the individual time stamps for in- and outfeed to derive the lead time.

To overcome these issues, it was decided to randomly select two samples from each working day (Monday-Friday) during day and evening shifts throughout the entire period under investigation. In total, this resulted in a total of 50 samples, which were used to estimate the lead time. Selecting only 50 samples was a sufficient compromise, since the main purpose was to examine the possibilities for collecting data in this way. Therefore, the approach constitutes a proof-of-concept that lead times can be extracted using the new system data in this way. Instead of manually picking samples in this way, a software could be developed to do automatically what was now done by hand.

The average lead time for the sample group was calculated, and the resulting average lead time was 3.37 hours. For compiled lead time results, see the later section 4.2.10 Accumulated factory lead time.

4.2.6 Final assembly

Data for final assembly originates from a new MES called CAMP_XM. The description of data extraction and processing from the cylinder head machining applies here, with no further explanations needed (due to CAMP data being structured in the same way).

Figure 4-21 shows lead times up to 20 hours. The data seems to be concentrated approximately in the interval of 1-2 hours. The histogram also revealed a spread of data, with some peaks visible around 8 and 16 hours.

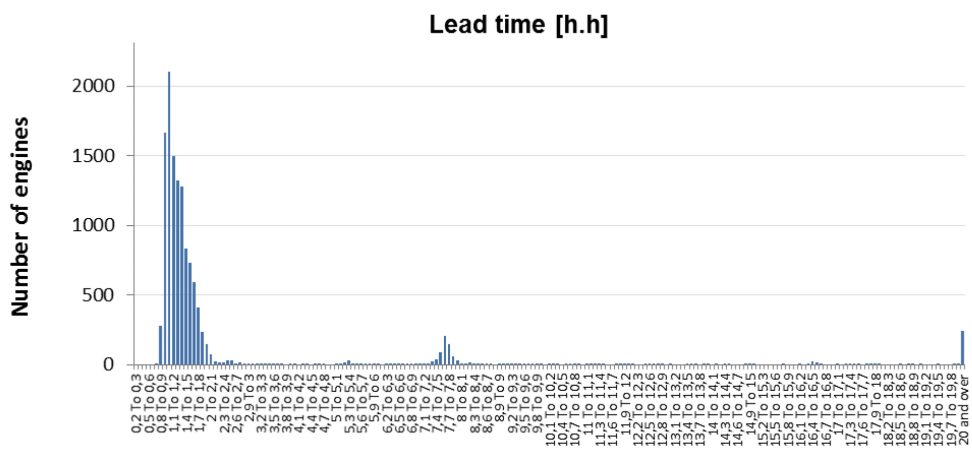


Figure 4-21 Final assembly lead times

Figure 4-22 shows a control chart where exceptional variation can be identified on four occasions. All occurred either on Fridays or Saturdays.

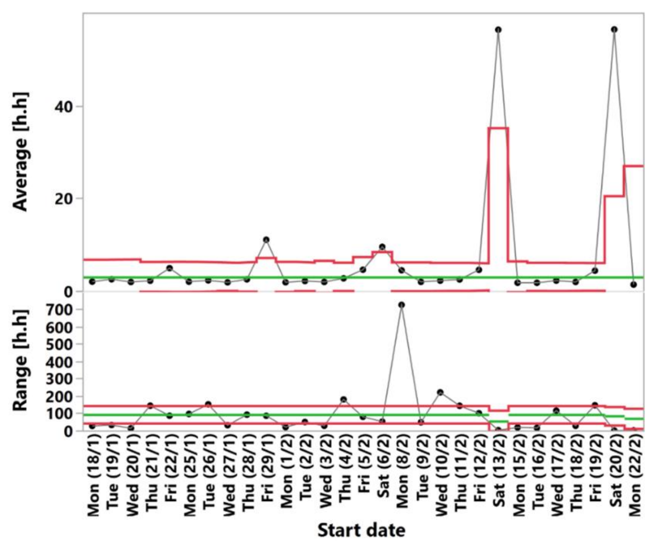


Figure 4-22 Control chart final assembly based on average lead time per start date

By looking at a bar graph displaying the number of completed engines each day in final assembly (the same chart as the one used in cylinder head machining, see Figure 4-8), it became evident that no engines were produced on Sundays (thus production was not run). A small number of engines were produced each Saturday except of one. As for cylinder head and base assembly, items seemed to be “trapped” in the system over weekends causing exceptional variation and subsequently longer lead times during those days.

The resulting lead time for this process stage was 3.01 hours. For compiled lead time results, see the later section of 4.2.10 Accumulated factory lead time.

DM1:

Figure 4-23 shows a control chart for DM1 based on average lead time per start date. Removing all lead times above 12 hours narrowed and lowered the natural process limits. No clear remnants of trapped data over weekends were visible when examining data points in DM1. From the range view, exceptional variation could only be identified on two occasions (Tuesday 19/1 and Friday 12/2).

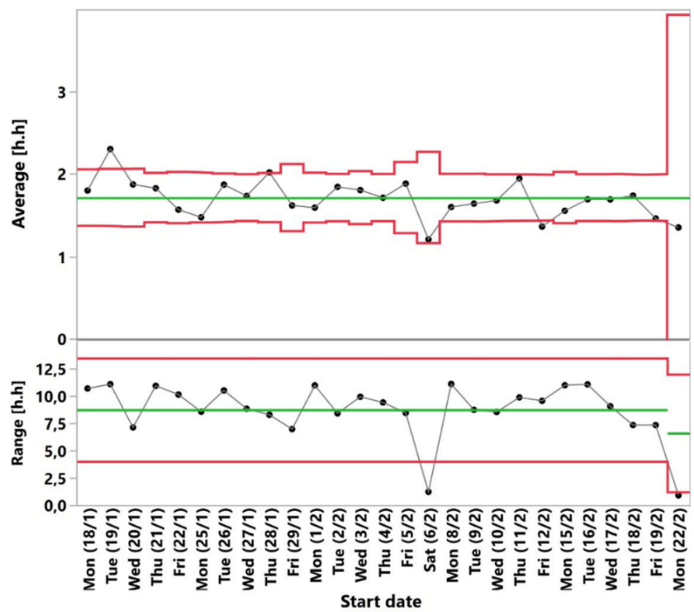


Figure 4-23 Control chart final assembly DM1 based on average lead time per start date

The resulting lead time for this process stage was 1.71 hours. For compiled lead time results, see the later section of 4.2.10 Accumulated factory lead time.

DM2:

Figure 4-24 shows the control chart for DM2 based on average lead time per start date. When compared to the control chart for DM1 (see Figure 4-23) it can be observed that lead times in DM2 are statistically stable in the range view, which is not the case in DM1. As in the case of previous DM2s the process limits have been narrowed and lowered.

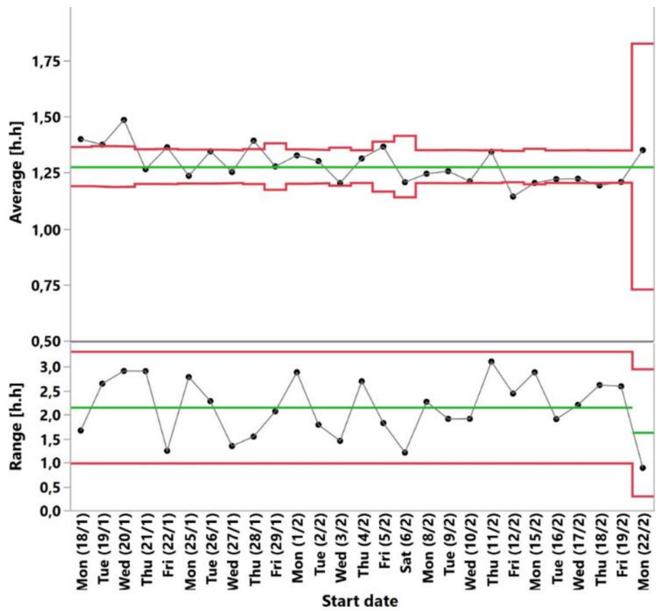


Figure 4-24 Control chart final assembly DM2 based on average lead time per start date

The resulting lead time for this process stage was 1.28 hours. For compiled lead time results, see the later section of 4.2.10 Accumulated factory lead time.

4.2.7 Assortment buffer

As was the case with the intermediate buffer, available data for the assortment buffer comprised of data extracted from an old MES. The MES for the assortment buffer is called KS1S.

The lead time extraction method previously used for the intermediate buffer was not feasible for the assortment buffer. The method used for the intermediate buffer utilized the outfeed from the stages before and after to derive the lead times. In this way, it was possible to avoid the old systems, but it required that preceding and following stages were both supported by new MESs. For the assortment buffer, this was not the case. The preceding process runs on a new MES, but the after following process was supported by an old MES.

The benefits of using the same method as for the intermediate buffer would be to be able to cover the whole period under investigation, and to overcome formatting issues with time and date.

Certain limitations come with the use of data from the old MES KS1S system in question. As for the intermediate buffer, a limited number of data lines could be saved and exported at one time. For the KS1S system used for the assortment buffer, the limit is 4000 lines (compared to the intermediate buffer that could store 5000 rows). Also in- and outfeed data was stored on separate rows, with date and time in unrecognizable form by Excel.

The data made available was from the 11th of April of 2016. For conformity reasons the decision was made to use the same sample size of 50, selecting from day and evening shifts, as with the intermediate buffer data giving an indicator to the average lead time.

The average lead time for the sample group was calculated, and the resulting average lead time was 0.42 hours. For compiled lead time results, see the later section of 4.2.10 Accumulated factory lead time.

4.2.8 Final storage

Data for final storage and the packaging and dispatch area was extracted and received together in one Excel file. The data originates form an old MES called MUS. Previous problems associated with old system data resurface once again here, meaning that the data for in- and outfeed was presented on separate rows and that date and time was stored in a format unrecognized by Excel.

For final storage and the packaging and dispatch area, data was available for the entire time period. For the same reasons as with the intermediate and assortment buffers a sample size of 50 was chosen with random samples taken from each working day, during day and evening shifts.

The average of the 50 samples was calculated and the resulting lead time for this process stage was 0.48 hours. For compiled lead time results, see the later section 4.2.10 Accumulated factory lead time.

4.2.9 Packaging and dispatch area

Due to the fact that data for the packaging and dispatch area was extracted from the organizations MUS system and processed together with data for the final storage, description of data extraction and processing from the final storage apply here, with no further explanations needed.

The average for the sample group was calculated, and the resulting lead time was 4.41 hours. For compiled lead time results, see the later section of 4.2.10 Accumulated factory lead time.

4.2.10 Accumulated factory lead time

Table 4-1 illustrates the total accumulated lead time along the critical path in the factory, as well as the relative distribution of lead time between processes, buffers and storages. All lead times were modified by multiplying them with an un-disclosed scaling factor in order to accommodate confidentiality. Relations in the data remained intact, but the actual lead times are not revealed.

The methods for data collection various, depending on the MES available at the particular stages along the path. In the raw material storage, no MES was available and lead time data had to be collected by manual observations utilizing production personnel for documentation (in this case 83 observations were collected over three days of production). Cylinder head machining, cylinder head buffer, base- and final assembly operate on new MES from which data was easily accessible. The intermediate buffer, assortment buffer, final storage and packaging and dispatch area operated on old MES, were it was proven much more difficult to collect the lead time data.

Table 4-1 Accumulated factory lead time

LEAD TIME (HOURS)	All data (unmodified)	DM1: Data < 12 hours	DM2: Data < 4 hours	Data through observations
Raw material storage				184,35
Cylinder head machining	11,48	5,33	3,32	
Cylinder head buffer	43,12			
Base assembly	4,76	3,43	2,83	
Intermediate buffer	3,37			
Final assembly	3,01	1,71	1,28	
Assortment buffer	0,42			
Final storage	0,48			
Packaging & dispatch	4,41			
Total Lead time (hours)	255,40	246,62	243,58	
Total Lead time (days)	10,64	10,28	10,15	

For the intermediate buffer, assortment buffer and packaging & dispatch area, the data received was too complicated to process. Therefore, 50 observations were randomly selected to provide an indication of the lead time for those stages. No data modifications were used for storages, buffers and the packaging & dispatch area, since there are no shifts there that explicitly effect lead times. Data modifications could possibly be developed for those areas as well, but with the need for additional data to understand the causes for exceptional variation.

When comparing the summarized lead time for all storages and buffers to the total lead time, it could be seen that 91 % of the time (231.7 hours out of 255.4 hours) was spent in storages and buffers. Focusing on the row of total lead times, the lead time decreased with 3.4 % between unmodified data and DM1. Between unmodified data and DM2 the decrease in lead time was 4.6 %. However, when focusing on the decrease in lead time for processes with data modifications some interesting aspects can be seen. The lead time for cylinder head decreased by 71 % between unmodified data and DM2. For base assembly the lead time decreased by 40.5 % between unmodified data and DM2 and in final assembly the decrease was 57.5 % between unmodified data and DM2.

5 Discussion

The chapter includes an elaborated discussion about data collection and process of lead time, lead time analysis, and future research recommendations.

5.1 Data collection and processing

The current structure of the organization and its MES made the extraction of lead time data difficult. Communication paths are long, requiring connections between multiple departments or employees in order to gain access to data sources. Furthermore, extraction of lead time data from the MES was carried out by external consultants, further extending communication paths.

Generally, manufacturing stages are supported either by new or old MES. The raw material storage was an exception, since it was not supported by any system. Through manual collection, 83 samples were documented over a period of three days. In order to gain further insight into lead times in raw material storage a longer lead time sampling study has to be carried out. Accumulating more data over a longer time period would make the lead time data more representative, thus providing sound ground for in-depth statistical analysis. Such data could help the organization to better analyse conditions for varying and long lead times. The data could also be used to construct and run DES models on raw material storage for optimization purposes, which have not been done before.

The new MES delivered data that was easily processed. However, data from the old systems came with certain limitations, which were not overcome since they required specialized IT competence. Two examples are that date and time formats were unrecognizable by Excel and in- and outfeed for each item was stored in separate lines. Since these issues could not be solved within the given timeframe, it was decided to be sufficient with respect to the purpose of the thesis to manually select 50 samples. The samples were manually processed to overcome the issues and estimate the lead time. For larger data sets, manual transformation becomes impractical. The fact that these problems could not be solved in time also demonstrates the extensive time consumption required for collection and data processing.

When discussing the issue of old system data with IT-department personnel, they seemed to possess the competence to solve the issues associated with old systems. Therefore, one of the largest obstacles for extracting and processing lead times from the old MES would be overcome with reasonable efforts. The reason for not being able to solve the issues was due to time constraints, relating both to the limited time for the thesis and to the high workload at the IT-department. The IT department could not spare enough resources, since they had to prioritize on-going projects in the organization.

For both new and old systems, lead time data is temporarily saved and exported to historical databases, activities which are done embedded in the MES (for more details on the IT systems structure, see Figure 4-4). This setup allows for rapid execution within the MESs. However, the structure is not suited for data collection and analytical purposes, and it also requires lead time data inquiries to be run overnight or weekends not to slow down the MES.

There are different ways of coping with this issue and to improve the collection of data from the MESs. Functionalities could be added to both new and old systems, which would enable easier and even automatic extraction of lead time data. Data inquiries could be auto-generated, scheduled to be executed overnight and sent to specific email addresses. Thereby, "Fresh lead time data" could be sent on daily basis to relevant departments, such as production- logistics-, and simulation engineering.

This can be accomplished in two different ways. The JMP statistical software could be connected to SQL databases (JMPb, 2016) (such as the organisations MES), allowing for direct importation, processing and analysis of data. The second option is that the IT-department could create a customized software, add a new database separated and connect it to existing MESs. Since data is already exported to an historical (but embedded) database, data could be exported elsewhere at the same time. With this setup, lead time data could be continuously extracted to a new database optimized for data mining and analytics (rather than rapid execution), thus improving collection and processing of data. The issue of having to run specific inquires overnight could be overcome by locating the database outside existing MESs.

From the process mapping and the overview picture of the various IT systems (Figure 4-3) it could be seen that several MESs are currently used in the organization, which make the IT environment highly complex. Historically, additional software or functionalities have often been added onto existing systems to solve specific IT-related issues, which over the years have increased the complexity further. From this viewpoint, it might not be ideal to keep adding additional software to current IT systems. However, the organization has decided to invest in new IT-systems, and are currently investigating the internal requirements for it. The two suggested methods for improving data collection and processing could be used as input when specifying requirements for those new systems.

In DES projects, input data management can take up as much as 40 % of the total time (Skoogh, 2011). As similar input data management, comprising data collection and processing, was required for the lead time analysis there seem to be huge potential savings by simplifying those activities.

Robertson and Perera (2001) introduced four methods for managing input data in simulation (see Figure 3-6). The same logic applies in the case of lead time analysis, where the simulation tool could be replaced by the JMP statistical software. In the first method, data is manually collected, processed and transferred. The second method is semi-automated, with data being manually collected and compiled but read automatically by the analytical tool (i.e. simulation or JMP). In the third method, all steps are automated but there is still an intermediate database. The fourth method allows for automated communication between the analytical tool and the ERP/MES. The time required for data collection and processing progressively declines from method one to four, at the expense of a more complex system setup.

The method used in the thesis corresponds to the second method for managing input data. Data was manually collected, processed and fed into an Excel spreadsheet. The spreadsheet was imported to the JMP statistical software, where the analysis was conducted. It is also relevant to mention here that lead times have never been collected, processed and analysed in this way and to this extent within the organization before.

Earlier in the discussion, two methods were proposed to improve lead time input data management in the organization. The first proposal was to create a customized software, which could automatically handle data inquiries and supply lead time data. The basic idea of the second method was the same, but instead of a customized software, the MES databases could be connected with JMP. When comparing these suggestions to Robertson and Pereras' (2001) input data management methods, it could be seen that they could both help the organization reach an automated setup, which means moving from input data management method two to three. Lead time data could be automatically collected, processed and fed from MES via an intermediate data warehouse to JMP, thus increasing efficiency.

Beside the potential time savings, another advantage could be the storing capacity and access to manufacturing data. Instead of having a limited spreadsheet, the intermediate database could be configured to store all sorts of manufacturing data. Beside lead time analysis, the database could be utilized by other organizational functions, such as simulation engineers, production engineers, lean coordinators and production managers for all sorts of analysis.

Automated communication between the analytical tool and the ERP/MES (corresponding to the fourth method), constitute the most efficient way for input data management (Robertson and Perera, 2001). However, the current structure of the IT systems means the fourth method cannot be reached. The main reason for that is the previously discussed issues of jeopardizing manufacturing execution when making requests to existing databases embedded in the MES. The methods introduced by Robertson and Perera constitute a good base for discussion on requirements of future IT systems in the organization.

5.2 Lead time analysis

Data was collected from MES and processed in Excel. Statistical analysis was conducted with the aid of JMP statistical software, using graphical representations to explore and thereby understand the manufacturing data. There are various graphical representations available, which could be used to analyse manufacturing data. Flow charts, bar charts, pie charts, and cause-and-effect diagrams are examples of quite simple and intuitive plots, which can be used as a base for manual analysis (Wheeler, 2003). The plots are intuitive in the sense that the results can be interpreted without advanced and specialized competence.

In many situations, the graphical representations might be in-sufficient due to the complexity embedded in the data. For instance, changing parameter settings based on what can be seen from a histogram might give un-expected results because of hidden relations and synergy effects between various parameters.

From this perspective, control charts provide a more systematic way of iteratively analysing and separating data (Wheeler, 2003). Unlike the ad-hoc analysis based on simpler graphics or lean thinking, control charts outline a powerful systematic approach for data analysis. In the thesis, simpler graphical representations were initially used for superficial analysis to better understand the data. This provided a sound base, which could then be complemented by the analysis based on control charts.

A drawback with control charts is that it requires specialized competence to interpret the results. The analyst must for instance understand upper- and lower process limits and routine- and exceptional variation. Thus, control charts might be more or less applicable depending on the context. It is a question whether the organization has the right competence or a willingness to learn or not.

Organizations tend to focus too much on improving the average performance to meet targets, instead of first addressing the variability to achieve output consistency (Hammersberg, 2016). There is a great risk of accepting variability as a naturally inherent process feature (Hopp & Spearman, 2008). In many cases, variability reduction might be a more efficient way to elevate performance compared to increasing the capacity (Ignizio, 2009).

Once a process is statistically stable, it becomes easier to efficiently improve the average performance. There are various means to proactively work with process variation. The analytical approach provided through the use of control charts is powerful, since it forces the analysts to identify and eliminate exceptional variation.

The coefficient of variability (CoV) or variability pooling can also facilitate the monitoring and reduction of process variation (Hopp & Spearman, 2008; Ignizio, 2009). The CoV is a unit-less measure, which makes the variation comparable between processes or manufacturing systems. The equation for CoV includes the standard deviation, which means an assumption about normal distribution is required. Since lead time data is not normally distributed, the CoV was not applicable to use in the thesis. Statistical transformation and variance stabilizing tools can be used to cope with this issue, but that was regarded to be outside the scope of the thesis.

In variability pooling, several sources of variability are combined to minimize their impact. One example of variability pooling is the use of generic- rather than dedicated buffers. From the process mapping, it was discovered that variability pooling was already used in the existing manufacturing system. For example, the assortment buffer is generic since the output from all final assembly modules are sent to the same buffer space.

The results from the control chart analysis in the thesis revealed that lead time was varying depending on the weekday. On regular weekdays (Monday-Friday) the lead time was substantially shorter and exposed to less relative variation compared to items partly or completely produced over weekends. This could be seen from the range views in the control charts. It was also discovered that this was caused by production being run slower, or sometimes not at all, over weekends.

Compared to analysing data in the form of condensed statistical measures (e.g. mean and standard deviation) JMP and control charts provide a more versatile picture It is important to visualize the data, instead of focusing solely on selected KPI values. The average lead time might not be representative for the factory or process studied.

The discovered relation between lead time and weekday might not be directly applicable for lead time optimization. Production will always be run slower over weekends, due to fewer operators being available. Still, discovering the connection was a prerequisite for the next iterative step of analysis. Since weekday had such a substantial impact on the lead time, other relations, trends and parameters of interest at first remained hidden underneath the variation caused by weekday. By identifying the weekday parameter, it became possible to separate the data and iterate the analysis in an attempt to discover additional root-causes for exceptional variation. In some processes, there could be dozens or hundreds of independent and dependent parameters causing exceptional variation, making this iterative approach necessary (Wheeler, 2003).

The revolutionary insight from the results residues in the proof-of-concept it provides for using statistics and control charts systematically to increase understanding of lead time data. Instead of stipulating that “the lead times are longer over weekends” based on gut feeling or personal experience, the relation can be confirmed with data. This is a crucial aspect, since the results become objective. No one can argue with proven facts. Thus, this way of analysing manufacturing data could potentially save time, since tedious discussions based on peoples’ opinions can be avoided.

The analysis in the thesis could have been extended further presuming access to additional data. When developing the data modifications, the intention was to make new rounds of analyses to examine each control chart and identify root-causes for exceptional variation.

For data modification one, it would have been interesting to examine root-causes further. For the ideal state in data modification two, it would also have been highly interesting to investigate the conditions required for high performance (with respect to lead time). Instead of focusing on reasons for long lead times, the idea with data modification two was to reverse the problem and instead investigate preconditions for short and stable lead times. However, to identify prerequisites for short lead time, additional data would have been needed.

For instance, the control chart for unmodified cylinder head machining data showed 5 points of exceptional variation in a row. One hypothesis could be that this was caused by production disturbances. However, without historical data about actual disturbances it becomes impossible to find evidence in the data to confirm the hypothesis. Thus, further analysis requires collection of additional data, such as executed shift and tact patterns, executed maintenance orders (preventive and emergency) and the current mix of variants in production. Experienced staff in the organization could also be utilized to suggest additional parameters of interest.

Lead time is explicitly linked to factory loading (Ignizio, 2009). Collecting historical data about the utilized capacity for the period of investigation could be of interest to aid the assessment of lead time and loading dependencies.

There are several functions in the organization, which rely on high-quality data for conducting different kind of analyses (Jonsson and Mattsson, 2009, Robertsson and Perera, 2001). Appropriate methods for systematic collection and processing of data could thus be of great value. When it comes to the lead time data, new insight on collection and processing could improve simulation analysis and optimization.

5.3 Future recommendations

As implied in the discussion, further analytical investigations could be done presuming access to additional data revealing the causes for outliers. The link between lead time and factory loading could also be of interest to study, as well as the correlation between lead time and maintenance. In a complex manufacturing system, there are many potential parameters of interest.

The analytical results from the thesis constitute a firm ground for further data analysis in the organization. The use of control charts could be seen as a demonstrator for how to systematically work with lead time data or other KPIs in the future. With the thesis work as a reference, similar extended studies could be conducted by the organization. In order to meet the lead time target, control charts could be constructed on regular basis to monitor lead times more continuously as a base for improvement discussions. Assuming root-causes for exceptional variation are found and eliminated to achieve a statistically stable factory, the analytical approach could be used for operational and strategic decisions on lead time related issues.

The organization could also complement the control charts through the use of design of experiments (DoE). Presuming access to relevant data, DoE could support the process of evaluating correlations between lead time and other parameters, as well as synergy effects among those parameters (Pyzdek, 2014). In this way, DoE could help the organization to identify conditions that cause varying or long lead times.

The development of IoT is creating smarter future factories with greater abilities to monitor and keep track of items using connected smart sensors (Farooq, et. al., 2015). The downside is the accumulation of big data, which is often too vast for organisations to process. By identifying influential parameters and creating more stable and predictable processes big data analytics becomes easier to master and more accurate predictions about future behaviour can be made.

The organization is currently assessing the possibilities for delivering engines in a sequenced one-piece flow, which means supplying engines in accordance with the production sequence in the car plants. For such a scenario, each engine in production would be dedicated to an existing customer order, which requires flexibility and short lead times. In order to enable such production, lead time analysis plays an important role. Accompanied with the use of big data analytics, IoT and automatic data collection and processing, production planning and execution could be supported to facilitate the mission towards “perfect customer sequence”.

5.4 Sustainability

With regards to the three pillars of sustainability: environment; economy and society, lead time reduction has the biggest impact on economy. By lowering lead time, tied up capital can be reduced resulting in saved cost for the organization. By altering the way procurement of raw material is conducted with regards to ordered quantity and frequency of deliveries the environment can be effected. If smaller more frequent deliveries are needed in order to lower the lead time in the raw material storage, then the environment would be negatively affected due to increased emissions of delivery trucks. If the situation is reversed however then the environment will benefit. Influence of reduced lead time on society is negligible.

5.5 Research generalisation

The lead time study could be extended in several ways to provide additional insight. Several delimitations to the lead time study were necessary due to the limited time frame of the thesis. The lead time study could be generalised by omitting some of these delimitations.

The lead time analysis was focused along one specific critical path through the factory, which was selected mainly for practical reasons. As there are several parallel flows in the manufacturing system, it would be desirable to understand where the longest accumulated lead time is. Since lead time had not been studied in this way before, the longest path through the factory was unknown beforehand. This meant that there was no guarantee that the selected path would be the truly critical one, since others could possibly limit the overall factory lead time.

In the future, the organisation could extend the lead time study by analysing additional parallel flows. Especially, lead time needs to be studied for the three other components (cylinder block, camshaft and crankshaft) to understand which of the four components that currently limits the lead time. The research could also be generalised by including both petrol and diesel variants (only petrol was considered in this case).

The lead times in raw material storages are likely affected by external factors, such as strategic purchasing decisions (e.g. safety stock levels), order quantities and delivery frequency from the suppliers. Therefore, the research could be complemented by investigations on supply-chain lead times. It is important to understand in what way logistic- and purchasing decisions affect the raw material storage and thus the overall factory lead time.

The thesis was conducted in one of the organisations engine factories, but it would also be highly interesting to investigate other engine factories within the organisation. Without detailed understanding of the specific manufacturing system configurations in the other factories, it is impossible to say whether lead time can be studied in the same way. Assessing the possibilities to make similar lead time analyses could possibly be done through process mappings or by consulting internal experts in the organisation.

Beside studies within the organisation, it could also be of high interest to generalise the research to other companies, both within and outside the automotive industry. There are some possibilities and limitations associated with such studies, but it is hard to assess the applicability without more in-depth investigations. Assuming lead time for different factories and companies could be measured, it would provide a sound base for benchmarking. As of now, it is hard to assess whether the factory lead time is competitive or not.

As previously stated management have set a target to decrease the overall factory lead time with 70 %. However, lead times had never been analysed in this depth and way before. The findings indicated that the total lead time is substantially longer than the target, and it might be relevant to evaluate whether the target is realistically achievable in the near future perspective.

In the thesis, it could be observed that even within the organisation the factory lead time was inconsistently defined, which lead to confusion. The same issue would probably arise when trying to generalise lead time studies to other factories, companies and industries. For instance, the degree of vertical integration could be different. Some companies might produce components in-house (like the engine factory studied in the thesis), while others would get all components from suppliers. In such cases, it would be hard to agree on generic definitions of factory lead time, which is required to enable lead time benchmarking.

5.6 Quality of research

The quality of research studies can be assessed from three aspects; reliability, replicability and validity (Bryman and Bell, 2011). First, reliability is the ability to repeat the results of the study, while replication concerns the ability to replicate the research method. Even though reliability and replicability are closely associated, they should not be confused to mean the same thing. Finally, validity assesses whether the measures represent the concept they are supposed to denote.

The authors believe that the lead time measurements are reliable, since statistical errors and fluctuations are accounted for through the use of statistical process control charts. When making additional extended studies in the raw material storage, the data tables could be cross-compared and the reliability of the thesis findings confirmed or dismissed.

The research is also considered repeatable, since efforts have been made to increase research transparency. In the thesis, lead time data was mainly processed in Excel, while JMP statistical software was used for the analysis part. Step-lists were created to assure that the processing and analysis was done in the exact same way for all stages along the critical path. These step-lists were also attached to the report as **Appendix E** and **Appendix F**.

One potential issue is that the replicability likely diminishes over time, since the manufacturing system is continuously changed. The issues and possible ways of collecting lead time data was closely associated with the structure of the IT systems. If the structure of the IT system changes, there are no guarantees that lead time can be collected in the same way in the future.

The authors believe that lead time measurements are representative for what they are intended to measure, namely the complete time from intake of raw material to outbound shipment of final engine products.

6 Conclusions

This chapter comprises the output to the three research questions introduced in the first chapter of the thesis.

Research question 1:

1. What is the current total lead time in the manufacturing system?
    a. How is the lead time distributed within the system?
    b. How does the lead time vary within the system?

The research has shown that it was possible to derive the lead time for each storage, buffer and process in the manufacturing process resulting in the total average lead time of 255.4 hours. By splitting the analysis up into several stages (covering each storage, buffer and process) it was possible to analyse the distribution of lead time within the manufacturing system. It could be seen that 91% of the lead time was spent in storages and buffers.

The research has also shown the variation of lead times within each stage (for new system data), especially depending on the weekday. Incidences were found where skewed lead times were presented due to work having been halted or run at a reduced tact on weekends. To countermeasure the issue two separate modifications were made to the data. The first excluded lead times over 12 hours in order to omit items where work was halted over weekends. The second excluded lead times over 4 hours in order to get as close to the ideal state as possible. As an example, the lead time in cylinder head machining was reduced by 71 % between unmodified data and the latter data modification.

Potentially, there are other parameters (such as the utilized factory capacity) contributing to variation, but unfortunately these relations could not be explored due to lack of additional data. The findings of the research open up the possibility to strategically analyse the entire factory from a lead time perspective. Thereby aiding in the decision making process when making changes and alterations to the processes to the means of lowering variation and creating more flexible and stable processes.

Research question 2:

2. How can lead time data be continuously extracted, processed and distributed efficiently within the organization?

Through analysing the MES the research has identified where sufficient data is available and where more representative data is needed. Lead times are extracted by different means depending on the MES at hand, which can vary between new, old and none-existing ones. Data from new systems is best representative allowing for in depth statistical analysis and visualization through control charts. For old system data, due to issues with data formatting, samples of 50 items were taken giving an indicator of lead times for those stages. For raw material storage where no MES data was available, 83 samples were taken through observations over the period of three days.

Further research is needed in raw material storage to adequately measure lead times for that stage, e.g. with a study extended over a longer time period than is done in the thesis. In order to continuously extract lead time data two means are proposed. The first involves adding functions to existing systems (through programming or additional software) specified to lead time extraction to a data warehouse optimized for data mining and analysis. The latter involves utilizing statistical analysis software, e.g. JMP that connects directly to the organizations MES system. In order for these solution to by viable the formatting issues with old system data has to be solved first, however the competence of which is believed to exist within the organizations IT department. Once data has successfully been extracted the use of statistical software like JMP can be used to automatically analyse, visualize and send out lead time information via email to relevant employees and departments. The organisation is currently investigating requirements for a potential new IT system. The suggested methods could also be used as input for when specifying requirements for such a system.

Research question 3:

3. What areas hold the biggest potential for lead time improvement and at what trade-off?

According to collected lead time data the biggest improvement potential is believed to exist in raw material storage. There, items are required to acclimatize for 24-48 hours before machining. However, they are on average stored for 184 hours which constitutes to 72% of the total lead time through the factory. Items are delivered to the factory several times a day and through better organization in procurement it might be possible to lower the lead time. The trade-off for lowering stocked material is however the danger of running out of material for the following process in the case of deliveries of raw material being late or cancelled.

Further research is needed with additional data (safety stock levels, maintenance and breakdown data, and etc.) for all the stages of the critical path before conclusions can be drawn and drastic changes made to the means of lowering lead time. The connection and interaction between stages of the critical path have to be taken into account as changes in one stage will effect both preceding and following stages. Simulation can e.g. be utilized to run different scenarios where the amount of items in storages and buffers are changes in order to acquire the optimal level for each stage.

7 References

ABB AB, (2016). *Internet of Things, Services and People* [Webpage] Available from: http://new.abb.com/about/technology/iotsp [Accessed: 2016-04-28]

Bellgran, M & Säfsten, K (2010). *Production development: Design and Operation of Production Systems.* Springer.

Berger, M., (2006). *Agile Manufacturing Systems – a Significant Quantum Leap for High Volume Production Approaches Flexibility Requirements as well as Fastest Reaction to Market Demands.* Published in: Reconfigurable Manufacturing Systems and Transformable Factories, pp 371-381. Springer Berlin Heidelberg New York. Available from: http://link.springer.com/book/10.1007%2F3-540-29397-3 [Accessed: 2016-04-28]

Berman, J. J., (2013). *Principles of big data: preparing, sharing, and analyzing complex information* Publisher: Elsevier, Morgan Kaufmann. Available from: http://www.sciencedirect.com/science/book/9780124045767 [Accessed: 2016-02-18]

Bryman, A., Bell, E., (2011). *Business Research Methods.* 3rd red. New York: Oxford University Press Inc.

Chen, F., Deng, P., Wan, J., Zhang, D., Vasilakos, A. V., Rong, X., (2015). *Data Mining for the Internet of Things: Literature Review and Challenges.* International Journal of Distributed Sensor Networks, Vol: 2015. Hindawi Publishing Corporation. Available from: http://www.hindawi.com/journals/ijdsn/2015/431047/ [Accessed: 2016-02-19]

Coletti, P., Aicher, T., (2011). *Mass Customization: An Exploration of European Characteristics.* Springer. Available from: http://link.springer.com/book/10.1007%2F978-3-642-18390-4 [Accessed: 2016-03-01]

Cwikla, G., (2014). *The methodology of development of the Manufacturing Information Acqusition System (MIAS) for production management.* Trans Tech Publications, Switzerland.

Duignan, J., (2016) *A Dictionary of Business Research.* Oxford University Press. Available from: http://www.oxfordreference.com.proxy.lib.chalmers.se/view/10.1093/acref/9780191792236.001.0001/acref-9780191792236-e-159 [Accessed: 2016-04-30]

Farooq, M. U., Waseem, M., Mazhar, S., Khairi, A., Kamal, T., (2015) *A Review on Internet of Things (IoT)* International Journal of Computer Applications, Vol: 113, No: 1. Available from: http://www.ijcaonline.org/archives/volume113/number1/19787-1571 [Accessed: 2016-02-18]

Hammersberg, P., (2016). Senior lecturer, Materials and Manufacturing Technology department and Technology Management of Economics department, Chalmers University of Technology *Discussion about industrial statistics* [Meeting] (Personal contact: 2016-03-18, 2016-03-29, 2016—3-31 and 2016-04-18)

Hopp, W., Spearman, M., L., (2008). *Factory Physics.* 3d ed. McGraw-Hill Boston.

Ignizio, J., P., (2009). *Optimizing Factory Performance: Cost-effective ways to achieve significant and sustainable improvement.* McGraw-Hill Companies, Inc.

JMPa Statistical Discovery SAS, (2016). *About us.* Available from: http://www.jmp.com/en_us/about.html [Accessed: 2016-04-28]

JMPb Statistical Discovery, SAS, (2016). *Using JMP* [Guide-book accessible through the JMP software help menu]

JMPc Statistical Discovery, SAS, (2016). *Basic analysis* [Guide-book accessible through the JMP software help menu]

JMPd Statistical Discovery, SAS, (2016). *Essential Graphing* [Guide-book accessible through the JMP software help menu]

JMPe Statistical Discovery, SAS, (2015). *Quality and Process Methods,* [Guide-book accessible through the JMP software help menu]. Version 12, SAS Institute Inc., Cary, NC, USA,

Jonsson, P., Mattsson, S-A., (2009) *Manufacturing, planning and control.* McGraw-Hill Higher Education.

King, P. L., King, J. S., (2015). *Value Stream Mapping for the Process Industries: Creating a Roadmap for Lean Transformation.* Productivity Press. Available from: http://library.books24x7.com.proxy.lib.chalmers.se/assetviewer.aspx?bookid=74089&chunkid=194696220&noteMenuToggle=0&leftMenuState=1 [Accessed 2016-02-16]

Konrad, R., (2014). *Fundamentals of Automotive and Engine Technology: Standard drives, Hybrid Drives, Brakes, and Safety System.* Bosch Professional Automotive Information, Springer Vieweg. Available from: http://link.springer.com.proxy.lib.chalmers.se/book/10.1007%2F978-3-658-03972-1 [Accessed: 2016-03-03]

Koren, Y., (2006). *General RMS Characteristics. Comparison with Dedicated and Flexible Systems.* Published in: Reconfigurable Manufacturing Systems and Transformable Factories, pp 371-381. Springer Berlin Heidelberg New York. Available from: http://link.springer.com/chapter/10.1007/3-540-29397-3_3 [Accessed: 2016-04-28]

Koren, Y., Shpitalni, M., (2010) *Design of reconfigurable manufacturing systems.* Journal of Manufacturing Systems, Vol: 29, Nr: 4. Elsevier Ltd. Available from: http://www.sciencedirect.com/science/article/pii/S0278612511000021 [Accessed: 2016-02-25]

Krumeich, J., Werth, D., Loos, P., Schimmelpfenning, J., Jacobi, S., (2014) *Advanced Planning and Control of Manufacturing Processes in Steel Industry through Big Data Analytics: Case Study and Architecture Proposal,* IEEE International Conference on Big Data. Available from: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=7004408&tag=1 [Accessed: 2016-02-17]

Larose, D. T., Larose, C. D., (2015) *Data Mining and Predictive Analytics.* 2d edition, John Wiley & Sons Inc., Hoboken, New Jersey.

Liker, J. K. and Meier, D. (2006). *The Toyota Way Fieldbook: A Practical Guide for Implementing Toyota's 4Ps.* McGraw-Hill, New York.

McKinsey & Company (2016-01). *Automotive revolution: perspective towards 2030 - How the convergence of disruptive technology-driven trends could transform the auto industry.* Advanced Industries January 2016. Available from: https://www.mckinsey.de/sites/mck_files/files/automotive_revolution_perspective_towards_2030.pdf [Accessed: 2016-04-30]

Minitab, (2016). *Minitab 17* [Webpage] Available from: https://www.minitab.com/en-us/products/minitab/ [Accessed: 2016-04-28]

Myatt, G., J., (2014). *Making sense of data I: a practical guide to exploratory data analysis and data mining.* John Wiley & Sons, Inc. Available from: http://site.ebrary.com.proxy.lib.chalmers.se/lib/chalmers/reader.action?docID=10891177&ppg=15 [Accessed: 2016-03-30]

Nywlt, J., Grigutsch, M., (2015) *Big Data Analytics Based on Logistical Models.* The Business and Economics Research Journal, Vol: 8, Issue 1, pp. 57-62. Available from: http://search.proquest.com/docview/1759511269?OpenUrlRefId=info:xri/sid:summon&accountid=10041 [Accessed: 2016-02-17]

Petersson, P., Johansson, O., Broman, M., Blucher, D., Alsterman, H., (2010). *Lean: turn deviations into success with examples from manufacturing, administration and health care.* 2nd edition. Part Media, Bromma.

Pyzdek, T., 2014. *The Six Sigma Handbook.* 4 Edition. McGraw-Hill Education. Online at: http://accessengineeringlibrary.com/browse/six-sigma-handbook-fourth-edition [Accessed 2016-03-29]

R Foundation, (2016). *What is R?* [Webpage] Available from: https://www.r-project.org/about.html [Accessed: 2016-04-28]

Randell L., G., Blomsjö, G., S., (2001) *Database driven factory simulation: a proof-of-concept demonstrator.* Proceedings of the 2001 Winter Simulation Conference

Robertson, N. H., Perera, T., (2001) *Feasibility for automatic data collection.* Proceedings of the 2001 Winter Simulation Conference.

Rother, M., Shook, J., (1999) *Learning to see: value stream mapping to add value and eliminate muda. Version 1.2.* The Lean Enterprise Institute, Brookline, Ma., USA.

Ryan, T., P., (2007). *Modern Engineering Statistics.* John Wiley & Sons, Inc., Hoboken, New Jersey.

Scheuermann, C., Verclas, S., Bruegge, B., (2015). *Agile Factory – An Example of an Industry 4.0 Manufacturing Process.* 2015 IEEE 3rd International Conference on Cyber-Physical Systems, Networks, and Applications. Available from: http://ieeexplore.ieee.org.proxy.lib.chalmers.se/xpls/abs_all.jsp?arnumber=7272683&tag=1 [Accessed: 2016-04-27]

Skoogh, A., (2011). *Automation of Input Data Management: Increasing Efficiency in Simulation of Production Flows.* Department of Product and Production Development, Chalmers University of Technology, Gothenburg, Sweden.

Summers, D.C., (2010). *Lean Six Sigma.* 1 Edition. Pearson Education, Inc., New Jersey, USA.

Tolio, T. (2009) Design of Flexible Production Systems: Methodologies and Tools. Springer-Verlag Berlin Heidelberg

Waldchen, D., (2014). *Towards a new order in the global automotive industry: How Asian companies catch up to their western peers* Anchor Academic Publishing. Available from: http://site.ebrary.com/lib/chalmers/reader.action?docID=10856556&ppg=10 [Accessed: 2016-03-01]

Wang, L., Alexander, C. A., (2015) *Big Data in Design and Manufacturing Engineering*. American Journal of Engineering and Applied Science, Vol: 8, Nr: 2. Available from: http://thescipub.com/abstract/10.3844/ajeassp.2015.223.232 [Accessed: 2016-02-12]

Westkämper, E., (2006) *Factory Transformability: Adapting the Structures of Manufacturing.* Published in: Reconfigurable Manufacturing Systems and Transformable Factories, pp 371-381. Springer Berlin Heidelberg New York. Available from: http://link.springer.com/chapter/10.1007%2F3-540-29397-3_19 [Accessed: 2016-04-28]

Wheeler, D. J., 2003, *Making Sense of Data – SPC for the service sector.* SPC Press, Knoxville, Tennessee.

8 Appendices

Appendix A: Table of Figures

Appendix B: Lead time data extraction from various IT systems (Interview Summary)

The following is a summary of an interview with a person who works with the Manufacturing Execution System at the IT department at the organization. The main purpose of the interview was to accumulate knowledge regarding extraction of lead time data from the various production IT systems used at the organization.

Both the IT system and the IT department are organized into three hierarchical sub-groups; ERP, MES and shop floor control. Extracting lead time data based on production time stamps stored in the IT systems at MES level, has never been done before. Generally, data extraction is done by accessing all CAMP systems separately (e.g. cylinder head, base assembly and final assembly module 3, respectively), which is heavily time consuming. Data is extracted through "data inquiries" to the system, which are commanding system orders to extract user specified data. For a specific time period, each data inquiry takes approximately two hours to execute in each system. Since the production lines run on separate CAMP systems, it means one data inquiry must be made for each line.

The data servers used in production are shared resources, with the densest data traffic during daytime when production is run on full speed. Therefore, data inquiries must be scheduled on weekends or nights to avoid overloading of the data servers. Otherwise, system overloads could possibly lead to shutdowns, which would also stop production scheduling and execution.

With regards to the raw material storage, there is no traceability in the IT systems. Lead time data must be manually measured, since time stamps cannot be extracted. The intermediate buffer, the assortment buffer, the engine dispatch and shipment area all operate on old IT systems (not CAMP like many production lines). The possibilities of extracting lead time data from those older systems are unknown, but are to be investigated by the IT department. For the intermediate buffer, lead time may be derived by cross-comparing data from base- and final assembly.

With regards to the new managerial direction on production development, with specified targets on lead time reduction, future ideas for continuous monitoring of lead time are of interest. Instead of running manually initiated data inquiries, it is possible to build and implement new software functionalities for automatic extraction of lead time data in the CAMP systems. Adding such functionalities to the CAMP systems could possibly be done in 2 weeks by external consultants. The software could automatically run lead time data inquiries each night, structure the data as an excel file and send it out to selected email addresses. The possibilities for customized data inquiries would be limited and the software "hardcoded" (not user-friendly). It is likely that users would soon ask for additional add-on functions, e.g. possibilities to customize the parts of the manufacturing system to be included.

The CAMP systems are constructed for effective and rapid communication (via virtual device) with the equipment PLCs. Therefore, CAMP is poorly constructed with regards to data reporting, visualization and analysis purposes. It is possible to incorporate such functionalities, but it has to be built by consultants.

CAMP are not only accessible for the IT department. There are tools available, so that external users can access data in CAMPs databases, but there is no interface available for lead time data questions. At the moment, it is only possible to access information on certain production activities, or from certain operations in production.

A tricky part is also that the database is so big that even if external users could easier access e.g. lead time data, the system would be severely slowed down by several users making data inquiries. It has already happened that production was stopped due to people looking at and inquiring data, which slowed down the system too much, causing IT system overload that shut down the entire system.

Appendix C: Raw material storage 1 (Interview Summary)

The following is *a summary of an interview with a person who works with inbound goods. The main purpose of the interview was to accumulate knowledge regarding the handling and registration of inbound goods.*

The interviewee manages a team of approximately 10 people responsible for unloading of inbound goods and outbound shipment of some specialized goods (not outbound deliveries of finalized engines). Roughly 60 truckloads of inbound goods are received each day and night. The goods reception process starts at the factory gate, where the truck driver submits goods documentation to the gate staff, which are part of the interviewee’s team. Information is manually entered into the logistics IT system based on the documentation received. The gate staff calls for a forklift driver to assist with unloading of the truck. The truck is released through the gate, material is unloaded and placed in storage. For the 4 Cs, material is only signed for at the gate. The forklift driver counts the number of unloaded pallets to confirm the order quantity. No barcode scanners are used.

A note called “flagga” is attached on each pallet and includes information about the material on the pallet. From the note, the following information is included: delivery note number, article number, supplier code and date. Total number of pallets of the same variant on the truck should be included, but is sometimes missing. The pallet note is removed by production staff as the material is fed to the production lines. In order to trace pallets back to the truck and get the time through the factory gate, the barcode on the top-left side of the pallet note is needed.

The raw material storage area is organized according to variants and date, and the principle of first-in-first-out (FIFO) is applied. FIFO means that the oldest material should be first fed to the lines. The machining operators need to identify the pile that contains the desired article number, then look at the date to pick the oldest material. The acclimatizing rule (24-48 hours) was confirmed and followed relatively well for cylinder heads due to high storage turnover.

The unloading of inbound goods team is only responsible for unloading trucks and moving the material into the raw material storage. Operators from production are responsible to feed material from raw material storage into the production lines. The trucks only carry one component (one of the four Cs), since each component comes from different suppliers. One truck may still contain several different variants of the same component.

Appendix D: Raw material storage 2 (Interview summary)

The following is a summary of an interview with a person who works with production planning. The main purpose of the interview was to accumulate knowledge regarding the organization and handling of the raw material storage.

A specific target temperature, which raw material has to have acquired before being released into machining, has not been established yet. However, work is currently being initiated where such targets will be measured and set for all machining processes. According to today’s instructions, during the winter, raw material needs to be stored for 24-48 hours in order to acclimatize due to cold transport. 48 hours for cylinder head and camshaft, and 24 hours for cylinder block and crankshafts. During the summer or if using heated transport (which are currently not used today) the 24-48 hour acclimatized period can be waived. However, the minimum acclimatizing time is 24 hours for all blanks. The importance of having items acclimatized to room temperature has to do with measurement requirements (extremely low tolerances) as metal expands when heated and detracts when cooled. The 24-48 hour limit is not derived from scientific calculations, merely a set limit derived from experience and knowledge about the process.

On occasion some aluminium items are released directly into machining, without having been acclimatized to room temperature. The reason for that is twofold. Firstly, aluminium is less affected by temperature with regard to expanding and detracting. Secondly, the process starts off with rough machining giving the item time to acclimatize, through the use of cutting fluid, before low tolerance machining commences at later stages in the process. Hence, under normal production conditions every item is acclimatized before being sent to machining. If there is a lack of material in storage the situation can arise where items are sent prematurely into machining without having been acclimatized for the full 24 hours. This has not resulted in a higher defect rate of finished items. The purpose of the raw material storage is not solely to serve as an acclimatizing holding cell, but equally important as a safety buffer. Hence heated transport or utilizing air-blowers of baths to warm the items up doesn’t necessarily result in items spending less time in storage.

Raw material for petrol cylinder-head/block and cam/crank-shaft arrives several times each day. The material planning system uses a parameter of 48-hour maximum storage time of each item, respectively, however the theoretical maximum storage capacity has not been calculated. If the situations are to arise where more than 48 hours of runtime material must be stored in the storage area the FIFO method is jeopardized due to the “right” material not being accessible. Otherwise material can be stored, for a short period of time, under a roof outside which is not optimal. The storage area is organized according to the arrival date and time of raw material (delivery note on each pallet). There is currently no IT-system that organizes what material is destined next for machining, according to FIFO. That responsibility falls on the production operator (forklift driver) feeding the machining process.

Appendix E: Step-list for data processing in Excel

<u>All data (unmodified)</u>

1. **“Freeze top row” with column names.**

2. **Hide all unnecessary columns.**
   - Needed columns are the following:
   - For “Cylinder head” data:
     - P:FIRSTSTATION (for diesel/petrol filtering purposes)
     - Q: STARTTIME
     - U: LAST_OP0250
     - Z: PARTNUMBER
     - AA: SERIALNUMBER
   - For “Cylinder head final storage” data:
     - P:FIRSTSTATION (for diesel/petrol filtering purposes)
     - S: FIRST_FL
     - V: LAST_FL
     - Z: PARTNUMBER
     - AA: SERIALNUMBER
   - For “Base assembly” data:
     - Q: STARTTIME
     - U: ENDTIME490A1
     - X: PARTNUMBER
     - Y: SERIALNUMBER
   - For “Final assembly” data
     - B: PARTNUMBER
     - C: SERIALNUMBER
     - C: FIRST_03S220
     - F: LAST_03S920

3. **Rename columns for easy navigation between documents.**
   Rename:
   - Columns: STARTTIME / FIRST_FL / FIRST_03S220
     → START DATE/TIME
     - Format cells: Change to → custom: yyyy-mm-dd h:mm:s
   - Columns: LAST_OP0250 / LAST_FL / ENDTIME490A1 / LAST_03S920
     → END DATE/TIME
     - Format cells: Change to → custom: yyyy-mm-dd h:mm:s
   - Column: FIRSTSTATION (Only for cylinder head)
     → Diesel/Petrol (OP2010=diesel / OP0010=petrol)

4. **Move “part number” & “serial number” to first two columns. (if necessary)**
    - Click on “serial number” column and move cursor to the left over “part number” column, now covering both columns.
        - Hold down shift button
        - Move cursor to edge of covered columns (cursor changes to a hand logo), click the edge and drag to the first column.

5. **Remove diesel data (if necessary).**
    - For “Cylinder head” and “Cylinder head final storage”:
        - Add filter to column (diesel/petrol)
        - Filter for only OP2010 (diesel)
            - Select all filtered OP2010 data
              Be careful to select all filtered data
              (Select top row, shift, command, and arrow down)
        - Delete rows
        - Remove filter
        - Hide column (diesel/petrol)

    - For “Base assembly”:
        - Add filter to column(PARTNUMBER)
        - Filter for 3139565, 3139578, 3139579, 3139581, 3139633
            - Select all filtered part number data
              Be careful to select all filtered data
              (Select top row, shift, command, and arrow down)
            - Delete rows
        - Remove filter

    - For “Final assembly”:
        - No filtering needed (line petrol dedicated)

6. **Remove NULL values (if necessary).**
    - Add filter to column(END DATE/TIME)
    - Filter for only NULL
        - Select all filtered NULL DATA
          Be careful to select all filtered NULLS
          (Select top row, shift, command, and arrow down)
        - Delete rows
    - Remove filter

7. Split up date and time from one column to two columns (for filtering purposes).

- Insert two columns after column(START DATE/TIME)
  - Name the first column: START DATE
    Format cells: Change to → Custom yyyy-mm-dd
  - Name the second column: START TIME
    Format cells: Change to →Time 13:30:55

  - In column (START DATE) insert formula =INT(Column(START DATE/TIME)).
    - Drag down for all rows (by double clicking)
  - In column(START TIME) insert formula = START DATE/TIME - START DATE.
    - Drag down for all rows (by double clicking)
  - Repeat same actions for column(END DATE/TIME)

8. Calculate lead time.

- Add three columns after column(END TIME)
  - Name first column: LEAD TIME
    Format cells: Change to → Time 37:30:55
  - Name second column: LEAD TIME (n) (n for number)
    Format cells: Change to → Number with three decimal places
  - Name Third column: LEAD TIME (d) (d for decimal)
    Format cells: Change to → Number with three decimal places
  - In column(LEAD TIME) insert formula = END DATE/TIME - START DATE/TIME.
    - Drag down for all rows (by double clicking)
  - In column(LEAD TIME (n)) insert formula = END DATE/TIME - START DATE/TIME
    - Drag down for all rows (by double clicking)
  - In column(LEAD TIME (d)) insert formula = Column(LEAD TIME (n)) * 24
    - Drag down for all rows (by double clicking)
- Hide column LEAD TIME (n)
  This results in lead time calculated in both hh:mm:ss and in hours with minutes and seconds in terms on decimal hours.

<u>Data modification 1 → All data < 12 hours</u>

1. **Removing data where work is halted (weekends)**
   - Add filter to column(LEAD TIME (d))
     - Click on filter and filter greater than 12
       Be careful to select all filtered data
       (Select top row, shift, command, and arrow down)
     - Delete rows
   - Remove filter

<u>Data modification 2 → All data < 4 hours</u>

2. **Removing data where work is halted**
   (weekends and nightshifts)
   (Getting as close to the ideal state as possible)
   - Add filter to column(LEAD TIME (d))
     - Click on filter and filter greater than 4
       Be careful to select all filtered data
       (Select top row, shift, command, and arrow down)
     - Delete rows
   - Remove filter

Appendix F Step-list for graph building in JMP

1. From the processed data in the Excel file, copy the following columns and insert from left to right in the empty data table:
- W: Start date (confirm date format yyyy-mm-dd)
- X: Start time
- AB: End date
- AC: End time
- Q: Part number
- AF: Lead time (decimal)

2. Double click the "column label", rename each column and change data type, modeling type and format when necessary. For each column, specify formats accordingly:
- Start date = Numeric, Nominal, yyyy-mm-dd
- Weekday start = default settings
- Start time [h:m:s] = Numeric, Nominal, h:m:s
- End date = Numeric, Nominal, yyyy-mm-dd
- Weekday end = default settings
- End time [h:m:s] =Numeric, Nominal, h:m:s
- Part number = Character, Nominal (default)
- Lead time [h.h] = Numeric, Continuous, Best

3. Add columns, which display the weekday of start and end date respectively (to enable later sub-grouping based on the day of the week to search for weekday patterns)
- Double click the column label of weekday
- For weekday start:
    - Right click the column label → Formula
    - Left box select "Start date"
    - Right box select "Date time" → "Day of week" → OK
    - Change "Value ordering" to make the days appear in logic order (by default Sunday =1, Monday =2 etc.).
        - Move Sunday to the end:
            - Stand in the data table → Right click the column label → Column info
            - Column properties → Value ordering → Select "1" → Move down (to bottom) → OK
    - Rename "1-7" to the actual day of week to make it more understandable straight from the data table.
        - Stand in the data table → Right click column label → Column info
        - Column properties → Value labels
        - In "value" enter "1", in label type "Sunday" (repeat for rest)
- For weekday end: Follow the same procedure as above, with one crucial difference:

- In the left box select "End date" (instead of Start date).

4. Build histogram chart to assess the distribution of the data set
- Of course possible to do in JMP, but it was harder to customize the plot to make it visual for the report. The histogram was therefore built in Excel.

5. Build "Number of engines completed on each day" for the time period.

Not needed for data modification 1 and 2

- Graph builder → Drag "end date" to X-axis
- Save as JMP report (only way to re-open the plot for later use or customization in JMP). Select embedded option.
- Add axis labels, customize to make the plots coherent for the report.
- If production is not run on weekends, "fill out" the plot with the days that are not represented in the graph (to visualize this in the plot)
  - Add the days at the bottom of the end date column (weekday auto generates). Leave remaining columns empty. Empty days are now added in the plot.
- Save as → Switch to JPEG → Select "300" to save the picture with higher resolution than default settings (clearer picture in the report).

6. Create a box plot to examine the distribution of data
- Graph builder → Lead time on Y-axis
- Customize axis labels, font etc.
- Save as JMP report and JPEG picture

7. Create "control chart – average and range on start date"
- If "empty days" were added (as specified in the last point in step 5), these days needs to be removed before building the control chart (otherwise there will be gaps in the control chart). Erase the rows and save the new data table.
- Analyse → Quality and Process → Control Chart Builder
  - Lead time on Y-axis (Does it reveal anything of interest?)
  - Start date (or test end date) on X-axis
    - Grouping on date clarifies the plot. However, the grouping may in some cases be misleading, so assess the consequences of grouping every time.
- Save as JMP report and JPEG 300 picture

8. Create “control chart – individual and moving range over time”

- Analyse → Quality and Process → Control chart builder
  - Lead time on Y-axis
  - Save as JMP report and JPEG 300 picture

9. Create “control chart – individual and moving range phased on weekday”

- Analyse → Quality and Process → Control chart builder
  - Lead time on Y-axis
  - Weekday on phase (drag to the top, not the X-axis).

General guidelines for saving in JMP

- Save as JMP reports (each plot is saved as an individual file)
  - Enables later re-opening of the files to continue the work, change or customize further.
  - Save as (file format .jrp)

- Save as high resolution JPEG file (each plot as a picture to be used for documentation)
  - Save as (switch to JPEG, select “300”)
  - Helps to save the pictures with higher resolution compared to default settings in JMP.
  - An alternative approach for creating pictures is to use the ESP file format, which is vector-based to enable picture scaling without pixels.
    - Tutorial available on: https://www.youtube.com/watch?v=jbfXDiYPuO