TECNO-LÓGICA
Revista do depto. de Química e Física, do depto. de Engenharia, Arquitetura e Ciências Agrárias e do Mestrado em Tecnologia Ambiental
http://online.unisc.br/seer/index.php/tecnologica

APPLICATION OF MULTIVARIATE CONTROL CHARTS FOR MONITORING AN INDUSTRIAL PROCESS

*Custodio da Cunha Alves[1], Elisa Henning[2] *, Robert Wayne Samohyl[3], Andrea Cristina Konrath[4], Olga Maria Formigoni Carvalho Walter[3]*

1 Departamento de Engenharia de Produção Mecânica, Universidade da Região de Joinville, 89219-710, Santa Catarina, Brasil.
2 Departamento de Matemática, Universidade do Estado de Santa Catarina, 89219-710, Santa Catarina , Brasil
3 Departamento de Engenharia de Produção, Universidade Federal de Santa Catarina, CEP, Santa Catarina, Brasil
4 Departamento de Informática e Estatística, Universidade Federal de Santa Catarina, CEP, Santa Catarina, Brasil.

**E-mail: elisa.henning@udesc.br*

Recebido em 24/05/2013
Aceito em 24/06/2013

ABSTRACT

The effective simultaneous monitoring of the many quality characteristics of a production process often depends on statistical tools that have become more and more specific. The goal of this paper is to investigate, via an industrial application, whether there are significant differences in sensitivity between the use of Multivariate Cumulative Sum (MCUSUM), Multivariate Exponentially Weighted Average (MEWMA) control charts, and Hotelling T^2 charts to detect small changes in the mean vector of a process. Machining process real data were used. A MCUSUM control chart was applied to monitor these two quality characteristics of this process simultaneously. A MEWMA chart was also applied. The result was compared to that of the application of the Hotelling T^2 chart, which showed that the MCUSUM and MEWMA control charts detected the change sooner. This study was essential to determine the best option between these three charts for the multivariate statistical analysis of this industrial process.

Keyword: Statistical Process Control; MCUSUM; Hotelling T^2; MEWMA; Monitoring.

1 Introduction

Multivariate Statistical Process Control consists of a number of powerful tools for problem solving and improvement of quality control by reducing variability in industrial manufacturing processes. Among these tools, the most commonly used statistical methods in industries are the multivariate control charts.

The most well-known current multivariate control chart used to monitor the mean vector of a process is the traditional Hotelling T^2 chart. In this chart, each data set or each value is entered individually and its relationship to other points is determined only by the chart [1]. Although extremely effective, it is not the only tool available for the multivariate analysis of industrial processes. In some cases, other types of multivariate control charts can complement or replace this chart with improved results. This is the case of multivariate control charts with memory, the subject of the present study. This kind of chart accumulates information from past and recent observations and thereby detects small changes in the parameters of a multivariate process with a lower value of average run length (ARL_1) than that obtained by the Hotelling T^2 charts for the same value of ARL_0 [2].

Furthermore, with the advance of technology, computer programs have facilitated further development and implementation of multivariate statistical methods in industries, such as tools for quality control. However, little attention has been given to the specific characteristics of the specific development of multivariate control charts. Fortunately, free computer programs, such as GNU R environment [3], are excellent alternatives to overcome this deficiency. The importance of these resources to complement some functions that are not available in commercial packages can be highlighted [4]. In this study, statistical analysis is supported by the R language.

This article is organized as follows: section 2 describes multivariate control charts. Section 3 provides information about the data used. Section 4 describes the application of the charts. Section 5 presents the conclusions and final considerations.

2 Multivariate Control Charts

There are situations in an industrial setting where, in the same process, it is necessary to simultaneously control two or more quality characteristics. Although the application of univariate control charts for each individual variable is a possible solution, it can lead to erroneous conclusions if the joint probability of Type I Errors (false alarms) and the correlation structure between the *p* variables in question are not considered [2].

Multivariate charts are developed using the measurements of the results of processes for multiple variables. These measurements are generally presented by subgroups of related items, identified in the literature as rational subgroups. However, in the chemical, petrochemical, mining, and other industries, an item can be considered as part of material collected instantaneously from a given flow or from a homogeneous batch of some product. In these situations, the subject of this study, process monitoring through individual observation, is recommended [2].

The development of multivariate control charts is separated into two phases. The first phase (phase I) consists of obtaining a representative sample of the data with the goal of determining the control limits, which is generally a retrospective study of the data. The second phase (phase II) consists of the monitoring of the process [5].

The performance of a control chart is typically evaluated through parameters related to the distribution of time required for the chart to emit a signal. The Average Run Length (ARL) is one of these parameters. In a control chart, ARL_0 indicates the average number of samples collected until the emission of a signal during the period under control, while ARL_1 represents the average number of samples collected until the emission of a signal that indicates an out-of-control situation [2, 6].

The most frequently used multivariate statistical control charts are the Hotelling T^2, the MCUSUM (Multivariate Cumulative Sum) and the MEWMA (Multivariate Exponentially Weighted Moving Average) charts. The main characteristic of MCUSUM and MEWMA charts is that they are sensitive to small and persistent changes in the process [2, 5].

This study involves the use of the MCUSUM and MEWMA charts for individual observations of a machining process, whose quality characteristics have sample size $n = 1$. In addition, a statistical analysis of the performance data is carried out to determine whether there is a significant difference in the sensitivity of these charts with the Hotelling T^2 chart, especially in terms of the detection of small changes in the mean vector [6].

2.1 Multivariate Cumulative Sum (MCUSUM) Control Chart

The multivariate cumulative sum control chart (MCUSUM) is a procedure that uses the cumulative sum of deviations of each random vector previously observed compared to the nominal value to monitor the mean vector of a multivariate process [4]. This chart was proposed by [7] and is an extension of the univariate CUSUM control chart. In this procedure, the scalar quantities are replaced by vectors. C_i is defined as

$$C_i = \sqrt{(\mathbf{S}_{i-1} + \mathbf{X}_i - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}^{-1}(\mathbf{S}_{i-1} + \mathbf{X}_i - \boldsymbol{\mu}_0)} \quad (1)$$

where $\mathbf{S}_i$ are the cumulative sums expressed as

$$\mathbf{S}_i = \begin{cases} \mathbf{0} & \text{if } C_i \le k, \\ (\mathbf{S}_{i-1} + \mathbf{X}_i)(1 - \frac{k}{C_i}) & \text{if } C_i > k, \end{cases} \quad (2)$$

where $\mathbf{X}_i$ is the vector of p-dimensional sample observations regarding the current sample unit; $\mathbf{S}_0 = 0$ and the reference value $k > \mathbf{0}$. The covariance matrix $\boldsymbol{\Sigma}$ can be known or estimated.

The quantity to be plotted on the control chart is

$$Y_i = \sqrt{\mathbf{S}_i'\boldsymbol{\Sigma}^{-1}\mathbf{S}_i} \; . \quad (3)$$

The method gives an out-of-control signal if $Y_i > h$, where h is the upper control limit.

2.2 Hotelling T^2 Control Chart

Among the existing multivariate charts, the Hotelling T^2 control chart is the one that is the most well-known chart in the literature. Hotelling [2, 8-9] was a pioneer in the research of multivariate quality control. During World War II, he used a multivariate control approach with data containing information about the location of bombers. Initially, Hotelling supposed that the variables in question followed a normal multivariate distribution with a mean vector $\overline{\mathbf{X}}$ and a covariance matrix ***S***. Thus, taking samples of size *n* for each of the *p* variables in question (to be monitored) and considering the estimates of the parameters, the equation for obtaining the estimates of the statistic T^2 is given by

$$T^2 = \left(\mathbf{X} - \overline{\mathbf{X}}\right)'\mathbf{S}^{-1}\left(\mathbf{X} - \overline{\mathbf{X}}\right), \quad (4)$$

where $\overline{\mathbf{X}}$ and $\mathbf{S}^{-1}$ represent, respectively, the estimates for the vector of means and the inverse of the covariance matrix. Expression (4) is used as a base for the construction of the T^2 Hotelling chart for individual observations [8, 9]. In the first phase, the control limits for [2] are based on the beta distribution, and are expressed by Equation (5)

$$UCL = \frac{(m-1)^2}{m} \beta_{\alpha,p/2,(m-p-1)/2} \,, \tag{5}$$

where $\beta_{\alpha,p/2,(m-p-1)/2}$ is the superior percentage point α to a distribution β with parameters *p / 2* and *(m-p-1)/2*. In the second phase, the new limits are established only to monitor future observations using the control limits shown in Equation (6).

$$UCL = \frac{p(m+1)(m-1)}{m^2 - mp} F_{\alpha,p,m,m-p} \,, \tag{6}$$

where *p* is the number of variables (quality characteristic) and *m* is the number of samples. If the value of T^2 exceeds the upper control limit (UCL), the process is said to be out of statistical control. The lower control limit (LCL) for the two phases is equal to zero.

A significant problem in the case of individual observations is the estimate of the covariance matrix of the process. In this article, the estimator of the covariance matrix $\mathbf{\Sigma}$ (7) that uses the difference between successive pairs of observations was used.

$$\mathbf{S} = \frac{1}{2} \sum \frac{\mathbf{v}\mathbf{v}'}{(m-1)} \,, \tag{7}$$

where $\mathbf{S}$ represents the estimate for the covariance matrix of the process and $\mathbf{v} = (x_{i+1} - x_i), i = 1,2,\ldots,m-1$ (see [2]).

2.3 Multivariate Exponentially Weighted Moving Average (MEWMA) Control Chart

The Multivariate Exponentially Weighted Moving Average Chart (MEWMA) [2,10] is a logical extension of the Exponentially Weighted Moving Average Control Chart (EWMA), similar to the MCUSUM. The main difference is that the MEWMA provides greater weight to more current information and less weight to older information. The MEWMA chart is defined as follows

$$\mathbf{Z}_i = r\mathbf{x}_i + (1-r)\mathbf{Z}_{i-1} \,, \tag{8}$$

where *r* is the diagonal matrix which contains the *p* weighting constants $\{r_i \in (0,1]\}$. $\mathbf{Z}_{i-1}$ is a *p*-dimensional vector related to the scores of the sample $i-1$ with $\mathbf{Z}_0 = \boldsymbol{\mu}_0 = \mathbf{0}$. The quantity plotted on the control chart is

$$T_i^2 = \mathbf{Z}_i' \mathbf{\Sigma}_{Zi}^{-1} \mathbf{Z}_i \,, \tag{9}$$

where $\mathbf{\Sigma}_{Zi}^{-1}$ is the inverse of the covariance matrix. The covariance matrix of Z_i is expressed by

$$\mathbf{\Sigma}_{Zi} = \frac{r[1-(1-r)^2]}{2-r} \mathbf{\Sigma} \,, \tag{10}$$

where $\mathbf{\Sigma}$ is the covariance matrix, which can be estimated or known. It is analogous to the variance of the univariate EWMA.

The process is considered under control if $T_i^2 < h$ in Equation (9). The value of *h* (upper control limit) depends on the desired ARL for the MEWMA chart. The ARL performance depends not only on the centrality parameter given by the Mahalanobis distance when *p* characteristics receive the same weight *r*. For this reason, the performance of this chart can be compared with other multivariate charts [10].

3 Performance of Multivariate Control Charts

The evaluation and comparison of different types of multivariate control charts are performed using statistical and economic performance indicators. The average number of samples collected up to the appearance of an out-of control signal (ARL), according to [2], is the most commonly used statistical indicator to evaluate the performance of a control chart and to make comparisons between different types of charts.

The ARL is a parameter that takes into account the probabilities of Type I and Type II errors. In statistical process control, a Type I error occurs if we conclude that the process is out of control when in fact it is not. A type II error occurs when we do not detect when the process is out of control. Therefore, to evaluate the parameters of a control chart, it is customary to study the behavior of the ARL. It is desirable that the ARL of the chart is large when the process is under control and quite small when the process is out of control. Accurate determination of the ARL is not always possible because the majority of control variables involve correlation. However, there are numerical methods for determining parameters that optimize control charts behavior such as the Integral Equation Method, Markov Chains, and Simulation.

In recent decades, great deal of research has been conducted on the improvement and application of numerical methods to obtain approximate parameters for evaluating the performance of the univariate control chart. However, when it comes to optimizing the parameters of a multivariate control chart, few studies have been developed, with the exception of the MCUSUM control chart by [7]. Lowry *et al.* [10] proposed a table for the MEWMA with ARL k e h using the Simulation Method for an under control ARL of 200 and *p* = 2; 3; and 4 quality characteristics.

Lee and Khoo [11] applied the Markov Chain Method for situations under control with parameters ARL, k and h for the MCUSUM chart for individual observations with *p* = 2; 3; and 4 quality characteristics under control for ARL of 100, 200, 370, 500, and 1000. The Integral Equation Method with Gaussian Quadrature was proposed by Alves [4] to optimize the ARL, *k*, and *h* in the MCUSUM control chart for individual observations with p = 2; 3; and 4 for quality characteristics under control for ARL of 200, 500, and 1000. This method involves the analytical derivation of an integral equation, whose numerical solution via Gaussian Quadrature enables the user to obtain the approximation solution of these parameters. This method is an excellent alternative for the optimization of the MCUSUM chart, and since it is more versatile, it provides better results for the value of the ARL and a faster calculation method compared to simulations and relative simplicity of implementation.

The performance in terms of the ARL for various size changes in the mean vector of the main multivariate control charts Hotelling T^2 , MEWMA [10] and MCUSUM [4] for p = 2; n = 1 and ARL_0 = 200 is shown in Figure 1.

As can be seen (Figure 1), the Hotelling T^2 chart is more sensitive to larger changes in the mean vector of the process. On the other hand, for small changes, the MCUSUM and MEWMA charts are more sensitive. The existing difference between the performance of these two charts for this situation (when p = 2 variables) is considered significant, as seen in Figure 1, for change sizes of $d \geq 1$, where the MCUSUM chart is more sensitive.

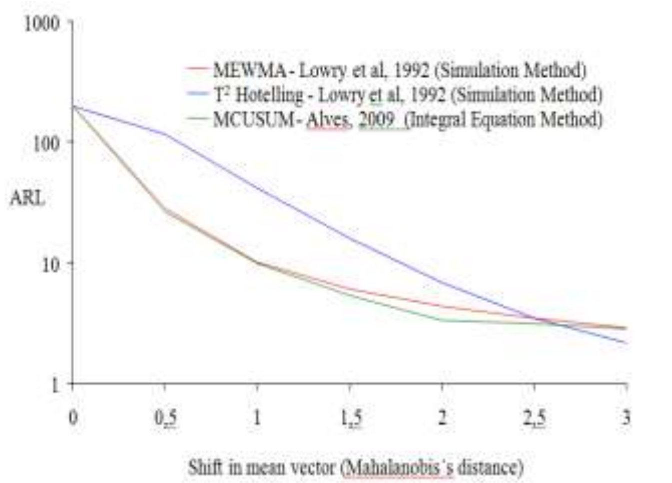

Figure 1 – ARL Performance of the MCUSUM, MEWMA and Hotelling T^2 charts (p = 2; n = 1 and ARL_0 = 200).

4 Methodology and Procedures

The real data used in this article are related to a machining process of a passenger vehicle engine block made of cast iron and machined in a metal works plant located in southern Brazil [4], pg. 192. A simplified diagram considered essential to the understanding of the quality characteristics of the machining process is presented (Figure 2).

According to Figure 2, the position of the center of each hole is defined by the distance to the XY axes with the tolerance delimited by a cylinder with a diameter of 0.16 mm. The axis of each cylinder is perpendicular to the plane formed by the XY. The center of hole 1 is 5 mm away from X, 103.25 mm away from Y , and 194.27 mm away from the center of hole 2 (D1,2).

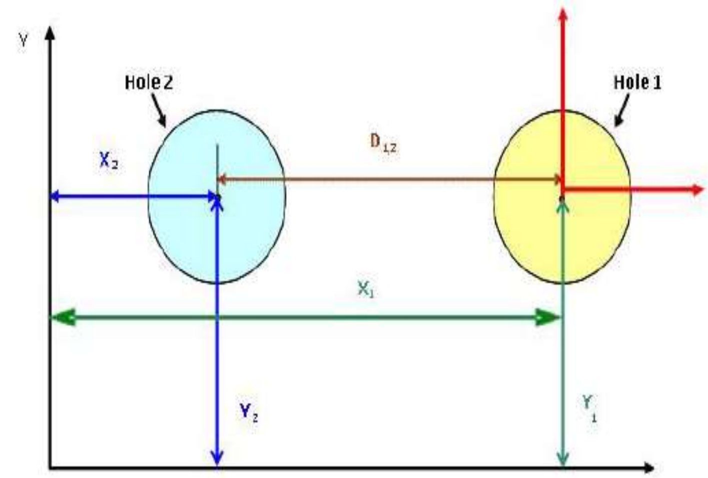

Figure 2 : Diagram of holes 1 and 2 of the engine blocks of passenger vehicles.

The drilling process of the engine block is done in two stages: the pre-drilling operation and the final drilling operation. In this process, the block holes 1 and 2 are references for placing the pieces in the attachment device in every machining operation performed in different machines. In the first machining operation, the block guides milling is performed using a hard metal drill and a reamer. After the initial operations, the final machining of the finished hole is done. The goal of this operation is both the nominal value of the diameter and the nominal value of the positional. The positional is used as a guide to do the final operations in the piece.

In this article, the charts are used only for the data referring to the final drilling operation of hole1. In this process, the quality characteristics to be monitored simultaneously are X_1: positional hole for the X coordinate; and Y_1, positional hole for the Y coordinate. The nominal values (specifications) are respectively 5 ± 0.08mm and 103.25 ± 0.08mm. The choice of this process is justified by the fact that we are dealing with a process in which we can monitor several quality characteristics with small variations in the mean vector of the process.

All of the analyses were performed using R [3] with the qcc [12], MSQC [13], and QRM [14] packages. MCUSUM and MEWMA routines developed in R [15] were applied to generate these charts. We adopted ARL_0 = 200, i.e. false alarm rate α = 0.005, and the reference value k = 0.5. The decision limit h = 5.493 was estimated with the Integral Equation Method [4].

5 Results and Analysis

The performance of the MCUSUM chart in terms of sensitivity in detecting small changes in the mean vector of this process is compared with the Hotelling T^2 charts under the same conditions, i.e. for $p = 2$; $n = 1$ and $ARL_0 = 200$ as shown in Figure 3. It can be seen that the MCUSUM chart (Figure 3) signaled a change in the mean vector at the 10th sample. The Hotelling T^2 control chart (Figure 4) detects the change only in the 18th sample.

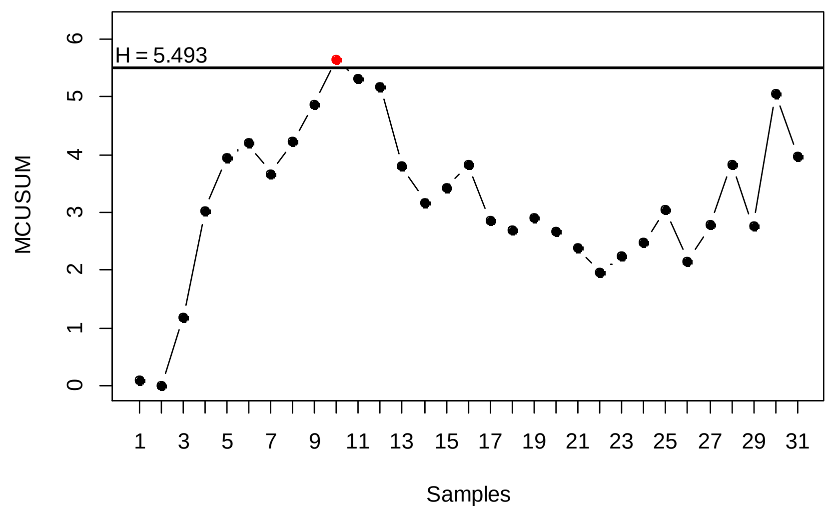

Figure 3: MCUSUM chart created from the process data.

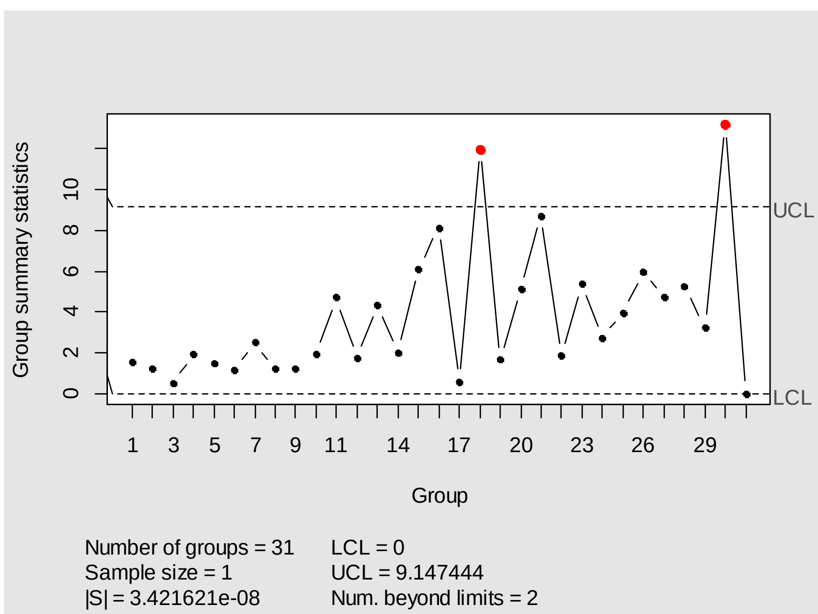

Figure 4: Hotelling T^2 chart created from the process data.

Figure 5 shows the MEWMA control chart applied to the data. It appears that it is more sensitive in detecting changes in the process than the Hotelling T^2 chart. For this case, the MEWMA has the same performance as the MCUSUM chart.

TECNO-LÓGICA
Revista do depto. de Química e Física, do depto. de Engenharia, Arquitetura e Ciências Agrárias e do Mestrado em Tecnologia Ambiental
http://online.unisc.br/seer/index.php/tecnologica

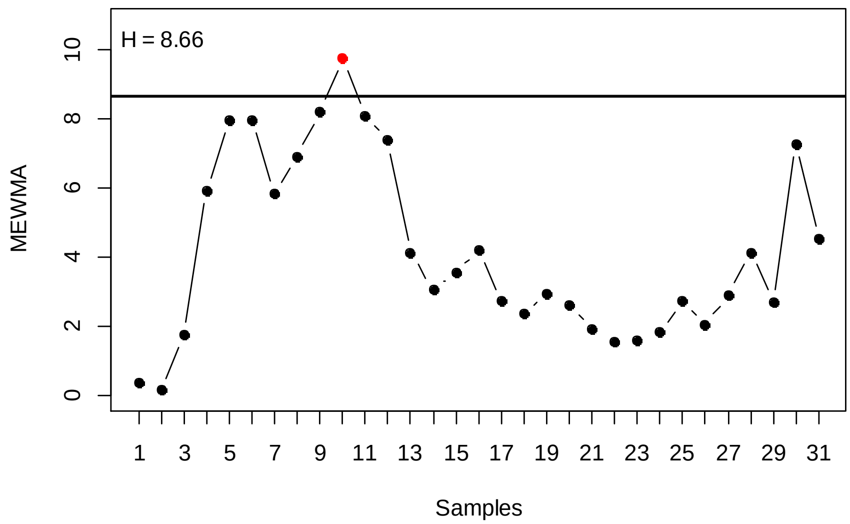

Figure 5: MEWMA chart created from the process data.

The application of these charts requires data with normal distribution and absence of autocorrelation. Normality was verified using Mardia´s test (*p-value* = 0.72711 for asymmetry; *p-value* = 0.2373 for kurtosis), and the normal probability plot [16]. It can be concluded that the data has a normal distribution (Figure 6) and it is not autocorrelated according to the correlograms (Figure 7.a and 7.b).

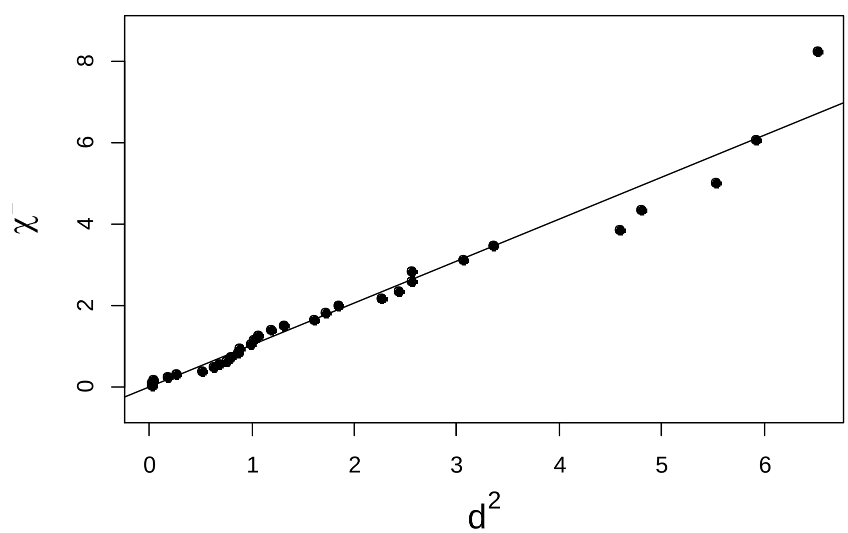

Figure 6: Normal probability plot.

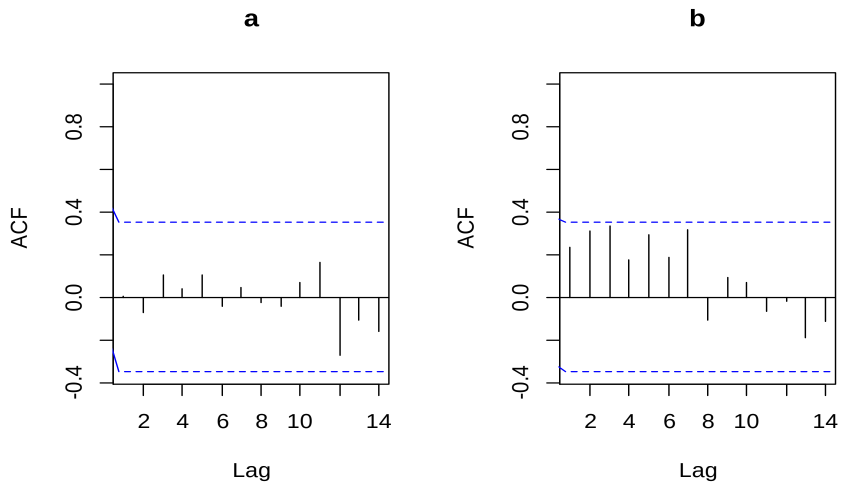

Figure 7: Autocorrelation of the variable X_1(a), autocorrelation of the variable Y_1 (b).

6 Conclusions

The MCUSUM and the MEWMA charts are alternatives to the Hotelling T^2 chart in situations where it is important to detect small changes in the parameters of a process. In such situations, these charts have better performance in terms of ARL and are, therefore, more sensitive than the Hotelling T^2 chart to detect a small change in the mean vector of the process allowing faster action. The results obtained in this study suggest that the MCUSUM chart is an excellent statistical tool for the monitoring of a machining process with multiple quality characteristics. The use of real data was essential for the multivariate control chart theory to be applied to a practical situation.

References

[1] BERSIMIS, S.; PSARAKIS, S.; PANARETOS, J.; Multivariate Statistical Process Control Charts: An Overview, Quality and Reliability Engineering International, vol. 23, p. 517-543, **2007**.

[2] MONTGOMERY, D. C. Introduction to Statistical Quality Control. Rio de Janeiro: LTC, 6th Ed, **2009**.

[3] R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org, **2013**.

[4] ALVES, C.C. The Method of Integral Equation with Gaussian Quadrature to optimize the parameters of the multivariate cumulative sum control chart. Doctorate Thesis, Federal University of Santa Catarina, Florianópolis, Brazil, **2009**.

[5] MASON, R.L., YOUNG, J.C.; Multivariate Statistical Process Control with Industrial Applications. ASA/SIAM: Philadelphia, **2002**.

[6] ALVES, C.C.; SAMOHYL, R.W.; A utilização dos gráficos de controle CUSUM para o monitoramento de processos industriais. In: XXIV National Production Engineering Meeting, Florianópolis, 2004, Brazil. **Anais…**Florianópolis, **2004**.

[7] CROSIER, R.B.; Technometrics, vol. 30, n. 3, p. 291-303, **1998**.

[8] LOWRY, C.; MONTGOMERY, D. C.; A IIE Transactions, vol. 27, p. 800-810, **1995**.

[9] TRACY, N.D.; YOUNG, J. C.; MASON, R.L.; Journal of Quality Technology, vol. 24, n. 2, p. 88-95, **1992**.

[10] LOWRY, C.A.; WOODALL, W.H.; CHAMP, C.W.; RIGDON, S.E.; A, Technometrics, vol. 34, p. 46-53, **1992**.

[11] LEE, M.H.; KHOO, M.B.; International Journal of Reliability, Quality and Safety Engineering, vol. 13, p. 479 – 497, **2006**.

[12] SCRUCCA, L. (2004). qcc: an R package for quality control charting and statistical process control, R News, vol. 4, n.1, 1, p. 11-17, **2004**.

[13] SANTOS-FERNANDEZ, E.; MSQC: Multivariate Statistical Quality Control. R package version 1.0.0. url://http://CRAN.R-project.org/package=MSQC, **2012.**

[14] PFAFF, B.; McNEIL, A.; QRM: Provides R-language code to examine Quantitative Risk Management concepts. R package version 0.4-8. URL: http://CRAN.R-project.org/package=QRMlib, **2012**.

[15] ALVES, C.C.; HENNING, E.; SAMOHYL, R.W.; O desenvolvimento de gráficos de controle MCUSUM e MEWMA em ambiente R como um procedimento alternativo para análise estatística de processos multivariados. In: XXVIII National Production Engineering Meeting. Rio de Janeiro, 2008, Brazil. **Anais**... Rio de Janeiro, **2008**.

[16] MINGOTI, S.A.; Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada. Belo Horizonte: Editora UFMG, **2005**.