

COMPARING CONTROL CHARTS FOR GAUSSIAN MEAN VECTORS WITH MEWMA AND SLIDING WINDOW SCHEMES

Denis Altieri de Oliveira Moraes

Departamento de Estatística, Universidade Federal de Santa Maria, 91105-900 – Santa Maria – RS, Brasil
daltieri@smail.ufsm.br

Fernando Luiz Pereira de Oliveira

Departamento de Estatística, Universidade Federal de Ouro Preto, 35400-000 – Ouro Preto – MG, Brasil
fernandoluiz@iceb.ufop.br

Luiz Henrique Duczmal

Departamento de Estatística, Universidade Federal de Minas Gerais, 31270-901 – Belo Horizonte – MG, Brasil
duczmal@est.ufmg.br

Frederico Rodriguez Borges da Cruz

Departamento de Estatística, Universidade Federal de Minas Gerais, 31270-901 – Belo Horizonte – MG, Brasil
fcruz@est.ufmg.br

Resumo: Para distribuições normais, as distâncias T^2 de Hotelling e MEWMA estão diretamente relacionadas à distância de Bhattacharyya. Essa relação provê importante uma informação a respeito do erro de classificação na forma de um limite superior de probabilidade, indicando o grau de sobreposição entre dois processos. Para demonstrar esse fato é conduzido um estudo simulado para o monitoramento do vetor de médias em um processo Gaussiano bivariado. As características do gráfico de controle de confiança proposto são utilizadas para comparar os efeitos da estimação do vetor de médias através do esquema MEWMA e janelas móveis, as quais têm pesos uniformes, linear e exponencialmente distribuídos. Os resultados demonstram que os gráficos de controle de confiança MEWMA são mais fáceis de calibrar e também apresentam menor efeito inercial para grandes mudanças.

Palavras-chave: Processos gaussianos pontuais, vetores de médias, controle estatístico de processos, parâmetro de não-centralidade, distancia de Bhattacharyya, janela móvel.

Abstract: *In this work we show that for normal distributions the Hotelling's T^2 and the multivariate exponentially weighted moving average (MEWMA) distances are directly related to the Bhattacharyya distance. This relationship provides important information concerning on the misclassification error probability as an upper bound on it. In fact, this useful information indicates the overlap degree between the in- and out-of-control processes. Therefore, the first purpose of this simulation study is to monitor the mean vector of a bivariate Gaussian process by means of an informative control chart based on probability bounds. Additionally, a comparison study is carried to measure the effects of estimating the actual mean vector by the MEWMA scheme and sliding window schemes, which are chosen to have uniform, linear and exponentially distributed weights. Results demonstrated that the confidence MEWMA control chart is easier to calibrate and shows less inertia for big shifts in the mean vector than the sliding window approach.*

Keywords: Gaussian point processes, mean vectors, statistical process control, noncentrality parameter, Bhattacharyya distance, sliding window.

1. Introduction

In many industrial problems the probability of misclassification is a subject of great interest, but the calculation is a difficult task even when the observed data is normal. Therefore, the idea of monitoring a process by calculating its probability to be in- or out-of-control is usually discarded. Recent advanced statistical techniques with applications to the \bar{X} and the S^2 control chart includes the univariate case (Faraz and Saniga, 2012), and the multivariate case, where a recent work also covers the global process monitoring by controlling the mean vector and covariance matrix simultaneously (Niaki and Memar, 2009).

Considering the process control of mean vectors only, the most utilized method to monitor big shifts is the Hotelling's T^2 control chart (Hotelling, 1947), while in the case of smaller shifts the multivariate exponentially weighted moving average (MEWMA) control chart (Lowry et al., 1992) is more popular for being simpler to implement when compared to its most famous concurrent, the multivariate cumulative sums (MCUSUM) control chart (Crosier, 1988). Although the methodology utilized in this work may be extend for the multivariate global process monitoring by probabilities, as an initial propose we only consider the process control of multivariate mean vectors.

If a closed-form for the error probability is not provided, one may seek either an approximate expression or an upper bound on the error probability. A closed form for the upper bound on the error probability is very useful for many reasons. Beyond of reducing the computational effort greatly, the evaluation of a simple formula may provide an insightful knowledge about the actual process state. Furthermore, the misclassification error increases significantly with the number of dimensions, reducing dramatically the standard confidence levels that the process is actually in-control (Fukunaga, 1990). Due to this fact, the evaluation of a probability measure instead of raw distances gives more valuable information about the price we have to pay for not knowing the alternative process state *a priori*. Focusing this objective, the monitoring of Gaussian mean vectors by means of a simple distance transformation that leads to a control chart directly based on probabilities is discussed in this work.

Additionally, another question that arises when the process is monitored for small shifts in the mean vector with MEWMA based control charts is the inertial phenomenon (Lowry et al., 1992), which delays change detection if the actual shift occurring in the mean vector is big and the chart is calibrated for very small shifts. To avoid this phenomenon and for practical reasons, the researcher may seek for alternative approaches to estimate the actual mean vector, including discarding totally old observations by the utilization of sliding window schemes. While the EWMA method accumulates information about all the previous observations into the actual mean vector, the sliding window approach discards past observations faster and gives most weight only to the recent past.

Many authors such as Hwang and Hubele (1993a, 1993b), Guh (2005, 2008) suggested a moving window approach as an essential tool for on-line pattern identification. However, two problems that can be addressed are how to choose the appropriate window size and how to deal with unnatural patterns in which a misalignment of the pattern in time is possible. Also, the identified pattern can be different in time from the training pattern (Hachicha and Gorbel, 2012) and dynamic window sizes should be more appropriated. However, dynamic sizes for the sliding window schemes are not covered in this work. As demonstrated in the experiments of this work, to keep the window size fixed reflects directly on the shift magnitude to be detected. Additionally, they do not have the property of reducing the inertial phenomenon for big shifts.

To provide an analysis on the effects of estimating the actual mean vector based on the MEWMA or sliding window approaches, tree types of sliding windows are proposed for comparison purposes. The proposed sliding window observation weights are uniform, linear and geometrically distributed. As the sliding window size is an important parameter, the probability control charts with the utilization of sliding window schemes can also provide

benchmarking criteria for existing and future developments. Additionally, the probability based control chart facilitates the comparative study by the standardization of the statistical distances into the 0-1 interval as upper bounds for the usual confidence levels.

In the following sections, the main properties of the noncentrality parameter traditionally used to monitor the mean vector by the Hotelling's T^2 and MEWMA control charts, as well the link with the upper bound on the error probability is described. Next, the experiments on the Hotelling's T^2 and the probability control chart performance with individual observations are compared. Further, the MEWMA and sliding window approaches comparison are carried out by the computation of the average run lengths (ARLs), or average time to signal (ATS) as the interval between observations is regular. Finally, some remarks and recommendations on the probability control chart and the effects of sliding window schemes are made.

2. Methodology

It is known that the performance measured by the average run length (ARL) of traditional control charts like the Hotelling's T^2 and MEWMA depends only on the noncentrality parameter, not depending on the shift's direction (Lowry et al., 1992). This distance is given by

$$d_t^2 = (\mathbf{X}_t - \mathbf{M}_0)^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{X}_t - \mathbf{M}_0), \quad (1)$$

where \mathbf{X}_t , \mathbf{M}_0 and $\boldsymbol{\Sigma}_0$ are the observed vector, the in-control mean vector and the in-control covariance matrix, respectively. The decision rule gives an out-of-control signal as soon as $d_t^2 > h_1$, where h_1 is a specified threshold that leads to a pre-specified false alarm rate, usually defined in terms of the ARL.

In his original paper in 1947, Hotelling suggested the utilization of d^2 instead of d to avoid the labor of extracting the square root, but as the computational power has massively increased in the last decades it is almost no matter anymore. Thus, to maintain clear the effect on the in-control limits, in this work d is used for experiment comparisons varying in the 0-4 range, which is the Hotelling's distance squared root.

While the Hotelling's T^2 considers the global process monitoring by outlying observations that are outside the in-control boundaries, the MEWMA statistic considers the entire process to be out-of-control as soon as

$$z_t^2 = (\mathbf{M}_t - \mathbf{M}_0)^T \boldsymbol{\Sigma}^{-1} (\mathbf{M}_t - \mathbf{M}_0) > h, \quad (2)$$

where h is the pre-specified threshold to achieve a desired ARL and \mathbf{M}_t is the mean vector estimated with past and current information by a MEWMA scheme, such that

$$\mathbf{M}_t = (1 - \lambda) \mathbf{M}_{t-1} + \lambda \mathbf{X}_t. \quad (3)$$

and $0 < \lambda \leq 1$. Observe that when $\lambda = 1$, the MEWMA distance reduces to the Hotelling's distance.

The noncentrality parameter is very popular in the pattern recognition field (Therrien, 1989), also known as Mahalanobis distance which is derived from the more general cases, the Bhattacharyya and Chernoff bounds (Fukunaga, 1990). Particularly, the Bhattacharyya bound gives a simplified closed-form expression to compute an upper limit on the Bayes error for the case of normal distributed processes such as

$$\varepsilon_u = \sqrt{P_1 * P_2} \int \sqrt{p_1(X) * p_2(X)} dX = \sqrt{P_1 * P_2} e^{-\mu(1/2)}, \quad (4)$$

where

$$\mu(1/2) = \frac{1}{8} (\mathbf{M}_2 - \mathbf{M}_1)^T \left(\frac{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}{2} \right)^{-1} (\mathbf{M}_2 - \mathbf{M}_1) + \frac{1}{2} \ln \frac{|\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2|}{2 \sqrt{|\boldsymbol{\Sigma}_1| |\boldsymbol{\Sigma}_2|}}. \quad (5)$$

The term $\mu(1/2)$ is called the Bhattacharyya distance and is used as an important separability measure between two normal distributions, where \mathbf{M}_i and Σ_i , $i = 1,2$, are the mean vector and covariance matrix of each class. This distance is composed of two terms, the first one carrying the information about the process difference in the mean vectors, and the second part corresponding to the difference in the covariance matrices.

Rao (1947) explained that this distance is an explicit function of the proportion of overlapping individuals in the two populations. Rao (1949) also commented that Bhattacharyya had developed a perfectly general measure defined by the distance between two populations based on a metric of Riemannian geometry, with the angular distance between points representing the populations in a unit sphere.

In the case of single-hypothesis tests, like in statistical process control (SPC) problems, the out-of-control state is generally undetermined. Then, instead of utilizing equation (4) which supposes two known processes, it is more interesting to evaluate only the upper bound for the Type I error, which refers only to the known process and is given by

$$\varepsilon_{u1} = \sqrt{P_2/P_1} \int \sqrt{p_1(X) * p_2(X)} dX = \sqrt{P_2/P_1} e^{-\mu(1/2)}. \quad (6)$$

Also, as this work is focused in the monitoring of mean vectors only, the assumption of equal covariance matrices reduces the Bhattacharyya distance to the noncentrality parameter, except by a constant, assuming the form

$$\mu(1/2) = \frac{1}{8} (\mathbf{M}_t - \mathbf{M}_0)^T \Sigma_0^{-1} (\mathbf{M}_t - \mathbf{M}_0). \quad (7)$$

where \mathbf{M}_t is the mean vector estimated at the instant t , \mathbf{M}_0 is the in-control mean vector and Σ_0 is the in-control covariance matrix.

This simplified form preserves all the known properties of the Hotelling's T^2 and MEWMA control chart with respect to the performance measured by the average run length. The first and second order moments for the Bhattacharyya distance for the case of equal covariance matrices are easily deduced from the results on the d^2 statistic as presented by Fukunaga (1992). The simulated experiments presented in the following section agree with the theoretical values for the first and second moments of the Bhattacharyya distance as well for the Hotelling's T^2 with high precision.

2.1. Probability control charts

The theoretical results presented above provide a different look in the process monitoring by transforming the statistical raw distances and their respective in-control boundaries into probability values as standard patterns. First, if there is no special reason to weight the in- and out-of-control process differently, the processes are equally weighted in equation (6), thus reducing the upper bound on the Type I error to $\exp(-\mu(1/2))$. Different weights for the processes will result in a scale modification to be further analyzed, but still preserving the 0-1 domain.

Observe that when the process is actually in-control, the estimated mean vector, or individual observations, must show no significant difference from the in-control standard error levels. This leads to an upper bound of ε_1 that is near to one because the in-control and current processes are completely overlapped. When the mean vector shifts to the out-of-control state, the upper bound on ε_1 decreases indicating less overlapping among the processes. By other hand, if the complementary probability is taken, it indicates an upper bound on the confidence level which is closer to zero, meaning that the current process is not being apart from the in-control state.

Based on those appointments, the probability control chart when individual observations are compared to the in-control mean vector is taken as the standard level for the different ways of estimating the mean vector. This approach can be viewed as the MEWMA chart with $\lambda = 1$, or a sliding window chart with unitary window size. For this reason, this control chart is identified by the SW1 code (sliding window of unitary size). This control

chart is a simple scale transformation of the Hotelling's T^2 by the use of Bhattacharyya distance, triggering a signal when

$$p_t = 1 - \exp \left[-\frac{1}{8} (\mathbf{X}_t - \mathbf{M}_0)^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{X}_t - \mathbf{M}_0) \right] > h^*, \quad (14)$$

where h^* is the in-control upper limit to achieve a desired ARL_0 .

If the individual observed vector is switched by a mean vector, it is possible to utilize the MEWMA or sliding window schemes to its estimation. Equation (3) is utilized to estimate \mathbf{M}_t in the case of an MEWMA based control chart. For all methods utilizing mean vector estimates instead of individual observations, the probability control chart triggers an out-of-control signal as soon as

$$p_t = 1 - \exp \left[-\frac{1}{8} (\mathbf{M}_t - \mathbf{M}_0)^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{M}_t - \mathbf{M}_0) \right] > h^*. \quad (15)$$

where h^* is chosen to achieve a desired ARL_0 .

For all cases of sliding window schemes, only the observation vectors inside current window are weighted and the mean vector is given by

$$\mathbf{M}_t = \sum_{i=t-k+1}^t w_i^* \mathbf{Y}_i, \quad (16)$$

with $\sum_{i=t-k+1}^t w_i^* = 1$. The uniform sliding window approach equally weights all the observations inside the window of size k with w_i^U given by

$$w_i^U = \frac{1}{k}, \quad (17)$$

$i = t - k + 1, \dots, t$. The linear sliding window approach gives more weight to the most recent observation and decreases linearly the weight of older vectors as

$$w_i^L = \frac{\frac{j}{k}}{\sum_{j=1}^k \left(\frac{j}{k} \right)}. \quad (18)$$

In the exponential sliding window scheme, the weights for the observation vectors inside in the window are distributed by

$$w_i^E = \frac{j^\varphi}{\sum_{j=1}^k (j^\varphi)}. \quad (19)$$

where φ is a smoothing factor between 0 and 1. When $\varphi = 1$, the exponentially weighted window converges into the uniform window. The smoothing factor φ utilized for the exponentially weighted window is fixed as 0.7 as it decay below 0.5 after two steps. The calculation of individual weights for the three proposed sliding window schemes of size 4 is illustrated in Table 1.

Table 1. Weights computation for sliding window schemes with size 4

<i>Window Position</i>	<i>t-3</i>	<i>t-2</i>	<i>t-1</i>	<i>t</i>	Σ
Uniform distribution	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{4}{4}$
Uniform Weights	0.250	0.250	0.250	0.250	1
Linear Smoothing	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{3}{4}$	$\frac{4}{4}$	$\frac{10}{4}$
Linear Weights	0.100	0.200	0.300	0.400	1
Exponential Smoothing	0.7^4	0.7^3	0.7^2	0.7^1	1.77
Exponential Weights	0.135	0.193	0.276	0.395	1

The control chart calibration procedure was carried in two steps to achieve an $ARL_0 = 200$ for all control charts. The first step adjusts a linear regression models in the form $d^2 = a + b * \ln(ARL)$. This procedure gives an approximate first estimative of in-control thresholds for each chart. The second step in the calibration procedure iteratively adjusts the threshold by interpolation. Next section illustrates the functionality of the proposed control chart and analyses the comparative experiments.

3. Results and discussion

The first part of the experiments compares the Hotelling's T^2 and the SW1 control chart, which performs a scale transformation of the Hotelling's distance. Figure 1 part (a) shows the

signal pattern for the case of no change in the mean vector, $d = 0$. Parts (b) and (c) shifts the mean vector process at time $t = 201$ to the distances $d = 3$ and 6 , respectively. In the scatter-plot below the control chart, the out-of-control observation vectors are marked with red dots in the scatter-plot, while the in-control dots are black. The vertical line in the middle of the chart delimits the change point. The horizontal dashed lines are the in-control thresholds for the pre-defined $ARL_0 = 200$. Given in probability value, the in-control upper limit for the SW1 chart is $h_{SW1}^* = 0.7362$. The correspondent in-control noncentral distance that holds for an $ARL_0 = 200$ in the Hotelling's T^2 control chart is $d = 3.265$, which is a scale transformation of h_{SW1}^* .

Observing Figure 1 (c), notice that most of the out-of-control observation vectors do not overlaps the in-control region, resulting in probability values closer to 1. This indicates that the confidence level converges to 1 when the processes are not overlapped. This characteristic pattern does not happen with the Hotelling's T^2 statistic because it has no bound for maximum values, making the interpretation of out-of-control signals difficult to evaluate.

A more detailed summary of the raw distances and their equivalent confidence levels are given in Table 2, where \bar{d}^2 and \bar{p} are average values and $Sd(*)$ is the standard deviation from 100.000 sample replications of size 10. Note that the simulated experiments confirm with high precision the parameters of the Hotelling's T^2 statistic (d^2). As expected, the ARL for both charts are exactly equal, demonstrating that the transformation of the Hotelling's T^2 by the Bhattacharyya distance into probabilities does not modify the ARL performance.

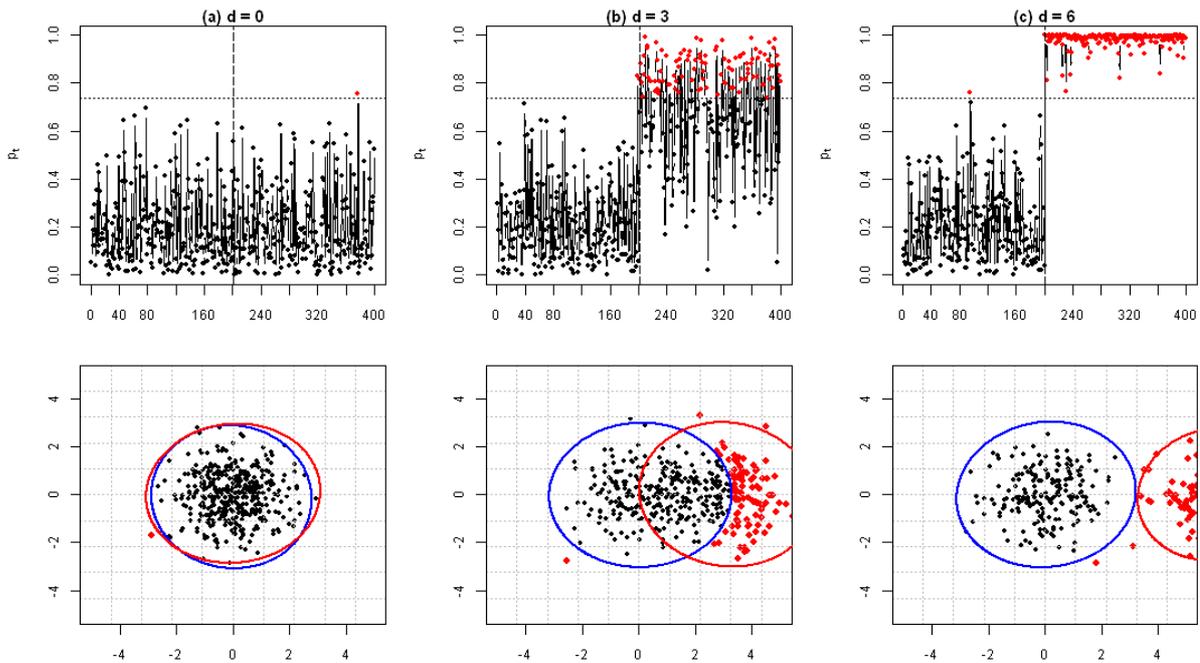


Figure 1: Confidence control chart for individual vectors with scatter plots

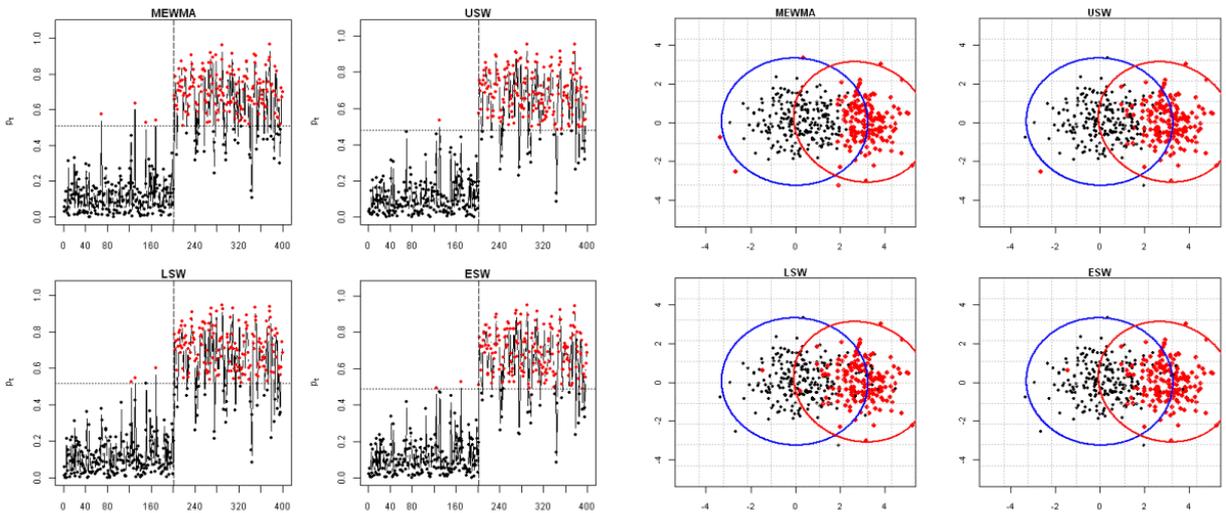
When the MEWMA based control chart utilizes $\lambda = 1$, it performs identically to the SW1 control chart as a standard confidence maximum level to protect the global in-control process region. Also notice that both USW and ESW control charts with sliding windows of any size performs equally when $\varphi = 1$.

Figure 2 is composed of four sets of control charts in part (a) and their respective two-dimensional scatter plots in part (b). The observation of this figure shows the reducing effect on the confidence levels for all control charts. For the MEWMA based control chart with $\lambda = 0.7$, the transformed in-control limit is $h_{MEWMA.7}^* = 0.5086$ ($d = 2.3842$). Also is possible to notice a reduction in the in-control limits of the USW, LSW and ESW charts, which are $h_{USW}^* = 0.4811$ ($d = 2.2908$), $h_{LSW}^* = 0.5166$ ($d = 2.4115$) and $h_{ESW}^* = 0.4901$ ($d = 2.3842$).

= 2.3212). This reduction indicates that the chart become sensitive to smaller changes in the mean vector, and does not matter the individual vector's distances.

Table 2: Summary of the Hotelling's T^2 and Confidence statistics with ARL comparison

d	\bar{d}^2	Sd(d^2)	ARL	$\bar{p}(\%)$	Sd(p)(%)	ARL	
0.0	2.000	1.850	200.6	20.00	15.74	200.6	
	0.006	0.006	0.634	0.001	0.000	0.634	
0.5	2.251	2.070	118.8	21.97	16.93	117.7	
	0.007	0.007	0.376	0.001	0.001	0.372	
1.0	3.001	2.642	43.1	27.62	19.57	43.1	
	0.009	0.008	0.136	0.001	0.001	0.136	
1.5	4.252	3.407	16.0	36.13	21.96	16.0	
	0.013	0.011	0.051	0.001	0.001	0.051	
2.0	6.003	4.263	7.0	46.39	22.98	7.0	
	0.019	0.013	0.022	0.001	0.001	0.022	
2.5	8.253	5.163	3.6	57.19	22.30	3.6	
	0.026	0.016	0.011	0.002	0.001	0.011	
3.0	11.004	6.086	2.2	67.49	20.13	2.2	
	0.035	0.019	0.007	0.002	0.001	0.007	
3.5	14.270	7.039	1.5	76.51	16.97	1.5	
	0.143	0.070	0.005	0.002	0.001	0.005	
4.0	18.021	7.987	1.2	83.86	13.39	1.2	
	0.180	0.080	0.004	0.003	0.000	0.004	
h(ARL ₀ =200)		10.66			73.62		



(a) Control Charts (b) Confidence Ellipses
Figure 2: Confidence control charts with $\lambda = 0.7$, SW2 and $\phi = 0.7$

Despite the fact that the control charts become more sensitive to smaller shifts in the mean vector, it is interesting to note a drawback of USW, LSW and ESW schemes that allows to some extreme values clearly out-of-control to be considered in-control. In the same manner, many vectors that could be considered in-control are marked as out-of-control dots. This happens because the observation vector at instant $t - 1$ receives too much weight in the SW approach for the actual observation vector compensate unless it is an outlier in the opposite side of the out-of-control process. The MEWMA based control chart seems to avoid this problem in its respective scatter-plot performing a good differentiation between in- and

out-of-control vectors. Such behaviors are due to the fact that the MEWMA scheme accumulates all the past information in the current mean vector while the SW scheme does not.

When the window size increases to 4 (SW4), the MEWMA based control chart has the λ parameter decreased from 0.7 to 0.4 for comparison purpose, marked as MEWMA.4. Figure 3 illustrates the four confidence control charts standard patterns for a shift of magnitude $d = 3$. The respective in-control limits are very closer each other and all them leads to completely separable processes, which are $h_{MEWMA}^* = 0.2747$ ($d = 1.6028$), $h_{USW}^* = 0.2677$ ($d = 1.5789$), $h_{LSW}^* = 0.3190$ ($d = 1.7530$) and $h_{ESW}^* = 0.3082$ ($d = 1.7168$). More detailed information concerning the mean and standard deviation of the transformed statistic for this set of control charts are given in Table 3, where it is interesting to note that the reduction on the in-control limits to the level of smaller distances. That reduction gives insight about what is the optimum distance that can be detected efficiently detected for each chart configuration, which is below $d = 3$ for the MEWMA.7 and SW2 charts and below $d = 2$ for the MEWMA.4 and SW4 charts.

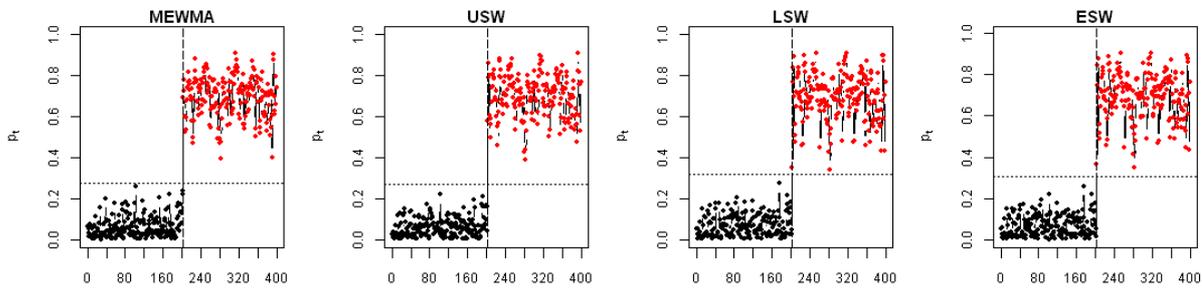


Figure 3: Confidence control charts with $\lambda = 0.4$, SW4 and $\phi = 0.7$

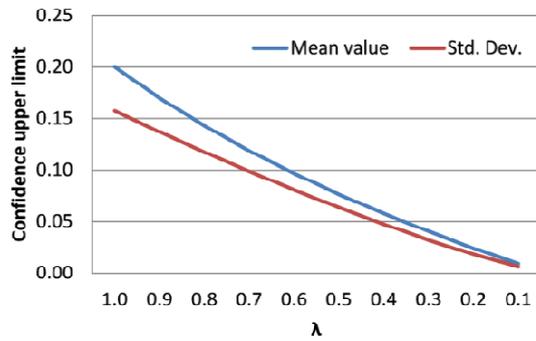
A detailed comparison between the confidence MEWMA based control chart baselines (mean values) and standard deviations for the in-control process with λ varying from 1 to 0.1 by 0.1 units is illustrated in Figure 4. To analyse the out-of-control behaviour of the proposed statistic, the mean vector is shifted with d varying in the 0-7 range by 0.5 units. This information on the first and second order moments of the proposed statistic is also valid to inspect in order to have an informative support for the decision maker. From the results presented in Figure 5 is easy to notice the interesting patterns of the out-of-control statistics for the confidence chart. Note that when the mean value decreases with the smoothing parameter the standard deviation have a point of maximum that is shifted to higher values.

The ARL comparisons between the MEWMA and SW control charts are given in Tables 4. In Figure 6 (a) is shown the ARL comparison of all control charts, while Figure 6 (b) the natural logarithm is taken to amplify the methods differences. Figure 7 splits the comparison into two groups. In Figure 7 (a) the SW1, MEWMA.7 and SW2 control charts are compared, while in Figure 7 (b) the comparison is between the SW1, MEWMA.4 and SW4 control charts. From that Figures is possible to notice the higher degree of inertia effect produced by the SW schemes.

The second set of control charts in Figure 7 (b) compares the SW1, MEWMA.4 and SW4 charts. Although those control charts performs better for shifts below $d = 2.0$, they have higher degree of inertial effect than the SW2 charts for shifts ($d = 4$). Again the USW4 has the best performance, which is comparable to the MEWMA.4 chart. With respect to the robustness against the inertia impact of large shifts, the USW approach seems to be more effective scheme. The LSW and ESW scheme performs worse in both cases when compared to the USW scheme for big shifts. While the differences between the SW schemes for small shifts are not evident in the SW2 charts, the LSW4 and ESW4 charts performs worse than the USW4 chart also for the smaller shifts.

Table 3: Summary of the MEWMA.4 and SW4 control chart statistics

	MEWMA.4		USW4		LSW4		ESW4	
d								
0.0	5.76	4.72	5.45	4.28	6.58	5.29	6.32	5.09
	0.04	0.03	0.04	0.03	0.05	0.04	0.04	0.04
0.5	7.84	5.94	7.61	5.65	8.64	6.64	8.39	6.44
	0.06	0.04	0.05	0.04	0.06	0.05	0.06	0.05
1.0	13.77	8.68	13.73	8.84	14.49	9.82	14.24	9.64
	0.10	0.06	0.10	0.06	0.10	0.07	0.10	0.07
1.5	22.72	11.73	22.88	12.67	23.25	13.62	22.99	13.51
	0.16	0.08	0.16	0.09	0.16	0.10	0.16	0.10
2.0	33.51	14.53	33.76	16.53	33.68	17.36	33.42	17.40
	0.24	0.10	0.24	0.12	0.24	0.12	0.24	0.12
2.5	44.92	16.78	45.04	20.01	44.55	20.64	44.26	20.88
	0.32	0.12	0.32	0.14	0.32	0.15	0.31	0.15
3.0	55.88	18.34	55.62	22.80	54.81	23.16	54.46	23.62
	0.40	0.13	0.39	0.16	0.39	0.16	0.39	0.17
3.5	65.63	19.13	64.78	24.73	63.76	24.76	63.35	25.45
	0.46	0.14	0.46	0.17	0.45	0.18	0.45	0.18
4.0	73.77	19.19	72.21	25.77	71.11	25.43	70.63	26.33
	0.52	0.14	0.51	0.18	0.50	0.18	0.50	0.19
$h(ARL_0=200)$	27.47		26.77		31.90		30.82	

**Figure 4: Mean value and standard deviation of the Confidence MEWMA CC for the in-control process with various λ 's**

Although the SW2 control charts performs better than the SW1 chart and similarly to the MEWMA.7 chart in Figure 7 (a) for small shifts, the inertial effect is visible for distances larger than $d = 3.0$. Ordering the schemes from the less to the most sensitive with respect to the inertial effect, the MEWMA.7 chart performs better, followed by the USW2, ESW and LSW charts.

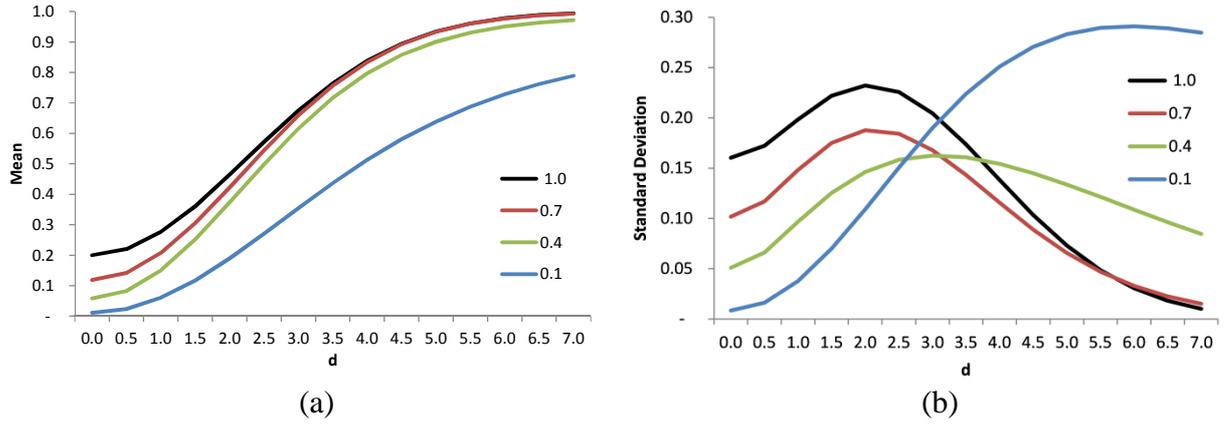


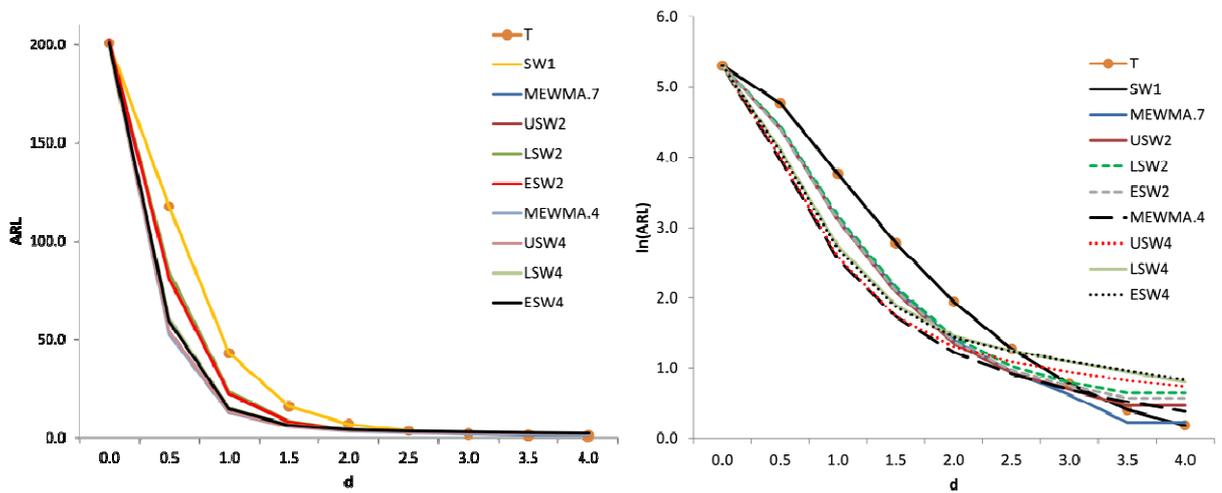
Figure 5: Mean value (a) and standard deviation (b) of the Confidence MEWMA CC for the out-of-control process with various d 's

4. Conclusion

In this work we propose a new manner of monitoring Gaussian mean vectors by the use of an upper bound on the confidence that the process is in-control. Instead the monitoring of the noncentrality parameter, we suggest the use of the Bhattacharyya distance and its relationship with the upper bound on the misclassification error. While the Hotelling's T^2 traditional distance has no maximum values, the proposed confidence control chart based on probabilities for individual observation vectors manifest a useful distinction between processes in the 0-1 range. In this case, when the out-of-control process became completely separable (not overlapped) from the in-control process, the proposed statistic converge to 1, not going to infinity.

Table: ARL comparison between MEWMA and SW control charts varying λ and the window size

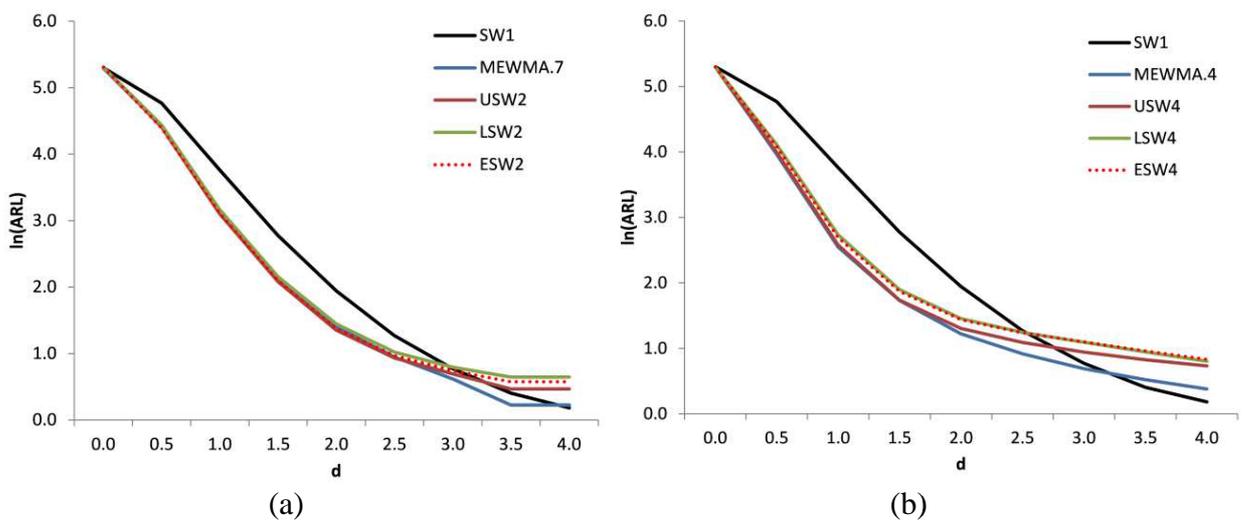
d	EWMA.7	USW2	LSW2	ESW2	EWMA.4	USW4	LSW4	ESW4
0.0	198.9	202.7	201.8	201.9	199.3	199.0	199.1	200.7
	4.45	4.53	4.51	4.51	4.46	4.45	4.45	4.49
0.5	83.2	82.0	84.6	81.0	52.6	54.4	61.1	59.2
	0.83	0.82	0.85	0.81	0.53	0.54	0.61	0.59
1.0	22.7	22.2	23.7	22.4	12.8	13.2	15.4	14.8
	0.23	0.22	0.24	0.22	0.13	0.13	0.15	0.15
1.5	8.4	8.0	8.6	8.1	5.6	5.7	6.7	6.5
	0.08	0.08	0.09	0.08	0.06	0.06	0.07	0.06
2.0	4.1	3.9	4.2	4.0	3.4	3.7	4.3	4.2
	0.04	0.04	0.04	0.04	0.03	0.04	0.04	0.04
2.5	2.6	2.5	2.8	2.6	2.5	3.0	3.5	3.5
	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
3.0	1.9	2.0	2.2	2.1	2.0	2.6	3.0	3.0
	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03
3.5	1.5	1.8	2.0	1.9	1.7	2.3	2.6	2.6
	0.01	0.01	0.01	0.01	0.02	0.02	0.03	0.03
4.0	1.3	1.6	1.9	1.8	1.5	2.1	2.2	2.3
	0.01	0.02	0.02	0.02	0.01	0.02	0.02	0.02



(a) ARL (b) ln(ARL)
Figure 6: ARL and ln(ARL) comparison for all control charts

Additionally, we show that the probability control chart for individual observation vectors can be extended to the more general case, the monitoring of small shifts by the use of MEWMA based control charts and control charts with sliding window schemes. In the same manner of the MEWMA method, instead of using individual observation vectors, the sliding window approaches are commonly used to estimate the actual mean vector for different purposes, including the on-line pattern recognition. We show the equivalence in performance measured by the ARL among the MEWMA based control charts and sliding window schemes for specific parameters.

While this equivalence holds for the small shifts in the mean vector, the sliding window approach shows to be more suitable to suffer the inertial effect for bigger shifts than the MEWMA based scheme. Indeed, in the same manner that a decrease in the weighting factor λ of the MEWMA chart provides an identification of smaller shifts, the increasing in the sliding window size corresponds to a control chart that is more effective to the detection of smaller shifts, but at the cost of more inertial effect for big shifts than the MEWMA based chart.



(a) (b)
Figure 7: ln(ARL) comparison: (a) SW1, MEWMA.7 and SW2 control charts (b) SW1, MEWMA.4 and SW4 control charts

Future work on this topic includes the monitoring of the covariance matrix of a Gaussian process by the use of probability based control charts, as well the global process

monitoring, i.e., the jointly monitoring of the mean vector and covariance matrix of a multivariate Gaussian process.

Acknowledgments

The authors would like to thank FAPEMIG (APQ-00613-12), CAPES-REUNI and CNPq for the support to this work.

5. References

- Crosier, R. B. Multivariate Generalizations of Cumulative Sum Quality-Control Schemes. *Technometrics*, 1998; **30**(3): 291-303.
- Faraz, A. & Saniga, E. Multiobjective Genetic Algorithm Approach to the Economic Statistical Design of Control Charts with an Application to Xbar and S² Charts. *Quality and Reliability Engineering International*, 2013; **29**: 407-415.
- Fukunaga, K. *Introduction to Statistical Pattern Recognition*. (2nd ed.). Boston: Academic Press, 1990.
- Guh, R. S. & Shiue, Y. R. An effective application of decision tree learning for on-line detection of mean shifts in multivariate control charts. *Computers and Industrial Engineering*, 2008; **55**(2): 475-493.
- Guh, R. S. & Shiue, Y. R. On-line identification of control chart patterns using self-organizing approaches. *International Journal of Production Research*, 2005; **43**(6): 1225-1254.
- Hachicha, W. & Ghorbel, A. A survey of control-chart pattern-recognition literature (1991-2010) based on a new conceptual classification scheme. *Computers and Industrial Engineering*, 2012; **63**: 204-222.
- Hotelling, H. Multivariate quality control - illustrated by the air testing of sample bombsights. *Techniques of Statistical Analysis*, 1947; 111-184.
- Hwang, H. B. & Hubele, N. F. Back-propagation pattern recognizers for control charts: Methodology and performance. *Computers and Industrial Engineering*, 1993b; **24**(2): 219-235.
- Hwang, H. B. & Hubele, N. F. X control chart pattern identification through efficient off-line neural network training. *IIE transactions*, 1993a; **25**(3): 27-40.
- Lowry, C.A., Woodall, W.H. & Rigdon, S. E. A Multivariate Exponentially Weighted Moving Average Control Chart. *Technometrics*, 1992; **34**(1): 46-53.
- Niaki, S.T.A. & Memar, A.O. A new statistical process control method to monitor and diagnose bivariate normal mean vectors and covariance matrices simultaneously. *The International Journal of Advanced Manufacturing Technology*, 2009; **43**: 964-981.
- Rao, C.R. The problem of classification and distance between two populations. *Nature*, 1947; 159, 30.
- Rao, C. R. On the distance between two populations. *Shankya: The Indian Journal of Statistics*, 1949; **9**: 246-248.
- Therrien, C. W. *Decision Estimation and Classification, an Introduction to Pattern Recognition and Related Topics*. John Wiley & Sons, 1989.