NOVA SCHOOL OF SCIENCE & TECHNOLOGY

MECHANICAL AND INDUSTRIAL ENGINEERING DEPARTMENT

ROBIN LOUIS ROBERT STREF COUEDEL

Bachelor in Sciences of Industrial Engineering and Management

STATISTICAL QUALITY CONTROL OF A PRODUCTION PROCESS OF INVISIBLE ZIPPERS

Dissertation to obtain the degree of Master in Industrial Engineering and Management

INTEGRATED MASTER IN INDUSTRIAL ENGINEERING AND MANAGEMENT

NOVA University Lisbon

Novembro, 2021

NOVA SCHOOL OF SCIENCE & TECHNOLOGY

MECHANICAL AND INDUSTRIAL ENGINEERING DEPARTMENT

STATISTICAL QUALITY CONTROL OF A PRODUCTION PROCESS OF INVISIBLE ZIPPERS

Dissertation to obtain the degree of Master in Industrial Engineering and Management

ROBIN LOUIS ROBERT STREF COUEDEL
Bachelor in Sciences of Industrial Engineering and Management

Supervisor: Professor Doctor Izunildo Fernandes Cabral, Invited Assistant Professor, NOVA University Lisbon

Co-supervisor: Professor Doctor Ana Sofia Leonardo Vilela de Matos, Associate Professor, NOVA University Lisbon

Júri:

Presidente: Professor Doctor Ana Paula Ferreira Barroso, Assistant Professor, NOVA University Lisbon

Arguentes: Professor Doctor António João Pina da Costa Feliciano Abreu, Coordinating Professor with Aggregation, Lisbon Higher Institute of Engineering (ISEL)

Vogais: Professor Doctor Ana Sofia Leonardo Vilela de Matos, Associate Professor, NOVA University Lisbon

INTEGRATED MASTER IN INDUSTRIAL ENGINEERING AND MANAGEMENT

NOVA University Lisbon
November, 2021

Statistical Quality Control of a Production Process of Invisible Zippers

I dedicate this work to my parents.
Without them nothing would be possible.

ACKNOWLEDGMENTS

First, I wish to show my appreciation to my supervisor, Professor Doctor Izunildo Cabral, who assisted me throughout this thesis, and without who this thesis would not exist. I also wish to thank him for all the knowledge with which he helped me throughout the project.

To my co-supervisor, Professor Doctor Ana Sofia Matos. I wish to thank her for all the knowledge with which she helped me throughout the making of this thesis and the project, and without who this thesis would not exist either.

And to the head of the project, Professor Doctor Rogério Puga Leal, who led us throughout this project, and with who I learned plenty about communicating with the client, both through written and in-person communication. I also want to sincerely thank the whole team of the project, Professor Doctor Rogério Puga Leal, Professor Doctor Ana Sofia Matos, and Professor Doctor Izunildo Cabral, and emphasize the importance of their knowledge on statistical concepts and their application to practical situations, from which I immensely learned during these months.

I also wish to show my sincere appreciation for all the company members involved in the project, and thank them for their sympathy, for their willingness to help us during the project, and for providing us with all their knowledge.

I sincerely thank the NOVA School of Science and Technology for the research grant they provided me.

A special thanks to all my friends who accompanied me on this journey. I could not help but mention Carlos Antunes, Victor Balin, Rita Bajanca, Felipe Garcia, and Daniel Castela. They helped me grow in so many ways, and the moments we lived together will never be forgotten. I also want to mention all my dear friends I met during my Erasmus program: Rio, Nadia, Óscar, Thiago, David, Natalia, and Jon.

Last but not least, I deeply thank all my family, for their affection and support throughout this entire path. To my parents, and my grandparents, for everything they taught me and for always being a reference to me. To my brothers, for bringing all their energy and making my days brighter.

"If knowledge can create problems,
it is not through ignorance that we can solve them."
(Isaac Asimov)

ABSTRACT

The present work exhibits the development and implementation of an innovative statistical control plan applied to the production process of invisible zippers, within the context of a project developed by the NOVA School of Science and Technology for a zipper producer. The plan focused on two main axes: reevaluating currently applied sampling plans and identifying critical characteristics in the various stages of the process, with the development of a proposal for control charts for each of the stages and implementation of the respective charts. An important part of the plan was the implementation of a design of experiments to optimize critical processes.

Consequently, an integrated approach was implemented to define and solve the problem. At first, a complete definition and description of the process was executed through a visual representation with flowcharts. Then, critical points of the process were identified, which led to a preliminary implementation of control charts, planification of a design of experiments, and execution of several hypothesis tests.

Even though, as of the redaction of this study, no improvement on the process was achieved, several crucial conclusions were reached over its behavior following the implementation of the statistical tools. Some important conclusions were the out-of-control state of the process on some important characteristics, and strong presence of internal variability in the process. As a result, a design of experiments was considered the best approach for improvement, and its full planification has been achieved, as it is currently being performed.

As for the sampling plans, a necessity to reduce end-of-line inspections was identified and is expected to be enabled by the improvements arising from the design of experiments. On the other hand, the reception sampling plan was identified as insufficient, and is to be reviewed.

Keywords: Quality Improvement, Design of Experiments, Statistical Process Control, Control Charts

RESUMO

O presente trabalho expõe o desenvolvimento e implementação de um plano de controlo estatístico inovador aplicado ao processo de produção de fechos invisíveis, no contexto de um projeto desenvolvido pela Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa para um produtor de fechos de correr. O plano centrou-se em dois eixos principais: a reavaliação dos planos de amostragem atualmente aplicados e identificação de características críticas nas várias fases do processo, com o desenvolvimento de uma proposta de cartas de controlo para cada uma das fases e implementação das respetivas cartas. Uma parte importante do plano foi a implementação de um desenho de experiências para a otimização de processos críticos.

Consequentemente, foi implementada uma abordagem integrada para definir e resolver o problema. No início, foi realizada uma definição e descrição completa do processo através de uma representação visual com fluxogramas. De seguida, foram identificados pontos críticos do processo, o que levou à implementação preliminar de cartas de controlo, planificação de um desenho de experiências e execução de vários testes de hipóteses.

Apesar de, à data de redação deste estudo, não se ter alcançado uma melhoria do processo, alcançaram-se várias conclusões cruciais sobre o seu comportamento. Algumas conclusões importantes foram o estado fora de controlo do processo em certas características importantes, e a forte presença de variabilidade interna no processo. Como resultado, o desenho de experiências foi considerado a melhor abordagem para a sua melhoria, e a sua planificação completa foi efetuada, sendo que as experiências se encontram de momento a decorrer.

Quanto aos planos de amostragem, foi identificada a necessidade de reduzir as inspeções de fim de linha. Por outro lado, o plano de amostragem de receção foi identificado como insuficiente, e deverá ser revisto.

Keywords: Melhoria da Qualidade, Desenho de Experiências, Controlo Estatístico do Processo, Cartas de Controlo

INDEX

FIGURE INDEX

TABLE INDEX

LIST OF ABBREVIATIONS

ANOVA	Analysis of Variance
ANSI	American National Standards Institute
ARL	Average Run Length
ASQ	American Society for Quality
ASTM	American Society for Testing and Materials
CCD	Central Composite Design
CL	Central Line
CSS	Civil Communications Section
CUSUM	Cumulative Sum
CWQC	Company-Wide-Quality-Control
DoE	Design of Experiments
DMAIC	Define, Measure, Analyze, Improve and Control
EWMA	Exponentially Weighted Moving Average
FMEA	Failure Mode and Effect Analysis
ISO	International Organization for Standardization
JUSE	Union of Japanese of Scientists and Engineers
LCL	Lower Control Limit
OFAT	One Factor At a Time
PDCA	Plan, Do, Check, Act
QFD	Quality Function Deployment
R	Range
RSM	Response Surface Methodology
S	Standard Deviation
SCAP	Supreme Commander of the Allied Powers
SPC	Statistical Process Control
TQM	Total Quality Management
UCL	Upper Control Limit
VBA	Visual Basic for Applications

1. INTRODUCTION

1.1. Background and Motivation

Industrial producers have been pioneers in the use of statistical quality control tools, to continuously improve the quality of their products and processes, and guarantee conformance with established standards. In fact, statistical control tools and methodologies allow to evaluate processes, by observing their stability, computing their parameters, and analyzing whether they satisfy the established standards. If it is determined that the process is not operating at a sufficiently high level, these tools and methodologies can also help improve the process. They can do so by reducing its variability, increasing its robustness and/ or adjusting its medium output to a desired level, among others, which in term results in a decrease of nonconformities and increase of client satisfaction.

The present case study is associated with NOVA School of Science and Technology, who provided a research grant to work within the project. This project was developed by the NOVA School of Science and Technology for a zipper producer.

The company presents a strong culture of quality and continuous improvement. Therefore, the goal of the project was to elaborate an innovative statistical control plan applied to the production process of invisible zippers, to improve the process and ultimately reduce nonconformities. This goal was achieved through the implementation of several statistical quality control tools and methodologies, namely control charts, design of experiments, and hypothesis testing.

1.2. Objectives

The main objective of this work is to elaborate an innovative statistical control plan applied to a production process of invisible zippers. The implementation goal of this plan is to improve the process and ultimately reduce nonconformities.

The plan will focus on two main axes. The first one is the reevaluation of the currently applied sampling plans, from the raw materials reception to shipping, with particular emphasis on the Quality Wall. The second axis focuses on the identification of critical characteristics in the various stages of the process, with the development of a proposal for control charts for each of the stages and implementation of the respective charts.

The approach will be divided into two parts. One consists in the development and implementation of a statistical process control plan for the production process. This part includes the review of the current sampling plans and implementation of control charts. The other involves the utilization of design of experiments for the optimization of processes that may be considered critical. For an integrated approach, these two parts will go hand in hand and complete each other.

Therefore, this study will cover the main statistical quality control tools and methodologies utilized towards the process improvement, through exhaustive research on the subject. It will then cover their practical implementation to the process and an analysis of the results.

1.3. Research Process

At first, it is important to gather a full understanding of the process, its components, and the problem to be able to tackle it in the most appropriate way and properly apply the statistical quality control tools and methodologies. Consequently, the process will be analyzed and thoroughly described, with the close cooperation of the company team, composed of engineers and management of the company. This will be undertaken through presentation and clarification of questions by the company team, analysis of documentation, as well as by direct process observation. This stage will be supported by the utilization of several quality tools, such as flowcharts and Pareto charts. The results will then be validated with the company team.

Then, there will be a definition of which tools to utilize or methodologies to implement, as well as where on the process. Control charts will be implemented to assess the current state of the process, via the study of its parameters and variability, which will allow to observe whether the process is in statistical control or if it needs adjustment. Design of experiments will then be utilized to improve the process as needed, by determining the ideal configuration of inputs/ parameters at certain stages of the process. Hypothesis testing may also be utilized to clarify simpler assumptions, if required. An inspection plan might be implemented if a necessity for the revision of a sampling plan is identified and agreed upon.

Finally, there will be an evaluation of the results, to assess the implementation's impact and draw future recommendations that can further help improve the company's process.

1.4. Document Structure

The present dissertation is divided into 5 chapters, that intend to provide the reader with the necessary context and demonstrate the practical work that was elaborated.

In chapter 1, there is a small introduction of the context within which the work took place and motivation that led to its execution. Then, the objectives and research process utilized to achieve them are presented. Finally, the document structure is described.

In chapter 2, a thorough and exhaustive literature review is undertaken, to give the reader an explanation of every theoretical concept needed to understand the practical implementation. There is a presentation of the history of quality and its definition by some of the most important figures in its history, who developed it and disseminated its use. Then, there

is a review and theoretical explanation of every statistical quality control tool and methodology utilized during the study.

In chapter 3, the company is briefly presented, as well as its culture, and the products to which the tools and methodologies were applied. A summarized description of the processes, and their verification/ inspection points is also provided.

In chapter 4, there is a description of all the steps that were executed for the practical implementation of hypothesis tests, control charts and design of experiments, as well as other tools that might have been utilized. It also contains an explanation of every assumption, and presentation and interpretation of results.

Chapter 5 contains all the conclusions drawn from this work, as well as some future recommendations for the continuation of the work developed during the case study.

2. LITERATURE REVIEW

2.1. Quality Management

2.1.1. What is Quality?

Quality is an abstract concept that has many definitions. Some of the most prominent figures in the evolution of quality define it as the adaptation of a product to customer needs (Juran 1998), conformance to requirements (Crosby 1980), "the loss a product causes to society after being shipped, other than losses caused by its intrinsic functions" (Taguchi and Clausing 1990), or the "degree of requirements satisfaction given by a set of intrinsic characteristics" (ISO 9000:2015).

The perception of what quality is varies between the consumer and the producer. Consumers will evaluate quality based on functional characteristics, price, and appearance, while producers evaluate it based on technical requirements, quality assurance, and technical assistance. How quality is measured also varies across phases of the product lifecycle. First, projects must integrate needs and expectations of customers (Conception Quality). From there, mensurable specifications must be designed and products/ services must comply with these specifications (Manufacturing/ Service Delivery Quality). Then, there must be a measure of performance of the product/ service compared to customers' expectations (Quality in Use). Finally, it is important to measure the efficiency in contact with clients (Relationships' Quality) (Pereira and Requeijo 2012).

For Juran (1998), Quality is the "fitness for use" of the product/ service. All human institutions produce goods or services for human beings. Users are the human beings who receive the benefits of the product, either by consuming it or using it in a process to create another product. As such, products/ services should respond to the overall needs of users in price, delivery date, and fitness for use.

Fitness for use is defined by characteristics or features users consider beneficial such as the taste of food, lifetime and beauty of clothing, timeliness of public transports, etc. Juran defines quality characteristics as all features of the product, materials or processes needed to achieve fitness for use (Juran 1998).

Juran believed that quality costs could not be reduced to zero, and there was a balance where the costs of failure added to the cost of prevention was minimal. This meant the ideal

was not having a 100% conforming product, as it implies large prevention costs, but rather balancing these with the costs of failure (Juran 1998).

Crosby (1980), a notable driver for Quality in the United States of America in the 1970s and 1980s, defined it as conformance to requirements. "If a Cadillac conforms to all the requirements of a Cadillac, then it is a quality car. If a Pinto conforms to all the requirements of a Pinto, then it is a quality car." (Crosby 1980). This perfectly defines Crosby's vision on quality, as he thought quality and luxury of a product were unrelated. He also estimated that the common understanding that quality is immeasurable was an illusion. For him, products have technical requirements and the conformity to these requirements defines the quality of the product. These requirements need to be clearly defined and measurements should be undertaken continuously to ensure conformance to said requirements.

As for Taguchi, quality is measured by the costs that quality problems cause to society. Taguchi considers that a problem which occurs after product delivery will always cost more than having good quality, as failures imply many costs such as transportation of the product for reparation, reparation itself, and loss of reputation. Taguchi considers quality as a virtue of the design, and that a meticulously designed product will be more robust and durable, adapting to the worst conditions of use. During the production phase, the emphasis should not only be on keeping the process within certain tolerance limits, but also constantly seeking to reduce these tolerances and have products with lowest possible variation from the target (Taguchi and Clausing 1990).

2.1.2. History of Quality

Although the development of reflection around quality has been major during the last century, the concept was already vastly utilized during the history of mankind. There are historic records of early civilizations that show a care in selecting the appropriate materials for given tasks and definition of appropriate dimensions for several tools (Pereira and Requeijo 2012).

Quality was always an important factor in trade. In fact, for a product to sell, it has always been important for it to correspond to customer needs. In early stages of population development, craftsmen embraced all the production process, from product conception to selling, including the inspection process. As population clusters grew, demand increased, and craftsmen workshops appeared. These were managed by masters who verified the work of assistants who, in turn, verified the work of apprentices (Pereira and Requeijo 2012). Craftsmen started organizing in unions, denominated guilds, by the end of the 13th century. These guilds developed rules for product and service quality. They inspected the products and marked flawless ones with a specific mark or symbol, which worked as a proof of quality. Most craftsmen also had their own mark, which represented their reputation. Together, these marks/ symbols were used as a proof of the product's quality by customers. At that time, clients played an active role in product inspection during the act of buying (ASQ, n.d.; Fisher and Nair, 2009).

The industrial revolution, with the rise of mass production, totally changed how production was executed. There were no more craftsmen, as they became factory workers, and

started working on specialized tasks. From that point on until the end of the 19th century, quality management was a production function, and inspection was mostly performed at the end of production lines to guarantee products were shipped with decent quality. This changed with the works of Frederick Winslow Taylor, who invented the Taylor system, a new management approach that separated production from planning to increase productivity. However, the separation of quality responsibility from production resulted in a decrease in quality levels, as it further diminished the importance of inspection departments (Pereira and Requeijo, 2012; ASQ, n.d.; Fisher and Nair, 2009).

During World War I, a vast amount of military equipment was failing. As such, studies were undertaken to find the source of the problem, reaching the conclusion that specifications compliance was a big problem. This propelled quality into the Inspection Phase, in which size and importance of quality inspection departments grew largely and inspections were performed with higher frequency (Pereira and Requeijo 2012).

Between the two World Wars, there was a development of statistical techniques that enabled quality control by sampling, which allowed for quality control during production, instead of controlling quality on the final product. This marked the phase of Statistical Quality Control. Walter A. Shewart, who worked at Western Electric, employed statistical techniques to control processes and minimize defective products, and developed the concept of Statistical Quality Control and control charts, as well as their sampling procedures. He also characterized the two sources of variation: common cause and special cause variation. As such, Shewart's works were fundamental for the development of the concept of quality as we know it today (ASQ, n.d.; Fisher and Nair, 2009).

Harold F. Dodge and Harry G. Romig also had a significant contribution, with the development of acceptance sampling and the concept of consumer's and producer's risk. Acceptance sampling constituted an alternative to the inspection of every product and is now mainly utilized in raw materials and final product inspection (Pereira and Requeijo, 2012; Fisher and Nair, 2009).

Tomas Bata is another individual who is cited as an important figure in the development of quality in the beginning of the 20th century. He created the Bata-System of management and introduced changes in management philosophy and techniques. He practiced quality management in his enterprises before World War I, and Bata enterprises and Japanese executives already shared contacts and exchanged knowledge before World War II (Fisher and Nair 2009).

Despite all these developments in the beginning of the century, there was a small impact on American companies, as these contributions were implemented by few (Pereira and Requeijo 2012).

During World War II, sampling inspection found an increased usage to answer the need of diminishing time and resources associated to inspection, while still maintaining a high level of quality. The US military developed sampling tables for their suppliers, which are the basis of the sampling tables used today, and also offered training in Walter Shewart's statistical quality control techniques (ASQ n.d.).

After the War, there was a slow development of quality in the western countries. In 1946, the American Society for Quality Control (current American Society for Quality) was created. During the 1950s, the importance of product durability increased, which led to more analysis towards improvements in reliability of products. Many companies started to have departments of quality control that, on top of the normal inspection tasks, applied statistical techniques in response to the increasing complexity of production processes. Another important development of that decade was the introduction in the USA of Total Quality by Armand Feigenbaum (Pereira and Requeijo 2012).

A new phase of Quality Assurance, where it was important for organizations to prove that they could satisfy clients requirements, started and prevailed until the 1980s. That led to an emphasis on the quality of work methods, documentation, and control methods (Pereira and Requeijo 2012).

The development of quality in Japan was vastly different after the War, and greatly contributed to its development worldwide. Japan was going through an economic crisis, as the war left a wide extent of the country in ruins. There was no equipment, few companies functioning, and much of the top management either dead or in jail. As such, there was a need for high quality products at the lowest possible costs. In 1945, the Supreme Commander of the Allied Powers (SCAP), led by General Douglas McArthur, was tasked with the process of post-war reconstruction. The Civil Communications Section (CCS) of the SCAP was set up to establish a communications industry in Japan. Homer Sarasohn and Charles Protzman, working for the CCS, concluded that there was a need to teach company leaders about industrial management. Thus, they prepared the Principles of Industrial Management manual (Sarasohn and Protzman, 1948 as cited in Fisher and Nair, 2009), and started the CCS Management Seminars based upon it (Fisher and Nair 2009).

The Union of Japanese of Scientists and Engineers (JUSE) was founded in 1946 and led by Ichiro Ishikawa. This structure was also fundamental in the development of quality in Japan in the post-war era, and taught many courses, often collaborating with the CSS. There were some disagreements between the two entities, as the JUSE wished to start teaching statistical quality control early while the CSS judged it was still inappropriate. For Sarasohn and the CSS, statistics were merely a tool in the implementation of a good quality system. Indeed, he believed it was essential for company leaders and managers to understand quality as a whole, including all of its management aspects, before diving deeper into statistical concepts. He believed, as well, that Japanese leaders and managers were not yet ready to grasp the statistical concepts before evolving in other areas. As a result, instruction on statistical control tools were delayed and the first courses took place without mention to these subjects, which were later introduced. Other important figures in the development of quality in Japan were Deming, who was invited by the CSS and conducted several courses for top management in Japan, and Juran, whose work was first disseminated through the Japanese community by Deming and later came in person to teach seminars for top and middle management (Fisher and Nair 2009).

As a result, there was a vast implementation of quality principles throughout Japanese companies, at all levels of organizations. Japanese companies adopted a strategy called

Company-Wide-Quality-Control (CWQC), based on the Total Quality principles. Japanese products became highly competitive, possessing high quality while remaining at low costs. That propelled them and, around the beginning of the 1970's, they even became references in the global market, as they presented higher quality than their competitors. The rest of the world later benefited from the advancements of quality achieved in Japan. Genichi Taguchi, a Japanese engineer, was crucial in the development of tools such as the Design of Experiments (DoE) or the Taguchi Methods. Kaoru Ishikawa was also very important, as he emphasized a set of tools that are nowadays known as the seven basic quality tools. The JUSE also developed its set of tools, the seven new management and planning tools, oriented towards innovation, communication and planning (Pereira and Requeijo, 2012; Fisher and Nair, 2009; ASQ, n.d.).

Towards the end of the 20th century, globalization and technological evolution transformed markets, demand, and corporate culture. Globalization led to an increased competition and to increased expectations from customers. Companies were forced to focus more on customer satisfaction, as customer loyalty became an important part of business. These changes led to higher preoccupation with quality. An event that exacerbated the focus on quality, in the West, was an NBC news report of 1980 titled 'If Japan Can ... Why Can't We?'. Consequently, there was a big rise in the quality management consulting activity from that point on (Pereira and Requeijo, 2012; Fisher and Nair, 2009).

This change of scenery led to an increased utilization of several tools, techniques and methods that had been developed throughout the 20th century, such as the 14 quality tools, Quality Function Deployment (QFD), Benchmarking, Failure Mode and Effect Analysis (FMEA), DoE, Taguchi Methods, or Statistical Process Control (SPC) (Pereira and Requeijo 2012).

In the beginning of the 1980s, many companies in western countries started implementing Total Quality Management (TQM). TQM is a management philosophy based on the notion that every person in the organization plays a role in the quality of the products to answer to or exceed expectations from customers and stakeholders (Pereira and Requeijo 2012). It is a system aimed at the management of enterprises, which assembles principles, frameworks and a plan for implementation (Fisher and Nair 2009).

One of the main components of TQM is continuous improvement, a systematical process that allows for the consistent and gradual concretization of goals/ objectives set by the organization, and is based on the PDCA cycle (Plan, Do, Check, Act). The utilization of statistical techniques plays an important role in the implementation of continuous improvement. Both DoE and Taguchi Methods are useful tools to determine which controllable inputs significantly influence the measured quality characteristics (Plan). They also allow to determine the desired level of the inputs to reduce quality characteristics' variation and adjust their average to the desired value (Plan). After determining the desired levels for the controllable inputs, experimenting these values in the real process (Do) and observing the results (Check), statistical control can be implemented to solidify the improvements (Act) (Pereira and Requeijo 2012).

In 1987, the ISO 9000 series of standards, the first international reference for Quality Systems, were created. They underwent some modifications in 1994, 2000, and 2015 but are still used today and are important certifications for Quality Management Systems, the system that is responsible for quality within an organization. The certification assures that this system comprehends the required organizational structure, responsibilities, processes, and resources to implement a satisfactory quality policy (Pereira and Requeijo 2012).

In the same year, Six Sigma philosophy was developed by Motorola. It later disseminated in North America in the mid-90's. Hahn et al. (1999) define Six Sigma as a "highly disciplined and statistically based approach for removing defects from products, processes, and transactions" which involves every person in the organization. It focuses on reducing process variation and increasing process control, aiming to reduce defects below a level of only 3.4 defects per million opportunities. One can notice that maintaining specification limits six sigma away from the average of a process with a normally distributed variation does not yield 3.4 million defects per million opportunities. However, for this calculation, an incontrollable movement of 1.5 sigma of the average of processes on the long term is assumed. To achieve such a low number of defects, Six Sigma experts rely on the use of the DMAIC (define, measure, analyze, improve and control) approach and a set of tools such as SPC, FMEA, process mapping, among others. Six Sigma originated its own terminology, by attributing levels to its practicians, such as Champions, Master Black Belts, Black Belts, Green Belts, and Yellow Belts (Hahn et al., 1999; ASQ, n.d.).

Some practicians consider the Six Sigma methodology similar to, or even an evolution of TQM. Contrastingly, others consider it as only a problem-solving approach, distinct from TQM, which is a system that includes statistical tools and problem-solving approaches, yet is a broader, complete system to manage enterprises (Fisher and Nair 2009).

Six Sigma is often combined with Lean, which focuses on reduction of waste, as well as work standardization and flow. The Lean Six Sigma philosophy focuses on process improvement and aims to reduce variation, waste, and cycle time to create a competitive advantage. It seeks the involvement of all members of the organization and values defect prevention over defect detection (ASQ n.d.).

2.1.3. Current State of Quality

Nowadays, quality has greatly expanded. It is not only utilized in the industry anymore, with a wide utilization in the service industry. It has expanded to other fields of study, being not only taught to statisticians and engineers, but also to economists, managers, sociologists, medical staff, etc., and has shown useful in diverse areas such as innovation, IT knowledge management, and more. The involvement in quality throughout organizations has also largely increased, driven by the quality philosophies implemented over time which highlighted the importance of participation of each member of the organization for the improvement of quality. As a result, responsibility and knowledge about quality have spread through organization, with members of every area possessing knowledge on quality tools and methodologies (Pereira and Requeijo 2012).

There remains, to manage quality, a considerable utilization of TQM and Six Sigma, as well as many other techniques and methods, such as the 14 quality tools, SPC, DoE, to name a few (Pereira and Requeijo 2012).

Quality is nowadays highly recognized and awarded, as there are a considerable number of international and national awards, as well as levels of recognition for quality. Some of the most important quality awards are the Deming Prize, created in Japan by the JUSE in 1951, the Malcolm Baldridge National Quality Award, created by the US Congress in 1987, or the Australian Quality Awards, created in Australia and later renamed Australian Business Excellence Awards. These are given to both big companies and SMEs in several sectors and reward companies that show a good level of quality and implementation of quality methodologies (Pereira and Requeijo, 2012; Fisher and Nair, 2009).

2.2. Quality Tools

Quality tools are a set of 14 tools that were developed during the 20th century and have had a big contribution for the structured solving of several quality problems, showing useful for the continuous improvement of processes. Some of them help to identify improvement opportunities, others to identify and eliminate non-value added activities, or to reduce process variability (Pereira and Requeijo 2012).

The 14 quality tools are divided in two subgroups: basic quality tools, and quality management and planning tools. The 7 basic quality tools, that were first emphasized by Kaoru Ishikawa, are the cause-and-effect diagram (also called Ishikawa or fishbone diagram), check sheet, control chart, histogram, Pareto chart, scatter diagram, and stratification. Stratification is commonly replaced by the flowchart or run chart (ASQ n.d.). The 7 quality management and planning tools, which were developed by the JUSE, are the affinity diagram, interrelationship diagram, tree diagram, matrix diagram, matrix data analysis, arrow diagram, and process decision program chart (ASQ n.d.).

Montgomery (2012) presents his own list of tools, which he calls the magnificent seven, with a couple differences to the 7 basic quality tools. His list includes the stem-and-leaf plot, and the defect concentration diagram, while it excludes the flowchart. He emphasizes the importance of these tools in the context of a DMAIC implementation.

In the following sections, there will be a short introduction of all quality tools that are used in the case study. For more information on the remaining basic quality tools, see Ishikawa (1976), and on the remaining quality management and planning tools, see Tague (1995).

2.2.1. Flowchart

A flowchart is a visual tool defined as "a picture of the separate steps of a process in a sequential order" (Tague 2005). It is a generic tool that can be utilized to represent all types of processes, from a manufacturing process to a project plan. A flowchart generally includes sequences of actions, inputs and outputs of the process, as well as decision points. It might also include further information, such as people or time involved at various steps, or process measurements. It allows to develop an understanding of the process in study, and to build a common ground for communication, as it is easy to understand (Tague 2005).

As such, within the scope of statistical process control, it can prove useful in the beginning of the implementation of statistical quality control tools or methodologies in a process, to better define it and then make the best decisions regarding what to implement, and where. An important step to build a flowchart is to gather opinions from all personnel involved in the process (notably operators and supervisors, but also suppliers and customers) that is under study, and to verify the final flowchart with them, to ensure no mistake is made. This is crucial, as the flowchart might define the rest of the SPC implementation.

For more detailed information on different types of flowcharts, see Tague (1995)

2.2.2. Check Sheet

A fundamental factor when implementing any statistical control technique is to utilize correct data. If the data is not correct, results will not be either, however good might be the implementation of the techniques. Henceforth, the use of check sheets is particularly relevant, as it is a tool that allows to gather meaningful data, adequate to the goals for which that data is collected (Ishikawa 1976).

Therefore, check sheets are useful to collect data in the early stages of process improvement. They can be used to summarize all historical defect data on the product or process studied. By gathering data in a time-oriented manner, it helps to identify trends that can lead to root causes of the problems (Montgomery 2012). Check sheets both synthetize the data for further analysis and are easy to use by operators (Ishikawa 1976).

2.2.3. Pareto Chart/ Pareto Diagram

The Pareto chart, or Pareto diagram, is a histogram of attribute data. Within the scope of statistical process control, it is often used to classify defects. It may be used in the sequence of check sheets implementation. A Pareto chart classifies data according to frequency, by listing the item with higher frequency on the left and the item with lower frequency on the right. The other items are arranged in a descendent order of frequency. It helps synthetize data and gain visibility over what are the most frequent problems. These can, then, be the ones towards which a solving attempt is directed in a first approach (Ishikawa, 1976; Montgomery, 2012).

Pareto charts, however, do not identify the most important defects. A solution to this problem can be the use of a weighting scheme to modify the frequency count, or to increment the Pareto chart with another cost or exposure Pareto chart. For further discussion on this matter, see Montgomery (2012).

2.2.4. Tree Diagram

A tree diagram starts with one item that sequentially branches into one or more items. It helps to break down one initial idea or problem into finer levels of detail. Its sequential approach helps to go, step by step, from generalities to specifics. It can be used for many purposes, such as finding the root cause of a problem, developing logical steps to reach an objective, or evaluating issues against potential solutions, to name a few. As does the flowchart, it allows to build a common ground for communication (Tague 2005).

Within the scope of statistical process control implementation, it can prove effective to reveal potential causes, once a specific defect or issue has been identified, for example via the utilization of a Pareto chart. The cause-and-effect diagram, or Ishikawa diagram, can also serve this purpose. By first identifying cause categories and then sequentially identifying specific causes, potential root causes are finally reached. It can also contain further detail, and rank these causes regarding plausibility of impacting the problem (Montgomery 2012).

For more detail on how to build a tree diagram, and different purposes it can be used for, see Tague (1995).

2.2.5. Control Chart

Control charts are used for process control. They plot controlled process/ product characteristics over time. In opposition to simpler check sheets or histograms, they allow to study not only the changes in data over time, but also the changes introduced by modifications in the factors that affect the process, such as materials, workers, or equipment. Their analysis can show whether the process is out-of-control or presents abnormal behavior, giving valuable indications on the necessity of corrective actions. The use of control charts allows for distinction between common cause variation and special cause variation, as well as monitoring of processes once special cause variation has been eliminated and the process is considered in statistical control (Pereira and Requeijo, 2012; Ishikawa, 1976).

Control charts are further described in section 2.4.

2.3. Hypothesis Testing

A statistical hypothesis is an assumption about a population, generally in regard to a parameter of that population. Hypothesis tests (also called tests of hypotheses or tests of significance), are used to verify whether those assumptions are correct. They were mainly created to avoid the search and attribution of causes to a variation where there was simply random variation, not needing a cause to be justified. The parameters of a process are generally unknown and will often change over time. An example of this can be a measurement of the weight of a given sample of products, which is bigger than the mean. Hypothesis tests can be useful in this regard, as performing a test would allow to uncover whether that increase is due to a real variation in the process or if that is unlikely. Furthermore, hypothesis tests can also be utilized to compare populations parameters with each other and take several deductions, such as if their means are different, for instance (Dudewicz, 1998; Montgomery, 2012).

The assumption of the hypothesis test is denominated null hypothesis and denoted H_0. The test will determine, through the analysis of a sample of data, whether the null hypothesis is rejected or not. Samples must be evaluated carefully, as an average of observations that differs from the null hypothesis might just be part of the population and its natural variation. On the other hand, one result might correspond to many different hypotheses. These are managed by addressing the two types of error that might occur while performing the test. The first error is the type I error, denoted α, which consists of the rejection of the hypothesis when it is true and should not have been rejected. The second is the type II error, denoted β, associated

with not rejecting the hypothesis when it is false and should have been rejected (Dudewicz 1998).

Before performing a hypothesis test, the significance level is determined. This significance level is the level of risk of a type I error that the practitioner is willing to take. Commonly, tests are performed with a significance level of $\alpha = 0.05$, and often with $\alpha = 0.1$. Then, an acceptance region is determined, which is where $100 * (1 - \alpha)$ percent of the population falls. It the average of the sample is located within that region, the null hypothesis is not rejected. If the opposite happens, it is rejected (Dudewicz 1998).

The type II error is harder to determine, and tests are generally performed without its knowledge. It is important to note, nonetheless, that an increase in the sample size decreases the risk of type II error (Toutenburg and Shalabh 2009). It is possible to control the type II error, however that requires not fixing the sample size in advance, which is more complex. For more information on this topic, see Dudewicz (1998).

It can be complicated to draw conclusions from hypothesis tests. One of the outcomes is fairly indicative. If the null hypothesis is rejected, there is a $100 * (1 - \alpha)$ percent probability that it is, indeed, incorrect. However, the opposite outcome of not rejecting the null hypothesis does not necessarily prove that the assumption was correct, as multiple hypotheses could be accepted for the same sample, even though the population has only one real average value (Dudewicz 1998).

Therefore, to properly evaluate the results of the test, two approaches can be followed. First, confidence limits on the sample result can be calculated. If the confidence interval is large, the lack of rejection of the null hypothesis must be handled with caution, as there is a big risk that the assumption is not true. If it is small, then there can be a strong confidence that the assumption is, in fact, true. For more information on the calculation of confidence limits and interval, see Dudewicz (1998).

Another approach is to compute the p-value, which is indicative of how likely it is for the sample to assume a value equal or further to the assumption. As such, it gives an indication of how likely it is for the assumption to be true (Dudewicz, 1998; Montgomery, 2012).

Hypothesis tests can be either one-tailed or two-tailed. One-tailed tests are used to determine whether the studied parameter is higher or lower than a given number, or if a population's parameter is higher than the one from another population. Two-tailed tests are used to determine whether the studied parameter is equal to a given number, or if the parameter is different between two populations (Dudewicz, 1998; Montgomery, 2012).

There are many different types of hypothesis tests. Some are aimed at individual samples, others at two samples, or more than two samples. Hypothesis tests for individual samples are used when there is a need to know the parameters of a population. They can be used to infer on the mean of the population, its variance, proportion, to name a few. Hypothesis tests for more than one sample are utilized when there is a need to compare several populations with each other. They can be used to infer on the differences between means, variances, proportions, etc. of the populations. When comparing populations, it is also important to consider whether samples are matched or independent.

In this section, there is a summarized approach of each hypothesis test that was utilized during the case study. For more details on other types of hypothesis tests, see Dudewicz (1998) and Montgomery (2012)

2.3.1. T-Test for Two Populations Means with Matched Samples and Unknown Variance

The t-test for two populations means with matched samples is utilized for continuous variables defined by a normal distribution. Samples are matched when each observation of the first sample is matched with the same observation of the second sample (i.e. the first observation of the first sample is matched with the first observation of the second sample, etc.). For instance, a study where a group of people takes a medication, and there is an observation of their blood levels of a given nutrient before (first sample) and after (second sample) the treatment has matched samples. Another example applied to the industry could be the following. If there is a need for an experiment to determine the effect of two different cutting tools that need to cut a metallic plate on the porosity of the cut surface, an experience could be to cut 10 metallic plates with one tool and 10 metallic plates with another, and thus perform an independent samples test. However, the lack of homogeneity between observations would inflate the experimental error and make it harder to detect differences between samples. Another solution would be to select 10 plates and first apply the cutting process with a given cutting tool on one side of the plate, then with another cutting tool on the other side. By doing so, observations would be paired and there would be less variability, hence making it easier to detect differences between cutting tools (Montgomery 2012).

The different possible hypotheses for this test are given in equation 2.1, and the test statistic in equation 2.2.

$$\begin{aligned} &H_0: \mu_1 = \mu_2 \;\; vs. \;\; H_1: \mu_1 \neq \mu_2 \\ &H_0: \mu_1 \leq \mu_2 \;\; vs. \;\; H_1: \mu_1 > \mu_2 \\ &H_0: \mu_1 \geq \mu_2 \;\; vs. \;\; H_1: \mu_1 < \mu_2 \end{aligned} \tag{2.1}$$

$$t_0 = \frac{\overline{D} - (\mu_1 - \mu_2)}{\frac{S_D}{\sqrt{n}}} \tag{2.2}$$

$D = X_1 - X_2$ - Difference between populations
X_1, X_2 – Distributions of populations 1 and 2
μ_1, μ_2 – Means of populations 1 and 2
S_D – Sample standard deviation for population D
n – Sample size

Then, the value of t_0 can be compared with $t_{\alpha/2;n-1}$ in the case of a two-tailed test, or $t_{\alpha;n-1}$ in the case of a one-tailed test. In a two-tailed test, if $t_0 > t_{\alpha/2;n-1}$, the null hypothesis H_0 is rejected. In a right tailed test, if $t_0 > t_{\alpha;n-1}$, the null hypothesis is rejected, while in a left tailed test, it is rejected when $t_0 < -t_{\alpha;n-1}$. Another option to evaluate whether the null hypothesis is rejected is to compute the p-value, which can be done using equation 2.3. If the p-value < α, the null hypothesis is rejected.

$$P = 2 \times P(T_{n-1} > |t_0|)$$
$$P = P(T_{n-1} > t_0) \quad (2.3)$$
$$P = P(T_{n-1} > |t_0|)$$

2.3.2. Analysis of Variance for Means Comparison

There are situations where one should test more than two hypotheses. This happens often in process improvement, where for instance there will be a test between five types of raw material to see which one benefits the process the most. The analysis of variance with fixed effects model, or model I, can be used in these situations.

Another type of analysis of variance model is the analysis of variance with random effects, or model II, which is utilized when the specific levels that are used for the factor are of no particular interest and chosen at random, as the objective is to define the total variability. For more information on this model, see Toutenburg and Shalabh (2009).

The first step of the ANOVA is collecting the data. Each level of the parameter of study will have n observations. The observations can be characterized by the linear statistical model of equation 2.4.

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, n \end{cases} \quad (2.4)$$

y_{ij} – jth observation of level i
μ – Overall mean
τ_i – Effect of level i
ε_{ij} – Random error component

The hypothesis of the test is given in equation 2.5. The null hypothesis is that all the effects are null, which in term signifies that all means are equal.

$$H_0 : \tau_1 = \tau_2 = \cdots = \tau_i = 0$$
$$H_1 : \tau_i \neq 0 \text{ for at least one } i \quad (2.5)$$

Let the parameter of study/ factor be denoted A. To verify this hypothesis, the next step is to compute the sum of squares of the parameter in study (SS_A), the total sum of squares (SS_T) and the error sum of squares (SS_{Error}). These computations are shown in equations 2.6 to 2.8.

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij})^2 - \frac{(y_{..})^2}{an} \quad (2.6)$$

$$SS_A = \sum_{i=1}^{a}\frac{(y_{i.})^2}{n} - \frac{(y_{..})^2}{an} \quad (2.7)$$

$$SS_T = SS_A + SS_{Error} \quad (2.8)$$

$y_{i.} = \sum_{j=1}^{n} y_{ij}$ – Sum of all observations of level i
$y_{..} = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}$ – Sum of all observations

The sums of squares are then utilized to create an ANOVA table, such as Table 2.1.

Table 2.1 - ANOVA table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
A	SS_A	$a - 1$	$MS_A = \frac{SS_A}{a-1}$	$F_0 = \frac{MS_A}{MS_{Error}}$
Error	SS_{Error}	$a(n-1)$	$MS_{Error} = \frac{SS_{Error}}{a(n-1)}$	
Total	SS_T	$an - 1$		

The null hypothesis is rejected if $F_0 > F_{\alpha;a-1;a(n-1)}$. Another approach that can be used is to calculate the p-value, with the formula of equation 2.9. If p-value < α, the null hypothesis is rejected.

$$P = P(F_{a-1;a(n-1)} > F_0) \tag{2.9}$$

In complement of this analysis, a residual analysis should also be undertaken, as it allows to assess model adequacy. The residuals from an ANOVA analysis are the difference between observations and the average for that specific level, which is shown in equation 2.10.

$$e_{ij} = y_{ij} - \hat{y}_{i.} = y_{ij} - \bar{y}_{i.} \tag{2.10}$$

The obtained residuals should be analyzed by plotting them in a normal probability plot, in a residuals vs. experiments sequence and in a residuals vs. predicted values plot. These plots allow to test if the model is correct, and if the errors are normally and independently distributed.

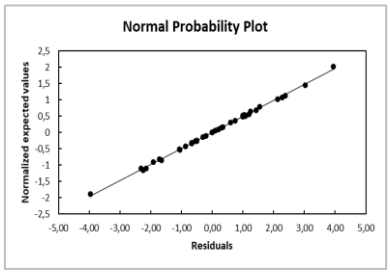


Figure 2.1.A - Normality

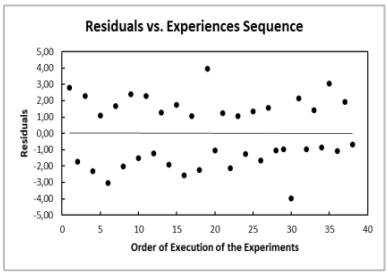


Figure 2.1.B - Independence

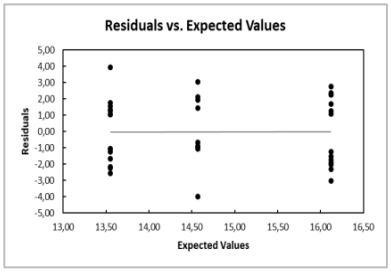


Figure 2.1.C - Homoscedasticity

Figure 2.1 - Residual plots

The normal probability plot allows to test the normality hypothesis by plotting the normalized expected values vs. the residuals. If the plotted values follow a straight line obtained through a linear regression of the data, the normality hypothesis is verified. The residuals vs. experiments sequence plot allows to test for residuals independence. If there is no correlation between the residuals and the order of the experiment, residuals are independent. Finally, the residuals vs. predicted values plot allows to test for the homoscedasticity principle by verifying whether there is any abnormal tendency on the plot. Examples of the three plots are presented in Figure 2.1 (Pereira and Requeijo 2012).

The ANOVA allows to conclude whether all means are equal or not. However, if the null hypothesis is rejected, it does not state which mean(s) is/ are different from the others. To do so, several methods can be utilized. One of them is the Fisher Least Significant Difference (LSD) method. This method can be used to test if each pair of levels is significantly different, testing $H_0: \mu_i = \mu_j$ (Montgomery 2012).

Each pair of means is significantly different if $|\bar{y}_i - \bar{y}_j| > LSD$, which can be obtained with equation 2.11.

$$LSD = t_{\alpha/2,N-a}\sqrt{MS_{Error}(\frac{1}{n_i}+\frac{1}{n_j})} \tag{2.11}$$

After doing the calculation for each pair of means, it is possible to differentiate which levels are significantly different from each other. This, in term, supports the decision of choosing which one(s) is/ are the best for the performance of the process.

2.4. Statistical Process Control

2.4.1. What is Statistical Process Control?

Dr. Walter A. Shewart of the Bell Laboratories was the originator of SPC and control charts. He created it while attempting to reduce variation in products of the Western Electric Company, and first published about it in 1924. Dr. Harold F. Dodge, a colleague of Shewart at the Bell Laboratories, conducted many other applications of control charts at the same company, and published the ASTM Manual on Presentation of Data in 1935. This manual describes many types of control charts and is currently known as ASTM MNL 7. He also played important part in the creation of the standards on control charts, which were revised into today's standards from ANSI and ASQ (Wadsworth 1998).

SPC is the application of the 14 quality tools to control process inputs (independent variables), monitor process outputs (dependent variables), and apply changes to improve the process (ASQ n.d.). SPC's most important tool is the control chart. As such, and due to there already being a sufficient definition of other quality tools in section 2.2, this section will mainly focus on control charts. The implementation of SPC helps to reduce variability and determine whether the process can produce within pre-defined specifications. It focuses on early problem detection and resolution. It is very useful nowadays, as strong competition in the industry and clients' standards require low-variability, stable processes, that produce

within the established specifications. The implementation of SPC helps to solve all these problems (Pereira and Requeijo 2012).

2.4.2. Process Variation

Shewart observed that variation occurs everywhere, including in manufacturing processes. A process always has an output that is variable to a certain degree. This variation is due to several factors that can be related to equipment, raw materials, manpower, environment (temperature, humidity, pressure, luminosity, etc.), methods (how well operations are defined, whether methods are adjusted or not to needs), or metrology (how well is the characteristic measured, related to measuring equipment and training of workers undertaking the measurements). The study of said variation and its reduction are the main motor of quality improvement, which can be greatly assisted by the use of control charts (Pereira and Requeijo, 2012; Wadsworth, 1998).

To improve a process, it is important to make a correct distinction between common cause variation and special cause variation. The first is variation that affects a process in statistical control. This variation is composed by small events which cause minor fluctuations that are not important for the output or might be uneconomic to correct. It follows a pattern and can be approximately described through statistical distributions, often following a normal distribution. On the other hand, special cause variation is made of sporadic fluctuations, generally assignable to a specific event or cause, that do not follow a statistical distribution, while also causing more significant variation than common cause variation. Whenever there is special cause variation, the process is out of control. Therefore, after correct distinction between the two types of variation, it is important to eliminate the latter. Control charts help to distinguish these two types of variation, as special cause variation is indicated by observations that fall out of the control limits or present specific patterns (Pereira and Requeijo, 2012; Wadsworth, 1998).

When special cause variation is detected, it is important to gather any useful information. This includes any significant events that might have happened before the problem occurred (lack of normal raw material, unusual noises of the equipment, etc.), the possible causes, and the actions taken before its appearance (utilization of raw material from another supplier, adjustment to the equipment or even specific movements of the operator and manner of operating the equipment before the occurrence). Its elimination generally requires operational adjustments that are not complicated. This approach not only helps to solve the current problem, but also to prevent future problems, which is a cornerstone of continuous improvement. It is advised to measure any improvement actions to confirm the improvement and identify any side effects that can result from the modifications. Recalculating control charts limits might be required, as adjustments can alter variation (Oakland 2008).

After the process is in statistical control, it is possible to estimate process parameters and determine whether it has capability to produce within specifications. Even after the process is in statistical control and producing within specifications, it is important to seek the reduction of common cause variation, in a continuous improvement approach. Common

cause variation reduction can be achieved through the use of quality tools or methods such as DoE or Taguchi methods (Pereira and Requeijo 2012).

2.4.3. Control Charts' Implementation Benefits

Control charts have a big impact in the observation of process variation. They help to distinguish common cause variation from special cause variation, to eliminate special cause variation when they are first implemented and to quickly correct issues when the process is in statistical control, by identifying special cause variation as soon as it appears. It does so by showing the amount and nature of variation in a sequential manner, keeping track of which observations were linked to that variation, which enables the interpretation of patterns and detection of changes in the process. This, in term, helps to identify the cause of variation and understand where changes must be undertaken, which can prevent disproportionate investments in uninformed attempts to correct problems (Pereira and Requeijo, 2012; Wadsworth, 1998).

As a result, control charts improve consistency, as well as prediction capacity, as they help to reach a process in statistical control. Afterwards, process parameters and capability can be determined. As the process is predictable, quality levels and associated costs are known and stable, which is valuable information both for the producer and the clients (Pereira and Requeijo 2012).

Control charts also help to prevent the occurrence of non-conforming product, by promptly detecting abnormal outputs. Its implementation is greatly beneficial for continuous improvement, by helping to identify the impact of changes in the process. This, in term, facilitates a reduction of unit costs, by reducing the amount of non-conforming product (both through prompt elimination of special cause variation and process improvement) and increasing productivity (Pereira and Requeijo 2012).

It also benefits from being a simple tool, as it can be filled by operators and easily interpreted by them. This allows for better autonomy for the operators, as they can interpret when there is a need for corrective actions and eventually perform small adjustments as required. On top of that, it provides a common language for discussions about the process, as it is easily interpretable by any stakeholder (Pereira and Requeijo 2012).

Ishikawa (1976) states that control charts can be used for purposes other than merely process control, as he describes that control charts are well suited to analyze a process and examine the nature of the influence of different factors. For example, in a comparison of two supplier's raw material, the utilization of control charts can reveal the parameters of the process with each supplier's material, which can be used in the final decision. However, some methodologies that have since been popularized, such as DoE, presented in section 2.5, are more efficient for said purpose.

The coupling of control charts with control devices and specialized software can also yield great benefits, when controlling automated processes. This enables continuous control of quality characteristics and automated analysis of data, which can include automatic generation of alerts and a preliminary analysis of plausible causes. While process control still requires the analysis of problems and identification of often unexpected variables causing an

excess in variation, this can greatly help engineers and managers in their attempt to stabilize and control processes (Wadsworth 1998).

2.4.4. Preparation for Control Charts Utilization

Not all processes are perfect candidates for SPC implementation. Its implementation is recommended on repetitive processes that produce big amounts of similar products. It is also recommended that the products or the process have easily mensurable characteristics. This will guarantee an easier implementation and increase the likelihood of the chart being representative of the process. Furthermore, it is also recommended to choose a process and characteristic with a high inspection rate and above desired rejection rate. The implementation will be more impactful on these processes, as it greatly reduces inspection needs, as well as defects rate. The utilization of a Pareto chart can be useful for that purpose, as it can help to identify which characteristics generate the most nonconforming items. Another important consideration is the point in the process where control is undertaken, which should be early enough to prevent nonconformities and additional reprocessing work (Berk and Berk, 2000; Wadsworth, 1998).

Some actions must be taken for a proper implementation of control charts. It is important to give proper training to any personnel involved in the process where SPC is to be implemented (operators, managers, etc.). This training must not only focus on teaching how to use and interpret control charts, but also on explaining the advantages of its implementation, as it is crucial that stakeholders feel involved in the process and are motivated to make the required changes in their work habits. It is also important to explain the role of each stakeholder in the implementation (Pereira and Requeijo 2012).

It is important to properly define the process, focusing on how it works, interacts with other processes, and what are the factors that influence it (personnel, equipment, materials, methods, environment, measurement system) (Pereira and Requeijo 2012). The utilization of tools such as flowcharts can be useful for this phase. As processes evolve in time and workers change, it is rare for one single person to have an exhaustive knowledge on a process. Therefore, building a flowchart with the intervention of all the personnel that work on the process (operators, supervisors, engineers, etc.) is a good option for its proper definition (Berk and Berk 2000).

This brings us to another important factor: teamwork. Good communication is critical for continuous improvement and problem resolution, as it allows for facts and data to flow properly throughout the organization. Besides interdepartmental communication, it is also important to allow all members of the organization to report issues and suggestions without fear or intimidation, and without trying to blame anyone for the presented issues. Another important aspect of teamwork is bringing together diverse skills, knowledge, and experience, as many problems require solutions that can only be achieved through the common participation of personnel with diverse competences. For instance, defining the process requires this common participation (Oakland 2008).

Afterwards, the moment/ place of the process where the implementation is made, and the controlled characteristics must be defined. These characteristics should be the most critical

for the quality of the process. It is essential to gather an understanding of what the critical parameters of the product/ component are (dimension, weight, hardness, elasticity), and what are the parameters of the process that influence it (temperature, humidity, force applied) (Berk and Berk 2000). Some aspects that must be considered during this step are the clients' requirements, areas/ parameters where a lot of non-conforming products are detected and/ or there is a low efficiency and correlation between characteristics. The utilization of techniques/ tools such as DoE, Taguchi methods or even Ishikawa diagram can be greatly useful during this phase. Then, it is important to define what information has to be gathered, but also where, how and under what conditions it must be gathered, and what measurement system (equipment and methods) is utilized to measure the chosen characteristics (Pereira and Requeijo 2012).

Another crucial action that must be taken before and during the implementation is reducing as much as possible factors that could impact process variation. For instance, there must be precaution towards not mixing batches of raw material, not undertaking any unnecessary adjustments of the equipment, or introducing any new operators (Pereira and Requeijo 2012).

The proper execution of all these steps before the implementation can avoid many problems. An especially important one is resistance to change, which can lead to a lack of effort from the personnel. The proper involvement and training of all involved personnel is generally a solution. Another problem can be difficulty in maintaining the charts or improper utilization of the charts by only detecting special cause variation when points are out of the limits and not detecting other relevant patterns. A solution that might help overcome this obstacle is the implementation of software that detects special cause variation in the charts. This reduces the interpretation responsibility on the operators and engineers, and facilitates the early resolution of problems (Berk and Berk 2000).

2.4.5. Statistical Basis of Control Charts

A control chart is a graph that shows the evolution of a statistic (ω) referring to a given characteristic over time. On this graph, the average of the statistic is represented by the central line (CL). The upper control limit (UCL) and lower control limit (LCL) define the interval within which the measurements must be located when the process is in statistical control. The pattern must also be perfectly random between the control limits. If any points outside the control limits or certain specific patterns show up, special cause variation is present, and the process is out of control (Pereira and Requeijo 2012).

Control charts may be of two types. Control charts for variables are applied to quality characteristics that are measurable on a continuous scale, and chart the central tendency and variability of the process. These provide the maximum amount of information per item inspected. On the other hand, control charts for attributes are utilized for quality characteristics that are inspected on a go no-go basis, and thus provide an inferior amount of information (Montgomery, 2012; Wadsworth, 1998). As the variables where control charts were applied in the case study were measurable on a continuous scale, there is only a more detailed description of control charts for variables, which can be found in section 2.4.6. For more information

on control charts for attributes, see Montgomery (2012), Pereira and Requeijo (2012) and Wadsworth (1998).

There are usually two phases during the implementation of control charts. In Phase I, the process' parameters are still unknown. Therefore, data is collected and analyzed to identify whether special cause variation is present and, if so, eliminate it and estimate the process parameters. Phase I is referred by Ishikawa (1976) as "process analysis" phase, as the charts built allow to observe whether an in-control state has been achieved. Phase II is the process control phase, where the known parameters are utilized to monitor process behavior (Montgomery 2012).

After the conclusion of Phase I, it is possible to calculate process capability, by using the estimation of the parameters of the process.

Wadsworth (1998) states that the center line can be a desired value. This is, however, not recommended, as the main purpose of the control chart is to analyze and improve on the current state of the process and measured characteristics, and the selection of a center line that is not representative of the average of the samples might hinder this task.

In Phase II, the final chart determined in Phase I is built (with the computed CL, UCL and LCL) and points of collected samples are sequentially added to the chart. After every addition to the chart, there is a verification of whether the points show the presence of special cause variation. If there is special cause variation, the cause must be analyzed, and corrective actions undertaken.

The limits of a Shewart control chart, assuming that the values of the statistic ω follow a Normal distribution $N\sim(\mu_\omega, \sigma_\omega^2)$, are given by equation 2.12.

$$\begin{aligned} UCL_\omega &= \mu_\omega + 3\sigma_\omega \\ CL_\omega &= \mu_\omega \\ LCL_\omega &= \mu_\omega - 3\sigma_\omega \end{aligned} \tag{2.12}$$

The limits of Shewart control charts are at a distance of $\pm 3\sigma$ from the central line. This implies that, when the process is in control, the probability of any point being between the limits is of 99.73%. Anytime a point is outside the control limits, it is assumed that special cause variation is present. Therefore, there is a 0.27% risk of false alarm, which means a 0.27% risk for a point to be outside the limits when the process is in control. This risk of false alarm is a risk of type I error, denominated α. By adopting this criterion in a process in statistical control, there will be an error approximately every 370 points. This value is denominated in-control ARL (Average Run Length). On the other hand, β is the probability of type II error, which in this case is the probability that the control chart does not indicate the process is out-of-control when it is, in fact, out-of-control for a given sample (Pereira and Requeijo, 2012; Wadsworth, 1998).

ARL is defined as the average number of data points/ samples before there is a point outside the control limits. It has been the most utilized measure to estimate the performance of a control chart. There are two types of ARL. In-control ARL (ARL_0) is the average number of points before a false alarm, while out-of-control ARL (ARL_Δ) is the average number of points, when the process is out of control, before a point indicates that the process is indeed

out-of-control. It is desirable that the in-control ARL (ARL_0) is as high as possible to minimize false alarms and that out-of-control ARL (ARL_Δ) is as small as possible to quickly identify when the process is out of control. In-control and out-of-control ARL can be calculated with equation 2.13 (Pereira and Requeijo 2012).

$$ARL_0 = \frac{1}{\alpha}$$
$$ARL_\Delta = \frac{1}{1-\beta} \tag{2.13}$$

The choice over what distance in terms of σ the control limits have to the central line is up to the user and does not necessarily need to be the indicated $\pm 3\sigma$. However, it is important to consider that an increase in distance will increase both in-control ARL and out-of-control ARL, which means big values of $\pm\sigma$ might lead to insufficient detection when the process is out-of-control. On the other hand, decreasing the distance will decrease both ARLs, which leads to a higher occurrence of false alarms. Shewart chose this balance of $\pm 3\sigma$ because it implies a small risk of looking for problems that do not exist (false alarms). Even though it also implies an appreciable risk of not detecting small shifts in the parameter under study, he considered this to be less costly than to look for problems that do not exist. However, this is not necessarily always the best balance and it must be adjusted to each case (Pereira and Requeijo, 2012; Wadsworth, 1998).

Another aspect to consider is the sampling size and frequency. The ideal would be to have frequent large samples; however, this is generally not feasible from an economic standpoint. The remaining options are small samples with high frequency or large samples with low frequency. Currently, the industry tends to prefer the first, as it allows to rapidly detect trend shifts (Montgomery 2012). The type of control charts is also a variable that influences this decision, as control charts for variables require smaller samples than control charts for attributes (Wadsworth 1998).

Furthermore, the collection of samples is also a question mark. There are two general approaches to rational subgrouping: taking samples of consecutive units of production, or taking random samples of all process output over the sampling interval. The first one is preferable if the objective is to measure variability and detect special cause variation between samples, while also providing a better estimation of the standard deviation of the process for control charts for variables. The latter is preferable when the objective is to make decisions about the acceptance of all products manufactured since the last sample (Montgomery 2012).

As previously mentioned, not only the samples that are outside the limits are considered indicators of special cause variation. Any kind of cyclic pattern or trend shift within the limits is also indicative of special cause variation. There are a set of rules that identify special cause variation in Shewart's control charts. These are presented in Table 2.2, as seen in ISO 7870-2:2013 (2013). The first four rules are denominated Western Electric Rules, as they were advocated in 1956 in the Western Electric Company's quality control program. The other rules have since been developed by several practitioners. It is important to mention that whenever an indicator of special cause variation is detected, the situation must be carefully inspected to

relate this special variation to an assignable cause. This requires experience and knowledge of the process (Montgomery, 2012; Wadsworth, 1998).

Many of these rules are based on the division of the control chart in three zones: zone A, zone B and zone C. These zones can be obtained by dividing each of the halves of the chart in three equal parts, such as shown in Figure 2.2. As can be seen, zone A is the third of the upper or lower part of the chart with is the furthest from the Center Line, while zone C is the closest to the Center Line and zone B is simply located between the two.

The use of these rules increases the ability to detect smaller shifts. There should be a careful use of these rules, however, as the use of several rules increases probability of type I error and decreases in-control ARL. Some rules can even be bad indicators of tendency shifts and increase false alarm rate, which is problematic. Therefore, it is important to carefully decide which rules are to be applied/ verified (Montgomery, 2012; Wadsworth, 1998).

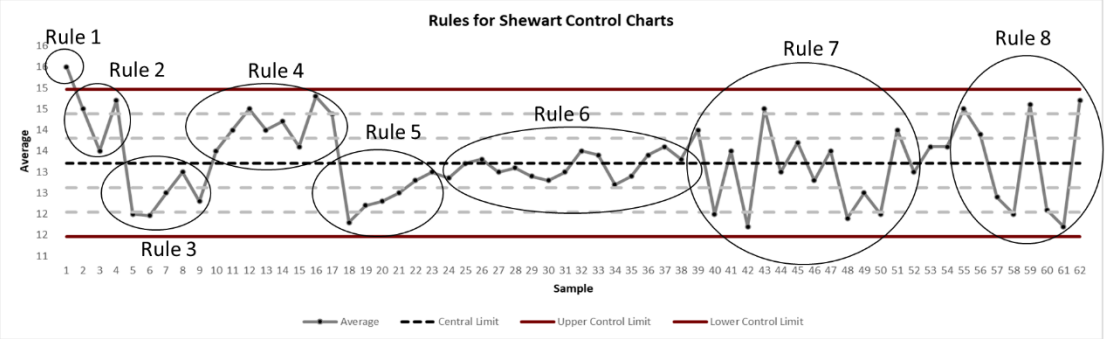


Figure 2.2 - Rules for Shewart's control charts

Table 2.2 – Rules for Shewart's control charts (Adapted from Montgomery, 2012)

No.	Rule
1	One or more points outside the limits
2	Two out of three consecutive points in zone A (on the same side of the chart)
3	Four out of five consecutive points in zone B (on the same side of the chart)
4	Eight consecutive points on the same side of the central limit
5	Six consecutive points steadily increasing or decreasing
6	Fifteen consecutive points in zone C
7	Fourteen consecutive points alternating up and down
8	Eight consecutive points outside zone C

Shewart control charts are deemed helpful in Phase I of control chart implementation, as they are simple to construct and interpret, and allow to unveil both large and sustained shifts as well as outliers. In Phase I, there is not a big preoccupation over ARL, as the biggest concern is elimination of special cause variation, and so the discussed rules are useful in this step. In Phase II, special cause variation usually results in smaller process shifts. Therefore,

the utilization of Shewart control charts is less recommended in Phase II, as they only detect big shifts. The utilization of some of the rules presented is also not a great solution, as they drastically increase the false alarm rate, and ARL is already a concern in this phase. The cumulative sum (CUSUM) and exponentially weighted moving average (EWMA) control charts are deemed to be more useful in this context.

2.4.6. Control Charts for Variables

Control charts for variables are generally either $\bar{x}$ and R, or $\bar{x}$ and s charts. $\bar{x}$ charts plot the sample average, while R charts plot its amplitude and s charts its standard deviation. However, in cases where slight shifts in the mean are relevant, CUSUM or EWMA charts might be utilized. $\bar{x}$ and R charts have had a traditionally more frequent use, as their calculations are easier and the concept of amplitude, which is simply the largest value of the sample subtracted by the smallest one, is easier to understand and explain than the concept of standard deviation. Nowadays, this is not a problem, as computers allow for easier calculations and there is a more widespread knowledge on basic statistical concepts. Both $\bar{x}$ and R, and $\bar{x}$ and s charts are useful and interchangeable for small samples (below 10 observations per sample). However, $\bar{x}$ and s charts are highly recommended when dealing with bigger samples (10 observations or more per sample), or when the sample size is variable (Montgomery, 2012; Wadsworth, 1998).

Since some of the characteristics over which control charts were implemented during the case study had sample sizes larger than 10, and to maintain coherence over all control charts, $\bar{x}$ and s charts were utilized. Therefore, there is only a more detailed explanation of calculations for these charts. For more information on other types of control charts for variables, see Montgomery (2012), Pereira and Requeijo (2012), and Wadsworth (1998).

In real context, it is rare to know μ and σ for the studied characteristic. Hence why, in phase I, they must be estimated from samples or from recent historical data, and then adapted as special cause variation is eliminated. Each sample should contain a given number of observations, taken under the same technical conditions, and at the same time or from the same lot. After special cause variation is eliminated, it is possible to reach a final estimate of process parameters when in-control.

The steps of Phase I for Shewart's control charts for variables are the following:

1) Choose the represented characteristic.
2) Develop a control plan that states the sample size, sampling frequency, measuring equipment and method.
3) Select the chart type.
4) Collect m samples of dimension n during a defined period, that must reach a total of $N \geq 100$, N being the total of measurements ($N = m \times n$). It is also important that $m \geq 20$. These samples might be from historical data.
5) Compute the statistic for each of the m samples.
6) Compute the central line and control limits.
7) Verify the existence of special cause variation.

8) Eliminate the data points that are affected by special cause variation/ show the existence of special cause variation.
9) Repeat steps 6 to 8 until there is no more special cause variation.
10) Build the final control chart when there is only common cause variation.

The estimators of μ and σ are obtained by computing the average of the averages and the average of the standard deviations of each sample, as shown in equations 2.14 and 2.15 (Ishikawa, 1976; Montgomery, 2012).

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \cdots + \bar{x}_m}{m} \tag{2.14}$$

$$\bar{s} = \frac{1}{m}\sum_{i=1}^{m} s_i \tag{2.15}$$

$i = 1, 2, \ldots, m$ - samples
$\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m$ – average of sample i
m – total number of samples

For further calculations, it is necessary to use constants that are tabulated for an ARL of 370, which means control limits of $\pm 3\sigma$. A simplified version is presented on Table 2.3, while a more complete version is presented in Appendix A. To calculate control limits for $\bar{x}$ and s charts, the formulas from equations 2.16 and 2.17 are respectively utilized.

$$\begin{aligned} UCL &= \bar{\bar{x}} + A_3\bar{s} \\ CL &= \bar{\bar{x}} \\ UCL &= \bar{\bar{x}} - A_3\bar{s} \end{aligned} \tag{2.16}$$

$$\begin{aligned} UCL &= B_4\bar{s} \\ CL &= \bar{s} \\ UCL &= B_3\bar{s} \end{aligned} \tag{2.17}$$

Calculations for charts with a variable sample size are slightly different. For further information on the topic, see Montgomery (2012).

Table 2.3 - Control chart constants (adapted from Pereira and Requeijo, 2012)

n	A	A2	A3	c4	B3	B4	B5	B6	d2	d3	D1	D2	D3	D4
2	2.121	1.880	2.659	0.7979	0	3.267	0	2.606	1.128	0.853	0	3.686	0	3.267
3	1.752	1.023	1.954	0.8862	0	2.568	0	2.276	1.693	0.888	0	4.358	0	2.574
4	1.500	0.729	1.628	0.9213	0	2.266	0	2.088	2.059	0.880	0	4.498	0	2.282
5	1.342	0.577	1.427	0.9400	0	2.089	0	1.964	2.326	0.864	0	4.918	0	2.114
6	1.225	0.483	1.287	0.9515	0.030	1.970	0.029	1.874	2.534	0.848	0	5.078	0	2.004
7	1.134	0.419	1.182	0.9594	0.118	1.882	0.113	1.806	2.704	0.833	0.204	5.204	0.076	1.924
8	1.061	0.373	1.099	0.9650	0.185	1.815	0.179	1.751	2.847	0.820	0.388	5.306	0.136	1.864
9	1.000	0.337	1.032	0.9693	0.239	1.761	0.232	1.707	2.970	0.808	0.547	5.393	0.184	1.816
10	0.949	0.308	0.975	0.9727	0.284	1.716	0.276	1.669	3.078	0.797	0.687	5.469	0.223	1.777

2.5. Design of Experiments

Experiments are a part of the scientific method, which is an iterative model. As not every process can be characterized by simple theoretical or mathematical models, or their

knowledge is not complete enough to do so, experiments are necessary to unveil information on these processes. The current state of knowledge of a subject leads to questions, which in term lead to experiments attempting to solve these questions. The iterative loop consists of assumptions or hypotheses, design of the experiments, data produced through the experiments and its analysis and interpretation, which lead to new assumptions, experiments, and results. This loop continues to go on until a satisfying level of knowledge is reached (Hunter et al., 1998; Toutenburg and Shalabh 2009).

Throughout this process, there is also uncertainty caused by experimental variability, errors or noise caused by uncontrollable or unidentified factors. Statistics provide a language and logic that deals with all these factors and facilitates the realization and interpretation of the experiments. It aims at improving their effectiveness and productivity, and helps the practitioner to produce information-rich data and to better analyze it. Nonetheless, the crucial component of drawing conclusions from this analysis still remains a responsibility of the practitioner (Hunter et al., 1998; Toutenburg and Shalabh, 2009).

Experiments play a crucial role in process development and process improvement. DoE is a statistical tool that is meant to facilitate and provide a statistical background to the execution of the experiments. Henceforth, it is particularly useful to improve quality and productivity in processes and products. It can be applied at the early stages of product development, to help design new products, or for the improvement of existing product designs and manufacturing processes. It is an important part of the DMAIC process, mostly in the improve step. Its correct implementation can improve product's manufacturability, reliability, and performance. It can also be extremely useful in process development and improvement (Montgomery 2012). The methodology was originally developed in the early 1920s to improve crop yields at the Rothamsted Agricultural station in the UK. It took experimental design a few decades to reach the industry and substitute the less complete One Factor At a Time (OFAT) methodology (Antony et al. 2020).

Genichi Taguchi, a Japanese engineer, greatly contributed to the development of DoE in the manufacturing industry, as he developed many of the methods utilized today. He emphasized the concept of "loss function", which explains that the furthest a quality characteristic is from its target due to variability, even if within specifications, the higher will the customer dissatisfaction be, and thus the higher will the costs be (Tague 2005).

Experimental Design, or DoE, is the formal plan used to conduct the experiments, which consist of a test or battery of test runs in which the input variables, denominated factors, are purposefully changed to observe the effect on the process output. Factors may be quantitative, such as temperature, length and weight, or qualitative, such as machines and operators (Hunter et al. 1998). DoE considers not only the individual effects of each factor, but also their interactions with each other, as factors occasionally function together synergistically or antagonistically to affect the output. This, and the fact experimental design is more thorough and explores every combination of the chosen values (levels) for the considered factors, as well as needs fewer trials, which leads to lower costs, are the main advantages of the implementation of DoE over the previously utilized technique of OFAT (Raffaldi and Kappele 2011).

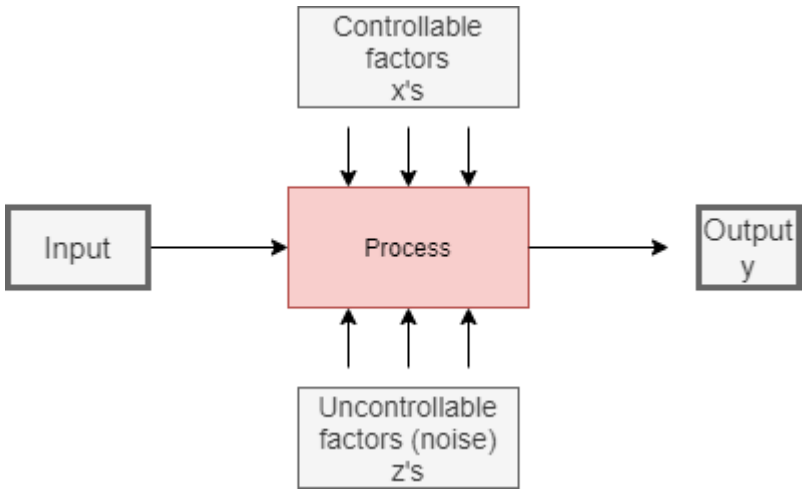


Figure 2.3 - Process model

The process is viewed as a "combination of machines, methods and people that transforms an input material into an output product" (Montgomery 2012). While some variables are controllable (x_1, x_2, …, x_n), others are uncontrollable (z_1, z_2, …, z_m). However, if the experiment requires it, these uncontrollable variables can sometimes be controlled. The uncontrollable factors are often called noise. Another situation that might present itself is one of factors that are controllable, however do not matter for the experiment. These are generally held constant for the duration of the experiment, to avoid it interfering with the results (Montgomery, 2012; Hunter et al., 1998). A visual representation of a process is presented in **Erro! A origem da referência não foi encontrada.**.

The output product has one or more characteristics that can be observed to assess it. The measured output, denominated response, varies in function of the objectives of the DoE implementation. Some outputs taken into consideration must also be indicative of product specifications or some important product characteristics. In cases where it would not be beneficial to promote an output in spite of another, several responses can be measured (Raffaldi and Kappele 2011).

An important consideration when running the experiments is to guarantee randomization of the runs. It decreases the effects of uncontrollable variables on the results, as they have more chances to balance themselves, while improving the validity of the estimates of experimental error variance. Another benefit from randomization is that it enables the application of statistical tests of significance and computation of confidence intervals. If it is impossible or too costly to randomize a factor, a procedure denominated blocking can be utilized. Another crucial component is replication. By replicating experiments, the effects of uncontrollable variables are further balanced out, and errors in measurement may be more easily detected (Hunter et al., 1998; Tague, 1995).

To sum it up, experimental design is the plan that includes the definition of the different factors, whether they are controlled or not, what are their levels, what is/ are the response(s), the sequence of the experiments, as well as how many replicates are performed. It also defines the type of design utilized to analyze the data.

Experiments can have several objectives. Some can be aimed at differentiating most significant input variables, at adjusting the process average to a desired value, others at determining which set of variables results in a better output, which one yields the lowest variability, or even at improving robustness, by finding which set minimizes the most the effects of the uncontrollable variables. Experiments can be made with one or several of these objectives in mind. Therefore, DoE can be utilized in process development or process improvement (Tague, 1995; Montgomery, 2012). Different types of DoE are suited for different objectives. Some types of DoE, relevant for this case study, are presented from sections 2.5.2 to 2.5.5. For more information on other types of DoE, see Montgomery (2012), Montgomery (2019), Toutenburg and Shalabh (2009), and Hunter et al. (1998).

Experimental design is closely interrelated with SPC. If a process presents a low capability, despite being in statistical control, the utilization of DoE can be useful to reduce variability, thus improving capability. SPC is a passive statistical method, as the process is observed and there is a waiting for useful information that may lead to changes in the process. However, this necessity for change is only detected when special cause variation is present. This means that SPC does not produce much useful information for change when a process is under control. On the other hand, DoE is an active statistical method, as it can be used to improve processes even when in statistical control (Montgomery 2012).

Experimental design can also be used to bring a process under control. If a process is out of control and has numerous input variables, the implementation of DoE can help to understand which are the relevant variables (Montgomery 2012).

Experimental design is reportedly used in many scientific fields. A study performed in 2018, based on more than a million scientific publications, showed that the main fields where DoE was applied were medicine (18%), engineering (10%), and biochemistry (10%). However, it is also applied in less expected fields such as agriculture, mathematics, arts, or psychology (Durakovic 2017).

Experimental design is mostly used in the manufacturing industry. However, it has also been implemented successfully in the service industry. A question that may arise is why DoE is not utilized more frequently in the service industry. Even though it can produce good results, there are some barriers to its implementation. There is a lack of awareness of the experimental design methods in these industries, difficulty in accurately measuring service's performance, influence of human behavior, high presence of noise factors (queuing, location, friendliness, etc.), difficulty in defining and controlling inputs and defining outputs to include intangible dimensions of a service, and difficulty to accurately define service processes (Antony et al. 2011).

Even with these challenges, DoE can produce great results in the service industry, when properly applied. Publications demonstrate its successful use in many sectors, such as the education sector, helping to evaluate teaching effectiveness in the UK higher education sector (Antony et al. 2019), healthcare sector, helping to reduce patient's waiting time in a cancer pharmacy (Arafeh et al. 2014), marketing sector, to increase the effectiveness of direct mail programs (Bell, Ledolter, and Swersey 2006) or to improve online customer experience (Bleier, Harmeling, and Palmatier 2019), or logistics sector, to evaluate order picking performance

trade-offs (Chackelson et al. 2013), to name a few. A study published in 2020, based on 26 articles, showed that the main sectors of implementation of DoE in the service industry were healthcare (27.6%), retail (24.1%) and logistics (17.2%) (Antony et al. 2020).

Today, many software packages are able to perform a lot of the calculation involved in this technique. These greatly facilitate the analysis of data, however the practitioner still needs to ensure the correct design of the experiment, as well as the proper analysis of the results. Aside from the computing capability, there are several other advantages to the use of software and computers for the utilization of experimental design, described as early as 1985. Computers can store the information on many types of designs, so that these can be easily available for implementation. Some software can prompt users to consider aspects of good experimental design they might not have thought of, such as restricted randomization, blocking, etc, which allows the design to be more complete and better match the objectives of the study. Hence why, adequate software can also be a motor of development for inexperienced practitioners, helping them to develop better experimental strategies (Snee 1985). The use of simulation to perform experiments has been a widespread practice, as it is a valuable alternative when physical experimentation is not available, or the costs associated with it, whether in time or resources, are too high. On the other hand, DoE can also be applied to simulations to improve them (Jankovic, Chaudhary, and Goia 2021).

2.5.1. Methodology for the implementation of Design of Experiments

Several methodologies have been presented in the community to help obtaining a successful implementation of DoE. Dave Doehlert, whom methodology is presented in an article by John Raffaldi and William D. Kappele (2011), and Montgomery (2019) both developed 7 step methodologies for this purpose. Hunter et al. (1998) also presents a 7 step methodology.

They present many similarities in their approach. Even though the content of each individual step varies, all focus on a pre-experimental planning phase, where the process is studied, and inputs and responses are defined, and then present steps for the development part of the DoE where results are obtained and analyzed. As Montgomery is a more knowledgeable source on this subject, his methodology is presented in this section. However, Dave Doehlert's and Hunter's methodologies address topics Montgomery does not. Dave Doehlert emphasizes model testing, stating it as an important part of his methodology. For more information on this topic, see Raffaldi and Kappele (2011).

Hunter's methodology, on the other hand, is more distinguished and fairly different in the organization of its steps. It emphasizes the importance of a quantitative approach during the whole process, from the definition of the problem to the analysis and conclusions, and recommends to try to avoid too many subjective interpretations. It also focuses more on the management aspect of organizing the experiments. Hunter mentions the importance of defining the schedule for the experiments, as well as necessary time, costs, machines, manpower, instrumentation, and more. He also mentions the importance of meticulously defining each step of the experimentation, to make it as easy to follow as possible. For more information on this methodology, see Hunter et al. (1998).

The steps of Montgomery's methodology are presented in the following points. Steps 2 and 3 do not need to follow the specific given order and can be undertaken simultaneously.

1. **Recognition and statement of the problem.** First, it is important to take into consideration that it is not always obvious that a problem where DoE can be applied exists. It is also not easy to properly define the problem and come to a generally accepted statement of it. In this phase, a team approach is recommended, and it is important to gather input from all involved parties and stakeholders: engineering, quality, manufacturing, management, marketing, customers, and operators. There should also be a clear definition of the objectives of the experiment.
2. **Selection of the response variable.** The selected response variable must provide valuable information about the process. This response will commonly be the average and/ or standard deviation of the selected output characteristic. Multiple responses are common. It is also important to define how measurement will be undertaken.
3. **Choice of factors and levels.** In this phase, the factors that will be studied during the experimentation, their ranges of variation, as well as the specific levels under study are defined. Process knowledge is key during this step. It is recommended to explore all potentially relevant factors and not being overly influenced by earlier experience. It is also recommended to keep a low number of factor levels. After selecting the factors, ranges over which the factors will be varied, and specific levels utilized for the runs must be defined. The utilization of tools such as brainstorming, flowcharts or an Ishikawa Diagram to gather and summarize information on causes for the problem and/ or factors of the experiment can be useful. Even though this step is anterior to the choice of an experimental design, having an idea of the desired design methodology is important. For the case of a screening experiment, the utilization of wider levels, meant to represent opposites of possible operating conditions, is recommended. This allows for differences between levels to be more obvious and for a better diagnosis of the factor's effect. In the case of an optimization study, levels are tightened around the level that resulted in the best response in the screening experiment (Tague 2005).
4. **Choice of experimental design.** This step involves decisions regarding selection of sample size, number of replicates, as well as an adequate run order, and determination of whether blocking or other randomization restrictions are applied. It also involves the choice of the specific experimental design and of the empirical model used to describe the results. This choice will depend on the objective of the experience defined in step 1.
5. **Performing the experiment.** While running the experiment, it is important to ensure that everything is being done according to plan, as errors in this phase can invalidate the experiment. Therefore, it is important to plan the verifications that need to be undertaken during the experiment. Coleman and Montgomery (1993) suggest that a few trial runs or pilot runs are often helpful as an

intermediate phase before this one. Based on that, choices made in steps 1-4 can be reviewed and adjusted.

6. **Statistical analysis of the data.** Statistical analysis helps to reach objective results and conclusions. Even though it cannot prove the effects of the factors, it can attach a level of confidence to a statement and/ or measure the likely error in a conclusion. Software packages can prove useful in this step, as they perform most of the calculation and can present graphics for data analysis and interpretation. Procedures such as hypothesis testing and confidence intervals are also useful to analyze the data. In this step, data is also analyzed through the utilization of the empirical model defined in step 4. Residual analysis and model adequacy checking must be undertaken to determine the validity of the experiment. Tague (1995) recommends the use of a Pareto chart to analyze the relative significance of the effects.
7. **Conclusions and recommendations.** Experimenters should draw practical conclusions from the results of the experiment and recommend further actions. Graphical methods are useful to present the results to stakeholders. Follow-up runs and confirmation testing might also be performed to validate conclusions. It is important to keep in mind that the experiment should be iterative, and the results from one experiment should be used to formulate new hypotheses, as there is generally not a perfect knowledge of all the required parameters on the first experiment.

2.5.2. Factorial Experiments

Factorial design is utilized when there are multiple factors of interest in an experiment, with two or more levels each, and interactions are relevant. It can be used to analyze the statistical significance of different factors, determine whether there are interactions and their significance, as well as to unveil an order of significance for the influence of the factors on the response. These are frequently employed in an engineering or manufacturing context. In this design, factors are varied together, which means that in a complete replicate of the experiments all possible combinations of factor levels are investigated. The total number of combinations is simply obtained by multiplying the number of levels for each factor (Toutenburg and Shalabh, 2009; Hunter et al., 1998; Montgomery, 2012).

Such as the ANOVA with a single factor, there is a factorial model with random or mixed effects, for when the levels utilized are not important and selected at random. For more information on this subject, see (Toutenburg and Shalabh 2009).

The effect of each factor is called a main effect and consists in the change obtained in the response by modifying the level of the factor. When a factor has two levels, one will be denoted "-" and the other "+", denominated low and high, respectively. The main effect of a factor A with two levels is the difference between the average response at the high level and the average response at the low level, such as presented in equation 2.18.

$$A = \bar{y}_{A^+} - \bar{y}_{A^-} \tag{2.18}$$

Often, the difference between the levels of one factor changes with a modification of the levels of other factors. When that happens, there is an interaction between the factors. When an interaction is present, the interpretation of the main effect can be misleading, as it can even mask the main effect. Hence why, knowledge of the interaction can sometimes be more relevant than of the main effect. As referred earlier, the OFAT methodology, which is still utilized by some companies as an alternative to DoE to improve processes, does not take into account these interactions. By changing each factor individually, the best combination might not be found if there is an interaction.

The ANOVA is often utilized to statistically analyze the data obtained from factorial experiments. The effect of each factor and its interaction can be defined by the sum of squares (SS_X) of their observations. The total sum of squares (SS_T) is the sum of these sum of squares plus the sum of squares of residual errors (SS_{Error}), and its value, for a factorial with 2 factors, is given by equation 2.20.

$$SS_T = SS_A + SS_B + SS_{AB} + SS_{Error} \tag{2.19}$$

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk})^2 - \frac{(y_{...})^2}{abn} \tag{2.20}$$

The results are then summarized in an ANOVA table, such as Table 2.4.

Table 2.4 - ANOVA table for a two-factor factorial

Source of Variation	**Sum of Squares**	**Degrees of Freedom**	**Mean Square**	$\boldsymbol{F_0}$
A	SS_A	$a - 1$	$MS_A = \frac{SS_A}{a-1}$	$F_0 = \frac{MS_A}{MS_{Error}}$
B	SS_B	$b - 1$	$MS_B = \frac{SS_B}{b-1}$	$F_0 = \frac{MS_B}{MS_{Error}}$
AB	SS_{AB}	$(a-1)(b-1)$	$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$	$F_0 = \frac{MS_{AB}}{MS_{Error}}$
Error	SS_{Error}	$ab(n-1)$	$MS_{Error} = \frac{SS_{Error}}{ab(n-1)}$	
Total	SS_T	$abn - 1$		

Then, a hypothesis test is undertaken for each effect and interaction. To test for no significant factor effects, nor significant interaction, the obtained F_0 value for each source of variation is compared with the tabular value of F at a given significance level α. If $F_0 > F_\alpha$, the null hypothesis is rejected, and it is possible to conclude that the respective effect/ interaction is significant. If the null hypothesis is not rejected, the effect is considered insignificant.

In complement of this analysis, a residual analysis should also be undertaken, as it allows to assess model adequacy. The residuals from a two-factor factorial are the difference between the observations and the corresponding cell averages, as shown in equation 2.21.

$$e_{ijk} = y_{ijk} - \hat{y}_{ij.} = y_{ijk} - \bar{y}_{ij.} \quad (2.21)$$

The obtained residuals should be analyzed by plotting them in a normal probability plot, in a residuals vs. experiments sequence and in a residuals vs. predicted values plot, such as shown in section 2.3.2, Figure 2.1.

2.5.3. The 2^k Factorial Design

The 2^k factorial design is a factorial design with k factors at 2 levels each. Its name comes from the fact that each replicate has 2^k runs. This factorial design is useful in process development and improvement, while also being relatively simple (Montgomery, 2012; Hunter et al., 1998).

Each factor has a "high" or "+" level, and "low" or "-" level. The design can be represented by a design matrix. In this matrix, each row corresponds to a run of the experiment, with its combination of levels for the factors and their interactions. The level of interactions corresponds to the product of the level of their factors in that run. Each run can also be represented by a single notation, which is a series of lowercase letters. In each run, the factors that have their letter present in the notation are set at the "high" level. If the letter is not present, they are set at the "low" level. Table 2.5 gives the design matrix for a 2^2 factorial design. Some other types of notations can be seen in Hunter et al. (1998).

Table 2.5 - Design matrix for 2^2 factorial design

Run	A	B	AB
(1)	-	-	+
a	+	-	-
b	-	+	-
ab	+	+	+

The contrast of each factor is the sum of its n observations on each run, multiplied by its contrast coefficient in that run, which can be observed on the design matrix (+1 for factor A in run *a* and -1 in run *b*, for instance). If we take the 2^2 design presented in Table 2.5, the contrast of A would be

$$Contrast_A = a + ab - b - (1)$$

The main effect of each factor or interaction can then be obtained using equation 2.22.

$$Effect_X = \frac{Contrast_X}{n2^{k-1}} \quad (2.22)$$

Another method for the estimation of the effects is the utilization of Yates' Algorithm, which can be seen in Hunter et al. (1998).

To obtain the sum of squares for any factor or interaction, equation 2.23 is utilized.

$$SS_X = \frac{(Contrast_X)^2}{2^k \times n} \quad (2.23)$$

The analysis of variance is then completed by computing SS_T and SS_E through the utilization of equations 2.20 and 2.19, respectively, and developing the subsequent ANOVA table, similar to Table 2.4.

The next step is the analysis of residuals. To obtain the residuals, it is first indicated to fit a regression model to the data. A general form for the regression model is presented in equation 2.24.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \quad (2.24)$$

$\beta_0, \beta_1, \beta_2, \ldots, \beta_k$ – regression coefficients
$x_1, x_2, \ldots, x_k$ – factors and interactions
ε – random error term

The fitted regression model is obtained via the substitution of all coefficients by their estimated value and does not include the random error term.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_k x_k$$

$\hat{\beta}_0$ is the average of all observations, and the estimates for each other coefficient $\hat{\beta}_j$ are one-half the effect estimate for the corresponding factor, which are obtained by using equation 2.22. The value for each run can be obtained by substituting each x_j by the level of the corresponding factor at that specific run.

The residuals are then obtained by applying equation 2.21, where $\hat{y}_{ij.}$ is the value of $\hat{y}$ from the prior equation on the run corresponding to the levels of both factors on the calculated residual. Afterwards, the normal probability plot, a residuals vs. factor levels plot for each factor and residuals vs. predicted values plot are made and conclusions upon the validity of the experiment are drawn.

2.5.4. 2^{k-p} Fractional Factorial Design

An increase in the number of factors in a 2^k factorial design leads to a drastic increase in the number of runs in each complete replicate. In fact, for each new factor, the number of runs duplicates. For instance, a 2^5 factorial design would require 32 runs. In such design, only 5 degrees of freedom would be attributed to main effects, while 10 would be attributed to two-factor interactions, and the 16 left would be attributed to higher-level interactions (interactions of 3 factors or more). If the assumption that certain interactions are negligible can be made, a 2^{k-p} fractional factorial design, which corresponds to a $\frac{1}{2^p}$ fraction, can be used. For instance, there can be the need to use a factorial design when the experimenter wishes to study 8 factors with only 32 runs. A 2^k factorial experiment with 8 factors would normally require 256 runs, which means the needed fraction would be $32/256 = 1/8 = 1/2^3$. So, the fractional design would be a 2^{8-3} design.

This design allows to gather information on the main effects and, depending on the resolution of the fractional design, some or all low-level interactions. This means that to implement fractional factorial design there must be an assumption that higher-level interactions are negligible (which is generally the case). Therefore, fractional design allows to study a big

number of factors with a reduced number of experiments. This makes it well suited for screening experiments, which are performed when there is a high number of factors of interest, and with the purpose of determining which factors significantly influence the response, to then investigate further these factors in subsequent experiments (Montgomery, 2012; Hunter et al., 1998). It is important to note that not only significant factors matter. In fact, results that show which factors are insignificant are also especially valuable, as they can be set at more economical levels or left uncontrolled, thus resulting in cost savings (Tague 2005).

There are several ways of performing a fractional factorial design. One of them is performing only a fraction of the runs. Unfortunately, by doing so, an important portion of information is lost. Another possibility is to equate one or several factors to interactions of other factors. The substituted interactions are denominated generators and the factors that substitute these interactions denominated generated factors. By doing so, a full factorial experiment is performed, and only certain interactions are lost. As less information is lost with the second method, it was utilized during the case study, and is discussed in further detail.

The design matrix obtained when performing a fractional factorial design is similar to the one shown on Table 2.6.

Table 2.6 - Design matrix for 2^{3-1} fractional factorial design

Run	**A**	**B**	**C ≡ AB**
(1)	-	-	+
a	+	-	-
b	-	+	-
ab	+	+	+

An important parameter to look at when performing a fractional factorial design is the design resolution, which catalogs designs according to the aliases they produce. There are three important design resolutions:

1) **Resolution III designs.** Main effects are not aliased with each other; however, main effects are aliased with two-factor interactions and two-factor interactions may be aliased with each other.
2) **Resolution IV designs.** Main effects are not aliased with each other nor with two-factor interactions, but two-factor interactions are aliased with each other.
3) **Resolution V designs.** Main effects and two-factor interactions are not aliased with any other main effect or two-factor interactions, but two-factor interactions are aliased with three-factor interactions.

Resolution III and IV designs are especially useful in screening experiments. Resolution IV is the preferable, as it will provide almost complete information on main effects and some information on two-factor interactions.

Another important element of a fractional factorial design is the identity element *I*, also called defining relation of the design. The defining relation is composed by all the words equal to the grouping of the letters of each generated factor with its generator, as well as the product

of these obtained identity columns. For example, in a 2^{6-2} fractional factorial design with $E \equiv ABC$ and $F \equiv BCD$:

$$I \equiv ABCE \equiv BCDF \equiv ADEF$$

To find the alias of an effect, one can multiply it by each word of the defining relation. For instance, the aliases of factor A are

$$A \equiv BCE \equiv DEF \equiv ABCDF$$

The selection of generators must allow for the highest possible resolution. A table with a set of maximum design resolutions for 2^{k-p} fractional factorial designs is presented by Montgomery (2019).

2.5.5. Response Surface Methodology

Response surface methodology (RSM) is used for process improvement. The principle of RSM is to plot variations of the fitted regression models, which is also called response surface, to analyze which regions of the graph produce the best results in the measured output characteristics.

To analyze the results of the experiment, the response can be approximated to a first-order model, such as the model presented in equation 2.24, or to a second-order model, presented in equation 2.25. The first-order model is utilized when the data can be fitted to a linear function. On the other hand, the second-order model is utilized when the data presents curvature. In practice, it is common to utilize the first-order model when we estimate to be far from the optimum operating conditions, and the second-order model when we are in the vicinity of the optimum.

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i} \sum_{<j=2}^{k} \beta_{ij} x_i x_j + \varepsilon \tag{2.25}$$

Even though it is unlikely that a polynomial function would correctly represent the true relationship over independent variables' entire space, it works well for fractions of it, which are usually tested during an experiment.

The method of steepest ascent consists in sequentially moving in the direction of the maximum increase of the response (or decrease if the objective is to minimize the response). By observing the fitted regression model, the increase and decrease in values of the variables that is optimum for the desired increase/ decrease in the response are quickly found. The obtained vector is called the path of steepest ascent. For instance, it might be observed that the path of steepest ascent follows a vector of an increase of 1 in factor A and a decrease of 2 in factor B. Then, there is a sequential movement along this path, which means a sequential realization of new experiences along this path, until there is no more increase in the response. This process can be repeated by fitting a new first-order model and renewing the execution of all the steps until the experimenter feels that the process is near the optimum.

After reaching the vicinity of the optimum, the response can be fitted to a second-order model to approximate it from the real optimum, as true response surface generally contains curvature. Then, the central composite design (CCD), rotatable design, or other designs can be used to find the optimum response (see Montgomery (2012)).

Despite not being applied during the case study, response surface methodology could be applied subsequently for a second DoE, after the screening experiment is finished, hence why it is presented in this section.

3. CHARACTERIZATION OF THE COMPANY

In this chapter, the enterprise in which the present study was undertaken, the company will be introduced, as well as the way it is organized, its mission and management.

The present study was undertaken in the production line dedicated to an invisible zipper, in close interaction with Quality and Production areas, who were crucial to the realization of the study.

In this chapter, the company will be presented, as well as its areas of Quality and Production.

3.1. The Company

The company produces fastening products that can be found on many garments, footwear, and fashion accessories of prestigious national and international brands, such as:

- Zippers.
- Plastic Hardware (buckles, strap adjusters, loops, etc.).
- Hook and Loop products (contact and other textile tapes, injection, etc.).
- Snaps and Buttons.

Its zippers come in several sizes and colors and are adapted to each client's needs, as its Research & Development department works directly with clients to ensure that their requirements are met.

3.1.1. The Company's Strategy

The company is based on a strong management philosophy that focuses on quality and continuous improvement. It is continuously aiming at improving its processes to give its clients the best value possible, and thus constantly attempts to reduce the amount of non-conforming parts or increase the stability of the process. It considers quality should be fully integrated, extending to all organization levels and every phase of the products' lifecycle. As such, attaining excellent quality, with a minimal cost, is a responsibility of every worker. The company is also involved in its clients' success, thus advises them over which products are best suited for their application, as well as gives technical recommendations, and information on technical specifications of the products.

The company is also highly committed to the environment, as it continuously seeks to improve its processes to, in term, reduce its utilization of resources and its pollution derived from the production of CO2 emissions, waste, and wastewater.

Within this scope, the company launched an offer for a project to improve the quality of its production line and eliminate the need for their Quality Wall, a workstation at the end of the production line where each item is fully inspected. The aim of the project was implementing several quality tools/ techniques to reduce the number of nonconformities in the production line of an invisible zipper. The project was awarded to a research team from the NOVA School of Science and Technology. The case study was developed within the context of the project.

3.2. The Product

The company produces several types of zippers for varied applications. The object of the study was an invisible zipper, divided into two sub-products. One is resistant to fire and not meant to be opened after installation, and therefore have a slider without any pull. On the other hand, the second sub-product is resistant to both fire and light and is meant to be opened after installation, hence why it has a slider with a pull and lower operating force specifications.

3.3. Production Areas

After entering the premises, raw materials go through several areas to be transformed into the final product and shipped to the client. The three main areas responsible for production are the Dyeing Area, Assembly Area, and Quality Wall.

In the Reception Area, boxes of continuous unbleached chain are received. The chain looks like a continuous zipper without slider, as can be seen on Figure 3.1. The chain's lateral tension is tested, to ensure conformity of the incoming product. For this purpose, one-meter samples are collected from five non-sequential boxes. The sampling procedure is independent from the lot size. The samples are then divided in several pieces and tested at the Quality laboratory, where they are subjected to a lateral tension test, and lateral traction is applied until the samples break. If the sample average is above the specification, the sample is compliant. If all samples are compliant, the lot is compliant.

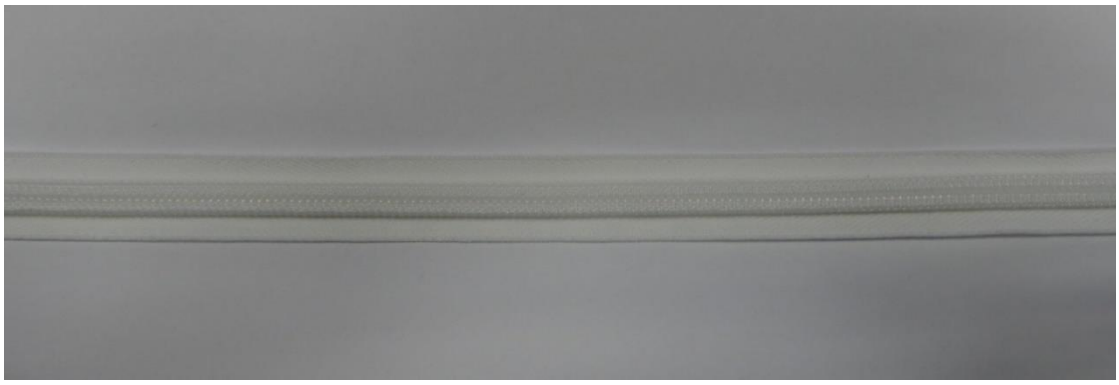


Figure 3.1 - Crude chain

There is a procedure for the verification of the order lot when lateral tension tests are non-compliant. Fortunately, this procedure has never been applied. It aims at identifying non-compliant boxes within the lot. Samples from boxes with production dates close to the non-compliant boxes are tested, until a conclusion is reached over the time period in which non-compliant boxes were manufactured. The situation is then reported to the supplier, with the identification of the non-compliant boxes. Finally, the other boxes are cleared and utilized in production. Independently from test results, additional samples are also kept to guarantee that posterior tests can be run in case of problems with the lot further along production or at the client.

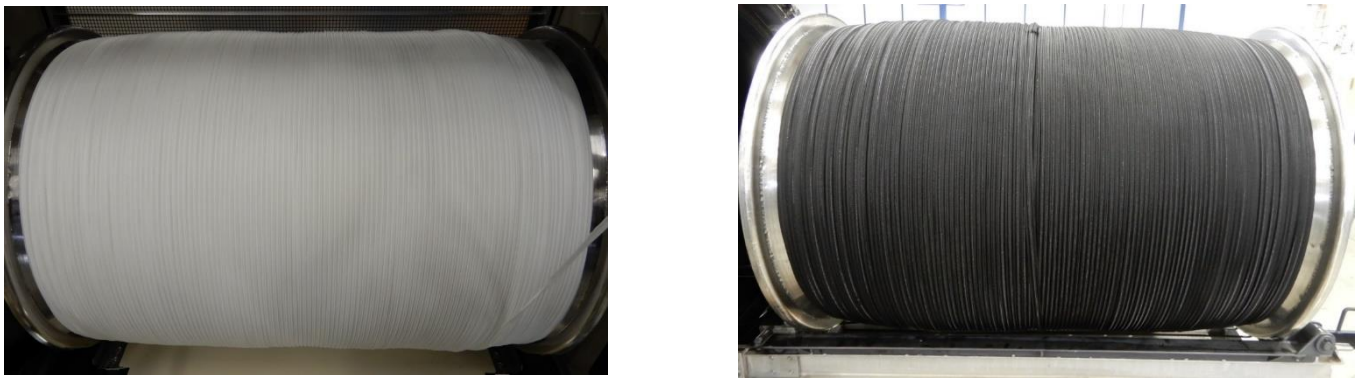


Figure 3.2 - Crude chain coil (left) and dyed chain coil (right)

In the Dyeing Area, several boxes of chain are joined and coiled before entering a dyeing machine. Figure 3.2 shows coils of crude chain and dyed chain. Each coil corresponds to a dyeing lot. Then, the chain is uncoiled, lubricated, dried, and ironed, before being sent to the Assembly Area. In the Assembly Area, all the operations that transform the chain into zippers are executed: spacing, welding of bottom, marking, cutting, slider insertion, and top stop crimping. The product is finally inspected at the end of the final assembly machine, and at the Quality Wall, before being shipped to the client. The final zipper part is shown on Figure 3.3.

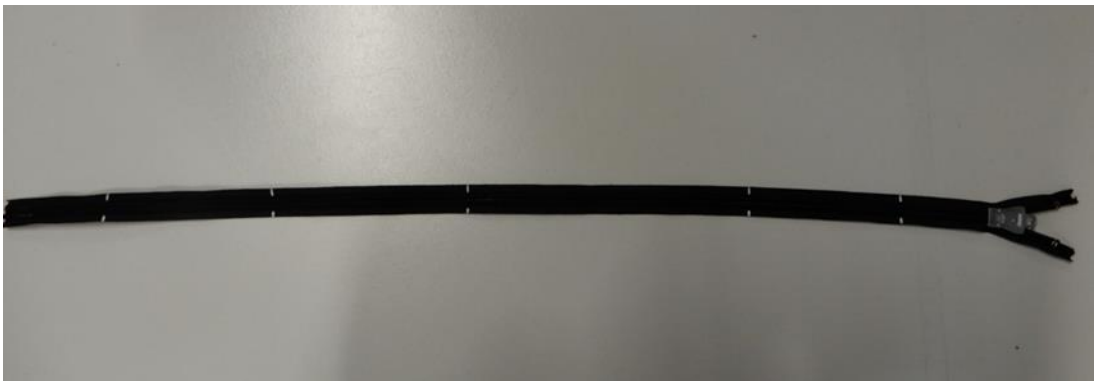


Figure 3.3 - Zipper part

Occasionally, tainted chain is ordered, in response to a specific need or problem in production. In that situation, it goes through all tests (color, flammability, and operating

characteristics) that the coil lots go through, and only passes shortly through the Dyeing Area for ironing.

3.3.1. Dyeing Area

The Dyeing Area is where the unbleached chain is dyed and chemicals are applied to it, to attribute it its adequate operating conditions. It is mainly composed by coiling machines, dyeing machines, and drying machines (that also uncoil, lubricate, and iron the chain). It has two intermediate storage areas, one for coils that await entrance in a dyeing machine or validation of color, and another for boxes of uncoiled chain that await quality validation (flammability, operating force, and lateral tension tests). It also has an area for dye "recipe" preparation, which consists in a group of separated doses of products (powder dyes, chemicals) that are introduced in the dyeing machine during different phases of its cycle. Finally, two small areas are dedicated to flammability and color testing, close to which a drying machine, meant for the samples to be tested, is located.

It is important to note that there is a dyeing phase within the Dyeing Process. As each phase might also be denominated of process, such as the coiling process, for example, the distinction between the two will henceforth be made through the usage of capital letters. Therefore, Dyeing Process refers to the wider process that groups all processes of the Dyeing Area, and dyeing process corresponds to the individual process undertaken by the dyeing machine where the chain is, in fact, dyed.

Another important aspect to refer is that production is limited by the cycle time of the dyeing machines, which dye a coil of chain approximately every 6h. As a result, 5 to 6 dyeing lots are produced every day.

As referred, quality tests are undertaken during the Dyeing Process. The color test verifies whether the lot's color is close to specification, as well as if it is close to the color of the previous lot. A sample from the outer part of the coiled chain is introduced in a color verification machine, which compares the chain tonality with the specified color specification. The machine returns the deviation from the specification. If the deviation is superior to the specification, or the deviation from the anterior lot is superior to a second specification, the color is non-compliant. However, passing this test does not grant compliance, as the machine is initially aimed at metallic pieces, which results in less precise measurements on textiles. Therefore, there is an additional step, where an operator proceeds to a visual verification of the sample. The final decision over the lot's compliance is taken by the operator. Non-compliant lots are rare. However, in that occurrence, the color can be adjusted through a second dyeing of the coil lot.

The flammability test, meant to ensure low flammability of the product, is necessary due to the utilization of the product in the automotive industry. Samples from the outside, middle and inside of the coil are taken for this test after the drying process, where the chain is uncoiled. One end of the samples is lit on fire. If the flame reaches a certain point/ length across the sample, the lot is non-compliant. If a lot is non-compliant, and non-urgent, it undergoes testing at an external laboratory. In cases where the lot is urgent or still non-compliant after external tests, it goes through a flammability treatment bath.

The flammability and color test are undertaken in the Dyeing Area laboratories. Operating characteristics tests (lateral tension and operating force), on the other hand, take place in the Quality laboratory. Two-meter samples from the outer and inner part of the coil, which are in term divided into ten separate pieces each, five for each test, are taken to perform the tests.

The lateral tension test ensures zippers are resistant enough to sustain the lateral forces they will suffer during utilization without breaking. This test is undertaken with a sample size of 10. Samples are introduced in the lateral tension test machine, and lateral traction is applied until the samples break. If the sample average is above the specification, the lot is compliant.

The operating force test guarantees that operators of the client and/or users (depending on the zipper type) will be able to manipulate the zipper and normally proceed to its opening and/or closing, taking into account ergonomic requirements. Samples are opened and closed in a test machine that measures the operating force at every instant, and then computes its average over the opening or closing. Even though each opening and closing results from the same piece of zipper, each one is considered an individual observation. This means that the sample size for this test is 20. If the sample average is below specification, the lot is compliant. There is a different specification for each sub-product.

Occasionally, lateral tension or operating force tests are non-compliant. When that happens, the chain goes through a second lubrication, which usually solves the problem.

3.3.2. Assembly Area

The Assembly Area is where the continuous chain suffers all the physical modifications to be transformed into an operational zipper part. In this area, the chain elements are spaced to start separating each zipper, bottom stops are soldered, the chain is marked, and, finally, cut into individual zippers, followed by the insertion of sliders, and crimping of top stops. As previously specified, there are different types of sliders for the two sub-products. Finally, there is a complete inspection of each zipper part, and nonconformities are registered. One of the lines benefits from automated verification through a visual inspection machine. However, some verification is still required, as the machine often produces false negatives, mainly due to folding or wrinkling on the zipper, which impedes the visual system to properly verify the characteristics.

When entering the assembly process, at the spacing machine, staples are overlapped with yellow tape. These staples result from junctions between chain segments, used when joining the initial chain segments of different boxes to form the continuous chain that is coiled, for example. The yellow tape is later filtered in the final assembly, to avoid any zipper with staples to reach the client.

For each machine, the article code is inserted, which dictates the machine's parameters. At the exit of each machine, there is a verification of the first and last zippers of the production lot, to verify that everything is going as planned and there is no mistake in machine parametrization. Each time the machine stops, there is also an extra verification of zipper parts following the one which caused the stopping. Each characteristic of the zipper is verified during the inspection at the end of the production line, to guarantee that no nonconformal

product reaches the client. The following elements are inspected: slider, top and bottom stops, marks, length, chain tonality and aspect (folded or wrinkled), elements, extensions. There is a specifically more thorough verification on the first units from each production lot, to guarantee that there is no problem with machine parametrization or operation.

3.3.3. Quality Wall

At the Quality Wall, all units of product are verified once again, before being packed and leaving to the shipping area. Once again, each element is inspected: slider, top and bottom stops, marks, length, chain tonality and aspect (folded or wrinkled), elements, extensions. There are some additional punctual verifications, such as manual verification of the operating force, that are applied whenever the operator feels it is necessary to verify it or are thoroughly verified to answer specific problems with which production is temporarily dealing. When nonconforming parts are identified, they are reworked, if possible, or switched for a conforming part from leftovers of a previous production lot.

The Quality Wall is used to provide one more barrier and avoid any nonconformal product to reach the client, as automotive manufacturers have very strict inspection of incoming products. One of the objectives of the project was to eliminate the need for this area.

3.4. Non-Production Areas

There are several areas not directly responsible for production who, however, play an essential role in the manufacturing process. From these areas, the Quality, Environment and Occupational Health and Safety Management System was most involved in the project, hence why it is presented in the following section.

3.4.1. Quality, Environment and Occupational Health and Safety Management System

At the company, quality is viewed as a necessity, and an integral part of the company. It influences all levels of the organization and all phases of the product's lifecycle. One of the main responsibilities of the Quality, Environment and Occupational Health and Safety Management System is to guarantee a level of excellence for the company's products. It both regulates quality by performing a high level of inspection and by continuously driving and implementing improvements to processes that, in term, lead to a better stability and diminish the number of nonconformities. Its inspections focus on several types of quality characteristics.

1. **Operating Characteristics.** It is important that the products respect certain specifications regarding operating force and lateral tension. These are verified during the Dyeing Process, with the usage of specific machines that quantify them.
2. **Safety Characteristics.** As the zippers are intended for the automotive industry, it is important that they respect safety measures. The flammability test, which is executed during the Dyeing Process, ensures the fireproof condition of the zippers.

3. **Visual Characteristics.** It is important that the color is within specifications and that the color from a coil lot is similar to the previous coil lot, so that products that end up in the same shipping lot have similar colors. It is both verified with the aid of a specialized machine and visually.
4. **Nonconformities.** Nonconformities are detected during the final inspection and at the Quality Wall. These characteristics are attributes, being characterized as OK/ NOK. There are a lot of different inspected types of nonconformities: slider presence and position, tonality, bottom-stop and top-stop presence and position, marking, just to name a few.

4.
CASE STUDY:
INVISIBLE ZIPPERS PRODUCTION LINE

The objective of the project contracted between the company and the NOVA School of Science and Technology was to elaborate an innovative statistical control plan, through an integrated approach and the implementation of several complementary quality tools, to improve a process of invisible zipper production. This chapter contains a description of the methodology utilized throughout the study, and a detailed approach of the statistical methods utilized.

It is important to note that the project was undertaken in close collaboration with a team of managers and engineers from the company, which will henceforth be denominated the company team.

Even though the study did not follow a documented approach, such as a DMAIC, a general approach was followed and is presented in Figure 4.1.

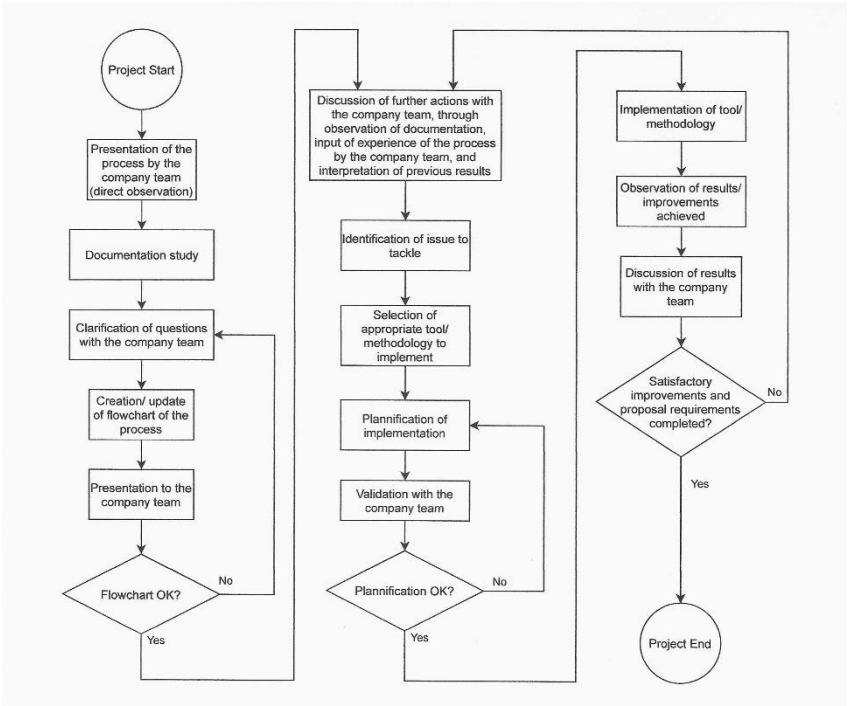


Figure 4.1 - General approach of the study

4.1. Process Definition

Initially, the priority was describing the process in detail, to fully grasp where improvements could be made and where to apply different tools or methodologies. Thus, it was decided to build a flowchart, to study and understand the process, as well as different issues and nonconformities faced at each step of the process, measurement capabilities and inspection frequencies, among others. This flowchart gathered several inputs. First, members of the company team presented the process, by showing and explaining every step that took place in the factory. Another input was the study of documentation, made available by the company team, which consisted in a compilation of vertical flowcharts, control plans, process of failure mode and effects analysis documents, inspection lists, and a list of nonconformities and their historical record of occurrence. This allowed to gather precious information on processes, inspection, and nonconformities. The flowchart was then appropriately reviewed with the company team to ensure its completeness and correctness. The subsequent flowchart, which served as a basis during the decision-making process, is not presented for confidentiality reasons.

Another key element was the list of nonconformities and record of occurrence. This list came in the form of a check sheet with data from each production line and contained a Pareto chart of all aggregated nonconformities for the past few months, presented in Figure 4.2. This data, along with discussions with the company team, allowed to understand and differentiate which issues were the most important, both in terms of proportion and consequences.

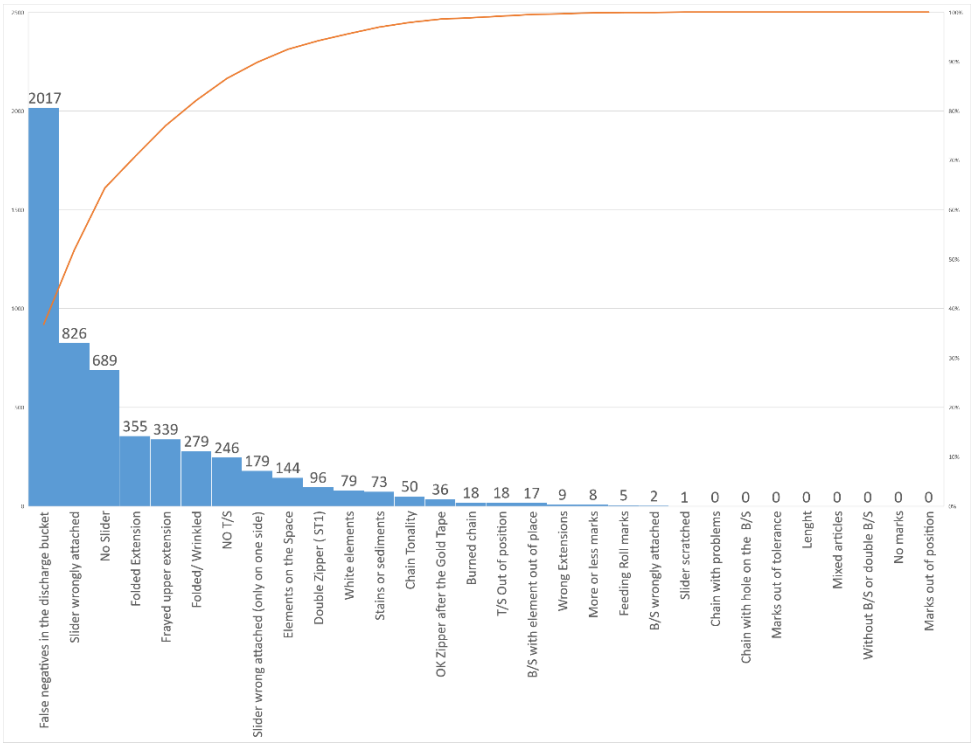


Figure 4.2 - Nonconformities Pareto chart

There are a few notable aspects in the Pareto chart. The chart shows a high number of false negatives in the discharge bucket. These false negatives are a result of the automated inspection on one of the production lines. These are not a nonconformity in themselves, however they are shown to be correlated with the folding and wrinkling of the zipper, as these issues prevent the machine from properly verifying the characteristics, mainly the marks. One can observe that folded or wrinkled nonconformities are far inferior to the occurrence of false negatives, which is justified by two factors. First, the inspection of folding or wrinkling of the zipper is only visual, therefore depends on the sensitivity of the operator over what is too much for the client. Second, and most important, it is not abnormal for a zipper to present a slight wrinkling or folding and it does not significantly affect its performance. Thus, this small alteration, which might hinder the visual verification of the machine, may not be a problem for the client. A folded/ wrinkled zipper is presented in Figure 4.3.

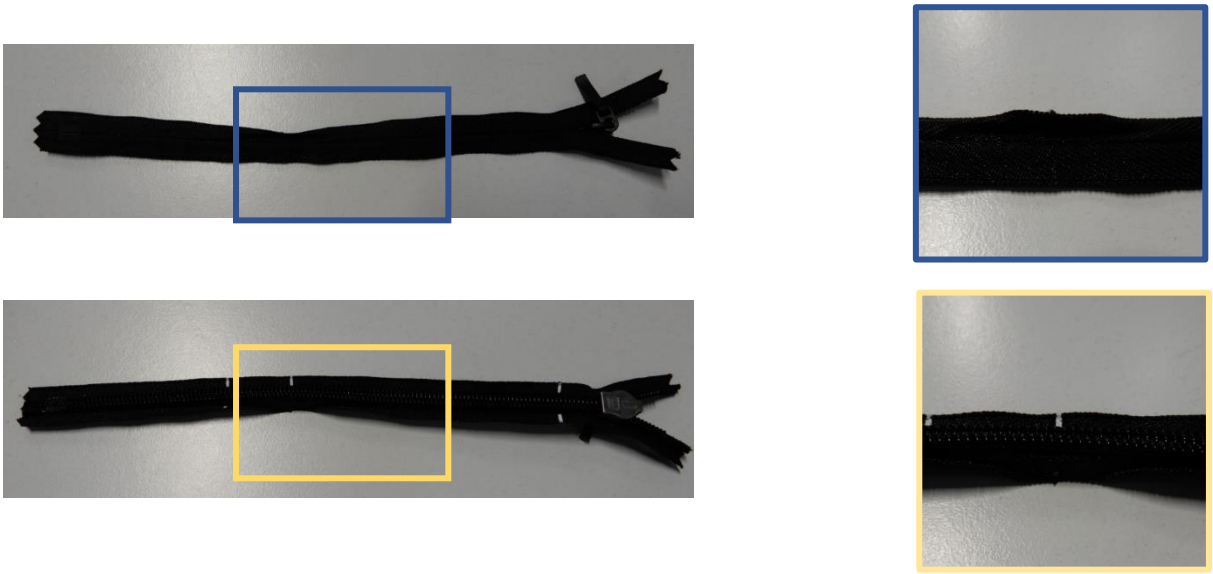


Figure 4.3 - Folded/ wrinkled zipper with top view (top) and bottom view (bottom)

Another relevant aspect is the occurrence of slider and extension nonconformities. It was decided not to initially investigate these issues, as there was a focus on intervening in the most initial steps of the process, since that is the best way to control quality over the whole process, and these nonconformities appear at the end of the assembly process. The origin of these nonconformities was also unclear, and there was an assumption that they could be influenced by curvature or wrinkles/ folds in the zipper, mainly for the extension nonconformities. An example of a zipper presenting curvature is shown in Figure 4.4.

One of the objectives of the project was to improve its quality by reducing the amount of nonconformities. To do so, it was necessary to target a process with high probability of incidence on nonconformities and as early as possible in the production process. It had been observed by the company team that the coiling and dyeing processes strongly influence the curvature and wrinkling of the chain. In fact, the coiling process was suspected to introduce a good amount of curvature and folding/ wrinkling, due to the torsion it applies to the chain and the added pressure of the compacted weight of the chain on itself. On the other hand, in the dyeing process, the chain is subjected to high pressures and temperature, and many chemicals. It is suspected that it can originate several problems regarding chain tonality, curvature,

and folding/wrinkling, which can be aggravated in the occurrence of differences of pressure within the coiled chain deriving from the coiling process.

Figure 4.4 - Curved zipper

As a result, this situation was considered ideal for the implementation of a DoE, as there was a need for a reduction of nonconformities, and there were two excellent candidates: the coiling and dyeing processes. Even though the curvature and folding/ wrinkling are not always considered nonconformities, as zippers which present these attributes are only considered nonconforming if the attribute is highly accentuated, they were suspected to lead to other issues further in the production line and increase the occurrence of other nonconformities. For instance, these could create problems during the marking process, thus originating marking nonconformities, or affect the extensions, as previously mentioned. Additionally, these are undesirable attributes aspect-wise.

Before making a choice regarding the next steps for the process improvement, it was decided that preliminary control charts would be implemented, to observe the current state of the process.

4.2. Control Charts

As the Dyeing Process is believed to introduce most, if not all, of the variability and alteration the characteristics suffer, control charts were implemented on the operating characteristics, controlled at the end of this process. More precisely, they were implemented on the peaks and average of the operating force, and lateral tension characteristics.

The charts were developed with historical data for the operating characteristics. The data was divided into several versions of the raw material, as versions had been changing often due to attempts from the supplier to improve the process. There was not enough data on the latest version, so data from the prior version was utilized. It was chosen to not mix data from the different versions to minimize external variation/ noise in the control chart. The data indicated whether the entries came from the beginning or end of the coil, however it was initially decided that it would not be separated in those subgroups, as it was considered that characteristics at the beginning and end of the coil should be similar. Another assumption was that the operating force was not significantly different during the opening and closing of the zippers, which was later proved wrong by the analysis of residuals from the ANOVA test performed earlier. Even though the data did not state whether the entries were from opening

or closing, a closer look would show that the entries alternated between opening and closing data for each piece of zipper tested.

As a result, the initial control charts were made with 37 samples of 20 observations for the two characteristics of the operating force, and 38 samples of 10 observations for the lateral tension. The total number of measurements $N_{OF} = 37 \times 20 = 740$ for the operating force and $N_{LT} = 38 \times 10 = 380$, were both far superior to the minimum recommended number of 100 measurements. The number of samples for both characteristics were also far higher than the minimum recommended of 20.

As both operating force and lateral tension are measured on continuous scales, control charts for variables were utilized. More specifically, $\bar{X}$ and S charts were utilized, as the control charts were made with computing capacity, which meant there was no need for manual calculation, and sample sizes were bigger than 10.

As there were a considerable number of charts to build (considering each filtering of out-of-control points), a Visual Basic for Applications (VBA) code was developed on a macro-enabled Microsoft Excel™ sheet, to properly treat the data. The code synthetizes the data, iteratively removes all out-of-control points, verifying the rules for Shewart charts and indicating the violated rule to the user, produces the charts and computes the percentage of out-of-control points after all have been removed. On the first iteration, only data points violating rule 1, that is data points outside the control limits, are removed.

Each iteration of the code generates a table such as the one in Figure 4.5. Each line corresponds to a sample, and the columns correspond to the number of the sample (1), the average of the sample (2), which constitutes the data point for the $\bar{X}$ chart, the control limits of the $\bar{X}$ chart (3), the standard deviation of the sample (4), which constitutes the data point for the S chart, the control limits of the S chart (5) and the violated rule number (6) associated with the chart designation ("Med" for the $\bar{X}$ chart and "DP" for the S chart), in case any rule was violated by the data point. The out-of-control samples are identified by a red line.

1	2	3			4	5			6
Sample	Average	Central Limit	Upper Control Limit	Lower Control Limit	Standard Deviation	Central Limit	Upper Control Limit	Lower Control Limit	Rule
1	14,67	13,20	14,96	11,45	3,00	2,58	3,84	1,31	
2	16,56	13,20	14,96	11,45	4,14	2,58	3,84	1,31	Med Regra 1
3	13,42	13,20	14,96	11,45	2,68	2,58	3,84	1,31	
4	12,77	13,20	14,96	11,45	1,96	2,58	3,84	1,31	
5	13,84	13,20	14,96	11,45	2,36	2,58	3,84	1,31	
6	11,96	13,20	14,96	11,45	2,25	2,58	3,84	1,31	
7	11,43	13,20	14,96	11,45	2,12	2,58	3,84	1,31	Med Regra 1
8	10,77	13,20	14,96	11,45	2,24	2,58	3,84	1,31	Med Regra 1
9	13,21	13,20	14,96	11,45	3,44	2,58	3,84	1,31	
10	12,68	13,20	14,96	11,45	2,67	2,58	3,84	1,31	
11	16,41	13,20	14,96	11,45	1,98	2,58	3,84	1,31	Med Regra 1
12	14,20	13,20	14,96	11,45	2,37	2,58	3,84	1,31	
13	15,22	13,20	14,96	11,45	1,77	2,58	3,84	1,31	Med Regra 1
14	14,21	13,20	14,96	11,45	2,93	2,58	3,84	1,31	
15	12,60	13,20	14,96	11,45	3,31	2,58	3,84	1,31	
16	15,84	13,20	14,96	11,45	3,35	2,58	3,84	1,31	Med Regra 1
17	14,38	13,20	14,96	11,45	1,94	2,58	3,84	1,31	
18	11,80	13,20	14,96	11,45	1,84	2,58	3,84	1,31	

Figure 4.5 - Iterative table of data points and its elements

Each iteration also generates a second table, such as shown on Figure 4.6. This table contains information, for this iteration, on the average and standard deviation of the characteristic for all data points/ samples for the $\bar{X}$ and S charts (1), the sample size (2), the number of samples (3), the tabulated value of constant c4, for the indicated sample size, utilized for the calculation of the control limits (4), the constants and control limits of the S chart (5), followed by the constants and control limits of the $\bar{X}$ chart (6), and then the number of out-of-control samples and its proportion (7) related to the total number of samples from the first iteration. On the last iteration, this table also has rows for the number and proportion of total out-of-control samples across all iterations (8).

1	Average	12,118
	Standard Deviation	1,9018
3	n	10
	k	21
	c4	0,9727
	B4	1,7157
5	B3	0,2843
	LSCs	3,2629
	LICs	0,5406
	A3	0,9753
6	LSCx	13,973
	LICx	10,264
7	Out-of-control samples	0
	Proportion (%)	0
8	Total of out-of-control samples	16
	Proportion (%)	43,2

2

4

Figure 4.6 - Iterative table for chart limits calculation and general information and its elements

Finally, each iteration generates the $\bar{X}$ and S charts. For example, the charts obtained for the average operating force during the first, second and last iteration are presented in Figure 4.7. The other charts for the first, second and last iteration of the peaks of operating force and lateral tension are presented in Appendix B.

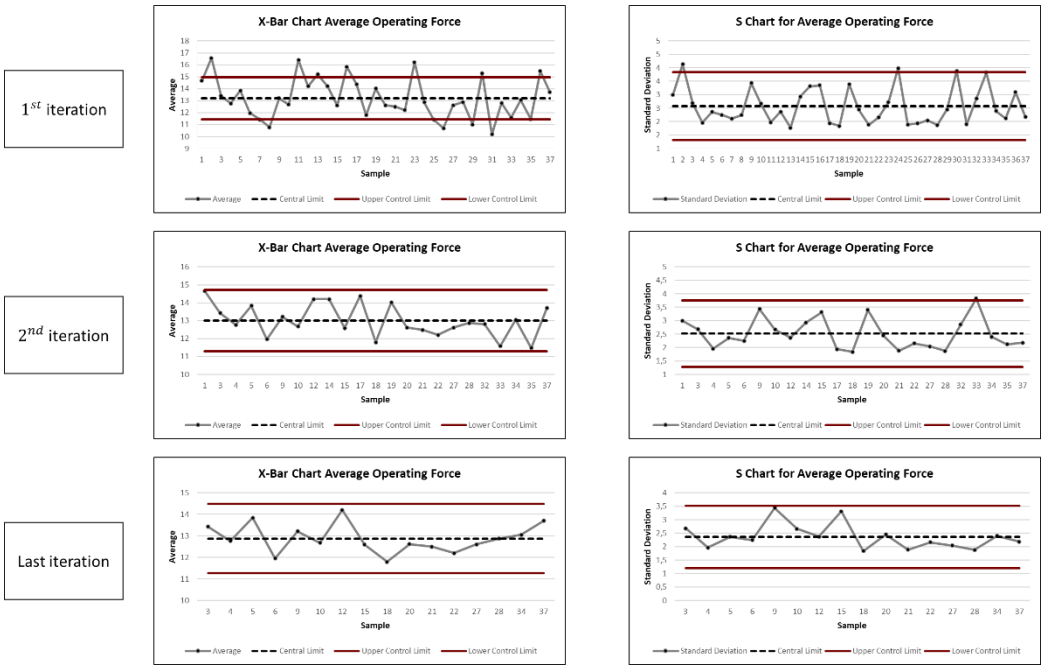


Figure 4.7 - Control charts for average operating force

Following this, it was noticed that each chart presented an exceptionally high number of out-of-control data points, with over 30% of out-of-control data points on the lateral tension charts and over 55% on the operating force charts, which made the process appear to be totally out-of-control. The percentage of out-of-control samples for each chart is presented on Table 4.1.

Table 4.1 - Percentage of out-of-control samples of control charts with combined data

Characteristic	Out-of-control samples (%)
Lateral Tension	31.6
Peaks of Operating Force	62.2
Average Operating Force	56.8

Due to this abnormally high percentage of out-of-control points, it was questioned what led to such a high number, and whether there was any neglected factor introducing variation. An assumption was formulated over a potential influence of significantly different data from the beginning and end of the coil. As a result, hypothesis tests were undertaken to conclude whether there was truly a significant difference between the beginning and end of the coil on those characteristics.

To do so, the means of the lateral tension, and of the peaks and average of operating force, were compared for the beginning and end of the coil. That gave place to three hypothesis tests. Each of these tests was executed to compare two paired samples, as each comparison was between observations of the beginning and end of the same coil. The outliers, which were the largely out-of-control points identified with the charts, were removed from the samples. All tests were performed with a significance level of $\alpha = 5\%$. The full results of the tests are shown in Appendix C and the p-value results of each test in Table 4.2.

Table 4.2 - Hypothesis test p-value results for comparison between means at the beginning and end of coil

Characteristic	P-value
Lateral Tension	0.1199
Peaks of Operating Force	0.0031
Average Operating Force	0.0003

The results show no significant difference between the beginning and end of the coil for the lateral tension. It is worth to note, however, that the p-value is of only 12%, which also does not give any confidence that the difference is not significant. On the other hand, the results show a clear significant difference for both the peaks of operating force and average operating force, with p-values of 0.31% and 0.03%, respectively. Therefore, it was decided to create new charts with separated data for the beginning and end of the coil for both the operating force and lateral tension characteristics.

As such, the VBA code was adapted to build control charts with separated data from the beginning and end of the coil. Thus, even though the number of samples remained the same, sample sizes were reduced in half, to 10 for the operating force characteristics, and 5 for

the lateral tension. This equaled a total number of measurements $N_{OF} = 37 \times 10 = 370$ for the operating force and $N_{LT} = 38 \times 5 = 190$, both still superior to the minimum recommended number of 100 measurements.

Even though the sample size for lateral tension was reduced to 5, it was considered that $\bar{X}$ and S charts would still be utilized, as these are not particularly counter-indicated for smaller sample sizes, and to maintain coherence and obtain a better comparison between the combined and separated data.

Each chart for the peaks and average of operating force and lateral tension was then generated, for both the beginning and end of the coil. The charts obtained for the average operating force, for the beginning of the coil, during the first, second and last iteration are presented in Figure 4.8 as an example. The other charts are presented in Appendix B.

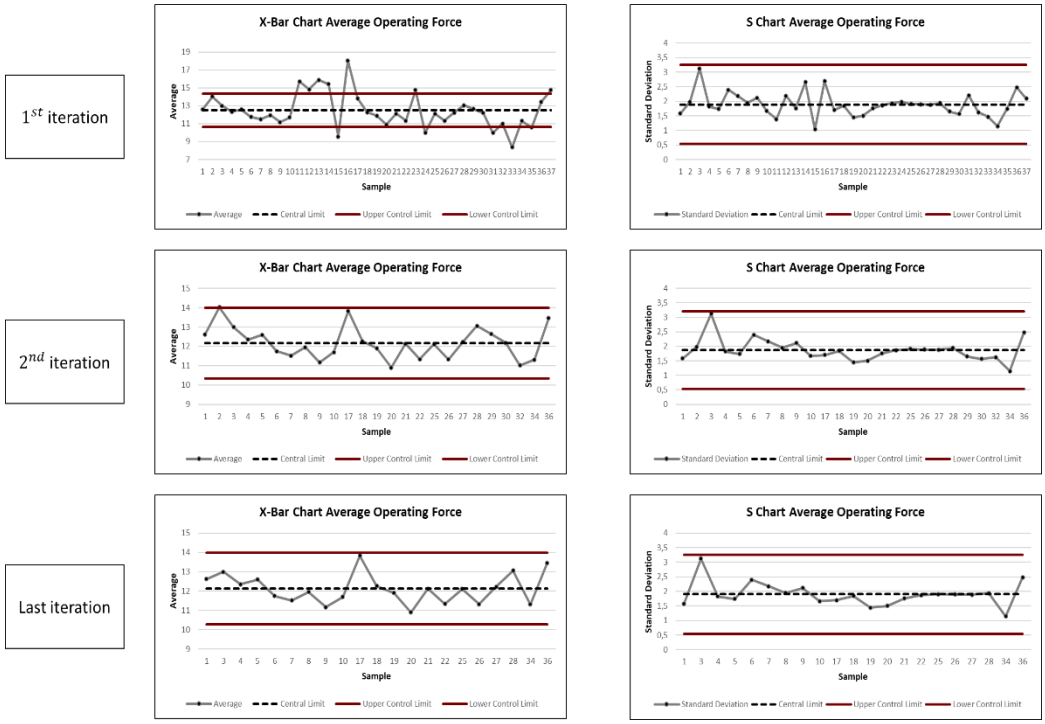


Figure 4.8 - Control charts for average operating force in the beginning of the coil

As a result, almost every chart presented an improvement. The percentage of out-of-control points for each measured characteristic with combined data, and separated data from beginning and end of the coil is presented in Table 4.3.

Table 4.3 - Percentage of out-of-control samples of control charts

	Out-of-control samples (%)		
Characteristic	**Combined Data**	**Beginning of Coil**	**End of Coil**
Lateral Tension	31.6	21.1	34.2
Peaks of Operating Force	62.2	35.1	45.9
Average Operating Force	56.8	43.2	37.8

4.3. Results of the Preliminary Phase and Discussion

The preliminary analysis of the process brought some interesting results, which are discussed in this section.

The analysis of the nonconformities Pareto chart identified slider and extension nonconformities as the main nonconformities of the process. However, it was decided not to initially investigate these issues, as these occurred in the last steps of the process, going against the focus on intervening in the most initial steps of the process, and the origin of these nonconformities was unclear. There was also an assumption that they could be influenced by curvature or wrinkles/ folds in the zipper, mainly for the extension nonconformities.

As such, even though curvature and folding/ wrinkling of the chain are not always considered nonconformities, it was concluded that applying a DoE, in an attempt to reduce these nonconformities, would be a good option.

The results of the hypothesis test performed during the control charts preliminary implementation, as shown in Table 4.4, allowed to observe a significant difference between the beginning and end of coil for the operating force characteristic, with a p-value of 0.31% and 0.03% for the peaks and average of operating force, respectively. These are both much lower than the test's significance level of 5%. As for the lateral tension characteristic, the results showed no significant difference between the beginning and end of the coil, with a p-value of 12%. It is worth to note, however, that the p-value is relatively low, which also does not give any confidence of the difference not being significant. For context, a test with a significance level of 15% would have concluded there was a significant difference between the beginning and end of the coil for this characteristic.

Table 4.4 - Hypothesis test p-value results for comparison between means at the beginning and end of coil

Characteristic	P-value
Lateral Tension	0.1199
Peaks of Operating Force	0.0031
Average Operating Force	0.0003

This test also demonstrated that the Dyeing Process, as was suspected, introduces variation in the operating force. It means that, even though there is a certain amount of variation which can be attributed to raw material, there is also a clear influence in the internal process, which in term signifies that at least a part of the out-of-control data points apparent in the control charts were due to the process itself. This further justifies the execution of a DoE in the Dyeing Process, more specifically in the coiling machine, to reduce the variability introduced by the process.

The preliminary control charts were implemented on the peaks and average operating force, and lateral tension characteristics. They were first implemented on the combined data for the beginning and end of the coil, and then these data were separated and distinct charts

created, deriving from the conclusions taken from the previously analyzed hypothesis test results. The results of the preliminary implementation of the control charts, as can be seen on Table 4.5, showed a significant proportion of out-of-control samples for the observed characteristics, which allowed to conclude an out-of-control state of the process for these characteristics. As can be seen, even though there are many out-of-control samples in every chart, the operating force characteristic present a worse behavior than the lateral tension characteristic. Furthermore, an improvement can be observed in the operating force charts when data from the beginning and end of coil is separated. This checks out with the previously analyzed hypothesis test results. However, there is an improvement only for the data from the beginning of the coil in the lateral tension charts, with a slightly worst result for the end of the coil. This does not indicate too much of a difference between data from the beginning and end of the coil for this characteristic, which also coincides with the hypothesis test results.

Table 4.5 - Percentage of out-of-control samples of control charts

	Out-of-control samples (%)		
Characteristic	**Combined Data**	**Beginning of Coil**	**End of Coil**
Lateral Tension	31.6	21.1	34.2
Peaks of Operating Force	62.2	35.1	45.9
Average Operating Force	56.8	43.2	37.8

Furthermore, through a visual analysis of the charts, shown in Appendix B, one can observe that the *S* charts had few out-of-control data points for both characteristics, which is indicative of there not being a big discrepancy in the variation within the samples, and that the problem lies in the variation from one sample to another.

It is worth noting that the residual analysis later performed during the ANOVA test showed a difference between the opening and closing values for the operating force. Therefore, a third version of the control charts, which would separate the opening and closing of the zipper on the operating force characteristics, could be undertaken to observe whether there is a new decrease in out-of-control samples.

Therefore, both the out-of-control state of the process and conclusion that the Dyeing Process introduces internal variation on the operating characteristics reinforced the need for a DoE. As such, it was decided to implement a DoE, either on the coiling or dyeing process. It was also decided that new control charts would be built after the DoE implementation to observe improvements and, only then, advance to phase II of the control charts.

4.4. Design of Experiments

4.4.1. Recognition and statement of the problem

After deciding a DoE should be implemented in the coiling or dyeing process, it was concluded that each run would represent a coil lot. However, many questions remained. The first one was to which process exactly would the DoE be applied. As both processes were

suspected of introducing a similar number of problems, the biggest factor for this decision was the ease of implementation. To grasp an understanding of the ease of implementation, the inputs of each process were studied. For this phase, in addition to the utilization of the company team knowledge and documentation, there was a direct observation of the Dyeing Process and interaction with operators.

The coiling process has several inputs. First, there is the crude chain inserted in the machine. It had been demonstrated previously that the characteristics of the crude chain were irregular, and even though new versions of the chain have since been developed by the supplier, its consistency remains a question. The chain is manufactured in the supplier's premises and the company has no control over it, which makes it a complicated input for the DoE. There is, however, one controllable aspect of the crude chain, which is the position at which it is inserted in the machine. The chain can be inserted with the elements facing inward (down) or outward (up) the coil. The other inputs of the coiling machine are its parameters: coiling tension, pitch, rotation speed, initial acceleration, and guide contact time. These are controllable and can be adjusted for each coil.

The dyeing process, on the other hand, is also influenced by various inputs. One of these inputs is the coil itself. This is controllable to a certain degree, as the parameters of the coiling machine and chain position influence the result, but it remains affected by the same uncertainties as the coiling process regarding the crude chain. The machine parameters, which are the temperature, pressure, and water flow, cannot be individually setup. Rather, an individual dyeing process requires several phases with different values of each parameter, hence why programs, which compile the parameters for all the phases of a dyeing process, are selected according to the sub-product type. Additionally, the machine does not fully control the parameters. For instance, if there is a shortage on water flow available, the water flow in the machine will be affected. In fact, the machine only guarantees parameters do not fall under a certain limit or go too far over the specified value, but they are slightly variable during the whole dyeing process. As such, individually controlling each parameter would be complicated and demand external control.

Another key component of the dyeing process is the dye "recipe". The dye "recipe" is composed of several separate doses of dyes and chemicals. The first step of its preparation is the mixture of various powder dyes into a recipient. Each powder, corresponding to a different color, is manually weighed on a balance. The balance is very precise, with a tolerance of $\pm 0{,}1\ g$. Then, several chemicals are weighed and placed into their own recipient. The types and quantity of chemicals depend on the program of the dyeing machine. The balance for the chemicals is much less precise. However, each chemical amounts for quantities of several hundreds of grams, and the specification tolerance is of $\pm 1\%$, for which the balance is sufficiently precise. Several questions arise from this dye "recipe" preparation. One concerns the manual weighing. As the weighing of the dyes requires a high level of precision, it is not impossible that the human factor may introduce some occasional errors. Another question regards the homogenization of the dyes, which occurs solely in the dyeing machine, and might not be perfect. It was suspected to be related with the white elements' nonconformity. A hypothesis test was suggested to test this assumption; however, it was decided to prioritize the planning

and realization of the DoE. In short, even though the dye "recipe" introduces some question marks, it is mostly controllable.

Finally, the ease of implementation pointed towards the coiling process, as the control of the machine parameters was simpler. Additionally, there was also a pre-existent interrogation over the ideal parametrization of the coiling machines, with some company factories in other countries presenting different default parametrizations. Therefore, it was decided that the DoE would be applied to the coiling process.

4.4.2. Choice of factors and levels

Then, arose the question of what the factors of the experiment would be. It was decided to include the three most important parameters for the machine setup: coiling tension, pitch, and rotation speed. The coiling tension consists in the tension applied to the chain during the process. This tension increases with the amount of chain that is coiled. An example of coiling tension setup can be observed in Table 4.6. The pitch is the horizontal distance between consecutive passages of the chain at the same angle of the coil. This means that, after the chain has gone 360º around the coil, the pitch is the horizontal distance between the two sequential parts of the chain. The rotation speed is simply the speed of the chain around the coil. The other machine parameters were held constant. Finally, the last factor was the position of the chain. The crude chain was an uncontrollable factor; however, it was considered that the randomization of the experiments would deal with it.

Some other factors needed to be held constant, such as the dyeing machine, coiling machine, and sub-product type, as these could eventually introduce small variations, which should be held minimal for the DoE purpose.

Table 4.6 - Levels for coiling tension in screening experiment

	Mass (g)	
Length	**Level 1**	**Level 2**
[0; 1000[	750	825
[1000; 2000[	850	935
[2000; 4000[	900	990
[4000; 7000]	1000	1100

Table 4.7 - Factor levels for all factors except coiling tension in screening experiment

Factor	**Level 1**	**Level 2**
Pitch (mm)	9	11
Rotation Speed (m/ min)	110	130
Chain Position	Up	Down

Afterwards, the specific levels for the factors were determined. These were thoroughly investigated with the company team to make sure that each level would be plausible of bringing improvements, while also maintaining a difference as wide as possible between the two levels, to guarantee that significant effects would be detected. The specified levels for the

coiling tension are presented in Table 4.6, while the levels for the other factors are presented in Table 4.7. Level 1 corresponds to the current setting for the process, while level 2 corresponds to the alternative setting chosen for the experiment.

4.4.3. Selection of the response variable

Afterwards, the question of how to measure the results of the DoE, or what output characteristics would serve as responses, was raised. Some of the first characteristics explored were the operating characteristics tests. These presented advantages and drawbacks. The operating characteristics are measured on a continuous scale, which are optimal to measure results. However, these characteristics were already within specifications. They were also complicated to measure with a high frequency, as each battery of tests accounts for around half an hour. Therefore, operating characteristics can only be measured on a small portion of the chain, measuring only two meters on a coil.

It was decided that the chain tonality would not be used as a response, as previous experience from the company team led to believe there was little correlation between the coiling process and the chain tonality. The flammability of the chain was also considered to present no correlation with the coiling process.

The issue was that the only other available indicators of results were the number of nonconformities. This was considered as far from ideal, as they are measured on a go no-go basis, which is less effective to measure results.

Another interrogation was how to measure curvature and folding/wrinkling. Even though these could be manually observed, a question was raised over how one would determine what was too much curvature or folding/ wrinkling for it to be considered a negative for the experience results. Another question was at what point in the process these characteristics should be verified and to what extent. Should one verify them at the end of the Dyeing Process, and then how would one define the inspection process, or should one verify each zipper at the end of the production lines.

A hypothesis that was then raised was whether curvature or folding/ wrinkling could be correlated with the continuous variation in the operating force. The company's operating force test machine returns a full graph of the continuous operating force of each observation. However, the first problem that was faced was that even though the graph was made available by the machine, there was no information about the variation on an individual observation. The machine only gave the average operating force for each observation and the global averages and standard deviation between all observations. It also did not give the specific data points of the graph, hence why analyzing the variation within each observation was not an option. Given these limitations, an experiment was performed to observe whether there was a correlation between average operating force and curvature or folding/ wrinkling.

This test compared three populations: normal zippers with no curvature nor wrinkling, zippers with curvature and no wrinkles, and zippers with folds/ wrinkles and no curvature. As there were more than two populations, it was decided to perform an ANOVA test.

The three populations for the ANOVA test were respectively denoted: normal, curvature, wrinkles. The test had the following null hypothesis:

$$H_0: \mu_{Normal} = \mu_{Curvature} = \mu_{Wrinkles}$$

The test was performed with a significance level of $\alpha = 5\%$. Data was collected by the company with sample sizes that allowed them to comfortably test the zippers without interfering with production. Each observed zipper produced two operating forces: opening and closing, which were not discriminated during the test. The collected data and corresponding results of the ANOVA test are presented in Table 4.8 and Table 4.9.

Table 4.8 - Data for ANOVA hypothesis test for curvature and wrinkles

Zipper Type	Y = Operating Force (N)	$Yi.$	$\overline{Yi.}$
Normal	18.9; 14.4; 18.4; 13.8; 17.2; 13.1; 17.8; 14.1; 18.5; 14.6; 18.4; 14.9; 17.4; 14.2	225.7	16.12
Curvature	15.3; 11; 14.6; 11.3; 17.5; 12.5; 14.8; 11.4; 14.6; 12.3; 14.9; 11.9; 15.1; 12.5	189.7	13.55
Wrinkles	13.6; 10.6; 16.7; 13.6; 16; 13.7; 17.6; 13.5; 16.5; 13.9	145.7	14.57

Table 4.9 – ANOVA table from ANOVA hypothesis test for curvature and wrinkles

Variation Source	SS	Freedom Degrees	MS	F_0	p-value
Between zipper types	46.81	2	23.40	5.56	0.008
Error	147.26	35	4.21		
Total	194.07	37			

As can be observed, the p-value of 0.8%, which is way below $\alpha = 5\%$, allows to conclude that there is a clear difference between the means of the three populations.

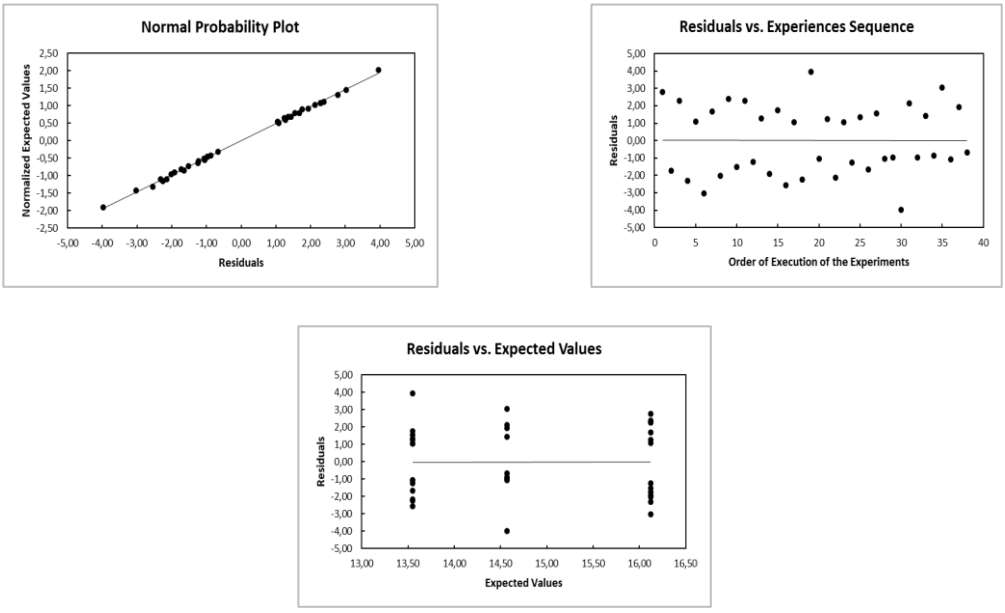


Figure 4.9 - Residual plots of ANOVA test

A residual analysis was performed on the data to guarantee the validity of the test. The results from the residual analysis are presented in Figure 4.9. The residuals vs. experiences sequence and residuals vs. expected values plots showed no specific pattern and demonstrated the independence and homoscedasticity of the data. The normal probability plot, however, showed an unusual pattern. Even though the data points follow the linear regression, there is an abnormal lack of points in the middle of the line, nearest to zero, where there should be a higher concentration. This derives from the utilization of both opening and closing operating force indistinctively. As the opening and closing forces are not equal, there are only points on both sides of the distribution and no points in the middle. However, as this pattern is balanced on both sides, it was considered to not affect the results of the test.

Table 4.10 - LSD test results

Pair	$\lvert \overline{Y_{I.}} - \overline{Y_{J.}} \rvert$	LSD
Normal/ Curvature	2.57	1.31
Normal/ Wrinkles	1.55	1.54
Curvature/ Wrinkles	1.02	1.54

Then, an LSD test was performed to observe which zipper types had significantly different means. The results of the LSD test are presented in Table 4.10. As one can observe, the zippers with curvature and with wrinkles were the only pair with no significant difference between their means, even though there is barely a significant difference between normal zippers and zippers with wrinkles. Below are presented the average operating forces of the different types of zippers, with a line that unites averages with no significant difference.

$\overline{Y}_{Curvature}$	$\overline{Y}_{Wrinkles}$	$\overline{Y}_{Normal}$
13.55	14.57	16.12

A conclusion that could be drawn from this data was that the normal zippers perform poorly regarding operating force. However, this seemed counter-intuitive, given that an abnormal folding/ wrinkling or curvature would intuitively result in a higher operating force. After reviewing the results with the company team, it was observed that this behavior was, in fact, abnormal. The cause seemed to be data collection. It was concluded that each sample of zippers corresponded to a different dyeing lot, which meant that these results could merely be a byproduct of variation between lots, potentially caused by the crude chain or by a variation introduced in the Dyeing Process, and not related to the actual observed characteristics. After evaluating whether there would be a reconduction of the experiment, it was concluded there was no need for it. In fact, the experiment allowed to observe the wide variance between lots, which pointed towards a low probability of the operating force being a viable measurement variable to evaluate the curvature and folding/ wrinkling characteristics.

The ANOVA test conclusions, allied with the conclusion, drawn through the hypothesis test during the control charts preliminary implementation, that there was internal variation within the coils, which meant that the measurement of the operating characteristics on a small

portion of the chain were not representative of the whole coil, lead to a decision such that the operating characteristics were not a viable response for the experiments. However, it was decided their data would still be collected during the experiment to verify whether any run resulted in an abnormally poor performance on these characteristics. It was also decided that operating characteristics would be measured before and after the dyeing process, to evaluate its effect on said characteristics. More specifically, two-meter samples would be taken after the coiling process, from the outside of the coil. These sample results would then be compared with the samples that are taken at the end of the Dyeing Process, from the outside of the coil. Another additional inspection introduced in the context of the DoE was to measure the operating characteristics on the crude chain. Five samples were removed from the crude chain for each run, to compare their results with the results obtained on the dyed chain.

As a result, the remaining option was to utilize the amount of nonconformities as a response, and find a replicable manner of evaluating curvature and folding/ wrinkling, to include them in the responses. Even though nonconformities are an attribute and, as such, not ideal to measure results, it was considered that, within the context and given that every run would correspond to a dyeing lot, they would be a reliable indicator for the results of the experiment, as each dyeing lot corresponds to approximately 6,000 zippers, which is enough to reliably observe results and draw conclusion from characteristics that are attributes.

The nonconformities were simply measured by their proportion of occurrence. However, there was initially not a thorough registration of all occurrences of nonconformity. Within the assembly area, if there were nonconformities before the final assembly machines, such as a spacing nonconformity at the spacing machine for instance, the corresponding zippers would be discharged without registration. The experience of the company team indicated this seldom happened. However, for the sake of the experiment, it was considered important to register each and every nonconformity. As such, there was a revision of the existing check sheets to adapt them to all assembly stages enable the registration of every nonconformity. Another problem with the current check sheets was that they did not keep the identification of the dyeing lot, which was indispensable for the experiment, as each run of the experiment corresponds to a dyeing lot. Finally, check sheets were developed for each workstation of the Assembly Process, to register each nonconformity that might happen at that specific workstation, and nonconformities were registered at all stages of the assembly line, at the final assembly inspection and at the Quality Wall.

Sequentially, a solution was reached over how to measure curvature and folding/ wrinkling. Curvature would be measured with a gauge, such as the prototype presented in Figure 4.10, at the Quality Wall. As can be seen, the gauge presents two stops towards which the zipper must be pushed. It is then verified whether the zipper fully covers the green zone. If it does, it is considered a zipper with curvature nonconformity. If some green still appears, it is considered to not have curvature.

The prototype later suffered some modifications. It was opted to not use circular, but rather linear stops, and have the green zone cover both the upper and lower direction of the plane. This allowed for the operators to measure the curvature much quicker, by not needing

to change the zipper's orientation, but only experimenting on both sides of the gauge with the same grip.

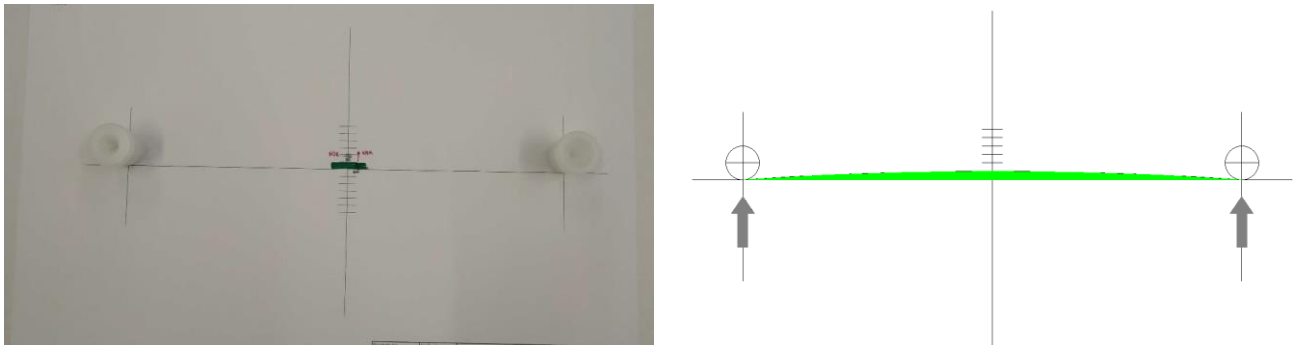


Figure 4.10 - Gauge for curvature attribute measurement

The operation of the gauge was observed during the realization of the experiments, which allowed to notice that the system presents some limitations regarding replicability, as there is some degree of subjectivity on zippers that present a curvature close to the defined maximum acceptable. This is justified by the fact that the grip of the operators on the zipper can alter its curvature, and that the green zone might appear visible from certain angles and be invisible from others. However, it was also observed that operators presented a careful manipulation of the zippers and a consistent criterion when evaluating the visibility of the green zone, which allows to conclude that this subjectivity is ultimately not significant for the results of the experiment.

As for folding/ wrinkling, the final decision was to measure it through the false negatives generated by the automated inspection at the end of the production line, since folding/ wrinkling prevents the machine from properly verifying marks and generates these false negatives, as previously stated. Arising from this, the proportion of false negatives was utilized as the response for folding/ wrinkling.

In short, the responses for the experiment were the number of nonconformities, number of zippers with a curvature nonconformity, and number of false negatives at the automated inspection machine. The objective of the experiment was to minimize each of these responses.

As nonconformities would be used as a measure, it was also necessary to keep the assembly line constant. A relevant aspect for the choice of the assembly line was the utilization of the automated inspection machine, which resulted in a decision of using the specific assembly line that included this inspection.

4.4.4. Choice of experimental design

Subsequently, the experimental design was chosen. It was determined that a screening experiment with a 2^{k-p} fractional factorial design would be implemented. More specifically, a 2^{4-1} design was selected. This design was intended to observe which of these factors were the most relevant and to grasp a first idea of their best levels. A second experiment would be performed afterwards if considered necessary. The fractional factorial design was chosen to diminish the number of experiments necessary, as each experiment would correspond to a coil lot, which takes about six hours to manufacture. Additionally, one replication of the

experiment was planned, to increase the reliability of the results and diminish the influence of the noise factors, which further increased the necessary time for the experiment. Yet another aspect that increased the experiment duration was the fact that some factors such as the dyeing machine, type of chain, and assembly line were held constant. All of these ultimately increased the necessity for a fractional factorial. Furthermore, a 2^{4-1} factorial was chosen, as it is a resolution IV design, which allows for a perfect observation of main effects and a good observation of two-way interactions. This totalized 16 runs for the experiment.

The utilization of a fractional factorial design raised a new question over the generated factor. It was decided that rotation speed would be the generated factor, as its interactions with other factors were not considered relevant, and the three-factor interaction between all other factors, the generator, was not considered significant either. The resulting design matrix is presented in Table 4.11.

Table 4.11 - Design matrix with standard order

	A	**B**	**C**	**D ≡ ABC**	**Response**	
Experience	**Coiling tension**	**Pitch**	**Chain position**	**Rotation speed**	**Replica 1**	**Replica 2**
1	750 to 1000	9	Up	110	Y_{11}	Y_{21}
2	825 to 1100	9	Up	130	Y_{12}	Y_{22}
3	750 to 1000	11	Up	130	Y_{13}	Y_{23}
4	825 to 1100	11	Up	110	Y_{14}	Y_{24}
5	750 to 1000	9	Down	130	Y_{15}	Y_{25}
6	825 to 1100	9	Down	110	Y_{16}	Y_{26}
7	750 to 1000	11	Down	110	Y_{17}	Y_{27}
8	825 to 1100	11	Down	130	Y_{18}	Y_{28}

Finally, the 16 runs were generated in a random order, to ensure randomization of the DoE. The resulting randomized design matrix is presented in Appendix D.

4.4.5. Performing the experiment

A trial run was performed before beginning the experiments, to ensure that everything was going as planned. In this trial run, there was a verification of the samples collection, factors parametrization (machine parametrization and chain position), measurement of nonconformities through the check sheets and of the curvature inspection at the Quality Wall. No error was detected, so the procedure was validated, and the runs were harmonized with the production and started.

As of the writing of this case study, the experiment is being performed. Hence why, results of the DoE will not be shown.

4.5. Sampling plan review and work in progress

During the analysis of the process, it was observed that there was no necessity of updating the sampling plans for the operating, safety, or visual characteristics. All these

characteristics were mostly within specifications and presented no complaints from clients. Even though the study showed a big variation of the measured characteristics, notably on the operating characteristics, within a coil of chain, and only a small portion of the chain is measured with the current sampling plan, it was concluded that a wider control would be inefficient. This conclusion was based on the fact that said characteristics inspections are time-demanding due to the available technology. The analysis of the inspection at the reception revealed, however, a need for a new inspection plan. Indeed, the collection of five one-meter samples for lots of up to almost a hundred thousand meters of chain, on a material that is proved to be inconsistent within the same lot, was judged insufficient. However, this was not deemed a priority, and the inspection plan at the reception has not been reviewed as of now.

A crucial part of the sampling process review focused on the Quality Wall, as it is a highly expensive part of the process. The necessity of the Quality Wall is to be reevaluated at the end of the implementation of the DoE, after analyzing whether there is a significant enough reduction of nonconformities that implies a reduction in the need for this area.

Several other tasks were deemed as not being a priority and, despite being foreseen, have yet to be undertaken. One of them is the test on the homogenization of the powder dyes. For this test, the color of coils dyed with a homogenized "recipe" would be compared with the color of coils dyed with a normal "recipe". A one-tailed randomized hypothesis test to compare the color deviation from the specification of the two independent samples would then be applied to the data, to observe whether the deviation is inferior on the coils with a homogenized dye "recipe".

Another pending task is the development of a tree diagram to dig deeper and attempt to reach the root causes of the other main nonconformities evidenced by the Pareto chart. When evaluating the plausible causes of the curvature and folding/ wrinkling, to see which processes were good candidates to attempt to reduce these issues, a compilation of these causes was grouped. A tree diagram was later developed on this basis. The diagram was then extended to the other main nonconformities. However, nonconformities would be reevaluated after the execution of the DoE. Therefore, this diagram was not considered a priority and, as such, not analyzed nor further developed in collaboration with the company team. As such, it remains in a development phase. The tree diagram is presented in Appendix E.

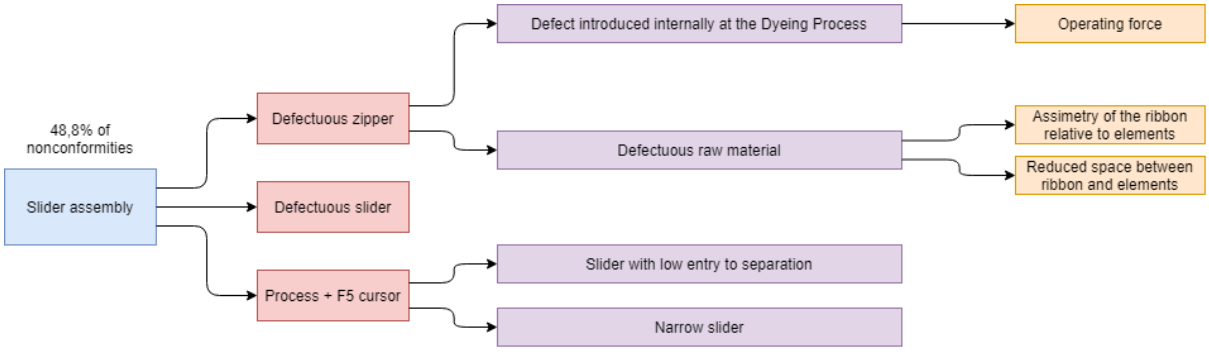


Figure 4.11 - Root cause tree diagram for slider insertion

The diagram led to some questions over the root causes of the slider nonconformities, which in term resulted in the formulation of several hypotheses, some of which have yet to

be tested, whereas the others have not been associated with any plausible test yet. The sub-diagram that led to these hypotheses is presented in Figure 4.11. It is worth reinforcing that there are two sliders for the invisible zippers, one for each sub-product. One of them is narrower and presents a higher occurrence of slider nonconformities. It also has a shorter entry to separation, hence why both are potential root causes.

The first test that can be formulated for the slider nonconformity aims at observing whether there is correlation between the operating force, more specifically a high operating force, and slider nonconformities. To test this, zippers with slider nonconformity and functional zippers would be drawn, ideally from production lots corresponding to the same coil lot, and their operating force would be measured with the operating force machine. A one-tailed randomized hypothesis test would then be applied to the data to compare the means of the two independent samples.

The second test would aim at determining whether the problem is related to defective sliders. To do so, sliders which have failed to be properly inserted would be carefully observed and tested for reinsertion. This could allow to easily identify the cause if sliders are, in fact, defective.

Both causes that lead to these tests are pointed as being relatively improbable by the company team. However, performing the tests would allow to fully discard them.

5. CONCLUSIONS AND FUTURE WORK

This chapter presents the conclusions associated with the present study and the results obtained through the implementation of the applied statistical control tools and methodologies. It also contains subsequent steps and propositions for future work.

5.1. Conclusions

The research methodology presented in this dissertation fits the proposed techniques of the literature review. Before implementing such techniques, it was necessary to build a basis and understand which variables influenced the process, hence why a study and synthetization of the process was required. To do so, material and knowledge from different sources was compiled into a flowchart of the process. This allowed to gather information on individual processes, raw material, equipment, inspection methodologies, and evaluate the influence of the human factor.

Through that analysis, it was possible to discern the complexity of a process that combines a rapid production rate of five to six coils of thousands of meters of fasteners, equating to 30,000 to 36,000 zippers a day, on a product that is inconsistent and might show a different behavior from one part of the coil to another. Additionally, many of its characteristics, such as wrinkling, folding, curvature, or its nonconformities are hard to measure on a continuous scale, and thus require to be measured on a go no-go basis. The mensurable characteristics, on the other hand, present a slow inspection rate, which only allows for the test of about the equivalent of 28 to 34 zippers a day.

Even though, as of the redaction of this study, no improvement on the process was achieved, several crucial conclusions were reached on the process behavior during the preliminary analysis phase and planification of the DoE, which was considered the best approach for improvement. Given the complexity of the process, its planification was time-consuming, hence why the DoE is being performed as of the writing of the study.

During the preliminary analysis, aside from the process definition, several $\bar{X}$ and S Shewart control charts for the operating characteristics were developed, to observe whether the process was in statistical control. At first, control charts with both data from the beginning and end of the coil were utilized, as there was no evidence of statistical difference between the two. After noticing an unusually high amount of out-of-control samples on the control

charts, the hypothesis of a such difference was tested. Three two-tailed randomized hypothesis tests for the mean of paired samples were performed, one for each characteristic. The p-values that resulted from the test are presented in Table 5.1. As can be seen, both operating force characteristics present a significant difference for the beginning and end of the coil. While there is no evidence to conclude that the lateral tension presents a significant difference between the beginning and end of the coil, the p-value is low, which implies there is also no evidence to indicate an absence of statistical difference.

Table 5.1 - Hypothesis test p-value results for comparison between means at the beginning and end of coil

Characteristic	P-value
Lateral Tension	0.1199
Peaks of Operating Force	0.0031
Average Operating Force	0.0003

Following this, control charts for the operating characteristics on the beginning and end of the coil were developed. These charts presented a lower amount of out-of-control samples, mainly for the operating force charts, as expected with the results of the hypothesis test. However, they still evidenced an out-of-control process, which reinforced the necessity of the execution of the DoE. The proportion of out-of-control samples for each characteristic and chart are presented in Table 5.2.

Table 5.2 - Percentage of out-of-control samples of control charts

	Out-of-control samples (%)		
Characteristic	Combined Data	Beginning of Coil	End of Coil
Lateral Tension	31.6	21.1	34.2
Peaks of Operating Force	62.2	35.1	45.9
Average Operating Force	56.8	43.2	37.8

It was decided to apply the DoE to the coiling process, which is the first process of the Dyeing Area, where the crude chain is coiled, before entering the dyeing process. Each run of the experiment is materialized by a coil lot. The controlled factors of the DoE are the coiling tension, pitch, rotation speed, and chain position. There are some held-constant factors, such as the zipper type, coiling machine, dyeing machine, and assembly line. Other factors that are uncontrollable, such as the raw material, or exact temperature, pressure, and water flow in the dyeing machine, are considered noise factors. The randomization of the DoE is expected to deal with these noise factors. The levels for the coiling tension, which increases along the coiling process, are presented in Table 5.3, while the levels for the other factors are presented in Table 5.4.

Table 5.3 - Levels for coiling tension in screening experiment

	Mass (g)	
Length	**Level 1**	**Level 2**
[0; 1000[	750	825
[1000; 2000[	850	935
[2000; 4000[	900	990
[4000; 7000]	1000	1100

Table 5.4 - Factor levels for all factors except coiling tension in screening experiment

Factor	**Level 1**	**Level 2**
Pitch (mm)	9	11
Rotation Speed (m/ min)	110	130
Chain Position	Up	Down

There are three distinct responses for the DoE. First, the total number of nonconformities, gathered through the use of check sheets, is evaluated. The second response consists in number of zippers that present curvature, which is evaluated with a customized gauge. The third and last response is the number of false negatives on the automated inspection, directly correlated with the wrinkling or folding of the ribbon of the zippers. The DoE aims at minimizing all responses. Additionally, flammability and color are measured as usual, to ensure no abnormal behavior is introduced with any parametrization. The same applies to operating characteristics. However, these are also evaluated before the dyeing process, to observe whether it introduces any alteration on said characteristics. Additionally, these characteristics are measured on the crude chain, to verify whether a correlation between the operating characteristics of the raw material and of the chain after the Dyeing Process exists.

The DoE is a 2^{4-1} fractional factorial design with 2 replicas, which totals 16 runs, with the rotation speed as the generated factor. Its design matrix is presented in Table 5.5. The runs order was later randomized to guarantee a minimization of influence of the noise factors.

Table 5.5 - Design matrix for the DoE

	A	**B**	**C**	**D ≡ ABC**	**Response**	
Experience	**Coiling tension**	**Pitch**	**Chain position**	**Rotation speed**	**Replica 1**	**Replica 2**
1	750 to 1000	9	Up	110	Y_{11}	Y_{21}
2	825 to 1100	9	Up	130	Y_{12}	Y_{22}
3	750 to 1000	11	Up	130	Y_{13}	Y_{23}
4	825 to 1100	11	Up	110	Y_{14}	Y_{24}
5	750 to 1000	9	Down	130	Y_{15}	Y_{25}
6	825 to 1100	9	Down	110	Y_{16}	Y_{26}
7	750 to 1000	11	Down	110	Y_{17}	Y_{27}
8	825 to 1100	11	Down	130	Y_{18}	Y_{28}

While deciding upon the responses for the DoE, a hypothesis test that aimed at determining whether there was a correlation between curvature or folding/ wrinkling of the ribbon and the operating force was executed. While the results were considered invalid due to the sampling of the data, this hypothesis test evidenced two aspects. First, it evidenced a wide difference in operating force between different coil lots. Second, and most important, it evidenced a statistical difference in the operating force for the opening and closing of zippers.

During the analysis of the process, the different sampling plans were studied. No necessity to review operating, visual or safety characteristics was observed. The reception sampling plan, however, was identified as insufficient, due to a low verification on large size lots. The reception inspection plan has not yet been reviewed, as a result of the prioritization of the DoE and control charts implementation over it.

Another key element of the sampling plans, the Quality Wall, where there is verification of many characteristics on all zippers, is expected to have its importance diminish following the implementation of the DoE. This might lead to an abolishment of the Quality Wall.

5.2. Suggestions for Future Work

The completion of the DoE, and analysis of the results, is a crucial task. Based on the results of the DoE, a decision will then have to be taken over the necessity of a second DoE, with less factors and more levels, as an attempt to identify the best combination for the coiling process. Consequently, nonconformities are expected to diminish, as well as the Quality Wall necessity, which will then need to be evaluated in collaboration with the company team.

After the DoE completion, the control charts need to be revisited. A reduction of out-of-control samples is expected, which should allow for the implementation of phase II control charts. This will lead to a controlled process and an easier detection of special cause variation, which will, in term, lead to an easier identification of root causes for these occurrences.

An aspect that was overlooked when developing the control charts, and later discovered through the residual analysis of the ANOVA test, was the significant difference between the opening and closing of zippers for the operating force. Even though it does not seem like that would have altered the conclusion that the process is statistically out-of-control, given the high proportion of out-of-control samples, the new charts for the operating force characteristics should separate the data for the opening and closing of the zippers.

Posterior to the completion of these tasks, the reevaluation of the inspection plan at reception is recommended, to guarantee a higher confidence in the identification of nonconformities in the raw materials, if these are to happen.

Furthermore, some other minor tests might be implemented to attempt to reduce some common nonconformities that might subsist after the DoE implementation. To do so, a tree diagram for root cause identification will be analyzed in collaboration with the company team, and subsequent tests will be decided upon. There are some foreseen tests on the relationship between homogenization of the dye "recipe" and color of the chain, as well as between operating force and slider insertion nonconformities, and defective sliders and slider nonconformities.

REFERENCES

Antony, J., S. Coleman, D. C. Montgomery, M. J. Anderson, and R. T. Silvestrini. 2011. "Design of Experiments for Non-Manufacturing Processes: Benefits, Challenges and Some Examples." *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* 225 (11): 2078–87. https://doi.org/10.1177/0954405411395857.

Antony, Jiju, Stavros Karamperidis, Frenie Antony, and Elizabeth A. Cudney. 2019. "Understanding and Evaluating Teaching Effectiveness in the UK Higher Education Sector Using Experimental Design: A Case Study." *International Journal of Quality and Reliability Management* 36 (2): 202–16. https://doi.org/10.1108/IJQRM-01-2018-0011.

Antony, Jiju, Elisabeth Viles, Alexandre Fonseca Torres, Taynara Incerti de Paula, Marcelo Machado Fernandes, and Elizabeth A. Cudney. 2020. "Design of Experiments in the Service Industry: A Critical Literature Review and Future Research Directions." *The TQM Journal* 32 (6): 1159–75. https://doi.org/10.1108/TQM-02-2020-0026.

Arafeh, Mazen, Mahmoud A. Barghash, Eman Sallam, and Alaa AlSamhouri. 2014. "Six Sigma Applied to Reduce Patients' Waiting Time in a Cancer Pharmacy." *International Journal of Six Sigma and Competitive Advantage* 8 (2): 105–24. https://doi.org/10.1504/IJSSCA.2014.064256.

ASQ. n.d. "7 Basic Quality Tools: Quality Management Tools | ASQ." Accessed October 31, 2021a. https://asq.org/quality-resources/seven-basic-quality-tools.

ASQ. n.d. "History of Quality - Quality Management History | ASQ." Accessed February 8, 2021b. https://asq.org/quality-resources/history-of-quality.

ASQ. n.d. "Six Sigma Definition - What Is Lean Six Sigma? | ASQ." Accessed February 13, 2021c. https://asq.org/quality-resources/six-sigma.

ASQ. n.d. "What Are the Seven New Management & Planning Tools? | ASQ." Accessed October 31, 2021d. https://asq.org/quality-resources/new-management-planning-tools.

ASQ. n.d. "What Is Statistical Process Control? SPC Quality Tools | ASQ." Accessed March 29, 2021e. https://asq.org/quality-resources/statistical-process-control.

Bell, Gordon H., Johannes Ledolter, and Arthur J. Swersey. 2006. "Experimental Design on the Front Lines of Marketing: Testing New Ideas to Increase Direct Mail Sales." *International Journal of Research in Marketing* 23 (3): 309–19.

Berk, Joseph, and Susan Berk. 2000. "Chapter 10 - Statistical Process Control: A Secret Weapon We Gave to Japan…." In *Quality Management for the Technology Sector*, edited by Joseph Berk and Susan Berk, 88–105. Woburn: Butterworth-Heinemann. https://doi.org/10.1016/B978-075067316-7/50010-3.

Bleier, Alexander, Colleen M. Harmeling, and Robert W. Palmatier. 2019. "Creating Effective Online Customer Experiences." *Journal of Marketing* 83 (2): 98–119. https://doi.org/10.1177/0022242918809930.

Chackelson, Claudia, Ander Errasti, David Ciprés, and Fernando Lahoz. 2013. "Evaluating Order Picking Performance Trade-Offs by Configuring Main Operating Strategies in a Retail Distributor: A Design of Experiments Approach." *International Journal of Production Research* 51 (20): 6097–6109. https://doi.org/10.1080/00207543.2013.796421.

Coleman, David E., and Douglas C. Montgomery. 1993. "A Systematic Approach to Planning for a Designed Industrial Experiment." *Technometrics* 35 (1): 1–12. https://doi.org/10.2307/1269280.

Crosby, Philip B. 1980. *Quality Is Free: The Art of Making Quality Certain*. 1st ed. New York: Mentor; Cover Bent edition.

Dudewicz, Edward J. 1998. "Basic Statistical Methods." In *Quality-Control Handbook*, edited by Joseph M. Juran and A. Blanton Godfrey, 5th ed., 44.1-44.112. New York: McGraw-Hill.

Durakovic, Benjamin. 2017. "Design of Experiments Application, Concepts, Examples: State of the Art." *Periodicals of Engineering and Natural Sciences (PEN)* 5 (3). https://doi.org/10.21533/pen.v5i3.145.

Fisher, N. I., and V. N. Nair. 2009. "Quality Management and Quality Practice: Perspectives on Their History and Their Future." *Applied Stochastic Models in Business and Industry* 25 (1): 1–28. https://doi.org/10.1002/asmb.756.

Hahn, Gerald J., William J. Hill, Roger W. Hoerl, and Stephen A. Zinkgraf. 1999. "The Impact of Six Sigma Improvement—A Glimpse into the Future of Statistics." *The American Statistician* 53 (3): 208–15. https://doi.org/10.1080/00031305.1999.10474462.

Hunter, J. Stuart, Mary G. Natrella, E. Harvey Barnett, William G. Hunter, and Truman L. Koehler. 1998. "Design and Analysis of Experiments." In *Quality-Control Handbook*, edited by Joseph M. Juran and A. Blanton Godfrey, 5th ed., 47.1-47.77. New York: McGraw-Hill.

Ishikawa, Kaoru. 1976. *Guide to Quality Control*. 2nd ed. Tokyo: Asian Productivity Organization.

ISO 7870-2:2013. 2013. *Control Charts — Part 2: Shewhart Control Charts*. Genève, Switzerland: International Organization for Standardization.

ISO 9000:2015. 2015. *Quality Management Systems — Fundamentals and Vocabulary*. Genève, Switzerland: International Organization for Standardization.

Jankovic, Aleksandar, Gaurav Chaudhary, and Francesco Goia. 2021. "Designing the Design of Experiments (DOE) – An Investigation on the Influence of Different Factorial Designs on the Characterization of Complex Systems." *Energy and Buildings* 250 (November): 111298. https://doi.org/10.1016/j.enbuild.2021.111298.

Juran, Joseph M. 1998. "How to Think About Quality." In *Quality-Control Handbook*, edited by Joseph M. Juran and A. Blanton Godfrey, 5th ed., 2.1-2.18. New York: McGraw-Hill.

Montgomery, Douglas C. 2012. *Design and Analysis of Experiments*. 8th ed. Hoboken, New Jersey: John Wiley & Sons, Inc.

Montgomery, Douglas C. 2012. 2012. *Introduction to Statistical Quality Control*. 7th ed. Hoboken, New Jersey: John Wiley & Sons, Inc.

Oakland, John S. 2008. "Chapter 13 - Designing the Statistical Process Control System." In *Statistical Process Control*, edited by John S. Oakland, 6th ed., 336–58. Oxford: Butterworth-Heinemann. https://doi.org/10.1016/B978-0-7506-6962-7.50016-X.

Pereira, Zulema Paula do Perpétuo Socorro Lopes, and José Fernando Gomes Requeijo. 2012. *Qualidade: Planeamento e Controlo Estatístico de Processos*. 2nd ed. Lisbon: Fundação da Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.

Raffaldi, John, and William D. Kappele. 2011. "Getting Started With Design of Experiments." Quality Magazine. March 25, 2011. https://www.qualitymag.com/articles/87730-getting-started-with-design-of-experiments?v=preview.

Sarasohn, Homer M., and Charles A. Protzman. 1948. *The Fundamentals of Industrial Management*. Civil Communications Section, GHQ, SCAP.

Snee, Ronald D. 1985. "Computer-Aided Design of Experiments—Some Practical Experiences." *Journal of Quality Technology* 17 (4): 222–36. https://doi.org/10.1080/00224065.1985.11978972.

Taguchi, Genichi, and Don Clausing. 1990. "Robust Quality." *Harvard Business Review*, January 1, 1990. https://hbr.org/1990/01/robust-quality.

Tague, Nancy R. 2005. *The Quality Toolbox*. 2nd ed. Milwaukee: ASQ Quality Press.

Toutenburg, Helge, and Shalabh. 2009. *Statistical Analysis of Designed Experiments*. 3rd ed. New York: Springer. https://doi.org/10.1007/978-1-4419-1148-3.

Wadsworth, Harrison M. 1998. "Statistical Process Control." In *Quality-Control Handbook*, edited by Joseph M. Juran and A. Blanton Godfrey, 5th ed., 45.1-45.29. New York: McGraw-Hill.

APPENDIXES

Appendix A - Constants for Control Charts

Table A.1 - Constants for control charts (Adapted from Pereira and Requeijo, 2012)

	Chart for Averages						Chart for Standard Deviations						Chart for Amplitudes			
Observations	Factors for Control Limits			Factors for Center Line			Factors for Control Limits			Factors for Center Line			Factors for Control Limits			
in Sample, n	A	A2	A3	c4	1/c4	B3	B4	B5	B6	d2	1/d2	d3	D1	D2	D3	D4
2	2.121	1.880	2.659	0.7979	1.2533		03.267		02.606	1.128	0.8865	0.853		03.686		03.267
3	1.732	1.023	1.954	0.8862	1.1284		02.568		02.276	1.693	0.5907	0.888		04.358		02.574
4	1.500	0.729	1.628	0.9213	1.0854		02.266		02.088	2.059	0.4857	0.880		04.698		02.282
5	1.342	0.577	1.427	0.9400	1.0638		02.089		01.964	2.326	0.4299	0.864		04.918		02.114
6	1.225	0.483	1.287	0.9515	1.0510	0.030	1.970	0.029	1.874	2.534	0.3946	0.848		05.078		02.004
7	1.134	0.419	1.182	0.9594	1.0423	0.118	1.882	0.113	1.806	2.704	0.3698	0.833	0.204	5.204	0.076	1.924
8	1.061	0.373	1.099	0.9650	1.0363	0.185	1.815	0.179	1.751	2.847	0.3512	0.820	0.388	5.306	0.136	1.864
9	1.000	0.337	1.032	0.9693	1.0317	0.239	1.761	0.232	1.707	2.970	0.3367	0.808	0.547	5.393	0.184	1.816
10	0.949	0.308	0.975	0.9727	1.0281	0.284	1.716	0.276	1.669	3.078	0.3249	0.797	0.687	5.469	0.223	1.777
11	0.905	0.285	0.927	0.9754	1.0252	0.321	1.679	0.313	1.637	3.173	0.3152	0.787	0.811	5.535	0.256	1.744
12	0.866	0.266	0.886	0.9776	1.0229	0.354	1.646	0.346	1.610	3.258	0.3069	0.778	0.922	5.594	0.283	1.717
13	0.832	0.249	0.850	0.9794	1.0210	0.382	1.618	0.374	1.585	3.336	0.2998	0.770	1.025	5.647	0.307	1.693
14	0.802	0.235	0.817	0.9810	1.0194	0.406	1.594	0.399	1.563	3.407	0.2935	0.763	1.118	5.696	0.328	1.672
15	0.775	0.223	0.789	0.9823	1.0180	0.428	1.572	0.421	1.544	3.472	0.2880	0.756	1.203	5.741	0.347	1.653
16	0.750	0.212	0.763	0.9835	1.0168	0.448	1.552	0.440	1.526	3.532	0.2831	0.750	1.282	5.782	0.363	1.637
17	0.728	0.203	0.739	0.9845	1.0157	0.466	1.534	0.458	1.511	3.588	0.2787	0.744	1.356	5.820	0.378	1.622
18	0.707	0.194	0.718	0.9854	1.0148	0.482	1.518	0.475	1.496	3.640	0.2747	0.739	1.424	5.856	0.391	1.608
19	0.688	0.187	0.698	0.9862	1.0140	0.497	1.503	0.490	1.483	3.689	0.2711	0.734	1.487	5.891	0.403	1.597
20	0.671	0.180	0.680	0.9869	1.0133	0.510	1.490	0.504	1.470	3.735	0.2677	0.729	1.549	5.921	0.415	1.585
21	0.655	0.173	0.663	0.9876	1.0126	0.523	1.477	0.516	1.459	3.778	0.2647	0.724	1.605	5.951	0.425	1.575
22	0.640	0.167	0.647	0.9882	1.0119	0.534	1.466	0.528	1.448	3.819	0.2618	0.720	1.659	5.979	0.434	1.566
23	0.626	0.162	0.633	0.9887	1.0114	0.545	1.455	0.539	1.438	3.858	0.2592	0.716	1.710	6.006	0.443	1.557
24	0.612	0.157	0.619	0.9892	1.0109	0.555	1.445	0.549	1.429	3.895	0.2567	0.712	1.759	6.031	0.451	1.548
25	0.600	0.153	0.606	0.9896	1.0105	0.565	1.435	0.559	1.420	3.931	0.2544	0.708	1.806	6.056	0.459	1.541

$For\ n > 25$

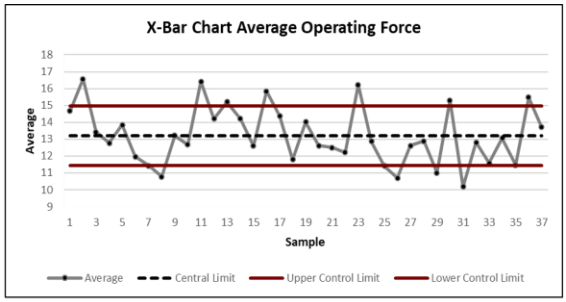
$$A = \frac{3}{\sqrt{n}} \qquad A_3 = \frac{3}{c_4\sqrt{n}} \qquad c_4 = \frac{4\,(n-1)}{4n-3} \qquad B_3 = 1 - \frac{3}{c_4\sqrt{2\,(n-1)}} \qquad B_4 = 1 + \frac{3}{c_4\sqrt{2\,(n-1)}}$$

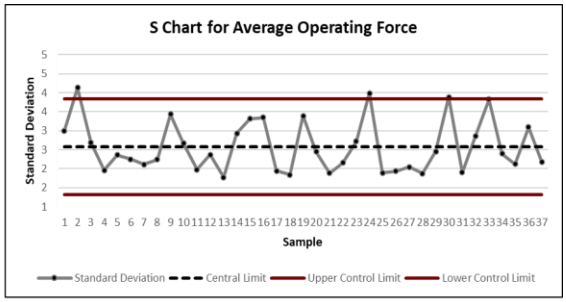
$$B_5 = c_4 - \frac{3}{\sqrt{2\,(n-1)}} \qquad B_6 = c_4 + \frac{3}{\sqrt{2\,(n-1)}}$$

Appendix B - Control Charts

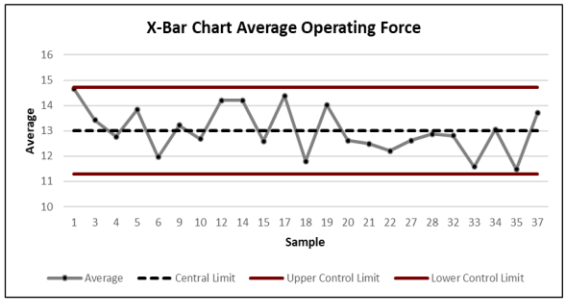
Average	12,88
Standard deviation	2,37
n	20
k	16
c4	0,9869
B4	1,490424
B3	0,509576
UCLs	3,53
LCLs	1,21
A3	0,679725
UCLx	14,49
LCLx	11,27
Out-of-control samples	0
Proportion (%)	0
Total of out-of-control samples	21
Proportion (%)	56,8

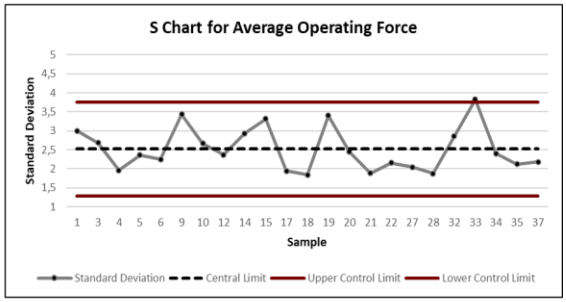
1^{st} iteration



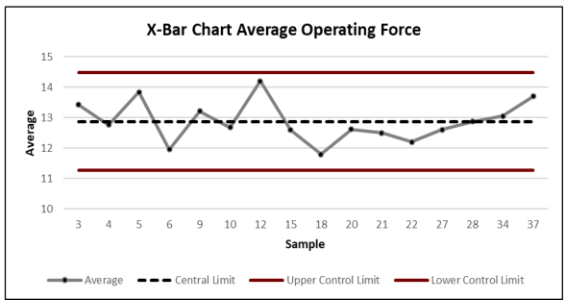


2^{nd} iteration





Last iteration



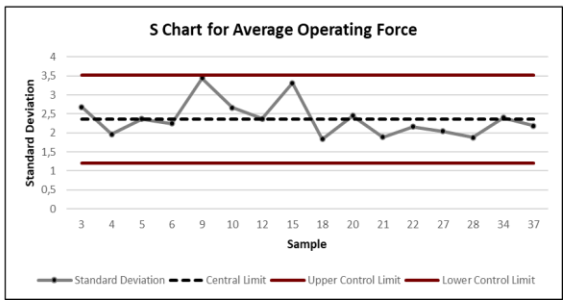


Figure B.1 - Control charts for average operating force

Average	15,37
Standard deviation	2,88
n	20
k	14
c4	0,9869
B4	1,4904
B3	0,5096
UCLs	4,29
LCLs	1,47
A3	0,6797
UCLx	17,33
LCLx	13,41
Out-of-control samples	0
Proportion (%)	0
Total of out-of-control samples	23
Proportion (%)	62,2

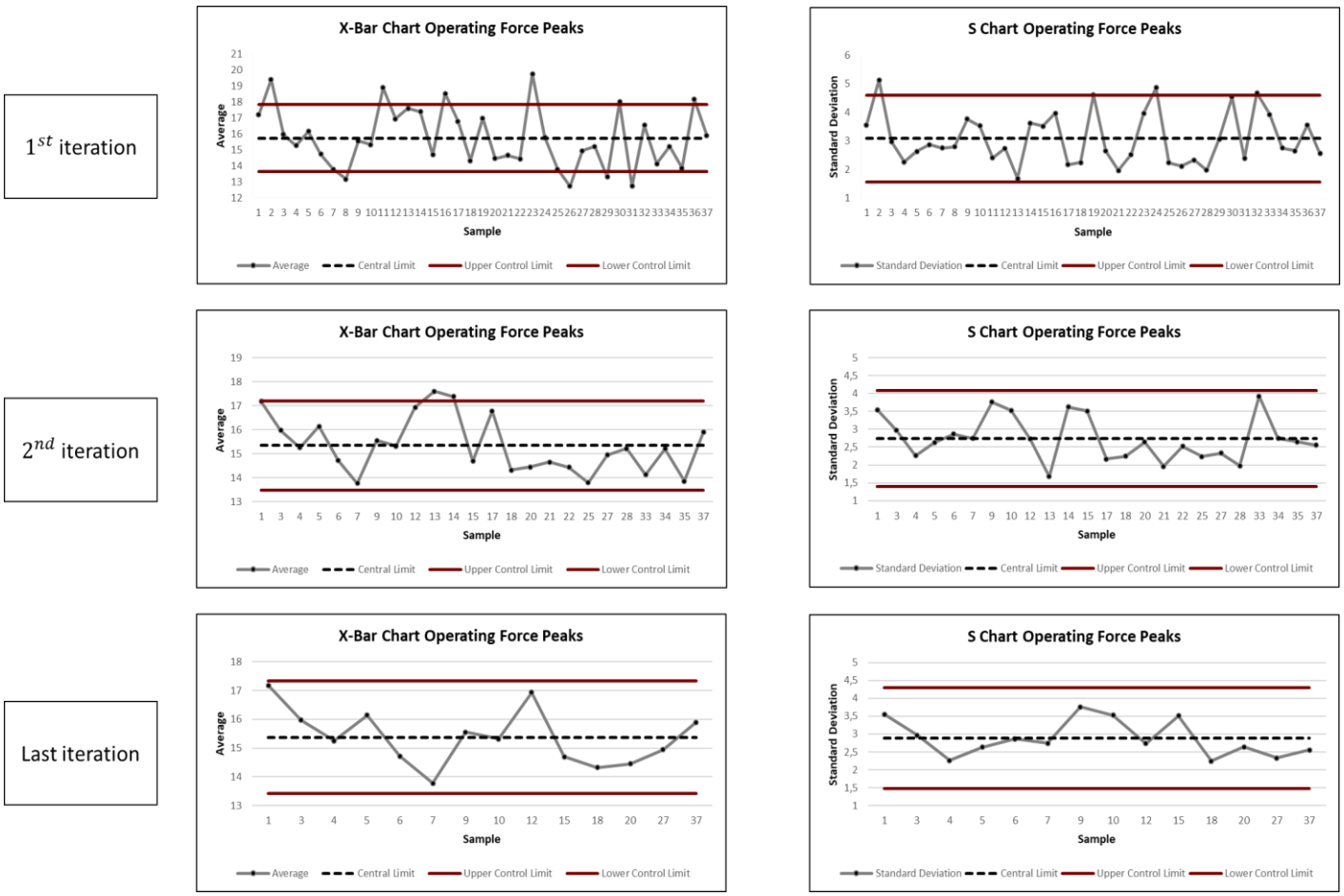
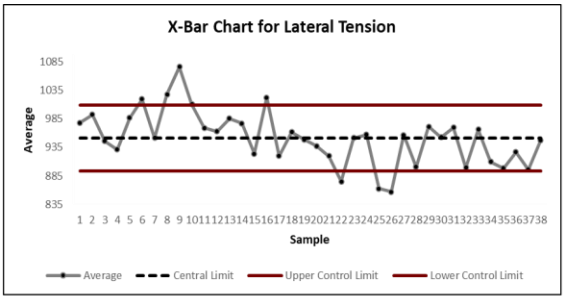
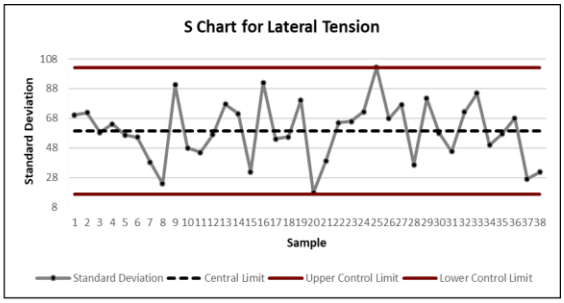


Figure B.2 - Control charts for peaks of operating force

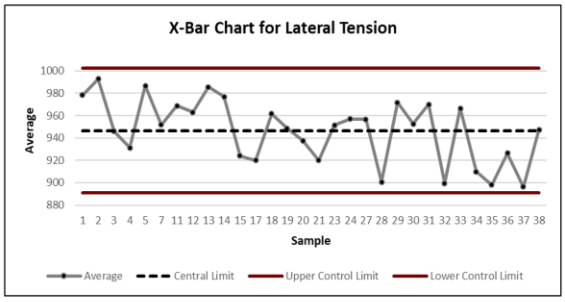
Average	952,57
Standard deviation	58,21
n	10
k	26
c4	0,9727
B4	1,715738
B3	0,284262
UCLs	99,87
LCLs	16,55
A3	0,975309
UCLx	1009,33
LCLx	895,80
Out-of-control samples	0
Proportion (%)	0
Total of out-of-control samples	12
Proportion (%)	31,6

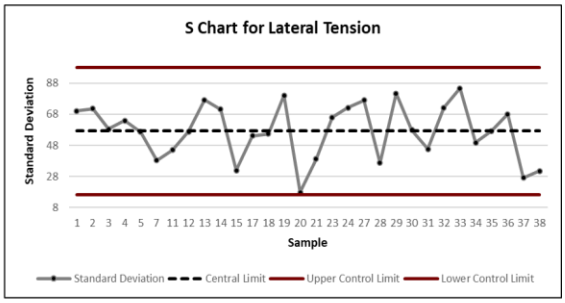
1^{st} iteration



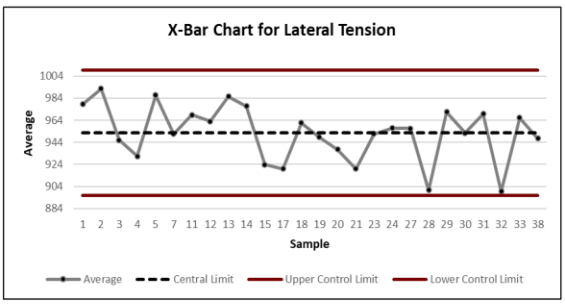


2^{nd} iteration





Last iteration



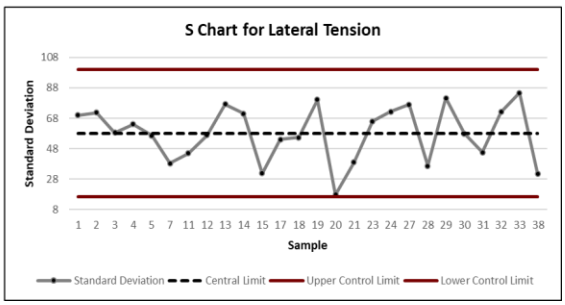


Figure B.3 - Control charts for lateral tension

Average	12,12
Standard deviation	1,90
n	10
k	21
c4	0,9727
B4	1,7157
B3	0,2843
UCLs	3,26
LCLs	0,54
A3	0,9753
UCLx	13,97
LCLx	10,26
Out-of-control samples	0
Proportion (%)	0
Total of out-of-control samples	16
Proportion (%)	43,2

Figure B.4 - Control charts for average operating force of the beginning of the coil

Average	14,04
Standard deviation	2,36
n	10
k	23
c4	0,9727
B4	1,715738
B3	0,284262
UCLs	4,05
LCLs	0,67
A3	0,975309
UCLx	16,3
LCLx	11,7
Out-of-control samples	0
Proportion (%)	0
Total of out-of-control samples	14
Proportion (%)	37,8

Figure B.5 - Control charts for average operating force of the end of the coil

Average	14,43
Standard deviation	2,17
n	10
k	24
c4	0,9727
B4	1,715738
B3	0,284262
UCLs	3,73
LCLs	0,62
A3	0,975309
UCLx	16,55
LCLx	12,31
Out-of-control samples	0
Proportion (%)	0
Total of out-of-control samples	13
Proportion (%)	35,1

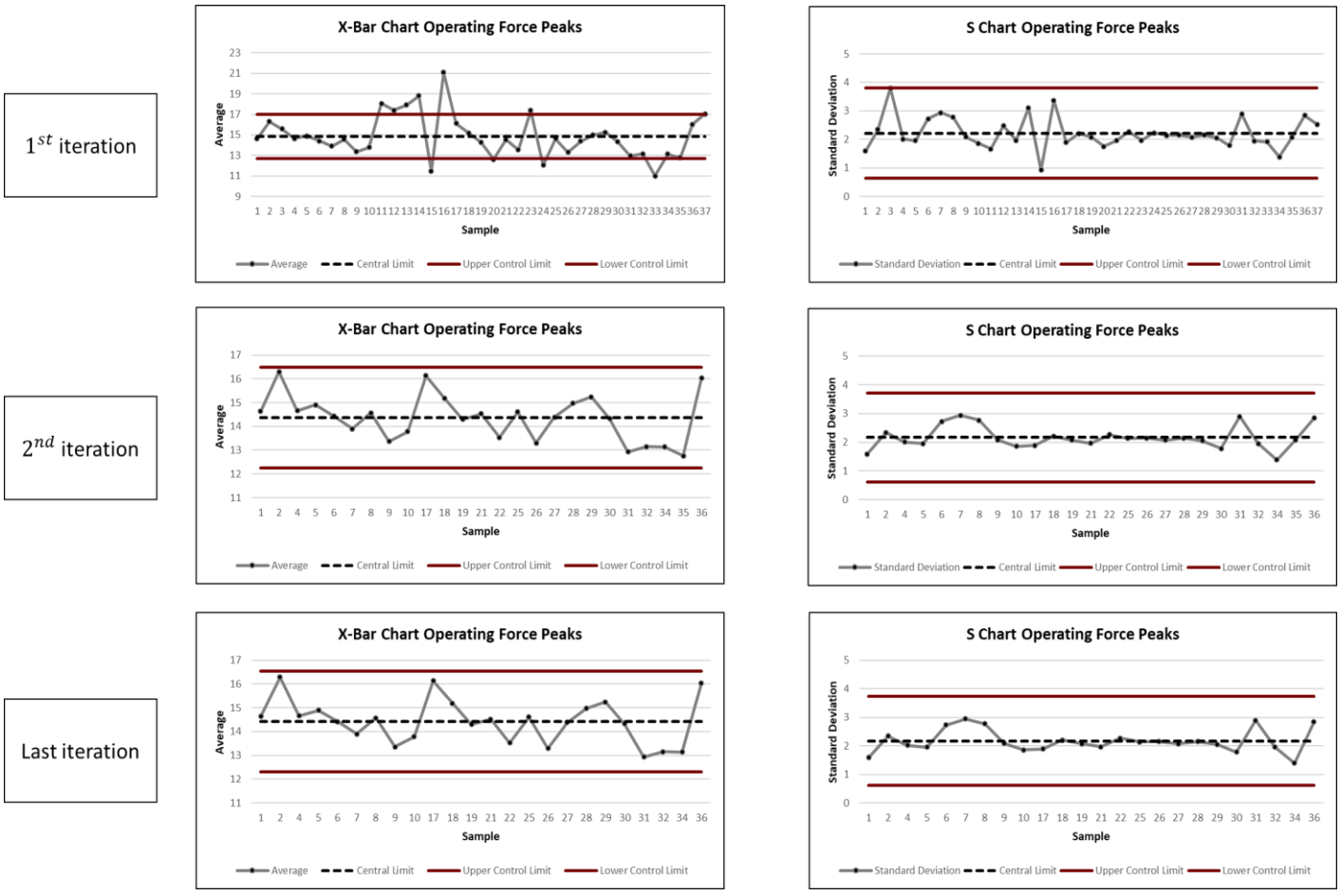


Figure B.6 - Control charts for peaks of operating force of the beginning of the coil

Average	16,26
Standard deviation	2,46
n	10
k	20
c4	0,9727
B4	1,715738
B3	0,284262
UCLs	4,22
LCLs	0,70
A3	0,975309
UCLx	18,66
LCLx	13,85
Out-of-control samples	0
Proportion (%)	0
Total of out-of-control samples	17
Proportion (%)	45,9

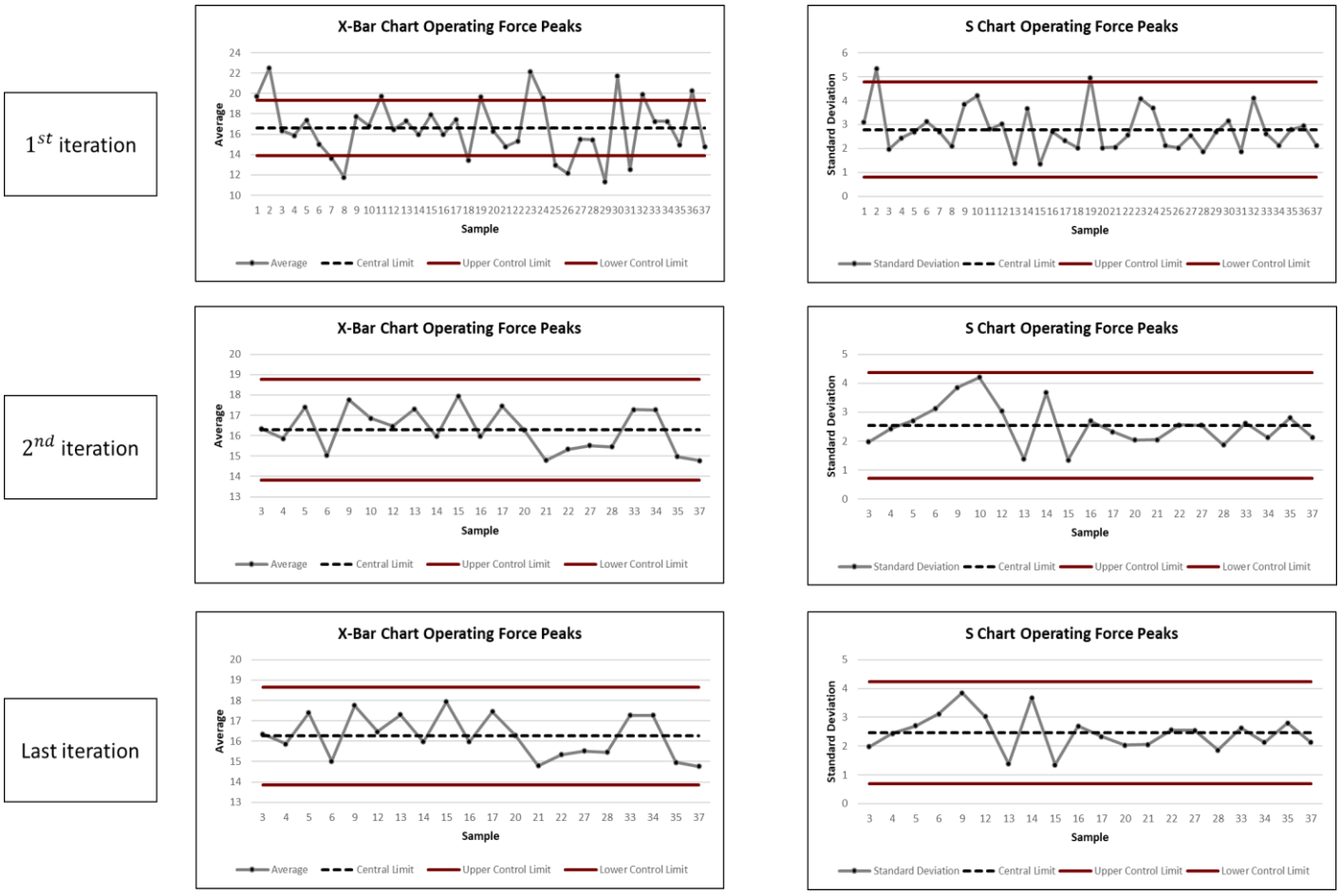


Figure B.7 - Control charts for peaks of operating force of the end of the coil

Average	940,11
Standard deviation	44,01
n	5
k	30
c4	0,94
B4	2,0889
B3	0
UCLs	91,93
LCLs	0
A3	1,4273
UCLx	1002,93
LCLx	877,30
Out-of-control samples	0
Proportion (%)	0
Total of out-of-control samples	8
Proportion (%)	21,1

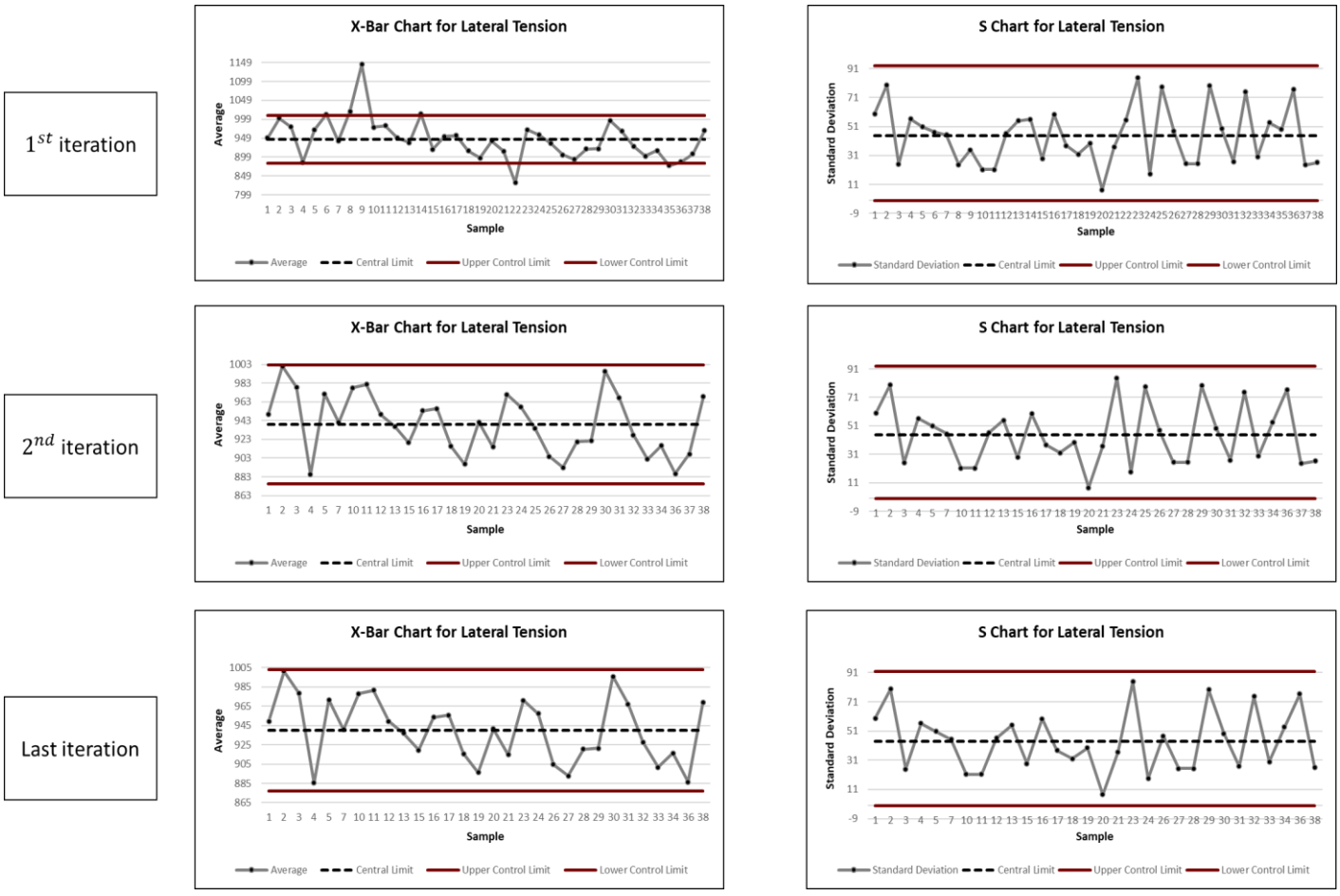


Figure B.8 - Control charts for lateral tension of the beginning of the coil

Average	957,12
Standard deviation	51,86
n	5
k	25
c4	0,94
B4	2,0889
B3	0
UCLs	108,32
LCLs	0
A3	1,4273
UCLx	1031,13
LCLx	883,11
Out-of-control samples	0
Proportion (%)	0
Total of out-of-control samples	13
Proportion (%)	34,2

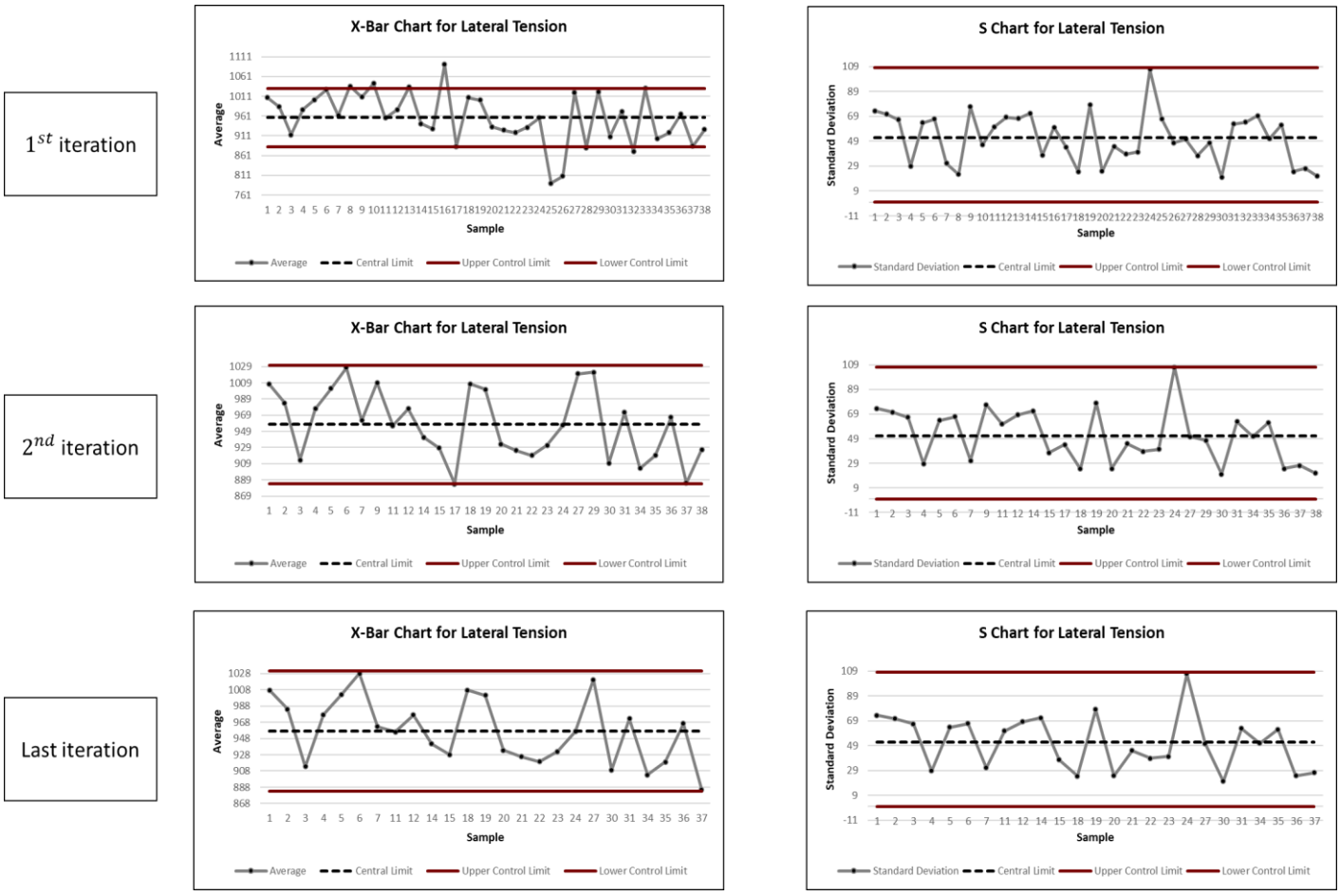


Figure B.9 - Control charts for lateral tension of the end of the coil

Appendix C - Hypothesis test results for comparison of averages between beginning and end of the coil

Table C.1 - Hypothesis test results of comparison between beginning and end of the coil for average operating force characteristic

	Average at beginning of coil	Average at end of coil
Average	11,92	13,59
Variance	0,71	1,90
Observations	18	18
t statistic	-4,4715	
p-value	0,0003	
t critical value	2,1098	

Table C.2 - Hypothesis test results of comparison between beginning and end of the coil for peaks of operating force characteristic

	Average at beginning of coil	Average at end of coil
Average	14,56	16,20
Variance	1,64	2,85
Observations	20	20
t statistic	-3,3850	
p-value	0,0031	
t critical value	2,0930	

Table C.3 - Hypothesis test results of comparison between beginning and end of the coil for lateral tension characteristic

	Average at beginning of coil	Average at end of coil
Average	944,65	961,73
Variance	1607,35	2647,32
Observations	33	33
t statistic	-1,5977	
p-value	0,1199	
t critical value	2,0369	

Appendix D - Randomized Design Matrix for the Design of Experiments

Table D.1 - Randomized design matrix

		A	B	C	D	
Experience	**Replica**	**Coiling tension**	**Pitch**	**Chain position**	**Rotation speed**	**Response**
13	2	750 to 1000	9	Down	130	
16	2	825 to 1100	11	Down	130	
12	2	825 to 1100	11	Up	110	
9	2	750 to 1000	9	Up	110	
8	1	825 to 1100	11	Down	130	
1	1	750 to 1000	9	Up	110	
6	1	825 to 1100	9	Down	110	
14	2	825 to 1100	9	Down	110	
3	1	750 to 1000	11	Up	130	
11	2	750 to 1000	11	Up	130	
10	2	825 to 1100	9	Up	130	
5	1	750 to 1000	9	Down	130	
4	1	825 to 1100	11	Up	110	
2	1	825 to 1100	9	Up	130	
15	2	750 to 1000	11	Down	110	
7	1	750 to 1000	11	Down	110	

Appendix E - Nonconformities Tree Diagram

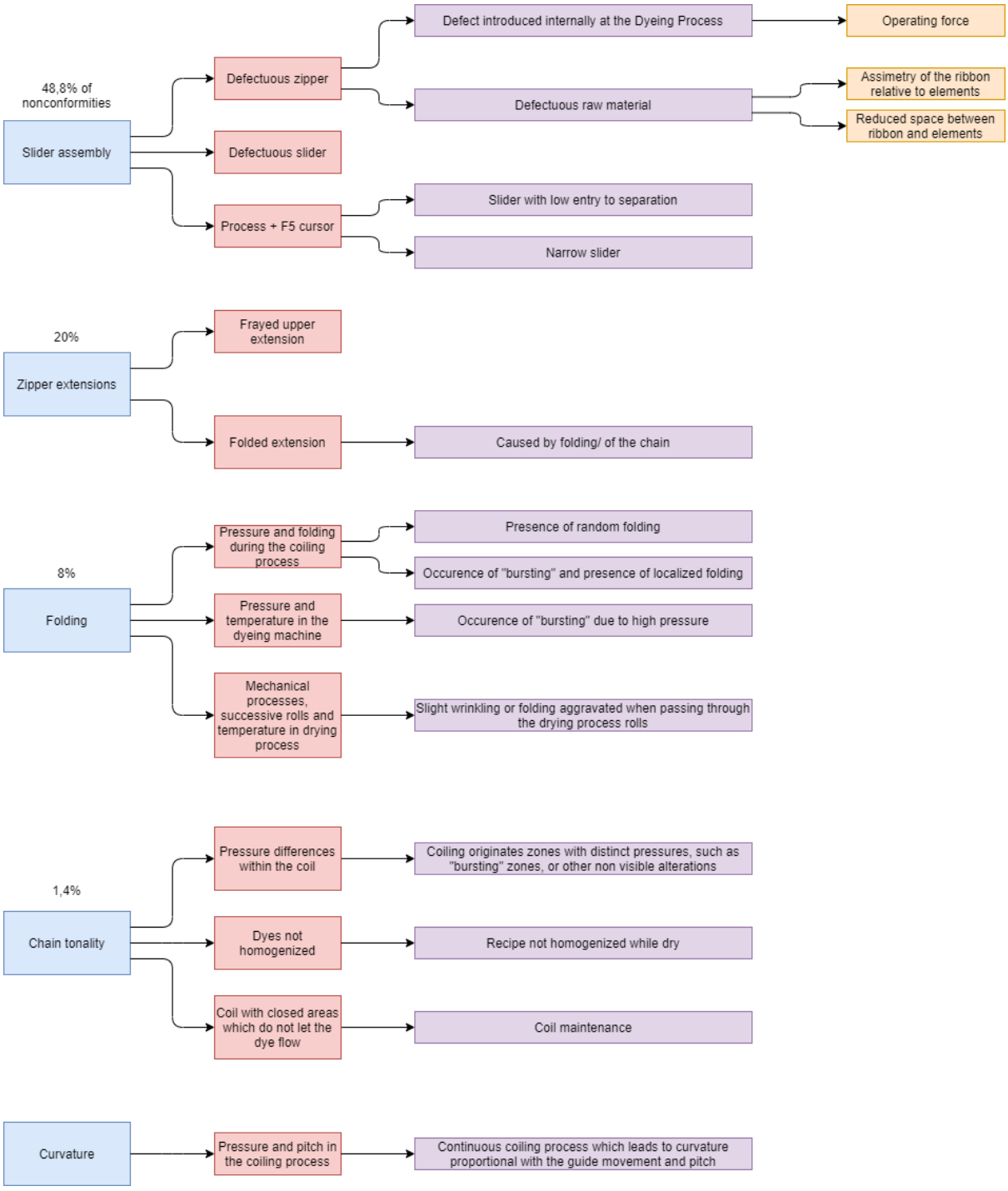


Figure E.10- Nonconformities tree diagram

ROBIN LOUIS ROBERT STREF COUEDEL

APPLICATION OF STATISTICAL CONTROL TOOLS IN THE PRODUCTION LINE OF AN AUTOMOTIVE INDUSTRY SUPPLIER

〈2021〉

NOVA SCHOOL OF SCIENCE & TECHNOLOGY