UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

Daily Morales

SHEWHART-TYPE LOCATION CONTROL CHART BASED ON $\hat{C}_p$ AND $\hat{C}_{pk}$ APPLIED TO PROFILE MONITORING

São Carlos - SP
2021

Daily Morales

SHEWHART-TYPE LOCATION CONTROL CHART BASED ON $\hat{C}_p$ AND $\hat{C}_{pk}$ APPLIED TO PROFILE MONITORING

Tese de doutorado apresentada ao Programa de Pós-graduação em Engenharia de Produção da Universidade Federal de São Carlos (UFSCar), para obtenção do título de Doutor em Engenharia de Produção

Orientador: Prof. Dr. Pedro Carlos Oprime

São Carlos - SP
2021

UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Engenharia de Produção

Folha de Aprovação

Defesa de Tese de Doutorado do candidato Daily Morales, realizada em 01/12/2021.

Comissão Julgadora:

Prof. Dr. Pedro Carlos Oprime (UFSCar)

Profa. Dra. Fabiane Letícia Lizarelli (UFSCar)

Prof. Dr. Mario Orestes Aguirre González (UFRN)

Prof. Dr. Jorge Alberto Achcar (USP)

Prof. Dr. Felipe Schoemer Jardim (UFF)

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.
O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Engenharia de Produção.

This work is dedicated to my wife Rosangela
and my children Gustavo and Sara
for their continuous and unparalleled
love, patience, and support.

Acknowledgements

First, I would like to thank God for His mercy and strength, carrying me in times when I thought of giving up.

I would like to thank Prof. Dr. Pedro Carlos Oprime for his invaluable supervision and support during my Doctorate Degree program. His guidance, knowledge, experience, and support were fundamental to the conclusion of this work.

I also would like to thank the engineer Samuel Bozzi Baco, who mediated access to the company's data and made valuable suggestions in the conduct of my work.

I would like to thank my friends and colleagues from the Production Engineering Department of the State University of Maringá, which "hold the fort" while I was away to do my doctorate.

Finally, I would like to express my gratitude to my wife and my children. Without their tremendous understanding and encouragement in the past few years, it would be impossible for me to complete my thesis.

Abstract

Through their statistical limits, statistical control charts are essential tools for process monitoring and control, helping to interpret its stability and detect deviations when the process is under the action of special causes. This sensitivity is an advantage for processes needing precise control but is a drawback where some slack in the process is permissible. Even though small attributable causes may be inherent to the process, their removal may become impossible, impractical, or expensive, and control charts need to be "desensitized" in detecting small shifts that are not necessarily of practical significance. The increasing availability of data made possible by the intensive use of sensors and data collectors, typical of Industry 4.0, have made of profile monitoring an up-and-coming area of research, and the location control chart a standing out alternative due to its applicability and simplicity, keeping all the data information observed in each location where the profile needs to be evaluated. This thesis discusses the need for desensitization of the location control charts to ensure that signs of an out-of-control process that do not present a high risk of producing non-conforming items are ignored. Using process capability indices $\hat{C}_p$ and $\hat{C}_{pk}$, this thesis aimed to propose and evaluate a location control graph model with expanded limits based on Shewhart-type control charts, considering its practical and economic significance. By comparing the results obtained for the location control chart using the traditional approach and the location control chart with expanded limits based on $\hat{C}_p$ and $\hat{C}_{pk}$ proposed in this thesis, was observed that the expanded control limits had performed better. There was a significant reduction in the likelihood of signaling that the process is out-of-control when this is irrelevant or, at least, does not present practical or economic significance. Since the location control chart developed does not aim to control the stability of the process but to prevent the process from producing items outside the specification limits, the results demonstrate that, from the economic point of view, it constitutes a more interesting alternative to the traditional approach.

Keywords: profile monitoring, process capabiblity indices, location control chart, practical significance.

Resumo

Através de seus limites estatísticos, os gráficos de controle estatístico são ferramentas essenciais para monitoramento e controle de processos, ajudando a interpretar sua estabilidade e detectar desvios quando o processo está sob a ação de causas especiais. Essa sensibilidade é uma vantagem para os processos que precisam de controle preciso, mas é uma desvantagem onde alguma folga no processo é permitida. Embora pequenas causas atribuíveis possam ser inerentes ao processo, sua remoção pode tornar-se impossível, impraticável ou cara, e os gráficos de controle precisam ser "dessensibilizados" na detecção de pequenos desvios que não são necessariamente de significado prático. A crescente disponibilidade de dados possibilitados pelo uso intensivo de sensores e coletores de dados, típico da Indústria 4.0, tornaram o monitoramento de perfis uma área de pesquisa atualizada e o gráfico de controle de localização uma alternativa de destaque devido a sua aplicabilidade e simplicidade, mantendo todas as informações observadas em cada local onde o perfil precisa ser avaliado. Esta tese discute a necessidade de dessensibilização dos gráficos de controle de localização para garantir que os sinais de um processo fora de controle que não apresentem um alto risco de produzir itens não conforme sejam ignorados. Usando índices de capacidade de processo $\hat{C}_p$ e $\hat{C}_{pk}$, esta tese teve como objetivo propor e avaliar um modelo de gráfico de controle de localização com limites expandidos baseados nas cartas de Shewhart, considerando o seu significado prático e econômico. Ao comparar os resultados obtidos para o gráfico de controle de localização usando a abordagem tradicional e o gráfico de controle de localização com limites expandidos com base em $\hat{C}_p$ e $\hat{C}_{pk}$ propostos neste tese, observou-se que os limites de controle expandidos tiveram melhor desempenho. Houve uma redução significativa na probabilidade de sinalizar que o processo está fora de controle quando isso é irrelevante ou, pelo menos, não apresenta significância prática ou econômica. Como o gráfico de controle de local desenvolvido não visa controlar a estabilidade do processo, mas evitar que o processo produza itens fora dos limites de especificação, os resultados demonstram que, do ponto de vista econômico, constitui uma alternativa mais interessante para a abordagem tradicional.

Palavras-chave: monitoramento de perfis. índices de capabilidade de processo. cartas de controle de localização, significância prática.

List of Figures

List of Tables

List of abbreviations and acronyms

ACC	Acceptance control chart
AQL	Acceptance quality level
ARL_0	Average Run Length for In-Control Process
ARL_1	Average Run Length for Out-Of-Control Process
CL	Control Limits
C_p	Process capability index for centered process
C_{pk}	Process capability index for non centered process
C_{pk_0}	Minimum allowable capability index
EWMA	Exponentially weighted moving average
FRCC	Functional Regression Control Chart
IC	In control process
INMETRO	National Institute of Metrology
k	Sigma limits of traditional Shewhart control charts
LSL	Lower specification limit
LCL	Lower control limit
LCL_{ACC}	Lower control limit of acceptance control chart
m	Number of process samples
n	Number of profiles per sample
OC	Out of control process
p	Number of observations per profile
PCI	Process capability index
PM	Profile monitoring
r_l	Lower limit electric motor's RPM operating range

r_u	Upper limit electric motor's RPM operating range
RPL	Rejectable process level
RPM	Revolutions per minute
SPC	Statistical process control
SPM	Statistical process monitoring
SQC	Statistical quality control
USL	Upper specification limit
UCL	Upper control limit
UCL_{ACC}	Upper control limit of acceptance control chart

Contents

1 Introduction

1.1 Context of the study

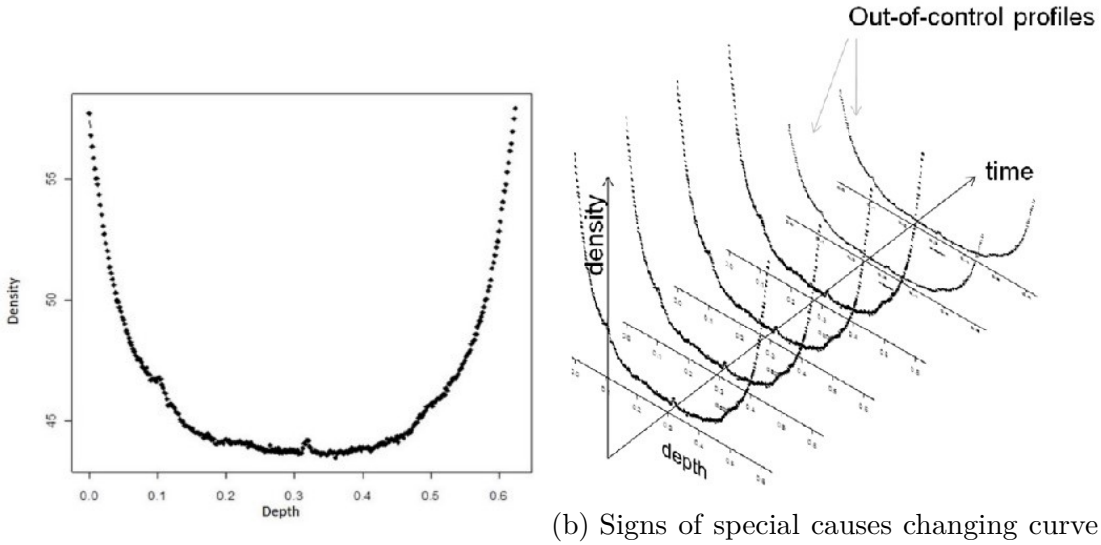
The quality of products and services is an important decision factor in most businesses. Regardless of whether the consumer is an individual, a corporation, a military defense program, or a retail store when the consumer is making purchasing decisions, he or she will be prone to consider quality with the same importance as the cost or the deadline for delivery (MARTIN; ELG; GREMYR, 2020; MONTGOMERY, 2019; RALEA et al., 2019; GARVIN, 1984).

The evolution of the production processes and the improvement of the technology of machines and production systems have made that the control procedures advanced significantly in recent times. The use of sensors, readers, and other forms of industrial automation, typical of Industry 4.0, enabled advances that not only increased speed and production volume but also allowed precise adjustments, reducing manufacturing variability, as well as obtaining data of several variables simultaneously. Such advances have enabled the gathering of large amounts of data, which provide the elements needed for product optimization and process improvements (SCHÜTZE; HELWIG; SCHNEIDER, 2018; KHAN et al., 2017; REIS; GINS, 2017; CIVERCHIA et al., 2017; REIS; SARAIVA, 2012).

Since the use of control charts proposed by Shewhart (1931), technological advances in industrial control procedures have greatly improved the quality of processes and products in the modern industry. The use of electronic data collectors facilitates the collection of data in a multitude of variables of all stages of production, and multivariate methods are necessary whenever one wants to monitor various quality variables and take advantage of any relationship between them. The use of these systems to control and monitor input, output, and process variables reduces overall variation and leads to improved accuracy, and lower production costs (GAO et al., 2020; SCHÜTZE; HELWIG; SCHNEIDER, 2018; HE; WANG, 2018; BERSIMIS; PSARAKIS; PANARETOS, 2007; MASON; YOUNG, 1998).

In the traditional applications of statistical process monitoring, the process performance considers measures of a single value or a vector of values of the quality characteristic in a given time or space. However, as Kang e Albin (2000) pioneeringly argue, there are situations in which the quality of the process or product can best be characterized by a functional relationship between the response variable, corresponding to the quality characteristic of interest, and one or more explanatory variables. (Figure 1a). In these

applications, a set of data represented by a curve (profile) is obtained for each sample. The statistical control of this profile is called profile monitoring (PM) and is used to monitor the stability of the curves over time and to determine if the changes in the curve pattern constitute signs of the presence of special causes in the process, as we can see in (Figure 1b) (NOOROSSANA; SAGHAEI; AMIRI, 2011).



(a) Profile.

(b) Signs of special causes changing curve pattern.

Figure 1 – Set of data represented by a profile.
Source: Noorossana, Saghaei e Amiri (2011)

Gao et al. (2020), He e Wang (2018), Maleki, Amiri e Castagliola (2018) and Megahed e Jones-Farmer (2015) state that the current trend in the industry involves collecting data about the quality and stability of the production process, defining a functional relationship between a response variable and one or more exploratory variables got from multi-sensors of machine tools and production systems. This functional relationship requires solutions for which traditional statistical control methods and process performance evaluation present unsatisfactory performance or may not even be applicable. The authors argue that the continuous evolution of equipment with integrated sensors, coupled with increased computing power, is increasing the ability to exploit advanced data analysis to monitor and control manufacturing processes.

Tools have been developed to provide processes capable of meeting customer expectations regarding cost and quality aspects, allowing the evaluation of process performance. Initially proposed by Kane (1986), capability indices are an essential concept in statistical process control, quantifying the relationship between actual process performance and the specification or tolerance limits of manufactured products (OPRIME et al., 2019; OPRIME; MENDES, 2017; AHMAD; ASLAM; JUN, 2016; PEARN; LIN, 2004).

Several authors such as Ghartemani, Noorossana e Niaki (2016), Keshteli et al. (2014b), Ebadi e Shahriari (2013) and Hosseinifard, Abbasi e Abdollahian (2011) point

out that among few studies on the evaluation of the capability of processes whose quality characteristic is a profile, most of them deal with profiles characterized by a linear relationship between the response variable and the exploratory variable.

With their statistical limits, statistical control charts have the role of helping to interpret the stability of the processes, and capability indices deal with the performance of these processes in terms of the quality of conformation with the design specifications of manufactured products. There has been a lot of research to study statistical methods that detect small deviations from the IC (In control process) model as quickly as possible (ATASHGAR; ABBASSI, 2020; YAO et al., 2020; AWAD, 2017; ABDELLA et al., 2016; ADIBI; BORROR; MONTGOMERY, 2015; JEN; FAN, 2014; KAZEMZADEH; NOOROSSANA; AMIRI, 2009; KAZEMZADEH; NOOROSSANA; AMIRI, 2008; ZOU; TSUNG; WANG, 2007). This high sensitivity to these small deviations can become a problem in many industrial processes. This problem has been identified and has been discussed (WOODALL; FALTIN, 2019). According to the authors, small attributable causes may be inherent to the process, and their removal may become impossible, impractical, or expensive. Thus, as the authors argue, control charts need to be "desensitized", pointing out that the current approaches of evaluating and comparing methods in control charting have high sensitivity in detecting small shifts that are not necessarily of practical significance.

This work contributes to the statistical process monitoring field through the use of control charts considering the economic and practical aspects. The proposed thesis introduces a new methodology for mathematical modeling and statistical interpretation of a control chart when a functional relationship defines the quality characteristic, and the process presents a high capability. This problem can be considered a problem of articulation, modeling, and decision-making in real systems, which is little explored in the literature and is not fully modeled in statistical terms. It is also noteworthy that the classical approach is not suitable for processes with a high capability index (C_p; C_{pk}). Shewhart control charts do not adequately monitor processes with small variability and high capability (outline situation of this research). This aspect is not considered in scientific articles, books, and recent works published in scientific journals, and the findings should make an important contribution to the field of production engineering.

1.2 Research question and objectives

According to Deming (2000), Juran e Godfrey (1998) and Feigenbaum (1991), who have developed the structured problem-solving methodology in the early years of the quality revolution, a *problem* is defined as a discrepancy between an existing standard or expectation and the actual situation. To improve a company's performance, it is necessary

to successfully solve problems that are causing dissatisfaction for internal or external customers. In this perspective, a knowledge gap can be considered a problem to be solved in the same way as an inaccurate or nonexistent procedure.

According to Diyoke, Okeke e Aniagwu (2016), an essential feature of an electric motor is the relationship between its torque and its revolution in revolutions per minute (RPM). In this case, the torque produced by each sampled electric motor is considered a response variable while their corresponding revolutions (RPM) is considered as an explanatory variable (Figure 2).

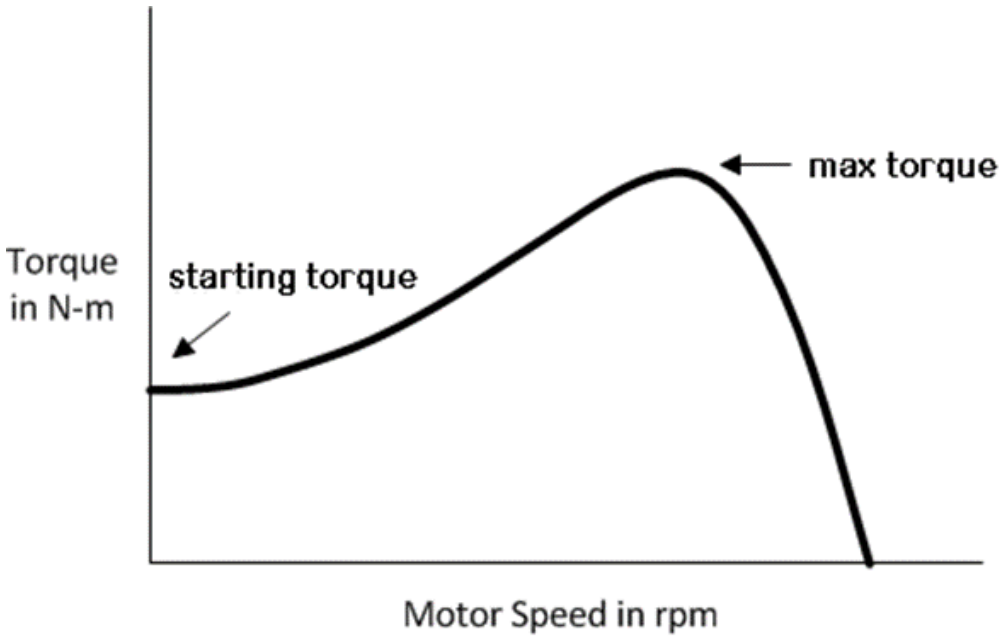


Figure 2 – Torque x RPM relationship of an electric motor.
Source: Diyoke, Okeke e Aniagwu (2016)

If the manufacturing process that produced the electric motor is under control, the profiles describing the relationship between torque and RPM should be similar and stable. In washing machines, the torque x RPM profile has a vital role in product performance and energy efficiency, impacting water and energy consumption per wash cycle (RAKHMAWATI et al., 2020). An eletric motor with some mechanical defect or other inadequacy will produce a different profile than the one that meets the design specifications and, as a consequence, may not meet the energy efficiency, water consumption, and electrical safety requirements set by the technical standards as defined by the National Institute of Metrology (INMETRO).

Although its importance is acknowledged, frequently, the functional relationship between torque and RPM is unknown or at least disregarded when determining the capability of the electric motor production processes, being used to define the product quality only the values of the starting torque and the maximum torque.

For this thesis, it is possible to define three aspects in which the research problem could be addressed: 1) a technological problem related to the imprecise procedure to determine the functional relationship between Torque and RPM for electric motors; 2) an academic problem related to the absence of a method or structure to manage this problem based on a statistical control chart; 3) a supportive method for decision making in economic terms, especially when the process has high capability indices. The research question this thesis intends to answer is: **Is a Shewhart-type functional control chart based on capability índices $\hat{C}_p$ and $\hat{C}_{pk}$ an acceptable control way to profile monitoring?**

1.2.1 General objectives

This thesis has as its theme the study of monitoring and analysis of process when a polynomial profile characterizes the quality characteristic. Its main objective is to propose, develop and evaluate a Shewhart-type functional control chart based on capability índices $\hat{C}_p$ and $\hat{C}_{pk}$ applied to profiles in real situations found in the industry where the process presents a high capability.

1.2.2 Specific objectives

- To do a review and analysis of the state-of-the-art research on profile monitoring;
- Develop a statistical model for control limits of the Shewhart-type functional charts based on the capability índices $\hat{C}_p$ and $\hat{C}_{pk}$;
- Evaluate the performance of the model developed using simulated data;
- Propose a profile monitoring framework based on process capability.

1.3 Motivation

The capability to detect tiny shifts in a profile is critical in high-precision production processes, and the rapid progress of precision manufacturing also indicates the importance of identifying minimal shift types of a process/product profile curve. Significant research and development effort has been made to search for new methods to monitor small shifts in the profile and minimal change in the process caused by special causes.

Nevertheless, when the process, despite having special causes, has a high capability and produces items whose quality is acceptable, and when the special causes are known, and their respective effects do not affect the performance and customer satisfaction, some flexibility or desensitization in the use of control chart may be required.

Problems related to have control chart signals associated with practical significance, not just statistical significance. have technical and economic relevance, and the development of new techniques for the analysis and determination of process performance, whose quality characteristic is a polynomial profile, will contribute to the advancement of the state-of-the-art in the analysis and optimization of products and industrial processes. This contribution is relevant, given the low availability of research on this topic found in the literature and its importance for decision-making procedures about the process.

The absence of adequate control limits that take into account the capability of the process results in wrong management decisions and financial losses, negatively impacting the performance and reliability of products and processes. There are few works on economic and practical significance applied to a functional relationship, and it is necessary to fill this gap. The completion of this work is expected to obtain significant results for production engineering, presenting a method that offers a rationale for practitioners to use engineering and subject matter knowledge so that signals have both practical importance and statistical significance.

1.4 Methodology

According to Bertrand e Fransoo (2002), this thesis is classified as a quantitative normative axiomatic research strongly based on real situations since it aims at the development of norms, policies, strategies, and actions that improve the current situation or available results in the literature. This research is based on models that prescribe a decision to the problem, seeking to find a solution to a new problem or to compare the performance of strategies that deal with the same problem (MORABITO; PUREZA, 2012).

The thesis will use the Modeling and Simulation method to operationalize the research, which according to Morabito e Pureza (2012), addresses the development, analysis, and testing of mathematical and symbolic descriptions of causal relationships between control variables and performance variables.

To better illustrate the research steps, Figure 3 shows the method used from the characterization of the problem until the evaluation of the results.

1.5 Thesis Structure

The rest of this text is organized as follows:

Chapter 2: This chapter presents a literature review regarding the profiles as the quality characteristic to be monitored and the different approaches in its monitoring,

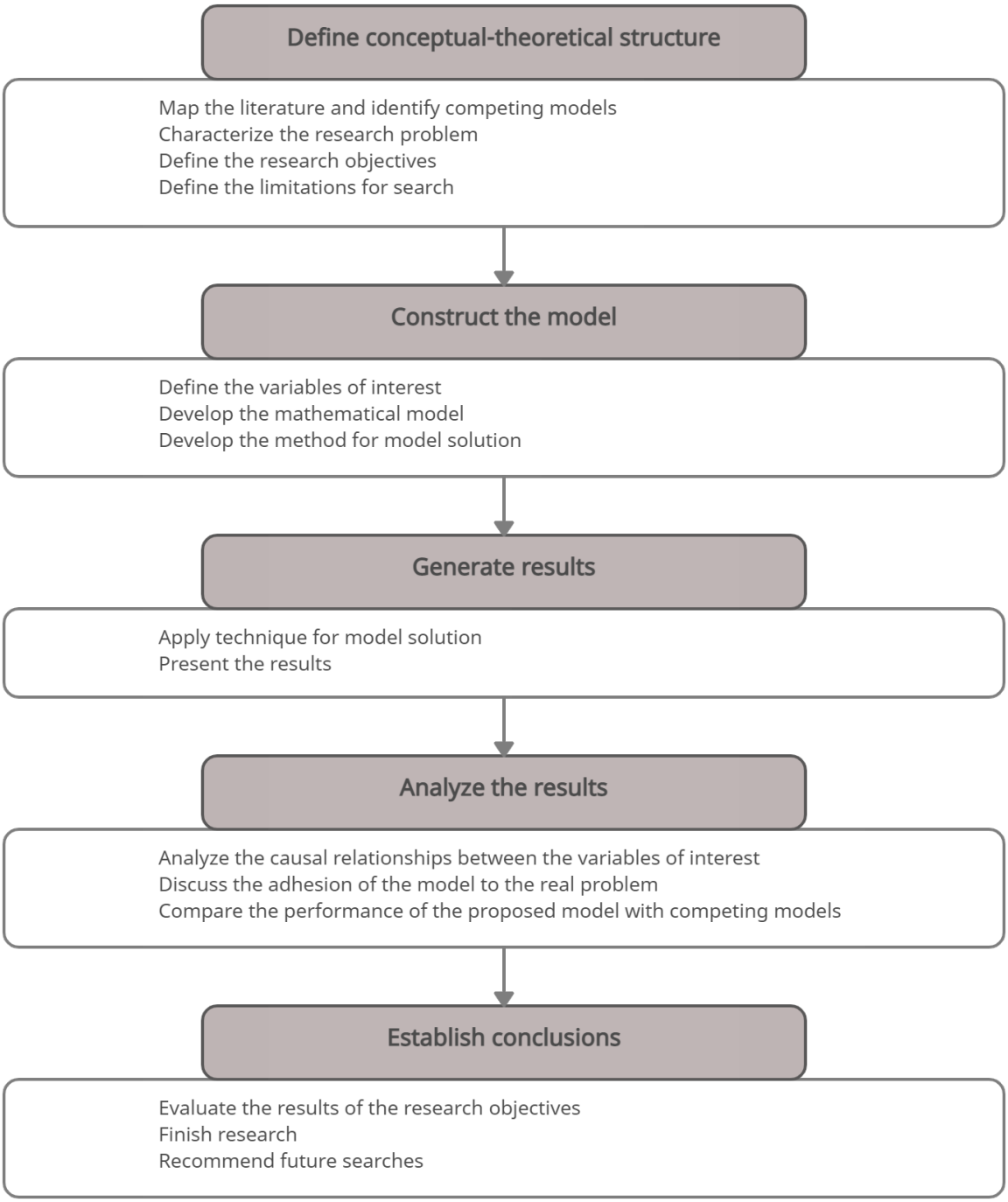


Figure 3 – Research steps.
Source: Adapted from Leoni (2015).

highlighting its increasing importance in Industry 4.0. It presents the location control chart as an alternative approach to profile monitoring, highlighting its applicability and simplicity. This chapter still discussed the adoption of relaxation measures in the monitoring of the process in those situations where it has high capability.

Chapter 3: This chapter initially presents a review of traditional process capabilities indices, extending their definition for profile analysis. Introduces the concept of traditional

univariate control charts based on $\hat{C}_p$ and $\hat{C}_{pk}$ indices, available in the literature, developing and modeling it to functional data.

Chapter 4: This chapter presents the approach adopted to treat data received from the company, arranging and transforming it consistently to allow its analysis and the development of the models to be simulated.

Chapter 5: This chapter develops and simulates location control chart models by adopting two approaches: Shewhart's traditional approach and the proposed functional control chart based on $\hat{C}_p$ and $\hat{C}_{pk}$ capability indices.

Chapter 6: This chapter presents the conclusions on the results achieved. Recognizing the limitations of the current study, it summarises the main research findings and explains the significance of the findings or contribution of the study. It ends by making recommendations for further research work.

2 Literature Review

Product and service quality plays an essential role in the success and performance of organizations, whether they are manufacturing or service. An organization that meets its customers' quality requirements at competitive prices can outperform its competitors, maintaining and conquering new markets. Many traditional statistical process control (SPC) concepts and techniques grew in response to the manufacturing environments prevalent several decades ago (GRIGG, 2020; MARTIN; ELG; GREMYR, 2020).

Different tools and methods of quality improvement and variability reduction was developed to be used in practice to improve process performance. The richness and variety of these applications have had a significant influence on the development of statistics as a discipline (i.e., change point detection, dating back to the pioneering work of Shewhart in the 1920s, developments in automatic process control, design of experiments, sequential analysis, reliability, among others) (PERDIKIS; PSARAKIS, 2019; HOCKMAN; JENSEN, 2016; HOERL; SNEE, 2010; NAIR; HANSEN; SHI, 2000).

However, as Soriano, Oprime e Lizarelli (2017) and Lizarelli et al. (2016) observe, advances made in the academic research world of SPC methods have not yet been fully incorporated in the manufacturing best practices. According to the authors, this may be a result of: i) cultural paradigms; ii) the types of groups used to promote and improve the field of statistical quality control, such as task force and semi-autonomous groups; iii) governance within groups (centralized or decentralized, use of specific methods or routines and monitoring of activities) and iv) understanding and applying new techniques and tools to problem identification and its solution.

Despite the gap between academic research and the adoption of best practices that exists in many companies pointed out by Soriano, Oprime e Lizarelli (2017), in many companies however, SPC is an effective element of the process control system, which comprises a set of monitoring techniques that have been successfully utilized for process monitoring and variation reduction in manufacturing applications.

Authors such as Rahali et al. (2021), He e Wang (2018), He e Wang (2017), Woodall e Montgomery (2014), have used the term statistical process monitoring (SPM) to refer to such developments and strategies because, in their opinion, it better reflects the application of the methods. The authors also emphasize that additional methods are also needed to visualize results when there are many data streams or a large amount of spatiotemporal data.

With the advance of technology and data collection methods, the range of ap-

plications of statistics has significantly expanded, and the outcome of the process can be as diverse as healthcare monitoring, image monitoring, risk analysis or spatiotemporal surveillance (HUBIG; LACK; MANSMANN, 2020; BUI; APLEY, 2018; KOOSHA; NOOROSSANA; MEGAHED, 2017; WOODALL et al., 2017).

SPM consists of a robust set of tools for understanding the variation of a process and characterizing its stability over time that help professionals improve the quality of products and services, achieving process stability, and reducing process variability. Several sources of variation can affect process performance. In particular, a process is said to be in control (IC) when all fluctuations are due to common causes of variability that cannot be removed without changing the process in some fundamental way, such as changing the equipment used or the raw material. Conversely, the process is said to be out of control (OC) when it operates under special causes of variation that must be identified and removed as soon as possible (MONTGOMERY, 2019; ALLEN, 2019; REIS; GINS; RATO, 2019; CAPIZZI, 2015; QIU, 2014).

As in traditional Statistical Process Control, Statistical Process Monitoring involves two phases: Phase I and Phase II. Phase I includes methods for understanding process behavior based on a historical baseline set of data. In-control parameter values for appropriate models are estimated in the retrospective Phase I and used to evaluate the stability of the process and design methods for on-going prospective monitoring in Phase II. In Phase II, we decide about the stability of the process relative to the Phase I baseline as each sample is collected over time. The objective of Phase II is to quickly detect any change in the process from its in-control state (MONTGOMERY, 2019; WOODALL; FALTIN, 2019).

However, the traditional Shewhart control chart does not consider whether a process has a high capacity or not. It recommends an interruption of the process when there is an indication of special causes to keep the production process at a stable level of variation. There are cases where it is not financially appropriate to intervene, even in the presence of special causes. Woodall e Faltin (2019), Oprime et al. (2019) and Oprime e Mendes (2017) ensure that small changes in the statistical parameter, for example, in the average, over time may have little or no practical importance, and the cost of false alarms discourages the use of control charts. When the benefits of interrupting the process in the presence of special causes are less than the costs, an excess of control can be considered (MONTGOMERY, 2019; AIAG/ASQC, 1991; WOODALL, 1985)

2.1 Profile Monitoring

In standard SPC applications, we traditionally sought to monitor the performance of a process or product by considering measurements on a single quality characteristic or a

quality characteristic vector at a given time or space. Nowadays, advanced manufacturing industries and more advanced technologies operate under much more complex and diverse conditions.

Nair, Hansen e Shi (2000) argue that these changes have important implications for the research directions in industrial statistics, not only in terms of identifying new problems and developing new methods but also in re-evaluating the paradigms that inspired previous approaches. Currently, large amounts of process and product quality data are collected regularly, made possible by advances in information technology and data acquisition. Many of these data have unique structures: images, functional data, marked point processes, and high-dimensional time series are frequent, among others (HUANG et al., 2020; REIS; GINS, 2017).

Advances in technology have enabled engineers and professionals to collect a large number of processes or product measurements from reconstructing the entire functional relationship to the process or product performance. This functional relationship is often called a profile, signature, or waveform. For each profile, it is assumed that p response variable values are measured along with the corresponding values of one or more explanatory or independent variables (NOOROSSANA; SAGHAEI; AMIRI, 2011).

A bibliometric analysis conducted shows that there has been an increase in publications on this subject since 2006, with 51.2% of articles published in the last five years (Figure 4).

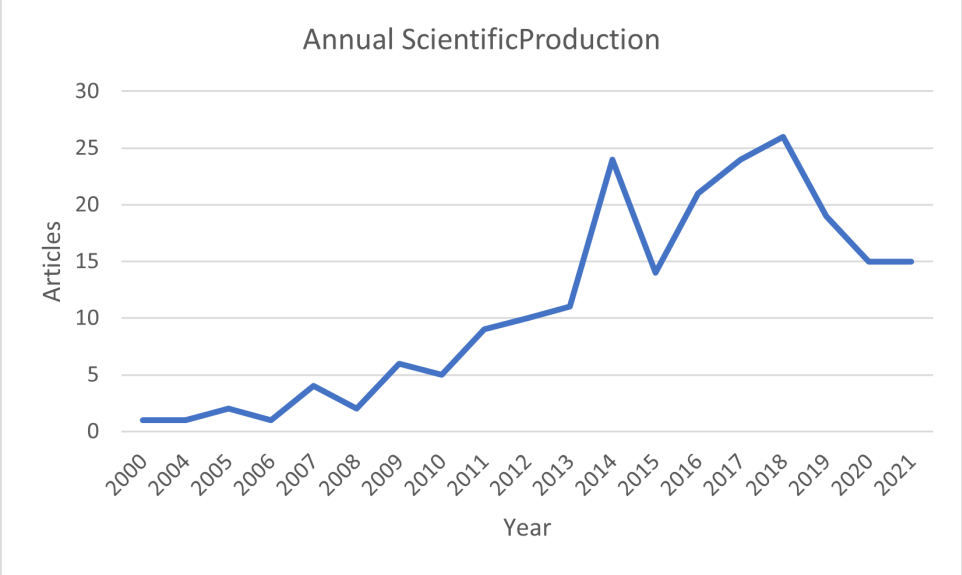


Figure 4 – Annual scientific production related to profile monitoring from 2004 to 2021. Source: Web of Science, 2021.

Figure 5 presents the relationship between the keywords of the articles analyzed on the topics: linear profiles, control charts, phase I analysis, statistical process control,

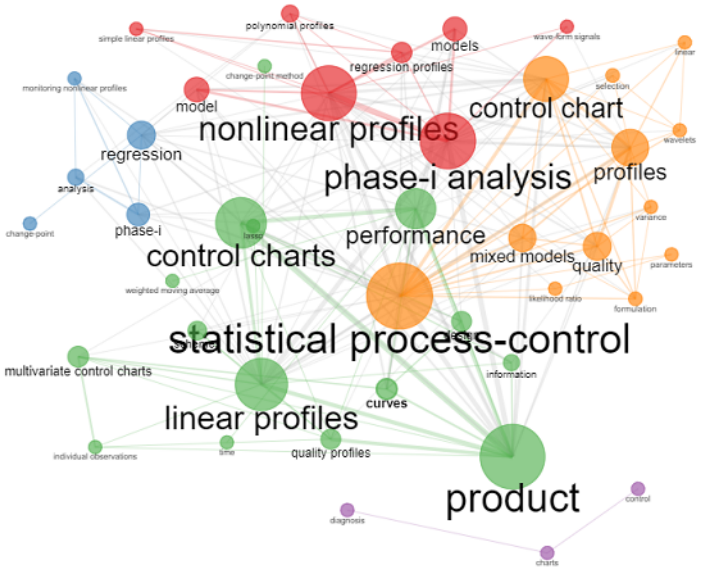


Figure 5 – Keyword co-occurrences diagram.
Source: Web of Science, 2021

polynomial profiles, and performance. The line thickness indicates how often two keywords were used together.

While process capability plays an essential role in evaluating the performance of a manufacturing process, Maleki, Amiri e Castagliola (2018), Woodall e Montgomery (2014) and Woodall (2007) pointed out that little attention was paid to evaluating process capability whose quality characteristic is a profile. More recently, other authors as Wu (2016), Keshteli et al. (2014a) and Hosseinifard e Abbasi (2012b) agreeing with him, have highlighted the inexistence, or at least, a small amount of articles published dealing with this important issue. It is possible to observe from Figure 5 that among the published works related to profile monitoring, few studies have contemplated process capability.

Starting with Kang e Albin (2000), many authors have discussed practical applications of profiles. In their paper, Kang e Albin (2000) use a linear profile to represent the relationship between the pressure and the amount of flow in a calibration problem in a semiconductor industry and a dissolution profile curve in the pharmaceutical industry. They propose two approaches. In the first, the function is described by its parameters, and these are monitored using the multivariate T^2 control chart. In the second, the function is viewed as a collection of points, and the differences or residuals between the reference points and the sample points are monitored using EWMA and R charts for the average residual.

Maleki, Amiri e Castagliola (2018), Woodall e Montgomery (2014), Noorossana, Saghaei e Amiri (2011), and Woodall et al. (2004) show an overview of some of the general issues involved in using control charts to monitor process and product quality profiles, and

relate the application of control charts to do functional data analysis reviewing applications involving linear profiles, nonlinear profiles, and the use of splines and wavelets. These authors pointing out that process and product profiles monitoring expands the area of statistical process monitoring to include a much broader array of statistical methods, and includes a much broader class of applications.

The current importance placed on profile analysis cannot be overstated, as pointed out by Fallahdizcheh e Wang (2021). Woodall e Montgomery (2014), Woodall (2007) and Woodall et al. (2004) who described it as "the most promising area of process control research". More recently, Maleki, Amiri e Castagliola (2018) reviewed the literature from the period 2008 up to 2018 presenting a conceptual classification scheme and classifying the papers in this area. These authors concluded that an increasing interest on diagnosis approaches has appeared by quality engineering researchers in recent years and the number of papers in this subarea has increased from 6 in 2008–2013 to 16 papers in 2014–2018.

According to Maleki, Amiri e Castagliola (2018), Woodall e Montgomery (2014), Noorossana, Saghaei e Amiri (2011), Woodall (2007) and Woodall et al. (2004), profile data are common in many applications. For example, pressure in a chamber is a linear function of flow setpoints in the calibration of a mass flow controller (KANG; ALBIN, 2000). The relationship between the torque of a car engine and its speed in revolutions per minute (RPM) is an important quality feature in the automotive industry (AMIRI; JENSEN; KAZEMZADEH, 2010). Tonnage signals over time in stamping processes are other forms of profile data (JIN; SHI, 2001).

Discussing the use of profile monitoring techniques for a data-rich environment, Schütze, Helwig e Schneider (2018), Reis e Gins (2017), Khan et al. (2017), Wang e Tsung (2005) argue that on-line sensors combined with a huge sample size have become a usual or ordinary thing in the modern manufacturing industry, due to the increasing complexity of processes and products, and the availability of advanced sensing technology. They exemplify their argument, citing as an example, a car factory in which an optical coordinate measuring machine (OCMM) measures from 100 to 150 points in each main assembly with a sample rate of 100%, for which efficient use of these data constitute a challenge. Zou, Zhang e Wang (2006) introduce a Likelihood Ratio Test (LRT) and EWMA control charts based on LRTs, designed to detect shifts in the intercept, slope, and standard deviation of linear profiles, when process parameters are unknown but some historical samples are available. They also provide two useful diagnostic aids to improve the practicality of the EWMA chart: one provides valuable information to process engineers concerning the time of the change, and the other decides which parameter has changed.

Centofanti et al. (2020) propose a new general framework for monitoring a functional quality characteristic when functional covariates are available, referred to as functional

regression control chart (FRCC). In particular, the quality characteristic is adjusted for the effects of the covariates employing multivariate functional linear regression model and then monitored by using jointly the Hotelling's T_2 and the squared prediction error (SPE) control charts built on its functional principal component decomposition.

Outside the industrial environment, Capezza et al. (2021) presents an application on modern ships, where the rapid development in data acquisition technologies produces data-rich environments in which measurements of the variable are continuously streamed and stored during navigation; therefore, they can be modeled as functional data or profiles. Using a real case study of a Ro-Pax vessel operating in the Mediterranean Sea, the authors demonstrate the applicability of the FRCC to the emissions of CO_2 (quality characteristic of interest) and the profiles of variables that have an impact on them (i.e., the covariates) explored in light of the new global and European level regulations on the monitoring, reporting, and verification of harmful emissions. In this paper, they show an application of the FRCC to answer, at the end of each ship voyage, the question: given the value of the covariates, is the observed CO_2 emission profile as expected?

Large amounts of time and money are invested in recalibrating the production process measuring system, even sometimes when the recalibration is not required. To optimize the calibration frequency and maintain a certain level of accuracy and precision, Gupta, Montgomery e Woodall (2006) propose two different schemes for simultaneous monitoring of the intercept, slope, and error standard deviation of a linear profile: one based on the classical calibration method monitoring the deviations from the regression line (referred to as NIST method), and the other based on individually monitoring the parameters of the linear profile (referred to as KMW method). The result of their study shows that the NIST method has a poor performance when compared with the KMW method.

Woodall (2007) highlighted the following important issues when monitoring profiles: i) the importance of carefully distinguish between Phase I and Phase II applications; ii) the decision regarding whether or not to include some between profile variation in common cause variation; iii) the use of methods capable of detecting any type of shift in the shape of the profile, and iv) the use of the simplest adequate profile model. In his paper, he arguments that parametric or nonparametric methods can be used in both Phases I and II. In many cases of profile monitoring, it is efficient to summarize the shape under control of the profile with a parametric model and monitor changes in the parameters of this model.The control charts are based on the estimated parameters of the model from successive profile data observed over time. For non-parametric methods, it is possible to alternatively monitor metrics that reflect the discrepancies between the observed profiles and a baseline profile established using historical data from Phase I.

Another point raised by Woodall (2007) is that, in many applications, the simple linear regression model is not sufficient to represent the shape of a profile, so more complicated methods are needed. To exemplify alternative and more appropriate approaches to situations like this, he lists Multiple and Polynomial Regression Models (KAZEMZADEH; NOOROSSANA; AMIRI, 2008; MAHMOUD, 2008), Nonlinear Regression Models (DING; ZENG; ZHOU, 2006; WILLIAMS; WOODALL; BIRCH, 2007), Mixed Models (JENSEN; BIRCH; WOODALL, 2008; MOSESOVA et al., 2006) and the use of Wavelets (REIS; SARAIVA, 2006; JEONG; LU; WANG, 2006) as possible alternatives to linear models. According to (CHICKEN; JR; SIMPSON, 2009), wavelet methods are usually recommended when the shape of the profiles is too complicated for linear and nonlinear models to work well.

Zou, Tsung e Wang (2007) propose a method to monitor a general linear profile that includes both a polynomial regression and a multiple linear regression model, using to illustrate the implementation of the proposed method a deep reactive ion etching (DRIE), taken from semiconductor manufacturing, which has a profile that fits a quadratic polynomial regression model. According to Qiu, Zou e Wang (2010), in the literature, most existing profile monitoring control charts require a fundamental assumption that the observations in the profile are independent of each other, and warns that this is generally invalid in applications. As they argue, in practice, the data within the profile is usually spatially or serially correlated. Thus, they propose a new control graph, which incorporates the local linear smoothing of the kernel in the exponentially weighted moving average (EWMA) control scheme.

An example from the automotive industry is presented by Amiri, Jensen e Kazemzadeh (2010), where one of the most important quality characteristics of an automotive engine is the relationship between the torque produced by an engine and the engine speed in revolutions per minute (RPM) (Figure 6). They model their problem as a second-order polynomial profile with autocorrelations structures within each profile, using a linear mixed model (LMM) approach to estimate polynomial regression parameters. To check the stability of the process and the existence or not of outlying profiles, they use in Phase I a T^2-based control procedure. Once a stable process is obtained, they estimate the mean vector and variance-covariance matrix of polynomial regression parameters to form a T^2 based control chart to monitor polynomial profiles in Phase II.

2.1.1 Simple Linear Profile

Acording Noorossana, Saghaei e Amiri (2011), in a linear profile data set with a single explanatory variable X and a response variable Y, data are m samples in form $\{(X_{i1}, Y_{i1}),\ i = 1, 2, \cdots, n_1\}, \{(X_{i2}, Y_{i2}),\ i = 1, 2, \cdots, n_2\}, \{(X_{im}, Y_{im}),\ i = 1, 2, \cdots, n_m\}$. In most research related to linear profile monitoring, it is assumed that the values for the

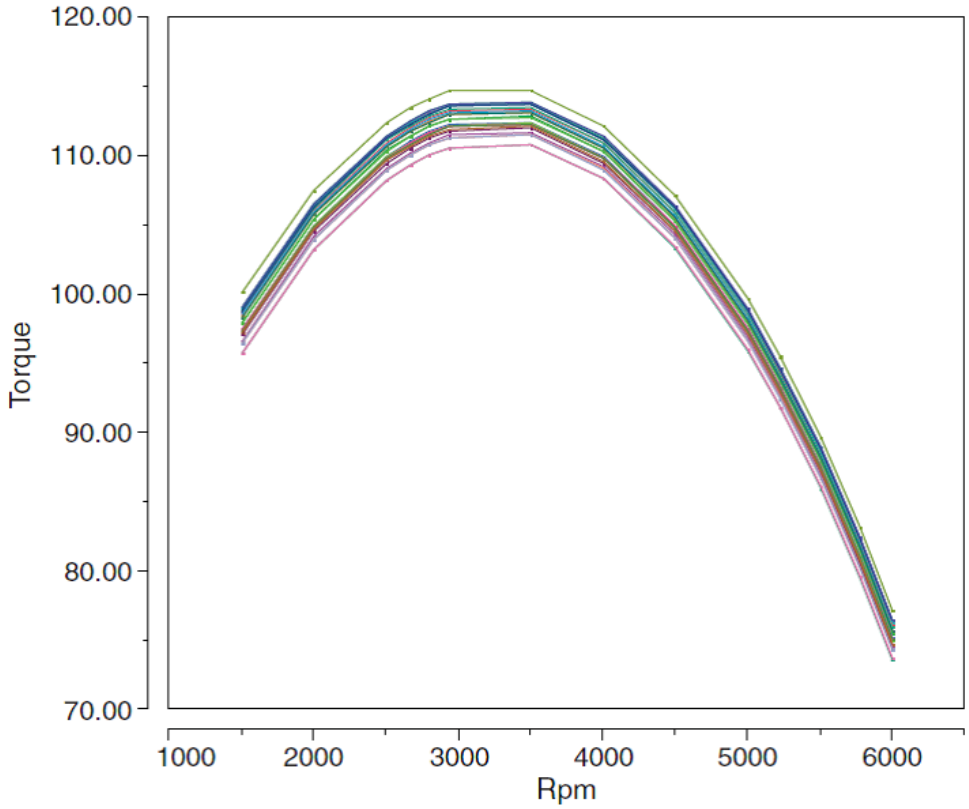


Figure 6 – Functional relationship between Torque and RPM for a car engine.
Source: Amiri, Jensen e Kazemzadeh (2010)

explanatory variable X are fixed, using the same set of values for each sample. The simple linear profile is a linear relationship between a response variable and one or more independent variables that must be monitored over time. In its simplest form, this relationship could be defined as follows:

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + \varepsilon_{ij},\ i = 1, 2, ..., n_j \ e \ j = 1, 2, ..., m \tag{2.1}$$

where Y_{ij} is the response variable corresponding to the value of i^{th} value of explanatory variable X to j^{th} sampled profile.

The parameters β_{0j} and β_{1j} correspond to the intercept and slope, respectively. The term ϵ corresponds to the error and follows a normal distribution with mean zero and variance equal to one. Kang e Albin (2000) used linear profiling to solve the calibration problem. Kim, Mahmoud e Woodall (2003) used an exponentially weighted moving average (EWMA) control diagram scheme to monitor simple linear profiles. If the process is under statistical control, that is, under the action of random causes only, the profile does not change and remains stable over time. However, if special causes are brought to bear on the process, this will lead to an increase in variability or a change in the mean profile, signaling that the process is out of control and investigative procedures and intervention

or corrective measures are required.

2.1.2 Multiple and Polynomial Profiles

More complicated models, such as multiple linear regression or polynomial regression rather than simple linear regression, should sometimes be used to model some real cases. For example, Amiri, Jensen e Kazemzadeh (2010) demonstrates that the functional relationship between the torque produced by a car engine and its speed in revolutions per minute (RPM), can be modeled by second-order polynomial regression. Parker e Finley (2007) discuss an example of a multiple linear regression calibration model. Zou, Tsung e Wang (2007) investigate a case of semiconductor fabrication that is modeled by a polynomial regression profile.

Assuming there are m samples in the form $(X_{1ij}, X_{2ij}, X_{pij})$, $i = 1, 2, \cdots, n_j$, $j = 1, 2, \cdots, m$ and $n_j > p$ with a response variable Y and explanatory variables X_1, X_2, $\cdots, X_p$. The model that relates the response variable and the explanatory variables for each sample is given by the expression:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \cdots + \beta_{pj}X_{pij} + \varepsilon_{ij} \tag{2.2}$$

where $i = 1, 2, \cdots, n_j$ e $j = 1, 2, \cdots, m$ or in its vector form:

$$\boldsymbol{Y}_j = \boldsymbol{X}_j\boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j \quad j = 1, 2, \ldots, m \tag{2.3}$$

where $\mathbf{Y}_j = \left(Y_{1j}, Y_{2j}, \cdots, Y_{n_jj}\right)^T$, $\boldsymbol{\varepsilon}_j = \left(\varepsilon_{1j}, \varepsilon_{2j}, \cdots, \varepsilon_{n_jj}\right)^T$, $\boldsymbol{\beta}_j = (\beta_{1j}, \beta_{2j}, \cdots, \beta_{pj})^T$ and matriz $\mathbf{X}_j$ is:

$$\boldsymbol{X}_j = \begin{bmatrix} 1 & x_{11j} & x_{21j} & \cdots & x_{p1j} \\ 1 & x_{12j} & x_{22j} & \cdots & x_{p2j} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n_jj} & x_{2n_jj} & \cdots & x_{pn_jj} \end{bmatrix}$$

Its assumed that the x values are known constants in each sample and ϵ_{ij} are independent, identically normal distribution with mean zero and variance σ_j^2. If the process is in statistical control, them the regression parameters are constant in each sample (NOOROSSANA; SAGHAEI; AMIRI, 2011).

When there is a k^{th}-order polynomial relationship between a response variable and one exploratory variable, Noorossana, Saghaei e Amiri (2011) propose an approach based on the generalized likelihood ratio test. In its method, it is assumed that there are m

samples of size n_j in the form of $[(x_{ij}, y_{ij}), i = 1, 2, \ldots, n_j, j = 1, 2, \ldots, m]$. Considering a step shift in one or more regression parameters after sample m_1, the following model is assumed:

$$\begin{aligned} Y_{ij} &= \beta_{01} + \beta_{11}x_{ij} + \beta_{21}x_{ij}^2 + \cdots + \beta_{k1}x_{ij}^k + \epsilon_{ij1} \\ i &= 1, 2, \cdots, n_j \quad j = 1, 2, \cdots, m_1 \\ Y_{ij} &= \beta_{02} + \beta_{12}x_{ij} + \beta_{22}x_{ij}^2 + \cdots + \beta_{k2}x_{ij}^k + \epsilon_{ij2} \\ i &= 1, 2, \cdots, n_j \quad j = 1, 2, \cdots, m \end{aligned} \tag{2.4}$$

where $\epsilon_{ij1} \sim N(0, \sigma_i^2)$ and $\epsilon_{ij2} \sim N(0, \sigma_2^2)$, considering the following null and alternative hypotheses,

$$\begin{aligned} H_0 : \beta_{01} = \beta_{02} = \beta_0, \beta_{11} = \beta_{12} = \beta_1, \cdots, \beta_{k1} = \beta_{k2} = \beta_k \\ \sigma_1^2 = \sigma_2^2, \cdots, = \sigma_k^2 \end{aligned}$$

$H_1 : H_0$ is not true

2.2 Approaches for profile monitoring

Significant analysis and discussion on the subject were presented by Noorossana, Saghaei e Amiri (2011). Colosimo e Pacella (2010) show that most of the approaches for profile monitoring proposed in the literature share a typical structure, which consists of: i) identifying a parametric model of the functional data; ii) estimating the model parameters, and iii) designing a multivariate control chart on the estimated parameters and a univariate control chart on the residual variance. The proposed approaches can then be classified regarding the type of application faced (i.e., calibration study, process signal, or geometric specification monitoring) or the modeling approach considered (mainly linear regression or approaches for multivariate data reduction such as principal/independent component analysis).

2.2.1 The location control chart

Proposed initially by Boeing (1998), the method consists of applying a traditional Shewhart control chart separately to each data point observed at a given location on the part, and where the control limits used for each location depend only on responses at that location (Figure 7).

According to Woodall et al. (2004), the location control chart can be considered the simplest approach for monitoring functional data, as it consists of applying a traditional control chart to the data observed at each given location, intending to combine simplicity

with the need to keep all the data information observed in each location where the profile needs to be evaluated. Colosimo e Pacella (2010) observed that the location control chart exhibits performance comparable to those achieved by the PCA-based approach. In a few cases, especially for production scenarios with correlated errors, the location control chart surprisingly outperforms both the regression-based and the PCA-based approaches. confirming that the simple location control chart can be considered a valuable alternative to parametric methods for profile monitoring.

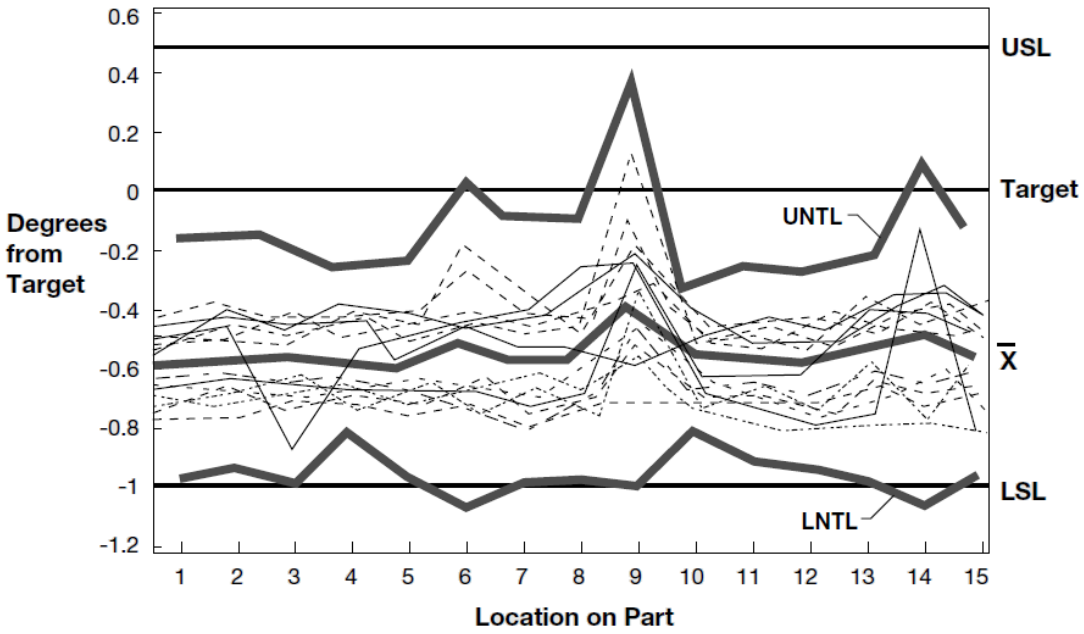
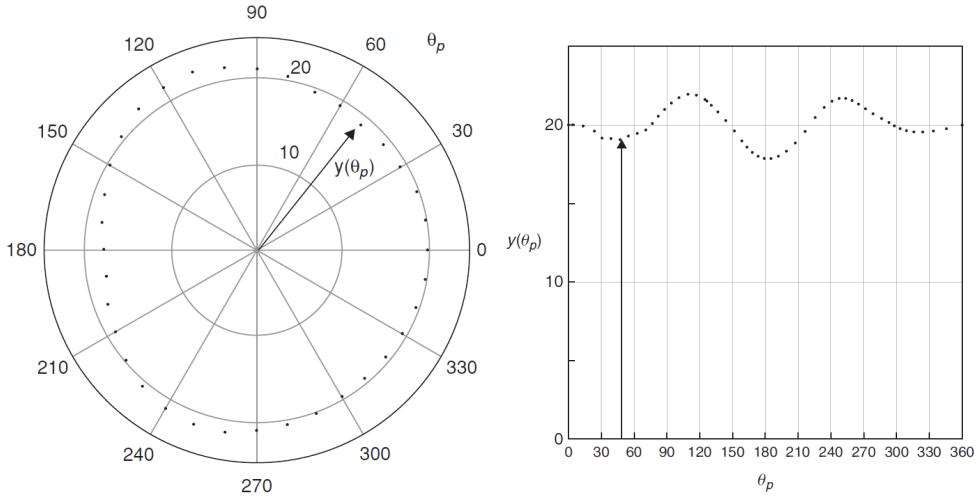


Figure 7 – Location control chart for flangle-angle measurements on 13 parts
Source: Boeing (1998)

Given a specific profile, an alarm is issued when at least one point, in the set of p observations, exceeds either the upper or the lower control limit at the specific location. It is worth noting that the control limits used at each location depend only on the responses at that specific position; thus, the main disadvantage of this method is that the multivariate structure of the data is ignored. An illustration of a location chart from Boeing (1998) is shown in Figure 7, where the response is the upper flange angle measured at $p = 15$ locations for $n = 13$ parts. The LSL and USL are the lower and upper specification limits, respectively. The $\widehat{LNTL}$ and $\widehat{UNTL}$ are the natural tolerance limits that are three standard deviations from the sample mean at each location (WOODALL et al., 2004).

The logic behind this approach is that if the observed profile is under control, the data observed at that specific location must remain within that range with a certain probability. On the other hand, when the process gets out of control, the control interval is likely to be violated at one or more locations. According to this method, an alarm is issued when at least one point, in the entire observed data set, exceeds the control limits (NOOROSSANA; SAGHAEI; AMIRI, 2011).



(a) Polar reference system. (b) Cartesian reference system.

Figure 8 – An example of circular profile.
Source: Colosimo, Semeraro e Pacella (2008)

Colosimo, Semeraro e Pacella (2008) present, as an example of functional data, a set of points measured in a machined profile subject to geometric specifications (Figure 8). According to authors, machined profiles and surfaces can be thought of as functional data if a subset of the coordinates, which are used to describe the position of each sampled point, can be represented as a function of one or more independent variables. In this case, each point collected in the machined profile is related to a specific spatial location. Given the simplicity of the location control chart, its use in practice can be justified in these production situations, even if attention should be paid to properly designing (Phase I) this tool.

2.2.2 Control limits of the location control charts

Assume we collect a group of n profiles, where each profile is a vector of p measurements observed at a fixed set of locations. The location control chart consists of applying a Shewhart's control chart separately to each data point observed at a given location on the part (i.e., by considering the mean and the standard deviation of the n data observed at that location and by computing the common $\pm k$ standard deviations from the sample mean). Given a profile, an alarm should be considered when at least one point, in the set of p observations, exceeds either the upper or the lower control limit (Figure 9).

Let $y_j(x)$ denote the data measured at a specific location of index x on the j^{th} profile, where $x = 1, 2, \ldots, p$ and $j = 1, 2, \ldots, n$. The control limits for the location of index x are as follows:

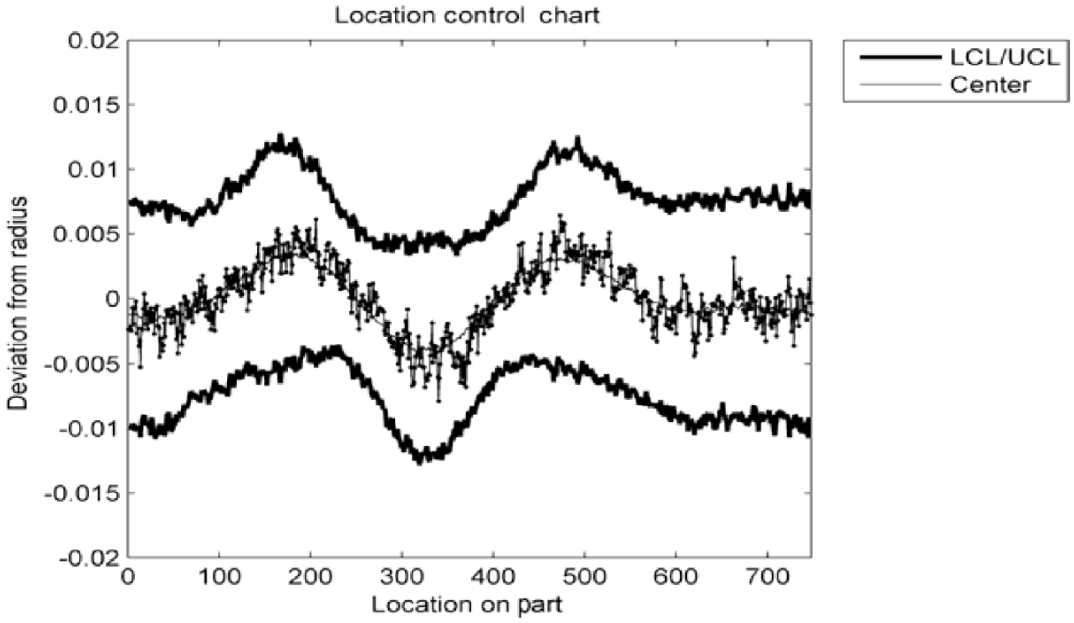


Figure 9 – One out of the profiles in 748 locations on part against the control limits of location control chart; vertical axis scale in mm.
Source: Noorossana, Saghaei e Amiri (2011)

$$\begin{aligned} UCL(x) &= \bar{y}(x) + Z_{\alpha/2}s(x) \\ CL(x) &= \bar{y}(x) \\ LCL(x) &= \bar{y}(x) - Z_{\alpha/2}s(x) \end{aligned} \tag{2.5}$$

where $\bar{y}(x) = \frac{1}{n}\sum_{j=1}^{n} y_j(x)$ and $s(x) = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}[y_j(x) - \bar{y}(x)]^2}$ are, respectively, the sample mean and the sample standard deviation of the data observed at location x, while $Z_{\alpha/2}$ represents the $(1 - \alpha/2)$ percentile of the standard normal distribution (NOOROSSANA; SAGHAEI; AMIRI, 2011).

The constant $k = Z_{\alpha/2}$ is computed, as in Shewhart's traditional approach, as a function of the required false alarm rate α. Since the fundamental ideas on which the control charts are based are similar, in structure, to a hypothesis test, we assume some statement, called the null hypothesis H_0. We also established the alternative hypothesis H_1, which is what we are trying to conclude (if the data supports it). Every time a new value is plotted on a control chart, a hypothesis is evaluated. The initial assumption is that the process is stable (under control). If, after drawing a point, we have enough evidence to reject this null hypothesis (we see a sign), we conclude that the alternative hypothesis (the process is out of control) is true. Ideally, we would correctly reject H_0 whenever the process was really out of control. However, if the process is really out of control and we have not detected it, we made a Type II error. The hypothesis statement can then be expressed as follows:

$$\begin{cases} H_0 : y_j(x) = y_{0j}(x) \; and \; \sigma_j(x) = \sigma_{0j}(x) \\ H_1 : y_j(x) \neq y_{0j}(x) \; and/or \; \sigma_j(x) \neq \sigma_{0j}(x) \end{cases} \quad for \;\; j = 1, 2, \cdots, n \; observations$$

and the probability of Error Type I (False Alarm) is calculated by

$$\alpha_j(x) = 1 - Pr\left[LCL(x) \leq \bar{y}_j(x) \leq UCL(x) \mid y_j(x) = y_{0j}(x);\; \sigma_j(x) = \sigma_{0j}(x)\right]$$

However, as the control limits are computed at each given location, P dependent control rules are simultaneously applied (i.e., simultaneous hypothesis testing for each value assumed by an exploratory variable). This implies that the percentile of the standardized normal distribution used to compute k should be corrected, because if multiple hypotheses are tested, the chance of observing a rare event increases, and therefore, the probability of incorrectly rejecting a null hypothesis (i.e., making a Type I error) increases. To this aim, the Bonferroni's rule for dependent events is used to attain an actual false alarm rate not greater than a predefined value.

2.2.3 Bonferroni's Rules

Let $\theta = (\theta_1, \theta_2, \cdots, \theta_k)'$ be a $(p \times 1)$ vector of parameters elements in an expected value vector or a variance-covariance matrix or partial regression coefficients in a general linear model. If separate two-sided confidence intervals are constructed for each of the p parameters, each with confidence coefficient $100\,(1-\alpha)\,\%$, and if A_i denotes the event that the interval for θ_i includes the actual value of θ_i, then, from the Bonferroni inequality, it follows that the probability $Pr\left[\bigcap_{i=1}^{p} A_i\right]$ that every interval includes the value of the parameter it estimates is ar least $(1 - p\alpha)$. Formally,

$$Pr\left[\bigcap_{i=1}^{p} A_i\right] \geq 1 - p\alpha \tag{2.6}$$

Thus the "family confidence coefficient" is at least $100\,(1-\alpha)\,\%$, whatever be the dependence among the statistics used in constructing the confidence intervals. If the confidence level for each separate interval is increased to $100\left(1 - \alpha p^{-1}\right)\%$, then the family confidence coefficient is at least $100\,(1-\alpha)\,\%$. The resulting confidence intervals are called *Bonferroni intervals* (ALT, 2006).

In obtaining Bonferroni intervals, it is not necessary that all the separate confidence coefficiens $[100\,(1-\alpha_i)\,\%,\; i = 1, 2, \cdots, p]$ be equal, only that $\sum_{i=1}^{p} \alpha_i = \alpha$. Thus if a few of parameter warrant greater interest than others, then the confidence coefficients for

thes parameters could be larger. Regardless of the allocation of the α_i's, the conservative Bonferroni intervals provide a viable alternative for achieving a family confidence coefficient of at least $(1 - \alpha)$.

Since that P dependent control rules are applied simultaneously, the Bonferroni rule for dependent events must be used to achieve an actual false alarm rate not greater than a predefined value. Therefore, let α denote the upper limit of the first type of probability error (false alarm probability); the value $\alpha = \alpha'/P$ is used to calculate the control limits of Equation 2.5. For example, assuming the standard value $\alpha' = 0.01$ as upper bound for the type I error probability and 748 locations (Figure 9), the value of $\alpha = 1.3369 \times 10^{-5}$ (i.e., $Z_{\alpha/2} = 4.354$) is used to calculate the control limits of Equation 2.5. (COLOSIMO; PACELLA, 2010).

2.3 Modified and Acceptance Control Charts

According to Oprime et al. (2019), Oprime e Mendes (2017), Holmes e Mergen (2000), there are processes that, due to their nature, present inevitable changes in the average value of the quality of interest characteristic, but are still capable of meeting the specifications established in the project. This situation occurs when the process standard deviation is too small concerning the tolerance width (i.e., the difference between the lower and upper specification limits). In usual terms of statistical process control, this process, while not necessarily under control, is capable of producing acceptable products that must be protected against rejection. Montgomery (2019) argues that when a high level of process capability is reached, it is sometimes useful to relax the vigilance level provided by the standard control charts.

In these situations, an alternative to standard control charts could employ modified control limits or acceptance control charts (ACC). The goal of ACC is different from that of the Shewhart control charts. The Shewhart control charts mainly aim to verify whether the process mean is stable over time, whereas ACC is concerned with maintaining the process mean at a specific range so that the nonconforming fraction does not exceed a desired value (ZHOU; GOVINDARAJU; JONES, 2019; OPRIME et al., 2019; OPRIME; MENDES, 2017; MOHAMMADIAN; AMIRI, 2013).

2.3.1 Modified Control Limits for $\bar{X}$ chart

The basic concept behind the first approach, the $\bar{X}$ chart with modified limits, is to allow the process mean to shift such that the fraction of nonconforming parts produced does not exceed a specified value δ. Hill (1956) and Freund (1957) present a general discussion of this technique. Montgomery (2019) also gives an extensive reference on this technique from the statistical theory aspect.

Assuming that the quality characteristic is normally distributed with an inherent within-sample variance σ^2. For a process with two-sided specification limits, in order to produce parts with a non-conforming fraction less than δ, the process mean μ is only allowed to shift within μ_L and μ_U, as shown in Figure 10.

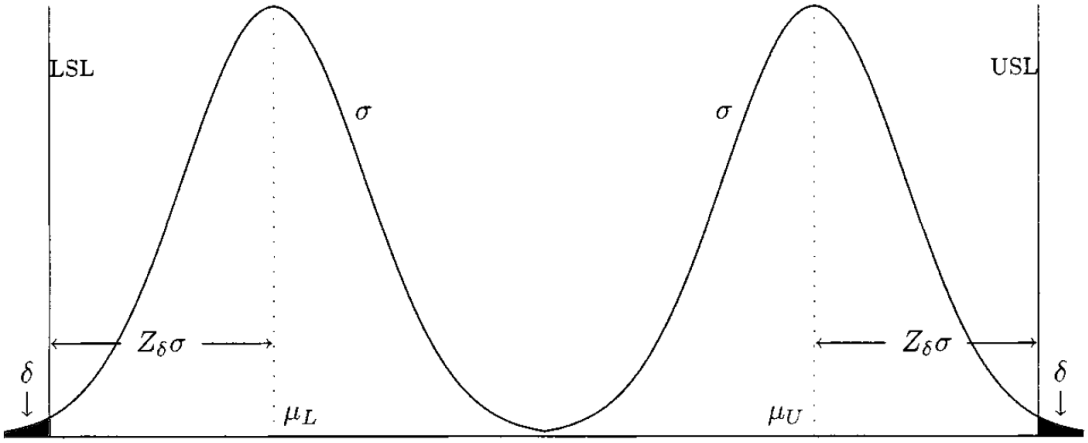


Figure 10 – Distribution of normal quality characteristic.
Source: Chang e Gan (1999)

Consequently, μ_L and μ_U are expressed as

$$\begin{aligned} \mu_L &= LSL + Z_\delta\sigma \\ \mu_U &= USL - Z_\delta\sigma \end{aligned} \tag{2.7}$$

where Z_δ is the upper $100(1-\delta)\%$ point of the standard normal distribution.

This chart, is based on a specified sample size n, a process fraction nonconforming δ, and Type I error probability α. In such chart, we interpret δ as a process fraction nonconforming that we will accept with probability 1-α being used in situations where the natural variability or "spread" of the process is considerably smaller than the spread in the specification limits, that is, C_p and C_{pk} is much greater than 1 (MONTGOMERY, 2019).

As shown if Figure 11, the lower and upper control limits can be derived with the probability α, of Type I Error as

$$\begin{aligned} LCL &= \mu_L - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \\ &= LSL + \left(Z_\delta - \frac{Z_{\alpha/2}}{\sqrt{n}}\right)\sigma \end{aligned} \tag{2.8}$$

$$\begin{aligned} UCL &= \mu_U + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \\ &= USL - \left(Z_\delta - \frac{Z_{\alpha/2}}{\sqrt{n}}\right)\sigma \end{aligned} \tag{2.9}$$

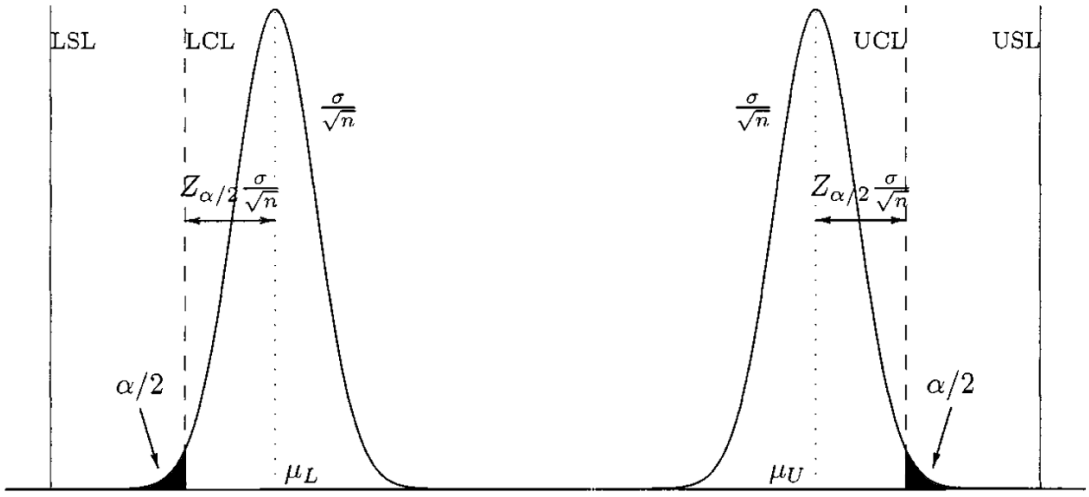


Figure 11 – Distribution of sample average of normal quality characteristic.
Source: Chang e Gan (1999)

2.3.2 Acceptance Control Charts

The second approach, the acceptance control chart (ACC), is to use a chart to monitor the fraction of nonconforming units or the fraction of units that exceed specifications. Developed originally by Freund (1957), ACC takes into account both the risk of rejecting a process operating at a satisfactory level (Type I error or α-risk) and the risk of accepting a process that is operating at an unacceptable level (Type II error or β-risk).

The modified limits are established, taking as input the APL_U and APL_L values, which are the largest and smallest permissible values of the mean, respectively (Figure 12). They are known as acceptance process levels ($APLs$) representing the interval at which the process mean can vary.

For the calculation of APL_U and APL_L, the specification limits, the standard deviation of the process, and acceptance quality level (AQL) are considered as shown below:

$$\begin{aligned} APL_U &= USL - Z_{AQL}\sigma \\ APL_L &= LSL + Z_{AQL}\sigma \end{aligned} \tag{2.10}$$

where USL and LSL are the upper and lower specification limits, σ represents the population standard deviation of the process and Z_{AQL} is the upper $100(1 - AQL)$ percentage point of the standard normal distribution related to the acceptable quality level (AQL) defined as the maximum of the rejected product proportions, whereas the process mean lies within the $APLs$ (MOHAMMADIAN; AMIRI, 2013).

In order to protect the proportion of rejected products against possible assignable

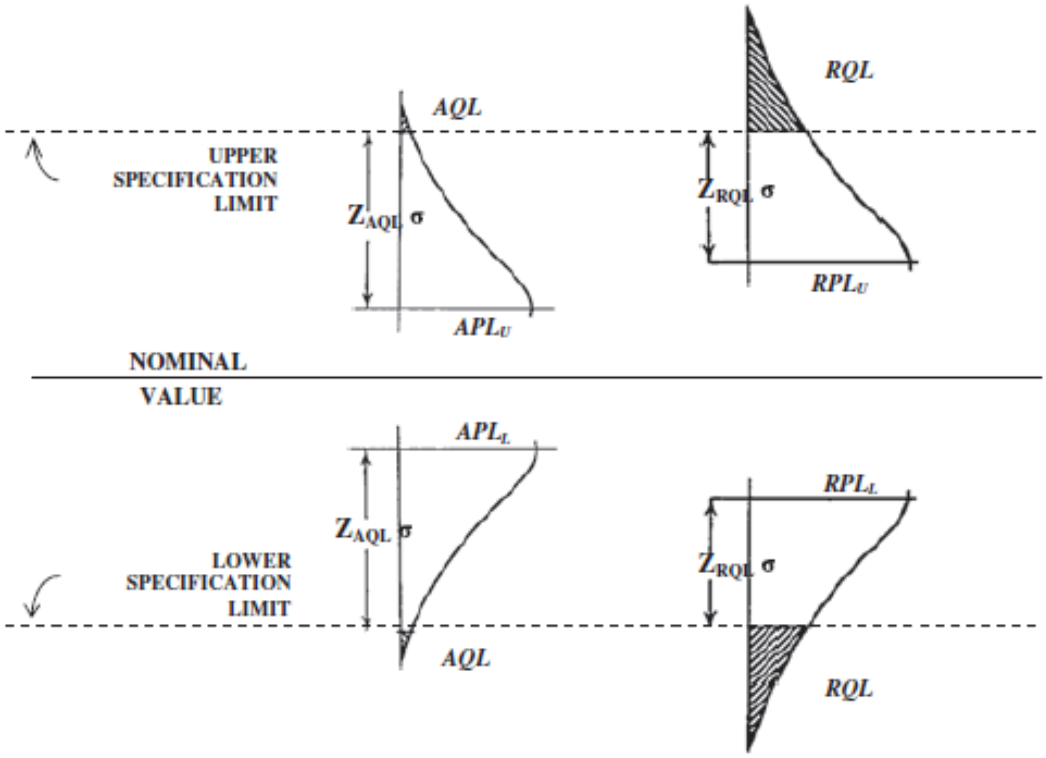


Figure 12 – AQL, $APLs$, RQL and $RPLs$ values in an Acceptance Control Chart
Source: Mohammadian e Amiri (2013)

causes, Mohammadian e Amiri (2013) propose a protective range known as the rejectable process levels (RPL_U; RPU_L), a process level which is rejectable and should be rejected most of the time. $RPLs$ are determined based on the specification limits, the standard process deviation, and the RQL, where RQL is defined as the maximum of rejected product proportion that can be tolerated. Therefore, RPL_U and RPU_L are calculated by expressions:

$$\begin{aligned} RPL_U &= USL - Z_{RQL}\sigma \\ RPL_L &= USL + Z_{RQL}\sigma \end{aligned} \tag{2.11}$$

where Z_{RQL} is a standard normal value associated with the probability RQL.

The region between APL and RPL is known as the zone of indifference because, in this region, the process can neither be accepted nor rejected (Figure 12).

The control chart is then plotted based on a sample size n, Type I error probability (α), and acceptable quality level (AQL). Thus, the control limits of ACC can be written as:

$$\begin{aligned} UCL_{ACC} &= USL - Z_{AQL}\sigma + Z_{\alpha}\frac{\sigma}{\sqrt{n}} \\ LCL_{ACC} &= LSL + Z_{AQL}\sigma - Z_{\alpha}\frac{\sigma}{\sqrt{n}} \end{aligned} \tag{2.12}$$

as represented in the Figure 13.

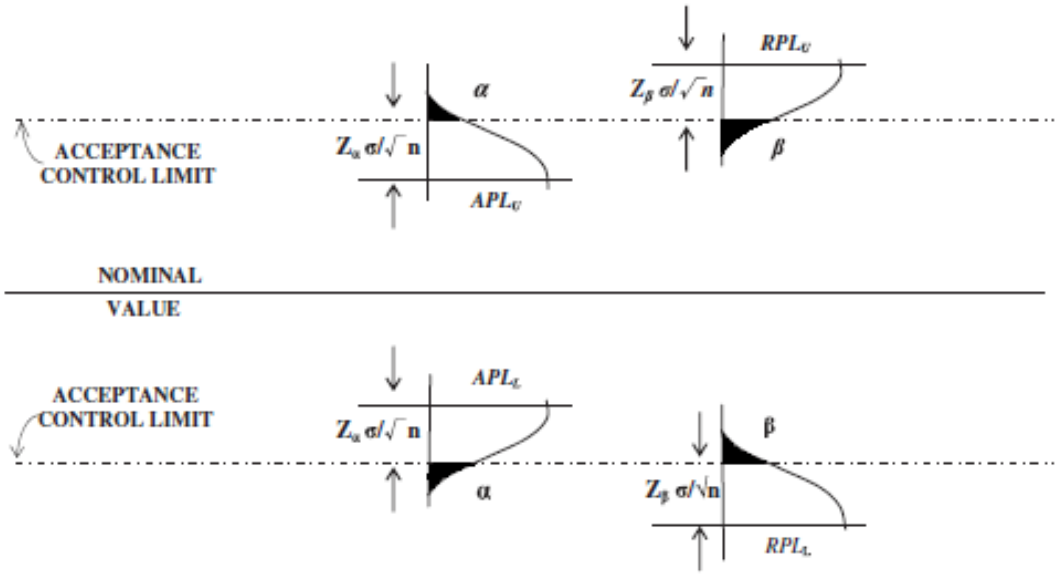


Figure 13 – Control limits and probability of Type I and Type II erros of ACC
Source: Mohammadian e Amiri (2013)

As indicated by Figure 13, we can determine α and $(1-\beta)$ to the UCL_{ACC} by Equations 2.13 and 2.14.

$$\begin{aligned} \alpha &= P\left(\bar{X} \geq UCL_{ACC} \mid \mu = APL_U\right) \\ &= P\left(\bar{X} \geq APL_U + Z_{\alpha}\frac{\sigma}{\sqrt{n}} \mid \mu = APL_U\right) \\ &= P\left(Z \geq Z_{\alpha}\right) = \int_{Z_{\alpha}}^{\infty} f(x)dx \end{aligned} \tag{2.13}$$

$$\begin{aligned} 1-\beta &= P\left(\bar{X} \geq UCL_{ACC} \mid \mu = RPL_U\right) \\ &= P\left(\bar{X} \geq APL_U + Z_{\alpha}\frac{\sigma}{\sqrt{n}} \mid \mu = RPL_U\right) \\ &= P\left(Z \geq Z_{AQL}\sqrt{n} + Z_{\alpha} - Z_{RQL}\sqrt{n}\right) = \int_{Z_{AQL}\sqrt{n}+Z_{\alpha}-Z_{RQL}\sqrt{n}}^{\infty} f(x)dx \end{aligned} \tag{2.14}$$

and to the LCL_{ACC} by Equations 2.15 and 2.16, respectively.

$$\begin{aligned}\alpha &= P\left(\bar{X} \leq LCL_{ACC} \mid \mu = APL_L\right) \\ &= P\left(\bar{X} \leq APL_L - Z_\alpha \frac{\sigma}{\sqrt{n}} \mid \mu = APL_U\right) \\ &= P\left(Z \leq Z_\alpha\right) = \int_{-\infty}^{-Z_\alpha} f(x)dx\end{aligned} \tag{2.15}$$

$$\begin{aligned}1-\beta &= P\left(\bar{X} \leq LCL_{ACC} \mid \mu = RPL_L\right) \\ &= P\left(\bar{X} \leq APL_L - Z_\alpha \frac{\sigma}{\sqrt{n}} \mid \mu = RPL_L\right) \\ &= P\left(Z \leq Z_{AQL}\sqrt{n} - Z_\alpha - Z_{RQL}\sqrt{n}\right) = \int_{-\infty}^{Z_{AQL}\sqrt{n}-Z_\alpha-Z_{RQL}\sqrt{n}} f(x)dx\end{aligned} \tag{2.16}$$

where $f(x)dx$ is standard normal probability function.

2.3.3 Functional approach to determine ACC limits

As in the capability index calculation, the existence of a functional relationship between the response variable y and the exploratory variable x imposes the need to consider the calculation of the limits of ACC over the entire range of values assumed by the exploratory variable (x_l, x_u). Thus, the limits UCL_{ACC} and LCL_{ACC} are expressed as $UCL_{ACC}(x)$ and $LCL_{ACC}(x)$, respectively, corresponding to the functional forms of the upper and lower control limits of the acceptance control chart for every value assumed by exploratory variable in range x_l and x_u, expressed in Equations 2.17.

$$\begin{aligned}UCL_{ACC}(x) &= USL(x) - Z_{AQL}\sigma(x) + Z_\alpha \frac{\sigma(x)}{\sqrt{n}} \\ LCL_{ACC}(x) &= LSL(x) + Z_{AQL}\sigma(x) - Z_\alpha \frac{\sigma(x)}{\sqrt{n}}\end{aligned} \tag{2.17}$$

where $x \in [x_l, x_u]$.

3 Process Capability Index $\hat{C}_p$ and $\hat{C}_{pk}$

A producer of goods will always attempt to persuade potential buyers that his (her) product is the best, being the leader in the field. Quality overkill may soon hinder industrial and business activities' progress and could have severe business cost and resource implications. Nevertheless, the widespread (and sometimes uncritical) use of Process Capability Indices (PCI) has led almost inadvertently to substantial quality improvements. However, better knowledge of their properties would occasionally prevent many wrong decisions regarding the process from being taken. In particular, quantifying a manufacturing process's "capability" is an essential initial step in any quality improvement program. Dictionaries usually define *capability* as: "the ability to carry out a task, to achieve an objective." As a rule, this activity usually involves an element of chance since the task may not be achievable every time, but we may be able to estimate what proportion of the time it can be achieved (WANG et al., 2021; MATSUURA, 2021; PEARN; KOTZ, 2006).

Process Capability Indices are widely used to determine whether a process can produce items within engineering/customer specifications limits, comparing the in-control process's output to the specification limits. It uses both the specification range and the process range to determine whether the process is capable (PERAKIS, 2021; AHMAD et al., 2019; AHMAD; ASLAM; JUN, 2016). The objective is to manufacture so that the measured value of the quality characteristic of interest (X) is $X = T$ (where T is the target value or nominal value) for each item produced. However, the reality is that X will end up being a "random" variable. It is often assumed that X is normally distributed with $X \sim N(\mu, \sigma^2)$.

The process's ability to produce close to the target or nominal value will depend on 1) the magnitude of σ, and 2) the relationship between μ and T. Ideally, we should have $\mu = T$. For a normal distribution, the interval of 6σ around μ contains all but 0.27% of the population. This value, called the capacity range, or just capacity, can give a general indication of the process's accuracy. Customers or manufacturers may wish to define a required level for product values x. This can take the form of a target-centered specification, $T \pm t$, or in the form of lower and upper specification limits (LSL, USL), the maximum acceptable range of key quality parameters Any item outside these limits will be considered scrap or need for rework, and cannot be sold or it does not work and must be classified as a nonconforming product. The use of specification limits allows the possibility that the midpoint m (between LSL and USL) is not the target. There may be situations where only one limit, LSL or USL, is needed.

The objective of process capability analysis is to estimate, monitor, and reduce the process's variability relative to process specifications (MONTGOMERY, 2019). The literature on process capability analysis is extensive and included univariate and multivariate processes where the quality measurements follow normal or non-normal distributions (AHMAD et al., 2019; FELIPE; BENEDITO, 2017; KOTZ; JOHNSON, 2002). More recently, authors as Oprime et al. (2019) and Oprime e Mendes (2017) have discussed the use of PCI in conjunction with control charts when the deviation of the process from the standard is small compared to the tolerance. In this case, the purpose is not to control but to approve the process.

Landim, Jardim e Oprime (2021) proposed a chart, named s^2 Modified Control Chart, where the process variance σ^2 is allowed to be larger than the in-control variance value σ_0^2 until a maximum value σ_{MAX}^2, as long as the process remains capable, only detecting genuine increases in process variance, which significantly increases the production rate of non-conforming items, preventing the practitioner from stopping the process and looking for attributable causes if only a small increase in process variation occurs. This is desirable in the sense that small increases in variance may not significantly affect the rate of non-conforming items being produced, and pausing the process incurs extra costs.

3.1 Traditional approach applied to determine capability index

Univariate method for measures process performance was suggested by Kane (1986) and investigated by other researchers such as Dianda, Quaglino e Pagura (2016), Wu, Pearn e Kotz (2009) and Chan, Cheng e Spiring (1988). Traditionally, the PCI most widely used in the evaluation of univariate processes are C_p and C_{pk} defined by Kane (1986) and expressed in Equations 3.1 and 3.2.

$$C_p = \frac{USL - LSL}{UNTL - LNTL} = \frac{USL - LSL}{6\sigma} \tag{3.1}$$

$$C_{pk} = \min\left\{\frac{USL - \mu}{UNTL - \mu}, \frac{\mu - LSL}{\mu - LNTL}\right\} = \min\left\{\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right\} \tag{3.2}$$

where μ represents the mean of process, σ is the process standard deviation, LSL is the Lower Specification Limit, and USL is the Upper Specification Limit, both determined externally.

The basic C_p gives only the magnitude of the overall process variation relative to the specification limits and cannot detect the process's shift from the specification center, so it cannot thoroughly explain the process performance, only indicating the potential

that the process has to meet the specifications. At the same time, the C_{pk} index considers the location of the process mean and the process variation relative to the specification limits, constituting a more reliable index of the real capacity of the process.

They may be set by management, the manufacturing engineers, the customer, or by-product developers/designers. $LNTL$ is the Lower Natural Tolerance Limit ($LNTL = \mu - 3\sigma$) and $UNTL$ is Upper Natural Specification Limit ($UNTL = \mu + 3\sigma$) (Figure 14). In a practical application, the mean μ and the process standard deviation σ is almost always unknown and must be replaced by an estimate, $\widehat{\mu}$ and $\widehat{\sigma}$, respectively, obtained from samples taken when the process is thought to be in control. For non-profile single variable, $\widehat{\mu}$, LSL, USL, $\widehat{LNTL}$ and $\widehat{UNTL}$ are considered as a single point (Figure 14).

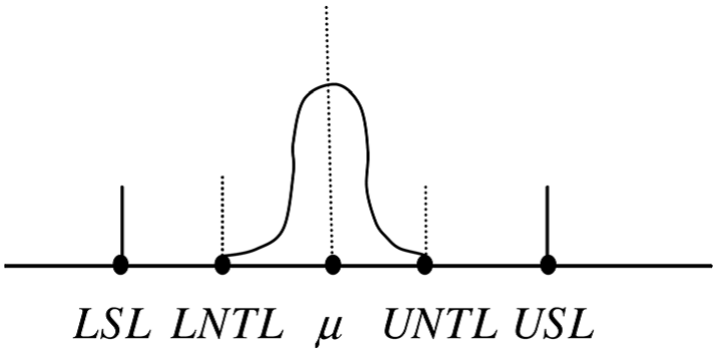


Figure 14 – Specification and Natural Tolerance Limits of a traditional univariate quality caracteristic.

3.2 Traditional approach applied to determine functional capability index

Although the capability of a process plays an essential role in evaluating the performance of a manufacturing process, Woodall (2007) points out that there is little attention to assessing the capability of processes whose quality characteristic is a profile. Keshteli et al. (2014b) reinforce that there are a few papers about the process capability index in profiles, and most of them are focused on the process capability index applied to linear profiles. However, as highlighted by the authors, in all of these methods, the response variables at different levels of the explanatory variable are considered, but the relationship in all ranges of the explanatory variable is neglected. They propose a new method to determine a process capability index for circular profile, where process capability index is defined as a functional form of the angle as explanatory variable. This results in measuring process capability in each level of all range of explanatory variable as well as determining a unique value of process capability index to give an overall judgment about process capability. The application of the proposed method is illustrated through a real case developed in an automotive industry.

Since that time, some studies have been carried out to introduce some process capability indices for profiles with both univariate and multivariate response data (MALEKI; AMIRI; CASTAGLIOLA, 2018). To exemplify the typical approach with which profile monitoring is handled, we can cite: Hosseinifard, Abbasi e Abdollahian (2011) use the Burr distribution to define a capability index applied to processes with a non-normal distribution; Hosseinifard e Abbasi (2012a) seeks to estimate the process capability index using the concept of nonconforming proportion applied to Phase I of a linear profile; Wang (2014) in turn, proposes two new process capability indices for linear profiles with unilateral specification, applying them in the process of producing electrolytic capacitors; Ghartemani, Noorossana e Niaki (2016) present a new approach to calculating the capability of linear profile processes by applying it to a leather industry, assuming that the explanatory variable assume fixed values while the specification limits are a linear relationships of the explanatory variable.

Another research conducted by Ghartemani, Noorossana e Niaki (2016) propose a method to estimate PCI of linear profiles, assuming functional relationships between the process specification limits and an explanatory variable. Based on multivariate PCI approach of Niavarani, Noorossana e Abbasi (2012), they assumed that the explanatory variable is fixed and the specification limits are a linear function of the explanatory variable. The results based only on simulation indicated appropriate performance of the proposed method, but deteriorate when coded values of explanatory variable is used.

In the profile monitoring, the existence of a functional relationship between the response variable y and the explanatory variable x imposes the need to consider the capability over the entire range of values of the explanatory variable $(x_l; x_u)$ (Figure 15).

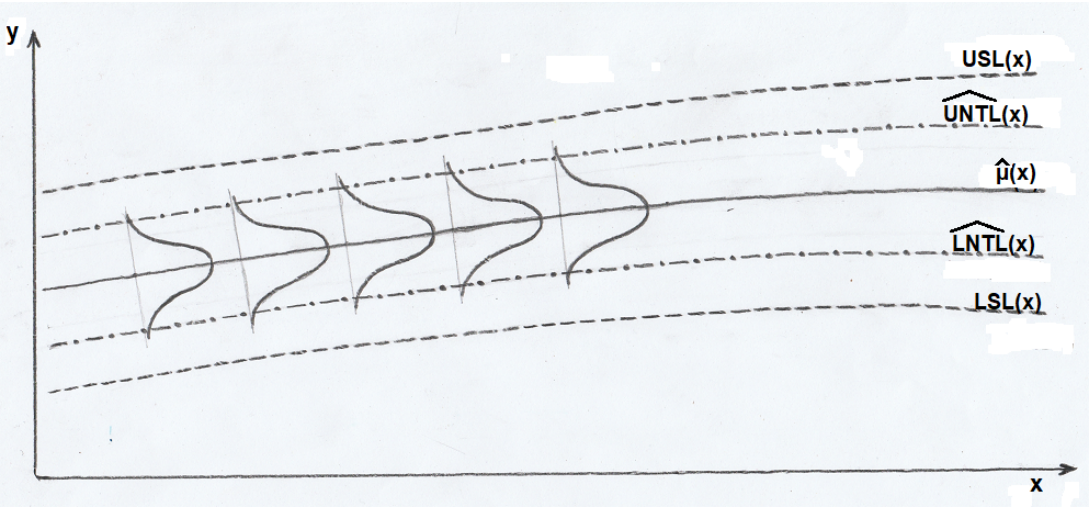


Figure 15 – Specification functional limits (LSL, USL) and Natural tolerance functional limits ($\widehat{LNTL}$, $\widehat{UNTL}$).

In this case, $\hat{\mu}$, $\widehat{LNTL}$, $\widehat{UNTL}$, LSL and USL, are expressed as $\hat{\mu}(x)$, $\widehat{LNTL}(x)$,

$\widehat{UNTL}(x)$, $LSL(x)$ and $USL(x)$, respectively, corresponding to the functional form of mean, natural limits and specification limits of the process. Keshteli et al. (2014a) propose to calculate $\widehat{C}_p$ and $\widehat{C}_{pk}$ by replacing functional forms of $\hat{\mu}$, LSL, USL, $\widehat{LNTL}$ and $\widehat{UNTL}$ in the traditional process capability index (Equations 3.1 and 3.2) obtaining the following equations:

$$\widehat{C}_p(x) = \frac{USL_y(x) - LSL_y(x)}{\widehat{UNTL}_y(x) - \widehat{LNTL}_y(x)}, x \in [x_l, x_u] \tag{3.3}$$

$$\widehat{C}_{pk}(x) = \min\left\{\frac{USL_y(x) - \hat{\mu}_y(x)}{\widehat{UNTL}_y(x) - \hat{\mu}_y(x)}, \frac{\hat{\mu}_y(x) - LSL_y(x)}{\hat{\mu}_y(x) - \widehat{LNTL}_y(x)}\right\}, x \in [x_l, x_u] \tag{3.4}$$

to evaluate the capability at each level of the explanatory variable. However, according to Keshteli et al. (2014a), it is necessary to have a unique value of the process capability index in all range of the explanatory variable in order to determine an overall judgment about process capability. Integrating the functional form of mean, natural limits and specification limits of the process over the entire range of values of the explanatory variable, that is, between x_l and x_u, it is possible to compute $\widehat{C}_p$(profile) by

$$\widehat{C}_p(profile) = \frac{\int_{x_l}^{x_u} [USL_y(x) - LSL_y(x)]\, dx}{\int_{x_l}^{x_u} \left[\widehat{UNTL}_y(x) - \widehat{LNTL}_y(x)\right] dx}, x \in [x_l, x_u] \tag{3.5}$$

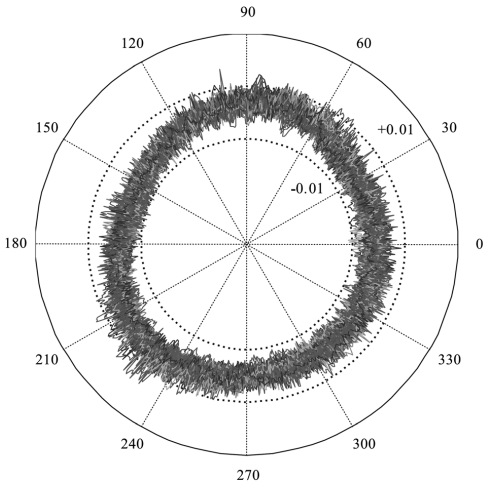
Similarly, the functional form of $\widehat{C}_{pk}$(profile) is computed by

$$\widehat{C}_{pk}(profile) = \min\left\{\frac{\int_{x_l}^{x_u} [USL_y(x) - \hat{\mu}_y(x)]\, dx}{\int_{x_l}^{x_u} \left[\widehat{UNTL}_y(x) - \hat{\mu}_y(x)\right] dx}, \quad \frac{\int_{x_l}^{x_u} [\hat{\mu}_y(x) - LSL_y(x)]\, dx}{\int_{x_l}^{x_u} \left[\hat{\mu}_y(x) - \widehat{LNTL}_y(x)\right] dx}\right\} \tag{3.6}$$

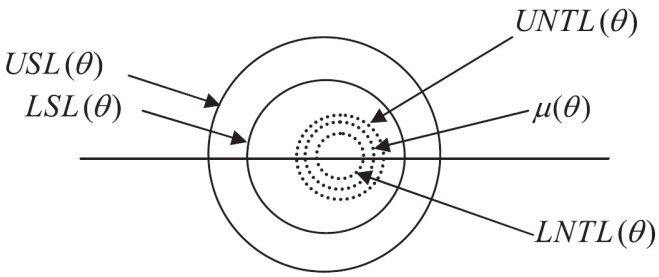
Machined or shaped profiles and surfaces can be treated as functional data if a set of coordinates, used to describe the position of each sampled point, can be represented as a function of one or more independent variables. When these independent variables are fixed, they can be treated as a counter or index of the sampled points.

Analyzing the problem of defining process capability for circular profile monitoring, Keshteli et al. (2014b) proposed an index for calculating capability for each level of the explanatory variable. Since natural limits of tolerance and specification of the circular profile have a functional shape, the process capability indices of the circular profile must have an angle-based functional shape (Figure 16). The authors then obtained a unique

value at each angle and, therefore, a unique value for the process capability indices at each angle.



(a) Polar representation of a circular profile



(b) C_p representation

Figure 16 – Process capability of a circular profile.
Source: Keshteli et al. (2014b)

The values of $UNTL_{R(\theta)}(\theta)$, $LNTL_{R(\theta)}(\theta)$, $USL_{R(\theta)}(\theta)$, $LSL_{R(\theta)}(\theta)$ e $\mu_{R(\theta)}(\theta)$ correspond to the functional form of the natural limits, specification limits, and the average of the traditional capability indices. Thus, Keshteli et al. (2014b) proposes to calculate C_p and do C_{pk} the following expressions

$$C_p(\theta) = \frac{USL_{R(\theta)}(\theta) - LSL_{R(\theta)}(\theta)}{UNTL_{R(\theta)}(\theta) - LNTL_{R(\theta)}(\theta)} \tag{3.7}$$

which allows calculate C_p for each value of the explanatory variable θ, where $\theta \in [0,\ 2\pi]$.

Integrating the expression of Equation 3.7 throughout the θ value range, it is possible to determine an overall capability index that is representative of process capability, as follows in Equation 3.8

$$C_p(profile) = \frac{\int_0^{2\pi} [USL_{R(\theta)} - LSL_{R(\theta)}]d\theta}{\int_0^{2\pi} [UNTL_{R(\theta)} - LNTL_{R(\theta)}]d\theta} \tag{3.8}$$

which, similar to the traditional C_p, only displays the potential ability of the process to meet specifications.

To calculate C_{pk}, we consider the "non-centrality" of the process around the average profile, thus obtaining the expression:

$$C_{pk}(\theta) = \min\left\{\frac{USL_{R(\theta)}(\theta) - \mu_{R(\theta)}(\theta)}{UNTL_{R(\theta)}(\theta) - \mu_{R(\theta)}(\theta)},\ \frac{\mu_{R(\theta)}(\theta) - LSL_{R(\theta)}(\theta)}{\mu_{R(\theta)}(\theta) - LNTL_{R(\theta)}(\theta)}\right\}, \quad \theta \in [0,\ 2\pi] \tag{3.9}$$

which determines the value of C_{pk} for each value of θ.

Equation 3.9 can, in turn, be used to determine a unique value for C_{pk} that simultaneously addresses the variability and "non-centrality" of the process, as we see in Equation 3.10

$$C_{pk}(profile) = \min\left\{\begin{array}{l}\dfrac{\int_0^{2\pi}\left[USL_{R(\theta)}(\theta) - \mu_{R(\theta)}(\theta)\right]d\theta}{\int_0^{2\pi}\left[UNTL_{R(\theta)}(\theta) - \mu_{R(\theta)}(\theta)\right]d\theta};\\ \dfrac{\int_0^{2\pi}\left[\mu_{R(\theta)}(\theta) - LSL_{R(\theta)}(\theta)\right]d\theta}{\int_0^{2\pi}\left[\mu_{R(\theta)}(\theta) - LNTL_{R(\theta)}(\theta)\right]d\theta}\end{array}\right\}, \quad \theta \in [0,\ 2\pi] \tag{3.10}$$

The authors point out that the proposed approach can easily be applied to other profile types, such as simple linear profiles, polynomial, and nonlinear profiles (KESHTELI et al., 2014b).

3.3 Sample distribution of $\hat{C}_p$ and $\hat{C}_{pk}$

Process capacity indices (PCI), such as C_p and C_{pk}, have been proposed in the manufacturing industry to provide numerical measures of whether a process can produce items within predefined specification limits (AHMAD; ASLAM; JUN, 2016). García-Díaz e Aparisi (2005) claim that it is common to find industries that have C_p and C_{pk}, greater than 2. In this case, such PCI can be used to design regions under control and out of control, making it possible to decide if it is necessary or not to stop the process for corrective actions. Thus, control charts for variables can be developed using PCI (C_p and C_{pk}) in order to establish and monitor the process, employing control charts based on the combination of two control mechanisms, namely, control limits (CL) and capability indices (SUBRAMANI; BALAMURALI, 2012).

Oprime et al. (2019) presents a mathematical development of the model for sample distribution of $\hat{C}_p$ and $\hat{C}_{pk}$ in the unconditional ARL function. According to Oprime et al. (2019), all these indices have a degree of uncertainty because they are estimated, being necessary to develop models that describe the probability distribution on the estimation of these indices.

Pearn e Lin (2004) investigated the natural estimator of the index $\hat{C}_{pk}$, showing that its distribution can be expressed as a mixture of the chi-square and the normal distributions under the assumption of normality, implementing the theory of hypothesis testing using the natural estimator of $\hat{C}_{pk}$.

The specification limits USL and LSL are the maximum value allowed for the product's characteristic, and the nominal specification value is the standard (ideal) value of the analyzed characteristic (N). Since the mathematical definitions of C_p and C_{pk} are given by 3.1 and 3.2, when $\mu_0 = N$ and the specification limits are equidistant from the nominal value, the maximum C_{pk} obtained has the same value as C_p (i.e. $C_{pk_{max}} = C_p = \frac{(USL-LSL)}{(6\sigma)} = \frac{(USL-\mu_0)}{(3\sigma_0)} = \frac{(\mu_0-LSL)}{(3\sigma_0)}$), wherein C_p is denoted as the potential capability and it does not consider the average location (OPRIME et al., 2019).

Pearn e Lin (2004) utilizing the identity $min\{a,b\} = (a+b)/2 - |a-b|/2$ present an alternatively index C_{pk} written as

$$C_{pk} = \frac{d - |\mu - m|}{3\sigma} \tag{3.11}$$

where $d = (USL - LSL)/2$ is half of the length of the specification interval, $m = (LSL + USL)/2$ is the mid-point between the lower and the upper specification limits. Replacing the process mean μ and the process stardard deviation σ by their conventional estimators $\bar{X}$ and S, which may be obtained from a process that is demonstrably stable (under statistical control), they defined the natural estimator $\hat{C}_{pk}$ as

$$\hat{C}_{pk} = \frac{d - \left|\bar{X} - m\right|}{3S} \tag{3.12}$$

Considering the r^{th} moment, and the first two moments as well as the mean and the variance of $\hat{C}_{pk}$ obtained by Kotz e Johnson (2002), under the assumption of normality, they define $K = (n-1)S^2/\sigma^2$, $Z = \sqrt{n}\left(\bar{X} - m\right)/\sigma$, $\xi = (\mu - m)/\sigma$, and $Y = |Z|$, rewritten $\hat{C}_{pk}$ as

$$\hat{C}_{pk} = \frac{\sqrt{n-1}\,(3C_p\sqrt{n} - Y)}{3\sqrt{nK}} \tag{3.13}$$

Kotz e Lovelace (1998 apud PEARN; LIN, 2004) argue that the construction of exact confidence intervals and test procedures for C_{pk} are complicated because the distribution of $\hat{C}_{pk}$ involves the joint distribution of two non-central t-distributed random variables, or as claimed by Pearn, Kotz e Johnson (1992), as a joint distribution of folded-normal and chi-square random variables. Under the normality assumption, K is distributed as χ^2_{n-1}, a chi-square distribution with $n-1$ degrees of freedom. Further, since Z is distributed as

the normal distribution $N(\sqrt{n}\xi, 1)$ with mean $\sqrt{n}\xi$ and variance 1, Y is distributed as the folded-normal distribution. Since $C_p = d/(3\sigma)$, and $|\xi| = 3(C_p - C_{pk})$, Pearn e Lin (2004) written probability density function of Y as

$$\begin{aligned} f_Y(y) &= \phi\left(y - \xi\sqrt{n}\right) + \phi\left(y + \xi\sqrt{n}\right) = \phi\left(y - |\xi|\sqrt{n}\right) + \phi\left(y + |\xi|\sqrt{n}\right) \\ &= \phi\left[y - 3(C_p - C_{pk})\sqrt{n}\right] + \phi\left[y + 3(C_p - C_{pk})\sqrt{n}\right], \quad y \geq 0 \end{aligned} \tag{3.14}$$

where $\phi(\cdot)$ is the probability density function of the standard normal distribution $N(0, 1)$.

The cumulative distribution function of $\hat{C}_{pk}$ can be obtained and expressed in terms of a mixture of the chi-square distribution and the normal distribution as presented by Vännman (1997). For $x > 0$

$$\begin{aligned} F_{\hat{C}_{pk}}(x) &= P\left(\hat{C}_{pk} \leq x\right) = P\left(\frac{\sqrt{n-1}\,(3C_p\sqrt{n} - Y)}{3\sqrt{nK}} \leq x\right) \\ &= 1 - P\left(\sqrt{nK} < \frac{\sqrt{n-1}\,(3C_p\sqrt{n} - Y)}{3x}\right) \\ &= 1 - \int_0^\infty P\left(\sqrt{nK} < \frac{\sqrt{n-1}\,(3C_p\sqrt{n} - Y)}{3x} | Y = y\right) f_Y(y)\,dy \\ &= 1 - \int_0^\infty P\left(\sqrt{nK} < \frac{\sqrt{n-1}\,(3C_p\sqrt{n} - y)}{3x}\right) f_Y(y)\,dy \end{aligned} \tag{3.15}$$

since K is distributed as χ^2_{n-1}, then

$$P\left(\sqrt{nK} < \frac{\sqrt{n-1}\,(3C_p\sqrt{n} - y)}{3x}\right) = 0, \tag{3.16}$$

for $y > 3C_p\sqrt{n}$ and $x > 0$. Hence:

$$\begin{aligned} F_{\hat{C}_{pk}}(x) &= 1 - \int_0^{3C_p\sqrt{n}} P\left(\sqrt{nK} < \frac{\sqrt{n-1}\,(3C_p\sqrt{n} - y)}{3x}\right) f_Y(y)\,dy \\ &= 1 - \int_0^{3C_p\sqrt{n}} P\left(K < \frac{(n-1)\,(3C_p\sqrt{n} - y)^2}{9nx^2}\right) f_Y(y)\,dy \end{aligned} \tag{3.17}$$

therefore

$$F_{\hat{C}_{pk}}(x) = 1 - \int_0^{3C_p\sqrt{n}} G\left(\frac{(n-1)\,(3C_p\sqrt{n} - y)^2}{9nx^2}\right) f_Y(y)\,dy \tag{3.18}$$

for $x > 0$ where $f_Y(y) = \phi[y + 3(C_p - C_{pk})\sqrt{n}] + \phi[y - 3(C_p - C_{pk})\sqrt{n}]$ and $G(\cdot)$ is the acumulative distribution function of the chi-square distribution χ^2_{n-1}.

Given the cumulative distribution function of $\hat{C}_{pk}$ expressed in Equation 3.18, Pearn e Lin (2004) present a test to determine if a given process is capable considering the following statistical hypotheses:

$$H_0 : C_{pk} \leq C \text{ (process is not capable)},$$
$$H_1 : C_{pk} > C \text{ (process is capable)}$$

Defining the decision making rule, test $\phi^*(x)$, as the following: $\phi^*(x) = 1$, if $\hat{C}_{pk} > c_0$; and $\phi^*(x) = 0$, otherwise. Thus, the test ϕ^* rejects the null hypothesis $H_0(C_{pk} \leq C)$ if $\hat{C}_{pk} > c_0$, with Type I error α, (i.e. the chance of incorrectly concluding an incapable process $C_{pk} \leq C$ as capable $C_{pk} > C$). As shown by Pearn e Lin (2004), if the values of α and C are given, the critical value c_0 can be obtained by solving the equation$P\left(\hat{C}_{pk} \geq c_0 \mid C_{pk} = C\right) = \alpha$ using available numerical integration methods. Set an specific value of C (the capability requirement), the p-value corresponding to c^*, a specific value of $\hat{C}_{pk}$ calculated from the sample data is

$$P\left(\hat{C}_{pk} \geq c^* \mid C_{pk} = C\right) = \int_0^{3C_p\sqrt{n}} G\left(\frac{(n-1)\left(3C_p\sqrt{n} - y\right)^2}{9n\left(c^*\right)^2}\right)\{\phi\left[y + 3(C_p - C_{pk})\sqrt{n}\right] + \phi\left[y - 3(C_p - C_{pk})\sqrt{n}\right]\}dy \quad (3.19)$$

Thus, given the capability requirement C, the process characteristic parameter C_p, sample size n, and risk α, the critical value c_0 can be obtained by solving the following equation:

$$\int_0^{3C_p\sqrt{n}} G\left(\frac{(n-1)\left(3C_p\sqrt{n} - y\right)^2}{9nc_0^2}\right)\{\phi\left[y + 3(C_p - C_{pk})\sqrt{n}\right] + \phi\left[y - 3(C_p - C_{pk})\sqrt{n}\right]\}dy = \alpha \quad (3.20)$$

3.4 $\bar{X}$ control chart with $\hat{C}_p$ and $\hat{C}_{pk}$ indices

Oprime et al. (2019) proposed a new design for an acceptance $\bar{X}$ control chart when the average of the process is unknown, and the standard deviation is known using expanded limits to create acceptable and unacceptable regions considering the capability index. The acceptance $\bar{X}$ proposed control chart considers the impact of the average

variation on the capability indices, monitoring the percentage of nonconformities and the C_{pk}, enabling the practitioners to make economic decisions, deciding when to stop the process for taking corrective measures. As stated by the authors, all processes can be classified according to the capability indices used to predict the performance of the process in meeting product specifications.

The framework proposed by Oprime et al. (2019), shown in Figure 17, has regions of control and out-of-control that considers average variations which impact on the capability indices and, consequently, on the fraction of nonconforming concerning the specifications of the product. The framework assumes that data has a normal distribution.

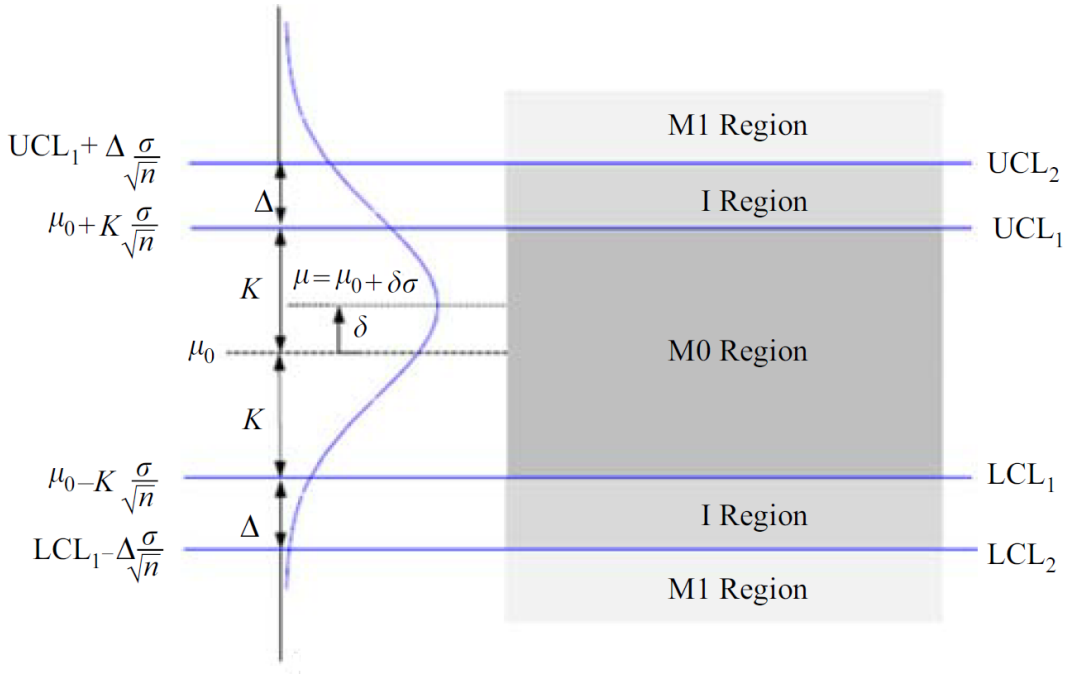


Figure 17 – Control chart design with expanded limits.
Source: Oprime et al. (2019)

The framework proposed presents three regions, as suggested by Woodall (1985): (M_0) when the process is in the acceptable condition; (M_1) when the process is in the unacceptable condition and (I) an indifferent region. In this new chart design built by Oprime et al. (2019), the first region (M_0) is given by the traditional Shewhart's control chart limits. The second region (I) is given by the expanded limits, considering C_p and C_{pk} (indiferent region). The third region (M_1) is when the average exceeds the expanded limits, indicating that there were unacceptable variations in the parameter. The unacceptable region was created considering the rule that a signal that exceeds the limits indicates an out-of-control process.

The $\bar{X}$ control chart is then designed defining CL as a function of the C_{pk} index, considering that C_{pk_0} is the lowest acceptable value for the process,

(1) When $C_{pk} < C_{pk_0} : CL = \bar{\bar{X}} \pm 3\left(\hat{\sigma}/\sqrt{n}\right)$ (traditional Shewhart's limits)
(2) When $C_{pk} \geq C_{pk_0} : CL = \bar{\bar{X}} \pm \left[\left(3\left(C_p - C_{pk}\right)\hat{\sigma}_0\right) + 3\left(\hat{\sigma}_0/\sqrt{n}\right)\right]$

In the new design proposed, the limits were expanded and out-of-control signals are now associated with the levels of the capability indices:

$$M_0 : \left[\bar{\bar{X}} - 3\left(\hat{\sigma}/\sqrt{n}\right) \leq \bar{X} \leq \bar{\bar{X}} + 3\left(\hat{\sigma}/\sqrt{n}\right)\right]$$

$$\begin{aligned} I : &\left[\left(\bar{\bar{X}} - 3\left(C_p - C_{pk}\right)\hat{\sigma}_0\right) - 3\left(\hat{\sigma}_0/\sqrt{n}\right) \leq \bar{X} < \bar{\bar{X}} - 3\left(\hat{\sigma}/\sqrt{n}\right)\right] \text{ and} \\ &: \left[\bar{\bar{X}} + 3\left(\hat{\sigma}/\sqrt{n}\right) \leq \bar{X} < \left(\bar{\bar{X}} + 3\left(C_p - C_{pk}\right)\hat{\sigma}_0\right) + 3\left(\hat{\sigma}_0/\sqrt{n}\right)\right] \end{aligned} \tag{3.21}$$

$$\begin{aligned} M_1 : &\left[\left(\bar{\bar{X}} - 3\left(C_p - C_{pk}\right)\hat{\sigma}_0\right) - 3\left(\hat{\sigma}_0/\sqrt{n}\right) > \bar{X}\right] \text{ or} \\ &: \left[\left(\bar{\bar{X}} + 3\left(C_p - C_{pk}\right)\hat{\sigma}_0\right) + 3\left(\hat{\sigma}_0/\sqrt{n}\right) < \bar{X}\right] \end{aligned}$$

Since both M_0 and I regions are considered as acceptable, the proposed design use only two regions, acceptable (M_0 and I) and unacceptable (M_1) region. It's considered that a signal beyond the proposed new limits indicates that the process capability is compromised by a variation of the average greater than acceptable (only the one-point rule exceeding the CL was used in the performance analysis of the control chart).

3.4.1 Functional $\bar{X}$ control chart with $\hat{C}_p$ and $\hat{C}_{pk}$ indices

As stated before, the existence of a functional relationship between the response variable Torque and the explanatory variable RPM imposes the need to consider $\bar{X}$ control chart with $\hat{C}_p$ and $\hat{C}_{pk}$ indices over the entire range of values of the explanatory variable RPM $(r_l; r_u)$.

This work introduces an average variation permissible for the control limits in the traditional $\bar{X}$ control chart when it may have little or no practical importance. Using this new design with the limits expanded, and the out-of-control signals associated with the levels of the capacity indices according to the regions defined by Equation 3.22 and shown in Figure 18 allow economic decisions about the process applying $\bar{X}$ control chart with $\hat{C}_p$ and $\hat{C}_{pk}$ indices separately to each data point observed at a given location.

However, as discussed before, since that P dependent control rules are applied simultaneously, the Bonferroni rule for dependent events must be used to achieve an actual false alarm rate not greater than a predefined value. Therefore, let α denote the upper

limit of the first type of probability error (false alarm probability); the value $\alpha = \alpha'/P$ is used to calculate the control limits of Equation 3.22.

3.4.2 Functional $\bar{X}$ control chart limits with $\hat{C}_p$ and $\hat{C}_{pk}$ indices

Assuming that the process produces items whose Torque distribution has a normal distribution at each RPM value considered within the operating range, that is, between r_l and r_u, lower and upper limit electric motor's RPM operating range, respectively, then there will be little danger of produce a defective item until the functional form of mean $\hat{\mu}(r)$ has shifted into rejectable process zone.

As previously stated in Section 2.2.2, the functional relationship between the response variable $Torque$ and the exploratory variable RPM imposes the need to consider the calculation of the control limits (CL) over the entire range of values assumed by the exploratory variable (r_l, r_u). Thus, the limits UCL and LCL are expressed as $UCL(r)$ and $LCL(r)$, respectively, corresponding to the functional forms of the upper and lower control limits of the acceptance control chart for every value assumed by exploratory variable in range r_l and r_u, expressed in Equations 3.22. For the proposed design in this work we will use only two regions, acceptable region (M_0) and unacceptable region (M_1), since the M_0 and indifferent zones I are both considered as acceptable (Figure 18).

Based on the previous framework developed by Oprime et al. (2019) to address a univariate problem, we extend and create a functional approach to be applied over the entire range of values assumed by the explanatory variable RPM.

$$
\begin{aligned}
M_0 &: \left[\bar{\bar{X}}(r) - k\left(\hat{\sigma}(r)/\sqrt{n}\right) \leq \bar{X}(r) \leq \bar{\bar{X}}(r) + k\left(\hat{\sigma}(r)/\sqrt{n}\right)\right] \\
I &: \left[\bar{\bar{X}}(r) - 3\left(\left(C_p(r) - C_{pk}(r)\right)\hat{\sigma}_0(r)\right) - k\left(\hat{\sigma}_0(r)/\sqrt{n}\right) \leq \bar{X}(r) < \bar{\bar{X}}(r) - k\left(\hat{\sigma}(r)/\sqrt{n}\right)\right] \\
&\text{and} \\
&: \left[\bar{\bar{X}}(r) + k\left(\hat{\sigma}(r)/\sqrt{n}\right) \leq \bar{X}(r) < \bar{\bar{X}}(r) + 3\left(\left(C_p(r) - C_{pk}(r)\right)\hat{\sigma}_0(r)\right) + k\left(\hat{\sigma}_0(r)/\sqrt{n}\right)\right] \\
M_1 &: \left[\bar{\bar{X}}(r) - 3\left(\left(C_p(r) - C_{pk}(r)\right)\hat{\sigma}_0(r)\right) - k\left(\hat{\sigma}_0(r)/\sqrt{n}\right) > \bar{X}(r)\right] \text{ or} \\
&: \left[\bar{\bar{X}}(r) + 3\left(\left(C_p(r) - C_{pk}(r)\right)\hat{\sigma}_0(r)\right) + k\left(\hat{\sigma}_0(r)/\sqrt{n}\right) < \bar{X}(r)\right]
\end{aligned}
\tag{3.22}
$$

where r corresponds to the values assumed by the exploratory variable RPM between r_l and r_u, k is the $Z_{\alpha/2}$ corresponding to k sigma limits of Shewhart control charts corrected by Bonferroni. In this work, it was assumed the standard value $\alpha^{'} = 0.0027$ as upper bound for the Type I error probability. So, the value $\alpha = 1,588x10^{-4}$ (i.e. $Z_{\alpha/2} = 3,777$)

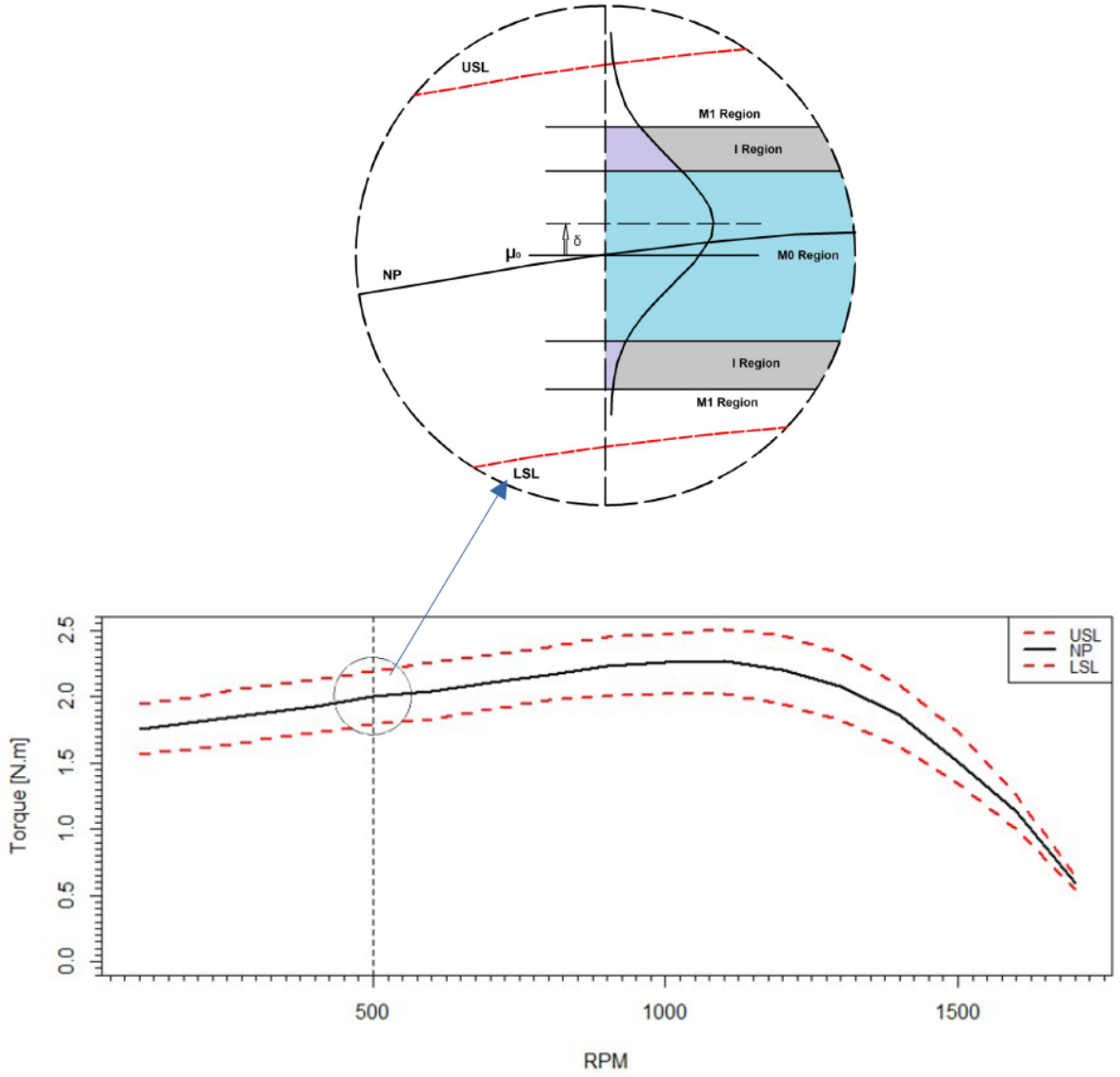


Figure 18 – (M_1): unacceptable condition, (I): indifferent region, and (M_0): traditional Shewhart's control chart limits at 500 RPM.

was used for designing the 17 control limits in Equation 3.22 over the region from r_l = 100 RPM to r_u = 1700 RPM.

Oliveira, Oprime e Jardim (2018) proposed a new design for an acceptance

Oprime et al. (2015) proposed a new design for an acceptance

4 Data Tidying and Computer Procedures

The development of this research and the formulation of mathematical models involved imply acquiring data from two distinct sources: i) their theoretical curves adjusted, and ii) real data of the functional relationship between the torque and the rotation in revolutions per minute (RPM) of sampled electric motors.

The data used to determine the functional relationship between the torque of electric motors and the correspondent rotation in revolutions per minute (RPM) were made available by the research and development sector of a Brazilian company's industrial plant. The company kindly assigned these data subject to the confidentiality terms established between this researcher and the company.

We used RStudio (RStudio Team, 2019), an Integrated Development Environment for the R language (R Core Team, 2020) to implement computational solutions. R is a programming language for statistical computing and graphics to clean, analyze, and graph data, and provides routines for exploratory and descriptive analysis of functional data (i.e., information summarized in the form of profiles, where each data point is the response observed in a given location, spatial or temporal). It is widely used by researchers from diverse disciplines to estimate and display results and by teachers of statistics and research methods. To develop the work, we use two additional packages: the *qcc package* to analyze the process capability and plot its respective graphs, and the *tydeverse package* used to plot *ARL* graphs.

The R code is available in Annex B. It was not possible to make the database available due to the confidentiality terms established by the company.

4.1 Modeling the theoretical electric motor operating curve

The theoretical operating curves of electric motors were made available by the company in the form of "pdf" files indicating the nominal Torque x RPM ratio and the respective upper and lower specification limits (Figure 19).

Because it was a "pdf" file, the first step was to extract numerical data from the information available in the form of an image. The following procedure was adopted:

1) obtaining the XY coordinates from the image

2) axis calibration by clicking known values to interpolate a coordinate system

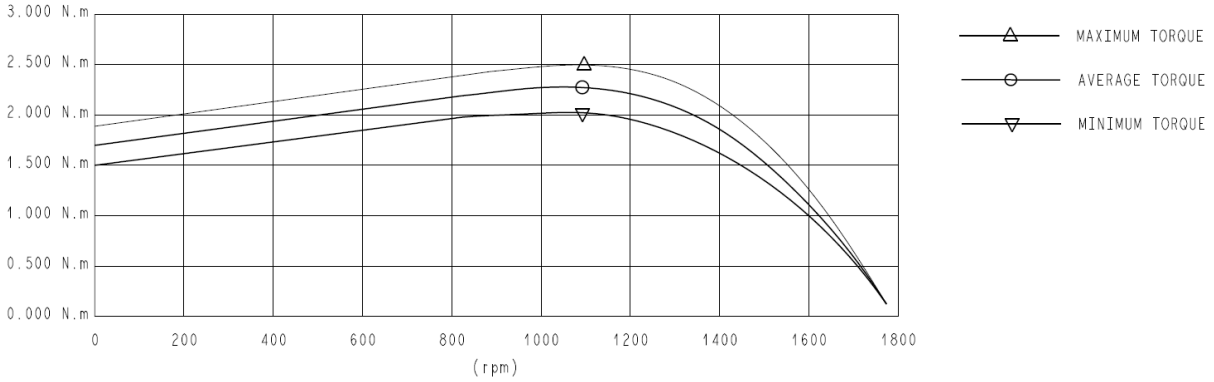


Figure 19 – Torque curve made available by the manufacturer of electric motors.

3) manually click each data point in a data series

This procedure was done using WebPlotDigitizer software, a free, open-source tool that can operate on a wide variety of image file types (ROHATGI, 2011). Seventeen pairs of points were obtained between the 100 and 1700 RPM range for the curves corresponding to the nominal profile (NP) and the lower (LSL) and upper (USL) specification limits and were saved in "csv" format (Table 1). These data were later used in the mathematical modeling of theoretical curves and the determination of process capability indices.

Table 1 – Theoretical values for Torque x RPM of electric motor Model 3, 220V

RPM	LSL	NP	USL
100	1.5612	1.7519	1.9446
200	1.6124	1.8156	2.0051
300	1.6730	1.8706	2.0755
400	1.7228	1.9232	2.1230
500	1.7916	1.9989	2.1886
600	1.8327	2.0433	2.2589
700	1.9033	2.1105	2.3138
800	1.9711	2.1620	2.3729
900	1.9988	2.2298	2.4449
1000	2.0164	2.2568	2.4699
1100	2.0180	2.2712	2.5061
1200	1.9489	2.2060	2.4594
1300	1.8183	2.0768	2.3299
1400	1.6239	1.8581	2.0885
1500	1.3458	1.5090	1.7359
1600	0.9979	1.1346	1.2549
1700	0.5444	0.5869	0.6349

4.1.1 Theoretical curve/model adjustment

The theoretical model of the Torque x RPM relationship shown in Figure 19 was nonexistent, or at least, it was unknown by this researcher. Such a theoretical relationship is of fundamental importance in developing this work since it makes it possible to determine the torque specification limits for each rotation's value of the electric motor and its capability index. The theoretical model also plays an essential role in implementing the simulations since it provides the base (Nominal Profile) from which we simulate the samples.

Clearly, Torque and RPM's relationship shown in Figure 19 is defined by a curve and not a line. In this case, as the relationship between Torque and RPM is non-linear, it will be modeled using polynomial regression, a particular case of multiple linear regression. In general, we can model the expected value of y as an k^{th} degree polynomial, yielding the general polynomial regression model, as follows:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \beta_k x_i^k + \epsilon_i,\ i = 1, 2, \cdots, n \tag{4.1}$$

where n is the pairs of observations $(x_i, y_i); i = 1, 2, \cdots, n$, k is the degree of the polynomial and ϵ_i is the random error associated with the response y_i.

To determine the model that best fits the theoretical curve data was used the R stats package (R Core Team, 2020). For each model, its parameters, the coefficient of determination R^2, and the coefficient of determination adjusted R^2_{adj} were calculated (Table 2).

Table 2 – Curve fitting model parameters of theoretical curve

Model	β_0	β_1	β_2	β_3	β_4	β_5	β_6	R^2	R^2_{adj}
Model 1	1.272e+00	2.584e-03	-1.645e-06	—	—	—	—	0.8567	0.8362
Model 2	1.872e+00	-9.253e-04	3.092e-06	-1.754e-09	—	—	—	0.9942	0.9929
Model 3	1.721e+00	4.124e-04	-3.484e-08	8.945e-10	-7.358e-13	—	—	0.9984	0.9979
Model 4	1.618e+00	1.645e-03	-4.302e-06	6.928e-09	-4.447e-12	8.246e-16	—	0.9993	0.9989
Model 5	1.686e+00	6.309e-04	4.258e-07	-2.764e-09	5.281e-12	-3.865e-15	8.685e-19	0.9994	0.9991

Graphically, it is possible to observe from the Figure 20 that the second and third-order models, Model1 and Model2, do not fit properly to the data. The fourth (Model3), fifth (Model4), and sixth (Model5) order models, in turn, have a much better fit quality. The choice between Model3, Model4, and Model5 made only based on the determination coefficient R^2 may be inappropriate since R^2 can be made artificially high by overfitting, that is, by including many terms in the model.

To determine which model terms are statistically significant, analysis of variance (ANOVA) between Models 3, Model 4, and Model 5 was performed using the anova function

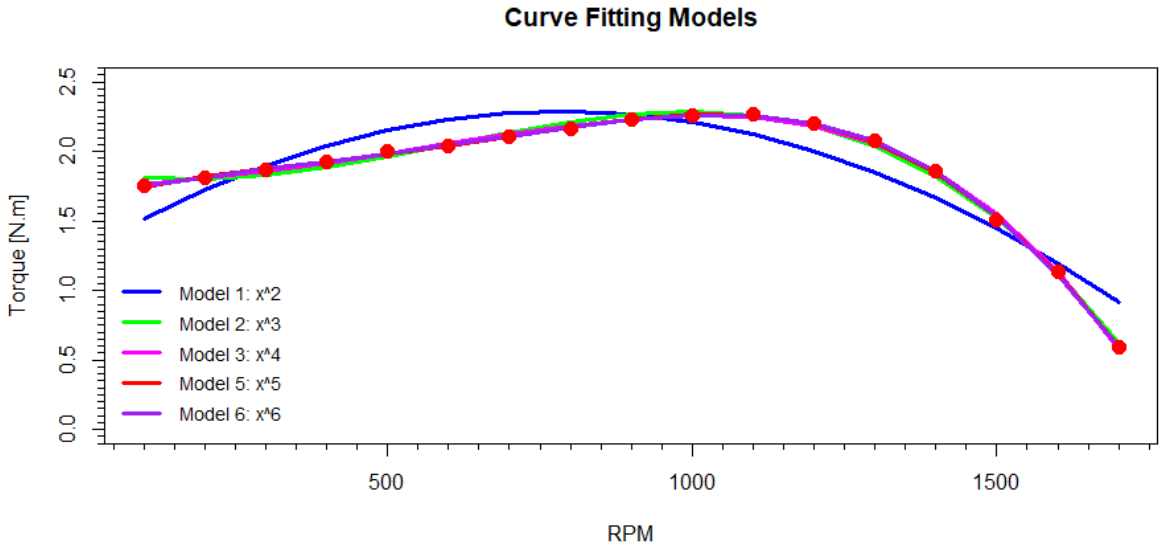


Figure 20 – Curve fitting models for theoretical curve of electric motor Model 3, 220V

Table 3 – Analysis of variance for curve model fitting.

Model	Res.DF	RSS	Df	Sum of Sq	F	Pr(>F)
Model1	14	0.44758	1			
Model2	13	0.01810	1	0.42948	2408.2570	2.981e-13
Model3	12	0.00501	1	0.01309	73.4274	6.419e-06
Model4	11	0.00227	1	0.00274	15.3706	0.002864
Model5	10	0.00178	1	0.00048	2.7149	0.130437

available in the stats package of R. The results obtained are shown in the Table 3.

The partial F-statistic given by the anova function measures the effect of the sum of squares of error by an increase in the extra variables in the equation. The test statistic value F decreases as the sum of squares of error decreases by the addition of extra variables to the equation. If the statistic F is not decreased, the extra variable cannot justify for its inclusion in the equation and hence the extra variables can be excluded.

It can be seen from the data in Table 3 that, at the 0.01 level of significance, both Model4 and Model5, aren't significantly better than Model3, which explains 99.84% of the variability observed in the response variable. However, as Walpole et al. (2007) points out, it should be noted that the insignificance of any coefficient does not necessarily imply that it does not belong in the final model. It merely suggests that it is insignificant in the presence of all other variables in the problem. Given candidate models of similar predictive or explanatory power, the simplest model is most likely to be the best choice.

So, for the development of the models, it was assumed the existence of m samples of size n_j in the form (x_{ij}, y_{ij}), $i = 1, 2, \cdots, n_j$, $j = 1, 2, \cdots, m$, where the subscript i

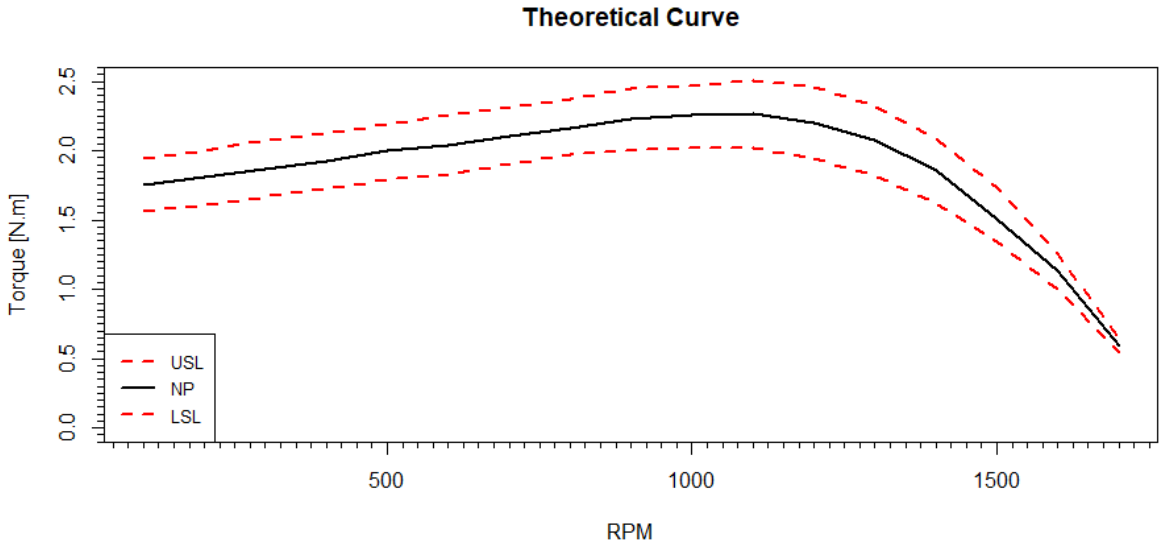


Figure 21 – Theoretical curve of electric motor Model 3, 220V

represents the i^{th} observation of each j^{th} sampled profile. In this research, it was assumed that there is a k^{th} order polynomial relationship between the response variable y and the exploratory variable x, as follows:

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + \beta_{2j}x_{ij}^2 + \beta_{3j}x_{ij}^3 + \beta_{4j}x_{ij}^4 + \epsilon_{ij} \tag{4.2}$$

where the errors ϵ_{ij} are independent and identically distributed (i.i.d), with the mean zero and variance σ_{ij}^2 for all i and j, (i.e., $\epsilon_{ij} \sim N(0, \sigma_j^2)$).

4.2 Modeling real Torque x RPM electric motors functional relationship

The company provided the actual data used in this research and which served as the basis for the formulation of mathematical models in the "csv" format. The data were obtained using a dynamometer developed by its engineering department. The specific characteristics and technical details of the dynamometer operation were not provided because its confidential technical information as outlined in the confidentiality terms established between the university, the company, and the researchers for data availability.

Three electric motor models at voltages 127 and 220 volts were evaluated for their Torque x RPM functional relationship. To ensure compliance with the confidentiality terms established, the identification of the models was masked. Table 4 presents the number of electric motors sampled for each model and voltage combination and their average number of observations.

Due to the inherent characteristics of the dynamometer data acquisition system, the number of observations obtained in each sample is variable. For example, in the case of Model 3, voltage 220 V, the number of observations ranged from 1897 (Sample 3) to 2055 (Sample 11), with the average number of observations shown in Table 4.

Table 4 – Electric motor models sampled

Model	Voltage [V]	Samples	Average Number of Observations
1	127	3	1945
1	220	3	1983
2	127	6	1969
2	220	7	1960
3	127	7	1987
3	220	11	1937

For unknown reasons, the dynamometer's behavior at the ends of the acquisition range presents an erratic behavior (Figure 22). For this reason, in this work will be considered only the values between 100 and 1700 RPM.

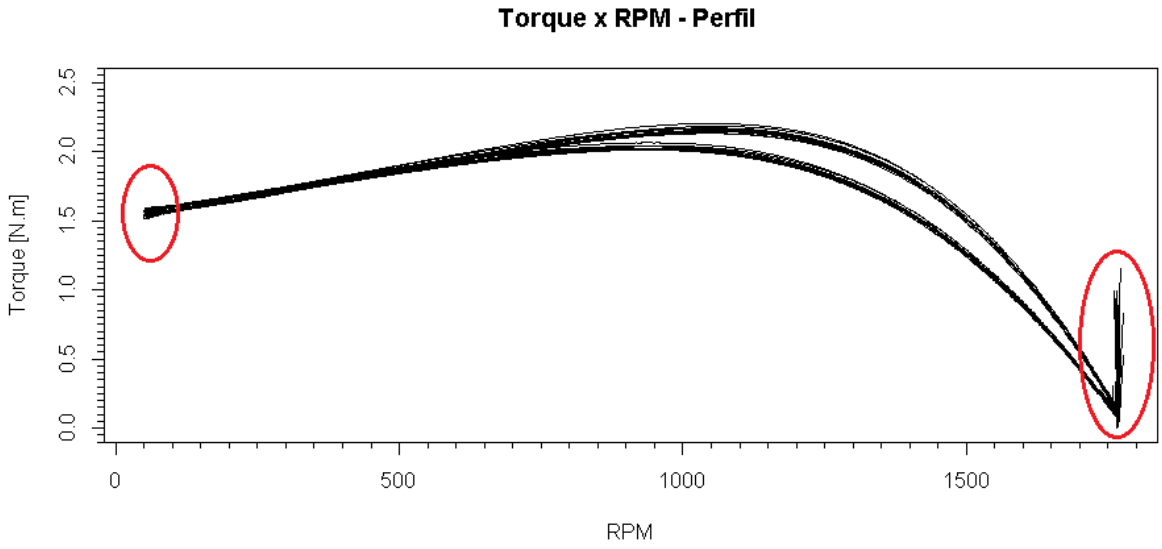


Figure 22 – Torque curves of Model 3 - 220 V, got from dynamometer system indicanting (in red) erratic behaviour on extremities of acquisition range.

The dynamometer data acquisition cycle comprises the period in which a load is applied to the shaft of the electric motor and the period in which the shaft is released. As a result, the empirical data can be represented by two torque curves (Figure 23). To express the functional relationship between Torque and RPM representative of the electric motor performance, only data related to the condition of the shaft under load will be considered.

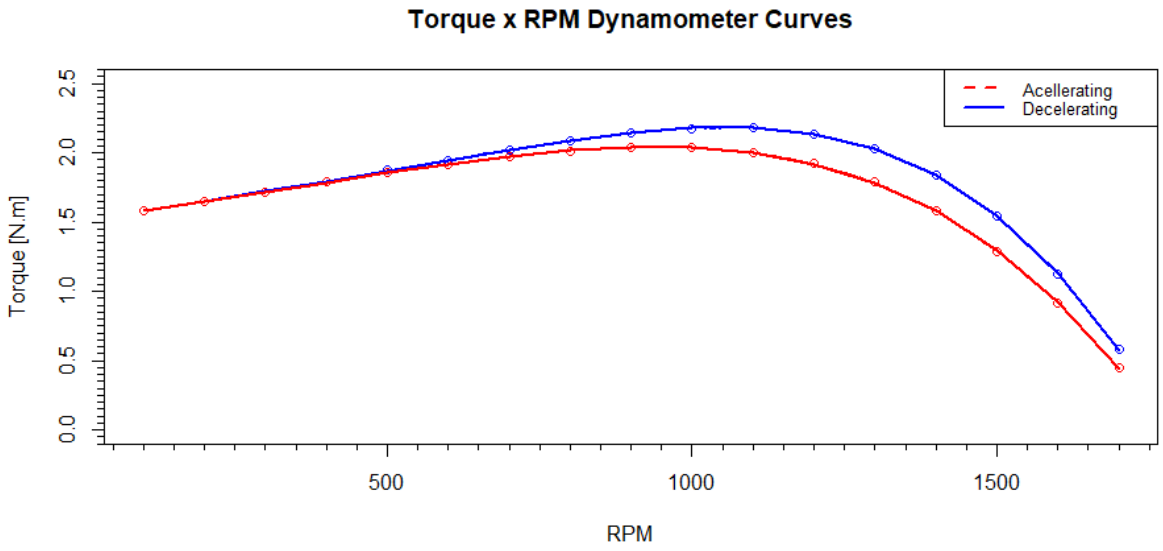


Figure 23 – Curves of Model 3 - 220 V, representing acellerating and decelerating torque conditions .

Another characteristic inherent to the dynamometer acquisition system was the non-regularity observed in the acquisition interval between successive samples, as noted in Figure 24.

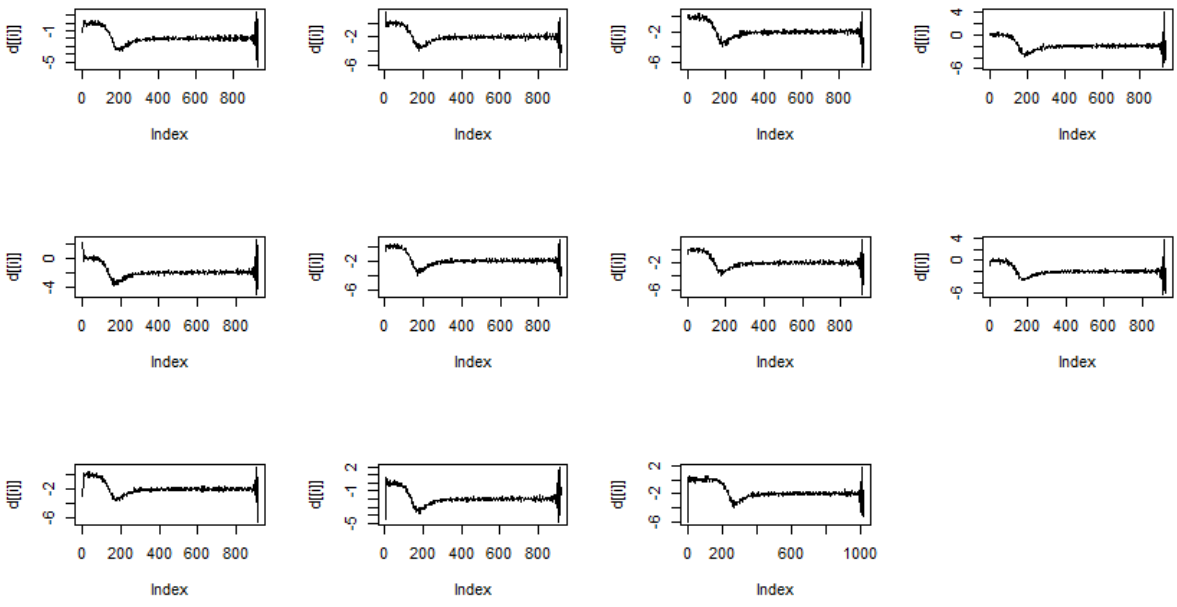


Figure 24 – Differences in successive observations obtained by the dynamometer for Model 3 - 220V.

Due to the irregular behavior observed in the dynamometer acquisition system at the extremities of the acquisition range, (i.e., values below 100 RPM and above 1700 RPM), these were disregarded in determining the functional relationship between Torque

and RPM for the sampled electric motors. To determine a unique set of values for the exploratory variable (RPM) applied to each sampled profile, an interval of 100 volts was established, totaling 17 points. As the dynamometer's signal acquisition system does not allow establishing the values at which the RPM was sampled, it was necessary to obtain the torque value for each of the RPM through a simple linear regression. This resulted in a set of 17 values between 100 and 1700 RPM and the corresponding torque for each sampled profile (Figure 25).

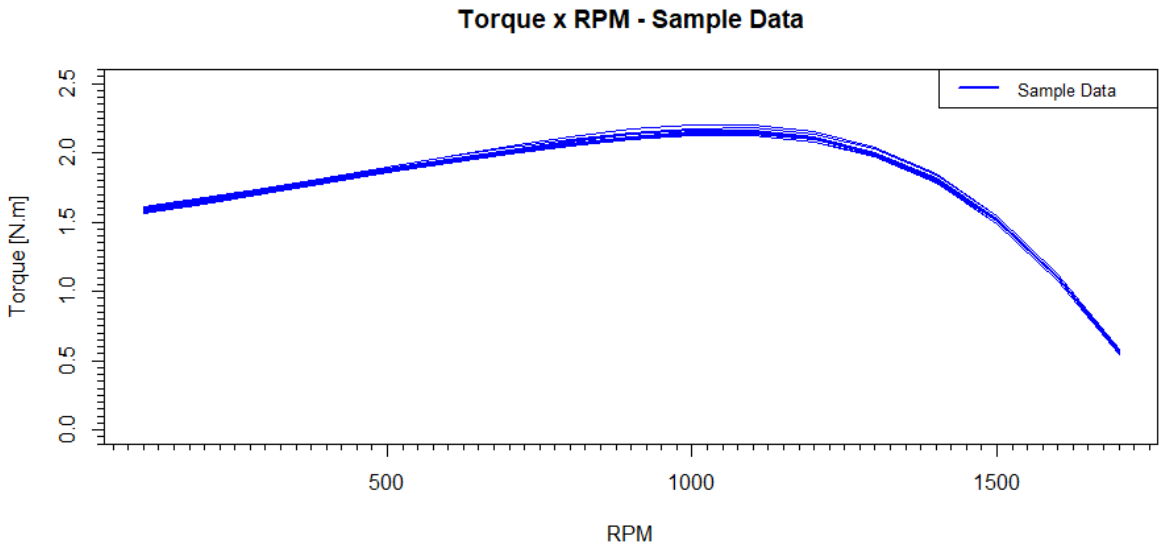


Figure 25 – Torque x RPM adjusted profiles for electric motor Model 3, 220 volts.

4.3 The Monte Carlo Computer Simulations Approach

We used to calculate the characteristics of the run-length (ARL) distribution. This method's popularity stems from the fact that, theoretically, no matter how complex the distribution of run-lengths is, computer simulations can almost always be used with relative ease to calculate precisely its distribution and their associated characteristics if the simulation is large enough (KONRATH; DONATELLI; HAMBURG-PIEKAR, 2006). In this work, we use 100000 simulations, as it is well known that the error of a run-length function can be limited by sufficiently increasing the size of the simulation.

For this work, we implemented the simulation using the statistical language programming R, running on a notebook equipped with Processor Intel(R) Core(TM) i3-8130U CPU 2.20GHz, RAM 12.0 GB, and running Windows 10 64-bit operating system.

Below, we show a computer simulation procedure's generic steps to calculate the run-length distribution for a two-sided control chart.

1. Specify the necessary parameters, such as the number of samples (m), number of profiles per sample (n) ; number of observations per profile (p), Type of error I assumed (α) and standard deviation, obtained from sample data (σ);

2. Simulate m samples with n profiles containing p observations centered on the nominal profile and standard deviation σ, assuming normal distribution and regression model defined in Equation 4.2;

3. Specify potential process capability (C_p) and minimum allowable capability (C_{pk}) and calculates the control limits for the simulated values using Bonferroni correction for 17 simultaneously hypothesis test;

4. Simulate n profiles containing p observations centered at δ standard deviations from the nominal profile and standard deviation σ, assuming normal distribution and regression model defined in Equation 4.2;

5. Determine the number N of profiles simulated presents at least one point outside of control limits;

6. Repeat steps 4 to 5 in total 100000 and calculates ARL as the proportion N/100000;

7. Repeat steps 4 to 6 varying δ from 0.25 to 2.50 incrementing by 0.25;

8. After obtaining a "data set" with 100000 observations of the run-length distribution, repeat steps 4 to 7 in total 100 times to get the characteristics of interest for the run-length distribution and descriptive statistics such as mean, standard deviation, median, and percentiles.

5 Results and Discussion

As stated before in Section 2.2.1, the location control chart consists of applying a Shewhart's control chart separately to each data point observed at a given location. For this work, each of the values assumed by the exploratory variable RPM in the speed range between 100 and 1700 RPMs will be considered a local point where, with a given probability, the data observed at that specific location should stay within the control limits.

The logic behind this approach is that, if the observed shape obtained from the functional relationship between Torque x RPM is under control, the data observed at that specific location must remain within the control limits with a certain probability. So, an alarm is issued when at least one point, in the whole set of data observed, exceeds the control limits.

To assess the relative performance of the proposed model, we develop and simulates location control chart models by adopting two approaches: Shewhart's traditional approach and the proposed functional control chart based on $\hat{C}_p$ and $\hat{C}_{pk}$ capability indices.

5.1 Design of the location control chart with traditional Shewhart approach

Since 17 dependent control rules are applied simultaneously, the percentile of standardized normal distribution used to compute $Z_{\alpha/2}$ in Equation 2.5 should be corrected. For that, we applied Bonferroni's rule for dependent events to attain a false alarm rate not greater than a predefined probability of False Alarm or Error Type I. Let $\alpha^{'}$ denote the upper bound of False Alarm rate, the value $\alpha = \alpha^{'}/17$ is used for designing the 17 control limits in Equation 2.5. In particular, assuming the standard value $\alpha^{'} = 0.0027$ as upper bound for the Type I error probability, the value $\alpha = 1,588x10^{-4}$ (i.e. $Z_{\alpha/2} = 3,777$) is used for designing the 17 control limits in Equation 2.5.

5.1.1 Phase I

In Phase I, a set of samples with n profiles is analyzed to assess the stability of the process and estimate the control state's parameters. During Phase I, the performance of control charts is defined in terms of the probability of deciding whether the process is in-control or not, that is, the probability of obtaining at least one statistic outside the control limits in the whole set of data observed using the set of m process samples.

It was considered 1000 simulations of Phase I control charting for different numbers of profiles, specifically $n = 25, 50, 75$, and 100, to determine the performance of the location control chart. The Run Length Average (ARL_0) was computed as the number of samples taken before at least one point exceeds the control limits, considering a global Type I Risk ($\alpha^{'}$) of 0.0027. The value of (ARL_0) was obtained through a Monte Carlo simulation representing the actual value of the (ARL_0).

To each simulation, it was assumed that for each location, the mean profile is the theoretical nominal profile (NP) obtained from Equation 4.1, and that variability is the variability means calculated in each location for sample profiles. In this case, the 17 values between 100 and 1700 RPMs of Model 3 - 220 V (Figure 26).

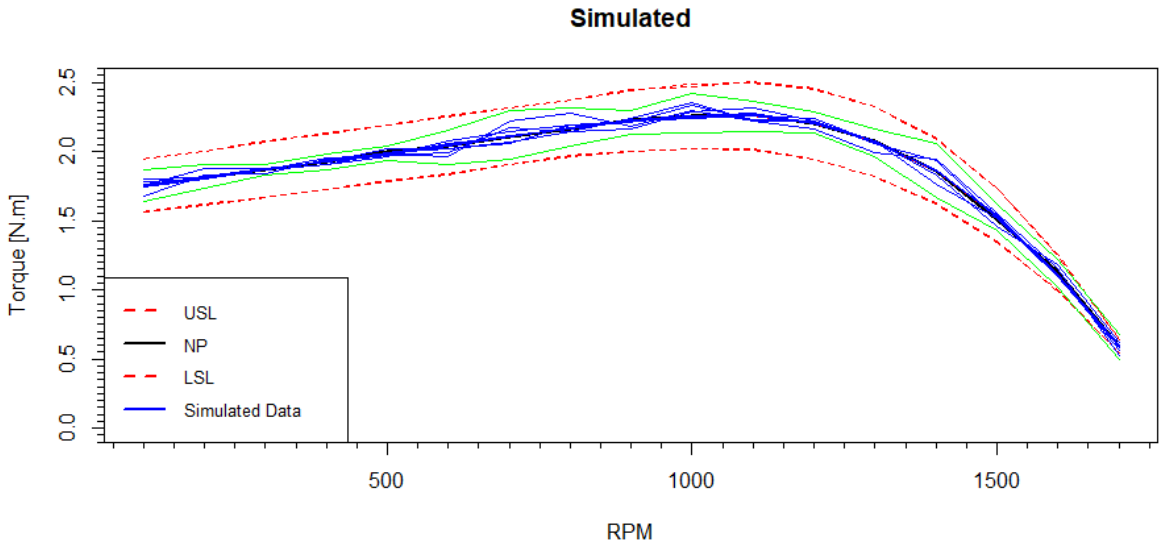


Figure 26 – Simulated profiles for electric motor Model 3, 220 volts.

Table 5 summarizes the actual Average Run Length (ARL_0) for In-Control Process for $n = 25, 50, 75$, and 100 profiles and $\alpha^{'} = 0.0027$. As these results demonstrate, the Average Run Length means is very low when compared with the nominal value 370,4, mainly when the number of profiles per sample is small. For example, for $n = 25$, ARL_0 is equal to 28,92, which corresponds to a Error Type I of 3,46%, much higher than the value of 0,27% associated with $\alpha = 0.0027$ of traditional Shewhart's control charts. As the number of profiles per sample increases, this value approaches the nominal value. Obviously, the run length is a random variable, because it is determined by the collected samples, which is random, with a geometric distribution (Figure 27).

For a given control chart, of course, the ideal situation is that its ARL_0 value is large and its ARL_1 value, when the process is out-of-control, is small. However, similar to the Type I and Type II error probabilities in the hypothesis testing context, this is not easy to achieve. To handle this issue, in the SPC literature, we usually fix the ARL_0 value

Table 5 – Phase I Simulation Results: Average Run Length for In-Control Process (ARL_0). (1000 simulations, $\alpha^{'} = 0.0027$)

n	1st Q.	Median	Mean	3rd Q.	sd
25	9,00	20,00	28,92	40,00	28,34
50	32,75	71,00	104,17	143,00	102,00
75	37,75	84,00	115,24	161,00	107,33
100	67,75	154,50	214,95	301,50	207,89

at a given level and try to make the ARL_1 value as small as possible.

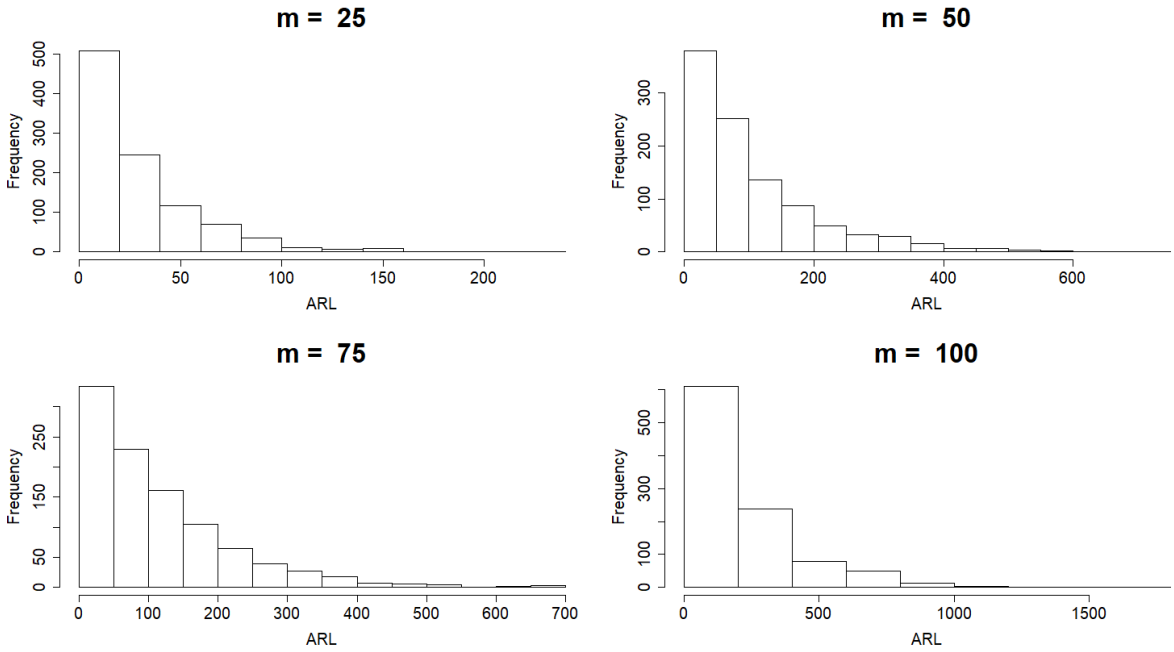


Figure 27 – Histograms of simulated distribution of ARL_0 for samples with m profiles. (1000 simulations, $\alpha^{'} = 0.0027$)

One possibility is to proceed a mathematical model's reparametrization, changing the $\alpha^{'}$ from its original value to another that reduces False Alarm Rate, increasing the ARL_0. As shown in Table 6, when the parameter $\alpha^{'}$ is changed from 0.0027 to 0.00135, the Average Run Length ARL_0 is increased, that is, the probability of Error Type I is reduced. Note that a reduced false alarm rate could seem an advantage at first. However, it means that control limits are too far from the centerline, thus, resulting in the ineffective detection of out-of-control profiles when they arise.

To manage this trade-off, it is important to understand how the proposed model behaves in the presence of special causes when the process is out of control.

5.1.2 Phase II

The objective in Phase II is to quickly detect any change in the process from its in-control state. The monitoring approaches are compared in terms of the Average Run

Table 6 – Phase I Simulation Results: Average Run Length for In-Control Process (ARL_0). (1000 simulations, $\alpha^{'} = 0.00135$)

n	1st Q.	Median	Mean	3rd Q.	sd
25	10,00	22,00	31,27	43,00	30,08
50	17,00	39,00	56,72	76,25	59,54
75	136,80	325,00	477,20	644,00	481,42
100	162,00	384,00	571,00	760,20	582,89

Length (ARL_1), where the run length is defined as the number of samples taken until an out-of-control signal is issued.

In order to evaluate performance in Phase II, occurrence of assignable causes are simulated. These out-of-controls are simulated assuming that the sample mean of the data observed at every location (Equation 2.5) has a shift of size $\delta = S\sigma$ (i.e the mean shift from μ_0 to $\mu_1 = \mu_0 + S\sigma$) at a given time point.

It was considered 1000 simulations of Phase II control charting for different numbers of profiles ($n = 25, 50, 75$, and 100) and different shifts ($\delta = 0.05, 0.10, 0.15, 0.20, 0.25$).The Average Run Length (ARL_1) was computed as the number of samples taken before at least one point exceeds the control limits, considering a global Type I Risk ($\alpha^{'}$) of 0.0027 (Figures 40-47).

Table 7 – Phase II Simulation Results: Average Run Length for Out-Of-Control Process (ARL_1). (1000 simulations, $\alpha^{'} = 0.0027$)

n	δ	1st Q.	Median	Mean	3rd Q.	sd
25	0.05	8,00	18,00	25,04	33,25	24,75
	0.10	5,00	12,00	16,42	22,00	16,78
	0.15	3,00	7,00	9,78	13,00	9,12
	0.20	2,00	4,00	6,16	8,00	5,82
	0.25	2,00	3,00	3,809	5,00	3,11
50	0.05	19,00	47,00	68,83	95,00	68,33
	0.10	9,00	20,00	28,21	38,00	27,75
	0.15	4,00	8,00	12,05	17,00	11,76
	0.20	2,00	4,00	5,70	8,00	5,22
	0.25	1,00	2,00	3,06	4,00	2,57
75	0.05	17,00	39,00	58,55	81,00	59,47
	0.10	6,0	13,00	19,10	27,00	18,82
	0.15	2,00	5,00	6,67	9,00	6,25
	0.20	1,00	2,00	2,84	4,00	2,37
	0.25	1,00	1,00	1,56	2,00	0,92
100	0.05	23,00	57,00	84,65	113,25	90,28
	0.10	6,00	14,00	19,06	26,00	18,46
	0.15	2,00	4,00	5,65	8,00	5,49
	0.20	1,00	2,00	2,00	3,00	1,59
	0.25	1,00	1,00	1,23	1,00	0,52

Table 7 summarizes the actual Average Run Length (ARL_1) for Out-Of-Control Process for $n = 25, 50, 75$ and 100 profiles, $\alpha^{'} = 0.0027$ and $\delta = 0.05, 0.10, 0.15, 0.20, 0.25$. Performance comparison is based on the ideal assumption that the in-control parameters are known.

As shown in Table 7, although the Average Run Length (ARL_0) for In-Control process using $\alpha^{'} = 0.0027$ is very low when compared with the nominal value $(370, 4)$, when the process is Out-Of-Control, the Average Run Length (ARL_1) rapidly decrease. For example, for $m = 75$, if the process has a shift of 0.25 standard deviation from de centerline, the ARL_1 mean is equal to $1, 56$, which corresponds to an Error Type II or a probability of non-detection β of $35, 9\%$. If instead $n = 100$ profiles had been used, the ARL_1 mean is equal to $1, 23$ and Type II Error would be $18, 7\%$, being such deviation detected up to the 2nd sample with 100% certainty.

From Table 8, it is possible to observe that when the parameter $\alpha^{'}$ is changed from 0.0027 to 0.00135 there are an increase of Error Type II. For example, for $n = 75$, if the process has a shift of 0.25 standard deviation from de centerline, the ARL_1 mean increase from $1, 56$ to 2.74, which corresponds to an increase of Error Type II β from $35, 9\%$ to $63, 5\%$. If instead $n = 100$ profiles had been used, the ARL_1 mean is equal to $1, 87$ that represents an increase from $18, 7\%$ to $46, 5\%$.

Table 8 – Phase II Simulation Results: Average Run Length for Out-Of-Control Process (ARL_1). (1000 simulations, $\alpha^{'} = 0.00135$)

n	δ	1st Q.	Median	Mean	3rd Q.	sd
25	0.05	8,00	18,00	26,05	36,00	25,92
	0.10	5,00	11,00	16,38	22,00	15,96
	0.15	3,00	7,00	10,45	14,00	10,66
	0.20	2,00	5,00	6,65	9,00	6,17
	0.25	1,00	3,00	3,92	5,00	3,39
50	0.05	11,00	26,00	36,99	51,25	36,16
	0.10	5,00	12,00	17,40	25,00	16,78
	0.15	2,00	5,00	7,35	10,00	6,63
	0.20	1,00	3,00	3,69	5,00	3,17
	0.25	1,00	2,00	2,18	3,00	1,53
75	0.05	69,00	145,00	205,00	279,00	199,32
	0.10	16,00	39,00	59,71	83,25	62,93
	0.15	5,00	12,00	18,14	26,00	17,40
	0.20	2,00	4,00	6,15	9,00	5,45
	0.25	1,00	2,00	2,74	4,00	2,19
100	0.05	66,00	139,00	199,80	266,00	189,53
	0.10	14,00	31,00	44,00	63,25	43,67
	0.15	4,00	9,00	12,33	17,00	11,91
	0.20	2,00	3,00	4,26	6,00	3,83
	0.25	1,00	1,00	1,87	2,00	1,27

As expected, we have the trade-off between Error Type I and Error Type II. Nevertheless, as it is possible to observe, the model's reparametrization to reduce the Type I Error, although it provokes an increase of the Type II Error, preserves the probability that the shift will be detected since for delta values greater than 0.25 up, the control chart will signal it to the 2nd sample with 100% certainty (Figure 28).

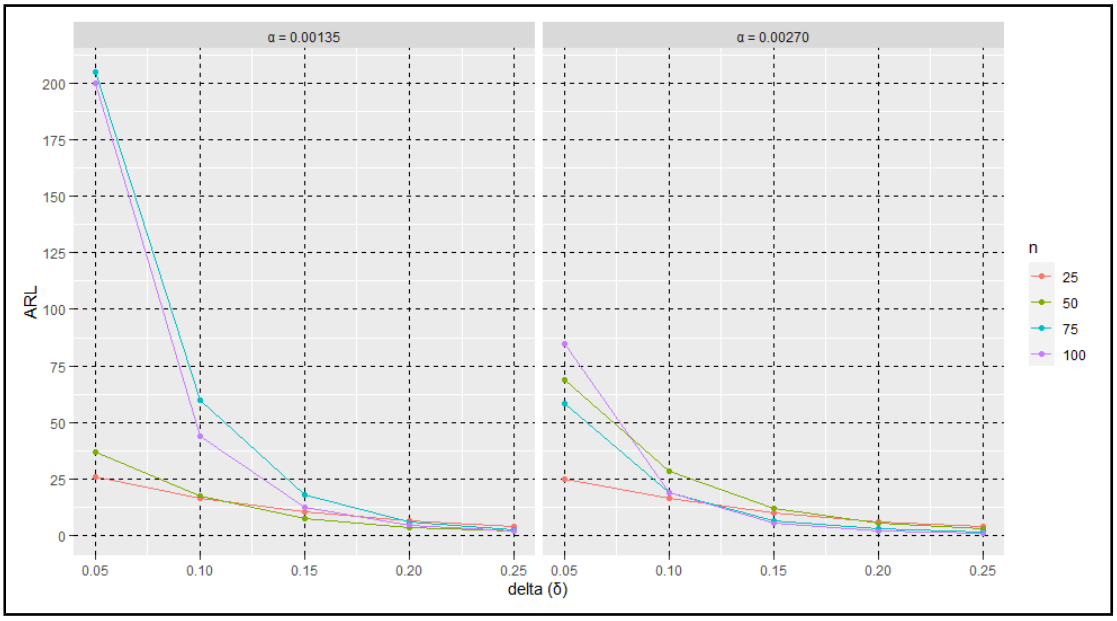


Figure 28 – Out-of-control Average Run Lenght (ARL_1)

5.2 Designing functional process capability indices

The company obtains electric motors from a supplier and incorporates them into its final product. For the appliance produced by the company to meet the requirements of operating performance and compliance with legislation, the torque curve of the electric motor must meet the design requirements. Such requirements are established on a theoretical curve that contains the nominal profile, and the upper and lower specification limits, all expressed by a curve. This experimentally obtained curve is available as an image, and there is no mathematical model to define it (Figure 19).

Initially, through the WebPlotDigitizer application, the XY pairs were obtained from the image of the theoretical curve corresponding to the functional relationship between the torque and rotations per minute of the electric motors. These data were imported into the R (R Core Team, 2020) statistical environment and the specification functional limits $USL(x)$ and $LSL(x)$ as well as the nominal profile $NP(x)$ were adjusted from the theoretical curve and calculated as:

$$USL(x) = b_{0USL} + b_{1USL}x + b_{2USL}x^2 + b_{3USL}x^3 + b_{4USL}x^4 \tag{5.1}$$

$$NP(x) = b_{0\,NP} + b_{1\,NP}x + b_{2\,NP}x^2 + b_{3\,NP}x^3 + b_{4\,NP}x^4 \tag{5.2}$$

$$LSL(x) = b_{0\,LSL} + b_{1\,LSL}x + b_{2\,LSL}x^2 + b_{3\,LSL}x^3 + b_{4\,LSL}x^4 \tag{5.3}$$

resulting in adjusted theoretical curve shown in Figure 29.

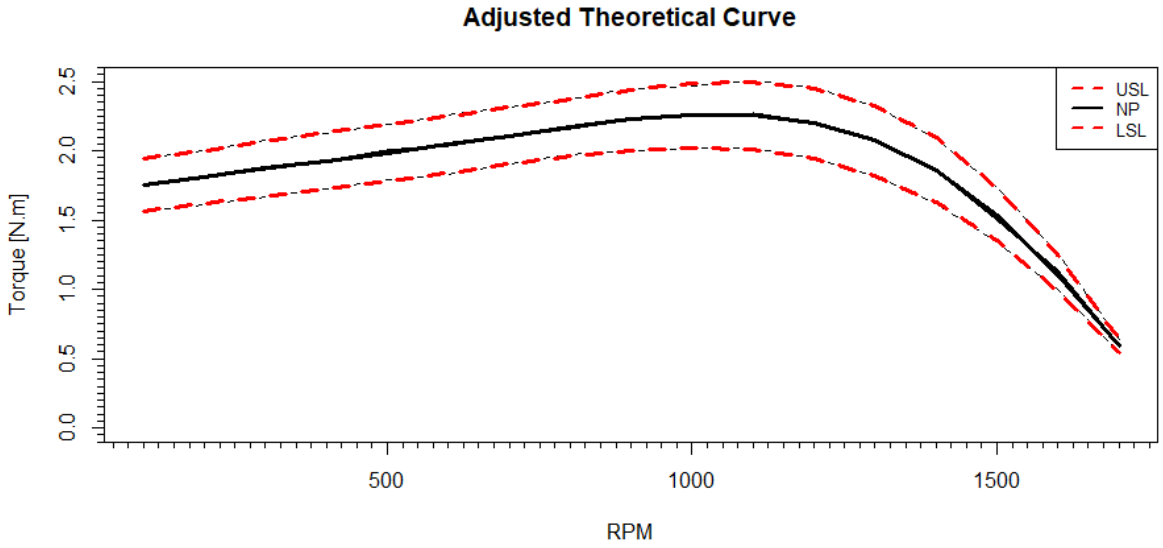


Figure 29 – Adjusted theoretical curve of electric motor Model 3, 220V

To determine the torque curves of the sampled electric motors, it was used a dynamometer developed by the company's engineering department. Data in the form "csv" file were imported into the R statistical environment and treated to filter the values at the extremities of the exploratory variable value range due to the anomalous dynamometer behavior at these points. The functional average $\hat{\mu}(x)$ was adjusted from the mean value of y_i at each level of the exploratory variable x_i. Similarly, the values of $\widehat{UNTL}(x)$ and $\widehat{LNTL}(x)$ from $\mu_i + 3\hat{\sigma}_i$ and $\mu_i - 3\hat{\sigma}_i$, respectively, were estimated at each value assumed by x_i, resulting in Equations 5.3 through 5.5 as follows:

$$\widehat{UNTL}(x) = b_{0\,UNTL} + b_{1\,UNTL}x + b_{2\,UNTL}x^2 + b_{3\,UNTL}x^3 + b_{4\,UNTL}x^4 \tag{5.4}$$

$$\widehat{\mu_y}(x) = b_{0\,\mu} + b_{1\,\mu}x + b_{2\,\mu}x^2 + b_{3\,\mu}x^3 + b_{4\,\mu}x^4 \tag{5.5}$$

$$\widehat{LNTL}(x) = b_{0\,LNTL} + b_{1\,LNTL}x + b_{2\,LNTL}x^2 + b_{3\,LNTL}x^3 + b_{4\,LNTL}x^4 \tag{5.6}$$

Table 9 – Parameters of adjusted models to functional profiles of electric motor Model 3, 220V

Functional Profile	Parameters of adjusted models				
	b_{0*}	b_{1*}	b_{2*}	b_{3*}	b_{4*}
$*USL$	1.88e+00	8.15e-04	-1.19e-06	2.15e-09	-1.17e-12
$*NP$	1.72e+00	4.12e-04	-3.48e-08	8.95e-10	-7.36e-13
$*LSL$	1.54e+00	2.17e-04	6.01e-07	7.87e-11	-4.19e-13
$*\widehat{UNTL}$	1.55e+00	8.48e-04	-7.33e-07	1.44e-09	-8.78e-13
$*\hat{\mu}$	1.50e+00	9.12e-04	-8.77e-07	1.51e-09	-8.83e-13
$*\widehat{LNTL}$	1.45e+00	9.75e-04	-1.02e-06	1.58e-09	-8.87e-13

where each parameter was obtained from a polynomial adjustment of order $k = 4$ (Table 9).

Applying Equations 5.1 to 5.6 in equations 3.3 to 3.6, we obtain for the process a $\widehat{C}_p = 3.68$ and a $\widehat{C}_{pk} = 1.88$, indicating that the production process of the sampled electric motor model (Model 3, 220V) has a high capability. The high capability presented compensates the fact that the profiles are not centered in relation to the nominal profile, which is evidenced by the significantly lower value of $\widehat{C}_{pk}$ when compared to the value of $\widehat{C}_p$ and shown in Figure 30.

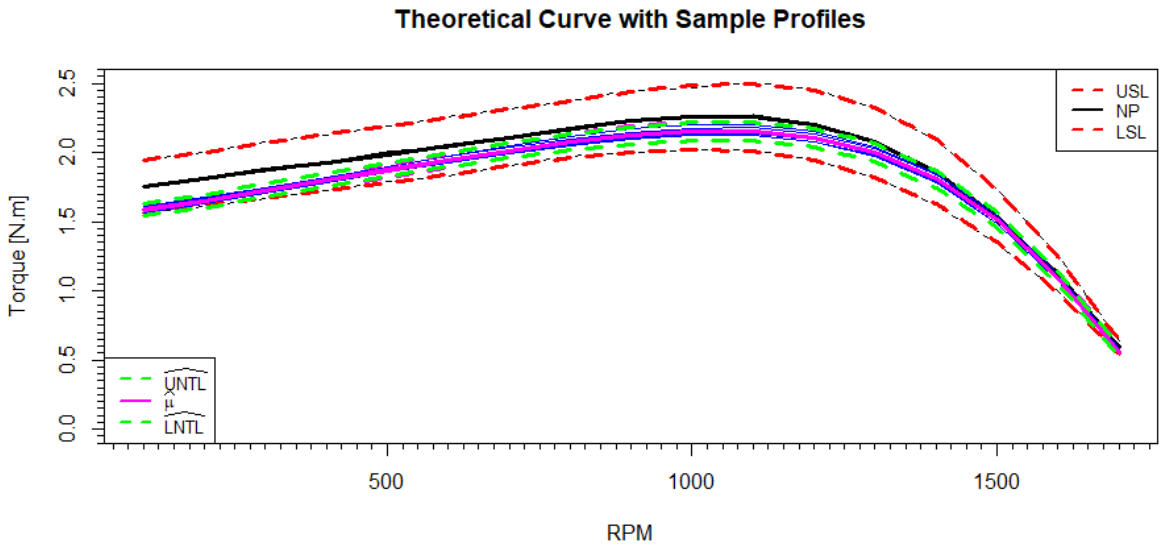


Figure 30 – Theoretical curve and natural tolerance limits for the adjusted profiles indicating high C_p

Although the process has a high potential capacity $\hat{C}_p$, it is clear from Figure 31 that $\hat{C}_{pk}$ has a very low value at the ends of the RPM range values. However, you can see that $\hat{C}_p$ and $\hat{C}_{pk}$ have a stable zone in the middle of the range. Therefore, one possibility is to divide the entire interval into t intervals and establish a weighting, assigning different

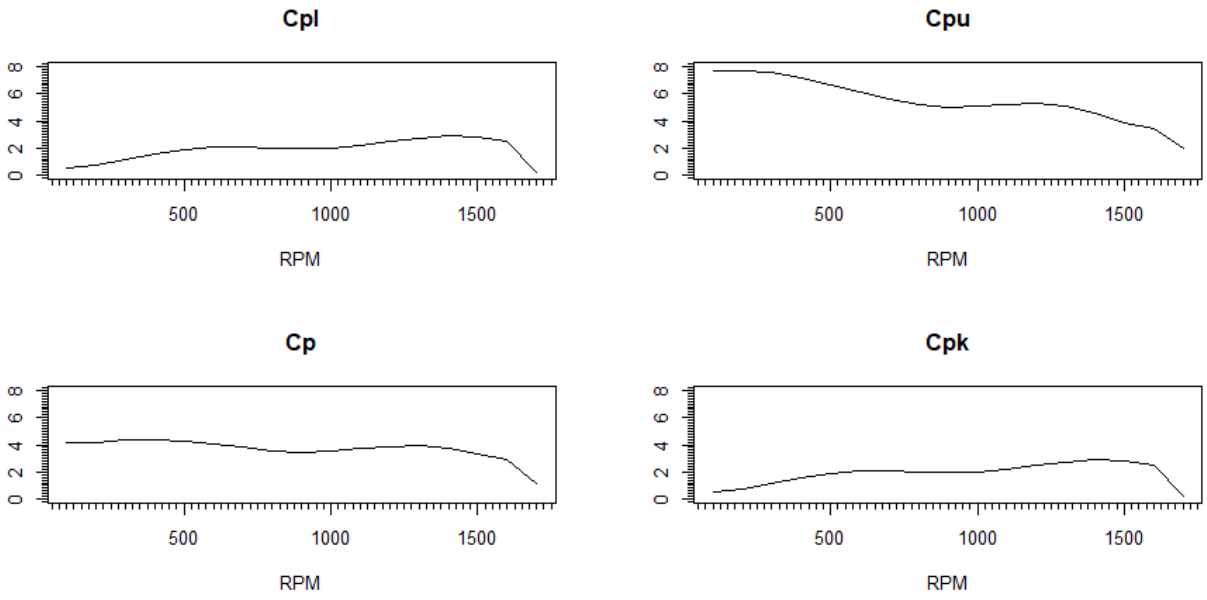


Figure 31 – Process Capability Index $\hat{C}_p$ and $\hat{C}_{pk}$

weights to each interval according to their importance from the point of view of the expected performance for the functional relationship between Torque and RPM. Such weighting would be established in order to guarantee compliance with the specifications in the most relevant range of the equipment's operational performance.

Such approach, must use engineering and subject matter knowledge, so that the weighting have both practical importance and statistical significance.

5.3 Designing functional $\bar{X}$ control chart limits with $\hat{C}_p$ and $\hat{C}_{pk}$ indices

In most situations where control charts are used, the focus is on monitoring or statistical control of the process, reducing variability, and continuously improving it. When a high level of process capability is achieved, as shown in Figure 32, it is sometimes helpful to relax the standard control chart's vigilance level. More often, when the natural dispersion of the process is much less than the dispersion allowed by the specification limits, some changes in the process level can be tolerated.

Although the process capability index is high, $\widehat{C}_p = 3.68$ and a $\widehat{C}_{pk} = 1.88$, due to the nature of the functional relationship between Torque and RPM in this study, it is possible to observe, based on Figure 31, that the process capability index ($\hat{C}_p$ and $\hat{C}_{pk}$) does not remain constant over the range of values considered for the independent variable (RPM).

This fluctuation in the values of $\hat{C}_p$ and $\hat{C}_{pk}$ should not be interpreted as a deterioration of the process, occurring basically for two reasons: a) the fact that the sample profile is not centralized in the nominal profile, mainly in the initial range of values of the independent variable, and b) because the tolerance range does not remain constant over the range of values considered for the independent variable, with a significant reduction of $\hat{C}_{pk}$ in its upper end, as shown in Figure 32.

Figure 32 shows that, despite the inevitable variability in the average value of the quality characteristic of interest, the process is still capable of meeting the specifications established in the project, especially in the range of values between 500 and 1500 RPMs where the process presents a high level of process capability, with $C_{pk} \approx 2$ or above.

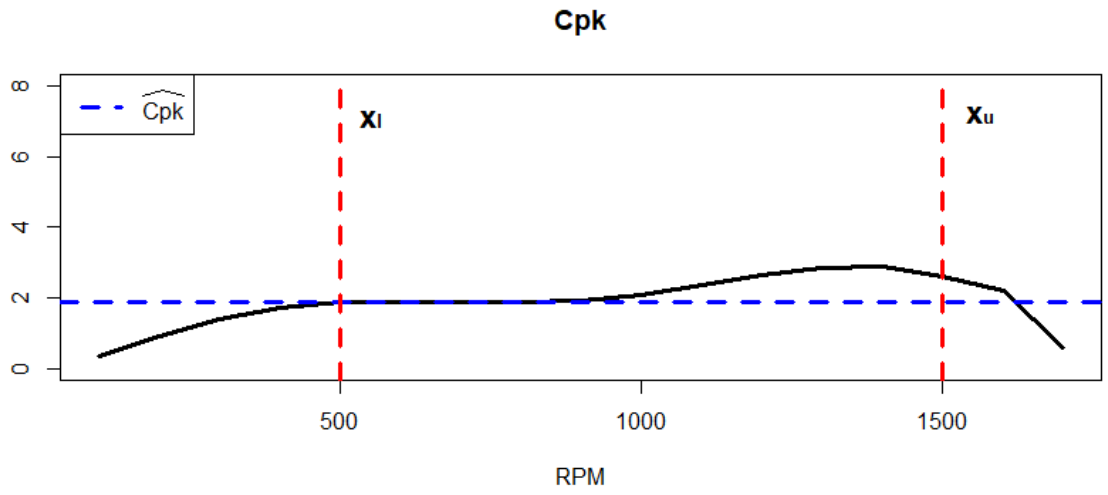


Figure 32 – $\hat{C}_{pk}$ process capability index over the electric motor's operating range.

So high level of process capability means that, in this state, the standard deviation of the average process values for each RPM is so small concerning the specification range that it is possible that the process, although not considered under control, may still be able to produce in an acceptable range (Figure 33).

Some shifts in process level must be expected and can be tolerated. These shifts usually result from an assignable cause that cannot be eliminated because of engineering or economic considerations. They often enter the system at infrequent or irregular intervals but can rarely be treated as random variance components. When shifts appear, the process may stabilize at a new level until the next event occurs. Between such disturbances, the process runs in control concerning inherent variability. If tolerance limits are satisfied, it not only may be uneconomic and wasteful of resources to control the process tightly, but it is very likely to be counterproductive to improve capability by reducing variability (WOODALL; FALTIN, 2019).

For the case under study, as shown in Figure 33, what is required is protection

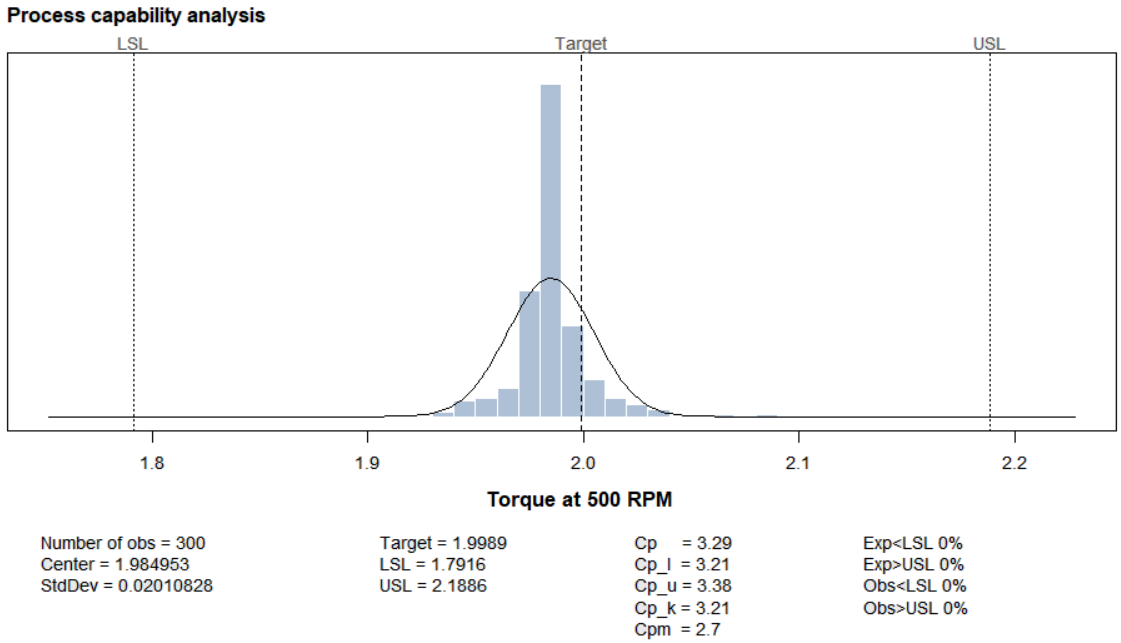


Figure 33 – Distribution of Torque at 500 RPM indicating a very capable process.

against a process shift that yields an undesirable percentage of items falling outside the specification limits.

The existence of a functional relationship between the response variable Torque and the exploratory variable RPM imposes the need to consider the calculation of the control limits (CL) over the entire range of values assumed by the exploratory variable $[x_l, x_u]$. For each RPM, there is a central zone M_0 (acceptable process) where the variable Torque is indisputably acceptable. The outer zone M_1 represents a process that is indisputably not acceptable. Between the inner and the outer zone are indifferent zones I where the process is acceptable but should be watched, and as the outer zone is approached, corrective action may be taken (Figure 18).

5.3.1 Numerical analysis of ARL

For the analysis, we consider only the one-point rule that exceeds the CL as defined in Equation 3.22. That is, if only one point for the response variable Torque correspondent to each explanatory variable from x_l to x_u, is located in the region M1, we consider the process not capable, and in fact, having an impact on the productive system with the production of items out of specification. Thus, the following hypothesis test was taken into consideration, as previously shown:

$$H_0 : C_{pk}(x) \leq C_{pk_0}(x) \,(\text{process is not capable}),$$
$$H_1 : C_{pk}(x) > C_{pk_0}(x) \,(\text{process is capable})$$

If H_0 is true, the process is not capable, and a location control chart with the traditional Shewhart approach should be used. If H_0 is false, it is necessary to decide the m and n parameters to conduct Phase I and Phase II. For this work, it was assumed that C_{pk_0}, the lowest acceptable value for the process, was 1.33 and 1.00, a target for processes 4 Sigma and 3 Sigma, respectively.

In the development of this thesis, we consider two more assumptions: 1) the shape of the profile does not change (i.e., the regression coefficients (β) of the curve determined by Equation 4.2 remain stable so that the profile moves as a block by δ; 2) the variability of the response variable Torque is known for each value of RPM exploratory variable considered.

For the scope of this work, it is considered that the process is not capable when, in fact, it impacts the production system with the consequent production of non-conforming items that do not meet the specification limits. The goal is to have control chart signals associated with practical significance, not just statistical significance, since a sample inconsistent with the assumed control model does not necessarily imply a process change of practical importance. This problem becomes more relevant as the sample size increases at each sampling point because then tiny changes in the process can be detected with high probability.

Considering, for example, the lowest acceptable value for the process (C_{pk_0}) as 1.33, we can formulate the following hypothesis:

$$H_0 : C_{pk}(x) \leq 1.33 \,(\text{process is not capable})\,,$$
$$H_1 : C_{pk}(x) > 1.33 \,(\text{process is capable})$$

However, given that 17 dependent control rules are simultaneously applied, we applied Bonferroni's rule for dependent events to attain a false alarm rate not greater than a predefined probability. If hypothesis H_0 is rejected, the process is considered capable.

In order to evaluate performance in Phase II, the occurrence of assignable causes is simulated adopting the Monte Carlos approaching shown in Section 4.3. These out-of-controls are simulated assuming that the sample mean of the data observed at each location (Equation 3.22) has a shift of $\delta = S\sigma$ (i.e the mean shift from μ_0 to $\mu_1 = \mu_0 + \delta\sigma$) at a given time point. It was assumed that for each sample, the shape of the profile remains the same (i.e., the coefficients of Equation 4.2 remains the same for all profiles simulated).

We use the equations shown in Section 3.4.2 to develop the analysis of the independent variables (number samples and the number of profiles per sample) and their effects on the ARL_1 to the control chart proposed. Introducing $\hat{C}_p$ and $\hat{C}_{pk}$ in the ARL_1 function, the number of samples taken until an out-of-control signal is issued, increased considerably

and the time needed to determine the actual number of simulations until obtaining an out of control process signal as well. It can be noted that simulating the distribution of run-lengths and their associated characteristics can be very time-consuming in the case of the in-control process because the distribution of run-lengths slopes steeply to the right, which means that there may be very long path values, which can take a long time to achieve.

For this reason, instead of determining the actual number of simulations until obtaining an out-of-control process signal, as done in Section 5.2, a very time-consuming procedure, especially when the process displacement (δ) related to the profile nominal is small, we adopted another strategy to implement the simulations, determining the proportion of profiles outside the control limits concerning a fixed number of simulations.

It was considered 100000 simulations of Phase II control charting for different numbers of profiles ($m = 10, 20$ and 40), diferent size samples ($n = 3, 5, 9$ and 15) and different displacement of the process in relation to the nominal profile ($\delta = 0.25, 0.50, 0.75, 1.00, 1.25, 1.50, 1.75, 2.00, 2.25$, and 2.50). The Average Run Length (ARL_1) was computed as $ARL_1 = 1/(1-\beta)$, considering a global Type I Risk ($\alpha^{'}$) of 0.0027. Each 100000 simulations was repeated 100 times in order to get an average to ARL_1 value, assuming that the Torque has a normal distribution with mean unknown and standard deviation known for each RPM.

Tables 10 to 12 presents ARL_1 values when torque mean is unknown and torque standard deviation is known for each RPM, and the lowest acceptable value for the process C_{pk_0} is 1.33, that is, a process where 99.38% of the electric motors produced have a torque curve that meets the specifications. For each Table, we consider a potential process capability (C_p) to produce electric motors that have a torque curve that meets the specifications equal to 2.00, 1.75, 1.50 and 1.33. The curves relating torque versus RPM are presented in the Figures 34 to 36, respectively.

Closer inspection of the Tables 10 to 12 and Figures 34 to 36 shows that when C_{pk} is higher than or equal to C_{pk_0}, the ARL_1 is high and, consequently, indicates a low probability of process interruption in search for special causes causing process instability, since the process meets the limits concerning the specifications of the product. These results corroborate the fact that the proposed control chart under analysis does not seek to monitor or control the stability of the process but to prevent the process from producing items outside the specification limits.

These tables are quite revealing in several ways. First, it is possible to observe that when the potential capacity of the process (C_p) is high, there is a small probability of a signal to be observed. For example, with $m = 20$, $n = 3$, $C_{pk_0} = 1.33$, and $C_p = 2.00$, if the process average had a displacement δ of 1.00 caused by some factor, the probability of

this average displacement is detected is only 0.169% (Table 11). In this condition, despite the displacement being one standard deviation, the process still presents a capability C_{pk} of 1.67, therefore above the minimum value tolerated for the process. This is good since, from the standpoint of practical significance, we want a signal to be given only when the process presents a high risk of producing defective items. When, however, the process experiences a more significant deviation, causing its C_{pk} to be reduced to the minimum value of acceptable capability for the process ($C_{pk_0} = 1.33$), a condition that begin to present a high risk of producing non-conforming items, the probability that such deviation be detected becomes equal to 3.125%. In this case, the decision to intervene in the process must use engineering and subject matter knowledge so that the weighting has both practical importance and statistical significance, evaluating whether attributable causes may be inherent to the process and if their removal is impossible, impractical, or expensive.

Second, there is a (rather) interesting outcome related to ARL_1 which has the following behavior: when C_{pk} is higher than or equal to C_{pk_0}, if n increases, the ARL_1 increases too. This result is somewhat counterintuitive because for the traditional Shewhart control chart, if you increases n, the power of a statistical test increases as well, resulting in a greater probability of correctly rejecting H_0. However, for C_{pk} less than C_{pk_0} there is an inversely proportional relationship resulting that as n increases the ARL_1 increases.

Third, when C_{pk_0} is equal to C_p, the behavior of the location control chart using expanded limits based on $\hat{C}_p$ and $\hat{C}_{pk}$, resembles the location control chart with the traditional Shewhart approach described in Section 5.1.

Overall, these results indicate that the performance of the location control chart with expanded limits based on $\hat{C}_p$ and $\hat{C}_{pk}$ indices presents a better performance than the traditional location control chart based only on Shewhart approach for situations where the difference between $\hat{C}_p$ and $\hat{C}_{pk}$ is large. However, considering the results of Tables 10 to 12, it can be observed that when the difference between $\hat{C}_p$ and $\hat{C}_{pk}$ is small, the proposed control chart presents a lower performance.

The results shown in Tables 13 to 15 and Figures 37 to 39 corroborate the findings of Tables 10 to 12 and Figures 34 to 36, indicating that when a high level of process capability is achieved, it can be helpful relax the level of control for variability and special causes reducing costs associated with unnecessary process stop. On the other hand, the practitioner can return to traditional control when the process presents less than an acceptable capability.

In this case, the difference between $\hat{C}_p$ and $\hat{C}_{pk}$ is greater than in Tables 13 to 15. For example, if $m = 20$, $n = 3$, $C_{pk_0} = 1.00$, and $C_p = 2.00$, when the process average had a displacement δ of 1.00 caused by some factor, the probability of this average displacement

is detected is only 0.0125% (Table 14). In this condition, despite the displacement being one standard deviation, the process exhibit a capability C_{pk} of 1.67, well above the lowest acceptable capability for the process (1.00) that corresponds to a 3 Sigma process, a process with 0.27% probability to produce non-conforming items.

Table 10 – Simulated ARL_1 using expanded limits for $m = 10$ and $C_{pk_0} = 1.33$.

C_p	C_{pk0}	n		δ									
				0.25	**0.50**	**0.75**	**1.00**	**1.25**	**1.50**	**1.75**	**2.00**	**2.25**	**2.50**
			C_{pk}	**1.92**	**1.83**	**1.75**	**1.67**	**1.58**	**1.50**	**1.42**	**1.33**	**1.25**	**1.17**
		3		3616.3	2029.9	1043.2	518.3	258.3	128.4	64.4	32.8	17.0	8.9
2.00	**1.33**	5		$> 10^5$	14566.2	4269.4	1365.3	453.1	153.8	54.1	19.8	7.8	3.3
		9		$> 10^5$	$> 10^5$	$> 10^5$	3756.4	629.5	118.6	25.1	6.3	2.1	1.2
		15		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	2243.5	194.9	23.1	4.2	1.4
			C_{pk}	**1.67**	**1.58**	**1.50**	**1.42**	**1.33**	**1.25**	**1.17**	**1.08**	**1.00**	**0.92**
		3		516.0	301.6	157.3	80.0	40.7	21.0	11.0	5.9	3.3	2.0
1.75	**1.33**	5		1848.2	662.0	227.5	79.0	28.0	10.8	4.5	2.1	1.3	1.0
		9		12032.8	1884.1	330.3	63.8	14.0	3.8	1.5	1.0	1.0	1.0
		15		$> 10^5$	$> 10^5$	7577.4	559.0	54.4	7.6	1.8	1.0	1.0	1.0
			C_{pk}	**1.42**	**1.33**	**1.25**	**1.17**	**1.08**	**1.00**	**0.92**	**0.83**	**0.75**	**0.67**
		3		75.3	45.0	24.2	12.8	6.9	3.9	2.3	1.5	1.1	1.0
1.50	**1.33**	5		99.2	38.7	14.6	5.9	2.7	1.4	1.1	1.0	1.0	1.0
		9		149.8	30.5	7.2	2.3	1.1	1.0	1.0	1.0	1.0	1.0
		15		1387.1	116.2	12.8	2.3	1.1	1.0	1.0	1.0	1.0	1.0
			C_{pk}	**1.25**	**1.16**	**1.08**	**1.00**	**0.91**	**0.83**	**0.75**	**0.66**	**0.58**	**0.50**
		3		21.0	12.8	7.1	4.0	2.4	1.5	1.1	1.0	1.0	1.0
1.33	**1.33**	5		15.2	6.6	2.9	1.6	1.1	1.0	1.0	1.0	1.0	1.0
		9		9.8	2.8	1.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		15		14.4	2.4	1.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0

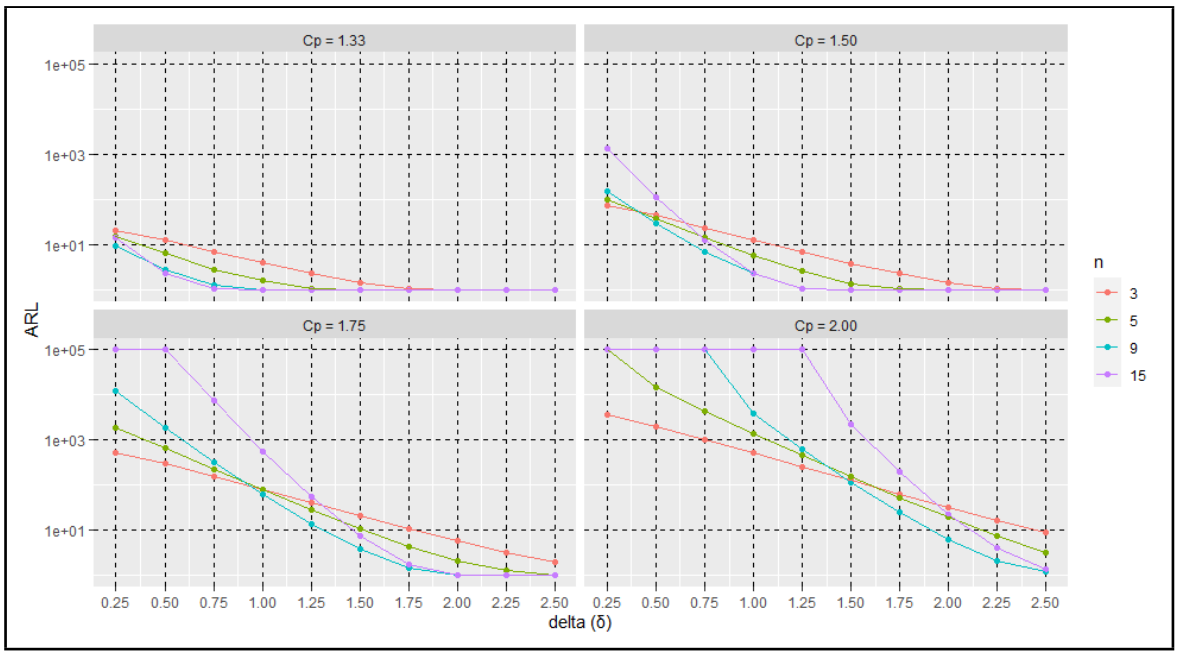


Figure 34 – Curves ARL_1 versus δ for $m = 10$ and $C_{pk_0} = 1.33$.

Table 11 – Simulated ARL_1 using expanded limits for $m = 20$ and $C_{pk_0} = 1.33$.

C_p	C_{pk0}	n		δ									
				0.25	**0.50**	**0.75**	**1.00**	**1.25**	**1.50**	**1.75**	**2.00**	**2.25**	**2.50**
2.00	**1.33**		C_{pk}	**1.92**	**1.83**	**1.75**	**1.67**	**1.58**	**1.50**	**1.42**	**1.33**	1.25	**1.17**
		3		4300.9	2361.2	1190.0	590.7	290.3	144.7	72.6	36.9	19.1	10.0
		5		$> 10^5$	13330.2	3883.9	1246.7	427.8	145.8	51.6	18.9	7.5	3.3
		9		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	4375.8	719.2	127.7	25.4	6.0
		15		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	9411.9	617.7	52.5	6.3	1.5
1.75	**1.33**		C_{pk}	**1.67**	**1.58**	**1.50**	**1.42**	**1.33**	**1.25**	**1.17**	**1.08**	**1.00**	**0.92**
		3		575.5	320.7	164.7	84.0	42.7	21.9	11.5	6.2	3.5	2.1
		5		1827.0	645.0	216.9	76.4	27.5	10.5	4.4	2.1	1.2	1.0
		9		$> 10^5$	$> 10^5$	4907.3	787.2	138.9	27.3	6.4	2.0	1.1	1.0
		15		$> 10^5$	$> 10^5$	17321.0	974.9	78.0	8.6	1.7	1.0	1.0	1.0
1.50	**1.33**		C_{pk}	**1.42**	**1.33**	**1.25**	**1.17**	**1.08**	**1.00**	**0.92**	**0.83**	**0.75**	**0.67**
		3		77.8	45.0	23.9	12.6	6.8	3.9	2.3	1.5	1.1	1.0
		5		99.1	37.7	14.2	5.7	2.6	1.4	1.0	1.0	1.0	1.0
		9		797.7	142.8	27.7	6.4	2.0	1.1	1.0	1.0	1.0	1.0
		15		1439.1	110.3	11.2	2.0	1.2	1.0	1.0	1.0	1.0	1.0
1.33	**1.33**		C_{pk}	**1.25**	**1.16**	**1.08**	**1.00**	**0.91**	**0.83**	**0.75**	**0.66**	**0.58**	
		3		20.8	12.4	6.9	3.9	2.3	1.5	1.1	1.0	1.0	1.0
		5		15.1	6.4	2.9	1.5	1.1	1.0	1.0	1.0	1.0	1.0
		9		24.6	5.9	1.9	1.1	1.0	1.0	1.0	1.0	1.0	1.0
		15		12.0	2.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

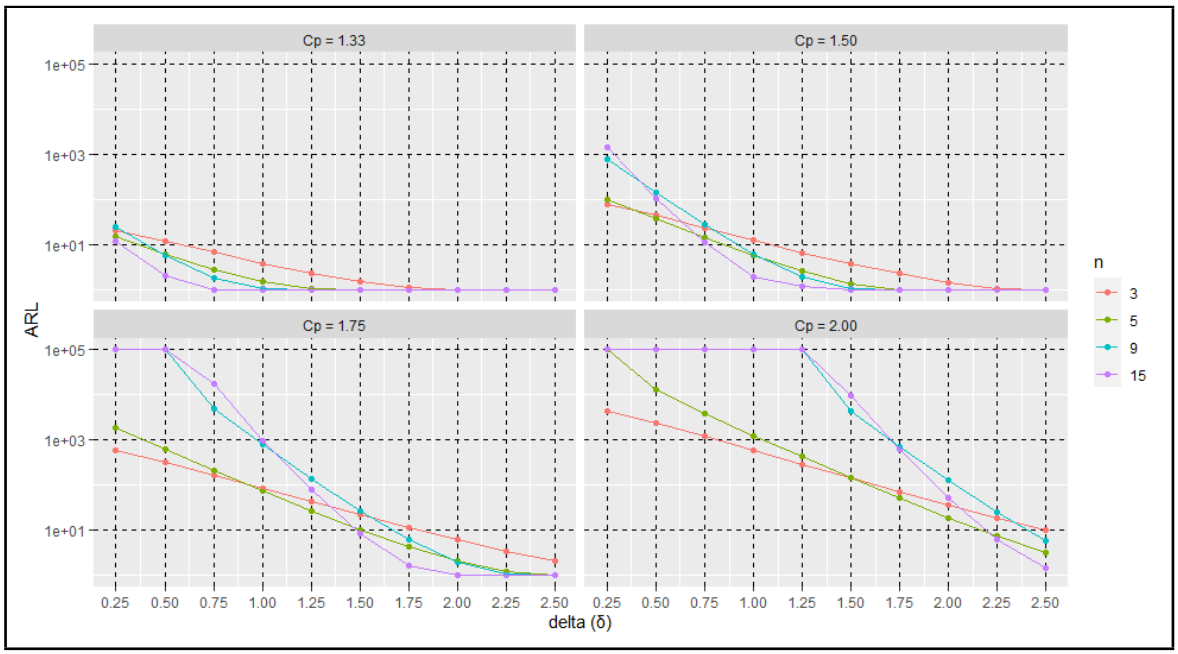


Figure 35 – Curves ARL_1 versus δ for $m = 20$ and $C_{pk_0} = 1.33$.

Table 12 – Simulated ARL_1 using expanded limits for $m = 40$ and $C_{pk_0} = 1.33$.

C_p	C_{pk0}	n		δ									
				0.25	**0.50**	**0.75**	**1.00**	**1.25**	**1.50**	**1.75**	**2.00**	**2.25**	**2.50**
2.00	**1.33**		C_{pk}	**1.92**	**1.83**	**1.75**	**1.67**	**1.58**	**1.50**	**1.42**	**1.33**	1.25	**1.17**
		3		3942.0	2224.5	1112.3	551.1	273.1	136.1	68.6	35.1	18.2	9.7
		5		$> 10^5$	$> 10^5$	$> 10^5$	7594.5	2453.5	812.5	268.7	91.9	32.9	12.4
		9		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	13400.8	1887.4	315.1	57.9	12.3	3.3
		15		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	9342.0	607.7	51.0	6.1	1.5
1.75	**1.33**		C_{pk}	**1.67**	**1.58**	**1.50**	**1.42**	**1.33**	**1.25**	**1.17**	**1.08**	**1.00**	**0.92**
		3		503.2	290.3	150.4	76.9	39.3	20.3	10.8	5.9	3.4	2.1
		5		6831.2	2385.3	774.9	260.3	89.8	32.3	12.2	5.0	2.3	1.3
		9		$> 10^5$	17635.0	2382.5	393.7	71.5	14.8	3.8	1.5	1.0	1.0
		15		$> 10^5$	$> 10^5$	18274.6	982.0	78.1	8.6	1.7	1.0	1.0	1.0
1.50	**1.33**		C_{pk}	**1.42**	**1.33**	**1.25**	**1.17**	**1.08**	**1.00**	**0.92**	**0.83**	**0.75**	**0.67**
		3		68.1	40.4	21.7	11.6	6.4	3.6	2.2	1.5	1.1	1.0
		5		223.4	85.1	30.8	11.7	4.8	2.3	1.3	1.0	1.0	1.0
		9		469.6	85.9	17.4	4.3	1.6	1.0	1.0	1.0	1.0	1.0
		15		1533.5	115.8	11.7	2.1	1.0	1.0	1.0	1.0	1.0	1.0
1.33	**1.33**		C_{pk}	**1.25**	**1.16**	**1.08**	**1.00**	**0.91**	**0.83**	**0.75**	**0.66**	**0.58**	**0.50**
		3		18.6	11.3	6.4	3.7	2.2	1.5	1.1	1.0	1.0	1.0
		5		25.7	10.8	4.5	2.2	1.3	1.0	1.0	1.0	1.0	1.0
		9		17.5	4.4	1.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		15		12.8	2.2	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

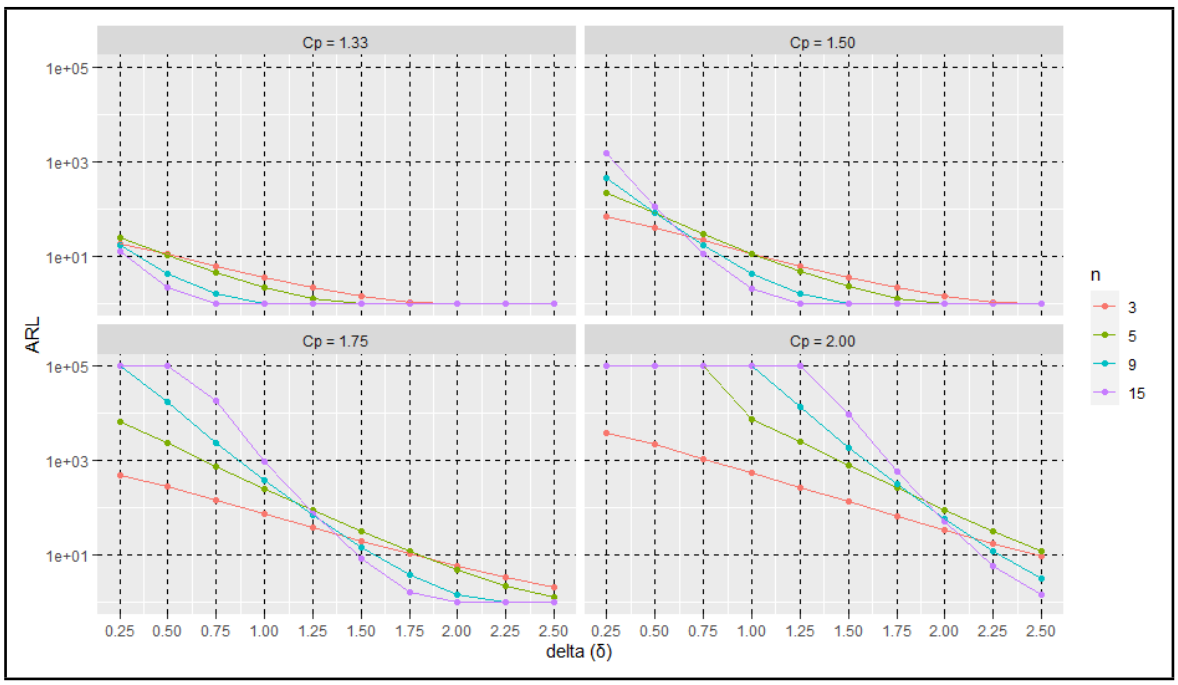


Figure 36 – Curves ARL_1 versus δ for $m = 40$ and $C_{pk_0} = 1.33$.

Table 13 – Simulated ARL_1 using expanded limits for $m = 10$ and $C_{pk_0} = 1.00$.

C_p	C_{pk0}	n		δ									
				0.25	**0.50**	**0.75**	**1.00**	**1.25**	**1.50**	**1.75**	**2.00**	**2.25**	**2.50**
2.00	**1.00**		C_{pk}	**1.92**	**1.83**	**1.75**	**1.67**	**1.58**	**1.50**	**1.42**	**1.33**	**1.25**	**1.17**
		3		$> 10^5$	$> 10^5$	13996.2	6519.7	3079.9	1468.5	715.6	350.8	173.0	87.1
		5		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	7914.4	2177.7	697.5	234.9	82.3
		9		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	4630.4	735.0	138.5	29.3
		15		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	5580.2	432.5
1.75	**1.00**		C_{pk}	**1.67**	**1.58**	**1.50**	**1.42**	**1.33**	**1.25**	**1.17**	**1.08**	**1.00**	**0.92**
		3		6870.3	3718.0	1907.8	952.2	467.7	231.3	115.2	58.1	29.6	15.3
		5		$> 10^5$	38144.1	11429.5	3547.0	1136.4	378.5	129.5	45.4	16.9	6.7
		9		$> 10^5$	$> 10^5$	$> 10^5$	16527.7	2372.7	415.8	79.9	17.6	4.7	1.8
		15		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	20508.0	1393.8	126.8	16.3	3.3
1.50	**1.00**		C_{pk}	**1.42**	**1.33**	**1.25**	**1.17**	**1.08**	**1.00**	**0.92**	**0.83**	**0.75**	**0.67**
		3		973.2	555.3	287.8	144.9	73.3	37.2	19.2	10.1	5.5	3.1
		5		4703.8	1721.0	566.3	193.2	67.2	24.4	9.4	4.0	1.9	1.2
		9		$> 10^5$	8223.4	1240.2	222.6	44.4	10.3	3.0	1.3	1.0	1.0
		15		$> 10^5$	$> 10^5$	$> 10^5$	4917.0	359.1	38.0	5.9	1.6	1.0	1.0
1.33	**1.00**		C_{pk}	**1.25**	**1.16**	**1.08**	**1.00**	**0.91**	**0.83**	**0.75**	**0.66**	**0.58**	**0.50**
		3		257.5	151.5	79.6	41.0	21.2	11.1	6.0	3.4	2.0	1.3
		5		623.7	234.5	52.1	29.7	11.2	4.6	2.2	1.3	1.0	1.0
		9		2299.7	409.2	77.4	16.7	4.4	1.6	1.1	1.0	1.0	1.0
		15		$> 10^5$	9322.0	655.8	61.1	8.1	1.8	1.0	1.0	1.0	1.0

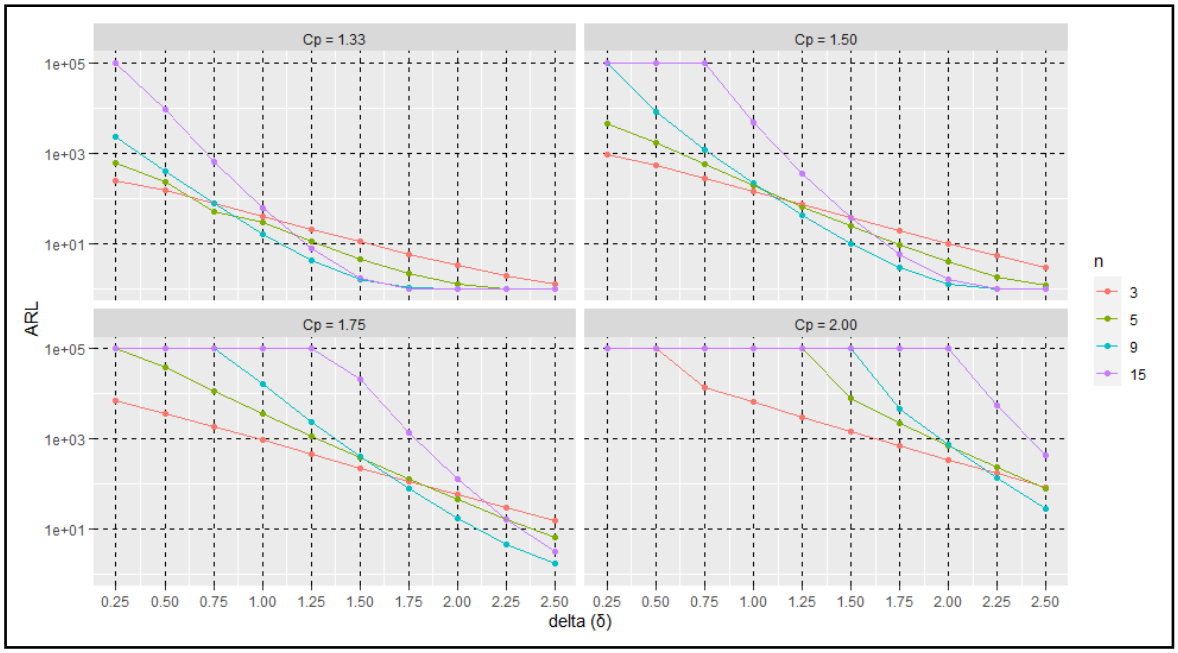


Figure 37 – Curves ARL_1 versus δ for $m = 10$ and $C_{pk_0} = 1.00$.

Table 14 – Simulated ARL_1 using expanded limits for $m = 20$ and $C_{pk_0} = 1.00$.

C_p	C_{pk0}	n		δ									
				0.25	**0.50**	**0.75**	**1.00**	**1.25**	**1.50**	**1.75**	**2.00**	**2.25**	**2.50**
2.00	**1.00**		C_{pk}	**1.92**	**1.83**	**1.75**	**1.67**	**1.58**	**1.50**	**1.42**	**1.33**	1.25	**1.17**
		3		$> 10^5$	$> 10^5$	$> 10^5$	7979.3	3864.9	1848.3	905.2	447.5	219.6	110.1
		5		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	22302.8	6330.5	1967.6	637.6	218.2	76.6
		9		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	3318.8
		15		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	3918.8
1.75	**1.00**		C_{pk}	**1.67**	**1.58**	**1.50**	**1.42**	**1.33**	**1.25**	**1.17**	**1.08**	**1.00**	**0.92**
		3		8675.8	4637.3	2264.1	1094.1	542.1	266.2	132.8	67.0	34.1	17.6
		5		$> 10^5$	$> 10^5$	10055.8	3059.6	1047.7	352.5	121.8	43.0	16.1	6.4
		9		$> 10^5$	$> 10^5$	4907.3	787.2	138.9	27.3	6.4	2.0	1.1	1.0
		15		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	7575.5	476.1	42.0	5.3
1.50	**1.00**		C_{pk}	**1.42**	**1.33**	**1.25**	**1.17**	**1.08**	**1.00**	**0.92**	**0.83**	**0.75**	**0.67**
		3		1089.7	603.4	309.9	156.3	78.5	39.8	20.5	10.8	5.9	3.3
		5		4957.4	1642.6	539.3	185.6	64.9	23.6	9.2	3.9	1.9	1.2
		9		$> 10^5$	$> 10^5$	$> 10^5$	4443.2	725.1	127.5	25.1	5.9	1.9	1.1
		15		$> 10^5$	$> 10^5$	$> 10^5$	12936.9	764.7	62.8	7.2	1.6	1.0	1.0
1.33	**1.00**		C_{pk}	**1.25**	**1.16**	**1.08**	**1.00**	**0.91**	**0.83**	**0.75**	**0.66**	**0.58**	**0.50**
		3		280.1	157.6	81.5	41.9	21.7	11.4	6.2	3.5	2.1	1.4
		5		631.9	227.7	79.5	28.8	11.0	4.5	2.1	1.3	1.0	1.0
		9		$> 10^5$	4295.4	699.0	123.8	24.4	5.8	1.9	1.1	1.0	1.0
		15		$> 10^5$	17345.8	926.9	73.6	8.1	1.7	1.0	1.0	1.0	1.0

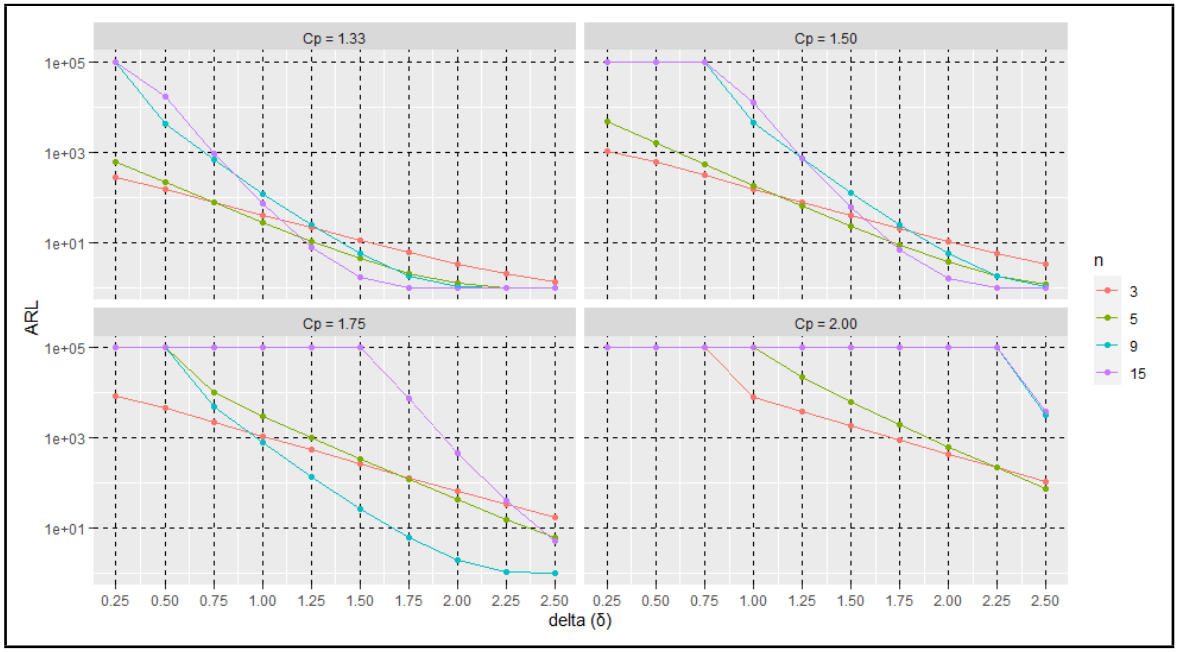


Figure 38 – Curves ARL_1 versus δ for $m = 20$ and $C_{pk_0} = 1.00$.

Table 15 – Simulated ARL_1 using expanded limits for $m = 40$ and $C_{pk_0} = 1.00$.

C_p	C_{pk0}	**n**		δ									
				0.25	**0.50**	**0.75**	**1.00**	**1.25**	**1.50**	**1.75**	**2.00**	**2.25**	**2.50**
2.00	**1.00**		C_{pk}	**1.92**	**1.83**	**1.75**	**1.67**	**1.58**	**1.50**	**1.42**	**1.33**	1.25	**1.17**
		3		$> 10^5$	$> 10^5$	17929.5	7916.0	3714.8	1809.5	898.2	441.9	217.1	108.7
		5		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	27217.9	7263.4	2295.4	755.0
		9		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	8075.1	1236.4
		15		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	3739.1
1.75	**1.00**		C_{pk}	**1.67**	**1.58**	**1.50**	**1.42**	**1.33**	**1.25**	**1.17**	**1.08**	**1.00**	**0.92**
		3		7713.2	4298.1	2119.8	1038.9	512.5	253.6	126.8	64.2	32.7	17.0
		5		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	7857.4	2429.8	773.3	256.6	88.2	31.6
		9		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	10591.3	1612.7	272.8	50.5	10.9
		15		$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	$> 10^5$	7331.4	461.3	40.4	5.1
1.50	**1.00**		C_{pk}	**1.42**	**1.33**	**1.25**	**1.17**	**1.08**	**1.00**	**0.92**	**0.83**	**0.75**	**0.67**
		3		964.0	551.7	282.6	143.9	72.6	37.0	19.2	10.2	5.6	3.2
		5		23342.1	7090.1	2271.6	748.9	251.1	87.2	31.3	11.8	4.9	2.3
		9		$> 10^5$	$> 10^5$	$> 10^5$	2060.9	341.6	62.8	13.1	3.5	1.4	1.0
		15		$> 10^5$	$> 10^5$	$> 10^5$	12443.5	756.4	62.2	7.2	1.6	1.0	1.0
1.33	**1.00**		C_{pk}	**1.25**	**1.16**	**1.08**	**1.00**	**0.91**	**0.83**	**0.75**	**0.66**	**0.58**	**0.50**
		3		244.1	141.5	74.1	38.2	19.9	10.6	5.8	3.3	2.0	1.4
		5		1957.1	697.6	236.1	82.0	29.3	11.2	4.6	2.2	1.3	1.0
		9		15867.7	2222.0	367.6	67.2	14.0	3.6	1.4	1.0	1.0	1.0
		15		$> 10^5$	16556.5	939.8	74.5	8.2	1.7	1.0	1.0	1.0	1.0

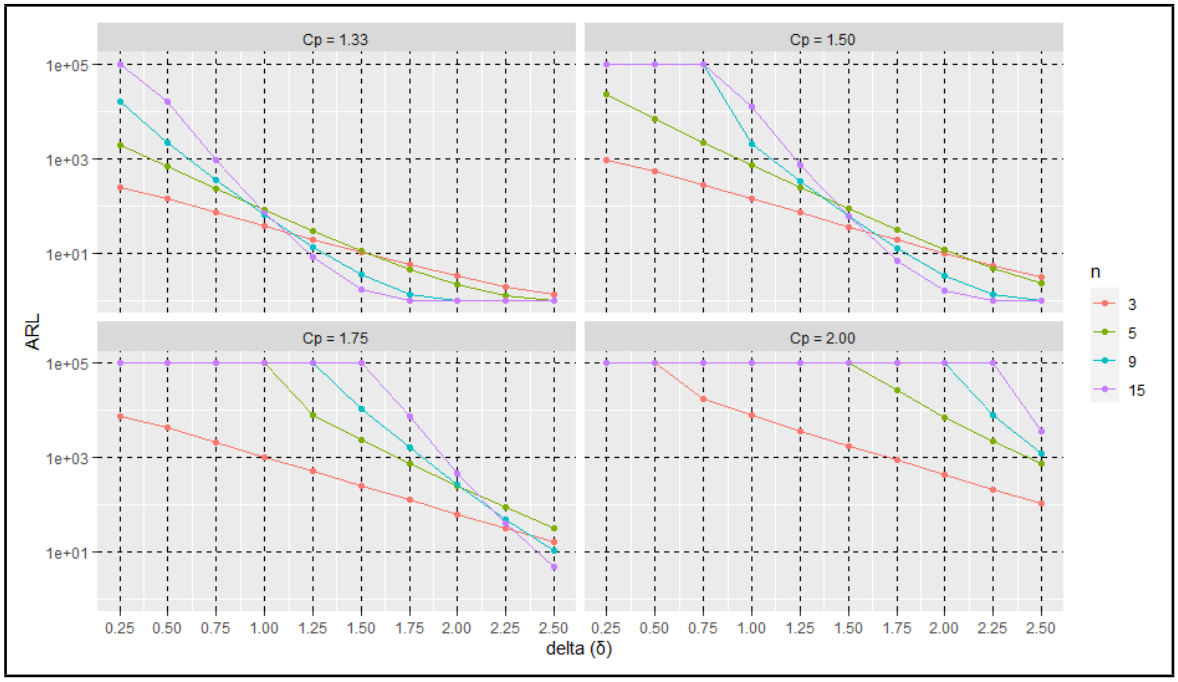


Figure 39 – Curves ARL_1 versus δ for $m = 40$ and $C_{pk_0} = 1.00$.

As mentioned in the literature review, when we have a situation where some slack in the process is generally allowed because the process is highly capable and where the specification limits are very wide with respect to process variation, attention should be directed to other issues more urgent, or in Woodall e Faltin (2019) words, of "greater practical significance". The selection of control limits, with their associated ARLs, reflects a judgment regarding the consequences of two types of errors that can be made: the false alarm and its consequent compromise of the credibility of the control tool, and on the other hand, the possibility of allowing a significant change in the process not to be detected in time.

In the daily production routines of companies and organizations, the decision to intervene in the process must use engineering knowledge and personal subject experience so that the weighting has both practical importance and statistical significance. The "owner of the process" would only want to react when a process change is significant enough to justify a reaction in practice. It is unrealistic to assume that any deviation, however small, should be detected as quickly as possible.

The use of a control region associated with the process capability indices $\hat{C}_p$ and $\hat{C}_{pk_0}$, presented by this thesis, so that only changes of practical importance are quickly detected, is analogous to hypothesis testing when the null hypothesis establishes an acceptable minimum value for the process' operation. The location control chart, proposed as an alternative to profile monitoring, explores the assumptions under Shewhart's systems thinking, expanding its traditional limits in order to "desensitized" it in detecting small shifts that are not necessarily of practical significance. An important aspect of being highlighted is its simplicity and applicability, keeping all the data information observed in each location where the profile needs to be evaluated.

6 Conclusion

Through their statistical limits, statistical control charts are essential tools in process monitoring and control, helping to interpret its stability. The capability indices, in turn, deal with the performance of these processes related to their capability of meeting the specifications established in the project. Many researches have been developed to design statistical methods able to detect small process deviations when the process is under the action of special causes, or as it is commonly said, out of statistical control.

This sensitivity is an advantage for processes needing precise control but is a definite drawback where some slack in the process is permissible. Although, small attributable causes may be inherent to the process, their removal may become impossible, impractical, or expensive. In reviewing the literature, it is possible to observe that few publications deal with the development of control charts that incorporate aspects of economic and practical significance to the control charts.

The literature review revealed that there has been a growing interest in monitoring and controlling the quality characteristics of products and processes characterized by a functional relationship between a response variable and one or more exploratory variables. This interest has been driven by the increasing availability of data made possible by the intensive use of sensors and data collectors, typical of Industry 4.0. This, combined with the fact that the shape of manufactured products is often an important aspect of quality and for which traditional control methods present unsatisfactory performance or may not even be applicable, have done profile monitoring an up-and-coming area of research. Among others approaches to profile monitoring, the location control chart presents as an alternative standing out for its applicability and simplicity, keeping all the data information observed in each location where the profile needs to be evaluated.

As with univariate statistical control charts, there is a need for desensitization of the location control charts in order to ensure that signs of an out-of-control process that do not present a high risk to produce non-conforming items, be ignored. Using process capability indices $\hat{C}_p$ and $\hat{C}_{pk}$, this thesis aimed to incorporate Shewhart-type location control chart models considering practical and economic significance when a profile defines the quality characteristic. The present thesis focused on the question: **Is a Shewhart-type functional control chart based on capability índices $\hat{C}_p$ and $\hat{C}_{pk}$ an acceptable control way to profile monitoring?**

This thesis addressed the research problem in three aspects: 1) modeling the functional relationship between Torque and RPM for electric motors based on data

given by a company; 2) modeling functional capability indices based on Torque x RPM relationship modeled before, and finally; 3) modeling expanded control limits based on capability indices to a location control chart. An interesting aspect in the development of the work was the need to treat the data provided by the company in such a way that they could be useful in the modeling and analysis stages. In a typical environment of the emerging Industry 4.0, characterized by large volumes of data, new ways to produce, organize, and analyze data need to be thought, since management decisions are only as good as the data on which they are based.

Based on the models available in the literature that address the monitoring for univariate statistics, this thesis has developed an acceptance control chart that incorporates practical and economic significance to the monitoring of profiles. Using the proposed approach by Boeing (1998) and Oprime et al. (2019), new functional control limits expanded was determined by including two new parameters, $\hat{C}_p$ and $\hat{C}_{pk}$, which appears to be the first study with a similar approach, constituting a theoretical gap in the statistical profile monitoring that was filled.

The built statistical model assumes that the average of the response variable is unknown and that its standard deviation is known at each value assumed by the exploratory variable, Other assumption assumed was that the profile shape keeps stable, that is, the regression coefficients of the profile remain the same. With the use of the R language, the performance analysis of the control charts was performed for the assumed assumptions.

By comparing the results obtained for the location control chart using the traditional approach and the location control chart with expanded limits based on $\hat{C}_p$ and $\hat{C}_{pk}$ proposed in this thesis, it was possible to observe that the expanded control limits have performed better. There was a significant reduction in the likelihood of signaling that the process is out-of-control when this is irrelevant or, at least, does not present practical or economic significance. Since the location control chart developed does not attempt to control the stability of the process but prevents the process from producing items outside the specification limits, the results demonstrate that, from the economic point of view, it constitutes a more interesting alternative to the traditional approach. In this way, the answer to the research question proves to be positive: **Shewhart-type functional location control chart based on capability indices $\hat{C}_p$ and $\hat{C}_{pk}$ constitute an acceptable control way to profile monitoring and presents an upper performance compared to the traditional functional approach.**

It is important to emphasize that the proposed method is similar to acceptance charts. The idea is that the product be produced from a process under statistical control at an acceptable level, but to assume that the process can present signs of out-of-control provided that the probability of producing items outside the specification limits be small,

determined by economic and practical aspects. The discoveries of this study have many practical implications. For example, provide a basis for action on the process to find the quality level with economically satisfying results for the company, using engineering knowledge and practical experience.

Future research

This is a fruitful area for more investigation, and various issues remain unanswered. Another study could evaluate the effect on the performance of the proposed location control chart if the standard deviation of the response variable is unknown. Another question that would deserve to be investigated would be the impact of variability in the profile regression coefficients, altering its form.

Another interesting issue is that the specification ranges are not constant throughout the entire extent considered for the exploratory variable. Further research could be undertaken to explore how a weighting for the control limits according to each value of exploratory variable could confer weight to those points where the characteristic of quality is more relevant, from the point of view of the product performance.

Finally, other interesting possibility would be to research the applicability of the proposed method in situations where the observed profile represents, not a functional relationship between two variables, but effectively a geometric shape as, by example, in the stamping processes.

References

ABDELLA, G. M. et al. Double ewma-based polynomial quality profiles monitoring. *Quality and Reliability Engineering International*, Wiley Online Library, v. 32, n. 8, p. 2639–2652, 2016.

ADIBI, A.; BORROR, C. M.; MONTGOMERY, D. C. Phase ii monitoring of polynomial and nonlinear profiles using a p-value approach. *International Journal of Quality Engineering and Technology*, Inderscience Publishers (IEL), v. 5, n. 2, p. 101–113, 2015.

AHMAD, L.; ASLAM, M.; JUN, C.-H. The design of a new repetitive sampling control chart based on process capability index. *Transactions of the Institute of Measurement and Control*, v. 38, n. 8, p. 971–980, 2016.

AHMAD, S. et al. Bibliometric analysis for process capability research. *Quality Technology & Quantitative Management*, Taylor & Francis, v. 16, n. 4, p. 459–477, 2019.

AIAG/ASQC. *Fundamental Statistical Process Control: Reference Manual*. Troy, MI: [s.n.], 1991.

ALLEN, T. T. *Introduction to engineering statistics and Lean Six Sigma: Statistical quality control and design of experiments and systems*. 3. ed. London: Springer-Verlag, 2019. ISBN 978-1-4471-7420-2.

ALT, F. B. Bonferroni inequalities and intervals. *Encyclopedia of statistical sciences*, Wiley Online Library, v. 1, p. 617–622, 2006.

AMIRI, A.; JENSEN, W. A.; KAZEMZADEH, R. B. A case study on monitoring polynomial profiles in the automotive industry. *Quality and Reliability Engineering International*, v. 26, n. 5, p. 509–520, 2010.

ATASHGAR, K.; ABBASSI, L. A new model to monitor very small effects of a polynomial profile. *International Journal of Quality & Reliability Management*, Emerald Publishing Limited, 2020.

AWAD, M. Fault detection of fuel systems using polynomial regression profile monitoring. *Quality and Reliability Engineering International*, Wiley Online Library, v. 33, n. 4, p. 905–920, 2017.

BERSIMIS, S.; PSARAKIS, S.; PANARETOS, J. Multivariate statistical process control charts: An overview. *Quality and Reliability Engineering International*, v. 23, n. 5, p. 517–543, 2007.

BERTRAND, J. W. M.; FRANSOO, J. C. Operations management research methodologies using quantitative modeling. *International Journal of Operations & Production Management*, MCB UP Ltd, v. 22, n. 2, p. 241–264, 2002.

BOEING. *Advanced Quality System Tools*. [S.l.]: The Boeing Company Seattle, WA, 1998.

BUI, A. T.; APLEY, D. W. Monitoring for changes in the nature of stochastic textured surfaces. *Journal of Quality Technology*, Taylor & Francis, v. 50, n. 4, p. 363–378, 2018.

CAPEZZA, C. et al. Functional regression control chart for monitoring ship co 2 emissions. *Quality and Reliability Engineering International*, Wiley Online Library, 2021.

CAPIZZI, G. Recent advances in process monitoring: Nonparametric and variable-selection methods for phase i and phase ii. *Quality Engineering*, v. 27, n. 1, p. 44–67, 2015.

CENTOFANTI, F. et al. Functional regression control chart. *Technometrics*, Taylor & Francis, p. 1–14, 2020.

CHAN, L. K.; CHENG, S. W.; SPIRING, F. A. A new measure of process capability: Cpm. *Journal of Quality Technology*, Taylor & Francis, v. 20, n. 3, p. 162–175, 1988.

CHANG, T. C.; GAN, F. F. Application of x-bar control chart with modified limits in process control. *Quality and Reliability Engineering International*, Wiley Online Library, v. 15, n. 5, p. 355–362, 1999.

CHICKEN, E.; JR, J. J. P.; SIMPSON, J. R. Statistical process monitoring of nonlinear profiles using wavelets. *Journal of Quality Technology*, Taylor & Francis, v. 41, n. 2, p. 198–212, 2009.

CIVERCHIA, F. et al. Industrial internet of things monitoring solution for advanced predictive maintenance applications. *Journal of Industrial Information Integration*, v. 7, p. 4–12, 2017.

COLOSIMO, B. M.; PACELLA, M. A comparison study of control charts for statistical monitoring of functional data. *International Journal of Production Research*, Taylor & Francis, v. 48, n. 6, p. 1575–1601, 2010.

COLOSIMO, B. M.; SEMERARO, Q.; PACELLA, M. Statistical process control for geometric specifications: on the monitoring of roundness profiles. *Journal of quality technology*, Taylor & Francis, v. 40, n. 1, p. 1–18, 2008.

DEMING, W. E. *The new economics for industry, government, education*. 2. ed. Cambridge: Mit Press, 2000.

DIANDA, D. F.; QUAGLINO, M. B.; PAGURA, J. A. Performance of multivariate process capability indices under normal and non-normal distributions. *Quality and Reliability Engineering International*, Wiley Online Library, v. 32, n. 7, p. 2345–2366, 2016.

DING, Y.; ZENG, L.; ZHOU, S. Phase i analysis for monitoring nonlinear profiles in manufacturing processes. *Journal of Quality Technology*, Taylor & Francis, v. 38, n. 3, p. 199–216, 2006.

DIYOKE, G. C.; OKEKE, C.; ANIAGWU, U. Different methods of speed control of three-phase asynchronous motor. *American Journal of Electrical and Electronic Engineering*, v. 4, n. 2, p. 62–68, 2016.

EBADI, M.; SHAHRIARI, H. A process capability index for simple linear profile. *The International Journal of Advanced Manufacturing Technology*, Springer, v. 64, n. 5-8, p. 857–865, 2013.

FALLAHDIZCHEH, A.; WANG, C. Profile monitoring based on transfer learning of multiple profiles with incomplete samples. *IISE Transactions*, Taylor & Francis, p. 1–69, 2021.

FEIGENBAUM, A. V. *Total quality control*. 3. ed. New York: McGraw-Hill, 1991.

FELIPE, D. de; BENEDITO, E. A review of univariate and multivariate process capability indices. *The International Journal of Advanced Manufacturing Technology*, Springer, v. 92, n. 5-8, p. 1687–1705, 2017.

FREUND, R. A. Acceptance control charts. *Industrial Quality Control*, v. 14, n. 4, p. 13–23, 1957.

GAO, R. X. et al. Big data analytics for smart factories of the future. *CIRP annals*, Elsevier, v. 69, n. 2, p. 668–692, 2020.

GARCÍA-DÍAZ, J. C.; APARISI, F. Economic design of ewma control charts using regions of maximum and minimum arl. *IIE Transactions*, Taylor & Francis, v. 37, n. 11, p. 1011–1021, 2005.

GARVIN, D. A. What does" product quality" really mean? *Sloan Management Review (pre-1986)*, Massachusetts Institute of Technology, Cambridge, MA, v. 26, n. 1, p. 25, 1984.

GHARTEMANI, M. K.; NOOROSSANA, R.; NIAKI, S. T. A. A new approach in capability analysis of processes monitored by a simple linear regression profile. *Quality and Reliability Engineering International*, Wiley Online Library, v. 32, n. 1, p. 209–221, 2016.

GRIGG, N. P. Redefining quality in terms of value, risk and cost: A literature review. *International Journal of Quality & Reliability Management*, Emerald Publishing Limited, 2020.

GUPTA, S.; MONTGOMERY, D.; WOODALL, W. Performance evaluation of two methods for online monitoring of linear calibration profiles. *International journal of production research*, Taylor & Francis, v. 44, n. 10, p. 1927–1942, 2006.

HE, P. Q.; WANG, J. Statistical process monitoring in the era of smart manufacturing. In: *American Control Conference*. [S.l.: s.n.], 2017. p. 4797–4802.

HE, P. Q.; WANG, J. Statistical process monitoring as a big data analytics tool for smart manufacturing. *Journal of Process Control*, v. 67, p. 35–43, 2018.

HILL, D. Modified control limits. *Applied Statistics*, JSTOR, v. 5, n. 1, p. 12, mar 1956.

HOCKMAN, K. K.; JENSEN, W. A. Statisticians as innovation leaders. *Quality Engineering*, Taylor & Francis, v. 28, n. 2, p. 165–174, 2016.

HOERL, R. W.; SNEE, R. Statistical thinking and methods in quality improvement: a look to the future. *Quality engineering*, Taylor & Francis, v. 22, n. 3, p. 119–129, 2010.

HOLMES, D.; MERGEN, A. E. Building an acceptance chart: These handy tools can help you meet specification limits. *Quality Digest*, 2000.

HOSSEINIFARD, S. Z.; ABBASI, B. Evaluation of process capability indices of linear profiles. *International Journal of Quality & Reliability Management*, v. 29, n. 2, p. 162–176, 2012.

HOSSEINIFARD, S. Z.; ABBASI, B. Process capability analysis in non normal linear regression profiles. *Communications in Statistics-Simulation and Computation*, Taylor & Francis, v. 41, n. 10, p. 1761–1784, 2012.

HOSSEINIFARD, S. Z.; ABBASI, B.; ABDOLLAHIAN, M. Performance analysis in non-normal linear profiles using gamma distribution. In: IEEE. *2011 Eighth International Conference on Information Technology: New Generations (ITNG)*. [S.l.], 2011. p. 603–607.

HUANG, T. et al. Statistical process monitoring in a specified period for the image data of fused deposition modeling parts with consistent layers. *Journal of Intelligent Manufacturing*, Springer, p. 1–16, 2020.

HUBIG, L.; LACK, N.; MANSMANN, U. Statistical process monitoring to improve quality assurance of inpatient care. *BMC Health Services Research*, Springer, v. 20, n. 1, p. 21, 2020.

JEN, C.-H.; FAN, S.-K. S. Profile monitoring of reflow process using approximations of mixture second-order polynomials. *Journal of Chemometrics*, Wiley Online Library, v. 28, n. 12, p. 815–833, 2014.

JENSEN, W. A.; BIRCH, J. B.; WOODALL, W. H. Monitoring correlation within linear profiles using mixed models. *Journal of Quality Technology*, Taylor & Francis, v. 40, n. 2, p. 167–183, 2008.

JEONG, M. K.; LU, J.-C.; WANG, N. Wavelet-based spc procedure for complicated functional data. *International Journal of Production Research*, Taylor & Francis, v. 44, n. 4, p. 729–744, 2006.

JIN, J.; SHI, J. Automatic feature extraction of waveform signals for in-process diagnostic performance improvement. *Journal of Intelligent Manufacturing*, v. 12, n. 3, p. 257–268, 2001.

JURAN, J. M.; GODFREY, A. B. *Juran's Quality Handbook*. 5. ed. New York: McGraw-Hill Professional, 1998.

KANE, V. E. Process capability indices. *Journal of quality technology*, Taylor & Francis, v. 18, n. 1, p. 41–52, 1986.

KANG, L.; ALBIN, S. L. On-line monitoring when the process yields a linear profile. *Journal of quality Technology*, Taylor & Francis, v. 32, n. 4, p. 418–426, 2000.

KAZEMZADEH, R.; NOOROSSANA, R.; AMIRI, A. Monitoring polynomial profiles in quality control applications. *The International Journal of Advanced Manufacturing Technology*, Springer, v. 42, n. 7, p. 703–712, 2009.

KAZEMZADEH, R. B.; NOOROSSANA, R.; AMIRI, A. Phase i monitoring of polynomial profiles. *Communications in Statistics—Theory and Methods*, Taylor & Francis, v. 37, n. 10, p. 1671–1686, 2008.

KESHTELI, R. N. et al. Developing functional process capability indices for simple linear profile. *Scientia Iranica. Transaction E, Industrial Engineering*, Sharif University of Technology, v. 21, n. 3, p. 1044, 2014.

KESHTELI, R. N. et al. Functional process capability indices for circular profile. *Quality and Reliability Engineering International*, Wiley Online Library, v. 30, n. 5, p. 633–644, 2014.

KHAN, M. et al. Big data challenges and opportunities in the hype of industry 4.0. In: IEEE. *2017 IEEE International Conference on Communications (ICC)*. [S.l.], 2017. p. 1–6.

KIM, K.; MAHMOUD, M. A.; WOODALL, W. H. On the monitoring of linear profiles. *Journal of Quality Technology*, American Society for Quality, v. 35, n. 3, p. 317, 2003.

KONRATH, A. C.; DONATELLI, G. D.; HAMBURG-PIEKAR, D. The application of monte carlos simulation to evaluate the uncertainty of control chart performance indices. 2006.

KOOSHA, M.; NOOROSSANA, R.; MEGAHED, F. Statistical process monitoring via image data using wavelets. *Quality and Reliability Engineering International*, Wiley Online Library, v. 33, n. 8, p. 2059–2073, 2017.

KOTZ, S.; JOHNSON, N. L. Process capability indices: a review, 1992-2000. *Journal of Quality Technology*, v. 34, n. 1, p. 2–12, 2002.

KOTZ, S.; LOVELACE, C. R. *Process capability indices in theory and practice*. [S.l.]: Arnold, 1998.

LANDIM, T. R.; JARDIM, F. S.; OPRIME, P. C. Modified control chart for monitoring the variance. *Brazilian Journal of Operations & Production Management*, v. 18, n. 3, p. 1–10, 2021.

LEONI, R. C. Gráfico de hotelling com esquemas especiais de amostragem para o monitoramento de processos bivariados autocorrelacionados. Universidade Estadual Paulista (UNESP), 2015.

LIZARELLI, F. L. et al. A bibliometric analysis of 50 years of worldwide research on statistical process control. *Gestão & Produção*, SciELO Brasil, v. 23, n. 4, p. 853–870, 2016.

MAHMOUD, M. A. Phase i analysis of multiple linear regression profiles. *Communications in Statistics—Simulation and Computation®*, Taylor & Francis, v. 37, n. 10, p. 2106–2130, 2008.

MALEKI, M. R.; AMIRI, A.; CASTAGLIOLA, P. An overview on recent profile monitoring papers (2008–2018) based on conceptual classification scheme. *Computers & Industrial Engineering*, Elsevier, v. 126, p. 705–728, 2018.

MARTIN, J.; ELG, M.; GREMYR, I. The many meanings of quality: Towards a definition in support of sustainable operations. *Total Quality Management & Business Excellence*, Taylor & Francis, p. 1–14, 2020.

MASON, R. L.; YOUNG, J. C. Why multivariate statistical process control? *Quality Progress*, v. 31, n. 12, p. 88–93, 1998.

MATSUURA, S. Bayes estimator of process capability index cpk with a specified prior mean. *Communications in Statistics-Theory and Methods*, Taylor & Francis, p. 1–13, 2021.

MEGAHED, F. M.; JONES-FARMER, L. A. Statistical perspectives on "big data". In: *Frontiers in Statistical Quality Control*. [S.l.]: Springer, 2015. v. 11, p. 29–47.

MOHAMMADIAN, F.; AMIRI, A. Economic-statistical design of acceptance control chart. *Quality and Reliability Engineering International*, Wiley Online Library, v. 29, n. 1, p. 53–61, 2013.

MONTGOMERY, D. C. *Introduction to statistical quality control*. 8th. ed. Hoboken, New Jersey: John Wiley & Sons, 2019. ISBN 978-1-119-39930-8v.

MORABITO, R.; PUREZA, V. Modelagem e simulação. In: MIGUEL, P. A. C. (Ed.). *Metodologia de pesquisa em engenharia de produção e gestão de operações*. 2. ed. Rio de Janeiro: Elsevier, 2012. p. 169–198.

MOSESOVA, S. A. et al. Profile monitoring using mixed-effects models. *Dept. Statist. Actuarial Sci., Univ. Waterloo, Waterloo, ON, Canada, Tech. Rep*, p. 06–06, 2006.

NAIR, V.; HANSEN, M.; SHI, J. Statistics in advanced manufacturing. *Journal of the American Statistical Association*, Taylor & Francis Group, v. 95, n. 451, p. 1002–1005, 2000.

NIAVARANI, M. R.; NOOROSSANA, R.; ABBASI, B. Three new multivariate process capability indices. *Communications in Statistics-Theory and Methods*, Taylor & Francis, v. 41, n. 2, p. 341–356, 2012.

NOOROSSANA, R.; SAGHAEI, A.; AMIRI, A. *Statistical analysis of profile monitoring*. 1. ed. Hoboken, New Jersey: John Wiley & Sons, 2011. ISBN 978-0-470-90322-3.

OLIVEIRA, B. K. S. de; OPRIME, P. C.; JARDIM, F. S. Desenvolvimento de um modelo estatístico para medir o desempenho do gráfico de aceitação a partir dos índices de capabilidade do processo. *Cadernos do IME-Série Estatística*, v. 45, p. 01, 2018.

OPRIME, P. C. et al. Acceptance x-bar chart considering the sample distribution of capability indices, Cp and Cpk: A practical and economical approach. *International Journal of Quality & Reliability Management*, v. 36, n. 6, p. 875–894, 2019.

OPRIME, P. C.; MENDES, G. H. de S. The x-bar control chart with restriction of the capability indices. *International Journal of Quality & Reliability Management*, Emerald Publishing Limited, 2017.

OPRIME, P. C. et al. Method for determining the control limits of nonparametric charts for monitoring location and scale. *Gestão & Produção*, SciELO Brasil, v. 23, p. 146–164, 2015.

PARKER, P. A.; FINLEY, T. D. Advancements in aircraft model force and attitude instrumentation by integrating statistical methods. *Journal of aircraft*, v. 44, n. 2, p. 436–443, 2007.

PEARN, W. L.; KOTZ, S. *Encyclopedia and handbook of process capability indices: a comprehensive exposition of quality control measures*. [S.l.]: World Scientific, 2006.

PEARN, W. L.; KOTZ, S.; JOHNSON, N. L. Distributional and inferential properties of process capability indices. *Journal of Quality Technology*, Taylor & Francis, v. 24, n. 4, p. 216–231, 1992.

PEARN, W. L.; LIN, P. C. Testing process performance based on capability index cpk with critical values. *Computer and Industrial Engineering*, v. 47, n. 4, p. 351–369, 2004.

PERAKIS, M. Some new methods for comparing the capability of two processes using process capability index cpm. *International Journal of Quality & Reliability Management*, Emerald Publishing Limited, 2021.

PERDIKIS, T.; PSARAKIS, S. A survey on multivariate adaptive control charts: Recent developments and extensions. *Quality and Reliability Engineering International*, Wiley Online Library, v. 35, n. 5, p. 1342–1362, 2019.

QIU, P. *Introduction to Statistical Process Control, 1st Edition*. [S.l.]: Chapman and Hall/CRC, 2014.

QIU, P.; ZOU, C.; WANG, Z. Nonparametric profile monitoring by mixed effects modeling. *Technometrics*, Taylor & Francis, v. 52, n. 3, p. 265–277, 2010.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2020. Disponível em: <https://www.R-project.org/>.

RAHALI, D. et al. Evaluation of shewhart time-between-events-and-amplitude control charts for correlated data. *Quality and Reliability Engineering International*, Wiley Online Library, v. 37, n. 1, p. 219–241, 2021.

RAKHMAWATI, T. et al. The important level of washing machine quality dimensions in 4.0 industrial era based on the perception of a laundry business: A preliminary investigation. In: IOP PUBLISHING. *IOP Conference Series: Materials Science and Engineering*. [S.l.], 2020. v. 722, n. 1, p. 012048.

RALEA, C. et al. Looking to the future: Digital transformation of quality management. In: FACULTY OF MANAGEMENT, ACADEMY OF ECONOMIC STUDIES, BUCHAREST, ROMANIA. *Proceedings of the INTERNATIONAL MANAGEMENT CONFERENCE*. [S.l.], 2019. v. 13, n. 1, p. 121–132.

REIS, M. S.; GINS, G. Industrial process monitoring in the big data/industry 4.0 era: from detection, to diagnosis, to prognosis. *Process*, v. 5, n. 3, p. 35, 2017.

REIS, M. S.; GINS, G.; RATO, T. J. Incorporation of process-specific structure in statistical process monitoring: A review. *Journal of Quality Technology*, Taylor & Francis, v. 51, n. 4, p. 407–421, 2019.

REIS, M. S.; SARAIVA, P. M. Multiscale statistical process control of paper surface profiles. *Quality Technology & Quantitative Management*, Taylor & Francis, v. 3, n. 3, p. 263–281, 2006.

REIS, M. S.; SARAIVA, P. M. Prediction of profiles in the process industries. *Industrial and Engineering Chemistry Research*, v. 51, n. 11, p. 4254–4266, 2012.

ROHATGI, A. *WebPlotDigitizer*. 2011.

RStudio Team. *RStudio: Integrated Development Environment for R*. Boston, MA, 2019. Disponível em: <http://www.rstudio.com/>.

SCHÜTZE, A.; HELWIG, N.; SCHNEIDER, T. Sensors 4.0–smart sensors and measurement technology enable industry 4.0. *Journal of Sensors and Sensor systems*, Copernicus GmbH, v. 7, n. 1, p. 359–371, 2018.

SHEWHART, W. *Economic Control of Quality of Manufactures Product/50th Anniversary Commemorative Issue*. Milwaukee: Asq Quality Press, 1931. ISBN 9780873890762.

SORIANO, F. R.; OPRIME, P. C.; LIZARELLI, F. L. Impact analysis of critical success factors on the benefits from statistical process control implementation. *Production*, SciELO Brasil, v. 27, 2017.

SUBRAMANI, J.; BALAMURALI, S. Control charts for variables with specified process capability indices. *International Journal of Probability and Statistics*, v. 1, n. 4, p. 101–110, 2012.

VÄNNMAN, K. Distribution and moments in simplified form for a general class of capability indices. *Communications in statistics-theory and methods*, Taylor & Francis, v. 26, n. 1, p. 159–179, 1997.

WALPOLE, R. E. et al. *Probability & statistics for engineers and scientists*. 8th. ed. Upper Saddle River: Pearson Education, 2007.

WANG, F.-K. Measuring the process yield for simple linear profiles with one-sided specification. *Quality and Reliability Engineering International*, Wiley Online Library, v. 30, n. 8, p. 1145–1151, 2014.

WANG, K.; TSUNG, F. Using profile monitoring techniques for a data-rich environment with huge sample size. *Quality and reliability engineering international*, Wiley Online Library, v. 21, n. 7, p. 677–688, 2005.

WANG, S. et al. Robust process capability indices and statistical inference based on model selection. *Computers & Industrial Engineering*, Elsevier, v. 156, p. 107265, 2021.

WILLIAMS, J. D.; WOODALL, W. H.; BIRCH, J. B. Statistical monitoring of nonlinear product and process quality profiles. *Quality and Reliability Engineering International*, Wiley Online Library, v. 23, n. 8, p. 925–941, 2007.

WOODALL, W. H. The statistical design of quality control charts. *Journal of the Royal Statistical Society: Series D (The Statistician)*, Wiley Online Library, v. 34, n. 2, p. 155–160, 1985.

WOODALL, W. H. Current research on profile monitoring. *Production*, SciELO Brasil, v. 17, n. 3, p. 420–425, 2007.

WOODALL, W. H.; FALTIN, F. W. Rethinking control chart design and evaluation. *Quality Engineering*, Taylor & Francis, v. 31, n. 4, p. 596–605, 2019.

WOODALL, W. H.; MONTGOMERY, D. C. Some current directions in the theory and application of statistical process monitoring. *Journal of Quality Technology*, v. 46, n. 1, p. 78–94, 2014.

WOODALL, W. H. et al. Using control charts to monitor process and product quality profiles. *Journal of Quality Technology*, Taylor & Francis, v. 36, n. 3, p. 309–320, 2004.

WOODALL, W. H. et al. An overview and perspective on social network monitoring. *IISE Transactions*, Taylor & Francis, v. 49, n. 3, p. 354–365, 2017.

WU, C.-W.; PEARN, W.; KOTZ, S. An overview of theory and practice on process capability indices for quality assurance. *International journal of production economics*, Elsevier, v. 117, n. 2, p. 338–359, 2009.

WU, X.-f. An assessment approach for process capability in simple linear profile. In: SPRINGER. *Proceedings of the 22nd International Conference on Industrial Engineering and Engineering Management 2015*. [S.l.], 2016. p. 613–620.

YAO, C. et al. A phase ii control chart based on the weighted likelihood ratio test for monitoring polynomial profiles. *Journal of Statistical Computation and Simulation*, Taylor & Francis, v. 90, n. 4, p. 676–698, 2020.

ZHOU, X.; GOVINDARAJU, K.; JONES, G. Acceptance control and guardbanding for error-prone individual measurements. *Quality and Reliability Engineering International*, Wiley Online Library, v. 35, n. 2, p. 517–534, 2019.

ZOU, C.; TSUNG, F.; WANG, Z. Monitoring general linear profiles using multivariate exponentially weighted moving average schemes. *Technometrics*, Taylor & Francis, v. 49, n. 4, p. 395–408, 2007.

ZOU, C.; ZHANG, Y.; WANG, Z. A control chart based on a change-point model for monitoring linear profiles. *IIE transactions*, Taylor & Francis, v. 38, n. 12, p. 1093–1103, 2006.

Annex

ANNEX A – Histograms

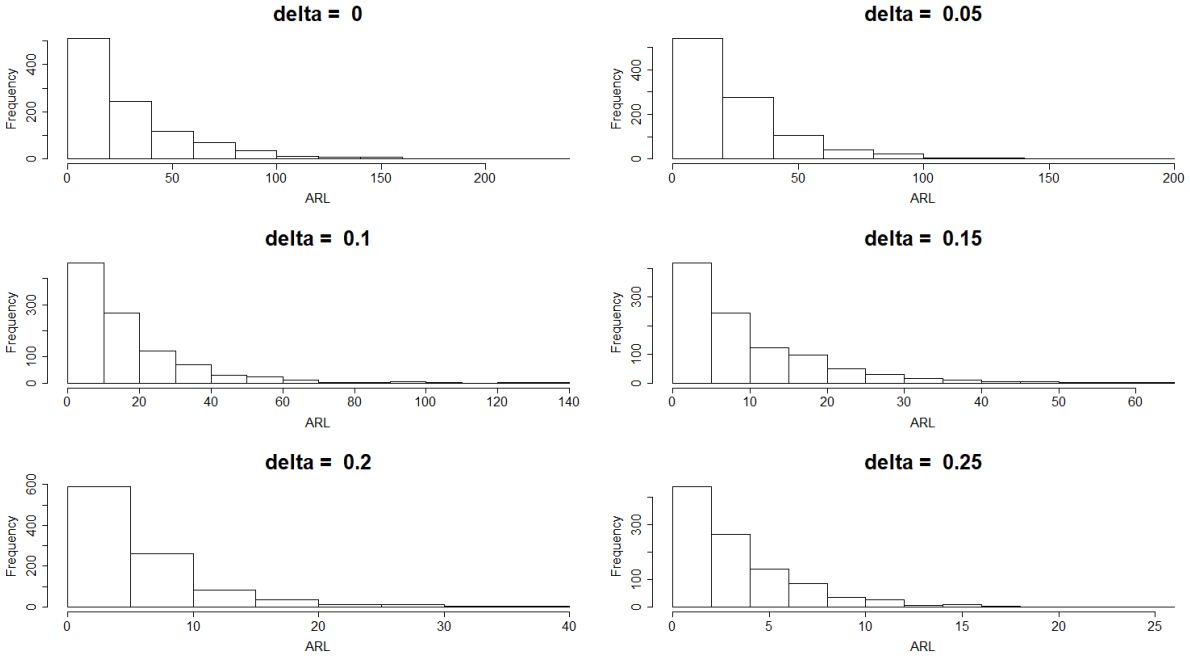


Figure 40 – Histograms of simulated distribution of ARL for samples with $m = 25$ profiles. (1000 simulations, $\alpha^{'} = 0.0027$)

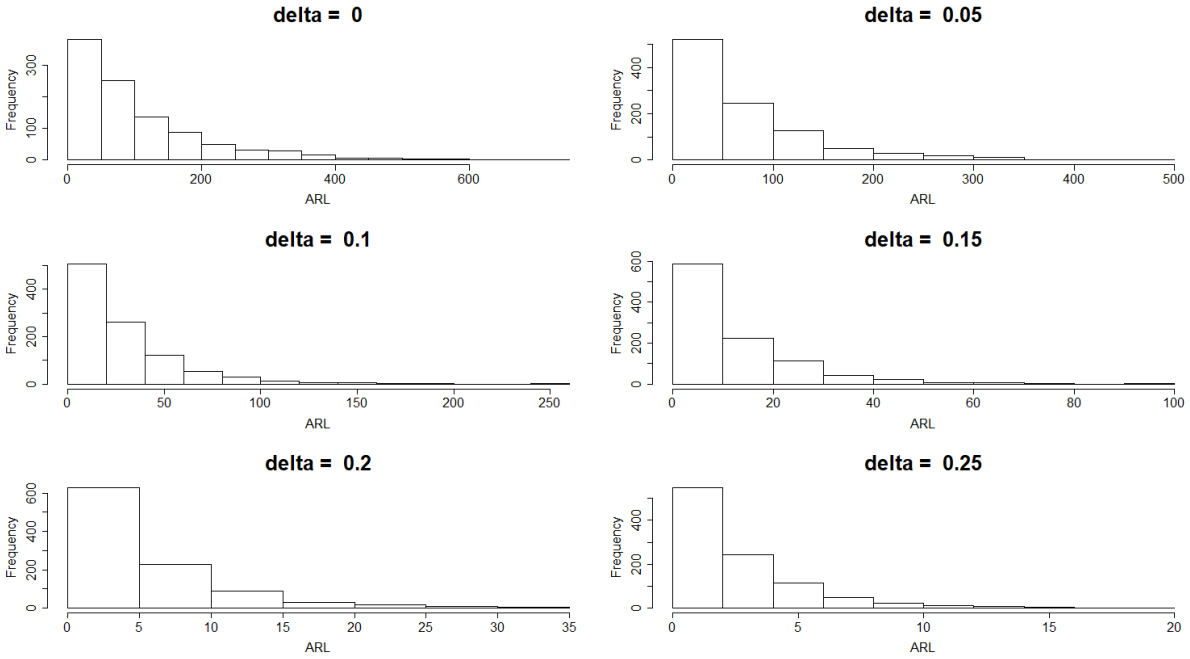


Figure 41 – Histograms of simulated distribution of ARL for samples with $m = 50$ profiles. (1000 simulations, $\alpha^{'} = 0.0027$)

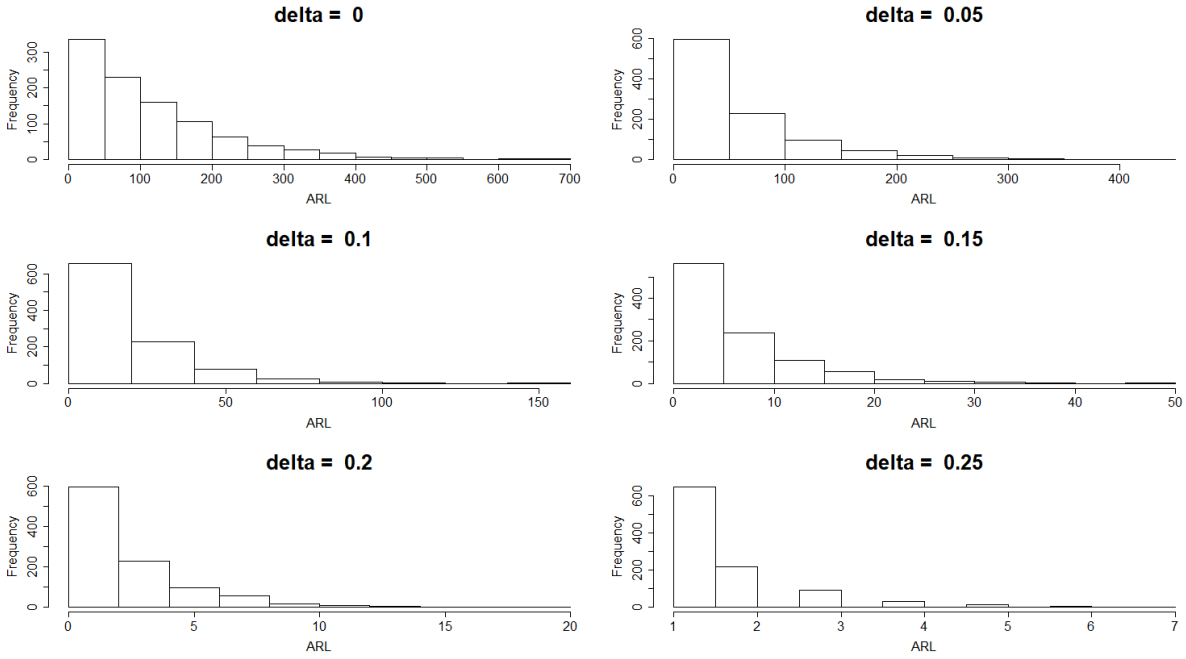


Figure 42 – Histograms of simulated distribution of ARL for samples with $m = 75$ profiles. (1000 simulations, $\alpha^{'} = 0.0027$)

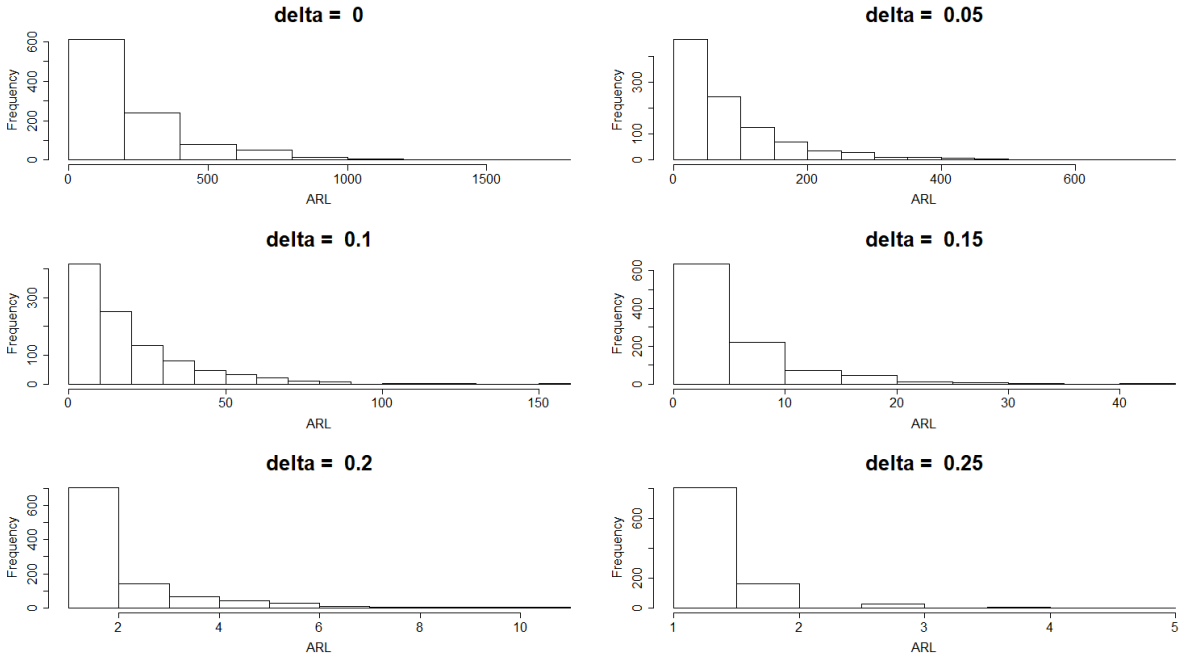


Figure 43 – Histograms of simulated distribution of ARL for samples with $m = 100$ profiles. (1000 simulations, $\alpha^{'} = 0.0027$)

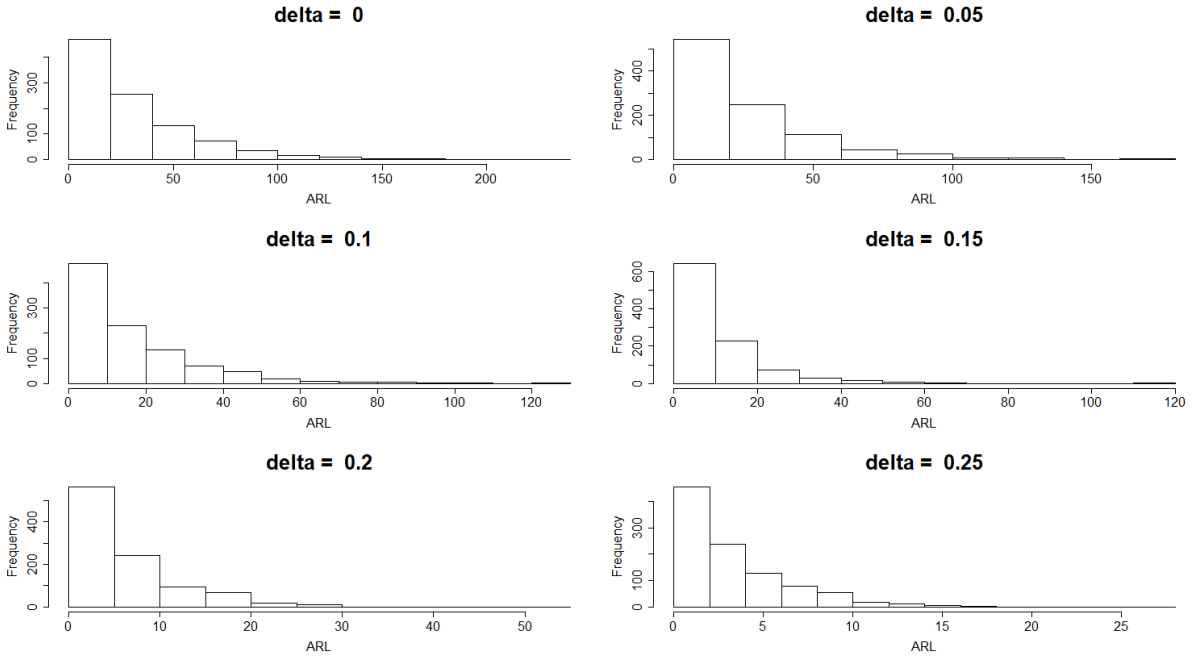


Figure 44 – Histograms of simulated distribution of ARL for samples with $m = 25$ profiles. (1000 simulations, $\alpha^{'} = 0.00135$)

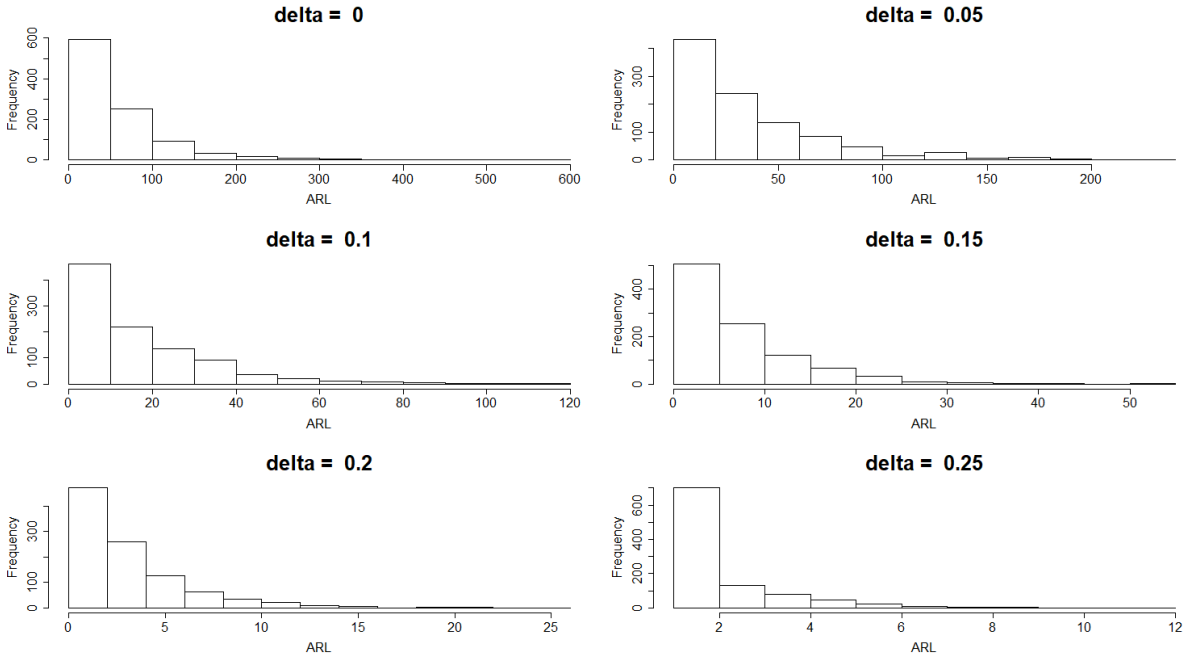


Figure 45 – Histograms of simulated distribution of ARL for samples with $m = 50$ profiles. (1000 simulations, $\alpha^{'} = 0.00135$)

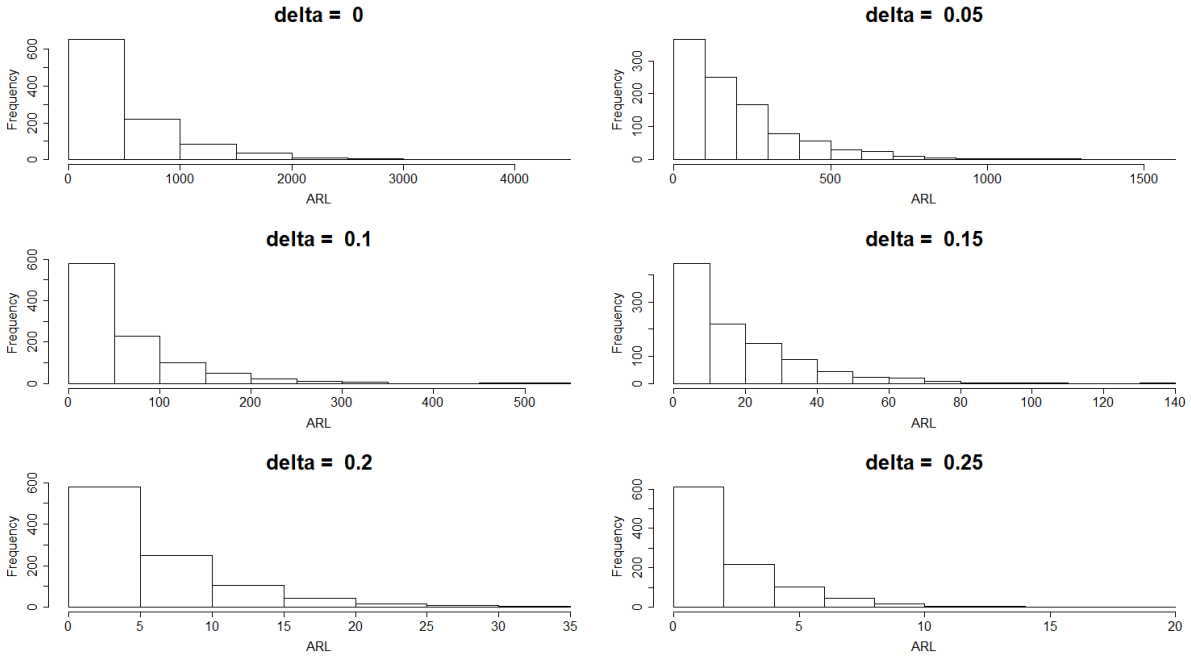


Figure 46 – Histograms of simulated distribution of ARL for samples with $m = 75$ profiles. (1000 simulations, $\alpha^{'} = 0.00135$)

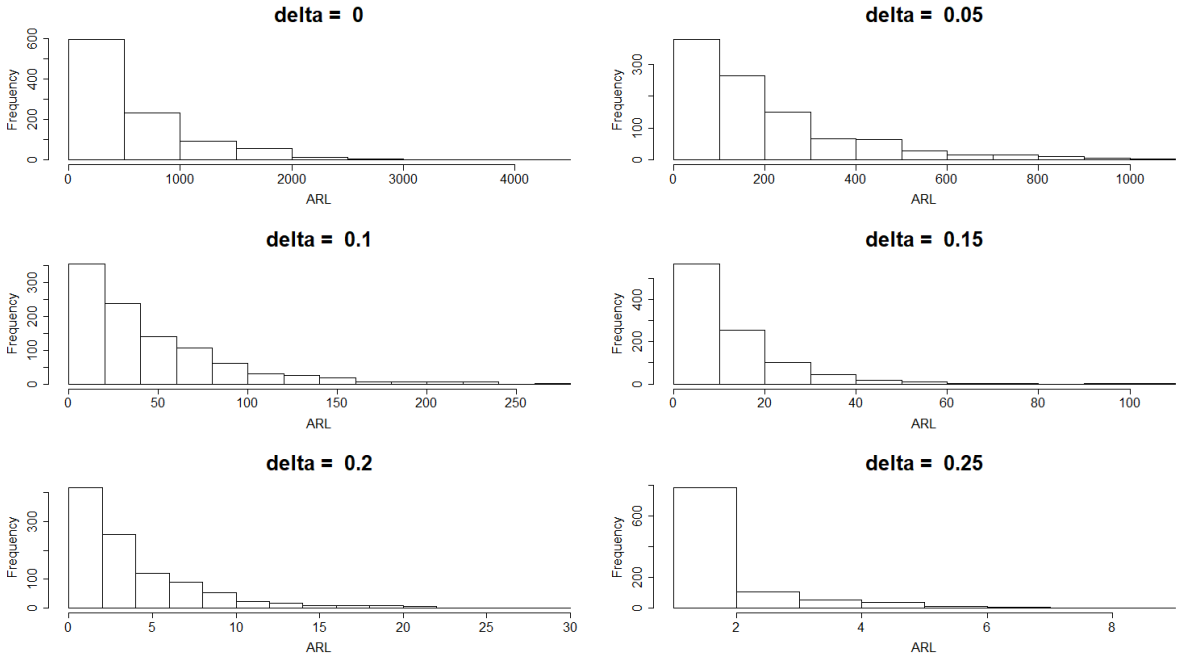


Figure 47 – Histograms of simulated distribution of ARL for samples with $m = 100$ profiles. (1000 simulations, $\alpha^{'} = 0.00135$)

ANNEX B – R Code

```
##########################################
#
# R Code to simulate ARL (Tables 10 - 15)
#
##########################################

rm(list = ls()) #Clears environment variables

#Load required libraries
library(tictoc)
library(Hmisc)
library(rstudioapi)

#Set the location of the .R file as the session's working
   directory
setwd(dirname(getSourceEditorContext()$path))

##Defines which curve (R: Braking or A: Release) will be used
s <- "R" # "R": Shaft braking; "A": Shaft release

#Calculates the Torque corresponding to the values of the
   explanatory variable
#by applying linear regression to the sample values immediately
   below and after
#the applied indices (ex. 100, 200, ..., 1700)
calcTor <- function(index,x1,y1,x2,y2){
 return(((y2 - y1)*((index - x1)/(x2 - x1)) + y1))
}

file.ca <- "./Data/XXXXXXXXX/220V/XXXXXXXXXX/XXXXXXXXXXX.csv" #
   Sample data
file.ct <- "./Data/XXXXXXXXX/220V/XXXXXXXXXXXX.csv" #Theoretical
   data

#Read the sample data (ca)
ca <- read.csv(file.ca, header = FALSE)
#Read theoretical curve data (ct)
```

```
ct <- read.csv2(file.ct, header = TRUE)

#Splits the data into two sets: RPM increase and decrease
ca.R <- ca[1:which.min(ca$V1),,]  #RPM Decrease
ca.A <- ca[which.min(ca$V1) + 1:(length(ca$V1) - which.min(ca$V1)
   ),,] #RPM Increase

#Initializes torque and rotation vectors
rot.R <- c()
rot.A <- c()
tor.R <- c()
tor.A <- c()

index <- seq(from = 100, to = 1700, by = 50)

if (s == "A") {
 cv <- ca.A
} else if (s == "R") {
 cv <- ca.R[order(ca.R$V1),,]
}

for (idx in index) {
 for (i in 1:length(cv$V1)) {
  if (cv$V1[i] < idx &  cv$V1[i + 1] > idx) {
   # print(i)
   rot.R <- c(rot.R,idx)
   tor.R <- c(tor.R,calcTor(idx,cv$V1[i],cv$V2[i],cv$V1[i + 1],cv
      $V2[i + 1]))
  }
 }
}

rt <- ct$RPM[1:17] #Theoretical curve rotation data
tq.LM <- ct$LM[1:17] #Theoretical curve torque data (LM)
tq.LIE <- ct$LIE[1:17] #Theoretical curve torque data (LIE)
tq.LSE <- ct$LSE[1:17] #Theoretical curve torque data (LSE)

obs <- 17 #Number of observations per profile
intervalo <- ((max(index) - min(index)))/(obs - 1)
x <- seq(100, 1700, by = intervalo)

#Fits a 4th-order regression model to the LM of the theoretical
```

```
   curve
fit.LM <- lm(ct$LM~poly(ct$RPM,4,raw=TRUE))

a0.LM <- fit.LM$coefficients[1]
a1.LM <- fit.LM$coefficients[2]
a2.LM <- fit.LM$coefficients[3]
a3.LM <- fit.LM$coefficients[4]
a4.LM <- fit.LM$coefficients[5]

y.LM <- a0.LM + a1.LM*x + a2.LM*x^2 + a3.LM*x^3 + a4.LM*x^4

tic() #Start counting time
##############################################
set.seed(1234) #Initializes a pseudo-random number generator

n_sims     <- 100000 #Number of simulations
n_samples  <- 20 #Sample number
n_profiles <- 3 #Number of profiles per sample
n_obs      <- 17 #Number of observations per profile
alpha      <- 0.0027 #Type I Error assumed
sigma      <- 0.018 #Standard deviation obtained from sample data

Cp  <- 2.00 #Cpk_max (Potential process capability)
Cpk <- 1.33 #Cpk0 (Minimum allowable capability)

#sigma <- mean((tq.LSE-tq.LIE)/(6*Cp)) #Determinado a partir dos
   LE para um Cp determinado

T <- array(0,c(n_samples,n_profiles,n_obs)) #Initialize the array
k <- abs(qnorm((alpha/n_obs)/2)) #CC opening factor corrected by
   Bonferroni

#Simulates n_samples (samples) with n_profiles (profiles)
   containing n_obs (observations)
#centered on the nominal profile (LM) with standard deviation
   sigmaT
for(i in 1:n_samples){
 for (j in 1:n_profiles) {
  T[i,j,] <- a0.LM + a1.LM*x + a2.LM*x^2 + a3.LM*x^3 + a4.LM*x^4
     +
  rnorm(length(x),0,abs( rnorm(1,0,sigma)))
 }
```

```
}

mean_T  <- vector(length = n_obs) #Initialize Torque vector for/
   each RPM
sd_T    <- vector(length = n_obs) #Initialize sd vector for/each
   RPM
UCL_T   <- vector(length = n_obs) #Initialize UCL_T vector for/
   each RPM
LCL_T   <- vector(length = n_obs) #Initialize LCL_T vector for/
   each RPM

#Calculates the Control Limits for the simulated values (with
   mean LM and sigma standard
#deviation)
for(i in 1:n_obs) {
 #mean_T[i] <- y.LM[i] #Centered on Nominal Profile
 mean_T[i] <- mean(T[,,i]) #Centered on Sample Average
 sd_T[i]   <- sd(T[,,i])
 UCL_T[i]  <- mean_T[i] + 3*(Cp - Cpk)*sd_T[i] + k*(sd_T[i]/sqrt(
    n_profiles))
 LCL_T[i]  <- mean_T[i] - 3*(Cp - Cpk)*sd_T[i] - k*(sd_T[i]/sqrt(
    n_profiles))
}

#View the Specification Limits, Control and Profiles of a given
   sample
plot(x,y.LM, ylim = c(0,3), type = "l", xlab = "RPM", ylab = "
   Torque[N/m]")
lines(x,UCL_T, type = "l", col = "red")
lines(x,LCL_T, type = "l", col = "red")
lines(x,tq.LSE, type = "l", col = "red")
lines(x,tq.LIE, type = "l", col = "red")
for(i in 1:n_samples ) {
 for(j in 1:n_profiles) {
  lines(x,T[i,j,], type = "l", col = "blue")
 }
}

par(mfrow = c(2,3))

n_rep <- 100
results <- list()
```

```
shifts <- seq(from = 0, to = 2.5, by = 0.25)
# mean_arl <- seq(1:length(shifts))
# median_arl <- seq(1:length(shifts))
# sd_arl <- seq(1:length(shifts))
# min_arl <- seq(1:length(shifts))
# max_arl <- seq(1:length(shifts))
#arl <- vector(length = n_sims)
arl <- vector(length = length(shifts))
T_Sim <- matrix(0, nrow = n_profiles, ncol = n_obs) #Initialize
   the array

for(r in 1:n_rep) {
 signal_global <- 0
 for (delta in shifts) {
  signal <- 0
  for (n_s in 1:n_sims) {
   for (n_p in 1:n_profiles) {
    T_Sim[n_p,] <- a0.LM + a1.LM*x + a2.LM*x^2 + a3.LM*x^3 + a4.
       LM*x^4 +
    rnorm(length(x),0,abs( rnorm(1,0,sigma))) + delta*sigma
   }
   condition <- (colMeans(T_Sim) > UCL_T | colMeans(T_Sim) < LCL_
      T)
   if(any(condition)) {signal <- signal + 1}
  }
  p <- signal/n_sims
  arl[delta*(10/2.5) + 1] <- 1/p
 }
 results[[r]] <- arl
 print(r)
}

results <- matrix(unlist(results), ncol=length(arl),byrow = TRUE)
summary(results)

toc()#End counting time

#CV <- apply(X = results, MARGIN = 2, FUN = sd)/apply(X = results
   , MARGIN = 2, FUN = mean)

##########################################
#
```

```
# R Code to Plot ARL (Figures 33 - 38)
#
##########################################

library(ggplot2)
library(readxl)

Results <- read_excel("G:/My Drive/Doctoral Thesis Research/R
   Codes/Simulation Results/Simulations Expanded Limits/Results.
   xlsx",
sheet = "Amostra 40 - 100")
Results$n <- as.factor(Results$n)

head(Results)

p <- ggplot(data = Results, mapping = aes(x = delta, y = ARL,
   color = n)) +
geom_point() +
theme(panel.grid.major = element_line(color = "black",
size = 0.5,
linetype = 2)) +
geom_line() +
facet_wrap(facets =  vars(Cp)) +
scale_y_log10() +
scale_x_continuous(breaks = c(0.0, 0.25, 0.50, 0.75, 1.00, 1.25,
   1.50, 1.75, 2.00, 2.25, 2.50)) +
labs(x = "delta (\u03b4)", color = "n")
theme_bw ()

to_string <- as_labeller(c('1.33' = "Cp = 1.33",
'1.5' = "Cp = 1.50",
'1.75' = "Cp = 1.75",
'2' = "Cp = 2.00"))

p + facet_wrap(facets =  vars(Cp), as.table = TRUE, labeller = to
   _string)


##############################################
#
# R Code to Plot Process Capability Analysis (Figures 31 - 38)
#
```

```
##################################################

rm(list = ls())
library(Hmisc)
library(readxl)
library(plyr)
library(readr)
library(pracma)
library(qcc)

library(rstudioapi)
setwd(dirname(getSourceEditorContext()$path))

file.ct <- "./Data/XXXXXXXXXXX/220V/XXXXXXXXXXXXXXX.csv"
ct <- read.csv2(file.ct, header=TRUE)

rt <- ct$RPM[1:17]
tq.LM <- ct$LM[1:17]
tq.LIE <- ct$LIE[1:17]
tq.LSE <- ct$LSE[1:17]

par(mfrow=c(1,1))
plot(rt,tq.LM, main="Theoretical Curve with Sample Profiles",
xlab="RPM", ylab="Torque [N.m]", type = "n", lwd = 2, col = "
   black",
xlim=c(100,1700), ylim=c(0.0,2.5))
lines(rt,tq.LIE, type= "n", lty = 2, col = "black")
lines(rt,tq.LSE, type= "n", lty = 2, col = "black")
#lines(ss, lwd = 2)
minor.tick(nx = 20, ny = 10)

fit.LM <- lm(ct$LM~poly(ct$RPM,4,raw=TRUE))

a0.LM <- fit.LM$coefficients[1]
a1.LM <- fit.LM$coefficients[2]
a2.LM <- fit.LM$coefficients[3]
a3.LM <- fit.LM$coefficients[4]
a4.LM <- fit.LM$coefficients[5]

x.LM <- seq(100,1700,by=100)
y.LM <- a0.LM + a1.LM*x.LM + a2.LM*x.LM^2 + a3.LM*x.LM^3 + a4.LM*
   x.LM^4
```

```
lines(x.LM,y.LM, col="black", lty=1, lwd=3)

fit.LIE <- lm(ct$LIE~poly(ct$RPM,4,raw=TRUE))

a0.LIE <- fit.LIE$coefficients[1]
a1.LIE <- fit.LIE$coefficients[2]
a2.LIE <- fit.LIE$coefficients[3]
a3.LIE <- fit.LIE$coefficients[4]
a4.LIE <- fit.LIE$coefficients[5]

x.LIE <- seq(100,1700,by=100)
y.LIE <- a0.LIE + a1.LIE*x.LIE + a2.LIE*x.LIE^2 + a3.LIE*x.LIE^3
   + a4.LIE*x.LIE^4
lines(x.LIE,y.LIE, col="red", lty=2, lwd=3, )

fit.LSE <- lm(ct$LSE~poly(ct$RPM,4,raw=TRUE))

a0.LSE <- fit.LSE$coefficients[1]
a1.LSE <- fit.LSE$coefficients[2]
a2.LSE <- fit.LSE$coefficients[3]
a3.LSE <- fit.LSE$coefficients[4]
a4.LSE <- fit.LSE$coefficients[5]

x.LSE <- seq(100,1700,by=100)
y.LSE <- a0.LSE + a1.LSE*x.LSE + a2.LSE*x.LSE^2 + a3.LSE*x.LSE^3
   + a4.LSE*x.LSE^4
lines(x.LSE,y.LSE, col="red", lty=2, lwd=3)

legend(x="topright",legend=c("USL", "NP", "LSL"), col=c("red","
   black","red"),
lty=2:1, lwd = 2, cex=0.8)

calcTor <- function(index,x1,y1,x2,y2){
 return(((y2-y1)*((index-x1)/(x2-x1))+y1))
}

mydir <- "Data/XXXXXXXXXX/220V/XXXXXXXXXXX"
myfiles <- list.files(path=mydir, pattern="*.csv", full.names=
   TRUE)
n <- length(myfiles) #numero de amostras (arquivos "csv")

#le os dados amostrais (ca)
```

```
ca <- lapply(myfiles, read.csv, header = FALSE)

ca.R <- vector("list", n)
cv.R <- vector("list", n)
ca.A <- vector("list", n)
cv.A <- vector("list", n)

for(i in 1:n){
 ca.R[[i]] <- ca[[i]][1:which.min(ca[[i]]$V1),,]
 ca.A[[i]] <- ca[[i]][which.min(ca[[i]]$V1)+1:(length(ca[[i]]$V1)
    -which.min(ca[[i]]$V1)),,]
}

rot.R <- vector("list", n)
rot.A <- vector("list", n)
tor.R <- vector("list", n)
tor.A <- vector("list", n)

index <- seq(from = 100, to = 1700, by = 100)

for(i in 1:n){
 cv.A[[i]] <- ca.A[[i]]
 cv.R[[i]] <- ca.R[[i]][order(ca.R[[i]]$V1),,]
}

for(j in 1:n){

 for (idx in index){
  for (i in 1:length(cv.R[[j]]$V1)){
   if (cv.R[[j]]$V1[i] < idx &  cv.R[[j]]$V1[i+1] > idx){
    # print(i)
    rot.R[[j]] <- c(rot.R[[j]],idx)
    tor.R[[j]] <- c(tor.R[[j]],calcTor(idx,cv.R[[j]]$V1[i],cv.R[[
       j]]$V2[i],cv.R[[j]]$V1[i+1],cv.R[[j]]$V2[i+1]))
   }
  }
 }

 for (idx in index){
  for (i in 1:length(cv.A[[j]]$V1)){
   if (cv.A[[j]]$V1[i] < idx &  cv.A[[j]]$V1[i+1] > idx){
    # print(i)
```

```
    rot.A[[j]] <- c(rot.A[[j]],idx)
    tor.A[[j]] <- c(tor.A[[j]],calcTor(idx,cv.A[[j]]$V1[i],cv.A[[
       j]]$V2[i],cv.A[[j]]$V1[i+1],cv.A[[j]]$V2[i+1]))
   }
  }
 }

}

m <- n
x <- seq(100,1700,by=100)
y <- matrix(0,m,length(x))


for (i in 1:m){
 y[i,] <- tor.R[[i]] #Define (A)celeracao ou (R)educao
 #y[i,] <- a0.LM + a1.LM*x + a2.LM*x^2 + a3.LM*x^3 + a4.LM*x^4 +
    a5.LM*x^5 + a6.LM*x^6 + rnorm(length(x),0,abs(rnorm(1,0,0.03))
    )
}

abline(v= c(seq(100,1600,100)), col = "black", lwd = 1, lty = 2)

################################################
y.mean <- apply(y,2,mean)
y.sd <- apply(y,2,sd)

fit.LMd <- lm(y.mean~poly(x,4,raw=TRUE))

a0.LMd <- fit.LMd$coefficients[1]
a1.LMd <- fit.LMd$coefficients[2]
a2.LMd <- fit.LMd$coefficients[3]
a3.LMd <- fit.LMd$coefficients[4]
a4.LMd <- fit.LMd$coefficients[5]

set.seed(1234)
n_obs <- 17
intervalo <- ((max(index)-min(index)))/(n_obs-1)

delta <- 0
alpha <- 0.0027
sigma <- 0.018
```

```
m <- 20
num_sim <- 100

x <- seq(100,1700,by=intervalo)
y <- matrix(0,m,length(x))
k <- abs(qnorm((alpha/n_obs)/2))

#-------------------------------------------------

n_samples <- 15
simulation <- array(0,dim = c(n_samples, m, n_obs))


#Simulate m profiles centered on the average profile of each
   sample (LMd)
for (i in 1:num_sim){
 for (j in 1:m){
  y[j,] <- a0.LMd + a1.LMd*x + a2.LMd*x^2 + a3.LMd*x^3 + a4.LMd*x
     ^4 + rnorm(length(x),0,abs(rnorm(1,0,sigma))) + delta*sigma
 }
 simulation[i,,] <- y
}

#Simulates m profiles centered on the nominal profile  of the
   theoretical curve (LM)
# for (i in 1:num_sim){
 #   for (j in 1:m){
  #     y[j,] <- a0.LM + a1.LM*x + a2.LM*x^2 + a3.LM*x^3 + a4.LM*
     x^4 + rnorm(length(x),0,abs(rnorm(1,0,sigma))) + delta*sigma
  #   }
 #   simulation[i,,] <- y
 # }

#Plota as curvas(perfis) de CADA amostras (com m perfis e n
   pontos)
#centralizada no perfil nominal
amostra <- 1
for (i in 1:m){
 lines(x,simulation[amostra,i,],type="l",col="blue")
}

RPM500 <- c()
```

```
for(i in 1:n_samples){
 for(j in 1:m){
  RPM500 <- c(RPM500, simulation[i,j,5])
 }
}

sample_size <- m

number_of_samples <- n_samples

sample <- rep(1:number_of_samples, each = sample_size)

df <- data.frame(RPM500, sample)

RPM <- qccGroups(RPM500, sample, data = df)
q <- qcc(RPM, type="xbar", nsigmas=3, plot=TRUE)
processCapability(q, spec.limits=c(1.7916,2.1886), target=1.9989,
    )
```