Monitoring the Mean Vector and the Covariance Matrix of Bivariate Processes

Marcela Aparecida Guerreiro Machado
Departamento de Produção,
Universidade Estadual Paulista – UNESP, Guaratinguetá, SP, Brazil
E-mail: marcela@feg.unesp.br

Antônio Fernando Branco Costa
Departamento de Produção,
Universidade Estadual Paulista – UNESP, Guaratinguetá, SP, Brazil
E-mail: fbranco@feg.unesp.br

Abstract

This paper proposes the joint use of two charts based on the non-central chi-square statistic (NCS statistic) for monitoring the mean vector and the covariance matrix of bivariate processes, named as the joint NCS charts. The expression to compute the *ARL*, which is defined as the average number of samples the joint charts need to signal an out-of-control condition, is derived. The joint NCS charts might be more sensitive to changes in the mean vector or, alternatively, more sensitive to changes in the covariance matrix, accordingly to the values of their design parameters. In general, the joint NCS charts are faster than the combined T^2 and $|S|$ charts in signaling out-of-control conditions. Once the proposed scheme signals, the user can immediately identify the out-of-control variable. The risk of misidentifying the out-of-control variable is small (less than 5.0%).

Keywords: non-central chi-square statistic, covariance matrix, mean vector, bivariate processes

Introduction

Control charts are often used to observe whether a process is in control or not. When there is only one quality characteristic Shewhart control charts are usually applied to detect process shifts. The power of the Shewhart control charts lies in their ability to separate the assignable causes of variation from the uncontrollable or inherent causes of variation. Shewhart control charts are relatively easy to construct and to interpret. As a result, they are readily implemented in manufacturing environments.

However, there are many situations in which it is necessary to control two or more related quality characteristics simultaneously. Hotelling (1947) provided the first solution to this problem by suggesting the use of the T^2 statistic for monitoring the mean vector of multivariate processes. If compared with the use of simultaneous $\overline{X}$ charts, the T^2 chart is not always faster in signaling process disturbances, see Machado and Costa (2008a). Many innovations have been proposed to improve the performance of the T^2 charts. Recently, Costa and Machado (2007) studied the properties of the synthetic T^2 chart with two-stage sampling. Costa and Machado (2008a) considered the use of the double sampling procedure with the chart proposed by Hotelling.

The first multivariate control chart for monitoring the covariance matrix Σ was based on the charting statistic obtained from the generalized likelihood ratio test. For the case of two variables, Alt (1985) proposed the generalized variance statistic $|S|$ to control the covariance matrix Σ.

Control charts more efficient than the $|S|$ chart have been proposed. Recently, Costa and Machado (2008b, 2008c), Machado and Costa (2008b) and Machado et al. (2008) considered the VMAX statistic to control the covariance matrix of multivariate processes. The points plotted on the VMAX chart correspond to the maximum of the sample variances of the p quality characteristics.

There are a few recent papers dealing with the joint control of the mean vector and the covariance matrix of multivariate processes. Khoo (2005) proposed a control chart based on the T^2 and $|S|$ statistics for monitoring bivariate processes. The speed with which the chart signals changes in the mean vector and/or in the covariance matrix was obtained by simulation. The results are not compelling, once the proposed chart is slow in signaling out-of-control conditions. Chen et al. (2005) proposed a single EWMA chart to control both, the mean vector and the covariance matrix. Their chart is more efficient than the joint T^2 and $|S|$ in signaling small changes in the process. Zhang and Chang (2008) proposed two EWMA charts based on individual observations that are not only fast in signaling but also very efficient in informing which parameter was affected by the assignable cause; if only the mean vector or only the covariance matrix or both.

In practice, the speed with which the control charts detects process changes seems to be more important than their ability in identifying the kind of change. For the univariate case, the use of the non-central chi-square statistic (NCS statistic) for monitoring the mean and the variance of processes simultaneously has been more effective than the joint use of the $\overline{X}$ and R statistics in detecting process changes, see Costa and Rahim (2004, 2006); Costa and De Magalhães (2005, 2007) and Costa et al. (2005).

In this article, we consider the joint use of two charts based on the NCS statistic for monitoring the mean vector and the covariance matrix of bivariate processes. The proposed scheme, named as the joint NCS charts, is an alternative to the joint use of the T^2 and $|S|$ charts. The NCS charts are recommended for those who aim to identify the out-of-control variable and the T^2 and $|S|$ charts are recommended for those who aim to identify the nature of the disturbance, that is, if the assignable cause changes the process mean vector or the covariance matrix.

The success of the recently proposed charts for monitoring the covariance matrix (see Costa and Machado (2008b, 2008c), Machado and Costa (2008b) and Machado et al. (2008)) was the motivation to design new charts to control both the mean vector and the covariance matrix.

The paper is organized as follows. In the second and third sections we present the joint NCS charts and the T^2 and |S| charts, respectively. The joint charts are compared in the fourth section. The mathematical development to obtain the power of the joint NCS charts is in the Appendix. An example is also presented to illustrate the application of the proposed scheme. Finally, the last section concludes the paper, presenting an analysis of the main results.

The Joint NCS Charts

The process is considered to start with the mean vector and the covariance matrix on target ($\mu = \mu_0$ and $\Sigma = \Sigma_0$), where $\mu'_0 = (\mu_x; \mu_y)$ and $\Sigma_0 = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix}$. The occurrence of the assignable cause changes the mean vector from μ_0 to $\mu'_1 = (\mu_x + c\sigma_x; \mu_y + d\sigma_y)$ and/or the covariance matrix from Σ_0 to $\Sigma_1 = \begin{pmatrix} a^2\sigma_x^2 & ab\sigma_{xy} \\ ab\sigma_{xy} & b^2\sigma_y^2 \end{pmatrix}$. The correlation $\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$ is not affected by the assignable cause.

When the joint NCS charts are in use, samples of size n are taken from the process at regular time intervals. Let X_i and Y_i, $i = 1, 2, 3,\ldots, n$ be the measurements of the variables X and Y. Let $\bar{X} = (X_1 + \ldots + X_n)/n$ and $\bar{Y} = (Y_1 + \ldots + Y_n)/n$ be the sample means of the variables X and Y, and let $e(x) = \bar{X} - \mu_x$ and $e(y) = \bar{Y} - \mu_y$ be the difference between the sample means and the target values of the process means. The NCS statistics are given by:

$$T(x) = \sum_{i=1}^{n} (X_i - \mu_x + \xi(x)\sigma_x)^2,$$
$$T(y) = \sum_{i=1}^{n} (Y_i - \mu_y + \xi(y)\sigma_y)^2, \qquad (1)$$

for $\rho \geq 0$, we define:

$$\text{If} \begin{cases} e(x) \geq 0 \text{ and } \begin{cases} e(y) < 0 & \xi(x) = \delta \text{ and } \xi(y) = -\delta \\ e(y) \geq 0 & \xi(x) = \xi(y) = \delta \times \delta_1 \end{cases} \\ \qquad\qquad\qquad \text{than} \\ e(x) < 0 \text{ and } \begin{cases} e(y) < 0 & \xi(x) = \xi(y) = -\delta \times \delta_1 \\ e(y) \geq 0 & \xi(x) = -\delta \text{ and } \xi(y) = \delta \end{cases} \end{cases} \qquad (2)$$

and for $\rho < 0$, we define:

$$\text{If } \begin{cases} e(x) \geq 0 \text{ and } \begin{cases} e(y) < 0 & \xi(x) = \delta \times \delta_1 \text{ and } \xi(y) = -\delta \times \delta_1 \\ e(y) \geq 0 & \xi(x) = \xi(y) = \delta \end{cases} \\ \qquad\qquad\qquad\qquad \text{than} \\ e(x) < 0 \text{ and } \begin{cases} e(y) < 0 & \xi(x) = \xi(y) = -\delta \\ e(y) \geq 0 & \xi(x) = -\delta \times \delta_1 \text{ and } \xi(y) = \delta \times \delta_1 \end{cases} \end{cases} \tag{3}$$

where the parameters δ and δ_1 are positive constants.

When the variables are correlated ($\rho \neq 0$) and $\delta_1 = 1$ (that is, $\xi(x)$ and $\xi(y)$ are discrete variables assuming only two values, $\pm\ \delta$), the joint NCS charts signal changes in the covariance matrix very fast; however, they are slow in signaling changes in the mean vector. The overall performance of the NCS charts improves when $\xi(x)$ and $\xi(y)$ assumes more than two values, for instance $\pm\ a_1$ and $\pm\ a_2$, with $a_2 < a_1$.

Based on that and following Costa et al. (2005), we propose the use of two design parameters, δ and δ_1, with $\delta = a_1$ and $\delta \times \delta_1 = a_2$. If the variables are positively correlated, the best overall performance is reached with $|\xi(x)| = |\xi(y)| = a_2$ (or a_1 if $\rho < 0$), for the cases in which $e(x)$ and $e(y)$ are both positive or both negative. Otherwise, $|\xi(x)| = |\xi(y)| = a_1$ (or a_2 if $\rho < 0$).

If $T(x)$ and/or $T(y)$ falls beyond the control limit CL, the joint NCS charts signal an out-of-control condition, reminding that two NCS charts are used, one for monitoring the X variable and another for monitoring the Y variable. In the Appendix we obtained the expression (A2), which gives the probability of signaling for the joint NCS charts.

The T^2 and |S| Charts

In the next section we compare the joint NCS charts with the joint T^2 and |S| charts. The T^2 chart was introduced by Hotelling (1947) and it is the most common chart used to control the mean vector of multivariate processes.

Consider that two correlated characteristics are being measured simultaneously and, when a sample of size n is taken, we have n values of each characteristic and the $\overline{X}$ vector, which represents the sample average vector for the two characteristics.

The charting statistic

$$T^2 = n\left(\overline{\mathbf{X}} - \mu_0\right)' \Sigma_0^{-1} \left(\overline{\mathbf{X}} - \mu_0\right) \tag{4}$$

is called Hotelling´s T^2 statistic. When the process is in-control, T^2 is distributed as a chi-square variate with two degrees of freedom, that is, $T^2 \sim \chi^2_{p=2}$. Consequently, the control limit for the T^2 chart is $CL = \chi^2_{p=2,\alpha}$, where α is the type I error. When the process is out-of-control, T^2 is distributed as a non-central chi-squared distribution with two degrees of freedom and with non-centrality parameter $\lambda = n(\mu - \mu_0)' \Sigma_0^{-1}(\mu - \mu_0)$, that is, $T^2 \sim \chi^2_{p=2}(\lambda)$.

The first multivariate control chart for monitoring the covariance matrix Σ was based on the charting statistic obtained from the generalized likelihood ratio test (ALT, 1985). For the case of two variables, Alt (1985) proposed the generalized variance |S| statistic to control the covariance matrix Σ. S is the sample covariance matrix

$$\mathbf{S} = \begin{bmatrix} s_x^2 & s_{xy} \\ s_{xy} & s_y^2 \end{bmatrix}.$$

When the process is in-control $\frac{2\cdot(n-1)\cdot|\mathbf{S}|^{1/2}}{|\Sigma_0|^{1/2}}$ is distributed as a chi-square variable with $2n - 4$ degrees of freedom.

Consequently, the control limit for the |S| chart is:

$$CL = \frac{\left(\chi^2_{2n-4,\alpha}\right)^2 \cdot |\Sigma_0|}{4\cdot(n-1)^2} \tag{5}$$

Comparing the Joint Charts

The average run length (*ARL*), which is defined as the average number of samples before a sample point outside the control limits, has been the one of the most important properties associated with the statistical process control charts. Knowledge of the *ARL* for a particular assignable cause (that changes the mean vector and/or the covariance matrix of multivariate processes) allows us to design more effective control charts. When the process is in-control, the *ARL* measures the rate of false alarms. A chart with a larger in-control *ARL* (ARL_0) indicates lower false alarm rate than other charts. A chart with a smaller out-of-control *ARL* indicates a better ability of detecting process shifts than other charts.

The correlation coefficient has a minor influence on the NCS charts performance, see Table 1. For example, considering a and b = 1.25 and c and d = 0.0, if ρ changes from 0.0 to 0.7, the *ARL* increases from 15.9 to 17.5. Tables 2 through 6 provide the *ARL* for the joint NCS charts and for the joint T^2 and |S| charts, where $\rho = 0.0; \pm 0.5; \pm 0.7$, a and b = 1.0; 1.25; 1.5 and c and d = 0.0; 0.5; 0.75; 1.0. A type I risk of 0.5% is adopted. One can see from these tables that the joint NCS charts compete in performance with the joint T^2 and |S| charts. We selected the values of δ and δ_1 in Tables 2 through 6 based on the overall performance of the NCS charts.

Tables 2 through 6 were built considering three different values of a and b and four different values of c and d. The orthogonal array is, in this case, made up of 144 combinations; however, these tables present only one half of the orthogonal array. The explanation is that the symmetric cases ($a = w_1$, $b = w_2$, $c = w_3$ and $d = w_4$) and ($a = w_2$, $b = w_1$, $c = w_4$ and $d = w_3$), with $w_1, w_2 \in \{1,\ 1.25,\ 1.5\}$ and $w_3, w_4 \in \{0,\ 0.5,\ 0.75,\ 1.0\}$, lead to the same *ARL*.

Table 7 shows the effect of δ on the ARL value of the joint NCS charts. Larger values of δ are better for detecting changes in the mean vector with $a = b = 1.0$, and worse for detecting changes in the covariance matrix with $c = d = 0.0$. For example, when $a = b = 1.0$ and $c = d = 0.5$, the *ARL* value decreases from 50.3 to 20.3 as δ increases from 0 to 2.0.

Volume 5, Number 1, 2008, pp. 47-62

Table 1 - Influence of ρ on the *ARL* values for the NCS charts ($\delta = 0.8$; $\delta_1 = 1.0$).

								n = 5							
			c	0	0	0.5	0.5	0	0.75	0.5	0.75	0.75	0	1.0	1.0
			d	0	0.5	0	0.5	0.75	0	0.75	0.5	0.75	1.0	0	1.0
ρ	*a*	*b*													
0.0[1]	1.0	1.0		200.0	41.1	41.1	22.6	14.4	14.4	11.3	11.3	7.7	6.0	6.0	3.3
0.5[2]				200.0	40.8	40.8	24.5	14.3	14.3	12.1	12.1	8.4	5.8	5.8	3.7
0.7[3]				200.0	40.2	40.2	25.7	14.1	14.1	12.7	12.7	8.9	5.7	5.7	4.0
0.0	1.25	1.0		29.5	18.8	11.5	9.6	10.5	6.0	6.9	6.9	4.5	5.2	5.2	2.5
0.5				29.6	19.5	11.2	10.0	10.7	6.0	7.5	5.8	4.9	5.2	3.5	2.8
0.7				29.8	20.1	11.4	10.3	10.7	5.9	7.9	5.9	5.2	5.3	3.5	3.0
0.0	1.5	1.0		8.1	7.3	5.2	4.8	5.6	3.5	4.0	3.4	3.1	3.8	2.5	2.0
0.5				8.2	7.3	5.2	4.9	5.8	3.6	4.5	3.5	3.3	3.9	2.5	2.3
0.7				8.2	7.4	5.2	5.0	5.9	3.6	4.6	3.5	3.4	3.9	2.5	2.3
0.0	1.25	1.25		15.9	8.8	8.8	6.1	5.3	5.3	4.3	4.3	3.4	3.3	3.3	2.1
0.5				16.5	8.9	8.9	6.7	5.4	5.4	4.7	4.7	3.8	3.3	3.3	2.4
0.7				17.5	9.2	9.2	7.1	5.4	5.4	4.9	4.9	4.0	3.3	3.3	2.5
0.0	1.25	1.5		6.8	4.7	5.2	3.9	3.4	3.8	3.0	3.1	2.5	2.4	2.7	1.8
0.5				7.0	4.7	5.4	4.1	3.4	4.0	3.2	3.4	2.8	2.5	2.8	2.0
0.7				7.3	4.8	5.5	4.4	3.4	4.1	3.3	4.1	3.0	2.4	2.8	2.1
0.0	1.5	1.5		4.4	3.5	3.5	2.9	2.7	2.7	2.4	2.4	2.1	2.2	2.2	1.6
0.5				4.7	3.6	3.6	3.2	2.8	2.8	2.7	2.7	2.4	2.2	2.2	1.8
0.7				5.1	3.8	3.8	3.4	2.9	2.9	2.8	2.8	2.5	2.2	2.2	1.9

[1] *CL*=29.4; [2] *CL*=29.3; and [3] *CL*=29.2.

Table 2 - *ARL* values for the joint T^2 and $|S|$ charts and NCS charts ($\rho = 0.0$; $\delta = 0.8$; $\delta_1 = 1.0$; $CL = 29.4$).

							n = 5							
		c	0	0	0.5	0.5	0	0.75	0.5	0.75	0.75	0	1.0	1.0
		d	0	0.5	0	0.5	0.75	0	0.75	0.5	0.75	1.0	0	1.0
a	*b*													
1.0	1.0		200.0	48.9*	49.2	19.9	16.7	16.8	9.5	9.5	5.4	6.8	6.6	2.3
			200.0	41.1**	41.1	22.6	14.4	14.4	11.3	11.3	7.7	6.0	6.0	3.3
1.25	1.0		42.6	22.2	17.5	11.0	11.2	8.9	6.7	6.3	4.4	5.2	4.7	2.1
			29.5	18.8	11.5	9.6	10.5	6.0	6.9	6.9	4.5	5.2	5.2	2.5
1.5	1.0		15.3	10.7	9.0	6.7	6.9	5.8	4.8	4.6	3.5	4.0	3.7	2.0
			8.1	7.3	5.2	4.8	5.6	3.5	4.0	3.4	3.1	3.8	2.5	2.0
1.25	1.25		15.6	9.9	9.8	6.8	6.2	6.3	4.7	4.7	3.5	3.9	3.9	2.0
			15.9	8.8	8.8	6.1	5.3	5.3	4.3	4.3	3.4	3.3	3.3	2.1
1.25	1.5		15.3	5.5	5.8	4.4	4.1	4.3	3.4	3.4	2.8	3.0	3.0	1.8
			6.8	4.7	5.2	3.9	3.4	3.8	3.0	3.1	2.5	2.4	2.7	1.8
1.5	1.5		4.4	3.6	3.6	3.1	3.0	3.0	2.6	2.6	2.3	2.4	2.4	1.7
			4.4	3.5	3.5	2.9	2.7	2.7	2.4	2.4	2.1	2.2	2.2	1.6

*T^2 and $|S|$ charts; **NCS charts.

Table 3 - *ARL* values for the joint T^2 and |S| charts and NCS charts (ρ = 0.5; δ = 1.2; δ_1= 0.75; *CL* = 32.6).

							n = 5						
		c	0	0	0.5	0.5	0	0.75	0.5	0.75	0.75	0	1.0	1.0
		d	0	0.5	0	0.5	0.75	0	0.75	0.5	0.75	1.0	0	1.0
a	*b*													
1.0	1.0		200.0	38.4*	38.5	34.8	10.8	10.8	15.9	15.9	10.6	4.1	4.2	4.1
			200.0	33.0**	33.0	29.5	11.0	11.0	13.8	13.8	9.9	4.6	4.6	4.2
1.25	1.0		41.6	17.7	14.1	15.8	7.6	6.6	10.1	8.7	7.1	3.5	3.5	3.5
			29.6	15.8	10.2	11.3	8.1	5.1	8.3	6.3	5.6	4.0	3.0	3.0
1.5	1.0		14.6	8.9	7.7	8.4	5.2	4.6	6.4	5.7	5.0	3.0	2.9	3.0
			8.0	6.5	4.9	5.3	4.7	3.3	4.6	3.8	3.5	3.0	2.3	2.4
1.25	1.25		15.9	8.7	8.6	8.6	4.9	4.8	6.1	6.1	4.9	2.9	2.9	2.9
			16.7	8.0	8.0	7.3	4.5	4.5	5.0	5.0	4.1	2.8	2.8	2.5
1.25	1.5		7.3	5.1	5.2	5.1	3.5	3.6	4.1	4.2	3.6	2.5	2.4	2.4
			6.9	4.4	4.8	4.4	3.0	3.4	3.3	3.6	3.0	2.2	2.4	2.1
1.5	1.5		4.3	3.5	3.5	3.4	2.7	2.7	2.9	2.9	2.7	2.1	2.1	2.1
			4.6	3.4	3.4	3.3	2.6	2.6	2.7	2.7	2.4	2.0	2.0	1.8

*T^2 and |S| charts; **NCS charts.

Table 4 - *ARL* values for the joint T^2 and |S| charts and NCS charts (ρ = 0.7; δ = 2.0; δ_1 = 0.7; *CL* = 45.75).

							n = 5						
		c	0	0	0.5	0.5	0	0.75	0.5	0.75	0.75	0	1.0	1.0
		d	0	0.5	0	0.5	0.75	0	0.75	0.5	0.75	1.0	0	1.0
a	*b*													
1.0	1.0		200.0	20.6*	20.2	39.7	5.7	5.6	16.5	16.5	12.9	2.3	2.3	5.0
			200.0	21.2**	21.2	34.1	6.9	6.9	15.2	15.2	11.1	3.0	3.0	4.7
1.25	1.0		39.1	11.9	9.8	17.1	4.6	4.2	10.6	8.7	8.2	2.2	2.2	4.1
			33.6	11.5	8.7	13.3	5.2	4.1	9.1	6.9	6.3	2.7	2.4	3.4
1.5	1.0		13.1	6.9	5.8	8.3	3.6	3.4	6.4	5.6	5.3	2.0	2.1	3.3
			9.4	5.6	4.6	6.0	3.6	2.9	5.1	4.1	3.9	2.2	2.0	2.6
1.25	1.25		15.7	6.9	6.8	9.2	3.5	3.5	6.2	6.2	5.4	2.0	2.0	3.3
			19.3	6.9	6.9	8.7	3.6	3.6	5.7	5.7	4.8	2.1	2.1	2.9
1.25	1.5		7.2	4.4	4.5	5.3	2.8	2.8	4.0	4.2	3.8	1.9	1.8	2.7
			8.1	4.2	4.6	5.2	2.7	2.9	3.8	2.9	3.5	1.9	1.9	2.4
1.5	1.5		4.3	3.1	3.1	3.6	2.3	2.3	3.0	3.0	2.9	1.7	1.7	2.2
			5.5	3.4	3.4	3.9	2.3	2.3	3.1	3.1	2.9	1.7	1.7	2.1

*T^2 and |**S**| charts; **NCS charts.

On the other hand, when $c = d = 0.0$ and $a = b = 1.25$, the *ARL* value increases from 13.8 to 22.3 as δ increases from 0 to 2.0.

Table 8 shows the effect of δ_1 on the *ARL* value of the joint NCS charts. In general, larger values of δ_1 are better for detecting changes in the mean vector when both variables are affected by the assignable cause and smaller values of δ_1 are better for detecting changes in the mean vector when only one variable is affected by the assignable cause.

Table 5 - *ARL* values for the joint T^2 and |S| charts and NCS charts ($\rho = -0.5$; $\delta = 1.2$; $\delta_1 = 0.75$; $CL = 32.6$).

							n = 5							
		c	0	0	0.5	0.5	0	0.75	0.5	0.75	0.75	0	1.0	1.0
		d	0	0.5	0	0.5	0.75	0	0.75	0.5	0.75	1.0	0	1.0
a	*b*													
1.0	1.0		200.0	34.9*	34.9	6.7	10.6	10.6	3.2	3.2	1.9	4.1	4.1	1.1
			200.0	33.0**	33.0	10.1	11.0	11.0	5.0	5.0	3.2	4.6	4.6	1.6
1.25	1.0		41.6	17.7	14.1	5.0	7.6	6.6	2.8	2.8	1.9	3.5	3.5	1.2
			29.6	15.8	10.2	5.5	8.1	5.1	3.7	3.4	2.5	4.0	3.0	1.5
1.5	1.0		14.6	8.9	7.7	3.9	5.2	4.6	2.5	2.5	1.8	3.0	2.9	1.2
			8.0	6.5	4.9	3.5	4.7	3.3	2.7	2.5	2.0	3.0	2.3	1.4
1.25	1.25		15.9	8.7	8.6	3.8	4.9	4.8	2.5	2.5	1.8	2.9	2.9	1.2
			16.7	8.0	8.0	4.0	4.5	4.5	2.7	2.7	2.0	2.8	2.8	1.4
1.25	1.5		7.2	5.0	5.3	3.0	3.5	3.6	2.2	2.2	1.7	2.5	2.5	1.2
			6.9	4.4	4.8	2.9	3.1	3.4	2.2	2.2	1.8	2.2	2.4	1.3
1.5	1.5		4.3	3.5	3.5	2.4	2.7	2.7	1.9	1.9	1.6	2.1	2.1	1.2
			4.6	3.4	3.4	2.4	2.6	2.6	1.9	1.9	1.6	2.0	2.0	1.2

*T^2 and |S| charts; **NCS charts.

Table 6 - *ARL* values for the joint T^2 and |S| charts and NCS charts ($\rho = -0.7$; $\delta = 2.0$; $\delta_1 = 0.7$; $CL = 45.75$).

							n = 5							
		c	0	0	0.5	0.5	0	0.75	0.5	0.75	0.75	0	1.0	1.0
		d	0	0.5	0	0.5	0.75	0	0.75	0.5	0.75	1.0	0	1.0
a	*b*													
1.0	1.0		200.0	20.6*	20.2	2.9	5.7	5.6	1.6	1.6	1.2	2.3	2.3	1.0
			200.0	21.2**	21.2	3.9	6.9	6.9	2.2	2.2	1.5	3.0	3.0	1.1
1.25	1.0		39.1	11.9	9.8	2.6	4.6	4.2	1.6	1.6	1.2	2.2	2.2	1.0
			33.6	11.5	8.7	3.0	5.2	4.1	2.0	1.9	1.4	2.7	2.4	1.1
1.5	1.0		13.1	6.9	5.8	2.4	3.6	3.4	1.6	1.7	1.3	2.0	2.1	1.0
			9.4	5.6	4.6	2.4	3.6	2.9	1.8	1.7	1.4	2.2	2.0	1.1
1.25	1.25		15.7	6.9	6.8	2.3	3.5	3.5	1.6	1.6	1.2	2.0	2.0	1.0
			19.3	6.9	6.9	2.5	3.6	3.6	1.7	1.7	1.4	2.1	2.1	1.1
1.25	1.5		7.4	4.3	4.5	2.1	2.8	2.8	1.5	1.5	1.2	1.9	1.8	1.0
			8.2	4.2	4.6	2.1	2.7	2.9	1.6	1.6	1.3	1.9	1.9	1.1
1.5	1.5		4.3	3.1	3.1	1.9	2.3	2.3	1.4	1.4	1.2	1.7	1.7	1.0
			5.5	3.4	3.4	1.9	2.3	2.3	1.5	1.5	1.3	1.7	1.7	1.1

*T^2 and |S| charts; **NCS charts.

Table 9 presents the values of *Pv*, which corresponds to the probability of the control chart signaling that the assignable cause affects the mean and/or the variance of the *X* variable (given by $T(x) > CL$) when in reality it affects the mean and/or the variance of the *Y* variable. From this table we can observe that the probability of the chart erroneously signaling is small (less than 5.0 %). The *Pv* values are obtained by the expression (A3) in the Appendix.

Table 7 - The influence of δ on the *ARL* of the joint NCS charts ($\rho = 0.5$).

					δ_1 = **1.0**			
			δ	0	0.5	0.7	1.0	2.0
			CL	18.33	24.10	27.50	33.50	61.12
a	*b*	c	d					
1.0	1.0	0	0	200.0	200.0	200.0	200.0	200.0
		0	0.5	76.7	47.8	42.6	38.7	34.9
		0.5	0.5	50.3	28.5	25.5	23.1	20.3
		0	1.0	12.0	6.8	6.1	5.6	4.9
		1.0	1.0	7.2	4.2	3.9	3.5	3.1
1.25	1.0	0	0	24.5	26.5	29.0	31.1	38.2
		0	0.5	21.3	19.5	19.4	19.6	20.7
		0.5	0	13.7	11.7	11.6	11.3	12.1
		0.5	0.5	12.8	10.5	10.3	9.8	10.3
		0	1.0	8.8	5.9	5.5	5.0	4.6
		1.0	0	4.9	3.8	3.6	3.5	3.4
		1.0	1.0	4.1	3.0	2.8	2.7	2.6
1.25	1.25	0	0	13.8	15.2	16.8	18.1	22.3
		0	0.5	9.7	8.8	9.1	9.2	9.9
		0.5	0.5	7.8	6.9	6.7	6.7	7.2
		0	1.0	4.3	3.5	3.4	3.3	3.3
		1.0	1.0	3.0	2.4	2.4	2.3	2.3

Table 8 - The influence of δ_1 on the *ARL* of the joint NCS charts ($\rho = 0.5$).

				δ = **1.0**			
			δ_1	0.5	0.75	1.0	2.0
			CL	26.85	29.15	33.50	61.00
a	*b*	c	d				
1.0	1.0	0	0	200.0	200.0	200.0	200.0
		0	0.5	34.8	35.7	38.7	38.1
		0.5	0.5	50.6	28.6	23.1	20.5
		0	1.0	5.1	4.9	5.6	6.3
		1.0	1.0	6.6	4.2	3.5	3.1
1.25	1.0	0	0	28.5	27.4	31.1	42.1
		0	0.5	15.5	16.2	19.6	24.4
		0.5	0	10.5	10.4	11.3	14.2
		0.5	0.5	14.6	10.8	9.8	10.1
		0	1.0	4.2	4.3	5.0	6.1
		1.0	0	3.2	3.1	3.5	4.6
		1.0	1.0	4.1	3.0	2.7	2.6
1.25	1.25	0	0	16.2	15.5	18.1	23.6
		0	0.5	8.2	7.9	9.2	11.9
		0.5	0.5	8.8	7.0	6.7	7.5
		0	1.0	2.9	2.8	3.3	4.4
		1.0	1.0	3.1	2.5	2.3	2.3

Volume 5, Number 1, 2008, pp. 47-62

Table 9 - *Pv* values for the joint NCS charts (%) (ρ = 0.5; δ = 1.2; δ_1 = 0.75; *CL* = 32.6).

				***n* = 5**			
		c	0	0	0	0	0
		d	0.5	0.75	1.0	1.5	2.0
a	*b*						
1.0	1.25		5.0	3.0	1.9	1.2	1.0
1.0	1.5		2.0	1.6	1.4	1.0	0.9
1.0	1.75		1.2	1.1	0.9	0.8	0.7
1.0	2.0		0.9	0.9	0.8	0.7	0.6

Example

In this section we provide an example to illustrate the use of the joint NCS charts. When the process is in-control, the mean vector and the covariance matrix are given by $\mu_0' = (0,0)$ and $\Sigma_0 = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$, respectively.

We initially generate 5 samples of size $n = 5$ with the process in control. The remaining 5 samples were simulated considering that the assignable cause changed the mean and the variability of X, that is, $c = 1.0$ and $a = 1.25$.

Table 10 presents the data of X, Y, $T(x)$, $T(y)$, T^2 and $|S|$. Figure 1 shows the joint NCS charts with design parameters $\delta = 1.2$, $\delta_1 = 0.75$ and $CL = 32.6$ (according to Table 2, $\alpha = 0.5\%$). Figure 2 shows the joint T^2 and $|S|$ charts. The joint NCS charts signal an

Table 10 - Values of X, Y, $T(x)$, $T(y)$, T^2 and $|S|$.

Sample		**Observations**									
		1	2	3	4	5		*T*(*x*)	*T*(*y*)	T^2	|S|
1	*X*	0.53	–1.83	0.20	0.89	–0.80		10.96	20.06	5.45	0.51
	Y	–0.27	–1.71	–0.10	–1.30	–1.55					
2		–1.63	–0.86	–0.25	0.78	–1.30		15.70	11.31	2.47	0.13
		–0.54	–0.51	–0.93	0.14	–0.93					
3		–0.27	0.12	–0.10	–0.73	–1.05		9.41	5.90	0.93	0.06
		–0.66	0.71	–0.18	0.22	–0.43					
4		–0.27	–0.68	–0.59	–1.60	–0.23		13.66	11.97	2.55	0.09
		–1.01	–0.31	–0.17	–1.38	0.18					
5		–0.07	0.77	2.02	–0.56	0.15		17.75	14.66	3.34	0.54
		0.48	–0.27	0.07	–1.66	–0.39					
6		2.20	0.05	1.07	–0.22	0.38		21.72	10.27	3.38	0.39
		0.52	–0.85	1.07	–0.73	–0.17					
7		0.83	0.82	0.76	1.93	–1.15		21.87	9.68	3.10	0.47
		0.97	–0.11	–0.44	–0.60	–0.20					
8		1.25	3.63	1.68	–0.10	–2.12		39.68	9.94	5.27	1.17
		0.14	0.53	0.54	–0.12	–1.29					
9		2.59	1.90	0.80	0.68	–1.87		32.00	27.93	16.89	1.39
		–0.77	–0.51	–0.81	–0.34	–2.64					
10		–0.14	1.01	1.13	1.52	2.33		31.27	13.30	10.95	0.58
		–0.71	–1.61	0.59	0.37	0.35					

out-of-control condition at sample 8. We can observe from Figure 1 that the variable X was the responsible for the out-of-control signal. The joint T^2 and $|S|$ charts signal an out-of-control condition at sample 9. According to Figure 2, the T^2 chart was the responsible for the signal; however, it is not possible to identify the variable that had the parameter(s) affected by the assignable cause.

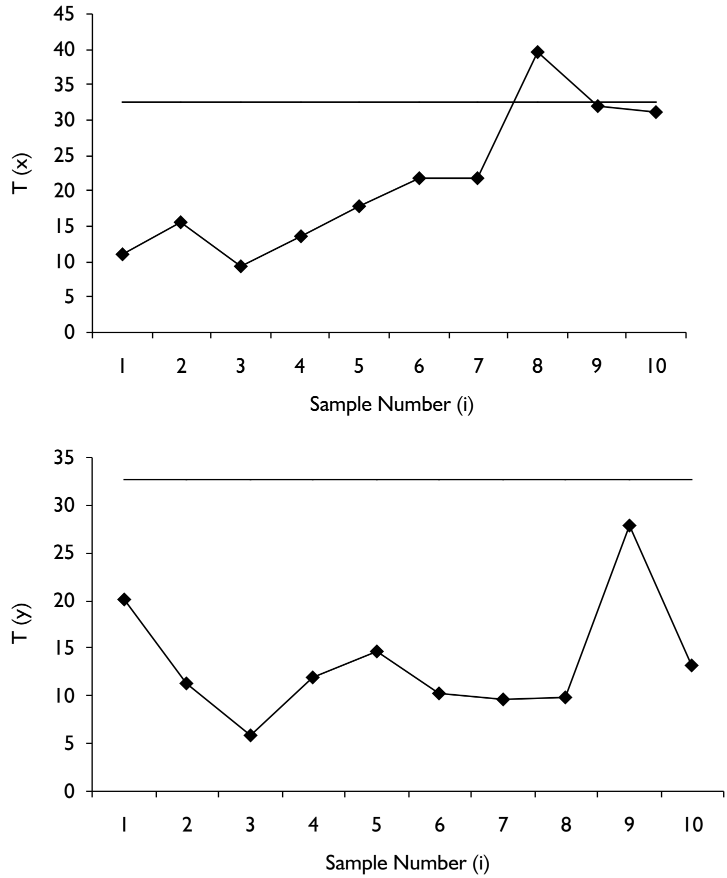

Figure 1 - Joint NCS charts – example.

Conclusions

In this article it is proposed the joint use of two charts based on the non-central chi-square statistic (NCS statistic) for monitoring the mean vector and the covariance matrix of bivariate processes. The way the NCS statistics were defined allowed to obtain the expression to compute *ARL*, which is defined as the average number of samples the joint charts need to signal an out-of-control condition. The joint NCS charts might be more sensitive to changes in the mean vector or, alternatively, more sensitive to changes in the covariance matrix, accordingly to the values of their design parameters. The proposed scheme is an alternative to the joint use of the T^2 and $|S|$ charts which, in general, is faster

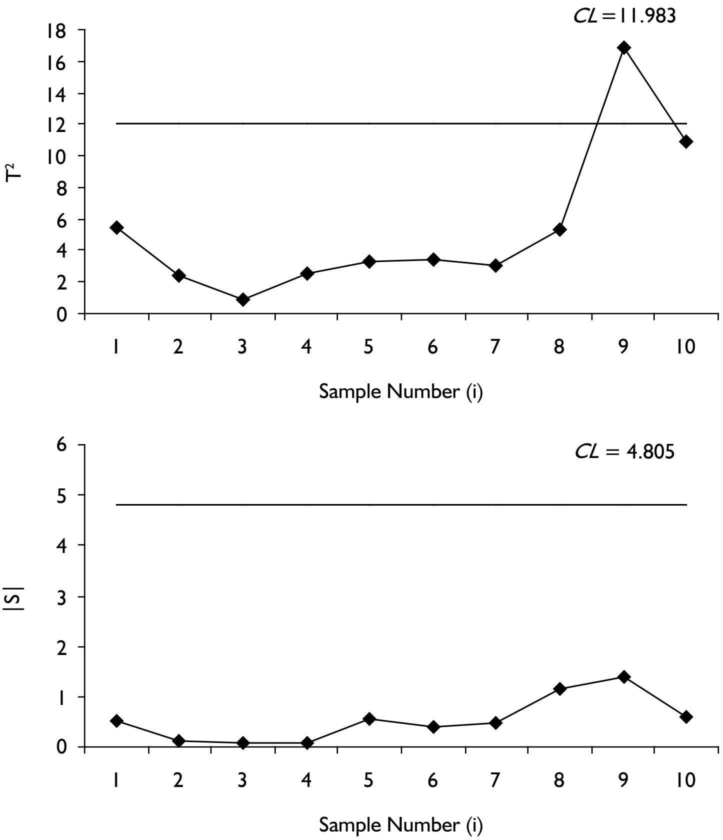

Figure 2 - Joint T^2 and |S| charts – example.

in signaling out-of-control conditions. The joint NCS charts are recommended for those who aim to identify the out-of-control variable instead of the parameter that was affected by the assignable cause: if only the mean vector or only the covariance matrix or both. The risk of the joint NCS charts misidentify the out-of-control variable is small (less than 5.0%).

Acknowledgements

This work was supported by CNPq – National Council for Scientific and Technological Development, Project 307744/2006-0 - and FAPESP - The State of São Paulo Research Foundation, Research Project 2006/00491-0. The authors would like to thank the Editor who handled the review process for this article, as well as the two anonymous referees who carefully read the earlier draft and made many constructive suggestions.

References

Alt, F.B. (1985), Multivariate Quality Control. In: Kotz, S., Johnson, N. L., eds. Encyclopedia of Statistical Science, Vol. 6, No.1, pp. 110-122.

Chen, G. & Cheng, S.W. & Xie, H. (2005), "A New Multivariate Control Chart for Monitoring Both Location and Dispersion", Communications in Statistics-Simulation and Computation, Vol. 34, No. 1, pp. 203-217.

Costa, A.F.B. and De Magalhães, M.S. (2005), "O Uso da Estatística de Qui-Quadrado no Monitoramento de Processos", Gestão & Produção, Vol. 12, No. 2, pp. 271-277.

Costa, A.F.B. and De Magalhães, M.S. (2007), "An Adaptative Chart for Monitoring the Process Mean and Variance", Quality and Reliability Engineering International, Vol. 23, No. 4, pp. 821-831.

Costa, A.F.B. and Machado, M.A.G. (2007), "Synthetic Control Chart with Two-Stage Sampling for Monitoring Bivariate Processes", Pesquisa Operacional, Vol. 27, No. 1, pp. 117-130.

Costa, A.F.B. and Machado, M.A.G. (2008a), "Bivariate Control Charts with Double Sampling", Journal of Applied Statistics. Vol. 35, No. 7, pp. 809-822.

Costa, A. F. B. and Machado, M. A. G. (2008b), A New Chart Based on Sample Variances for Monitoring the Covariance Matrix of Multivariate Processes. International Journal of Advanced Manufacturing Technology, in press.

Costa, A. F. B. and Machado, M. A. G. (2008c), A New Multivariate Control Chart for Monitoring the Covariance Matrix of Bivariate Processes. Communications in Statistics-Simulation and Computation, in press.

Costa, A.F.B. and Rahim, M. A. (2004), "Monitoring Process Mean and Variability with One Non-Central Qui-Square Chart", Journal of Applied Statistics, Vol. 31, No. 10, pp. 1171-1183.

Costa, A.F.B. and Rahim, M. A. (2006), "The Non-Central Chi-Square Chart with Two Stage Samplings", European Journal of Operation Research, Vol. 171, No. 1, pp. 64-73.

Costa, A.F.B. & De Magalhães, M. S. & Epprecht E.K. (2005), "The Non-Central Chi-Square Chart with Double Sampling", Brazilian Journal of Operations & Production Management, Vol. 2, No. 2, pp. 21-37.

Hotelling, H. (1947), Multivariate Quality Control, Illustrated by the Air Testing of Sample Bombsights, Techniques of Statistical Analysis, pp. 111-184.

Khoo, B.C. (2005), "A New Bivariate Control Chart to Monitor the Multivariate Process Mean and Variance Simultaneously", Quality Engineering. Vol. 17, No. 1, pp. 109-118.

Machado, M.A.G. and Costa, A.F.B. (2008a), "The Use of Principal Components and Univariate Charts to Control Multivariate Processes", Pesquisa Operacional, Vol. 28, No. 1, pp. 173-196.

Machado, M.A.G. and Costa, A.F.B. (2008b), "The Double Sampling and the EWMA Charts Based on the Sample Variances", International Journal of Production Economics. Vol. 114, No. 1, pp. 134-148.

Volume 5, Number 1, 2008, pp. 47-62

Machado, M.A.G. & De Magalhães, M.S. & Costa, A. F. B. (2008), "Gráfico de Controle de VMAX Para o Monitoramento da Matriz de Covariâncias", Produção, Vol. 18, No. 2, pp. 222-239.

Microsoft Fortran Power Station 4.0. (1995), Professional Edition with Microsoft IMSL Mathematical and Statistical Libraries, Microsoft Corporation, Washington, USA.

Mood, A.M. & Graybill, F.A. & Boes, D.C. (1974), Introduction to the Theory of Statistics, McGraw-Hill.

Zhang, G. and Chang, S.I. (2008), Multivariate EWMA Control Charts Using Individual Observations for Process Mean and Variance Monitoring and Diagnosis. International Journal of Production Research, in press.

Biography

Marcela Aparecida Guerreiro Machado holds an MSc in Mechanical Engineering and is now a PhD student in the Department of Production at UNESP – São Paulo State University, Brazil. Her main areas of interest are statistical quality control and design of experiments. She has published papers in the Journal of Applied Statistics, International Journal of Production Economics, Communications in Statistics, *Pesquisa Operacional*, *Produção* and International Journal of Advanced Manufacturing Technology. She was the recipient of a best paper award in celebration of the 40th SBPO.

Dr Antonio Fernando Branco Costa is an Associate Professor in the Department of Production at UNESP – São Paulo State University, Brazil. He was a postdoctoral fellow in the Center for Quality and Productivity Improvement at University of Wisconsin, Madison, USA. His current interest is in statistical quality control. He has published almost half hundred papers in the Brazilian Journal of Operations & Production Management, Gestão e Produção, Produção, Pesquisa Operacional, Journal of Quality Technology, European Journal of Operational Research, IIE Transactions, Journal of Applied Statistics, International Journal of Production Economics, International Journal of Production Research, Journal of Quality Maintenance in Engineering, Quality Technology and Quantitative Management, International Journal of Advanced Manufacturing Technology, Quality and Reliability Engineering International and Communications in Statistics. He is an active reviewer for several journals and an ASQ Certified Quality Engineer. He was the recipient of an IIE Transactions best paper award.

Submitted: 28 May, 2008.
Accepted: 20 October, 2008.

Appendix: the Probability of *T* (*x*) and/or *T* (*y*) Exceeding the Control Limit.

When the process is in-control the covariance matrix is given by $\Sigma_0 = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{pmatrix}$. The assignable cause changes the mean vector from μ_0 to $\mu_1' = \mu_x + c\sigma_x; \mu_y + d\sigma_y$ and/or changes the covariance matrix from Σ_0 to $\Sigma_1 = \begin{pmatrix} a^2\sigma_x^2 & ab\sigma_{xy} \\ ab\sigma_{xy} & b^2\sigma_y^2 \end{pmatrix}$. We consider that the assignable cause does not affect the correlation between *X* and *Y*, given by $\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$.

Let X_i and Y_i, $i = 1, 2, 3, \ldots, n$, be the measurements of the variables *X* and *Y* arranged in groups of size $n > 1$. Let $\bar{X} = (X_1 + \ldots + X_n)/n$ and $\bar{Y} = (Y_1 + \ldots + Y_n)/n$ be the sample means of the variables *X* and *Y*, and let $e(x) = \bar{X} - \mu_x$ and $e(y) = \bar{Y} - \mu_y$ be the difference between the sample means and the target values of the process means. The two-non central statistics are given by:

$$T(x) = \sum_{i=1}^{n} (X_i - \mu_x + \xi(x)\sigma_x)^2,$$

$$T(y) = \sum_{i=1}^{n} (Y_i - \mu_y + \xi(y)\sigma_y)^2,$$

after some manipulation, we have that:

$$\frac{T(x)}{a^2\sigma_x^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{a^2\sigma_x^2} + \left(Z_{\bar{x}} + \frac{c + \xi(x)}{a}\sqrt{n} \right)^2$$

$$\frac{T(y)}{b^2\sigma_y^2} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{b^2\sigma_y^2} + \left(Z_{\bar{y}} + \frac{d + \xi(y)}{b}\sqrt{n} \right)^2$$

where $Z_{\bar{x}} = \sqrt{n}\left(\frac{\bar{X} - \mu_x - c\sigma_x}{a\sigma_x} \right)$ and $Z_{\bar{y}} = \sqrt{n}\left(\frac{\bar{Y} - \mu_y - d\sigma_y}{b\sigma_y} \right)$.

Consequently:

$$Ps(Z_{\bar{x}}, Z_{\bar{y}}) = \Pr\left[\left(T(x) < CL\sigma_x^2\right) \cap \left(T(y) < CL\sigma_y^2\right) \,\middle|\, Z_{\bar{x}}, Z_{\bar{y}} \right] =$$

$$= \Pr\left[\left(\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{a^2\sigma_x^2} < CL_x \right) \cap \left(\frac{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}{b^2\sigma_y^2} < CL_y \right) \,\middle|\, Z_{\bar{x}}, Z_{\bar{y}} \right]$$

where $CL_x = \frac{CL}{a^2} - \left(Z_{\bar{x}} + \frac{c + \xi(x)}{a}\sqrt{n} \right)^2$ and $CL_y = \frac{CL}{b^2} - \left(Z_{\bar{y}} + \frac{d + \xi(y)}{b}\sqrt{n} \right)^2$.

If *X* and *Y* are normally distributed we have,

$$[y_i - (\mu_y + d\sigma_y)] \,|\, x_1, x_2, \ldots, x_n \sim N\left(\rho \frac{b\sigma_y}{a\sigma_x}[x_i - (\mu_x + c\sigma_x)]; b^2\sigma_y^2(1 - \rho^2) \right)$$

or

$$\frac{[y_i - (\mu_y + d\sigma_y)] + \rho \frac{b\sigma_y}{a\sigma_x}[(\mu_x + c\sigma_x) - \overline{X}]}{b\sigma_y\sqrt{1-\rho^2}} \Big| x_1, x_2, \ldots, x_n \sim N\left(\frac{\rho}{\sqrt{1-\rho^2}}(\frac{x_i - \overline{X}}{a\sigma_x}), 1\right)$$

As $\rho \frac{b\sigma_y}{a\sigma_x}[(\mu_x + c\sigma_x) - \overline{X}] = (\mu_y + d\sigma_y) - \overline{Y}$, we have (see Mood et al. (1974), page 168):

$$(\frac{y_i - \overline{Y}}{b\sigma_y})\frac{1}{\sqrt{1-\rho^2}} \Big| x_1, x_2, \ldots, x_n \sim N\left(\frac{\rho}{\sqrt{1-\rho^2}}(\frac{x_i - \overline{X}}{a\sigma_x}), 1\right)$$

consequently,

$$\frac{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}{b^2\sigma_y^2\sqrt{1-\rho^2}} \Big| x_1, x_2, \ldots, x_n = \sum_{i=1}^{n}\left(\frac{y_i - \overline{Y}}{b\sigma_y\sqrt{1-\rho^2}}\right)^2 \Big| x_1, x_2, \ldots, x_n \sim \chi^2_{n,(\rho^2/1-\rho^2)\chi^2_{n-1}}$$

as

$$\frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{a^2\sigma_x^2} = \sum_{i=1}^{n}\left(\frac{x_i - \overline{X}}{a\sigma_x}\right)^2 \sim \chi^2_{n-1}$$

then

$$P_S(Z_{\bar{x}}, Z_{\bar{y}}) = \int_0^{CL_x} g(t)dt = \int_0^{CL_x} \Pr\left[\chi^2_{n,(t\rho^2/1-\rho^2)} < \frac{CL_y}{(1-\rho^2)}\right]\frac{1}{2^{(n-1)/2}\,\Gamma[(n-1)/2]}e^{-t/2}t^{[(n-1)/2]-1}dt \quad \text{(A1)}$$

recalling that the notation $\chi^2_{n,(\rho^2/1-\rho^2)\chi^2_n}$ represents a non-central chi-square distribution with n degrees of freedom and non-centrality parameter given by $(\rho^2/1-\rho^2)\chi^2_n$. The subroutine CSNDF available on the IMSL Fortran library (1995) was used to compute the non-central chi-squared distribution function in expression (A1).

Finally, we have that the probability of $T(x)$ and/or $T(y)$ exceeding the control limit is given by:

$$p = 1 - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} Ps\left(Z_{\bar{x}}, Z_{\bar{y}}\right) f\left(Z_{\bar{x}}, Z_{\bar{y}}\right) dz_{\bar{x}}dz_{\bar{y}} \quad \text{(A2)}$$

where $f\left(Z_{\bar{x}}, Z_{\bar{y}}\right)$ is a standardized bivariate normal distribution function with correlation ρ.

During the in-control period $a = b = 1$, $c = d = 0$. As the false alarm risk of the control chart is a continuous decreasing function of CL, a grid search using (A2) allow us to obtain the value of CL that equates p to a specified false alarm risk (α), reminding that $ARL_0 = 1/\alpha$.

According to the expressions (A1) and (A2), the control limits depend on the correlation; however, the correlation has minor influence on the performance of the NCS charts.

The Pv values are given by:

$$P_v = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\int_{CL_x}^{\infty} g(t)dt\right] f\left(Z_{\bar{x}}, Z_{\bar{y}}\right) dz_{\bar{x}}dz_{\bar{y}} \quad \text{(A3)}$$

with $b > 1$ and/or $d > 0$ and $a = 1$ and $c = 0$.