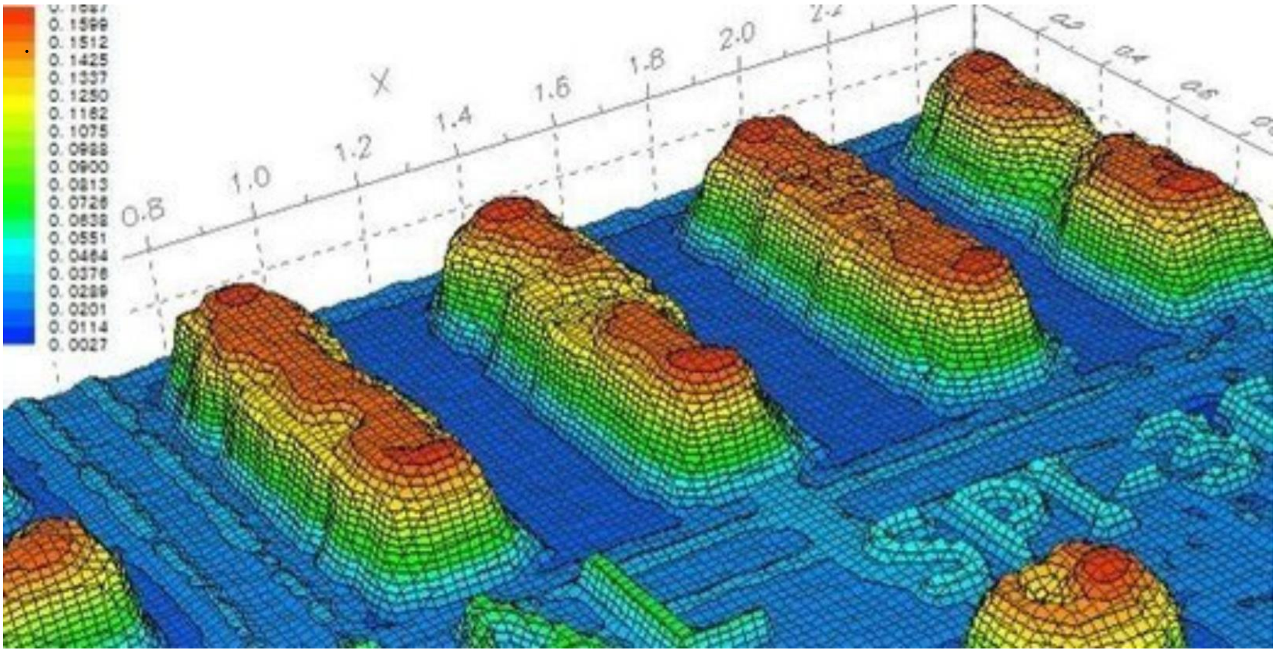
CHALMERS
UNIVERSITY OF TECHNOLOGY



Key Factors Influencing Scrap Levels at an Electronics Industry

A Six Sigma Project

Master's thesis in Quality and Operations Management

FREDRIK HANSSON
JOHANNES RAMBERG

DEPARTMENT OF TECHNOLOGY MANAGEMENT AND ECONOMICS
DIVISION OF SERVICE MANAGEMENT AND LOGISTICS

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2020
www.chalmers.se
Report No. E2020:032

REPORT NO. E2020:032

Key Factors Influencing Scrap Levels at an Electronics Industry

A Six Sigma Project

FREDRIK HANSSON
JOHANNES RAMBERG

Department of Technology Management and Economics
Division Service Management and Logistics
Chalmers University of Technology
Gothenburg, Sweden 2020

Key Factors Influencing Scrap Levels at an Electronics Industry
A Six Sigma Project
FREDRIK HANSSON
JOHANNES RAMBERG

Report no. E2020:032
Department of Technology Management and Economics
Division of Service Management and Logistics
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Image showing a 3D inspection of applied solder paste deposits (Surface Mount Process, 2020a).

Printed by Chalmers Reproservice
Gothenburg, Sweden 2020

Key Factors Influencing Scrap Levels at an Electronics Industry
A Six Sigma Project

FREDRIK HANSSON
JOHANNES RAMBERG

Department of Tachnology Management and Economics
Chalmers University of Technology

Abstract

The case company is a Swedish electronics manufacturer and desires to investigate potential root causes to process variations in their SMT-production. The case company further want to validate the data that they are collecting and also run the master thesis as a Six Sigma pilot project to investigate if Six Sigma should be implemented in a larger scale. The aim is therefore to investigate and understand why the case company experiences these process variations and to find potential root causes and at the same time run the project as a Six Sigma pilot study. With the project involving big data and the fact that Six Sigma lacks a method for big data analytics, the project provided new knowledge to bridge the gap between Six Sigma and big data analytics. Machine learning techniques are a useful way of bridging this gap and highly related to Industry 4.0.

The project was divided into three different phases according to Six Sigma methodology; the Define phase, the Measure phase and the Analyze phase with associated methods for each phase. The methods used were both qualitative and quantitative, with major focus on quantitative data. The qualitative data were used to support or match findings in the quantitative data. Quantitative data were retrieved from the SPI-machines in the SMT-production while qualitative data were gathered through observations, interviews and other activities. The most important metric to achieve high-quality products were solder paste volume and the results showed that factors related to the stencil design were very important and highly connected to variations. Another important factor was the human interaction and manual settings.

The findings were finally translated into managerial implications and recommendations for the case company to overcome the process variations.

Keywords: DMAIC, Six Sigma, Big Data, Solder Paste Inspection (SPI), Printed Circuit Board (PCB), Screen Printing Process, Surface-Mount Technology (SMT).

Acknowledgements

There has been many people involved throughout this master thesis. First, we would like to thank our supervisor at Chalmers University of Technology, Hendry Raharjo, for his extensive support, discussions and feedback throughout the master thesis, especially during the Analyze phase. Secondly, we would like to thank everyone from the case company that has supported our master thesis. We would especially like to thank the process engineer Patrik, who has been very engaging and supportive throughout the master thesis. The process engineers Andreas and Tomas, who have been supportive and engaged when we have had questions regarding the production process. The quality engineer Kristoffer, who has helped us with important graphs of the production history. We would also like to thank our supervisors at the case company, Kristoffer and Rakesh, for their continuous support and feedback. Lastly, we would also like to send our gratitude to all those that have been involved in this master thesis in one way or another.

Without your help, support and feedback, the successful outcome of the master thesis would not have been possible! For that we are truly grateful!

Fredrik Hansson & Johannes Ramberg, Gothenburg, May 2020

Contents

List of Glossaries

Board When producing PCBs, a board is the frame containing a certain amount of PCBs, depending on PCB size.

MSE Mean Squared Error, a statistics measure of the average squared difference between the actual and predicted value.

SMT Surface Mount Technology, a production method to mount electrical components on the surface of a PCB.

SPI Solder Paste Inspection, a machine that measures the solder paste deposit through different metrics.

PCB Printed Circuit Board, a plate with electrical conductors that facilitates electrical connections between electrical components.

Panel Another word used for PCB in the master thesis.

Solder paste A type of glue that is being put on the PCBs so that the electrical components can get stuck on the board.

Squeegee A type of blade that slides over the stencil with a certain angle and pressure to make the solder paste go through the apertures and onto the PCB. Also removes excess solder paste.

Stencil design A type of blueprint, made out of thin metal, used over the PCB in the manufacturing of the PCB to direct solder paste through the stencils apertures onto designated areas of the PCB.

List of Figures

List of Tables

1
Introduction

In this the chapter, an introduction to the master thesis is presented. First, a short background to the company and the project is presented followed by the problem statement, objective, aim and research questions of the master thesis. At the end of the chapter, delimitations will be presented. The company in this project will henceforth be referenced to as the case company

1.1 Background

Today's global, high-paced market with fierce competition and a global economy that are becoming more and more interconnected and integrated are pushing companies to become more and more adaptable and work with continuous improvements (Chen, 2008). In order to survive and thrive in this market setting, companies must produce high-quality products and services according to customer requirements, which might rapidly change due to trends and new needs, and constantly aim for the highest customer satisfaction (Kumar, Antony, Madu, Montgomery & Park, 2008). At the same time, the company must also strive for reducing costs, using fewer resources and reduce scrap levels as sustainability is a growing area that every company needs to work with (Ahuett-Garza & Kurfess, 2018; Chun & Bibanda, 2013).

The electronics manufacturing industry is characterised by innovation, fast development and fierce competition and to lie in the forefront of the industry is of the highest importance (Vault, 2020). Furthermore, research and development account for a large part of a company's spendings for the improvement of production systems due to the high competitiveness in the industry (Vault, 2020). The Swedish electronics manufacturing industry is characterised by high-cost production compared to many low-cost countries in Asia and Eastern Europe, but are valued for the high-quality products that are produced (Sundqvist, Hedman, Almström & Kinnander, 2012). Therefore, the utility degree of machinery as well as low scrap levels are an important competitive factor as well as customer flexibility and an overall effective production system (Sundqvist et al., 2012).

Related to this background, the case company is a global Swedish company, active in the electronics manufacturing industry. In Sweden, they have several different sites and among them are the production plant where this master thesis takes place. In the production plant, the production focuses on the production of printed circuit boards (PCB) and surface-mount technology (SMT), as well as the internal assembly lines of their final products.

Since the case company is active in an extremely competitive market, most of the money earned is going in to the development of new products to increase the number of market shares and customers. However, not enough attention has been paid to the increasing scrap levels and connected cost losses. The case company is therefore in a great need of fixing these problems to be able to stay profitable and competitive, finding the source to the variation and receiving recommendations on countermeasures. This creates a high sense of urgency that things must change and causes to the variation and high scrap levels must be found and countermeasures be implemented (Kotter, 2008). To face these challenges, the case company want to investigate if the they could use Six Sigma methodologies, since a Six Sigma strategy provide organisations with a methodology and toolbox to solve complex problems, increase internal quality, increase organisations profitability, market shares and customer satisfaction (Harry, 1998; Schroeder, Linderman, Liedtke & Choo, 2008). However, currently the case company uses a limited number of Six Sigma methods. Therefore, they initiated a master thesis project in a Six Sigma setting as a pilot to see if the methodology can be used further in the company.

1.2 Problem Statement

The current problem situation in the production concern high scrap levels in some of the production processes. Especially problematic is the case that most of the scrapping takes place late in the production process, after all value-adding steps has been performed to the product. After the implementation of several new quality check stations during the spring of 2019, scrap level increased dramatically and the company have few clues on what the cause is.

The case company's main focus has been to be able to deliver products to their customers in time rather than focusing on quality costs. However, the company have now realized that if they want to stay competitive and profitable, they need to make changes and improve their internal quality. This insight is highly interesting to the project team in different ways. Firstly, the sense of urgency increases motivation to find the underlying reasons to their quality issues. Secondly, the fact that the problem area has not been solved yet could indicate that the problem is complex. Hence, by approaching the problem through a Six Sigma project, the project team feel confident to be able to find and understand the yield variation and root causes behind the case company's high scrap levels at the SMT process.

1.3 Aim & Objective

The aim is to understand why the case company have yield variations at the SMT process, data collected by its quality control machines called solder paste inspection (SPI), and to find the root causes to these variations. In addition, they would view the master thesis as a pilot for running Six Sigma projects. Therefore, the objective is to help the case company understand the causes of yield variation and deliver validated and verified recommendations, through a Six Sigma project.

1.4 Research Questions

The following research questions have been formulated to achieve the objectives of the master thesis:

RQ1: How can Six Sigma be used to provide a systematic approach to understand the problem?

RQ2: What are the root causes to the yield variations and how can those variations be managed?

1.5 Research Approach

This master thesis takes place at the case company as previously stated. The master thesis, therefore, partially takes aim to approach and solve their stated problems. The master thesis will therefore be treated as a case study where the case company is the main focus area. But the master thesis also takes aim at contributing scientifically to investigate, compare and discuss potential gaps in the literature regarding the problem area to further expand current knowledge and research. These potential gaps will be discussed and related to the case study through the case company.

1.6 Delimitations

The master thesis has been delimited in several ways in order to scope the project down and to be able to provide the case company with practical and useful conclusions.

The case company has a few different productions and assembly lines. The project was, however, delimited to investigate the SMT process and its SPI. Further, the SPI measure various parameters after the solder paste has been applied to the PCB board. From this process, huge amounts of data are generated, data that also impact the further process steps. Therefore, the project concentrated its efforts on the processes related to the SPI-stations and the data produced from the SPI process.

Due to the limited amount of time for the master thesis, the task of implementing the recommended countermeasures lies upon the case company. Therefore, nor the Improve or the Control phase of the DMAIC model will be performed; Only the first three phases will be conducted, Define, Measure and Analyze phases.

2

Problem Background

In this chapter, a short description on what a PCB and solder paste is, the SMT process, and the most common solder paste printing defects will be presented. This is done to provide an overview of the SMT process, which will make the further chapters easier to follow and understand.

2.1 Printed Circuit Board

A PCB is the platform on which electrical and mechanical components are being placed upon and provides electrical connection between various components through tracks, pads etc. Typically, it is made of one or more layers of etched copper and non-conductive materials. Components are typically being soldered on top of the PCB on designated pads and can provide the PCB with many different functions and connections. (LaDou, 2006).

2.2 Solder Paste

The solder paste is the material or glue that mounts the components to the PCB board. The solder paste typically consists of around 90 % metal powder alloy and 10 % organic chemical material (Baluch & Minogue, 2007). The solder paste is critical to the assembly of the PCB and to the whole functionality since it provides mechanical, thermal and electrical functions to the PCB as a joining material for the different components on the PCB (Hwang, 1989). High quality of the solder paste is further essential since it provides function and performance to the overall PCB (Hwang, 1989).

2.3 SMT Process

To visualise the process and to get a better understanding of how an SMT process operates, how the solder paste printing process looks like and common solder paste defects, a brief summary is presented in Figure 2.1 and Figure 2.2.

The SMT process consists of three major parts; the solder paste printing process, or the screen printing process as it is also called, pick-and-place process and solder re-flow (Figure 2.1a). In the first step, the solder paste is applied on the PCB. In the second step, the PCB components are added onto the applied solder paste. In the third step, the PCB goes through an oven to melt the solder paste, the panel and the components together. Lastly, the PCB is cleaned and electronically tested before leaving the SMT process. The case company's SMT process uses the Squeegee Blade Method to apply solder paste in the screen printing process, as seen in Figure 2.1b.

Between the screen printing process and the pick-and-place process, the case company have installed a SPI quality check. The added SPI-station measure the output of the previous screen printing process. The SPI-station does this by sending out laser beams which the SPI-station uses to create a virtual 3D-model of the applied solder paste on the PCB (Surface Mount Process, 2020b). Thereafter, the SPI-station either scrap the PCB or let it proceed in the SMT process. IPC (2012) published the IPC-7527 standard to help organisations improve the quality of the screen printing process. Hence, if the applied solder paste is good or not good is decided accordingly to the IPC-7527 specification. The IPC-7527 standard mentions four crucial dimensions for a good solder paste deposit; its shape, offset, area and height.

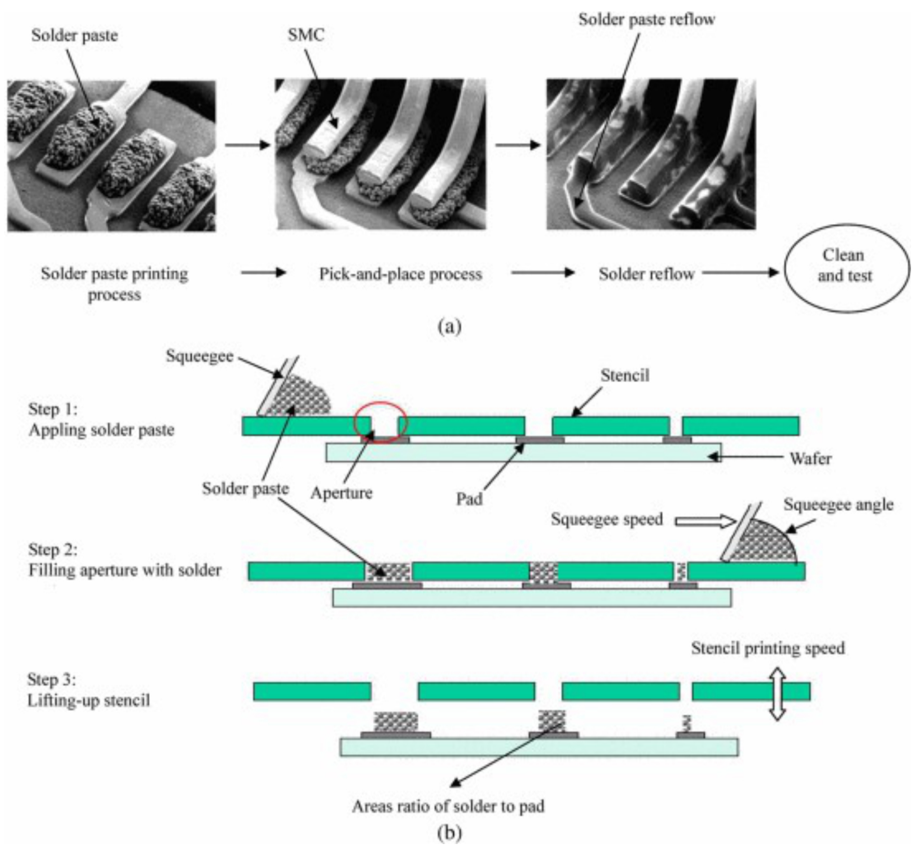


Figure 2.1: Process mapping for SMT line. (a) SMT line processes. (b) Screen printing process with the Squeegee Blade Method. (Li, Al-Refaie & Yang, 2008).

2.4 Solder Paste Printing Defects

Figure 2.2 below describes the most common defects related to the solder paste printing process. These are slumped print, scavenged print, bridged and peaking. A slumped print is mainly the result of a process operating at too high temperatures. The scavenged print occurs when the squeegee pressure is too high and the squeegee scoop the solder paste out rather than apply it. The bridging failure mode occurs due to poor board support and/or cleanliness of the stencil. A peaking defect may be the result of a too high separation speed between the stencil and the PCB. (Surface Mount Process, 2020a).

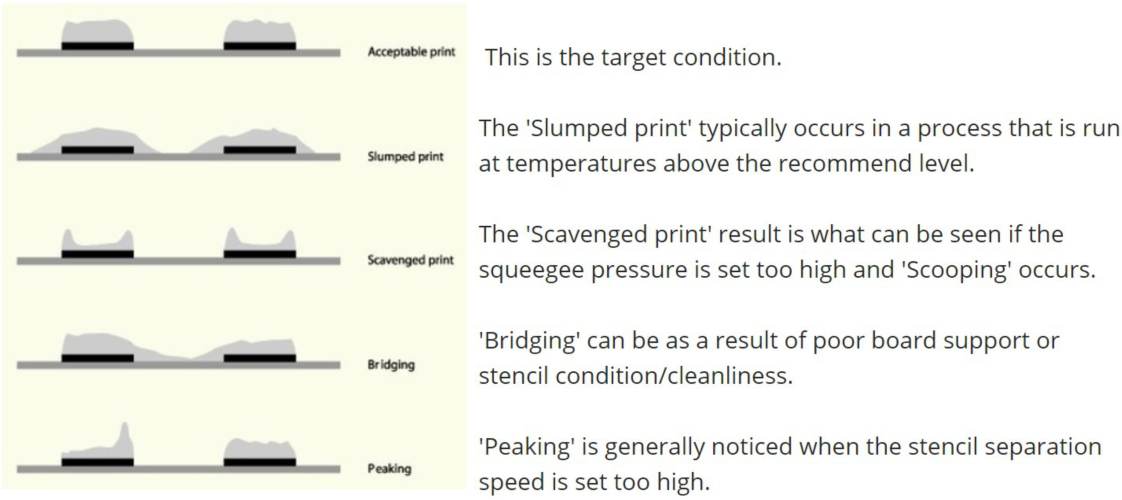


Figure 2.2: Common solder paste printing defects (Surface Mount Process, 2020a).

3
Methodology

In this chapter, the chosen methods for the master thesis are presented and discussed. The aim with the chosen methods is to fulfil the purpose and goals of the thesis and answer the research questions. The chosen research model is shown below in Figure 3.1.

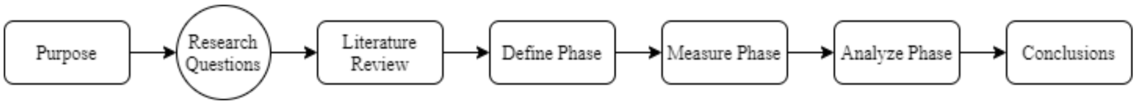


Figure 3.1: Research model.

3.1 Research Approach

The master thesis was based on the increased demand from the case company to increase their process quality and reduce variations. But in order to improve their process quality and reduce variations, they wanted to understand the root causes of the production quality issues. Therefore, the objective of the master thesis has been to investigate which parameters contributed to the process yield variations and which parameters that impacted the yield the most at the SMT production lines. The master thesis used both quantitative and qualitative data collection methods to achieve the project goals.

3.2 Data Collection

Primary data for the project were mainly quantitative data collected from the SMT-lines and more specifically from the SPI-machines. Further, this data were complemented with qualitative data through semi-structured interviews with people that possessed critical knowledge of the process and that could provide important input. A literature study was conducted as well to complement the project with previous conducted research that this master thesis will build upon and extend.

3.3 Literature Review

The aim of the literature review was to create an understanding and consensus about the research topic and previous research conducted within the area. It was conducted to create a theoretical background before the quantitative data gathering and following data analysis, supported by the qualitative data collected through interviews and input from key people. The literature review were further made to support the project's research and findings to be able to create an understanding on what research that has been conducted, the assessment of their findings and conclusions related to this project.

To find relevant literature, the scientific search engine Web of Science was used primarily and secondary Google Scholar were used as a complement. The search words that was used were primarily “SMT”, “DMAIC”, “Six Sigma”, “solder paste”, “squeegee”, “stencil”, “PCB” and “printed circuit board”.

3.4 Theoretical Framework

To deepen the understanding about the project area and the related production, methodology and statistical programmes used, a theoretical framework were conducted and is presented in chapter 4. The theoretical framework presents relevant literature about Six Sigma and the DMAIC method and its different steps, the Affinity-Interrelationship Method (AIM), the screen printing processes, recursive partitioning, bagging, boosting and random forest.

3.5 Quantitative Data

The quantitative data consisted of extracted data from all SMT-lines (SMT1 to SMT6) and specifically from the SPI-machines. The data consisted of information on parameters such as volume, area, height, offset, size, height, etc. and represented information on each position on every of the individual panels on the PCB board. The data were retrieved straight from the SPI-machines once a day and covered data of the last 50 produced PCB boards. Due to the amount and accessibility of the data, the period on which the data were collected was from the beginning of November 2019 until the end of January 2020. The data were then stored and accessible through a memory stick.

3.6 Qualitative Data

The qualitative data consisted of semi-structured interviews of key process engineers with extensive knowledge related to production and SMT-lines. The qualitative data were further extended with an AIM to further deepen the understanding of the underlying problems behind the high scrap levels and to compare the process engineers perceptions and understanding to the quantitative data. The semi-structured interviews were conducted to get important information around technical issues and get to know the operations and production better. Qualitative data were also collected continuously when needed though discussions either face-to-face or through Microsoft Teams. The gathering of qualitative data creates better trustworthiness for the quantitative data and was also relevant for the comparison of the two data types for the investigation of potential contradictions and similarities (Waller, Farquharson & Dempsey, 2016).

3.7 Data Analysis

Data analysis is an important component of the project to avoid potential uncertainties, misinterpretation and errors. The quality of the data is essential to achieve a trustworthy result and a result that you can rely upon (Wallen, 1996). Structuring and management of the data are also highly important to be able to assess the data in a correct way (Bryman, 2012). Therefore, the data were first translated from CSV-files into Excel, structured and then put into the statistical programmes JMP and Minitab for analysis. These programmes are useful both for data analysis and for visualizing the results. It was, however, soon realized that these programmes were not capable enough to handle the vast amounts of data that the project brought with it. Therefore, it was decided to use the programming language R instead due to its capability to handle, manage and analyse vast amount of data. Minitab were therefore only used to process less quantity data. To be able to use the data in R, the Excel-files were converted back to CSV-files again.

3.8 Discussion of Chosen Methods

The methods chosen for this project were based on the fact that vast amounts of data were to be analysed to find the root causes to the high scrap levels. Therefore, quantitative methods such as Measurement System Analysis (MSA), data cleaning, recursive partitioning, bagging, boosting and random forests were used to analyse the data. Due to the fact that this project included huge amounts of data, the choice of method for analysing data in the Analyze phase were the programming language R. R is especially good for handling large amounts of data and is also an excellent machine learning tool used to predict future data (James, Witten, Hastie & Tibshirani, 2013). Since neither Excel, JMP or Minitab could handle the full extent of this project's amount of data simultaneously, the choice of statistical program fell on R and Rstudio.

The use of both quantitative and qualitative methods were necessary to support the project and also help to explain errors, find connections and contradictions (Waller et al., 2016). The use of both methods enhanced the breadth of the project and allowed for a more holistic view of the problem situation (Schoonenboom & Johnson, 2017). Related to the qualitative methods, observations of the production and related production processes were first conducted before any interviews to not be influenced or coloured by information given by the interviewee. Performing interviews before observations might affect the focus and attention to important details that might not come up during the interview when performing observations afterwards (Waller et al., 2016). Therefore, interviews followed observations and later on, more observations were carried out but were limited to several round tours. However, extensive observations were not necessary to be able to carry out this type of project.

The most important information came from quantitative data and interviews.

Regarding chosen methods, there were several methods not chosen to be carried out, among them is-is not-analysis in the Define phase. Instead of is-is not-analysis, a critical-to-quality (CTQ) flowdown were conducted. Due to time constraints, the extent and aim of the project, the Improve and Control phases in DMAIC were not conducted.

3.9 Method Criticism

Criticism against the methods chosen for this project should also be discussed. When choosing both quantitative and qualitative methods, it is important to understand that the methods might divide the available resources and therefore it is important to assess the time and resources spent on each method (Bryman, 2006). It is also important to compare the different methods against each other continuously to find possible contradictions or misinterpretations (Bryman, 2006). Using both methods might also direct focus on things that one of the methods point out as important but result in missing out on important information from the other method. Regarding the qualitative methods, reliability and validity is important to assess as well and information retrieved from interviews might be subjective and biased (Golafshani, 2003).

Regarding data analysis, much time were spent on learning to use R, conduct data cleaning and to get the data together. This was very time consuming and potentially important parts might have been missed out due to this, both regarding lack of time and experience. As the progress of the project went on, it became clear that knowledge and experience in programming and data analysis were very important. To be able to carry out the project properly, more time might have been spent in the beginning to be more prepared when the data analysis phase started.

3.10 Thesis Trustworthiness

According to Waller et al. (2016), trustworthiness is based on the two components reliability and validity. The main issue for the thesis team, to ensure the trustworthiness of the project, was their lack of experience within the problem area. The thesis team took several different approaches to overcome this issue. A literature review was conducted to understand what studies that has been done before and their findings. Interviews with experienced process engineers were conducted to understand the case company's situation, but also regarding the SMT process in general. Also, discussions with the process engineers, the case company supervisors and the supervisors at Chalmers were conducted continuously throughout the master thesis and resulted in increased knowledge and experience within the problem field.

The case company shared their view of their quality issues in an objective and transparent way. However, there was no proof that the data given to the project team was valid. To overcome this issue, the thesis team conducted an MSA, together with one process engineer, to be able to trust the data before analysing the provided data.

By combining the trustworthiness of the data together with the experience gathered by the thesis team within the problem area, the thesis team could make the outcome of the master thesis reliable, valid and trustworthy.

4

Theoretical Framework

This chapter will present the theoretical framework that this project is based upon. First, a literature review is presented. Second, an overview of Six Sigma will be presented followed by more in depth explanation of the different phases related to DMAIC: Define, Measure and Analyze, excepting the Improve and Control phases. Methods that will be used under each phase will also be presented below

4.1 Literature Review

This section presents the literature review, covering previous research around the thesis topic, and will form a foundation for the relevance and need for this thesis.

4.1.1 Key Quality Characteristics

There has been a number of articles published concerning PCB manufacturing with SMT and solder paste printing processes, related to the topic of the thesis. Also, articles covering Six Sigma and the DMAIC approach related to PCB manufacturing has been found.

In the articles researched, all present different cases in different settings and have been investigating different things. Therefore, it is not surprising that all of them both have similar and different conclusions and recommendations on the important factors and optimal settings to a PCB manufacturing process.

4.1.1.1 Solder Paste Characteristics

The key quality characteristic "solder paste height" were mentioned in several articles as highly important for the outcome of the PCB manufacturing. Li et al. (2008) mentions that large variations in solder paste thickness are one of the main reasons of PCB failure related to manufacturing defects. Related to the manufacturing defects, around 60% of PCB defects can be derived from the solder printing process (Li et al., 2008; Khader, Yoon, & Li, 2017). The authors further mentions that much time are generally being spent on troubleshooting due to PCB defects and thus, achieving less variation in the solder printing process will not just result in monetary effects (Li et al., 2008). Tong, Tsung & Yen (2004) also stresses the importance of solder paste height as one of the key quality characteristics and so does Tsai (2011) and Mannan, Ekere, Ismail & Lo (1994). Controlling the solder paste height is critical in the printing process due to the fact that too little solder paste might cause open circuits between components and too much solder paste can result in bridging between the component pads, thus causing the PCB to fail (Tong et al., 2004).

Another article pointed out that the viscosity or quality of the solder paste used in the printing process is important for the outcome of the process (Asghar, Rehman, Aman, Iqbal, & Nawaz, 2019). The authors also mention that printing quality is strongly correlated to the quality of the solder paste. The process temperature also affects the viscosity of the solder paste and a temperature that is varying a lot, and the use of the same solder paste under different temperatures, will produce results with a lot of variation and result in difficulties controlling the printing process (Gopal, Rohani, Yusof & Bakar, 2006). The solder paste itself is a mix between a metal alloy powder and an organic chemical material, "flux" and the ideal proportion are around 90% metal alloy powder and 10% flux (Baluch & Minogue, 2007). Further, the quality of the solder paste is affected by other factors such as particle size, paste distribution on the PCB, rheology and temperature and viscosity changes during the printing process (Khader, Yoon & Li, 2017). Khader et al. (2017) proposes a proportion of 87,5% metal alloy powder and 12,5% flux as the ideal distribution but also mentions that this depends on the type of solder paste and the area of usage.

4.1.1.2 Squeegee Characteristics

Throughout more than half of the studied articles, Tong et al. (2004), Asghar et al. (2019), Mannan et al. (1994), Lau & Yeung (1997) and Khader et al. (2017) concluded that squeegee related characteristics had a significant impact on the screen printing process performance in one way or another. Mannan et al. (1994) conducted an experiment to investigate which squeegee parameters that affect the solder paste height the most and the result showed that the squeegee pressure onto the stencil was outlined as the most crucial. Further, Asghar et al. (2019), Lau & Yeung (1997) and Khader et al. (2017) also conducted experiments where they all concluded that squeegee pressure and speed was of high importance for the solder paste printing result. Asghar et al. (2019) concluded in their article that changing the applied squeegee pressure from 80N to 90N improved the solder printing process significantly.

Another factor that was shown to have a large impact on the solder paste printing result is the squeegee angle (Asghar et al., 2019; Marcoux, 1992). Applying different squeegee angles produce results with a lot of variation in the solder paste deposit (Asghar et al., 2019). Through further experiments, the angle of the squeegee and the material of the squeegee was also to have an impact on the screen printing process performance (Tong et al., 2004; Lau & Yeung, 1997). Hence, previous studies show that the four squeegee related parameters squeegee pressure, speed, angle and material affect the screen printing process performance.

4.1.1.3 Stencil Characteristics

According to many of the researched articles, stencil thickness and stencil design were among the important factors for achieving a good solder printing result (Tsai, 2011; Asghar et al., 2019; Farrell, Shea & Burlington, 2013; Khader et al., 2017). Farrell et al. (2013) conducted an experiment where 12 different stencils from three different vendors, each stencil consisting of different materials. They concluded, first of all, that the stencil design had a large impact on the result of the solder printing process and further that stencils with thermally cured nanocoating performed the best, but was also more expensive and increased the production lead time (Farrell et al., 2013).

Both aperture size and stencil thickness are affecting the solder paste deposit to a significant extent, and it is very important to adjust and match aperture size with the thickness of the stencil so that the process mean targets the nominal value of the process (Pan, Tonkay, Storer, Sallade & Leandri, 2004; Tsai, 2011). The aperture size in the stencil design is an important factor according to Khader et al. (2017) to achieve a good solder paste volume. In general, stencil thicknesses vary from 100 microns up to 200 microns and the choice of thickness should be based on the size of the aperture (Gopal et al., 2006). The authors Huang, Lin, Ying & Ku (2011) discusses that the ratio between aperture area and the aperture wall area, called area ratio, is highly important for the stencil printing result. According to the IPC-7525 standard, the minimum ratio is 0,66 but should preferably be much greater for a good printing result (IPC, 2007; Huang et al., 2011). Another ratio, called aspect ratio, is defined as the width of the aperture divided by the thickness of the stencil and according to the IPC-7525 standard, this ratio should be greater than 1,5 (IPC, 2007; Huang et al., 2011).

One defect that could occur in the screen printing process, is that the solder paste area bridging with another solder paste area. This bridging phenomena could be explained through either poor board support or/and poor stencil cleanliness (Solder Paste Process, 2020). Through the experiments by Lau & Yeung (1997), they concluded that the optimal interval before cleaning the stencil was five printings.

4.1.1.4 Temperature and Humidity

Both Li et al. (2008) and Lau & Yeung (1997) conducted analysis of variance (ANOVA) where they both concluded that the temperature does affect the screen printing process. The temperature is an important factor to control since running the process at temperatures above the recommended temperatures may lead to an increased amount of part produced with poor quality (Surface Mount Process, 2020). However, the two articles had different approaches to where the temperature was measured in the process. Li et al. (2008) invested the operating temperature in the screen printing process and its effect on the quality of the output, while Lau & Yeung (1997) investigated the temperature and humidity effect during the waiting time to the next process after the screen printing process in the SMT production process.

For the operating temperature, it was found that the optimal temperature was 20°C (Li et al., 2008). In their experiments, the operating temperature factor had the three levels 20°C, 23°C and 25°C where 20°C was shown to be the most optimal. Further, reducing the current operating temperature from 23°C to 20°C changed the process mean of the solder paste height from 170m to 150m. Moreover, the temperature reduction was one important factor to be able to reduce the variance of the solder paste height. For the temperature level during the waiting time, it was concluded by the authors that the temperature of 24°C was the optimal during the waiting time, out of the temperatures 22°C, 24°C and 26°C (Lau & Yeung, 1997). Further, Lau & Yeung (1997) concluded that a humidity level of 50% was the optimal during the waiting time, out of the humidity levels 50%, 60% and 70%. Furthermore, the article concluded that a product should not wait more than 60 minutes before entering the oven to ensure that the solder paste would not deform during the waiting time and thereby endanger the quality of the screen printing process.

4.2 The Six Sigma Methodology

The Six Sigma methodology was developed by Motorola Inc. in the 1980s as a consequence of the increasing threat from Japanese competitors (Linderman, Schroeder, Zaheer & Choo, 2003). The initial quality goal with Six Sigma was to decrease the defect per million opportunity (DPMO) down to 3,4 with a process variability of ± six standard deviations from the mean (Kwak & Anbari, 2006; Linderman et al., 2003). Further, a 3,4 DPMO results in a process yield of 99,99966%, while three standard deviations or Six Sigma results in a 66 810 DPMO and a process yield of 93.3%. Linderman et al. (2003) discussed that not all processes need or should achieve Six Sigma levels, since the higher the Six Sigma level, the more costly the process will be to maintain. Further, which Six Sigma level each process should aim for depends on its strategic importance and costs of the improvement.

Six Sigma have increased in popularity amongst organisations for its financial benefits, but also for its project approach and its structured methods (Arumugam, Antony & Linderman, 2016; Schroeder et al., 2008). In addition, Kwak & Anbari (2006) discussed that Six Sigma is a business strategy with a focus of improving the understanding of the customer requirements, productivity and the financial performance. To achieve the quality goals, Six Sigma uses a systematic and data-driven five phase process called DMAIC (Define, Measure, Analyze, Improve, Control) (Kwak & Anbari, 2006). The DMAIC process utilize both quantitative and qualitative tools and techniques to succeed with an increased process yield (Antony, Kumar, & Banuelas, 2006).

Schroeder et al. (2008) discuss that the Six Sigma concept and its use of tools and techniques are similar to other quality management concepts such as total quality management (TQM) and Lean, but what makes Six Sigma different is that it provide organisations with a new level of organizational structure. One big difference between Six Sigma and Lean is that Lean focus on speed and waste while Six Sigma focus on variation, defects and process evaluation (Antony, 2011). Further, what makes Six Sigma stands out is that Six Sigma not only explains what to do, but also how to do it (Walters, 2005).

4.2.1 Define Phase

This section describes the different parts included in the Define phase in the Six Sigma methodology. The purpose of the Define phase is to define what the problem really is, the customers perception, goals and an overview of the process involved. This is done mainly through qualitative data gathering and the preferred outcome of the Define phase is to build up a stable base from which to further investigate and approach the problem. Major tasks to be achieved in the Define phase are voice of the customer (VoC), process map, supplier, input, process, output, customer (SIPOC), CTQ flowdown and finally a project charter that defines the problem and the different parts involved.

4.2.1.1 Voice of the Customer

The VoC is used in Six Sigma to identify the key factors that the project will focus on and these key factors are based on how to satisfy the customers' requirements and expectations (Carleton, 2018). Further, by thoroughly understanding the VoC, the project team will be able to define the customer needs and expectations in measurable CTQ requirements. The developed CTQ requirements will operate as the main measures to define the effect of the Six Sigma project, by comparing the CTQs before and after the project. Thereby, there are several reasons why it is crucial for the Six Sigma project to fully understand the VoC. A poor understanding of the VoC may lead the project to solve issues that the customer does not value. Furthermore, not listening to the customer fully may result in delivering an outcome that the customer already knew. Hence, to understand the VoC is critical for the project to ensure that the project is focusing on the right problems from the customer point of view.

Carleton (2018) presented a seven step guide to identify the VoC and redefine the VoC into CTQ requirements. (1) Identify both the internal and external customers of the process. (2) Determine if it is possible to divide the different customers into segments according to their needs, geography and if they are internal or external. (3) Investigate if additional information is needed and what information is missing. (4) Establish a data collection plan. (5) Gather the VoC data and categorize the data into needs, wants, expectations and key issues. (6) Based on the collected VoC data, define CTQ requirements based on the VoC needs. Also, define specifications for the CTQs. (7) Decide which CTQ requirements the Six Sigma Project will focus on. Lastly, have a dialog with the customers and project sponsor to ensure that the chosen CTQ requirements are accurate.

4.2.1.2 Affinity-Interrelationship Method

The AIM method as described by Alänge (2009) was inspired of Shoji Shiba's 19 step by step approach. The AIM is used to analyse complex issues by repetitively performing the two management tools affinity diagrams and interrelationship digraphs (Alänge, 2009). Further, by using the affinity and interrelationship methods, qualitative data in form of thoughts and experiences from each employee can be collected, and consensus on the issue can be established within the project team. Hence, the most important outcome of the AIM is not to solve the issue, but to get a shared understanding within the project team of what the problem is and starts a discussion of potential root causes to the problem.

To complete an AIM activity, the AIM team goes through 10 steps. These 10 steps are divided into four stages. After finishing each stage the team stands in a ring and say "YO-ONE" and clap their hands once.

In stage one, the AIM team needs to decide what issue to analyse and to gather data. Step one is to define which issue the AIM should analyse. The issue should be translated into a sentence for the team to have as a research question. Step two is to warm-up and let everybody talk about the theme for up to one minute. Step three is to brainstorm 19-24 post-it notes of potential root causes to the issue that was chosen. Step three is also the end of stage one.

In stage two, the AIM team first ensure the quality of the collected data and is finished by a first level grouping activity. Step four is to clarify all of the post-it notes if needed. In step five, the post-it notes are grouped for the first time. The post-it groups may consist of maximum three post-it notes. If a post-it note is not grouped together with any other post-it, this post-it is called a lone wolf, which happens regularly. This method is called affinity diagram. Step five is the end of stage two.

In stage three, the AIM team perform a higher level grouping and decide how they are connected to each other. Step six is closely related to the grouping in step five, but on another level. In step six, the AIM team group the groups that was grouped in step five. The same rule applies in step six as in step five and the groups may consist of maximum three minor groups. Step six is repeated until there are 2-5 groups, excluding potential lone wolfs. Step seven is to show connections between the groups and how they affect each other. These connections take the form of an arrow, and they can only go in one direction. This part is the interrelationship digraph method. Step eight is to conduct a final layout of the groups and their connections. Step eight is the end of stage three.

Lastly, in stage four, the AIM team evaluate and make a conclusion of the question stated in step one. Step nine is to evaluate the outcome of the AIM. The evaluation is done by voting for the which post-it groups each AIM member thinks affects the AIM question the most. Step ten is the last step where the AIM team write a conclusion to the AIM question given the voting procedure. When the conclusion sentence is finished, all the team members that took part of the AIM activity sign the AIM paper with the name, date and place. (Alänge, 2009).

4.2.1.3 Process Map

The process map is a tool to graphically visualise a process, primarily in production. The process map shows each step in the process, it's input variables and output variables as well as non-value adding activities in the process. The input variables should be classified as either noise, control factors or standard operating procedures. To create a process map, it is of high importance to set the boundaries of the process and to define on what level it should be visualised on. This determines the complexity and the ease of interpretation and visualisation. To do this, a good way is to walk the entire process and document all production steps to be able to describe it efficiently. (Carleton, 2018).

4.2.1.4 CTQ Flowdown

The critical to (CTQ) flowdown tree tool is used to define DMAIC projects (Carleton, 2018). Further, the CTQ flowdown is a tool used to analyse a problems root cause by breaking the problem down into critical segments (LeanOhio, 2014). Furthermore, it supports the project to identify the largest posts for improvements. Hence, the CTQ flowdown tool is used to ensure that the project is working on critical issues for the business and its customers (Carleton, 2018). The CTQ flowdown tool does this by linking the customer requirements to the product or service characteristics.

4.2.1.5 SIPOC/Effective Scoping

SIPOC is a type of process map used to assess other important parts of the process. It is used to document the process at a higher level than for example a graphical process map, to create a visualisation of the process flow and what is in it (Carleton, 2018). As mentioned in 4.2.1, SIPOC stands for suppliers, inputs, process, outputs and customers. SIPOC helps define the boundaries of the process, identifies important actors as customers and suppliers and the different inputs and outputs of the process (Carleton, 2018). The one used for this project is an extended version developed by Peter Hammersberg, senior lecturer and Black Belt Champion at Chalmers University of Technology. It is extended to provide more in depth details on input, process and output since these often needs to be described more closely (Hammersberg, 2019a). The input section is extended with the description on what the inputs require from the system itself (Hammersberg, 2019a). The process section is extended to include competencies in the team handling the process and the jurisdiction to commit potential changes to the process (Hammersberg, 2019a). The output section is the one that has been most extended with descriptions on who uses the input, a list on requirements on the output and what one measure, the big Y, that should be understood and improved based on the output (Hammersberg, 2019a). Further included, based on the Y, what facts are behind the proposed improvement proposal and what other parameters that cannot get lost in the process.

The purpose behind SIPOC and effective scoping is to realize as early as possible what the problem really is and if it is related to something inside of the process or if it is entering with one of the inputs (Hammersberg, 2019b). Therefore, it becomes much more easy to elevate the right questions and issues to the relevant level that are able to handle the problem.

4.2.1.6 Project Benefit Assessment

The project benefit assessment is a tool used in the Define phase to identify and estimate hard and soft benefits of the project (Carleton, 2018). Further, by using the project benefit assessment a company can set a value of different projects. Hence, different projects can be compared to each other. Also, by being aware of the benefits of different projects, the company can develop a well-balanced project portfolio.

Hard benefits are defined as benefits that are visible in the financial accounts (Carleton, 2018). Further, the hard benefits are sorted into either (1) revenue enhancement or cost reduction or (2) working capital reduction. Revenue enhancement is benefits the project causes that leads to increased revenue such as increased sales. Cost reduction is benefits the project cause that leads to reduction of cost such as improved quality, reduced scrap, or increased efficiency. The ability for a company to pay its current liabilities with its current assets is called working capital (Investopedia, 2020). Furthermore, working capital reduction are benefits the project cause that leads to reduction of cost such as operating costs or inventory cost. Soft benefits are defined as benefits that are not visible in the financial accounts but valuable in other ways (Carleton, 2018). The soft benefits are sorted into either (1) cost avoidance or (2) non-financial benefits. Cost avoidance is benefits the project cause that leads to reduction of cost due to regulatory compliance and a decreased need to increase resources to meet increased demand. Non-financial benefits are benefits the project cause that leads to an increased value for the company that is not possible to set a value on, such as customer satisfaction, employee engagement and sustainability enhancements.

4.2.1.7 Project Charter

The project charter is the outcome of the Define phase that includes information of the projects business case, project problem, project goal statements, project scope, project plan and team members (Carleton, 2018). In other words, the project charter is a one-page summary of the whole Define phase.

The business case should include a baseline for the project and a clear linkage to the company's upper level strategy, identification of the customer needs and an initial estimation of the financial forecast of the projects impact. The problem statement is a description of the current state and the desired performance levels. The problem should be stated in measurable terms. The goal statements is a description of which goals and success criteria the project aims to achieve. It should be stated what the goals are and when they should be delivered. The project scope is a description of the process boundaries for the project. The project plan should include target dates of when the project aims to start and finish each phase. Lastly, the team members are a list of all team members and their roles and responsibilities. (Carleton, 2018).

4.2.2 Measure Phase

This section describes the different parts included in the Measure phase in the Six Sigma methodology. The purpose of the Measure phase is to collect relevant data to solve the stated problem in the Define phase. This is mainly done through quantitative data gathering. The outcome of the Measure phase should include relevant and enough quantitative data which will further be analysed in the Analyze phase. Major tasks in the Measure phase include p-diagram, data collection plan and an MSA.

4.2.2.1 P-diagram

The p-diagram, or parameter-diagram, is a tool used to describe and classify the parameters involved in the process as either noise factors, control factors or input signals (Carlson, 2020). This is done to see which parameters that can be affected and those who cannot get affected. The p-diagram thus helps focusing on what parameters that matters and can be controlled or changed and those that cannot be controlled due to complexity, cost or environmental factors. The p-diagram further gives a visualisation of the process and what parameters that affect the process (Juran, 1993).

Input signals are essential to the functionality of the process system and provides the functions, such as speed and pressure. Control factors are design parameters of the process that the engineering team have the possibility to change, such as stiffness and density. Noise factors are those factors that cannot be controlled by the engineering team, such as wear and environmental conditions. (Carlson, 2020).

4.2.2.2 Data Collection Plan

The data collection plan is done to assess what data should be measure, how much data should be measured and how often. The data collection plan can typically be structured in many different ways but should involve metrics, control and noise factors, the process measured, specification limits, frequency, sample size and storage (Carleton, 2018).

4.2.2.3 Measurement System Analysis

MSA, or measurement system analysis, is used to assess whether the measurement system measures correctly or not (Burdick, Borror & Montgomery, 2005; Carleton, 2018). The measuring might be a source of variation and thus it is highly important to investigate if the measurability is good or not since it allows the company to measure the true capability of their system (Carleton, 2018). By doing this, the company can also accept or reject produced units in an efficient way and trust that the quality of the system is at an adequate level. There are further two different types of MSAs depending on whether the data are variable or attribute (Carleton, 2018). In this project, the focus will be on variable data and will thus be presented below.

The two most important parameters in an MSA related to variable data is repeatability and reproducibility. Repeatability is related to variation due to gauge while reproducibility is related to variation due to operators (Burdick et al., 2005). These two parameters can also be translated to part-to-part variation and measurement system variation when using statistical programmes (Minitab, 2020). Repeatability is related to the measuring device, measuring the same part with the same measuring device repeatedly by one operator while reproducibility regards the same part being measured by the same measuring device but with different operators (Minitab, 2020).

There are different methods to do MSAs, including manual calculations based on known standards and calculation of the average standard, but the most commonly used are Gauge R&R where R&R stands for repeatability and reproducibility (Carleton, 2018). Gauge R&R is especially good to investigate if there are any excess variability in the system so that the true capability of the process can be assessed (Carleton, 2018). To conduct a Gauge R&R, there are several statistical programmes that might be used, and Minitab is one of them. When conducting a Gauge R&R in Minitab, there will be two important measures in the results section under percent contribution: Total Gauge R&R, with subsections repeatability and reproducibility and part-to-part variation (Minitab, 2020). Part-to-part variation involves the variability in measuring between different parts in the process (Minitab, 2020). The total variability in the process is thus broken down into these two measures and for an ideal measurement system, part-to-part variation should account for the major part of the total variation in the system (Minitab, 2020).

If the percent contribution (1) $< 1\% \rightarrow$ the system is acceptable; (2) 1% - $9\% \rightarrow$ potentially acceptable depending on the conditions (cost, how critical the measurement is, risks, etc.); and (2) $> 9\% \rightarrow$ the system is not acceptable, according to AIAG (Automotive Industry Action Group) (Minitab, 2020).

If the percent study variation and tolerance is (1) $< 10\% \rightarrow$ The system is acceptable; (2) 10% - $30\% \rightarrow$ Potentially acceptable depending on the conditions (cost, how critical the measurement is, risks, etc.); and (3)$> 30\% \rightarrow$ the system is not acceptable, according to AIAG (Minitab, 2020).

4.2.2.3.1 Percent Study Variation Percent study variation is another important metric in the results section in Minitab. This relates the measurement variation to the total variation in the system and should be as low as possible. Six standard deviations, related to sigma, should be used where it typically are defined as process variation, which also means that when data are normally distributed, approximately 99,73% of all data should fall within the six standard deviations or plus/minus three standard deviations from the mean. Percent tolerance can also be important and compare measurement variation to the set tolerance of the system. However, percent tolerance are better used if there are specifications that the variation should be tested against, otherwise percent study variation are much more appropriate. (Minitab, 2020).

4.2.2.3.2 Number of Distinct Categories Number of distinct categories equals to the number of distinctive groups of parts that the system can find in the measuring. The best result is, according to AIAG, if the number of distinct categories are five or more since this allows the system to distinguish between parts and the system are therefore also acceptable. If below two, the system cannot distinguish between parts and if exactly two, the system divides the parts into two categories, high and low. (Minitab, 2020).

4.2.2.3.3 MSA Charts Further in the results in Minitab, there are several graphical charts being part of the results. The R chart consists of ranges to show operator consistency, how well the operator measure. The chart shows the range between the smallest and largest measurement on how each operator measures each part. In the chart, there is also a centre line based on the average of the range between the measurements. Above and below the centre line, upper and lower control limits can be seen and are calculated based on subgroup variation. If an operator measures consistently, the data points will fall between the upper and lower control limit and the range will be small. However, if the operator measures inconsistently, the data points will fall outside of the control limits. (Minitab, 2020).

The X-bar chart regards repeatability and variation between parts. Just as the R chart, the X-bar chart also have data points regarding average measurement of each part, a centre line based on the overall average of the measurements and upper and lower control limits based on the number of measurements and repeatability. The X-bar chart on the other hand are much more focused on how much control there are or not in the measuring. Variation between part averages is the ideal here rather than variation from repeatability alone. If the data points are outside of the control limits, this shows that there are more variation between parts than for the actual measuring system itself. (Minitab, 2020).

Finally, there are three graphs that show operator by part interaction, measurements by operator and measurements by part. Ideally, these graphs should be highly identical and if the range between the data points related to each part of operator are big, this shows that the measuring is inconsistent. (Minitab, 2020).

4.2.3 Analyze Phase

This section describes the different parts included in the Analyze phase in the Six Sigma methodology. The purpose of the Analyze phase is to analyse the main data from the Measure phase and to find potential connections and solutions to the stated problem. This is done through quantitative analysis in R where multiple regression, bagging, boosting and random forests will be used.

4.2.3.1 Multiple Regression

Multiple regression is a method to explain and predict one variable from the basis of the values of two or more other variables (Carleton, 2018). The method accomplishes this by relating multiple x's, called predictor variables, to a y, called dependent variable. Further, Carleton (2018) discussed the importance of the dependent variable, the y, to be a continuous variable when conducting a multiple regression analysis and the use of the method to investigate if there are statistical connections between the y and the x's.

4.2.3.2 Recursive Partitioning

Recursive partitioning involves dividing the x's into as small and homogeneous subgroups as possible. The data that are being partitioned should ultimately belong to one class or one subgroup. Each partition or split of the data results in two nodes which ultimately will end up in a tree like structure turned upside down. The tree consists of decision nodes and terminal nodes where the decision nodes have successors, that is the terminal nodes. The terminal nodes, on the other hand, are the "leaves", the end nodes and these should represent the small, homogeneous subgroups mentioned above. A regression tree is part of recursive partitioning and includes numerical values related to the decision and terminal nodes. (Shmueli, Bruce, Yahav, Patel & Lichtendahl, 2018).

Visually, the regression tree first starts with a scatter plot consisting of a set of data. The scatter plot are then divided with horizontal and vertical lines, creating boxes or regions of different sizes. Each box or region represents a terminal node. (James et al., 2013).

4.2.3.3 Bagging

Bagging, or bootstrap aggregating, is a type of ensemble which means that models are combined to create a more powerful model (Shmueli et al., 2018). By taking many random data samples and averaging them against each other, performance stability is improved as well as reducing the risk of overfitting, having too many terminal nodes (Shmueli et al., 2018). Most often there are not more than one training data set and therefore bootstrapping is a good way of taking smaller data samples from the data training set and through bagging train the method to get a prediction of X and then take an average out of the predictions to create a prediction formula for future data (James et al., 2013).

4.2.3.4 Boosting

Boosting is a different variant of an ensemble and instead of improving the model through averaging different samples against each other, boosting improves areas in the data set that makes errors, by directing or forcing the model against those areas. A data sample containing errors are taken out from the data set and then the model should be fit towards that sample. (Shmueli et al., 2018).

4.2.3.5 Random Forest

Random forest is a type of bagging and includes combining many decision trees against each other to make better predictions (Shmueli et al., 2018). The method works similar to bagging but instead of just averaging data samples against each other, decision trees are put against each other, thus creating a forest of decision trees (Shmueli et al., 2018). The predictions from each tree are then combined to improve the overall prediction of the model (Shmueli et al., 2018). When the trees are averaged against each other variance are also being reduced (James et al., 2013).

5
Empirical Data

In this chapter, the findings and outcomes of the methods used will be presented. It is divided into three subchapters; Define phase, Measure phase and Analyze phase. In the Define phase, different methods will be used to define the problem and problem area. The Measure phase will mainly investigate if the project team can trust the data. Lastly, in the Analyze phase, the quantitative data will be analysed to investigate potential root causes to the yield variation.

5.1 Define phase

To fully understand the background of the problem area and if the current perception of the problem is based on facts, a number of tools was used. These were SMT production line observations, interviews, process map, understand VoC, CTQ flowdown, project benefit assessment, AIM, effective scoping and project charter.

5.1.1 Observations

When the project began in the middle of January, the project team arrived to the production site for the first time. The project then began with an introduction to the relevant people of the project as well as a round tour of the production site, showing every line and process step. Observations were throughout the entire production line to better assess every step of the production process. Every process step was explained by the project owner at the case company, with the help of both line operators and process engineers. Furthermore, access was given so that the project team could enter the production to do observations whenever necessary. The people involved in the project consisted mainly of process engineers with responsibilities for different areas in the production process.

5.1.1.1 Process map

From the observations in the production and as explained by the project owner and related personnel, a process map was made which is presented below in Figure 5.1.

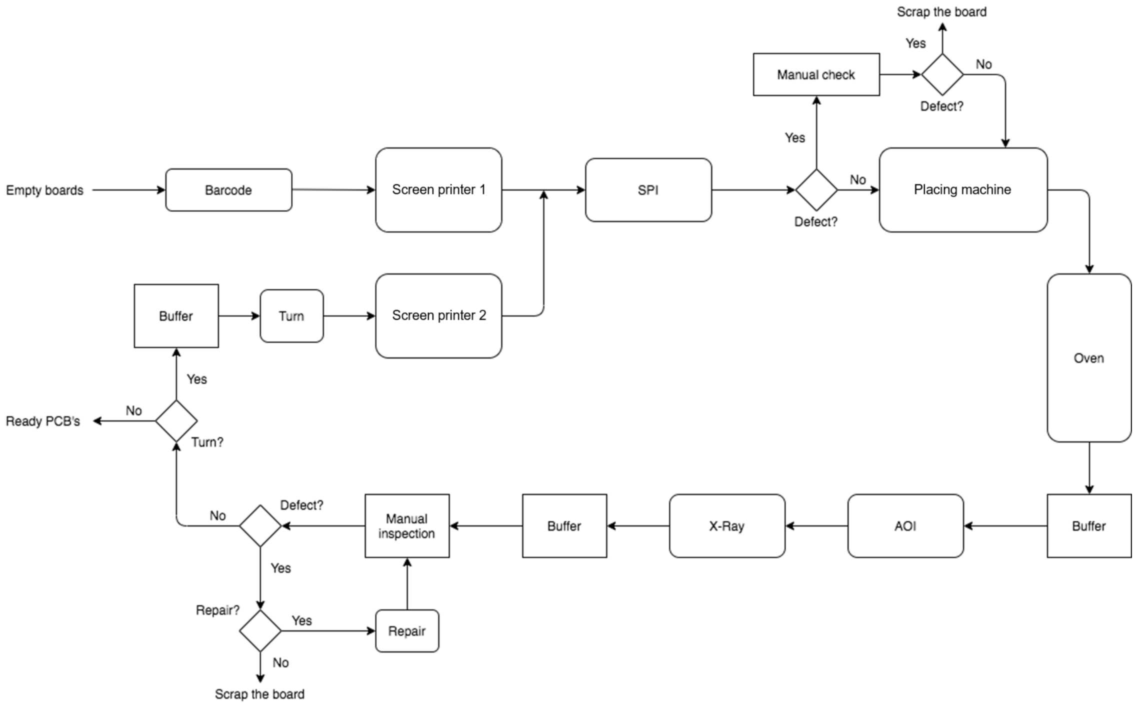


Figure 5.1: Process map of an SMT production line at the case company.

The process works as follows:

1. **Barcode:** An empty board enters the process and arrives to the barcode station. Here, the board gets a 3D-printed barcode to mark the board and give it an ID-number.
2. **Screen printer 1:** The board continues to the next station, the screen printer. Here solder paste is applied to the board. A stencil design, matching the current product produced, is laid on top of the board and solder paste is then applied through the aperture.
3. **SPI:** The board, now with solder paste, arrives to the SPI-station where it measure different parameters related to the solder paste. The SPI-station also checks whether the amount of applied solder pate meets the set tolerances.
   i **Yes:** If the amount of applied solder paste meets the set tolerances, it continues to the placing machine station.
   ii **No:** If not, it shows on a screen as defect and enters a manual check station. Here an operator manually checks whether it is defective or not. If defect, it gets scrapped. If not, it continues to the placing machine station.
4. **Placing Machine:** In this step, components are put into each assigned pad on the board.
5. **Oven:** The board enters an oven with different temperatures to melt the solder paste to make the components stick.
6. **Buffer:** The board now enters a buffer with space for up to 25 boards.
7. **AOI:** The board enters the Area Optical Inspection which is a method to optically inspect the board to check for potential defects. Measure if the right component is attached to the right place.

8. **X-Ray:** A x-ray is taken of the board to check for defects inside the board and to check if all electrical connections work.
9. **Buffer:** The board now enters a buffer with space for up to 25 boards.
10. **Manual inspection:** A manual inspection is then carried out to the board to check for defects.
    i If defect, the operator decides whether it should be repaired or not. If repaired, it returns for another manual inspection.
    ii If defect and not repairable, the board is scrapped.
    iii If not defect, it continues through the process.
11. **Turn:** If the board has just been manufactured on one side, then the board go to a buffer and then turned followed by the same, previous process on the untouched side. If both sides has been manufactured, it goes out of the process as a ready PCB.
12. **Screen printer 2:** Screen printer 2 works exactly the same as screen printer 1 but is another identical machine.

5.1.2 Interviews

Interviews were conducted with three key responsible process engineers in the production to get a better understanding of the production, the process and the underlying problems to the project. The first interview was done with person A, technical process engineer with the responsibility for the SPI-stations. The second interview was conducted with person B and person C, both responsible for the second line, which means that they are having more of an overlooking perspective.

5.1.2.1 Interview with technical process engineer

From the first interview with a technical process engineer, referred to as person A, it was made clear that the SMT lines deliver their PCBs internally to the assembly lines. The SMT lines are very repeatable which means that they are very good at producing the same products without the interference of noise factors. The human factor is more important according to person A, and this is where many of the defects occur through.

> *"A machine is better at repeating without being influenced by external factors, while the human factor may affect more. Different operators do things in their unique way and have their own routines. You end up in an OK-flow and then you maybe pass through a bad unit."*

The operators are not equally trained, have different routines, might get stuck in a flow of pressing "OK" when there appears to be a defect to the board and it needs a manual check and so on. Another problem is that if there is a defect to just one PCB on the board, the whole board are getting scrapped. Currently, tests are being conducted on SMT line 7 to make it possible to only scrap the defective PCB and not the entire board, thus saving both time, money and resources. As by now, repairs and rework are not conducted due to customer requirements and that it is perceived as safer to scrap than to risk deliver a defect product. The cost for scrapping is currently only calculated in material scrapping cost, not including other cost factors related to scrapping.

From the interview, it was also concluded that solder paste volume is the most important parameter with the baseline tolerance 100% with lower tolerance 40% and upper tolerance 180% on each pad on the board, all according to the IPC A610 class III standard (IPC, 2014). If the volume is calculated to be below 60%, an alarm sounds and if below 40%, the machine stops. On SMT6, a test is conducted where the operator should not be able to press OK on a board with solder paste volume below 40%. This will reduce potential human errors and further be implemented on all SMT lines. Regarding the upper tolerance, if the SPI measures are above 180%, there is a large possibility for something to be defective according to person A.

> *"Solder paste volume is the most important metric we use when deciding if scrapping a PCB or not. Too little solder paste after the screen printing process, cleaning intervals of the screen printing process and how they are maintained, may lead to the PCB not being OK in the SPI-station."*

The solder paste volume is the most important metric, the big Y, according to person A, and it involves how accurate they measure. Sometimes, an operator notices the process engineers that they have to approve many potentially defective boards during a short period of time. If it varies more than ±20% in solder paste volume, they do a bare board teach in order to "teach" the SPI on the measuring. Generally, there is little variation between the SMT lines but this is not something that are followed up and the same thing with the products themselves. The same board can be manufactured into several different final products and by doing so, more or smaller components are being placed on the boards so that they achieve different functions. To produce the boards and apply solder paste, different stencils are used which are possible to see in the data. Different SMT lines are further producing different products, which may result in some variation, that should also be possible to see in the data.

About potential causes of the high scrap levels and any possible connections between the parameters, person A suggested that too little solder paste after the screen printing could result in a potential defect. The cleaning of the screen printer might also be one factor as well as the type of solder paste used. There are also delays between two screen printings where the first print often gets defects. Related to the data and parameters measured, person A explained that they measure those parameters that physically can be measured and due to long experience in the industry measuring these parameters. There are also the IPC-A610 standard to take into account which sets high requirements on the products. Related to scrap, person A suggested that most boards are scrapped right after the SPI-station.

5.1.2.2 Interview with 2nd line process engineers

The second interview was conducted with person B and person C, process engineers with responsibility for the 2nd line, supporting escalated problems from production.

From the interview, both information that supported person A's view was presented but also new information. The quality is of highest importance for the case company since the PCBs are supposed to last for at least 15 years and they cannot contain any defects since they should support autonomous vehicles as well as people driving non-autonomous vehicles.

There are according to person B and person C, many reasons for a PCB board to be scrapped and potential scrap reasons are mostly described in the IPC A610 class III standard (IPC, 2014). That said, apart from defects related to solder paste volume and others, customer requirements are sometimes very fuzzy and unclear and it is difficult to get clear directions what actually customer requirements are. To further complicate things, rework are allowed to some products and to other products not. There are few standards regarding requirements and typically there are no customers that have the same product and thus the same requirements regarding quality controls, rework and scrapping.

> *"The biggest scrap post at the moment is too little solder paste volume on the PCBs. When the volume is below 40%, the PCB is not OK. One reason for this is clogged apertures in the stencils that prevents all the solder paste from being applied."*

Regarding scrapping, the largest scrapping post at the moment is too little solder paste through the stencil apertures. The reason behind this might be that the aperture gets clogged with time and too little solder paste gets applied to the PCB boards. If the solder paste volume reaches below 40% it is getting scrapped immediately. One problem is that on SMT1-5, the operator should scrap a board that has under 40% in solder paste volume, but the operator has the possibility to let it through, which happen sometimes. When there is too much solder paste, above 180%, it might be caused by operators doing manual work to the stencil aperture where solder paste has been applied manually. The operators have further learned that they might save time by not working according to set standards. Another potential defect might be caused by the solder paste flux that might cause the solder paste to become runny or create air pockets. There is currently a project running trying to solve this problem from the suppliers side.

> *"Operators have learned that it is possible to save time by working outside the standards. Also, how much solder paste that is added in the process is handled manually by the operator."*

The most important parameters are the volume while the shape (including height and area) is not as important. Further, offset might have an impact, but this does not often prove to be a problem due to the fact that the solder paste partially melts in the oven and therefore covers the pad. Volume is, therefore, the most important parameter that stands out from the rest according to person B and person C. There is no consensus on why they measure the many parameters that they measure today compared to the production in Canada for example, where only three parameters are measured. According to person C, it is because the customers are not as demanding as they are in Europe.

Related to the different sides of the board, the scrapping level looks the same but most attention are being paid to the underside of the board since there has been more components placed to the board when the process comes to the underside. Therefore, it is much more expensive to scrap the board and also the most expensive components are placed on the board as late as possible to avoid the board being scrapped with them placed on it. Older products, that are larger in size, also tend to have fewer problems while newer, smaller, sized products, tend to have more problems. Both person B and person C also stress the potential problem with the cleaning of the stencil designs and how clogged the apertures might be after a few cycles in production.

Person B and person C also brought up that analysing data is a quite new thing at the case company and that they just have started to look at it but it is hard to analyse and interpret. There are many different factors, some hard to assess, that impacts the final result. The biggest problem right now is that there is most often too little solder paste on the boards when the SPI measures and that it is unsure if they can trust the data from the SPI at all. Different boards can have different thicknesses from the supplier which could result in the results being misleading. It might be a good idea to look at the percentage solder paste that does not stick to the board and what is happening with that amount.

5.1.3 Voice of the Customer

The project team identified three main internal customers for the projects concerning scrapping levels at the SMT production lines. The production area team at the production site plan, follow and continuously improve the SMT production and flow. Therefore, the production quality function of the production area team for the SMT production lines is one of the main customers for the Six Sigma project. For each station at the SMT lines, process engineers are process owners of each process. This means that they are responsible for the process that they are process owners of. Therefore, it is in their interest for the project to investigate root causes for variation in the SMT production lines' stations. The third internal customer that was identified is the assembly lines where the final product is assembled. These assembly lines use the output from the SMT production lines, the PCBs, to assemble the final product. If there is any defect to the PCB, there is a high risk that an assembled product needs to be scrapped, hence resulting in scrapping costs and time loss for the case company and the assembly line. The primary external customer is further the customers of the final products at the assembly lines.

The production area team and the process owners both share an interest in reducing the SMT production line variations. The production area team needs a stable production to be able to produce according to plan and satisfy customer demands. One way of doing this is by reducing scrap and production stops. It is in the process owners interest to gain a more stable process. Also, it is in the whole production plant's interest to decrease variations in the SMT lines since they to lower production cost to gain more revenue for further investments to stay competitive in the long run.

The process engineer who is the process owner of the SPI-stations has collected data from each SPI-station at all SMT production lines daily since November 2019. Due to the huge amount of data, only the data from the latest 50 produced boards have been collected. The production area team also have the possibility to extract scrap cost and scrap attributes from the beginning of July 2019. Hence, data of the past is available. However, the project team needs to collect data to conduct MSA studies to investigate if the data can be trusted.

One of the main objectives of the master thesis is to investigate the yield and the production cost at the SMT production lines, see 1.2. Through a business analysis software program, the production area team is able to extract specific costs to SMT lines and its stations. Figure 5.2 below visualizes the breakdown of the scrapping cost at all of the SMT production lines. Further, it is based on scrapping figures, in the form of Pareto diagrams, from the production area team from 1st of September 2019 to 31st of January 2020. Hence, the figures represent the total SMT production costs during a period of five months. Scrapping costs of the screen printing processes 1 & 2 and SPI are considered as one area since the SPI-station controls the output of the screen printing processes. The total percentage of these scrapping cost is 20,7% of total costs. As can be seen in Figure 5.2, scrap cost placing machine is 32,8% of total costs, which is more than the total scrapping cost of screen printing process 1 & 2 and SPI. As can be seen in the process map, Figure 5.1, the screen printing process is followed by the placing machine process where the PCB components are added. However, in a dialog with the process engineers and team, it was decided to focus on the SPI and screen printing processes. The main reason for this choice was that a poor screen printing process may lead to poor connections between the solder paste and the components, which will lead to increased scrapping cost at the placing machine station and in further process steps. Hence, focusing on the SPI and screen printing process was decided together with the project stakeholders.

Through the CTQ flowdown tree, the project team was able to define the case company's key issue, expectations, needs and wants. The key issue that was identified was that the SMT production lines have high scrap levels due to variations in amount of solder paste that gets applied on the PCBs. Further, the production area team and process engineers expect a stable process and reliable and trustworthy measuring from the SPI-machines. Furthermore, the production area team needs less variations in applied solder paste volume to be able to reduce the scrapping levels. To do so, the production area team wants to better understanding potential root causes to the solder paste volume variations.

Finally, the VoC definition the project team developed: SPI need the right amount of solder paste on the PCBs in the screen printing processes at the SMT production lines.

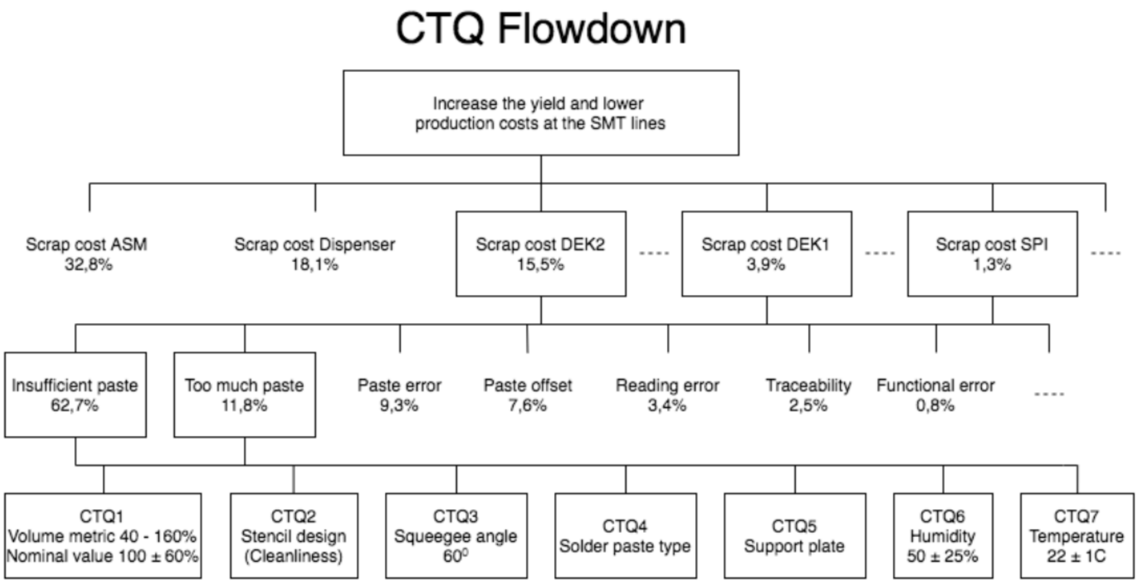


Figure 5.2: CTQ flowdown tree.

5.1.4 AIM

The AIM activity was performed by four process engineers with different responsibilities on the SMT production line; one process engineer with process ownership of the screen printing process, one process engineer with process ownership of the SPI-stations, one process engineer with process ownership of the AOI-station and one process engineer with responsibility of one entire SMT production line. Since the process engineers only were able to be away form their daily work for a limited amount of time, the AIM had to take a maximum of two hours. Due to the time limitations, the project team prepared the AIM question beforehand. The AIM question was "What were the reasons for the variations in applied solder paste that gets applied on the PCBs at the SMT production lines?". Also, the AIM activity was conducted in Swedish since all of the participants were Swedish and thus save time and avoid misunderstandings. Therefore, the original Swedish AIM results is presented in Appendix A, and a translated version presented in Figure 5.3.

Through the brainstorming and grouping steps, the AIM team produced five main groups and two lone wolfs. The five main groups were (1) maintenance, (2) educate operators, (3) cleanliness, (4) SPI-stations and (5) screen printing process. The two lone wolfs were (1) variances in thickness on the solder mask coating and (2) factorial tests with new solder paste to optimize the settings.

The conclusion was that the causes of the variations were that stencil design and wear may result in variations in solder paste volume. However, the AIM team did not agree completely. As the voting and the amount of arrows shows in Figure 5.3, the voting points and arrows are all over the place.

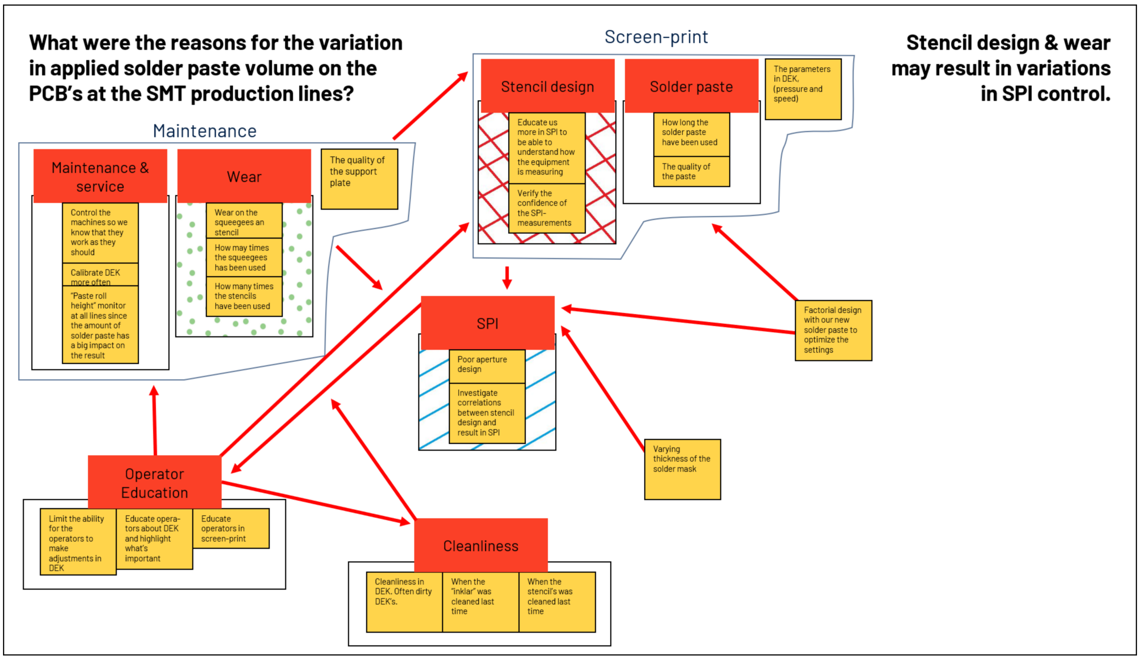


Figure 5.3: Translated AIM results, see the original in Appendix A.

5.1.5 Effective Scoping

The project team have conducted an effective scoping in order to define the process further. The outcome of the effective scoping is visualized in Figure 5.4 below. Further, the workflow of the effective scoping was divided into three sub-steps. First, the output and customer are defined. Secondly, the process is defined. Lastly, the supplier and inputs are defined.

The first steps in the effective scoping method is to define the output and the customer of the process, which are called Q1-Q6. The output of the screen printing process is a board with a certain amount of PCBs with solder paste applied to all the different areas where components will be added later in the SMT process. The customer of this process is viewed to be the SPI-stations since the SPI-stations measures the output of the screen printing process. For the SPI-stations to approve that the board to move to the next process step, the SPI-station controls if the solder paste volume, solder paste deposit offset (Y- and X-direction), solder paste deposit area and solder paste deposit height is within tolerances among other metrics. Step Q4 regards the choice one measure for the Six Sigma project to focus on. Since the CTQ flowdown tree, see Figure 5.2, showed that approximately 75% of the scrap reasons is due to insufficient or too much solder paste, the project team chose to focus on the solder paste volume metric. However, to ensure that the project team can trust the data collected, MSA studies are needed to analyse if the SPI-stations are trustworthy. Also, it is important to have in mind that constraints for the project team to reflect upon is that changes to improve the Y cannot increase the lead time at the SMT production line, change the layout of the SMT production line, or endanger the quality of the output.

Effective Scoping of continuous improvement projects

The sequence in itself, of questions Q1-Q4, Q5-Q7 and Q8-Q9 below, is key to facilitate consensus in the shift of an organisation's mindsets from push to pull, in accordance with the principles of Lean Six Sigma

Supplier	Input		Process	Output		Customer
Q8b. Who supplies the inputs?	Q8a. What are the inputs to the system?	Q9. What does the system require of the inputs?	Q7a. Team/project jurisdiction of changes	Q1. What comes out (of the physical flow) - OUTPUT?	Q3. What is required of the output from this particular user (List of big Y's and improvement proposals)	Q2. Who uses the output?
- *Board with several PCB's* - Supplier: Kanban-train & operators - *Solder paste* - Supplier: Kanban-train & operators - *Stencil fixture* - Supplier: Operators - *Squeegees* - Supplier: Operators - *Support plate* - Supplier: Operators - *Humidity & temperature* - Supplier: A climate control system	- Board with several PCB's - Solder paste - Stencil fixture - Squeegees - Support plate - Humidity & temperature	- *Board with several PCB's* - Requirement: The QR-code have to align with the program - *Solder paste* - Requirement: The QR-code have to align with the program - *Stencil fixture* - Requirement: The QR-code have to align with the program - *Squeegees* - Requirement: None requirement at the moment. However, QR code as the above are planned to be implemented - *Support plate* - Requirement: The QR-code have to align with the program - *Humidity & temperature* - Requirement: 50±25% humidity & 22±1°C	The proposed changes of the project will most likely not interfer with the current operations. The only thing that might conflict with other interests are if a big investment needs to be made, if the proposed change increase the cycle time of the process or changes that involves big changes in operations or organisation.	A PCB/board (with the applied volume of solder paste within the specification limits as well as components attached)	**Y1. Solder paste volume within tolerances** **Y2. Solder paste deposit offset within tolerances (Y- and X-direction)** **Y3. Solder paste deposit area** **Y4. Solder paste deposit height**	SPI uses the output from DEK1 & DEK2
			Q7b. What competences are needed in the team (WHO)?		**Q4. What ONE MEASURE (y) should be understood and improved? The y that scope the project and drive further exploration.** **Each small y has its own underlying system of influencing parameters, sometime overlapping. Use one template per y to reduce complexity** **Scope on y (not x - upstream) and don't proceed until Q1-Q4 is thoroughly understood!**	
			1. Process engineers at the SMT production lines 2. Process owners of the SMT production stations 3. AMG team		**Y1=y(x)** **Solder paste volume within the set tolerances are the main Y and cause of the high scrapping rates.**	
			Name of the underlying system that build up the y to be improved:		Q5. What is the baseline of the y and can that precis y be measured today (and can old data be trusted)? In other words: What is the **facts** behind the problem that form the base for our improvement promise? Show the data/proof of a problem!	
			DEK 1 & 2, SPI at SMT production lines, operators, process owners		- SPI measures the applied volume, the area covered by the solder paste, the height of the solder paste, and the offset in X and Y - Still unclear if the old data can be trusted, need to analyse this and to do tests - Facts exists on that insufficient paste is the main problem (data from production SCRAP costs)	
			From where is the physical output shipped?		Q6. What other Y can not be lost in the process (constraints)?	
			Internal assembly lines.		1. Changes cannot increase the lead time at the SMT production line 2. Changes cannot change the layout of the SMT production line 3. Changes in a reduction of scrapping costs cannot endanger the quality of the output	

Figure 5.4: Effective Scoping.

The second step in effective scoping is to define what authority the project team have to make changes to the process, which are called Q7a and Q7b. The proposed changes of the project will most likely not interfere with current operations. However, a few reasons that might conflict with other interests are if a big investment needs to be made, if the proposed change increases the cycle time of the process or changes that involves big changes in operations or organisation. For the Six Sigma project to be able to implement changes, it will need the competences from SMT process engineers, process owners of the SMT production stations and the production area team.

The third step in effective scoping is to define what inputs the process have and how they are supplied to the process, which are called Q8a-Q9. The screening process has several inputs, such as a board with several PCBs, solder paste, stencil fixture including a stencil, squeegee angle, squeegee speed, squeegee pressure, support plate, temperature and humidity. Mainly, the line operators are supplying all of the inputs, sometimes in combination with the Kanban train. However, temperature and humidity levels are supplied and controlled by a climate control system inside of the screen printing process. Further, for the process to be able to run, it has to accept the inputs through a QR-code system. Most of the inputs are controlled through a QR-code which has to align with the program settings for that specific input. However, for the temperature and humidity, different rules apply. In order for the screen printing process to run, the process require the temperature to be within a temperature level of 22±1ºC and a humidity level of 50±25% which are controlled by the climate control system.

5.1.6 Project Benefit Assessment

There are several benefits of this master thesis for the case company, as can be seen in Figure 5.5 below. The benefits are divided into hard and soft benefits.

Starting with the hard benefits, the project has an annual cost reduction potential of X SEK. This has been calculated from figures and data from which Figure 5.2 is based upon. Hence, scrapping costs are related to both insufficient solder paste and excessive solder paste. Further, the sum of insufficient and too much solder paste was divided by five to get a monthly cost; thereafter, the quotient was multiplied with 12 to get an estimated annual cost. Other hard project benefits are related to operating costs. By reducing the error of either insufficient or too much applied solder paste, the less time the production operators have to control and manage the process. Thereby, the operators can focus on other things on the production lines, for example directing efforts on the final check to avoid production stops which affects the total yield negatively.

The soft benefits for the project are divided into two subgroups: cost avoidance and non-financial benefits. First, if a product has a defect it is either scrapped or re-worked. Both cases may lead to a production stop in one way or another, taking time and resources from the production, mainly time and personnel. Hence, through reduction of production errors due to insufficient or too much applied solder paste, the case company can avoid cost due to scrap, rework and production stops. However, reducing scrapping levels is not only good for cost avoidance reasons but time and resource can be spent into other projects or to improve production further. It is also positive from a sustainability approach since less scrap means less resources that gets used in the process. Other non-financial benefits of the project is that the case company and its customers will increase the knowledge level of the SMT process and become benchmarks, both internally and externally.

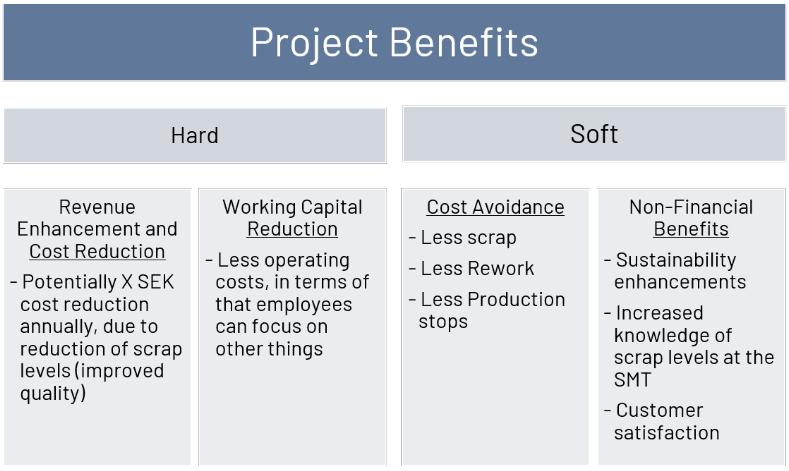


Figure 5.5: Project Benefits Assessment Matrix.

5.1.7 Project Charter

The output of the Define phase is the project charter, which is visualized in Figure 5.6. Through conversations with the production quality manager, to reduce the production cost is one of the five strategic objectives from the case company's global management. Hence, to reduce SMT production scrap costs is of great importance. The project problem statement is that the two scrap posts, insufficient and excessive solder paste, represent 75% of total scrapping costs in the screen printing processes with an estimated annual cost of X SEK. Hence, the scope of the project is limited to only investigating the screen printing processes and how the SPI-stations measures at the SMT production lines. The goal statement of the master thesis is to provide the case company with facts that explains the majority of the screen printing process variation and which parameters that affects the yield the most. Furthermore, these facts will result in recommendations for implementing countermeasures. Due to the time limit of the master thesis, the project plan is to spend approximately one month per phase. The project team consist mostly of process engineers but also the production quality side of the production area team.

Business Case	Problem/ Project Statement
In order to stay competitive and profitable, one out of the five strategic objectives from the case company's global management level is to reduce cost in general, which production scrap costs is a big part of. Initial financial forecast: reducing the SCRAP cost by 50 % will produce an annual savings of X SEK.	The scrap cost of insufficient and to much solder paste volume that gets applied in the Screen Printer 1 & 2 processes costs the case company X SEK annually. In the SPI, these scrap reasons are found and they represent 75% of the total scrap cost at this station.
Goal Statement	**Scope**
The goals of this project are: 1) to understand what parameters contributes to the process variations and high scrap levels, and what parameters affects the yield variation the most 2) to provide recommendations for implementing countermeasures	The scope of the project includes SMT lines 1-7, only looking at the scrap levels at the SPI. The SPI control that the volume from the previous process, Screen Printer 1 & 2, is according to specification. Hence, the process boundaries are from the Screen Printer 1 & 2 to the SPI station. Also, the project will only conduct the Define, Measure, and Analyze phase of the DMAIC model, due to time constrains.
Project Plan	**Project Team**
Define 2020-01-27 – 2020-02-28 Measure 2020-03-02 – 2020-04-03 Analyze 2020-04-13 – 2020-05-15	Case company supervisor A – Project owner Johannes Ramberg – Black Belt & master thesis student Fredrik Hansson – Black Belt & master thesis student Case company supervisor B – Green Belt & team member (quality contact) Process engineer A – SPI process owner Process engineer B – SMT process expert Process engineer C – SMT process expert Production Quality Department – Process owner

Figure 5.6: Project Charter.

5.2 Measure phase

The Measure phase consisted of two subphases. The first subphase will be to do an MSA to test the measurability of the process to assess if the data are trustworthy. The second subphase will include the collection and measuring of the main data. A p-diagram and a data collection plan were first made followed by the MSA.

5.2.1 P-diagram

A p-diagram for the screen printing process have been conducted, see Figure 5.7 below. Further, as viewed in Figure 5.7, the applied solder paste volume, in the screen printing process, is the Y as a function of the input signals, control factors and noise factors. To clarify, the output of the Y is measured by the following station in the SMT production process, the SPI-station. The output of the Y may either lead to an intended output of Y or an unintended one, where the intended output is a PCB with the right amount of solder paste and the unintended a PCB without the right amount of solder paste.

The input signals were divided into mass and energy categories. The mass category consists of material process input, such as the amount of solder paste and the PCB board. The energy category consists of the support plate vacuum, temperature, humidity and squeegee angle, speed and pressure.

The control factors, which the engineering team can change, are stencil thickness and type, aperture area, interval of the cleaning of the stencil and the separation distance and speed.

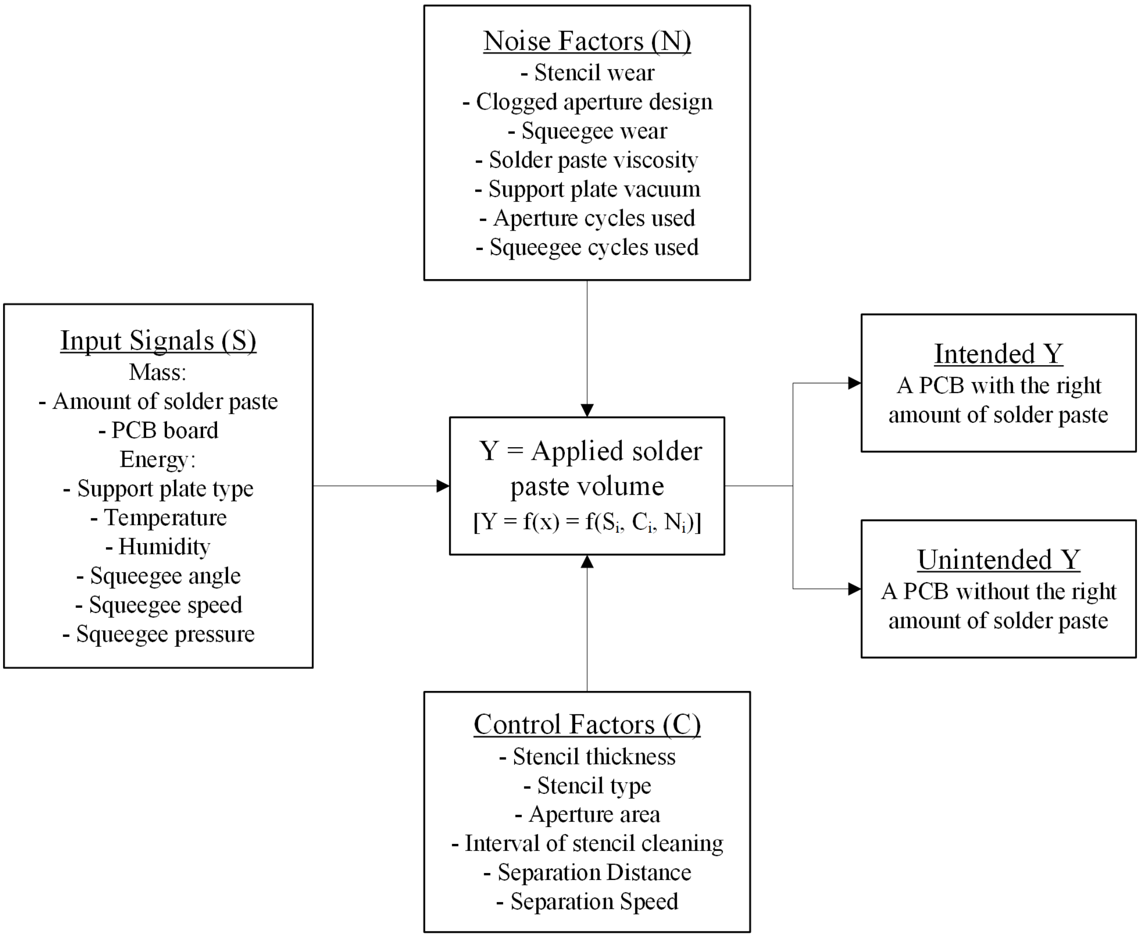


Figure 5.7: P-diagram of the screen printing process.

The project team identified six noise factors. The first two are regarding the wear of the stencil and clogged aperture design. Through interviews and the literature review, it was found that the condition of the stencil and aperture design may affect the outcome of the process. Regarding clogged aperture design, the more clogged, the less area for the solder paste to be applied on the PCB, which in turn will lower the amount of possible solder paste to be applied. The other noise factors found were the squeegee wear, solder paste viscosity, support plate vacuum and how many cycles the aperture design and squeegee are used.

5.2.2 Data Collection Plan

The data collection plan explains how the project team aims to gather the necessary data to be able to analyse the problem area and find possible root causes to the screen printing process variations.

The project team already had a lot of data gathered by the process owner of the SPI-stations. This data have been collected daily, since November 2019, from each of the six SMT lines during production stops. Due to the big amount of data, only the last 50 produced boards were chosen to be extracted from each SMT line since one produced board takes approximately 2 000-12 000 rows in Excel, depending on which product that is being produced. Hence, 50 boards correspond to about 100 000-600 000 rows in Excel. In this data from the SPI-stations, the project team had access to data of the screen printing process' output regarding the following:

- product name (part number)
- date and time produced
- if the applied solder paste is good or bad (SPI result)
- the applied solder paste volume in an absolute value and in percent
- solder paste height in an absolute value and in percent
- solder paste area in an absolute value and in
- solder paste offset in X and Y coordinates in an absolute value and in percent
- the size of the solder paste area in X and Y directions

When looking at the p-diagram, see Figure 5.7 above, the collected data corresponds to the intended and unintended Y. Further, the collected data is missing data regarding the input signals and control factors from the screen printing process. Hence, to be able to analyse the Y, the project team needed to collect data regarding the input signals and control factors of the screen printing process, such as:

- amount of solder paste added to the screen printing process
- solder paste viscosity
- support plate vacuum
- temperature
- humidity
- squeegee angle
- squeegee speed
- squeegee pressure
- stencil thickness & type
- aperture area
- the number of cycles the aperture have been used
- the number of cycles the squeegee have been used
- interval of stencil cleaning
- separation distance
- separation speed

5.2.3 MSA

The MSA was conducted on all of the seven SMT-lines when production had made a temporary stop. Twenty boards plus two reference boards were selected for the test. The boards all contained eight panels each. All of the boards initially had no solder paste applied so the 20 boards got a certain amount of solder paste put on them and the two reference boards were left empty. The boards were then marked with numbers ranging from 1-20 and the reference boards as number 26 and 27. Thereafter, the boards were run in the SPI-stations at SMT1-SMT7. The results from the test and the amount of data were however unsatisfactory due to the inability to analyse the results properly due to the extensiveness of the data. Additional tests were therefore conducted with only three boards and three replicates on all SMT-lines. These three boards were randomly chosen out of the 20 boards with a randomizing tool and resulted in boards 4, 12 and 17. The data were then put into Minitab for analysis.

5.2.4 ANOVA Gauge R&R (crossed) Settings

The settings for all of the MSA studies were the same, as shown in Figure 5.8 and 5.9 below.

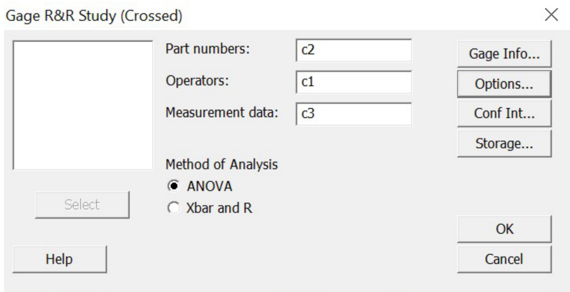


Figure 5.8: Gauge R&R (crossed) settings page one of two.

In Figure 5.8 above, part number, operator and measurement data was identified as column c2, c1, respective c3. The PadID was chosen as the part number since each PadID is unique with an individual applied solder paste volume. The PadID could be between 1 to 11 792, 1 474 PadIDs for each panel. To clarify, all panels are identical to one another, except that they have different amounts of solder paste applied on them. Hence, PadID 1 on panel 1 is corresponding to PadID 1 475 on panel 2. Therefore, PadID 1 and 1 475 can not be compared to each other in this MSA-study. The SPI-station was chosen as the operator since the SPI-station is the one that measures the output. The SPI column could be between 1-7, where 1 corresponds to the SPI-station on SMT line 1, 2 to the SPI-station on SMT line 2 and so on. Lastly, the volume in percentage was chosen as the measurement data since the amount of applied solder paste volume is the one metric, Y, the project team are investigating.

In Figure 5.9 below, the study variation, process tolerance, historical standard deviation and alpha to remove interaction term was identified. The study variation was set to six by default in Minitab. Further, it was set to six by the project team since it is needed to include 99.73% of the measurements. The process tolerance was set to 120 since the upper specification limit is 180 and the lower specification limit is 40. The project team chose to run the MSA-study without a historical standard deviation since the project team wanted to investigate the standard deviation in the test. Lastly, the alpha level of 0.05 was chosen as an appropriate level due to benchmarking with similar studies and due to that it is most commonly used.

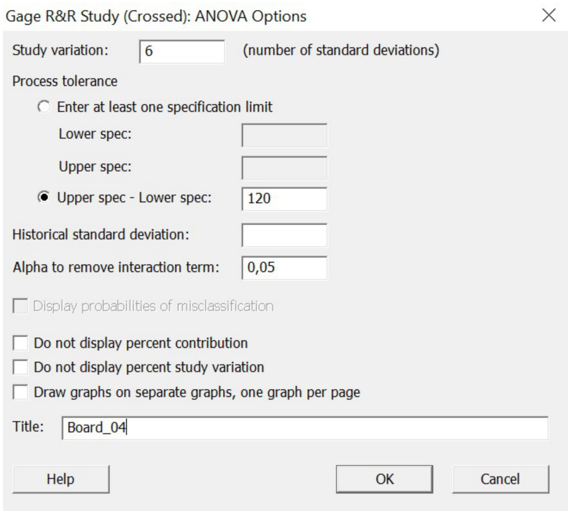


Figure 5.9: Gauge R&R (crossed) settings page two of two.

5.2.5 ANOVA Gauge R&R (crossed) results

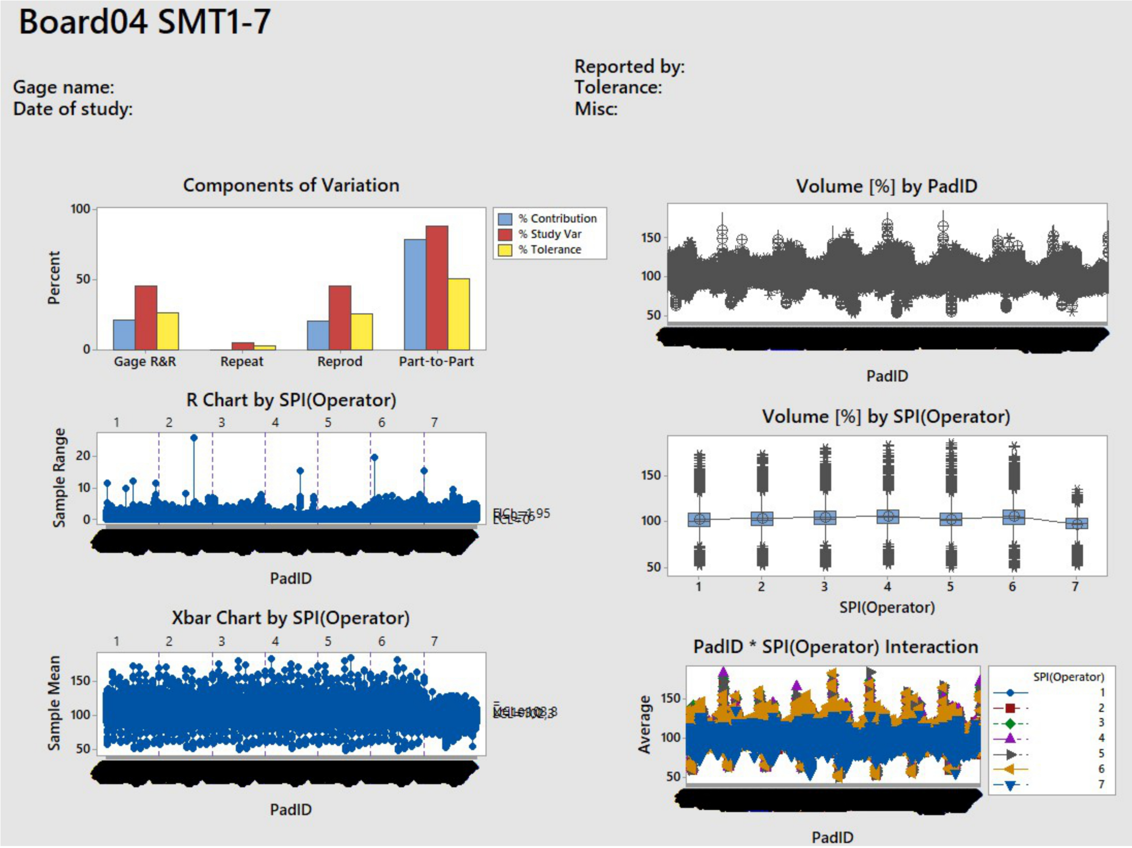
In this section, the results from the two different sub-studies will be presented. The first sub-study involves SMT1-7, and the second sub-study involves SMT1-6. SMT7 were removed from the second sub-study due to the deviations in results from the other SMT-lines.

5.2.5.1 SMT1-7

The test was made on the SPI-stations at all SMT-lines and included three boards and three replicates. The results from each board follow, see Figures 5.10 and Appendix B.

The total Gauge R&R percent contribution on board 4, 12 and 17 are 21.05%, 23.20% and 21.54% respectively which were all above the limit of 9% and thus unacceptable. Repeatability was very good for each board with repeatability percent contribution of 0.24%, 0.21, and 0.19% respectively which shows that repeatability was not a major contributor of variation. Reproducibility was not as good as the repeatability, with reproducibility percent contribution of 20.81%, 22.99% and 21.35% respectively which shows that reproducibility was the main contributor of variation between repeatability and reproducibility. Part-to-part variation as percent contribution were 78.95%, 76.80% and 78.46% respectively which were fairly low. The percent study variation was 45.88%, 48.17% and 46.41% respectively which were a result of the relatively large standard deviation.

The number of distinct categories for the different boards were all 2 which means that the system only can divide the groups of parts into high and low.



(a)

Results for: Board04

Gage R&R Study - ANOVA Method

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
PadID	11791	26423325	2241	38,70	0,000
SPI(Operator	6	1741770	290295	5012,82	0,000
PadID * SPI(Operator	70746	4096937	58	182,61	0,000
Repeatability	165088	52352	0		
Total	247631	32314384			

α to remove interaction term = 0,05

(b)

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	27,719	21,05
Repeatability	0,317	0,24
Reproducibility	27,402	20,81
SPI(Operator	8,204	6,23
SPI(Operator*PadID	19,198	14,58
Part-To-Part	103,955	78,95
Total Variation	131,675	100,00

Process tolerance = 120

Source	StdDev (SD)	Study Var (6 × SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	5,2649	31,5895	45,88	26,32
Repeatability	0,5631	3,3788	4,91	2,82
Reproducibility	5,2347	31,4082	45,62	26,17
SPI(Operator	2,8643	17,1859	24,96	14,32
SPI(Operator*PadID	4,3815	26,2892	38,18	21,91
Part-To-Part	10,1959	61,1751	88,85	50,98
Total Variation	11,4750	68,8498	100,00	57,37

Number of Distinct Categories = 2

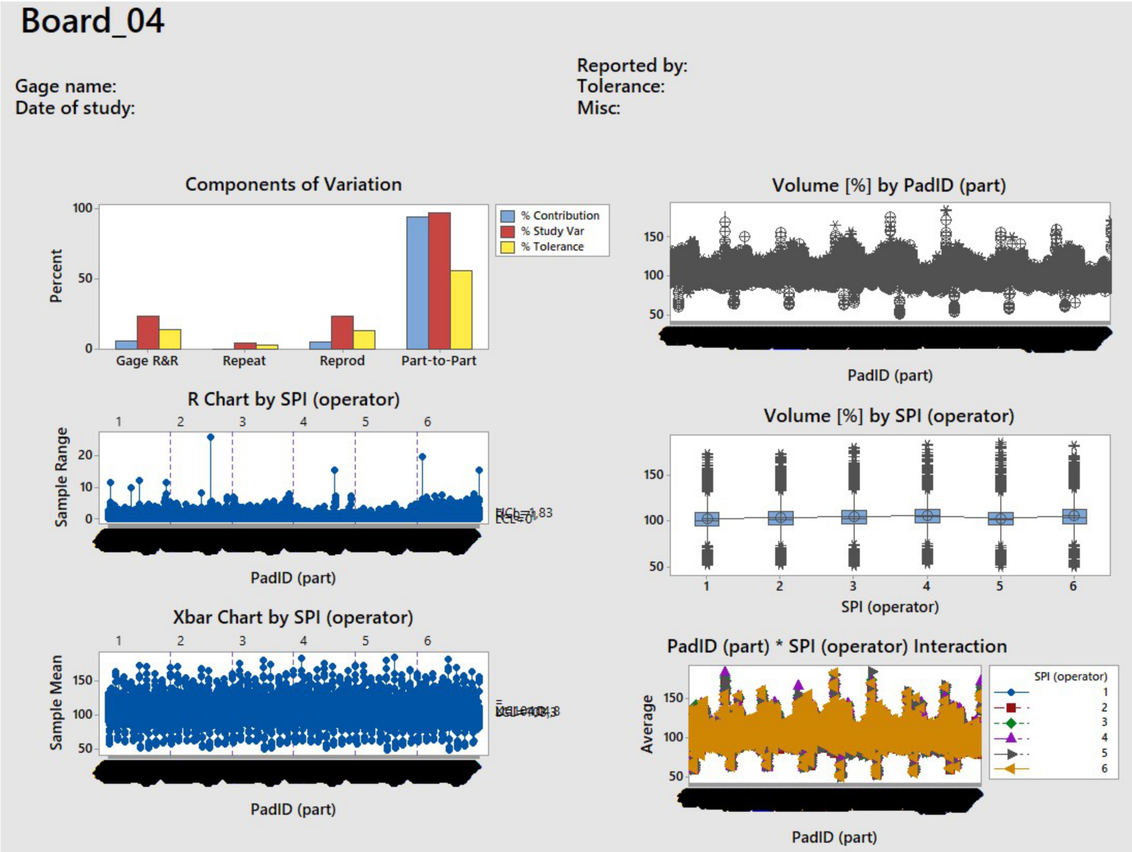
(c)

Figure 5.10: Results in Minitab from the Gauge R&R SMT1-7 of board 4. (a) Gauge R&R graphs. (b) Two-Way ANOVA results. (c) Gauge R&R results.

The R chart and X-bar chart contains 82 544, 11 792*7, rows and were therefore difficult to interpret visually. The range, however, were relatively small with the majority of data points between 0 and 5 and a few extremes reaching above 20 on board 4 and 12. Looking at the X-bar chart, the centre lines and the control limits were very narrow due to the low repeatability contribution to variation and the data points were out-of-control which shows high part-to-part variation.

5.2.6 SMT1-6

The test was made on the SPI-stations at the SMT1-6 lines and included three boards and three replicates. The results from each board follow, see Figures 5.11 and Appendix C.



(a)

Results for: Board 04

Gage R&R Study - ANOVA Method

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
PadID (part)	11791	26596123	2255,6	131,81	0,000
SPI (operato	5	276827	55365,5	3235,35	0,000
PadID (part) * SPI (operato	58955	1008878	17,1	60,69	0,000
Repeatability	141504	39900	0,3		
Total	212255	27921728			

α to remove interaction term = 0,05

(b)

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	7,457	5,66
Repeatability	0,282	0,21
Reproducibility	7,175	5,44
SPI (operato	1,565	1,19
SPI (operato*PadID (part)	5,610	4,26
Part-To-Part	124,362	94,34
Total Variation	131,819	100,00

Process tolerance = 120

Source	StdDev (SD)	Study Var (6 × SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	2,7307	16,3843	23,78	13,65
Repeatability	0,5310	3,1861	4,63	2,66
Reproducibility	2,6786	16,0715	23,33	13,39
SPI (operato	1,2508	7,5050	10,89	6,25
SPI (operato*PadID (part)	2,3686	14,2116	20,63	11,84
Part-To-Part	11,1518	66,9106	97,13	55,76
Total Variation	11,4812	68,8874	100,00	57,41

Number of Distinct Categories = 5

(c)

Figure 5.11: Results in Minitab from the Gauge R&R SMT1-6 of board 4. (a) Gauge R&R graphs. (b) Two-Way ANOVA results. (c) Gauge R&R results.

The total Gauge R&R percent contribution on board 4, 12 and 17 are 5.66%, 5.64% and 5.06% respectively which were all under the limit of 9% and thus acceptable according to circumstances. Repeatability was very good for each board with repeatability percent contribution are all 0.21% which shows that repeatability was not a major contributor of variation. Reproducibility was not as good as the repeatability, but still good, with reproducibility percent contribution of 5.44%, 5.43% and 4.84% respectively which shows that reproducibility was the main contributor of variation between repeatability and reproducibility. Part-to-part variation as percent contribution were 94.34%, 94.36% and 94.94% respectively which were high. The percent study variation was 23.78%, 23.76% and 22.48% respectively, all below 30% and thus acceptable during circumstances.

The number of distinct categories for the different boards were 5, 5 and 6 respectively which means that the system can distinguish between different groups of parts.

The range, however, were relatively small with the majority of data points between 0 and 5 and a few extremes reaching above 20 on board 4 and 12. Looking at the X-bar chart, the centre lines and the control limits were very narrow due to the low repeatability contribution to variation and the data points are out-of-control which shows high part-to-part variation.

5.3 Analyze phase

To be able to find the root causes of the variation of applied solder paste volume, several analysing tools were used. First, a grouping of variables will be presented to explain the variables used in the analysis. Secondly, the code and dataset that the analysis tools are based upon will be briefly explained. Thirdly, the results of the multiple regression, random forest and boosting analysis will be presented. Lastly, correlation graphs of important predictors will be presented.

5.3.1 Grouping of Variables

The analysis matrix that was used for the analysis consisted of 18 million rows with 35 columns. The matrix was based on the Excel-data files the project team received from the process engineers, which consisted of the resulting applied solder paste deposits, measured by the SPI-station, on the PCB board. This matrix was extended with information regarding the process values from the screen printing process, which were gathered through an internal database. Further, for the analysis in Rstudio to process faster, all Excel-files was saved as CSV-files. The 35 matrix columns were divided into three main groups: non-used columns, continuous data and attribute data, see Figure 5.12.

The purpose of the unused columns Job, PCBID, Barcode, PadID and Stencil Article Number was to use them as reference number to extract screen printing process values from the case company's internal database. The Stencil Stepped variable was not used since all values were the same for all rows. The Separation Distance and Speed were not used in the analysis since there were not data available.

The continuous data was further divided into two subgroups: data from the SPI-stations and data from the screen printing process. Volume(µm), Height(µm), Area(µm2) and Area(%) are all mathematically related to the Y of the project, Volume(%). Therefore, these columns of the matrix were excluded from the analysis. SizeX(mm) and SizeY(mm) are the intended width and length of the solder paste deposits. OffsetX and OffsetY are the deviations between the intended coordinates of the solder paste deposits and the actual outcome. PosX and PosY are the coordinates of the applied solder paste deposits. The variables Stencil Thickness, Temperature, Humidity, Squeegee Pressure, Squeegee Speed, Clean Rate 1 and Clean Rate 2 were all gathered through an internal database regarding process input values. The columns Aperture Wall Area and Aperture Area were both constructed by the project team. These two columns were excluded from the analysis since they are both mathematically related to the SizeX, SizeY and the Stencil Thickness variables.

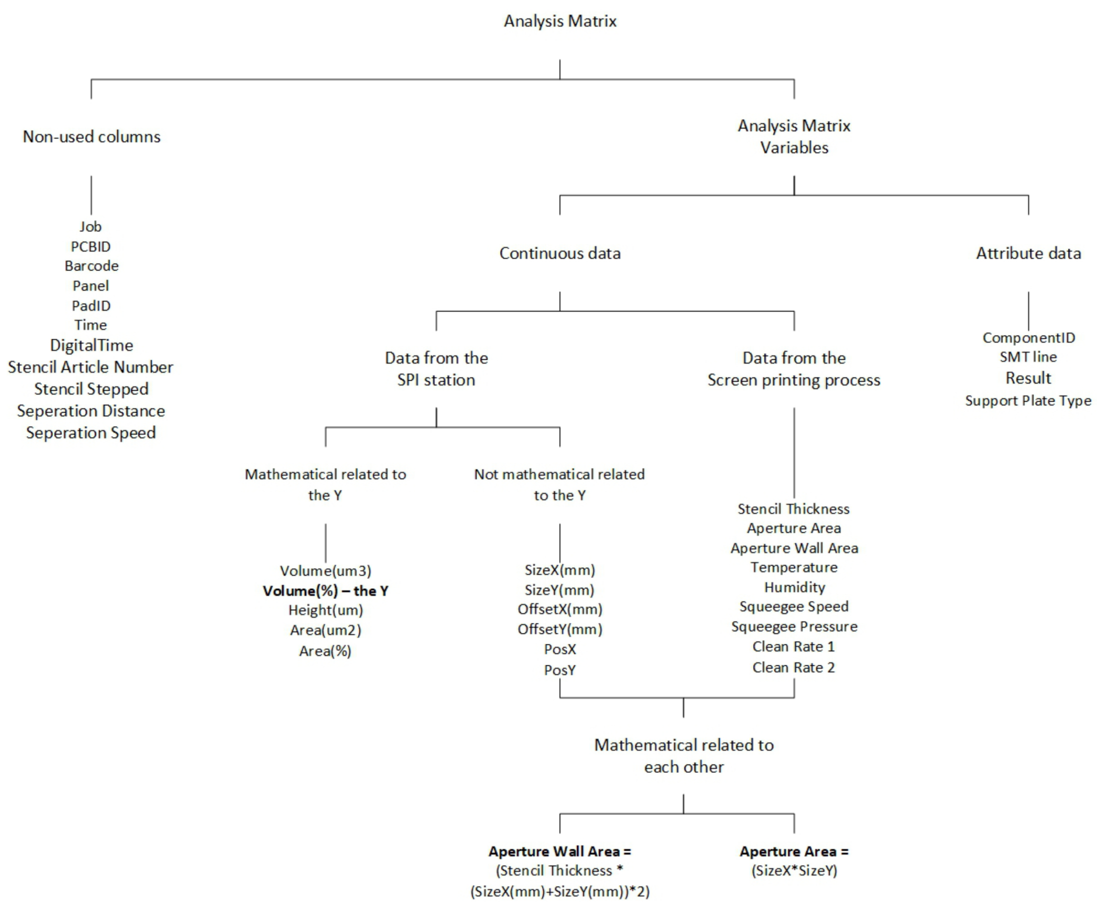


Figure 5.12: Grouping of Analysing Variables.

The attribute data of the analysis consisted of the four columns ComponentID, SMT line, Result and Support Plate Type. Depending on product type, the ComponentID column could consist of up to 100 different components on one single product. Therefore, the ComponentID consisted of more than 100 components in the analysis. In the project, data from six SMT lines were gathered and investigated; SMT lines 1, 2, 3, 4, 5 and 6. The Result column was the SPI-stations final judgement of each applied solder paste deposit. The Result could be classified as either “Good”, “Warning” or “Error”. Lastly, the Support Plate Type column was an article number of which support plate that was used during the screen printing process.

5.3.2 R code

The codes (see Appendix D) were written in the program R and were designed together with the supervisor from Chalmers University of Technology, Hendry Raharjo. The coding was done step by step and new codes were continuously added during supervision meetings. The major problem encountered was the lack of RAM-memory on the computers used and thus only a few CSV-files of data were used at the same time. Through switching to computers with extended RAM-memory of 32 GB, analysis of all of the data files could be conducted at the same time regarding different plots. When it came to random forests and boosting, the RAM-memory was still not enough so instead 60 out of 79 files were chosen to conduct random forests and boosting. The files excluded where files from product groups that contained many files as well as files from SMT2 and SMT5 which contained more, highly similar files than the other SMT-lines.

5.3.3 Multiple Regression

The multiple regression method explains the data matrix by relating the input variables to the Y-variable of the project, the Volume(%). The mean squared error (MSE) of the multiple regression analysis was 54,1. This means that the precision of the model is ± 7,36, taking the square root of the MSE.

5.3.4 Random Forest Analysis

Random forest is, as written in 4.2.3.5, a type of bagging where many decision trees are put besides one another and then used to make better predictions. In Figure 5.13, the result of the random forest can be seen and it presents a list of the 10 most important predictors for solder paste volume.

The number of trees were chosen to be 100 and can be related to Figure 5.14 where the MSE stabilized at around 100 trees and an MSE of 41,55. This means that the prediction error of the model is ± 6,45, taking the square root of the MSE.

	IncNodePurity <dbl>	vars <chr>
1	7458599	TemperatureC
2	5117465	SizeY
3	4497460	SizeX
4	4377234	SqueegeeSpeed
5	3673211	CleanRate1
6	3196268	HumidityPct
7	3180588	SqueegeePressure
8	2920481	S0009
9	2230570	X4
10	2221639	X5

Figure 5.13: Results of the Random Forest analysis. A table of the top 10 most important factors.

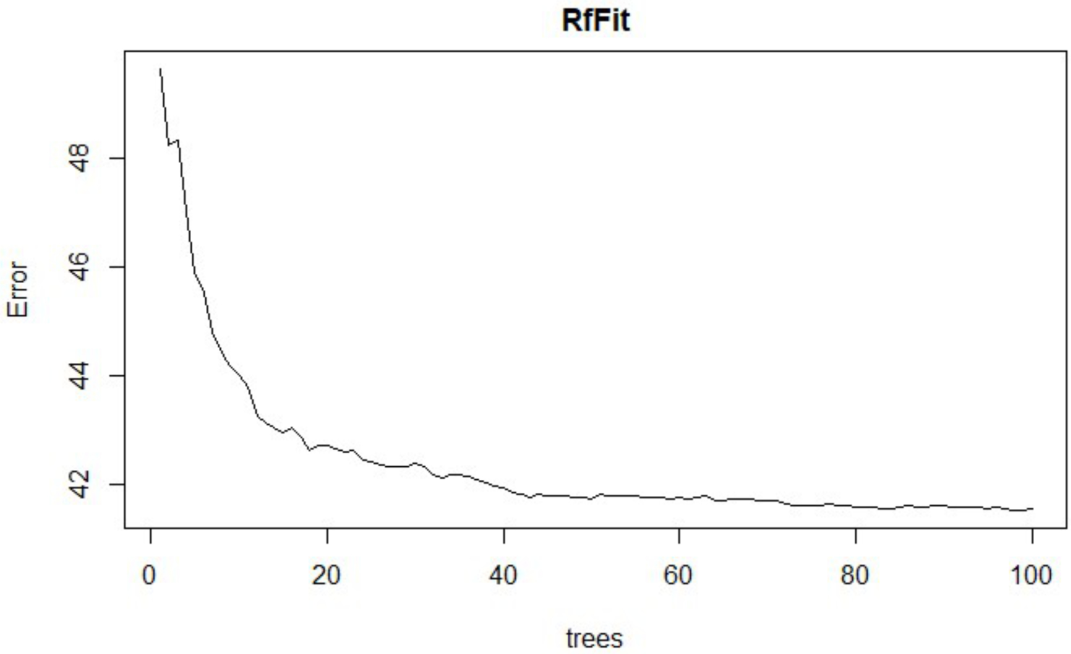


Figure 5.14: MSE of the Random Forest analysis.

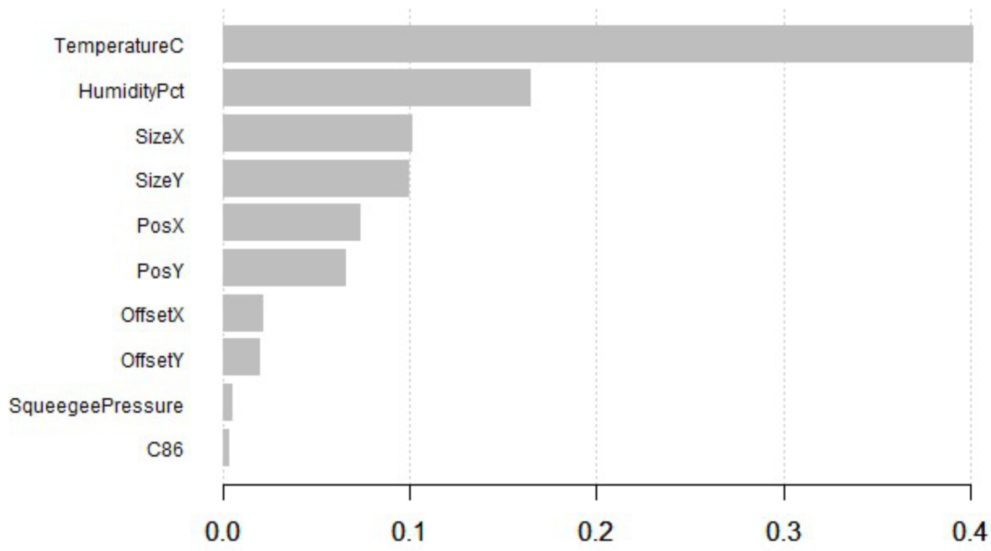
5.3.5 Boosting Analysis

After the random forest analysis, a boosting analysis was also made to investigate if a better prediction model could be achieved. To do this, a boosting analysis were first made and then followed by a tuning of the model. The boosting analysis followed by the tuning of the model resulted in an MSE of 24,45. This means that the prediction error of the model is ± 4,95.

As can be seen below in Figure 5.15a and 5.15b, the top predictors has now changed a bit as well as their impact as predictors when compared to the random forest analysis. Temperature is still the top predictor while Humidity has taken second place followed by SizeX and SizeY. Position and Offset, both X and Y are new predictors but especially Offset contributes to a minor extent. Clean Rate 1 has shrunk dramatically as well as Squeegee Speed and Squeegee Pressure.

Feature <chr>	Gain <dbl>	Cover <dbl>	Frequency <dbl>	Importance <dbl>
TemperatureC	0.402019515	0.009128862	0.00513620768	0.402019515
HumidityPct	0.165042388	0.004323932	0.00202554669	0.165042388
SizeX	0.101183102	0.028505576	0.03277045182	0.101183102
SizeY	0.099837291	0.028593166	0.01388946302	0.099837291
PosX	0.073605119	0.332992409	0.15141994957	0.073605119
PosY	0.065563712	0.279999070	0.11867016659	0.065563712
OffsetX	0.021309848	0.054585742	0.33050514654	0.021309848
OffsetY	0.019451220	0.055210916	0.31605762474	0.019451220
SqueegeePressure	0.005027046	0.002713269	0.00101277335	0.005027046
C86	0.003547916	0.001328831	0.00008267538	0.003547916

(a)



(b)

Figure 5.15: Results of the boosting analysis. (a) A table of the top 10 most important factors. (b) The top 10 importance factors visualized in a bar chart.

5.3.6 Correlation Graphs

Boosting showed a better prediction model compared to random forest. Therefore, correlation analysis between Volume(%) and the top predictors based on the boosting analysis will be presented in the following subsections. SizeX and SizeY has been changed to ApertureArea and ApertureWallArea and also includes ratio between them. Further, graphs depicting the ratio between SizeX and StencilThickness will also be presented. Lastly, graphs regarding PosX and PosY will be presented. The y-axis ranges from 0 up to just above 300 in Volume(%) solder paste and the x-axis varies depending on the chosen variable.

5.3.6.1 VolumePct vs. TemperatureC

In Figure 5.16 and 5.17 below, Volume(%) is on the y-axis and Temperature(°C) is on the x-axis. The x-axis describes the different temperatures in production when the products where manufactured. The figure also includes all SMT-lines and most of the data points are concentrated to the middle of the vertical bars and fewer data points on the extreme ends. There are more variation the lower the temperature with less variation the higher the temperature, especially when looking at the lower end of the bars with the data points that has a lower percentage of solder paste volume. In Figure 5.17, the temperature has been divided based on the different SMT-lines.

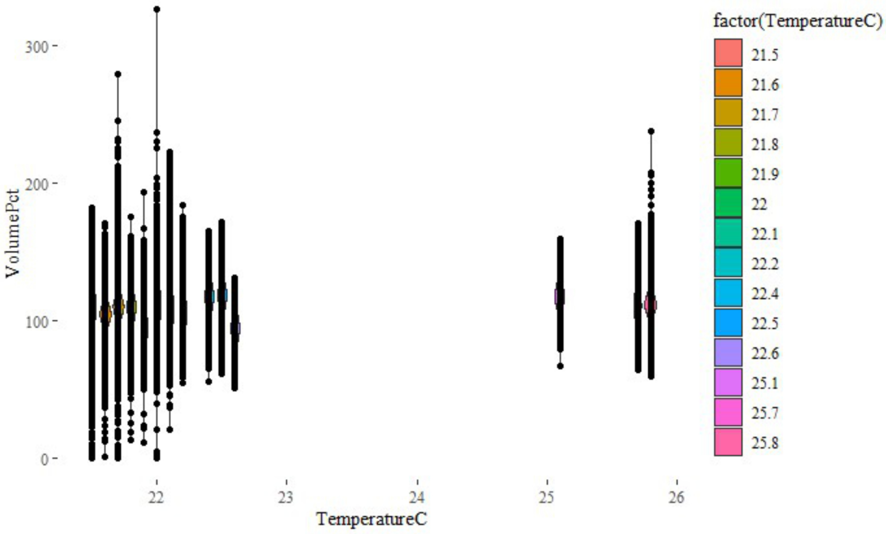


Figure 5.16: Scatter and violin plot showing the correlation between the Volume(%) and Temperature(°C).

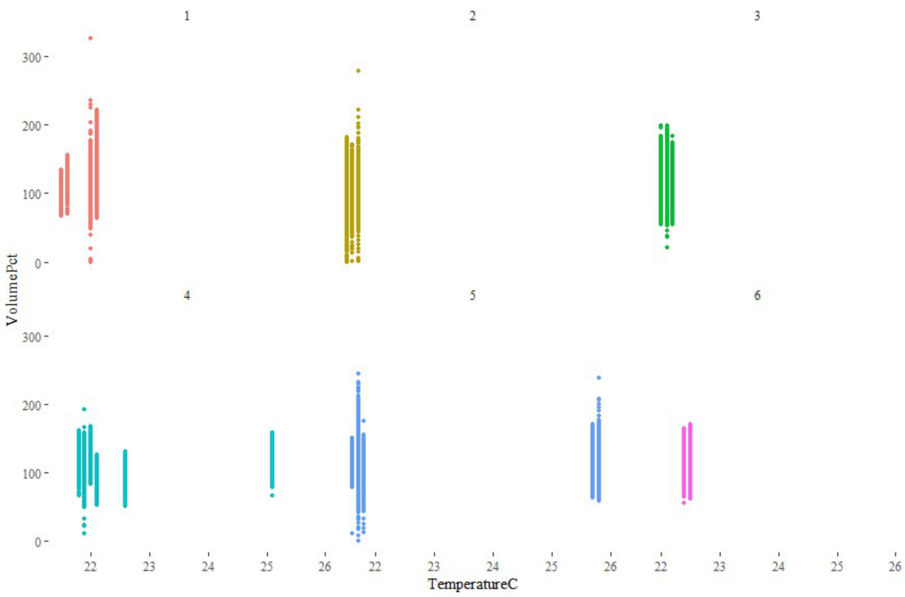


Figure 5.17: Scatter plot showing the correlation between the Volume(%) and Temperature(°C) for each SMT line.

5.3.6.2 VolumePct vs. HumidityPct

In Figure 5.18 and 5.19 below, Volume(%) is on the y-axis and Humidity(%) is on the x-axis. In Figure 5.18, with increased humidity, the variation also increases. Most of the variation tends to be equally distributed or in the lower parts of the bars. In Figure 5.19, SMT2 and SMT5 experiences the biggest variation while a humidity percentage of 36% experiences much variation and can be seen both in Figure 5.18 and 5.19.

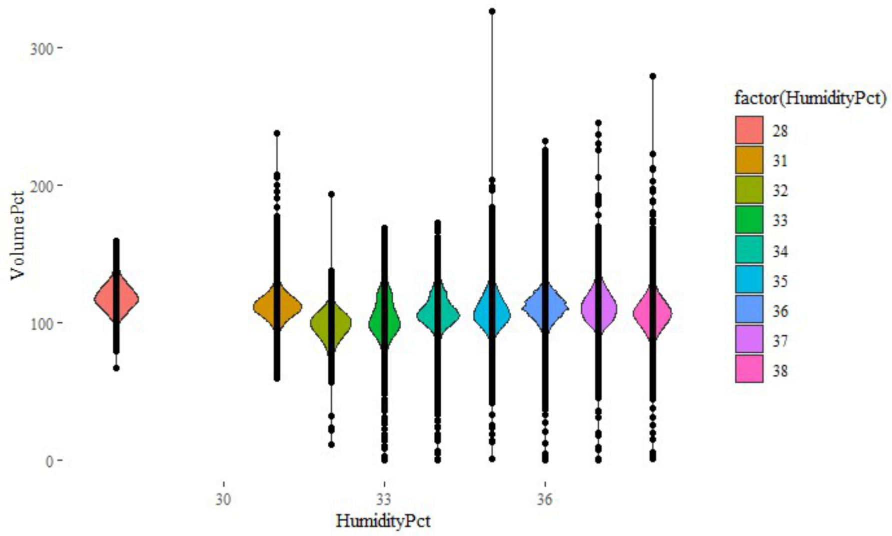


Figure 5.18: Scatter and violin plot showing the correlation between the Volume(%) and Humidity(%).

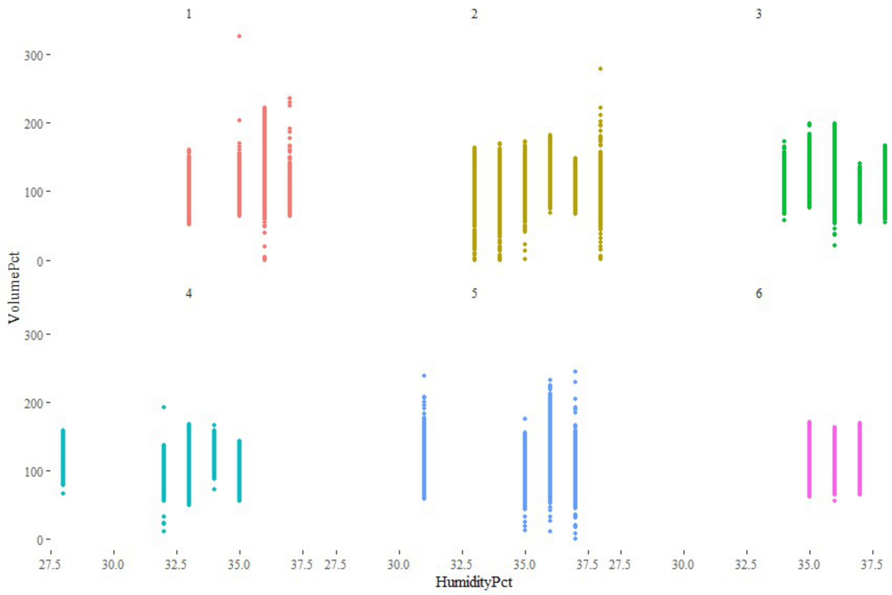


Figure 5.19: Scatter plot showing the correlation between the Volume(%) and Humidity(%) for each SMT line.

5.3.6.3 VolumePct vs. Aperture Area

In Figure 5.20 and 5.21 below, Volume(%) is on the y-axis and Aperture Area(mm2) is on the x-axis. The bigger the aperture area, the less variation, both upper and lower. Most of the data points are concentrated to the smaller aperture areas since most aperture sizes are small and therefore most of the variation is also present here. Figure 5.21 follows the same trend as 5.20 and variation is gradually getting less and less the bigger the aperture area.

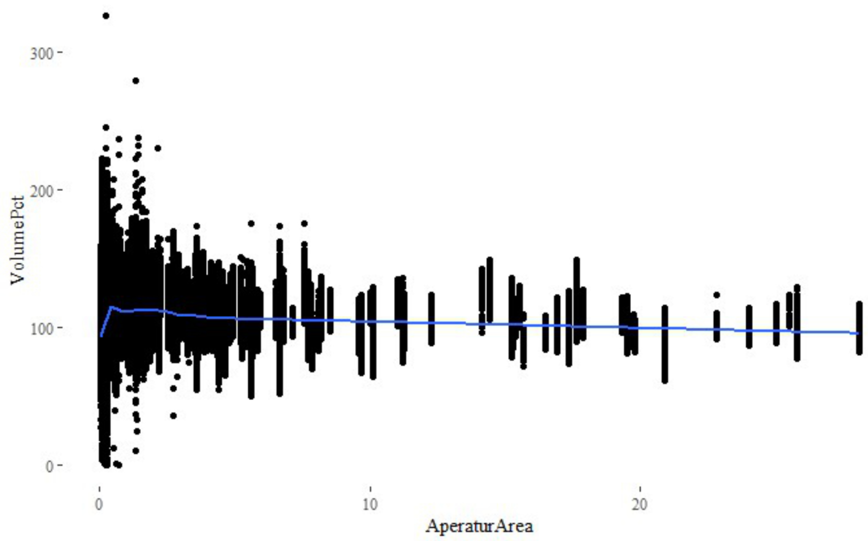


Figure 5.20: Scatter plot with a blue trend line showing the correlation between the Volume(%) and Aperture Area (mm2).

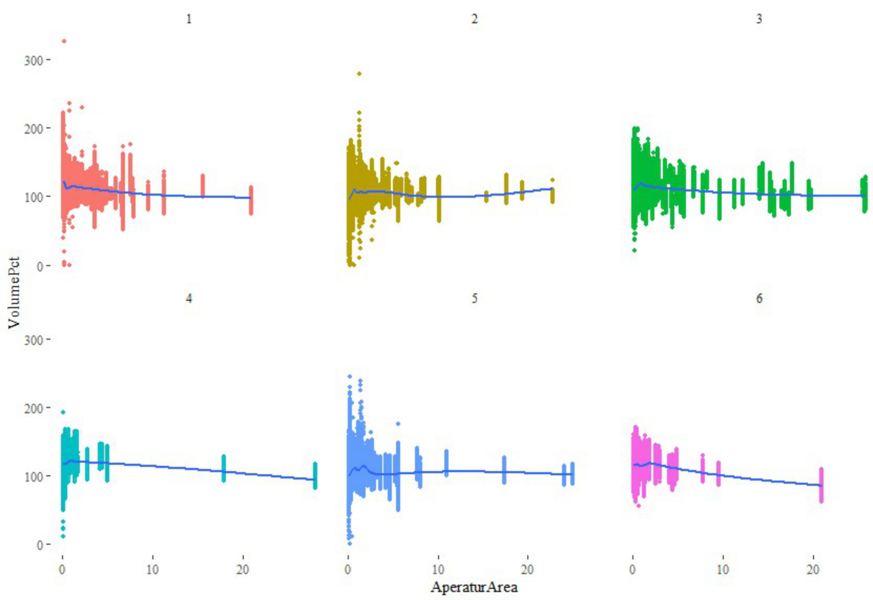


Figure 5.21: Scatter plot with a blue trend line showing the correlation between the Volume(%) and Aperture Area (mm2) for each SMT line.

5.3.6.4 VolumePct vs. Aperture Wall Area

In Figure 5.22 and 5.23 below, Volume(%) is on the y-axis and Aperture Wall Area(mm2) is on the x-axis. The aperture wall area is based on the circumference of the aperture and the thickness of the stencil. In Figure 5.22, the variation is high when the aperture wall area is small but decreases when the area becomes bigger. Both upper and lower variation from the baseline of 100 % solder paste volume decreases as well as the trend line.

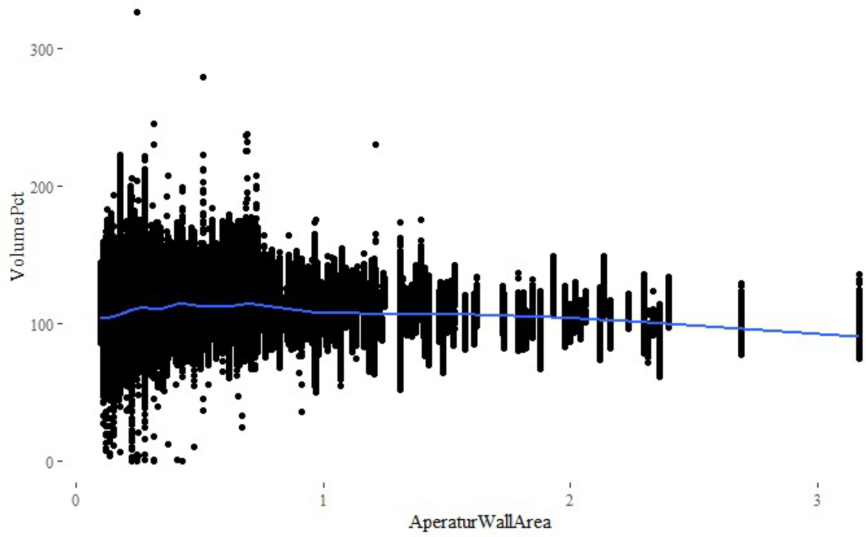


Figure 5.22: Scatter plot with a blue trend line showing the correlation between the Volume(%) and Aperture Wall Area (mm2).

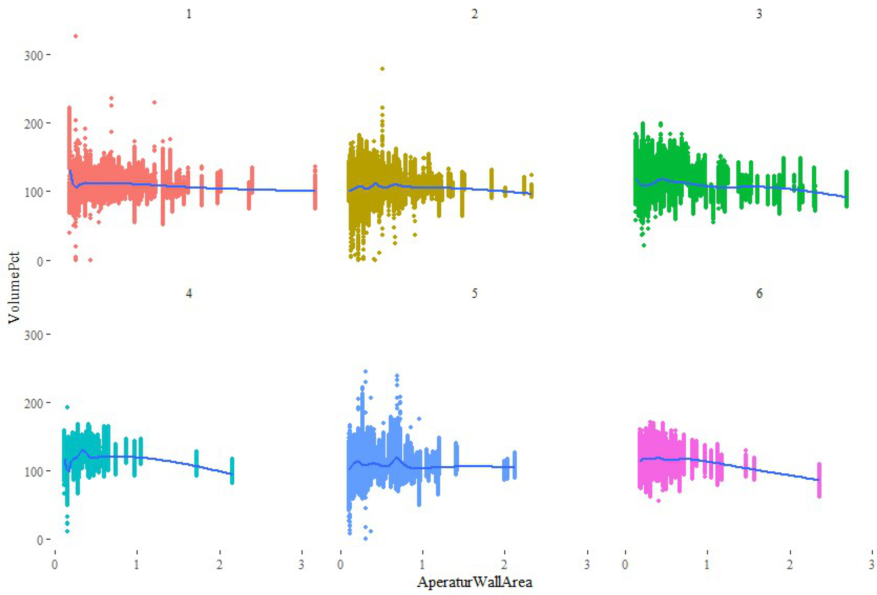


Figure 5.23: Scatter plot with a blue trend line showing the correlation between the Volume(%) and Aperture Wall Area (mm2) for each SMT line.

5.3.6.5 VolumePct vs. Area Ratio

In Figure 5.24 and 5.25 below, Volume(%) is on the y-axis and ratio of Aperture Area and Aperture Wall Area is on the x-axis. The bigger the ratio, the bigger the difference between aperture area and aperture wall area. Typically, when the aperture area is big and the aperture wall area is small, the ratio is big and vice versa. Both Figure 5.24 and 5.25 has the same ratio trends where volume percentage variation of solder paste decreases the bigger the ratio.

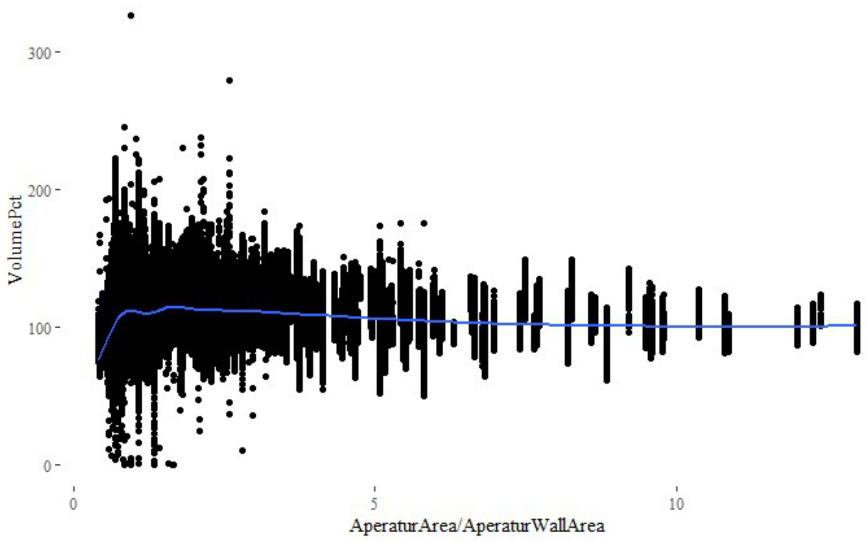


Figure 5.24: Scatter plot with a blue trend line showing the correlation between the Volume(%) and Aperture Area/Aperture Wall Area.

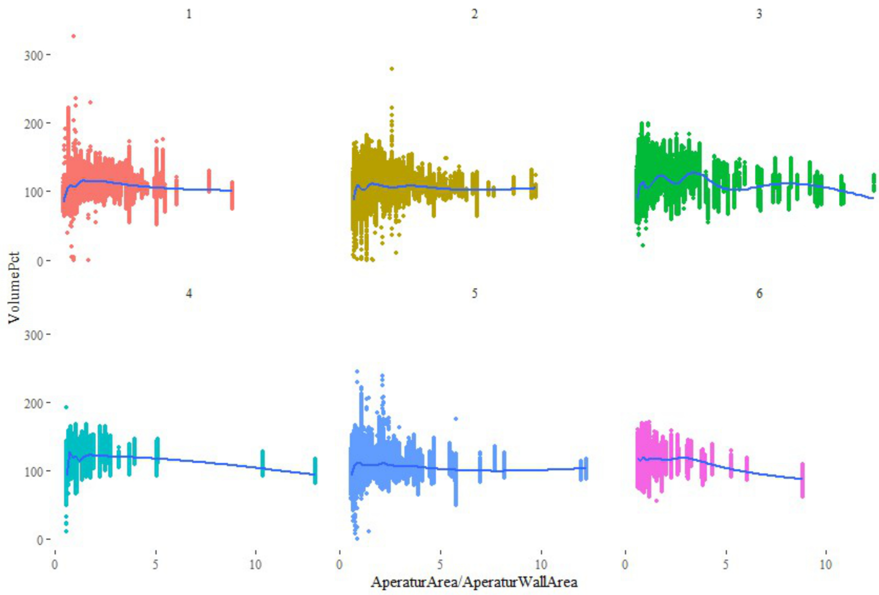


Figure 5.25: Scatter plot with a blue trend line showing the correlation between the Volume(%) and Aperture Area/Aperture Wall Area for each SMT line.

5.3.6.6 VolumePct vs. Aspect Ratio

In Figure 5.26 and 5.27 below, Volume(%) is on the y-axis and ratio of SizeX and Stencil Thickness is on the x-axis. The ratio, also called aspect ratio, takes the width of the aperture and divides it with the thickness of the stencil. In Figure 5.26 and 5.27, the trends are that the bigger the ratio, the lower the solder paste volume percentage variation.

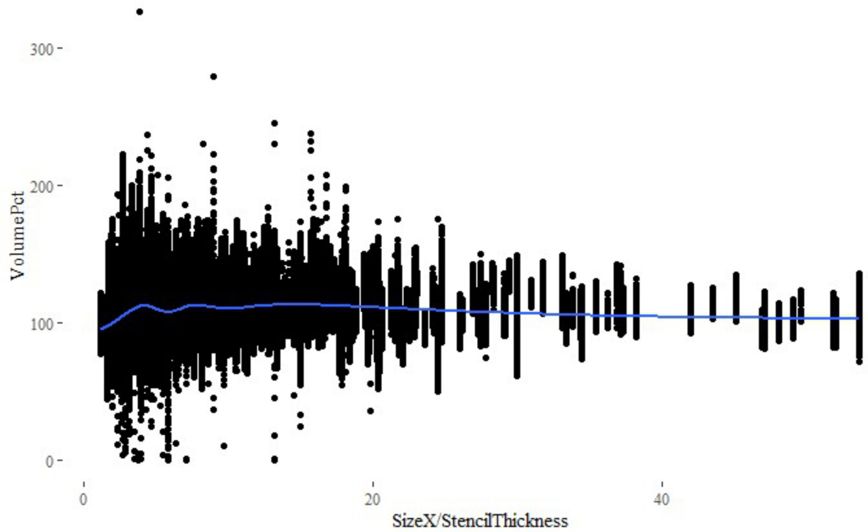


Figure 5.26: Scatter plot with a blue trend line showing the correlation between the Volume(%) and SizeX/Stencil Thickness.

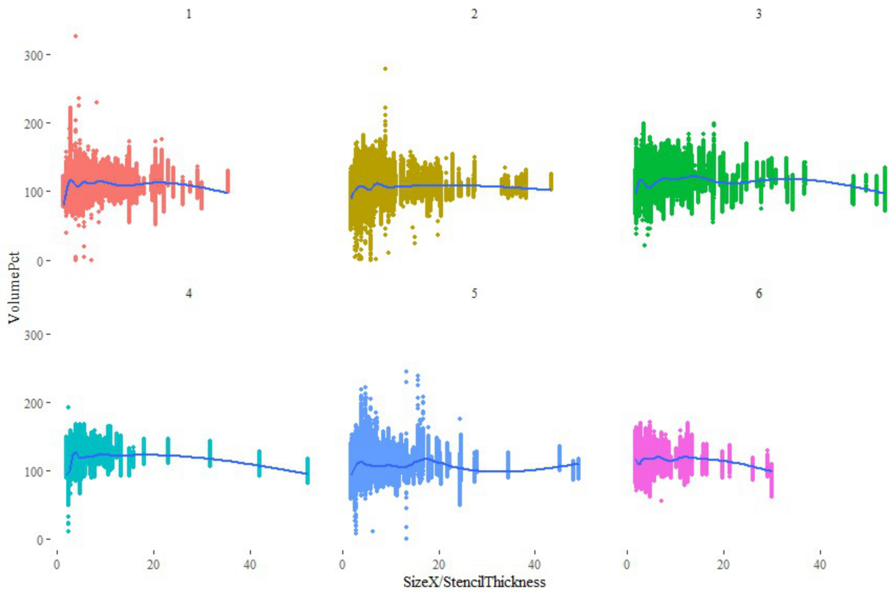


Figure 5.27: Scatter plot with a blue trend line showing the correlation between the Volume(%) and SizeX/Stencil Thickness for each SMT line.

5.3.6.7 VolumePct vs. PosX

In Figure 5.28 and 5.29 below, Volume(%) is on the y-axis and PosX is on the x-axis. In Figure 5.28 four tops reaching volume percentage above 200% and four minor tops with volume percentage reaching just below 200%. Looking at Figure 5.29, six out of these eight tops are related to SMT1, while the other two are shared by SMT2, SMT3. Also, SMT5 is showing a group of points peaking at the right end of the chart. SMT4 have several distinct downward groups of data points.

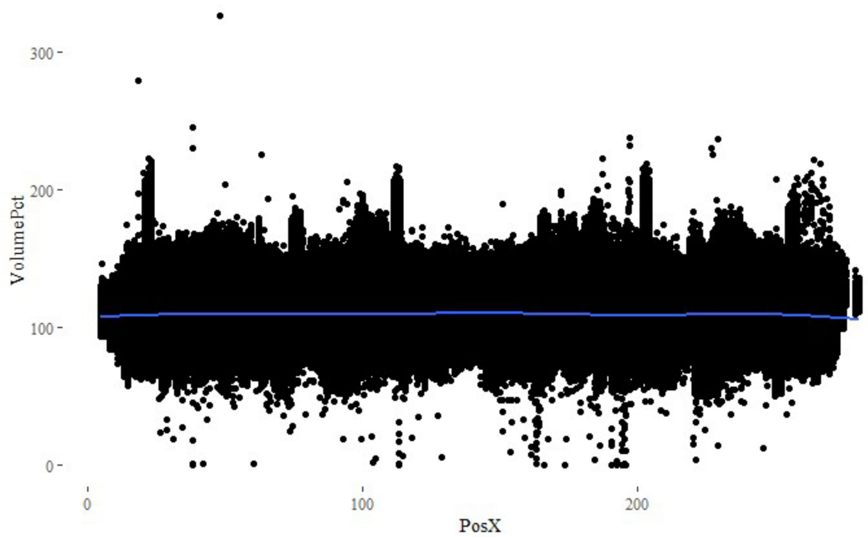


Figure 5.28: Scatter plot with a blue trend line showing the correlation between the Volume(%) and PosX.

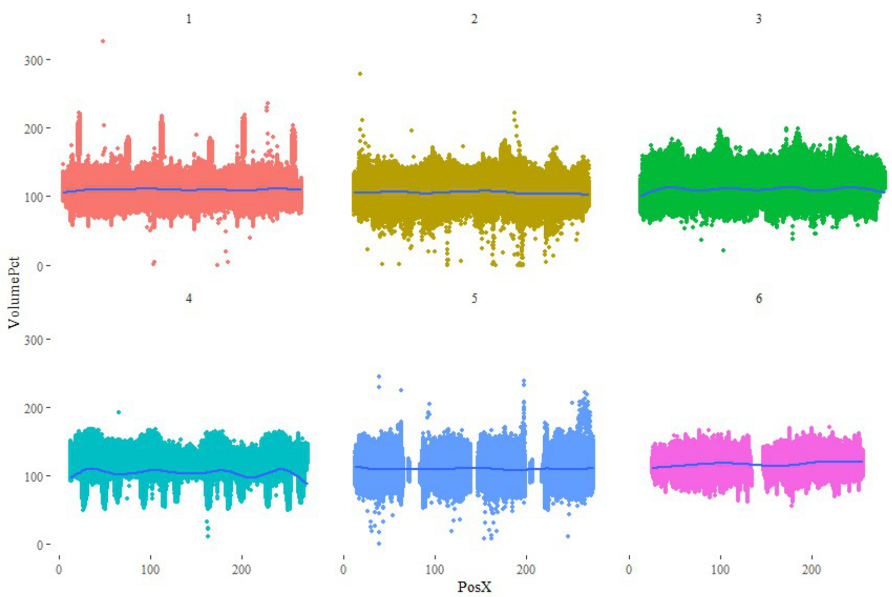


Figure 5.29: Scatter plot with a blue trend line showing the correlation between the Volume(%) and PosX for each SMT line.

5.3.6.8 VolumePct vs. PosY

In Figure 5.30 and 5.31 below, Volume(%) is on the y-axis and PosY is on the x-axis. In Figure 5.30 there are three distinct tops and two minor ones. When looking at Figure 5.31, these can be related to SMT1 for the three tops and SMT3 for the two minor tops. However, when looking at SMT4 in Figure 5.31, there are also three distinct downward groups of data points as well as a lot of downward variation at SMT2 and SMT5.

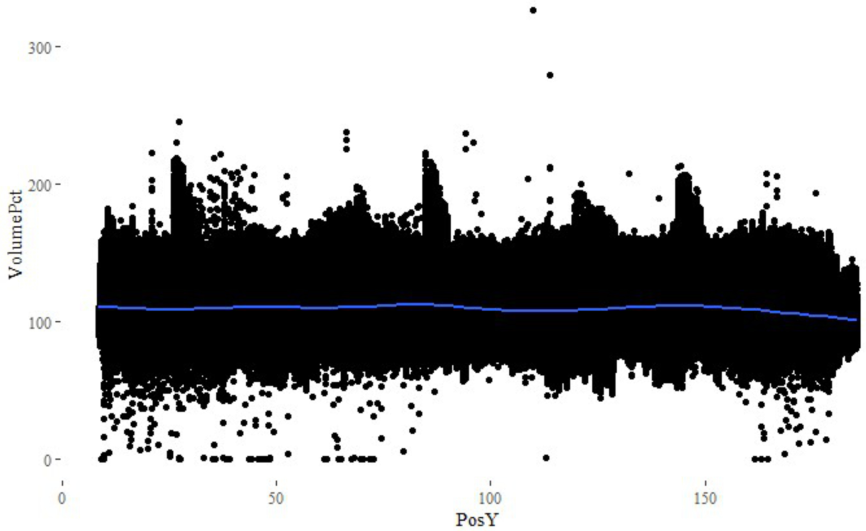


Figure 5.30: Scatter plot with a blue trend line showing the correlation between the Volume(%) and PosY.

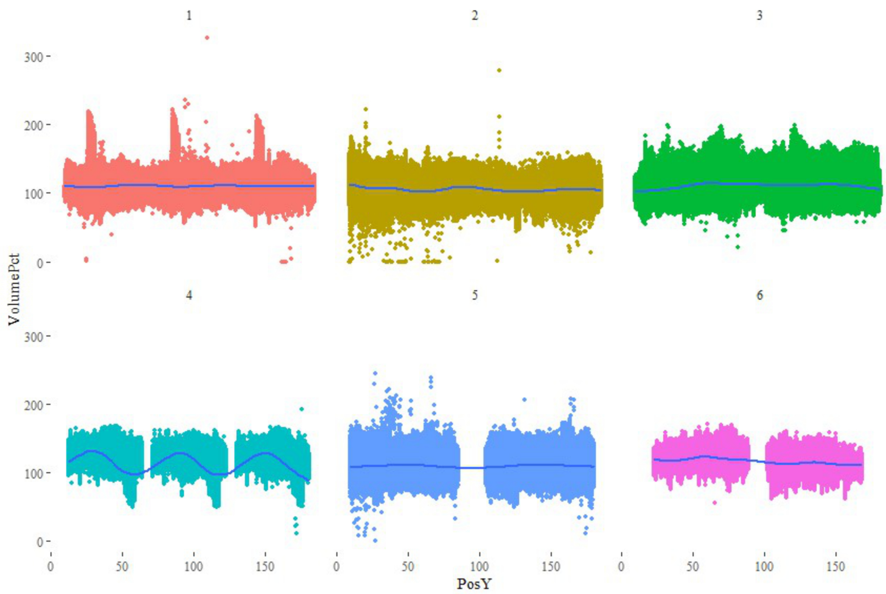


Figure 5.31: Scatter plot with a blue trend line showing the correlation between the Volume(%) and PosY for each SMT line.

6 Discussion

In this chapter, both qualitative and quantitative data will be discussed in combination with the literature to answer the research questions. Thereafter, first a discussion of the relevance of the master thesis will be presented followed by future research proposals.

6.1 Six Sigma Methodology

The master thesis was designed as a Six Sigma project for the case company in order to investigate if the methodology is something that the case company should consider to continue to work with and expand. Six Sigma methodologies include many useful tools and methods and have been used throughout the master thesis. Most importantly, Six Sigma is a business strategy that focuses on improving the understanding of the customers' requirements, the company's productivity and financial performance (Kwak & Anbari, 2006). To approach the issues stated by the case company, Six Sigma is a useful methodology to structure and break down the problem into sub-problems and find the root causes to solve the main problem (Schroeder et al., 2008). Through the use of both qualitative and quantitative tools and techniques, the case company can achieve support and compare data from different sources. Quantitative data are useful since it contains hard facts but gets even stronger with qualitative data that can support, direct or divert related to the quantitative data. Due to the vast amount of data that the case company creates in their processes and the high tolerances and standards on their products, they should consider putting more effort in minimizing variation, lower defects and do continuous process evaluations. For this purpose, Six Sigma is a highly useful methodology (Anthony, 2011). Apart from this, as described by Walters (2005), Six Sigma not only explains what to do, but also how to do it.

For the case company, the Define phase is especially important since this sets the entire setting of the project, why it should be done, if it should be done and address what the problem really is (Schroeder et al., 2008). It has to do with spending the time on the right things, not only doing the things right. Six Sigma methodologies are therefore important to help the case company investigate if a project should be done in an economic or strategic point of view. By spending time in the Define phase, greater understanding of the whole problem area is achieved and will also increase the knowledge of the bigger picture. It will help to scope the project down and to address the important things and do the right things to minimize waste, both time and resources.

6.2 Define Phase

The first step in the Six Sigma methodology is to understand the VoC. Interviews with process owners were conducted to get an overview of the SMT production and process as a whole. The same process engineers also guided the project team through the SMT production and explained the process in detail, see 5.1.1.1. The process engineers were chosen to be the interviewee and SMT tour guides since they are the ones in the case company that work the closest to the SMT-production every day. The combination between interviews and observations was important to get a holistic view of the production and the associated process and as recommended by Waller et al. (2016), interviews followed observations to make sure that the observations were not biased and so that the project teams experience from the observations would be included when conducting the interviews. After the interviews, more observations followed which enabled the project team to ask follow-up questions that was somewhat unclear or hard to understand from the interviews.

To further understand where the issue behind the process variation lies within the SMT-lines, the project team contacted the quality engineers in the production area team. The graphs retrieved from the production area team were of great importance to state the VoC. In the CTQ flowdown, see Figure 5.2, the biggest part of the scrap levels comes from the placing machine process after the SPI-station. However, in discussions with the stakeholders of the master thesis it was decided to focus on the screen printing process and the SPI-stations since they strongly believed that by reducing defects as early as possible, this will most likely result in less defects downstream. Identifying defects early also results in less economic damage since not many value-adding activities have been performed at that stage in the process. By translating the graphs into a CTQ flowdown, it was made clear that insufficient applied solder paste is the major reason for scrapping at the SPI-station, as the process engineers also stated during the interviews, see 5.1.2. Moreover, the CTQ flowdown was a great tool to visualize the decomposition of the problem area into more accessible parts. Thereafter, the AIM question was formulated according to the VoC.

The AIM activity is especially good for bringing people from different departments together to discuss their qualitative knowledge with each other and conduct brainstorming to address the stated question of the AIM (Alänge, 2009). Through the performed AIM, it was discovered that many people in the organisation had clues and ideas of the potential root causes of the problem area. During the AIM, consensus was achieved that the main issue in the SPI-stations is the variations of applied solder paste volume, but not what was causing it. Even though the AIM team lacked consensus and agreement of the cause of the issue, the AIM team concluded that the stencil design and wear may result in variations in the SPI-station. The people in the AIM team came up with both similar and different ideas to the process variations and it was clear that they had not been discussing the issue in a similar kind of setting before. The result, see 5.1.4, ended up in five different main groups: maintenance and service, operator education, cleanliness, screen printing inputs and SPI related issues. However, one issue with the AIM was that some of the input post-its were rather solutions to the main question than possible root causes. This led to every group being connected to each other, which was symbolized through the many arrows in Figure 5.3, too many arrows according to the standard by Alänge (2009). The first time a group perform a method like AIM, it may be a little weaker. But by doing it more frequently, a group of people will be quicker conducting the activity while still achieving a better outcome (Alänge, 2009). Further, the qualitative outcome from the AIM was used to guide the project team further but also to compare the qualitative data with the quantitative data. In future projects, the AIM activity could help the case company bringing different people together to brainstorm ideas but also to create consensus within teams and projects of what the issue actually is.

Effective scoping is a great tool to use early on in a project since conducting it the project will realize what the problem really is and what it is not (Hammersberg, 2019b). The baseline for effective scoping came from the output of the VoC. The remaining empty spaces of the effective scoping were gathered through interviews, screen printing process observations and discussions with process engineers. The effective scoping, see Figure 5.4, provided the project team with deeper understanding of the process and what affects the Y, but most importantly it helped scoping the project down. Further, the scoping of the project was crucial for the project not getting too comprehensive. The output of the effective scoping was further used to define the scope of the master thesis in the project charter.

As Carleton (2018) discussed, the project charter should include information of the projects business case, project problem, project goal statements, project scope, project plan and team members in maximum one page. This means that, by presenting only one page, everyone in the organisation would quickly be able to understand what the project is about, why it is needed, the aim with the project, who is going to work with the project and when it is planned to happen. In other words, the project chapter is a short summary of the whole Define phase. The purpose with the project charter was to use it to present a summary of the Define phase to the stakeholders, see Figure 5.6. Since not all stakeholders had followed every step of the master thesis, this tool was great to briefly explain the project for both those who have followed the progress of the project and those who have not. Further, by summarizing the project scope in this manner, it could help the case company to make short project summaries but also to compare different projects to each other.

6.3 Measure Phase

During the Measure phase, a P-diagram, see Figure 5.7, was first constructed to visualize the process, what the input signals were, what factors that were noise factors and thus could not be controlled, control factors and the intended and unintended output. The P-diagram gave the project team a good overview of the process and the different factors affecting the process. It also gave indications to what quantitative data that was needed to be gathered for the Analyze phase. The data collection plan that followed is important since this needs to be planned and executed accordingly to get sufficient and good quality data. Due to the vast amounts of data produced, limited data samples of the last 50 produced boards were retrieved from the SMT-processes which resulted in data files containing around 100 000-600 000 rows each.

The problem was, however, that the process engineers themselves had to retrieve the data manually and did so on unscheduled times. The data had recently begun to be retrieved from the process and procedures on how to collect the data were not in place. Many of the factors that the project team proposed as important, both noise factors and control factors were not measured or had to be retrieved from an old website with low functionality. During the course of the project, the project team discovered that many important factors needed according to Ashgar et al. (2019), Gopal et al. (2006) and Khader et al. (2017), were not measured or at least not stored or used properly, such as solder paste viscosity, squeegee and stencil wear, applied solder paste and solder paste left on the stencil after printing. This resulted in difficulties in accessing all the important data needed and also presents a bias in the project. There was also no proper storages for the data or a place that could provide accessibility to the data. Since the collection and usage of process data were quite new, it has not been worked with continuously and that is something that should be considered.

A more extensive data collection plan should be considered to be put into place, and procedures should be considered to be set on how to collect, store and access the data. The case company should also consider to work more continuously with the data and use it to their advantage to improve the SMT-processes. That is also why Six Sigma could be used when measuring data since it provides good methods and tools for doing so.

The MSA was an important part of the master thesis at the case company and regarded the validation and verification of the quantitative data. The case company knew in beforehand that repeatability was good at every SMT-line but were unsure regarding reproducibility between the SMT-lines, i.e. that the SPI-stations measured and returned the same data when measuring the same product. This was an important step to make sure that the data for the later analysis were trustworthy. The first MSA-test conducted on SMT-lines 1-7 produced unacceptable results and it was soon realised that SMT7 differed a lot from the rest, see Figure 5.10a and Appendix B. Therefore, it was decided to remove SMT7 from the MSA-study. SMT7 is a much newer line, not in production yet, with a newer SPI-machine and thus different settings. The focus instead turned to SMT1-6, the older lines and the lines from which the project data came from.

SMT1-6 produced better results and were considered as acceptable, see Figure 5.11a and Appendix C. The Gauge R&R were 5,66%, 5,64% and 5,06% for board 4, 12 and 17 respectively and are thus acceptable and well under the limit of 9%. Part-to-part variation as percent of contribution to the total variation is therefore high and lies around 94-95%, which means that the variation between the parts themselves accounts for almost all the variation. It is acceptable but not perfect and the ideal situation is a Gauge R&R below 1% (Minitab, 2020). However, this is difficult to achieve. It requires both time and resources, and will most likely not be worth it for the case company. Repeatability is good with 0,21% of the total variation while reproducibility accounts for the rest with 4,85% to 5,45% of total variation. The R chart is also showing relatively small range with the majority of the data points between 0 and 5 which is small considering the size of the data. There are a few extremes but that is acceptable due to the vast amounts of data points. The X-bar chart also shows high part-to-part variation with the centre lines and control limits narrowly placed which are caused by the low repeatability contribution to variation. Since the data points are out-of-control outside of the centre lines and control limits, this shows on high part-to-part variation which is good for the validity of the data. The conclusion from the MSA analysis is therefore that the data are valid, that the SMT-lines measure well internally and also between themselves. They can to a high degree measure the same data points and will satisfactory deliver valid and trustworthy data when retrieved. After part-to-part variation, reproducibility can be accounted for the majority of the variation, but that is totally natural since the SPI-machines are of different brands and age. To further enhance reproducibility, which the case company is about to do, changing the SPI-machines to newer versions might be a good idea due to the fact that the SPI-machine on SMT7 showed a result with less variation, lower range in the R chart and less variation in the sample mean in the X-bar chart. Another thing that is also important to mention is that there are many data points in the MSA-study and typically few data points are used in an MSA.

6.4 Analyze Phase

The master thesis used three analysis methods for analysing the quantitative data. They were multiple regression, random forest and boosting. The reason for using three analysing methods was to investigate which analysis that showed the best prediction of the Y, the applied amount of solder paste in volume percent, see 4.2.2.3. As stated by Carleton (2018), the multiple regression explains the dependent variable rather than predicting it. In this master thesis, multiple regression was used as a baseline for comparing the accuracy of the random forest and boosting prediction.

The MSE value was used to determine the accuracy of the data. Since the data included over 18 million rows, the number of data points was not an issue. Therefore, the importance lies on how well the models can predict new data. Further, the MSE value shows how precise the prediction from each analysis method is to predict the Y. When taking the square root of the MSE value, one gets a ± percentage value of how close the model is able to predict the Y. The lower MSE and square root MSE value, the better. As presented in the Analyze phase, see 5.3, the MSE of the multiple regression was 54,1 and the square root MSE of ±7,36, while the MSE of the random forest and boosting was 41,55 and 24,45 and the square root MSE of ±6,45 and ±4,95. This means that with boosting, the model can predict values with an error of on average ±4,95 which can be considered good.

There are, of course, other factors that was not included in the data that also affects the solder paste volume and the prediction of the model. To fine-tune it further, more data needs to be collected and included but considering the current data, it still has a considerably high validity and can predict further data, especially with the qualitative data as support. With this in mind, the results showed that both the random forest and boosting where more accurate in predicting the Y than the multiple regression. Since the boosting performed the best prediction model and the lowest MSE, its result was further used for correlation graphs.

Based on the results from the random forest and boosting analysis, the ten top predictors represent important factors from which the solder paste volume can be calculated from. The quantitative data are however not best on its own but should be complemented by literature and qualitative data. See Table 6.1 for a summary of the chapter 6.4, containing information regarding the literature, qualitative and quantitative data, triangulation and conclusion of each top predictor.

6.4.1 Area Ratio and Aspect Ratio

Two of the top predictors are the two most important ones, namely Size X and Size Y. From these two, aperture area, aperture wall area and two different ratios can be derived. A hypothesis was formed, based on the interviews, AIM and Huang et al. (2011), the smaller the area and aspect ratio, the greater the variation in solder paste volume is. This was based on the fact that the more aperture wall area, the more area that the solder paste could clog to. This is also highly related to the cleaning of the stencils, and it was mentioned in the interviews with the process engineers, see 5.1.2, that the apertures tend to be clogged after a few cycles in production. Due to the fact that the more clogged the apertures get, the less space for the solder paste to pass through the apertures and the less solder paste gets applied to the boards. This also connects to the main problem with solder paste volume that the most common scrapping cost related to the SPI-stations is insufficient solder paste volume. When looking at aperture area in Figure 5.20, most data points are close to zero which is natural since most aperture areas are small due to small component sizes. In this area, most variation is present.

With increased aperture area, the variation in solder paste volume also decreases. When looking at each SMT-line in Figure 5.21, the trend is the same with decreased variation the bigger the aperture area. The blue trend lines are also decreasing and moving towards the baseline of 100% solder paste volume. When looking at aperture wall area, it looks similar to aperture area, except the fact that the bigger the aperture area, the less variation in solder paste volume. The variation in low solder paste volume is decreasing quite rapidly with increased aperture wall area and the blue trendlines moves towards the solder paste volume baseline here as well. Due to this, the graphs have strengthened the hypothesis since the bigger the aperture area the less likelihood for clogged apertures and the more solder paste can get pasted on the boards.

From aperture area and aperture wall area, the area ratio can be derived with respect to solder paste volume. According to the IPC-7525 standard, the minimum area ratio is 0,66 but should preferably be greater (Huang et al., 2011). In Figure 5.24 and Figure 5.25, it is clear that most of the data points fall below 0,66 in area ratio and beyond that, variation quickly decreases. The bigger the area ratio, the less variation, and over an area ratio of approximately 3, there are no data points with less than around 50% solder paste volume. The same trends can be seen in Figure 5.25 where all SMT-lines are included that solder paste volume variation is decreasing and the blue trendlines are moving towards the baseline. The other ratio, called aspect ratio, is derived from Size X, that is the width in x-direction on the board and the stencil thickness. According to IPC-7525 standard, this ratio should be greater than 1,5 (Huang et al., 2011).

In Figure 5.26, most of the data points are greater than the aspect ratio of 1,5 which is good and is quite quickly decreasing with a increased ratio, that is when the width of the aperture is big and the stencil thickness is small. The same trends can once again be seen in Figure 5.27 including graphs on all SMT-lines. Both area and aspect ratio shows that the bigger the ratio, the smaller the variation in solder paste volume. More importantly, the data points with the lowest solder paste volume quickly vanishes with increased area and aspect ratio. This clearly strengthens the hypothesis that with a big aperture wall area, or more with a big area and aspect ratio, the less solder paste volume variation that also is concentrated towards the baseline of 100% solder paste volume.

According to Pan et al. (2004), Tsai (2011), Asghar et al.(2019), Farrell et al. (2013) and Khader et al. (2017), both aperture size and stencil thickness affects the solder paste deposit on the board and choosing and matching the right aperture size with the right stencil thickness is of the highest importance. The choice of stencil thickness should primarily be based on the aperture size and because of that, it is important that the thickness of the stencil can vary depending on the aperture (Gopal et al., 2006). In other words related to the case company, a good way of increasing the area and aspect ratio is to use stepped stencils with varying thickness depending on the aperture size. An interesting find in the data analysis related to this was that no stencils used in the retrieved data were stepped stencils when searched in the internal database. This indicates that the stencil thicknesses were the same across all the apertures on all stencils. According to the process engineers, different stencils are used for different products, and different products are produced at different SMT-lines.

Table 6.1: Summary of chapter 6.4.

Factor	Literature	Qualitative Data	Quantitative Data	Triangulation	Conclusion
Area & Aspect Ratios	Affects stencil printing quality. Huang et al. (2011)	Mentioned, but indirectly, as important in the qualitative data. (Process engineer B & C).	Size X third most important predictor in both boosting and random forest while Size Y was second in random forest and fourth in boosting.	All data types supports these factors as important.	The most important factors to investigate.
Cleaning Intervals	Affects the stencil design by preventing it from clogging. Lao & Yeung (1997)	Mentioned as an important factor to reduce variations. (Process engineer A, B & C)	Clean rate 1 fifth place in random forest and not present in boosting. Clean rate 2 not present at all.	Discrepancy between quantitative data and the two other data types.	Important factors that should be investigated further.
Temperature & Humidity	Affects overall quality and solder paste viscosity. Li et al. (2008), Lau & Yeung (1997), Gospal et al. (2006), Khader et al. (2017).	Only mentioned that a climate control system controls the temperature and humidity. (Process engineer B & C).	Temperature top one predictor in both boosting and random forest analysis. Humidity sixth most important in random forest and second top predictor in boosting. However inconclusive in the correlation graphs.	Literature and qualitative data supports these factors but inconclusive according to quantitative data.	Important factors that should be further investigated through DoE to find optimal intervals.
PositionX & PositionY	Not mentioned in the literature.	Mentioned as probably not important. (Process engineer B & C)	Not top predictor in the random forest analysis but fifth and sixth place in the boosting analysis.	Partial agreement on that these factors are not important.	Investigate if the solder mask might affect the the SPI measuring.
Squeegee Pressure & Speed	Important to find the optimal settings for the squeegee characteristics. Tong et al. (2004), Asghar et al. (2019), Mannan et al. (1994), Lao & Yeung (1997), Khader et al. (2017)	Not mentioned in the qualitative data.	Squeegee Pressure seventh place in random forest and ninth in boosting. Squeegee Speed fourth place in random forest and not present in boosting.	Partial agreement. Literature says important but qualitative and quantitative data contradicts.	Investigate further through DoE.
OffsetX & OffsetY	Not mentioned in the literature.	Mentioned as probably not important. (Process engineer B &C)	Not top predictor in the random forest analysis but seventh and eighth place in the boosting analysis.	Full agreement on that these factors are not important.	Not an important key factor and no discrepancy between data types.
Solder Paste Characteristics	Different solder paste types for different purposes. Viscosity is important. Asghar et al. (2019), Gopal et al. (2006)	Many different solder paste types are used at the case company. (Process engineer B &C)	Uncollected data.	Agreement between literature and qualitative data. No quantitative data however.	Collect data and conduct further analysis.

6.4.2 Cleaning Rate

Related to Aspect and Area ratio is the cleaning of the stencils, in the data referred to as cleaning rate 1 and 2. Cleaning rate 1 is a more basic cleaning while cleaning rate 2 is a more extensive cleaning of the stencil. Both process engineers and Lau & Yeung (1997) mentioned cleaning rate as an important factor. Through experiments conducted by Lau & Yeung (1997), they concluded that an interval of five consecutive printings between cleaning were the most optimal interval. This is in turn based on the solder paste type, the area and aspect ratio so the recommendation would be once again to do DoE with factorial experiments to investigate the optimal cleaning rate during operations.

The result from the random forest showed that cleaning rate 1 were among the top five predictors but when the boosting analysis was done, this showed that neither cleaning rate 1 nor cleaning rate 2 was among the top ten predictors for solder paste volume. Due to the importance of cleaning of the stencils from the qualitative data however, a further analysis was made. These were however inconclusive and could not be connected to solder paste volume in an efficient way. This also depended on the fact that there lacked data on between which boards cleaning 1 and cleaning 2 had taken place. This is further something that should be considered to be measured in the future by the case company.

6.4.3 Temperature and Humidity

As the data analysis proposed, temperature and humidity were considered as the two top predictors and they are important, mentioned by Li et al. (2008) and Lau & Yeung (1997). The interviewed process engineers did however not mention much about temperature and humidity except that it is controlled by a climate control system within the SMT-lines and that the temperature lies between 22±1ºC while the humidity level lies between 50±25%. Most of the data points related to temperature fell within the limit of 22±1ºC and showed a little bit less variation with increased temperature but that is hard to tell. On SMT4 and SMT5, temperatures were at around 25-26ºC but showed no more variation in the data than the rest. Regarding humidity it appears that with increased humidity, the variation in solder paste volume increases with the majority of the variation when the humidity lies on 36%. When looking at the different SMT-lines, most of the variation occurs on SMT-lines 2 and 5. Figure 5.16-5.19 appear to be rather inconclusive related to solder paste volume but these factors are still important, as indicated by both the data analysis and Li et al. (2008) and Lau & Yeung (1997).

The process temperature affects the solder paste to a high extent regarding viscosity and with changing temperatures or temperatures that are not fit to the solder paste, defects might occur (Gopal et al., 2006; Khader et al., 2017). Since the case company is about to change solder paste type, it is highly important to investigate at what temperatures and humidity that the new solder paste works the best. This might be done through a DoE with factorial experiments. Through experiments conducted by Li et al. (2008), it was concluded that the optimal operating temperature for solder printing was 20°C while Lau & Yeung (1997) concluded that when there occurred any waiting time between printing, the optimal temperature and humidity were 24°C and 50% respectively. Related to waiting time, the process engineers stated that there are delays sometimes between two printings were the first board often gets defects. With this said, temperature or humidity are most likely not the causes behind large variations in solder paste volume and most likely not the causes behind insufficient solder paste volume. However, related to the findings and Li et al. (2008) and Lau & Yeung (1997), a recommendation is to conduct DoE with factorial experiments to find the optimal operating temperature and humidity, especially when the type of solder paste is changed and when upgrading the SMT-lines with new machinery similar to SMT7. It might also be a good idea to investigate if temperature and humidity might be increased when delays occur between printings. It is also interesting that the range of the humidity is from 25% up to 75% and it might be wise to test the best humidity level related to the type of solder paste used in the process, especially since the data only included humidity's from 28% to 38% which is in the lower range of the tolerance.

6.4.4 Position X and Position Y

Based on the random forest analysis, Position X and Y appeared as the two next predictors after Size X and Y, however with low contribution as predictors. When looking at Figure 5.27-5.30, a few interesting trends could be seen. There were both groups of data points peaking upwards and downwards, especially on SMT-lines 1 and 4. The trendlines indicate however that it is few data points that belong to these groups and the majority of the data points are still around the baseline. It might be possible for the case company to proceed with these indications that on some products, especially on SMT-line 1 and 4, some positions might result in higher or lower amounts of solder paste volume, but it is also possible that it is random. However, according to the process engineers at 2nd line, SMT4 doesn't seem to have a random pattern but this is most likely due to the fact that the SPI-machine sometimes have difficulties knowing the reference points between the solder mask and the pads, especially around the edges.

6.4.5 Squeegee Pressure and Squeegee Speed

Squeegee pressure is important, however, together with squeegee speed, at least according to Tong et al. (2004), Asghar et al. (2019), Mannan et al. (1994), Lau & Yeung (1997) and Khader et al. (2017), but no conclusive connections to solder paste volume could be found. These are not parameters that the process engineers monitor either but is instead settings in the production according to IPC-standards.

6.4.6 Offset X and Offset Y

Offset X and Offset Y, these were predictors that had such a low contribution that these were left out of further data analysis, especially since the qualitative data does not support Offset X and Offset Y as contributors to low solder paste volume.

Potential offsets in x and y-direction is often also solved in the process steps after the SPI-station when the solder paste gets melted in the oven.

6.5 The Human Factor

Returning to cleaning and area and aspect ratio, these factors highly relate to each other and just as important the factors related to the apertures are, cleaning must be done frequently to avoid clogging of the apertures. But clogging also involves how much solder paste that is being put on the stencils before the printing. According to the process engineers, see 5.1.2, the amount of solder paste deposit on the stencils are being decided by the operators that can alter the settings depending on their experience and feeling about how much solder paste that is needed. That might very well result in too much solder paste being put on the stencils, or too little. This, in turn, might result in insufficient cleaning and the apertures will ultimately clog.

The case company has also recently changed the solder paste type, and Khader et al. (2017) stresses the importance of many characteristics related to solder paste that affects the printing quality, and by changing solder paste type, the operators must be aware that a different solder with all likeliness works differently from the previous one. The process engineers also mention that the operators often tend to take the easy way instead of following procedures, this due to lack of experience, different education and training as well as the fact that they can adjust many of the setting according to their own standards and experience. Many of the things that were highlighted in the AIM were also related to the training and education of the operators. These affect directly or indirectly many of the factors and the operators should be considered to receive more extensive training about the different process steps.

Another important human factor according to the process engineers is that operators are able to let boards through, even if there is a warning of a potential defect. Sometimes, operators get stuck in a flow of letting potential defects through. An alarm sounds when the solder paste volume gets lower than 60% and the machine stops if it gets below 40%. If an alarm sounds, the operator needs to do a manual check and the operator might instead just press pass, especially if he or she just has checked a few similar boards before that turned out “OK”. Tests have been conducted on SMT6 to remove the possibility to just let the boards through without a manual check and this are to be implemented on all SMT-lines in the near future. However, a recommendation to the case company is to remove as much potential of human errors as possible with taking away the possibility for the operators to adjust settings themselves.

6.6 Uncollected Data

Another important thing to address is uncollected data. Many of the factors that the project team came up with were not measured or just not stored in a database, see Table 6.2. Some factors were automatically recorded, some had to be retrieved and some were not measured at all. According to the process engineers, see 5.1.2.1, many factors are measured just because they can be physically measured and due to experience within the industry. However, no real manuals or information could prove why the factors are really measured, except for the IPC-A610 standard that sets high requirements on a few factors to measure. For example, see 5.1.2.2, at another department abroad, only three factors are being measured, and this due to customer requirements. At the case company, they check more than 12 factors, without any clue on why according to the process engineers. It seems that they are measured more on routine and experience, rather than to actually investigate what factors that matter. Many factors might be a remnant from older products produced but not necessary for new, updated products. One factor which data were not stored were the squeegee angle. According to the process engineers, the angle is the same on all squeegees.

Table 6.2: Summary of collected and uncollected data.

Measured/ Data Availability	*Easy*	*Difficult*
Yes	Humidity SizeX SizeY Squeegee angle Squeegee speed Squeegee pressure Stencil thickness Stencil type Support plate vacuum Temperature	None
No	Aperture cycles used Area Ratio* Aspect Ratio* Interval of stencil cleaning Solder paste type Squeegee cycles used	Cleanliness of stencil Clogged aperture design Solder paste viscosity Squeegee wear Stencil wear Solder paste before and after printing

** Area and Aspect ratios indirectly available.*

From the AIM, see 5.1.4, the participants stressed that data about wearing on the stencils and squeegees are not collected, how many cycles they have been used or how much solder paste that is applied on the stencils and how much solder paste is left afterwards. These are important parameters to look at as well as to collect information about between what boards a cleaning has been made. The case company should consider to overall work more with the collection and storage of data. This is according to the process engineers something new that they have not worked with for long so routines and infrastructure about the collection, storage and accessibility of the data has not been developed. Therefore, a data collection plan would be good for the case company to structure the data gathering and also address on how to store and access the data for analysis. The case company should also consider to work more with data overall and use it to improve their processes, reduce variations and thus save both time and resources that could be spent on better things.

6.7 Ishikawa Diagram

To summarize the discussion regarding the identified root causes of the variation in the Y, an Ishikawa diagram is presented in Figure 6.1.

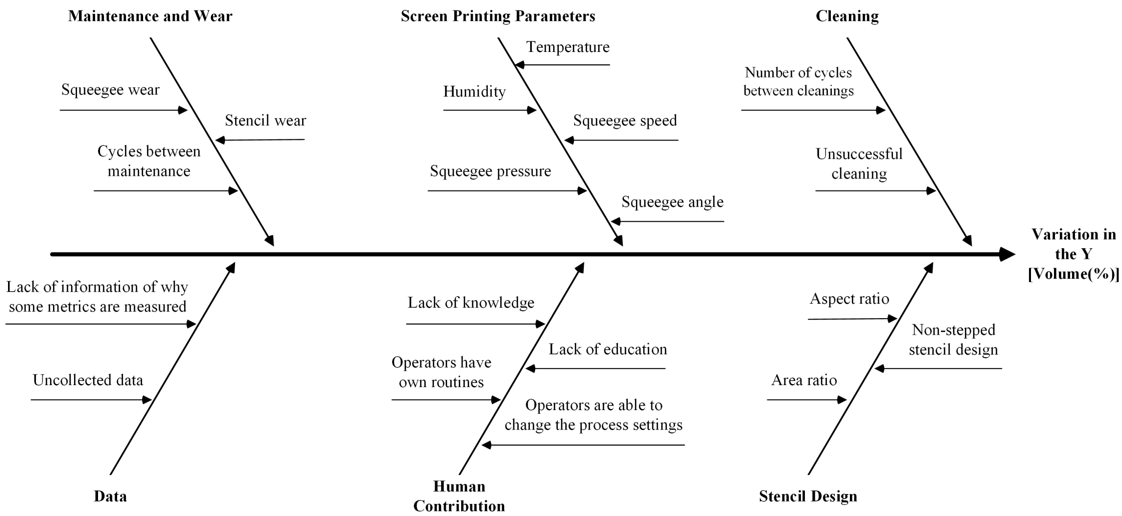


Figure 6.1: Ishikawa diagram of the variation in the Y.

6.8 Relevance to the Field of Research

When searching for literature related to Six Sigma and SMT, there was not much to be found. Some articles that were found were of older type and had different settings than this proposed project. Further, there has been several research articles concerning multiple regression analysis to investigate which parameters that are important. However, in today's production environment with big data, multiple regression is not powerful enough to handle this vast amount of data. Hence, creating a gap in the Six Sigma methodology and a new game plan is needed. Due to Six Sigma lacking a method for big data analytics (BDA), there is an interest from both researchers and practitioners within Industry 4.0 to solve this issue (Belhadi et al., 2020).

Gupta, Modgil Gunasekaran (2020) presented a holistic view of BDA in Lean Six Sigma for future research, where they mentioned machine learning as a research topic to handle BDA in Six Sigma. This since machine learning are able to learn patterns and make smart conclusions. Furthermore, little literature has been found around combining machine learning techniques with the issue regarding BDA using the Six Sigma methodology. In this thesis the machine learning techniques random forest and boosting was used to handle the BDA. These are useful methods to predict future data and by conducting this thesis, knowledge within the field of Six Sigma and Quality 4.0 is expanded.

Many of the researched articles findings and conclusions are in line with the findings and conclusions from this thesis but there are differences, especially related to the importance of area and aspect ratio and the qualitative side. These all present important gaps in the literature and through the setting of the master thesis as a case study, important research areas related to the field has been expanded. The Six Sigma approach also provides structure and depth to the analysis and the way on how to approach the stated problem. This research will therefore expand current knowledge and scientific research with new takings on Quality 4.0 related to Six Sigma.

6.9 Future Research

The project team suggests following up with the Improve and Control phase of this master thesis. Further, the project team suggest future research to conduct DoE with factorial experiments to further investigate the optimal settings of the factors, both those that are currently being measured and those that are not measured or stored. In addition, the project team recommends future DoE to be conducted with many different factors combined, not only with a few as current research has done. The case study conducted presents the complexity of the problem and that there are no shortcuts to be taken but all potential factors must be considered.

Future research would also contribute to the scientific field by including all of the five DMAIC phases and relating them both to machine learning and SMT-production. There are many gaps in the literature, as previously discussed, and continuing with a broader approach with Six Sigma related to SMT-production and machine learning would contribute to the expansion of current knowledge and research.

6.9.1 What is next for the case company

At the end of the master thesis, a presentation was held for the case company where the findings and recommendations were presented. The case company will consider continuing with the Improve and Control phases and investigate the optimal settings of the identified factors. The case company were satisfied with the findings and the following quote were mentioned:

> *"There were a few things that we already knew about, but most of it were new and very interesting. It is true what you are saying that many things are affecting the outcome of the solder paste volume, making the problem*

complex. We will definitely look into this, and you have created good ground for improvements. Well done.”

7

Conclusions & Recommendations

This master thesis has been investigating the yield variations at the SMT process at the case company. The aim was to, in a better way understand why the case company have yield variations and to find the root causes to these variations. In addition, they wanted to use the master thesis as a pilot for running future Six Sigma projects. Therefore, the objective was to help the case company understand the causes of yield variation and deliver validated and verified recommendations.

7.1 RQ1: How can Six Sigma be used to provide a systematic approach to understand the problem?

Six Sigma as a business strategy could help the case company to understand and improve customers' requirements, the company's productivity and financial performance. Six Sigma methodology helps to break down problems into sub-problems and thereby identify potential root causes to the main problem. Six Sigma not only explain what to do, but also how to do it. To approach the problem with the process variations, the solder paste volume metric were broken down to several components, structured and then analysed and compared with one another. Both quantitative and qualitative data are important to investigate if they match or not. Validity and verification of the data are also important, and the MSA conducted supported that the fact that the data can be trusted, both regarding repeatability and reproducibility.

SMT7 differed from the rest of the SMT-lines and a recommendation would be to investigate that further. The measurement system data from SMT1 to SMT6 were acceptable and considered valid. There are, however, possibilities for improvement but due to the fact that the measurement system is acceptable, it might not be worth the resources and time spent on it for minor improvements.

7.2 RQ2: What are the root causes to the yield variations and how can those variations be managed?

Looking at the whole picture, the identified root causes to the yield variation related to solder paste volume were discussed as the area ratio, aspect ratio, stencil cleaning intervals and the operators freedom to change the process settings. Furthermore, there are data that have not been collected, not measured, stored or inaccessible. These data should be considered to be collected, see Table 6.2. The area and aspect ratios are important metrics that could be used to lower variation at the SMT-lines. The ratios strengthened the hypothesis that the bigger the area and aspect ratio, the less solder paste volume variation becomes. Also, the cleaning interval is playing an important role in keeping the stencils from clogging and could be considered further.

The project team further recommend the case company to conduct DoE with factorial experiments to further investigate the optimal settings of the factors, especially when introducing new products, changing machines and inputs like solder paste. This could be done internally or through a future master thesis that continues on the work presented in this master thesis.

The project team recommends the case company to implement stepped stencils to more products and consider the area and aspect ratios when developing them. Further, the case company is recommended to consider minimizing the potential of human errors as much as possible by restricting the possibility for the operators to adjust settings themselves. The case company is recommended to also consider questioning why some things are being measured and that they are not just measured by routine or as remnants of old products that required other settings and measurements.

Lastly, the project team recommends the case company to consider collecting data regarding cleaning intervals, how it is connected to the data when occurred between boards and how efficiently solder paste is removed from the stencils. The case company should also consider collecting data on how much solder paste that is being put on the stencils and how much is left after printing to assess how much solder paste is left on the stencil afterwards. The data about the factors should also be considered to be stored in an accessible database and used to analyse and improve production and reduce process variation.

Bibliography

[1] Ahuett-Garza, H., & Kurfess, T. (2018). A brief discussion on the trends of habilitating technologies for Industry 4.0 and Smart manufacturing. *Manufacturing Letters*, 15 (Part B), 60–63.

[2] Alänge, S. (2009). The Affinity-Interrelationship Method AIM. A Problem Solving Tool for Analysing Qualitative Data Inspired by the Shiba “Step by Step” Approach.

[3] Antony, J. (2011). Six Sigma vs Lean: Some perspectives from leading academics and practitioners.*International Journal of Productivity and Performance Management*, 60(2), 185-190.

[4] Antony, J., Kumar, A., & Banuelas, R. (Eds.). (2006). *World class applications of six sigma.* Routledge.

[5] Arumugam, V., Antony, J., & Linderman, K. (2016). The influence of challenging goals and structured method on Six Sigma project performance: A mediated moderation analysis. *European Journal of Operational Research, 254(1),* 202-213.

[6] Asghar, R., Rehman, F., Aman, A., Iqbal, K., & Nawaz, A. A. (2019). Defect minimization and process improvement in SMT lead-free solder paste printing: a comparative study. *Soldering & Surface Mount Technology.*

[7] Baluch, D. & Minogue, G. (2007). Fundamentals of solder paste technology. Global SMT and Packaging. Oxfordshire, UK.

[8] Belhadi, A., Kamble, S. S., Zkik, K., Cherrafi, A., Touriki, F. E. (2020). The integrated effect of Big Data Analytics, Lean Six Sigma and Green Manufacturing on the environmental performance of manufacturing companies: The case of North Africa. *Journal of Cleaner Production*, 252, 119903.

[9] Bryman, A. (2006). Integrating quantitative and qualitative research: how is it done?. SAGE Publications.

[10] Bryman, A. (2012). Social research methods. Oxford University Press, USA.

[11] Burdick, R. K., Borror, C. M., & Montgomery, D. C. (2005). Design and analysis of gauge R&R studies: making decisions with confidence intervals in random and mixed ANOVA models (Vol. 17). SIAM.

[12] Carleton, S. A. (2018). The Black Belt Memory Jogger Second Edition: A pocket guide for Six Sigma DMAIC Success. Methuen, Massachusetts, USA: GOAL/QPC.

[13] Carlson, C. S. (2020). This Month’s Theme is Parameter Diagrams (P-Diagrams). Retrieved from: *https://www.weibull.com/hotwire/issue182/fmeacorner182.htm*

[14] Chen, C. C. (2008). An objective-oriented and product-line-based manufacturing performance measurement. *International Journal of Production Economics, 112(1),* 380-390.

[15] Chun, Y., & Bidanda, B. (2013). Sustainable manufacturing and the role of the International Journal of Production Research. *International Journal of Production Research, 51(23/24),* 7448–7455.

[16] Farrell, R., Shea, C., & Burlington, N. J. (2013, October). Stencil and Solder Paste Inspection Evaluation for Miniaturized SMT Components. In *Proceedings of SMTA International.*

[17] Golafshani, N. (2003). Understanding Reliability and Validity in Qualitative Research. *The Qualitative Report.*

[18] Gopal, S., Rohani, J. M., Yusof, S. M. & Bakar, Z. A. (2006). Optimization of Solder Paste Printing Parameters Using Design of Experiments (DOE). *Jurnal Teknologi, 43(A) Dis. 2006: 11–20,* Universiti Teknologi Malaysia.

[19] Gupta, S., Modgil, S., Gunasekaran, A. (2020). Big data in lean six sigma: a review and further research directions. *International Journal of Production Research*, 58(3), 947-969.

[20] Hammersberg, P. (2019a). Effective Scoping template. [PowerPoint-presentation]. Retrieved from: *https://chalmers.instructure.com/courses/3811/files/51177?module_item_id=10041*

[21] Hammersberg, P. (2019b). Scoping Challenges. [PowerPoint-presentation]. Retrieved from: *https://chalmers.instructure.com/courses/3811/files/51271?module_item_id=10053*

[22] Harry, M. J. (1998). Six Sigma: a breakthrough strategy for profitability. *Quality progress, 31(5),* 60.

[23] Huang, C., Lin, Y., Ying, K. & Ku, C. (2011). The Solder Paste Printing Process - Critical Parameters, Defect Scenarios, Specifications and Cost Reduction. *Soldering and Surface Mount Technology 23(4):211-223.*

[24] Hwang, J. (1989). Solder Paste in Electronics Packaging - Technology and Applications in Surface Mount, Hybrid Circuits, and Component Assembly. Van Nostrand Reinhold, New York.

[25] Investopedia (2020). How Do You Calculate Working Capital?. Retrieved from: *https://www.investopedia.com/ask/answers/071114/how-do-you-calculate-working-capital.asp*

[26] IPC (2007). IPC-7525: Stencil Design Guidelines. Illinois, USA: IPC.

[27] IPC (2012). IPC-7527: Requirements for Solder Paste Printing. Illinois, USA: IPC.

[28] James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013). An Introduction to Statistical Learning - with Applications in R. Springer, New York.

[29] Juran, J. M. (1993). Quality planning and analysis; from product development through use (No. 04; TS156, J8 1993.)

[30] Khader, N., Yoon, S. W. & Li, D. (2017). Stencil printing optimization using a hybrid of support vector regression and mixed-integer linear programming. *Procedia Manufacturing, 11,* 1809-1817.

[31] Kotter, J. P. (2008). A sense of urgency. Harvard Business Press.

[32] Kumar, M., Antony, J., Madu, C. N., Montgomery, D. C., & Park, S. H. (2008). Common myths of Six Sigma demystified. *International Journal of Quality & Reliability Management.*

[33] Kwak, Y. H., & Anbari, F. T. (2006). Benefits, obstacles, and future of six sigma approach. *Technovation, 26(5-6),* 708-715.

[34] LaDou, J. (2006). Printed Circuit Board Industry. *International Journal of Hygiene and Environmental Health, Volume 209(3),* 211-219.

[35] Lau, F. K., & Yeung, V. W. (1997). A hierarchical evaluation of the solder paste printing process. *Journal of Materials Processing Technology, 69(1-3),* 79-89.

[36] LeanOhio (2014). LeanOhio Green Belt: Transforming the Public Sector: Critical To Quality – CT Flowdown [PowerPoint]. Retrieved from: *https://www.lean.ohio.gov/Portals/0/docs/training/GreenBelt/GB_CT%20 Flowdown.pdf*

[37] Li, M. H. C., Al-Refaie, A., & Yang, C. Y. (2008). DMAIC approach to improve the capability of SMT solder printing process. *IEEE Transactions on Electronics Packaging Manufacturing, 31(2),* 126-133.

[38] Linderman, K., Schroeder, R. G., Zaheer, S., & Choo, A. S. (2003). Six Sigma: a goal-theoretic perspective. *Journal of Operations management, 21(2),* 193-203

[39] Mannan, S. H., Ekere, N. N., Ismail, I., & Lo, E. K. (1994). Squeegee deformation study in the stencil printing of solder pastes. *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part A, 17(3),* 470-476.

[40] Marcoux, P. P. (1992). Fine pitch surface mount technology: quality, design, and manufacturing techniques. Van Nostrand Reinhold, New York.

[41] Minitab (2020). Assessing Measurement System Variation. Retrieved from: *https://www.minitab.com/uploadedFiles/Documents/sample-materials/FuelInjectorNozzles-EN.pdf*

[42] Pan, J., Tonkay, G. L., Storer, R. H., Sallade, R. M., & Leandri, D. J. (2004). Critical variables of solder paste stencil printing for micro-BGA and fine-pitch QFP. *IEEE Trans. Electron. Packag. Manuf., vol. 27, no. 2, april,* 125-132.

[43] Schoonenboom, J. & Johnson, B. (2017). How to Construct a Mixed Methods Research Design. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie, Volume 69.*

[44] Schroeder, R. G., Linderman, K., Liedtke, C., & Choo, A. S. (2008). Six Sigma: Definition and underlying theory. *Journal of operations Management, 26(4),* 536-554.

[45] Shmueli, G., Bruce, P., Yahav, I., Patel, N. & Lichtendahl, K. (2018). Data Mining for Business Analytics - Concepts, Techniques, and Applications in R. John Wiley & Sons.

[46] Sundkvist, R., Hedman, R., Almström, P., & Kinnander, A. (2012). Improvement potentials in Swedish electronics manufacturing industry–Analysis of five case studies. *Procedia CIRP, 3,* 126-131.

[47] Surface Mount Process (2020a). Solder Paste Printing Process. Retrieved from: *http://www.surfacemountprocess.com/solder-paste-printing-process.html*

[48] Surface Mount Process (2020b). Solder Paste Inspection Process. Retrieved from: *http://www.surfacemountprocess.com/solder-paste-inspection-process.html*

[49] Tong, J. P. C., Tsung, F., & Yen, B. P. C. (2004). A DMAIC approach to printed circuit board quality improvement. *The International Journal of Advanced Manufacturing Technology, 23(7-8),* 523-531.

[50] Tsai, T. N. (2011). Improving the fine-pitch stencil printing capability using the Taguchi method and Taguchi fuzzy-based model. *Robotics and Computer-Integrated Manufacturing, 27(4),* 808-817.

[51] Vault (2020). Electronics Manufacturing. Retrieved from: *https://www.vault.com/industries-professions/industries/electronics-manufacturing*

[52] Wallén, G. (1996). Vetenskapsteori och forskningsmetodik. Lund: Studentlitteratur.

[53] Waller, V., Farquharson, K. & Dempsey, D. (2016). Qualitative Social Research: Contemporary Methods for the Digital Age. Sage Publications.

[54] Walters, L. (2005). Six Sigma: is it really different?. *Quality and Reliability Engineering International*, 21(3), 221-224.

A

Original AIM

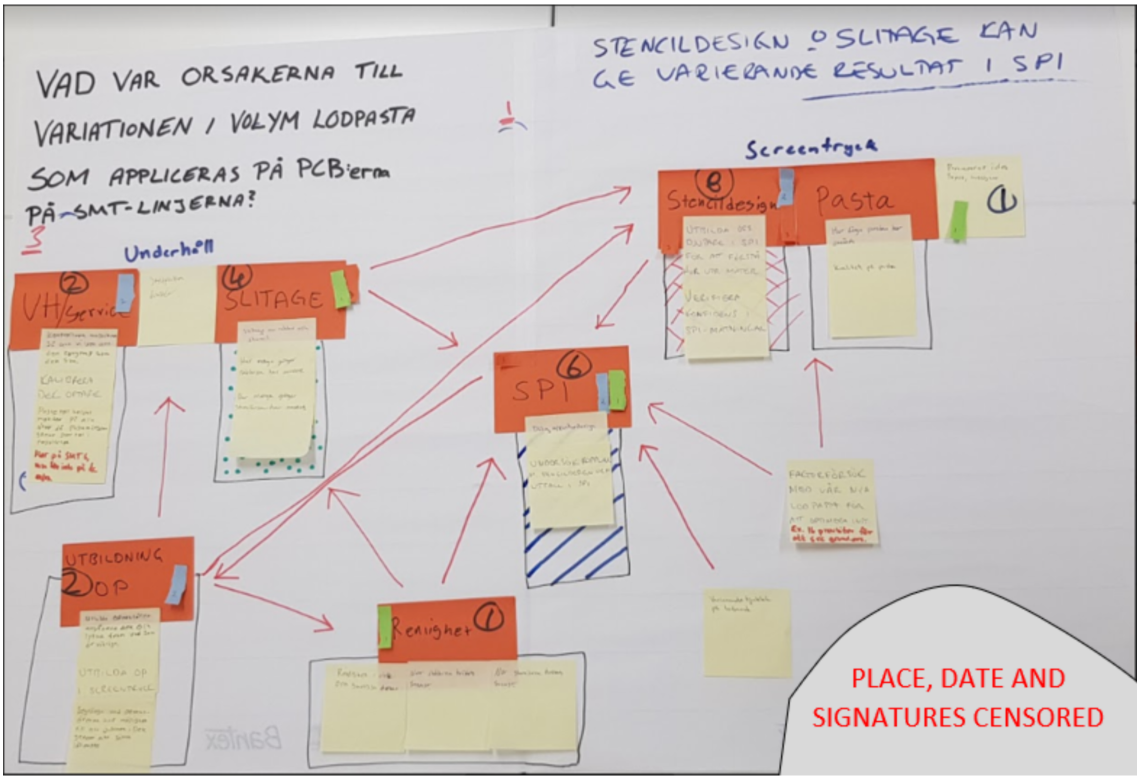
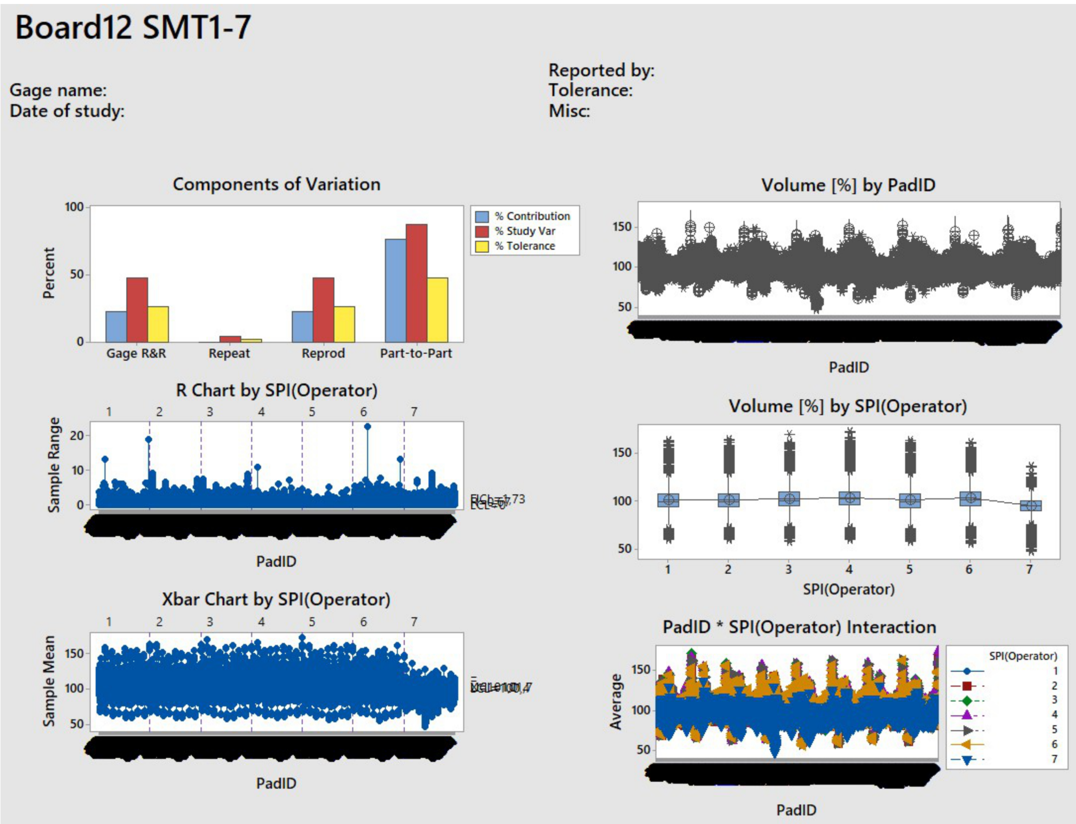


Figure A.1: Original AIM.

B

MSA Graphs for SMT1-7



(a)

Results for: Board12

Gage R&R Study - ANOVA Method

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
PadID	11791	23467038	1990	35,10	0,000
SPI(Operator	6	1857397	309566	5460,08	0,000
PadID * SPI(Operator	70746	4011033	57	224,35	0,000
Repeatability	165088	41719	0		
Total	247631	29377186			

α to remove interaction term = 0,05

(b)

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	27,816	23,20
Repeatability	0,253	0,21
Reproducibility	27,564	22,99
SPI(Operator	8,749	7,30
SPI(Operator*PadID	18,815	15,69
Part-To-Part	92,074	76,80
Total Variation	119,890	100,00

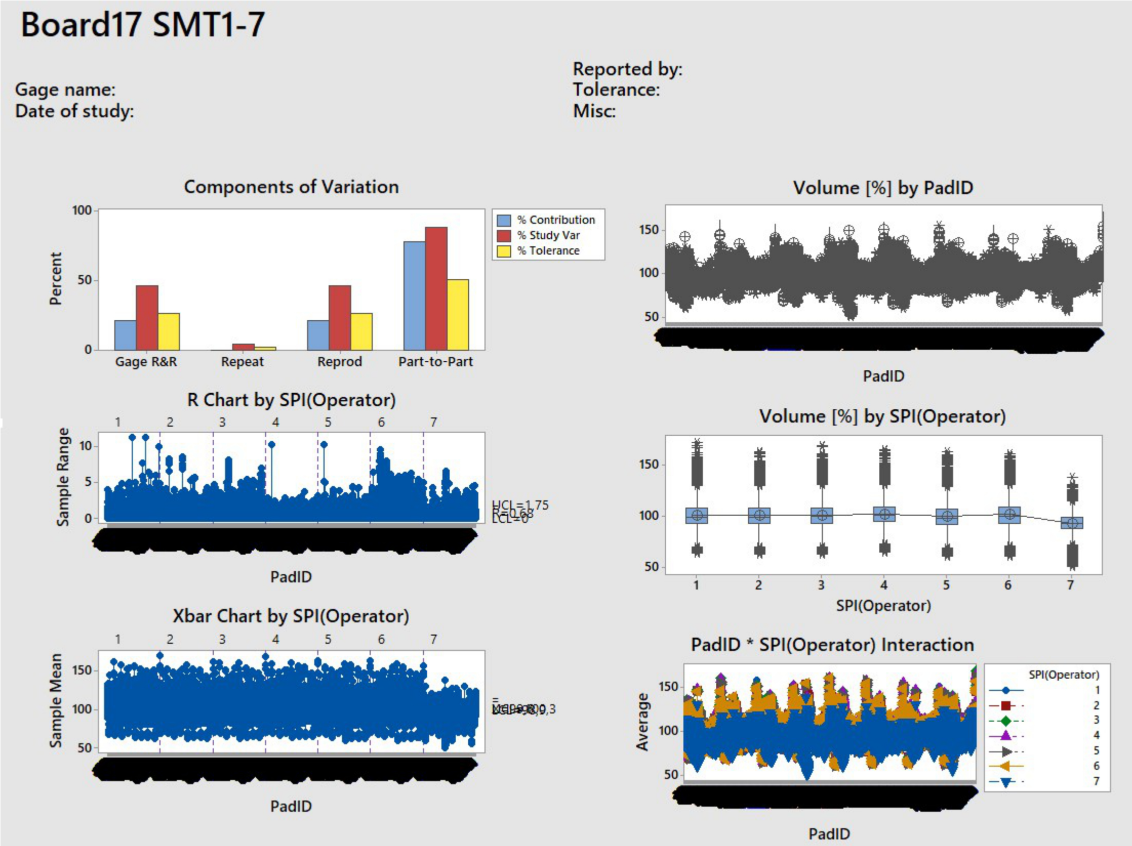
Process tolerance = 120

Source	StdDev (SD)	Study Var (6 × SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	5,2741	31,6447	48,17	26,37
Repeatability	0,5027	3,0162	4,59	2,51
Reproducibility	5,2501	31,5007	47,95	26,25
SPI(Operator	2,9579	17,7474	27,01	14,79
SPI(Operator*PadID	4,3376	26,0254	39,61	21,69
Part-To-Part	9,5955	57,5731	87,63	47,98
Total Variation	10,9494	65,6967	100,00	54,75

Number of Distinct Categories = 2

(c)

Figure B.1: Results in Minitab from the Gauge R&R SMT1-6 of board 12. (a) Gauge R&R graphs. (b) Two-Way ANOVA results. (c) Gauge R&R results.



(a)

Results for: Board17

Gage R&R Study - ANOVA Method

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
PadID	11791	26184256	2221	40,43	0,000
SPI(Operator	6	2088833	348139	6338,01	0,000
PadID * SPI(Operator	70746	3885988	55	219,94	0,000
Repeatability	165088	41229	0		
Total	247631	32200306			

α to remove interaction term = 0,05

(b)

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	28,316	21,54
Repeatability	0,250	0,19
Reproducibility	28,066	21,35
SPI(Operator	9,840	7,49
SPI(Operator*PadID	18,226	13,87
Part-To-Part	103,132	78,46
Total Variation	131,448	100,00

Process tolerance = 120

Source	StdDev (SD)	Study Var (6 × SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	5,3212	31,9275	46,41	26,61
Repeatability	0,4997	2,9984	4,36	2,50
Reproducibility	5,2977	31,7863	46,21	26,49
SPI(Operator	3,1368	18,8208	27,36	15,68
SPI(Operator*PadID	4,2692	25,6154	37,24	21,35
Part-To-Part	10,1554	60,9323	88,58	50,78
Total Variation	11,4651	68,7903	100,00	57,33

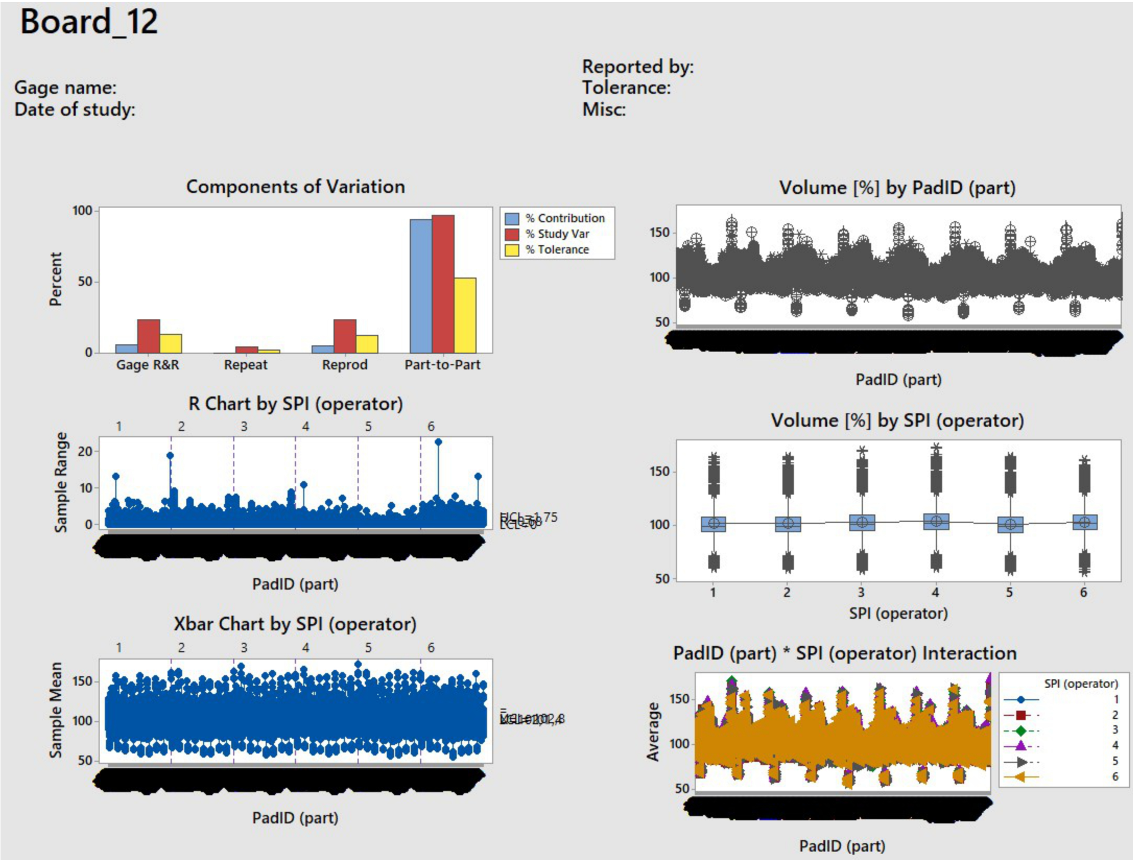
Number of Distinct Categories = 2

(c)

Figure B.2: Results in Minitab from the Gauge R&R SMT1-6 of board 17. (a) Gauge R&R graphs. (b) Two-Way ANOVA results. (c) Gauge R&R results.

C

MSA Graphs for SMT1-6



(a)

Results for: Board 12

Gage R&R Study - ANOVA Method

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
PadID (part)	11791	24006722	2036,0	124,56	0,000
SPI (operato	5	193763	38752,7	2370,83	0,000
PadID (part) * SPI (operato	58955	963656	16,3	65,01	0,000
Repeatability	141504	35576	0,3		
Total	212255	25199716			

α to remove interaction term = 0,05

(b)

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	6,711	5,64
Repeatability	0,251	0,21
Reproducibility	6,460	5,43
SPI (operato	1,095	0,92
SPI (operato*PadID (part)	5,365	4,51
Part-To-Part	112,204	94,36
Total Variation	118,915	100,00

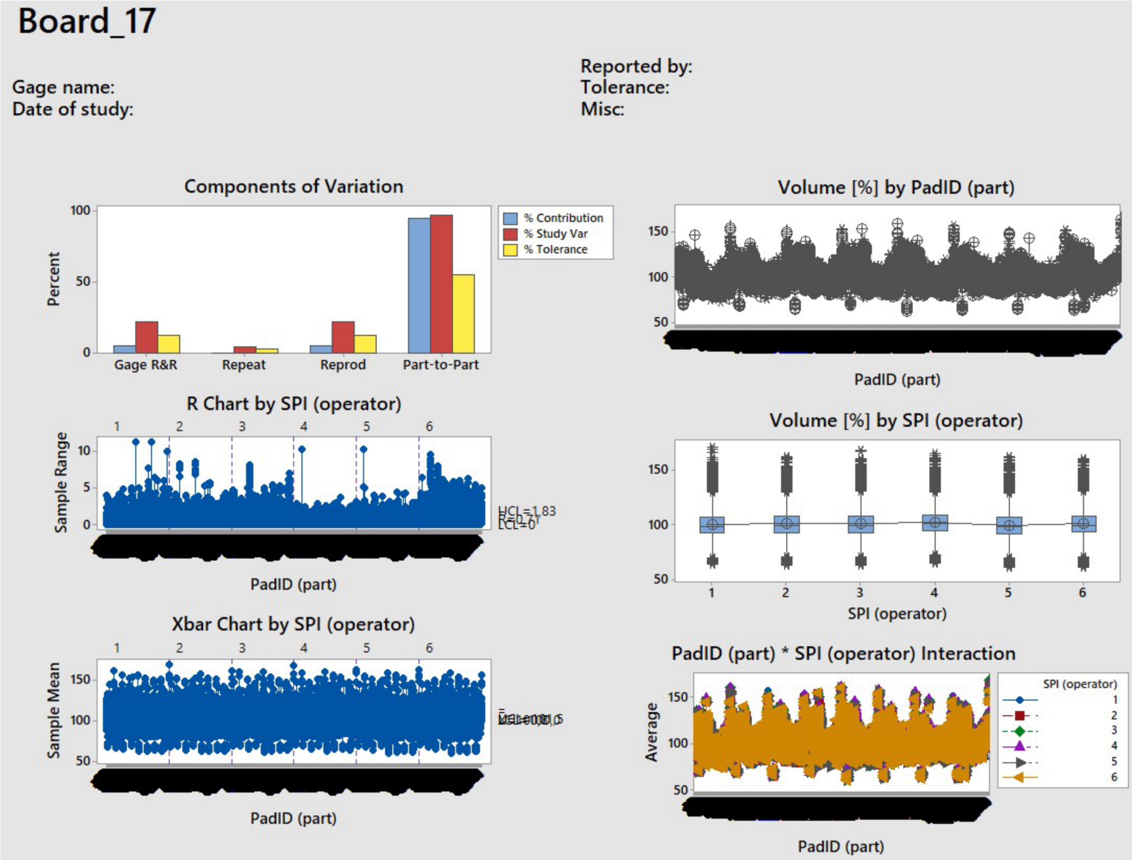
Process tolerance = 120

Source	StdDev (SD)	Study Var (6 × SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	2,5906	15,5435	23,76	12,95
Repeatability	0,5014	3,0085	4,60	2,51
Reproducibility	2,5416	15,2496	23,31	12,71
SPI (operato	1,0464	6,2785	9,60	5,23
SPI (operato*PadID (part)	2,3162	13,8971	21,24	11,58
Part-To-Part	10,5926	63,5559	97,14	52,96
Total Variation	10,9048	65,4290	100,00	54,52

Number of Distinct Categories = 5

(c)

Figure C.1: Results in Minitab from the Gauge R&R SMT1-6 of board 12. (a) Gauge R&R graphs. (b) Two-Way ANOVA results. (c) Gauge R&R results.



(a)

Results for: Board 17

Gage R&R Study - ANOVA Method

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
PadID (part)	11791	26141873	2217,1	131,36	0,000
SPI (operato	5	124386	24877,2	1473,99	0,000
PadID (part) * SPI (operato	58955	995014	16,9	62,61	0,000
Repeatability	141504	38145	0,3		
Total	212255	27299418			

α to remove interaction term = 0,05

(b)

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	6,508	5,06
Repeatability	0,270	0,21
Reproducibility	6,239	4,85
SPI (operato	0,703	0,55
SPI (operato*PadID (part)	5,536	4,30
Part-To-Part	122,235	94,94
Total Variation	128,743	100,00

Process tolerance = 120

Source	StdDev (SD)	Study Var (6 × SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	2,5511	15,3068	22,48	12,76
Repeatability	0,5192	3,1152	4,58	2,60
Reproducibility	2,4977	14,9865	22,01	12,49
SPI (operato	0,8383	5,0298	7,39	4,19
SPI (operato*PadID (part)	2,3529	14,1172	20,74	11,76
Part-To-Part	11,0560	66,3359	97,44	55,28
Total Variation	11,3465	68,0790	100,00	56,73

Number of Distinct Categories = 6

(c)

Figure C.2: Results in Minitab from the Gauge R&R SMT1-6 of board 17. (a) Gauge R&R graphs. (b) Two-Way ANOVA results. (c) Gauge R&R results.

D

R Code

```
```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

```{r InstallPackages}
library(tidyverse)
library(data.table)
library(randomForest)
library(xgboost)
library(ggplot2)
library(ggthemes)
```

```{r MergeCsvFiles}
filenames <- list.files(pattern=glob2rx("SMT?_*.csv"),
full.names=TRUE)
SolderData0 <- rbindlist(lapply(filenames,fread))
SolderData <- SolderData0
names(SolderData) <- make.names(names(SolderData),
unique = TRUE)
SolderData <- mutate_if(SolderData, is.character, as.factor)
```

```{r ScatterDiagram, fig.width=9, fig.height=6}
theme_set(theme_tufte())
p <- ggplot(SolderData, aes(x = ApertureArea,
y = VolumePct))
p1 <- p + geom_jitter(aes(color = as.character(SMTLine)),
width = 0, size = 1) + facet_wrap(~as.character(SMTLine))
+ guides(color = F) + geom_smooth()
#p2 <- p1 + geom_violin(alpha = 0.1) +
aes(color = as.character(SMTLine), alpha = 0.1)
+ guides(alpha = F)
plot(p1)
#ggsave("name.png", plot=p1, dpi = 300, height = 15,
width = 20, unit = "cm")
```
```

```
```

```{r CreateDummyCompID}
CompIDlist <- names(summary(SolderData$ComponentID))
for (i in 1:length(CompIDlist)) {
  CompID <- grep(CompIDlist[i],SolderData$ComponentID)
  SolderData[,ncol(SolderData)+1] <- 0
  SolderData[CompID, ncol(SolderData)] <- 1
}
NewColIndexStart <- ncol(SolderData0)+1
NewColIndexEnd <- ncol(SolderData0)+length(CompIDlist)
colnames(SolderData)[NewColIndexStart:NewColIndexEnd]
<- CompIDlist
```

```{r CreateDummyResult}
ResultList <- names(summary(SolderData$Result))
for (i in 1:length(ResultList)) {
  ResultID <- grep(ResultList[i],SolderData$Result)
  SolderData[,ncol(SolderData)+1] <- 0
  SolderData[ResultID, ncol(SolderData)] <- 1
}
NewColIndexStart <- ncol(SolderData0)+length(CompIDlist)+1
NewColIndexEnd <- NewColIndexStart+length(ResultList)-1
colnames(SolderData)[NewColIndexStart:NewColIndexEnd]
<- ResultList
```

```{r CreateDummyJob}
JobList <- names(summary(SolderData$Job))
for (i in 1:length(JobList)) {
  JobID <- grep(JobList[i],SolderData$Job)
  SolderData[,ncol(SolderData)+1] <- 0
  SolderData[JobID, ncol(SolderData)] <- 1
}
NewColIndexStart2 <- ncol(SolderData0)+length(CompIDlist)
+length(ResultList)+1
NewColIndexEnd <- NewColIndexStart2+length(JobList)-1
colnames(SolderData)[NewColIndexStart2:NewColIndexEnd]
<- JobList
```

```{r CreateDummySupportPlateType}
SupportPlateTypeList <-
names(summary(SolderData$SupportPlateType))
for (i in 1:length(SupportPlateTypeList)) {
```

```
  SupportPlateTypeID <- grep(SupportPlateTypeList[i],
  SolderData$SupportPlateType)
  SolderData[,ncol(SolderData)+1] <- 0
  SolderData[SupportPlateTypeID, ncol(SolderData)] <- 1
}
NewColIndexStart3 <- ncol(SolderData0)+length(CompIDlist)
+length(ResultList)+length(JobList)+1
NewColIndexEnd <- NewColIndexStart3+
length(SupportPlateTypeList)-1
colnames(SolderData)[NewColIndexStart3:NewColIndexEnd]
<- SupportPlateTypeList
```

```{r CreateDummySMT}
SolderData$SMTLine <- as.factor(SolderData$SMTLine)
SMTList <- names(summary(SolderData$SMTLine))
for (i in 1:length(SMTList)) {
  SMTID <- grep(SMTList[i],SolderData$SMTLine)
  SolderData[,ncol(SolderData)+1] <- 0
  SolderData[SMTID, ncol(SolderData)] <- 1
}
dim(SolderData)
colnames(SolderData)[155:160] <- SMTList
glimpse(SolderData)
```

```{r ReduceTheMatrix}
SolderDataReady<-SolderData[,-which(names(SolderData)
%in% c("Job", "PCBID", "Barcode", "Panel", "ComponentID",
"PadID", "Volume", "Height", "Area", "AreaPct", "Result",
"Time", "DigitalTime", "StencilThickness", "StencilStepped",
"StencilArticleNumber", "SupportPlateType", "SMTLine",
"SeparationDistance", "SeparationSpeed", "AperaturArea",
"AperaturWallArea"))]
names(SolderDataReady) <- make.names(names(SolderDataReady),
unique = TRUE)
SolderDataReady <- select_if (SolderDataReady, is.numeric)
SolderDataReady<-na.omit(SolderDataReady)
```

```{r TrainAndTestData}
set.seed(345)
SolderData <- SolderDataReady
TrainObs<-sample(nrow(SolderData),
round(nrow(SolderData)*.7,0)
,rep=F)
```

```
TrainData<-SolderData[TrainObs,]
TestData<-SolderData[-TrainObs,]
YColNo <- which(names(SolderData)=="VolumePct")
```
```

```{r MultReg}
options(scipen = 999)
OLSfit<-lm(VolumePct~., data=TrainData)
summary(OLSfit)
YhatOLS <- predict(OLSfit,newdata=TestData[,-YColNo])
Actual_Pred<-tibble (c1=YhatOLS, c2=TestData$VolumePct,
Err=c1-c2)
MSE_multReg <- mean((YhatOLS - TestData$VolumePct)^2)
MSE_multReg
```

```{r RandomForest}
NumTree <- 100
#Change this to a slighty bigger one e.g. 200 or 300
names(TrainData) <- make.names(names(TrainData),
unique = TRUE)
names(TestData) <- make.names(names(TestData),
unique = TRUE)
RfFit<-randomForest(VolumePct~.,data=TrainData,
mtry=sqrt(ncol(TrainData)),ntree=NumTree)
plot(RfFit)
RfPred<-predict(RfFit,newdata=TestData[,-YColNo])
mean((RfPred-TestData$VolumePct)^2)
imp<-data.frame(importance(RfFit))
imp$vars<-row.names(imp)
arrange(imp,desc(IncNodePurity))[1:10,]
```

```{r Boosting}
trainBoost<-sapply(TrainData,as.numeric)
testBoost<-sapply(TestData,as.numeric)
dim(testBoost)
dim(trainBoost)
boostFit<-xgboost(data=trainBoost[,-YColNo],
label=trainBoost[,YColNo],max.depth=1,nthread=4,nround=200,
verbose=0)
boostPred<-predict(boostFit,newdata=testBoost[,-YColNo])
mean((boostPred-TestData$VolumePct)^2)
df<-tibble(etas = seq(0.05,0.95,0.1),
           depths = seq(1,10,1)) %>% expand(etas,depths)
df$mse<-NA
```
```

```
for(i in 1:100)
  {boostCvFit<-xgb.cv(data = trainBoost[,-YColNo],
                      label = trainBoost[,YColNo],
                      max.depth=df$depths[i],
                      nround=100,
                      eta = df$etas[i],
                      objective = "reg:squarederror",
                      verbose=0, nfold = 5)
  bestFitCvError <- boostCvFit$evaluation_log[which.min
  (boostCvFit$evaluation_log$test_rmse_mean),
  'test_rmse_mean']
  df$mse[i]<-bestFitCvError^2
  }
df[which.min(df$mse),]
bestDepth<-df[which.min(df$mse),'depths']
bestEta<-df[which.min(df$mse),'etas']
boostFitBest<-xgboost(data = trainBoost[,-YColNo],
label=trainBoost[,YColNo],
max.depth=bestDepth, nround=100,eta = bestEta,
objective="reg:squarederror",verbose=0)
boostPredBest<-predict(boostFitBest,
newdata=testBoost[,-YColNo])
mean((boostPredBest-testBoost[,YColNo])^2)
importance <- xgb.importance(model = boostFit)
print(xgb.plot.importance(importance_matrix = importance,
top_n = 10))
```

DEPARTMENT OF TECHNOLOGY MANAGEMENT AND ECONOMICS
DIVISION OF SERVICE MANAGEMENT AND LOGISTICS
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY