

An EWMA control chart for the mean of individual streams in multiple stream processes

Gráfico de controle EWMA para as médias de fluxos individuais em processos multifluxo

Bruno Francisco Teixeira Simões¹ , Eugenio Kahn Epprecht² 

¹Universidade Federal do Estado do Rio de Janeiro – UNIRIO, Department of Quantitative Methods, Rio de Janeiro, RJ, Brasil. E-mail: bruno.simoes@unirio.br

²Pontifícia Universidade Católica do Rio de Janeiro – PUC-Rio, Department of Industrial Engineering, Rio de Janeiro, RJ, Brasil. E-mail: eke@puc-rio.br

How to cite: Simões, B. F. T., & Epprecht, E. K. (2021). An EWMA control chart for the mean of individual streams in multiple stream processes. *Gestão & Produção*, 28(3), e062. <https://doi.org/10.1590/1806-9649-2021v28e062>

Abstract: In a multiple stream process (MSP) a product is manufactured in a number of streams in parallel. The traditional tool for monitoring MSPs, the group control chart (GCC), does not take into account that typically the value of the quality variable in each stream is the sum of a component common to all streams and an individual component, of the particular stream. This may render the GCC ineffective in detecting shifts in the mean of individual streams. Based on this two-components model, we propose an exponentially weighted moving average (EWMA) GCC to monitor the means of the individual streams components. We optimize its design (minimizing the ARL for given shifts in the mean of a stream) and compare their ARLs with the ones of other existing charts devised for two-components MSPs. For this comparison, we needed to obtain optimal designs of these previous charts too, which were not available in the literature; this is an additional contribution of our work. The ARLs of the charts were obtained by simulation, with a number of runs sufficiently large to ensure precise results. The results show that the proposed chart outperforms the previous ones, becoming thus recommended for the statistical control of MSPs.

Keywords: Multiple stream processes; Group control chart; Components of variance; Exponentially weighted moving average.

Resumo: Em um processo de fluxo múltiplo (MSP) um produto é fabricado em vários fluxos em paralelo. A ferramenta tradicional para o monitoramento de MSPs é o gráfico de controle de grupo (GCG), que não leva em conta que tipicamente o valor da característica de qualidade em cada fluxo é a soma de um componente comum a todos os fluxos e um componente individual, do fluxo particular. Isso pode tornar o GCG ineficaz na detecção de mudanças na média de fluxos individuais. Com base neste modelo de dois componentes, propomos um GCG da média móvel ponderada exponencialmente (EWMA) para monitorar as médias dos componentes individuais dos fluxos. Otimizamos seu projeto (minimizando o ARL para determinados deslocamentos na média de um fluxo) e comparamos seus ARLs com os de outros gráficos destinados a MSPs de dois componentes. Para essa comparação, precisávamos obter os projetos ótimos também desses gráficos anteriores, que não estavam disponíveis na literatura; essa é uma contribuição adicional do nosso trabalho. Os ARLs dos gráficos foram obtidos por simulação, com um número de corridas

Received 12 Sep. 2020 - Accepted 18 Jan. 2021

Financial support: CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), project numbers 140933/2006-9, 302326/2008-1 and 308677/2015-3.



This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

suficientemente grande para garantir a precisão dos resultados. Estes mostram que o gráfico proposto supera os anteriores, tornando-se recomendado para o controle de MSPs.

Palavras-chave: Processos Multifluxo; Gráfico de Controle de Grupo; Componentes da variância; Médias Móveis Exponencialmente Ponderadas.

1 Introduction

Multiple stream processes (MSPs) are present in many industries, e.g. beverage, cosmetics and plastic industries, among others. They are characterized by the fact that the product, quality characteristics (quality variables) and specifications are the same in all streams, which ideally should have the same distribution, adjusted in the same target value and with the same spread. The first statistical process control (SPC) scheme specifically devised for MSPs was the group control chart (GCC), developed by the British during World War II and described by Boyd (1950). The idea behind a GCC is to aggregate in one control chart the information from all streams, making easier the assessment of the process state. Concretely: at every sampling time, n measures of a quality variable are made in each stream and the averages and ranges of the observations in each stream are calculated; in the group \bar{x} control chart only the smallest and largest averages are plotted and compared with the control limits; if they fall within the control limits the process is considered to be in control. The group R chart is similar (for small sample sizes there is no lower control limit and only the largest range is plotted). For several decades this was the only procedure in the literature for monitoring multiple stream processes. Nelson (1986) proposed a runs scheme to be used with group charts (with or without control limits): if a same stream yields the largest (or the smallest) reading more than r times in a row (where r is defined as a function of the number of streams of the process), this should be an indication that this particular stream has a higher (or lower) mean than the other ones.

Some further works on the monitoring of MSPs are Mortell & Runger (1995), Runger et al. (1996), Amin & Li, (2000), Lanning et al. (2002), Liu et al. (2008), Xiang & Tsung (2008), Meneces et al. (2008) & Epprecht et al. (2011a). For a comprehensive literature review of this topic, which describes these works with some detail & includes other references, see Epprecht (2015).

Mortell & Runger (1995) were pioneers in acknowledging the fact that many (if not most) MSPs have two separate sources of variation and thereby in such processes the value of the quality variable measured in time t in each stream should be represented by the sum of two random variables: a “mean level” common to all streams (and which may be stable or present some dynamics over time) plus an individual component responsible for the difference between the stream value and the common component value. This component will consist only of random variation about the common value, having null mean, in the case the process is in control and all streams are centered on the same value; on the other hand, a special cause affecting a stream may make the mean of its individual component differ from zero. An example of such a process is a filling process with a filling machine with many heads. The functioning of a common reservoir, pump and/or characteristics such as the density of the fluid are responsible for the common component level, while a clog or a fissure in some tube of one of the heads will affect only the output of that head. Mortell & Runger (1995) have shown the inefficiency of the traditional GCCs (and of Nelson’s runs scheme) in signalling special causes affecting the mean of a particular stream. They proposed then a statistic (R_t) that is sensitive to such kind of changes: the difference between the largest and the smallest stream averages in

time t . They have shown the effectiveness of this statistic and its superiority relative to the previous control schemes for MSPs. They also proposed and analyzed runs rules, CUSUM and EWMA versions of the chart.

For monitoring the means of the individual streams components under the two-component model Runger et al. (1996), based on a multivariate model and principal components decomposition, proposed a so-called S^2 chart on a statistic which turns out to be proportional to the sample variance between streams: S^2 equals $m-1$ times the sample variance of the streams averages, where m is the number of streams; it can also be interpreted as the squared norm of the orthogonal projection of the multi-stream vector on the complementary subspace of the first principal component (which is the sample mean of all streams). This chart showed similar performance to the R_t chart in the case of shifts in one stream but outperforms the R_t chart when more than one stream shifts. A MEWMA version of the chart was also proposed and briefly analyzed.

Other work based on the two-component model of MSP was Epprecht et al. (2011a). They proposed a *residuals* GCC where, in contrast with the classic GCC for the mean, the points plotted at each sampling time t are the smallest and the largest estimates of the individual streams components in time t rather than the smallest and the largest averages among all streams. This is a Shewhart-type GCC, and was shown to outperform the Shewhart-type versions of Mortell & Runger (1995) R_t chart and Runger et al. (1996) S^2 chart in the detection of mean shifts in one individual stream.

However, as mentioned, Mortell & Runger (1995) and Runger et al. (1996) had proposed and analyzed also EWMA versions of their charts, which were much faster in signalling shifts of small and moderate sizes (≤ 2 standard deviations) than their Shewhart-type counterparts. The fact that the residuals GCC proposed by Epprecht et al. (2011a) outperformed the Shewhart-type versions of the charts of those authors raises the question whether an EWMA version of the residuals GCC might outperform the EWMA versions of those charts as well. This directly motivates the proposal and analysis of such an EWMA version of the residuals GCC, that is, an EWMA GCC on the smallest and the largest estimates of the individual streams components. The definition, optimization and performance analysis of such chart is the object of the present paper.

We give expressions for computation of the EWMA statistics and of the control limits for such control chart, detail its operation, and optimize its design (that is, we find the values of the control limits and of the EWMA smoothing constant that minimize the out-of-control ARL of the chart) for a variety of cases (combinations of number of streams, sample size and size of shift in the mean of a stream for which the average run length — ARL — should be minimized); we optimize, as well, in the same sense, the EWMA versions of Mortell & Runger (1995) R_t chart and Runger et al. (1996) S^2 chart, for the sake of a fair performance comparison. To our knowledge, there was no optimization of the designs of the EWMA R_t and MEWMA S^2 control charts in the literature. Then we obtain the ARL profiles (ARLs for a range of shifts in the mean of an individual stream) of the three control schemes, for comparing their performances. The ARLs were obtained through simulation, due to the prohibitive complexity of an analytical approach, but with a number of runs large enough to yield precise results. As will be seen, the proposed chart outperforms its predecessors.

The stages for achieving the objectives of this research are in Table 1.

Table 1. Stages of the research and objectives of each one.

number	description of the stage	objective
1	Determine the expression of the EWMA statistics and control limits for the GEWMA- \bar{d} , chart	chart definition
2	Build and code a simulation model for computing the ARLs of the GEWMA- \bar{d} , EWMA R_1 and MEWMA S^2 charts. Do the same for the residuals GCC of Epprecht et al. (2011a)	model and program for computing ARLs
3	For different combinations of δ^* (specified size of shift in the mean of one stream), m (number of streams) and n (sample size: number of observations per stream at each sampling time), do a search to determine the optimal design, i.e., the value of the pair (λ, k) that minimizes the ARL_1 for the shift δ^* in the individual component of one stream, keeping at the same time the ARL_0 on a specified value. For each point (λ, k) in the search, the ARL_1 is computed with the simulation program. Do this for the three EWMA-type charts. For the Shewhart-type residuals GCC, just find the control limits that yield the specified ARL_0	optimization
4	For all the cases in Stage 3 above, run the simulation program with the optimal parameters of each chart to obtain their ARL profiles: their ARL_1 values for a range of shift magnitudes δ different from δ^* ; tabulate these profiles for comparing the performances of the charts and interpret the results	performance analysis and comparison

The rest of this paper is organized as follows: Section 2 describes, as a background, the two-component model proposed by Mortell & Runger (1995) and the statistics and charts proposed by them, by Runger et al. (1996) and by Epprecht et al. (2011a). Next, the EWMA chart we propose is described in Section 3. Then, in Section 4, we detail how the simulation was used for computing the ARLs of the three EWMA schemes considered. The design optimization problem is formally defined in Section 5, which also gives the optimal designs (the ones that give the minimum out-of-control ARL) found for the cases considered. Finally, the ARL profiles (ARL values for a range of shifts from 0.5 to 4 standard deviations) of the optimized charts and also of the residuals GCC are presented and compared in Section 6. Section 7 summarizes the conclusions of the analysis.

2 Background

As Mortell & Runger (1995), Runger et al. (1996) and Epprecht et al. (2011a), we consider here MSPs well represented by the following model:

$$x_{ij} = c_t + e_{ij} \tag{1}$$

where

x_{ij} is the j -th observation of the quality variable x in the i -th stream in sampling time t ; $i = 1, 2, \dots, m$, where m is the number of streams of the process, $j = 1, 2, \dots, n$, where n is the sample size (number of observations taken from each stream at sampling time t), and $t = 1, 2, 3, \dots$ are the indexes of the successive samples of the process in the monitoring phase;

c_t is a common component which represents the mean level of the process at time t ; it may either be a constant value or have some dynamics, exhibit autocorrelation, and therefore in the greatest generality its evolution over time is a stochastic process;

e_{tij} is the j -th observation of the individual i -th stream component in sampling time t ; we assume that when the process is in control, $e_{tij} \sim N(0, \sigma_0^2)$;

and, as implicit in this definition, we assume that the values of e_{tij} are iid; independent between themselves and over time; we also assume that the variables c and e are independent. Still according to this definition, when the process is in control, at any time t all streams have the same expected value (they are either centered in a same fixed value, or the expected value wanders but at any given time it is the same for all streams), so the e_{tij} 's can be seen as realizations of the random noise at stream i . On the other hand, when one of the streams goes out of control, its mean shifts and its e_{tij} 's have no longer null mean.

In the particular case of only one observation per stream at each sampling time ($n=1$), the index j is dropped.

From this model, it follows that

$$V(x_{tij}) = V(c_t) + V(e_{tij}) = V(c_t) + \sigma_0^2 \quad (2)$$

So, the classic \bar{x} group control chart (the traditional SPC tool for monitoring the mean of multiple stream processes), based on the m averages over j of the x_{tij} 's of each stream, would have control limits width proportional to a total standard deviation of $\sqrt{V(c_t) + \sigma_0^2/n}$. Then, unless the variance of the mean level, $V(c_t)$, is small relative to the variance of the individual components (intra-stream variance) σ_0^2 , this chart would be quite insensitive to shifts in the mean of one (or some) of the individual stream components e_i .

To overcome this problem, the three works cited that are based on this model of MSP propose monitoring separately the two components, c and e , with separate control charts. They also concur in that, for controlling the mean level, c_t should be estimated by the average of all the x_{tij} 's in time t , and that this statistic should be monitored using an appropriate SPC technique according to the nature of c_t — for example, a Shewhart, EWMA or CUSUM chart in case it exhibits no serial correlation, or some technique for monitoring autocorrelated processes otherwise. In any case, c_t is treated as one variable and hence will be monitored using some established (well-known) univariate process control scheme. For this reason, all these works focus (as we do) exclusively on the monitoring of the individual streams components e_i , $i = 1, 2, \dots, m$.

Mortell & Runger (1995), considering individual observations per stream, proposed the statistic

$$R_t = \max(x_{ti}) - \min(x_{ti}) \quad (3)$$

and analyzed a Shewhart-type control chart on R_t , as well as EWMA, CUSUM and runs rules versions of it. In case of more than one observation per stream at each sampling time, x_{ti} in Equation 3 should be replaced by the average (over j) of the n observations x_{tij} ($j = 1, 2, \dots, n$) in the stream. It is easy to see that in the second member of Equation 3 the mean level components (c_t) of $\max(x_{ti})$ and of $\min(x_{ti})$ cancel out, so the in-control standard deviation of R_t (and so the upper control limit for the R_t chart) is proportional to σ_0 , making the chart sensitive to shifts in the mean of a particular stream.

Runger et al. (1996) proposed an s^2 chart based on the statistic (in the notation of the present paper)

$$S^2 = \sigma_0^2 \sum_{i=1}^m (x_{ti} - \bar{x}_{ti})^2 \tag{4}$$

where \bar{x}_{ti} is the average of the m streams in time t and σ_0^2 is the in-control variance of x_{ti} (the same for all i), which, if not known, should be estimated from preliminary samples of the in-control process. Similarly to R_t , s^2 filters out the mean level component and is sensitive to shifts in the mean of one or some streams. Runger et al. (1996) also analyzed a MEWMA (Lowry et al., 1992) version of it. In this one, the values of each stream are separately smoothed (m EWMA values) and then combined in a s^2 statistic (similarly to the usual MEWMA procedure for multivariate processes). The case of more than one observation per stream at each sampling time is also accommodated in a similar way as in Mortell & Runger (1995), by considering the averages of the n observations in stream i in time t , and replacing σ_0^2 in Equation 4 by σ_0^2 / n .

In contrast with the R_t and the s^2 charts, which are based on a single statistic, Epprecht et al. (2011a) proposed controlling the means of the individual streams components by a *residuals* group control chart, based on the differences (or *residuals*) between the subgroup averages and the estimate of the mean level rather than on the subgroup averages themselves. In sampling time t , n observations x_{tij} ($j=1, 2, \dots, n$) are made at each stream i , $i=1, 2, \dots, m$. Next, the common component (mean level) in time t is estimated by the average of all observations in time t :

$$\hat{c}_t = \frac{1}{nm} \sum_{i=1}^m \sum_{j=1}^n x_{tij} \tag{5}$$

and the differences (or *residuals*)

$$d_{tij} = x_{tij} - \hat{c}_t \tag{6}$$

are calculated. Note that d_{tij} is an estimator of the individual component e_{tij} . Then the average difference in each stream i is computed as

$$\bar{d}_{ti} = \frac{1}{n} \sum_{j=1}^n d_{tij} \tag{7}$$

Alternatively, as can be seen by substituting Equation 6 in Equation 7, \bar{d}_{ti} can be computed by Equation 8

$$\bar{d}_{ti} = \frac{1}{n} \sum_{j=1}^n x_{tij} - \hat{c}_t \tag{8}$$

The maximum and minimum of the m values of \bar{d}_{ti} are plotted in a GCC with control limits appropriately set for the desired false-alarm rate. For details, the reader is referred to Epprecht et al. (2011a).

3 The EWMA group control chart proposed

We here propose an EWMA group control chart for the differences \bar{d}_{ti} (GEWMA- \bar{d} chart). At each sampling time, the sampling occurs exactly like in the case of the residuals GCC just described at the end of the previous section, and each \bar{d}_{ti} is calculated following the same steps (Equations 5, 6 and 7, or, alternatively, Equations 5 and 8). Then, each \bar{d}_{ti} is smoothed using the recursive Equation 9

$$Y_{ti} = \lambda \bar{d}_{ti} + (1 - \lambda) Y_{(t-1)i} \quad (9)$$

yielding m EWMA values, one for each stream. The smoothing constant $\lambda \in (0, 1]$ is an implementation parameter of the chart whose value has to be chosen by its user. Given the assumption that, with the process in control, $E(e_{tij}) = 0$, it follows, as shown in Epprecht et al. (2011a), that $E(\bar{d}_{ti}) = 0$ and this is the initial value for the EWMA statistic: $Y_{0i} = 0$. The maximum and the minimum Y_{ti} values in time t are plotted in a GCC with upper and lower control limits given by

$$UCL = k\sigma_0 \sqrt{\left(\frac{\lambda}{2-\lambda}\right) \left(\frac{m-1}{nm}\right)} \quad (10a)$$

$$LCL = -k\sigma_0 \sqrt{\left(\frac{\lambda}{2-\lambda}\right) \left(\frac{m-1}{nm}\right)} \quad (10b)$$

where the control limit factor k is also an implementation parameter of the chart. A point outside the control limits is considered as evidence of a shift in the mean of the corresponding stream. Although only the extreme EWMA values are plotted, in case of a signal it is advisable to verify if the EWMA value of any other stream is outside the limits.

The expressions for the control limits were derived as follows. Epprecht et al. (2011a) have shown that, when the process is in control, the assumption that $e_{tij} \sim N(0, \sigma_0^2)$ implies that every \bar{d}_{ti} is normally distributed with null mean and standard deviation equal to $\sigma_0 \sqrt{(m-1)/(nm)}$. As a consequence (using the well-known relation between the standard deviation of a normal variable and the standard deviation of its exponentially smoothed value), the standard deviation of every Y_{ti} is Equation 11:

$$\sigma_0 \sqrt{\left(\frac{\lambda}{2-\lambda}\right) \left(\frac{m-1}{nm}\right)} \quad (11)$$

and the control limits (see Equations 10) are just set at k standard deviations of the (null) expected value of the differences. Since the successive values of EWMA statistics exhibit serial correlation, k cannot be determined simply as a quantile of the standard normal distribution but should rather be obtained by search as the value that, for the given number

of streams of the process (m), sample size (n) and value of the smoothing constant λ , makes the in-control average run length match a specified ARL_0 value. This will be explained in detail in the next section.

4 Design of the chart – performance measures and their computation

The best choice of smoothing constant and control limits for the EWMA chart proposed depends on the number of streams of the process m , on the sample size n , on the value specified for the in-control average run length (ARL_0) and on the magnitude δ^* (specified) of the shift in the mean of an individual stream that it is important to detect with minimum delay. Given m and n , the criterion we adopt for the design of the chart is to choose the values of λ and k that, with fixed in-control average run length ARL_0 (specified), minimize the out-of-control run length (ARL_1) for the relevant shift δ^* specified. Minimization of the ARL_1 (when possible) for a relevant shift specified, under the constraint of a specified ARL_0 , is a popular criterion for the design of a control chart.

In the case of charts with memory such as EWMA and CUSUM charts, there are two possible scenarios to be considered for the definition of the ARLs, especially regarding ARL_1 : in the zero-state scenario, the process is already out of control when the monitoring starts; in the steady-state scenario, when the monitoring starts the process is in control and only after some time a special cause brings the process to an out-of-control state. The scenario has an impact on the ARL_1 because in the steady-state scenario when the special cause occurs the effect of the initial value of the EWMA statistic has already dissipated whereas in the zero-state scenario it is present. The result is that the zero-state ARL_1 is larger than the steady-state ARL_1 . However, the differences are not usually substantial, as the great bulk of literature on the performance of EWMA charts has shown.

We considered only the steady-state ARL_1 . The reason is that, even in case the monitoring starts with the process out of control, as soon as there is an alarm, the process should be investigated, the special cause eliminated, and when the monitoring is resumed, the process will be in control. The process should then be in control every time the monitoring is resumed after an alarm (including eventual false alarms), so the zero-state scenario is a transient situation, if it ever occurs.

The analytical computation of the ARLs (either in control or out of control) of the proposed chart is not only quite complex but would be prohibitively computer intensive. For univariate EWMA charts it is still feasible to use, for example (and most often) Markov chain models. With multiple stream processes, however, the \bar{d}_{ti} values of different streams are correlated (the proof is given in Epprecht et al., 2011a), which makes the different Y_{ti} values cross-correlated, too. The Markov chain model for more than two simultaneous EWMA charts of cross-correlated variables is extremely time consuming to compute (see, for example, Epprecht et al., 2011b). For this reason, we resorted to simulation.

In order to guarantee accurate simulation results, we used 10,000 runs for obtaining each ARL estimate. This guarantees a standard error of less than 1% of the $SDRL$ (standard deviation of the run length). The $SDRL$, in turn, is always smaller than the ARL . Indeed, in the case Shewhart-type charts, whose run lengths follow a geometric distribution, $SDRL = \sqrt{ARL(ARL-1)} < ARL$. The run length distributions of EWMA charts are less skewed and will have $SDRL < \sqrt{ARL(ARL-1)} < ARL$. This means that standard error of the ARL estimates would be of less than 1% of the estimate. In each simulation run, values were generated for every e_{ij} (c_i was kept constant and equal to zero, without loss of generality, because the charting procedure filters it out completely; as a result, the

simulated x_{tij} 's coincide with the e_{tij} 's; see Equation 1). At each simulated sampling time t , \hat{c}_t was calculated by Equation 5; next, Equations 8 and 9 were applied for each stream i , and the EWMA values ($Y_{\bar{d}_i}$) were compared with the control limits.

To obtain the steady-state ARL_1 's, a "warm-up" period of 50 samples (from $t=1$ to $t=50$) was allowed before inserting the simulated mean shift in one of the streams. In case of an alarm during the first 50 samples, the run was disconsidered and the "warm-up" period restarted (with $Y_{0i}=0$ for all streams).

The EWMA versions of the charts proposed by Mortell & Runger (1995) and Runger et al. (1996) were also simulated for purposes of performance comparison. The same series of x_{tij} values generated in each simulation run was used for all the three schemes.

With EWMA charts, the ARL_0 constraint fixes the relationship between the smoothing constant λ and the control limits width factor k ; so, given the former, the latter is uniquely determined (see for example Crowder, 1989, for curves of $k \times \lambda$ — also called *iso- ARL_0 curves* — for the univariate EWMA chart for monitoring the process mean). In the case of the GEWMA- \bar{d} chart, this relation will depend also on the number of streams (m). For illustration, Figure 1 shows the GEWMA- \bar{d} iso- ARL_0 curves for $m=20$ streams, for 3 different ARL_0 values. These curves were obtained by simulation and search: discretizing the range of possible values for λ , for each value of λ considered, a univariate search was conducted for the value of k that yielded the desired ARL_0 (computed by simulation). The dots on the curves in the Figure are the points (λ, k) thus obtained. Other points in the curve can be obtained by linear interpolation. So, each curve contains all points (λ, k) that give a same ARL_0 . These curves are helpful in the stage of optimization of the chart, described in detail the next section. In summary, given the specified ARL_0 , the goal is to find the pair (λ, k) in the respective curve that yields the shortest ARL_1 for a shift of interest in the mean in one stream.

For a fair performance comparison of the (optimized) EWMA- \bar{d} chart with Mortell & Runger's (1995) EWMA- R_t chart and Runger et al. (1996) MEWMA- S^2 chart, we must use the optimal designs of these charts too. As a prerequisite to the optimization, we determined, in the same way, iso- ARL_0 curves for the EWMA- R_t and the MEWMA- S^2 charts. Figures 2 and 3 exhibit the curves for 20-stream processes and 3 different ARL_0 values. Curves for other numbers of streams are given in Simões (2010) and are available upon request. To the best of our knowledge, there is no work in the literature that gives the $k \times \lambda$ curves for these charts.

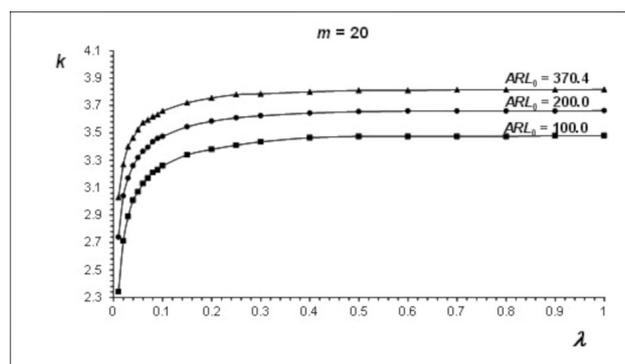


Figure 1. $k \times \lambda$ curves for the GEWMA- \bar{d} chart for 20 streams

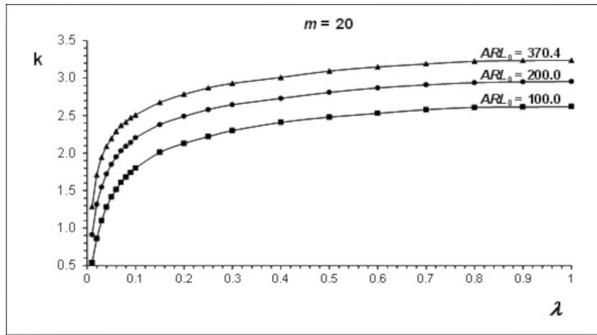


Figure 2. $k \times \lambda$ curves for the EWMA- R_i chart for 20 streams

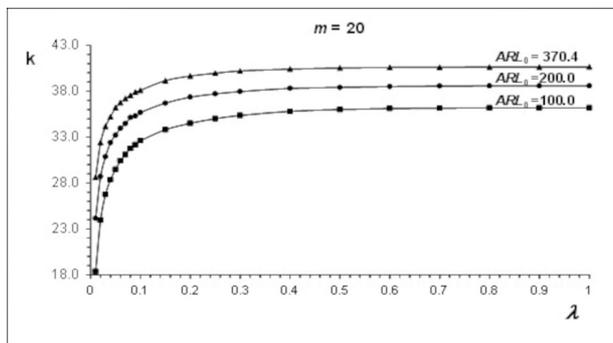


Figure 3. $k \times \lambda$ curves for the MEWMA S^2 chart for 20 streams

5 Optimal designs

With m , n , ARL_0 and δ^* given, we obtain the best design for the GEWMA- \bar{a} chart solving the following optimization problem (Equation 12):

$$\text{Min } ARL_1(\delta^*, \theta) \tag{12}$$

subject to (Equation 13)

$$ARL(\delta = 0, \theta) = ARL_0 \tag{13}$$

where:

$\delta = (\mu_1 - \mu_0) / \sigma_0$ is a shift in the in-control mean of an individual stream component e_i , represented in standard deviation units; note that, as we assume that the in-control mean of any individual stream component is zero (that is, $E(e_{ij}) = 0$), $\delta = \mu_1 / \sigma_0$.

δ^* is the smallest shift in the mean of the individual component of a stream that is considered relevant to detect with minimum delay (for which ARL_1 should be minimized); θ represents the set of parameters of the chart: $\theta = \{n, m, k(\lambda)\}$, where the notation $k(\lambda)$ indicates that, given the desired ARL_0 , k is uniquely determined by the respective iso- ARL_0 curve as a function of λ , as just seen in the preceding section.

With k thus determined as a function of λ , and with n and m as input data, λ becomes the only decision variable, so the optimal solution can be found by any univariate search

method. We chose to use the Golden Section Search method (Kiefer, 1953), programmed in Matlab™. For each value of λ , the ARL_1 for the δ^* specified was obtained by simulation. The search was considered to have converged when the last interval was smaller than 0.005. This termination criterion was based on previous observation that the ARL_1 as a function of λ is quite flat around the minimum. The midpoint of the last interval was then considered to be the optimal solution.

The optimal values of λ and k , under the constraint of $ARL_0 = 200$, are shown in the columns with headings λ^* and k^* of Table 2 (for $n = 1$) and Table 3 (for $n = 4$), for $m = 5, 10$ and 20 streams and $\delta^* = 0.5, 1.0$ and 2.0 . Larger values of δ^* were not considered because EWMA schemes are devised for the detection of small to moderate changes in the process parameters. Optimal designs for other numbers of streams in the process (namely, $m = 2, 3$ and 15) and also for other ARL_0 values ($ARL_0 = 100$ and 370.4) can be found in Simões (2010) and can be obtained by request to the first author.

As seen in these tables, we also optimized the designs of the EWMA- R_t and MEWMA- S^2 charts (for the same values of m, n, δ^* and ARL_0), for the purposes of performance comparison (which is the subject of the next section). The optimization problem statement and the solution method are the same for all charts; only the monitoring statistic and control limits differ. The monitoring statistics of the EWMA- R_t and MEWMA S^2 charts were described in Section 2. We now describe how their control limits are related to k , for full understanding of their designs in Tables 2 and 3.

Note that these charts have only one (upper) control limit (UCL), because their statistics are positive and only increases in their expected values indicate out-of-control situations. The UCL of the EWMA- R_t chart is given by Equation 14

$$UCL = \left(d_2 + kd_3 \sqrt{\frac{\lambda}{2-\lambda}} \right) \frac{\sigma_0}{\sqrt{n}} \quad (14)$$

Where, again, k denotes the control limit factor, and d_2 and d_3 are the mean and standard deviation of the relative range for normal data, which are tabulated in many statistical and SPC books. Note that, here, the “sample size” to which d_2 and d_3 correspond is the number of streams, m , rather than the number of observations per stream, n , in time t .

In the case of the MEWMA S^2 chart, k is the very control limit (that is, $UCL = k$), because when the process is in control the MEWMA statistic (W_t) has unit variance. Indeed, in sampling time t , the statistic is (using our notation for consistency) (Equation 15)

$$W_t = \frac{n(2-\lambda)}{\sigma_0^2 \lambda} \sum_{i=1}^m (Z_{ti} - \bar{Z}_t)^2 \quad (15)$$

where Z_{ti} is the exponentially smoothed value of the i -th stream and \bar{Z}_t is the sample average (over i) of the Z_{ti} 's. Since the variance of $\sum_{i=1}^m (Z_{ti} - \bar{Z}_t)^2$ is $(\sigma_0^2/n)\lambda/(2-\lambda)$, W_t has unit variance.

When the process is in control, W_t is distributed as a chi-squared variable with $m-1$ degrees of freedom; this is why the values of k for this MEWMA chart in Tables 2 and 3 increase noticeably with m . In Runger et al. (1996), in the expression of W_t the factor that appears pre-multiplying the summation is $\sigma_0^2 \lambda / (2-\lambda)$, which is clearly a typo.

The column k^* of Tables 2 and 3 gives also, in one row for each pair (m, n) , the control limit factor k of the Shewhart-type *residuals* GCC of Epprecht et al. (2011a).

The control limits for this chart are set at $\pm k\sigma_0\sqrt{(m-1)/(mn)}$. Note that this expression corresponds to Equations 10a and 10b in the particular case $\lambda=1$. Indeed, the Shewhart chart can be seen as a particular case of an EWMA chart with $\lambda=1$. With λ thus fixed (and considering n given), the constraint on ARL_0 determines uniquely the value of k , so (unlike with the EWMA charts) there is no possibility of optimizing the Shewhart chart for a given δ^* ; this is why for this chart the values of δ^* are left blank and there is only one value of k for each pair (m, n) , while for the EWMA charts λ^* , k^* and the ARL_1 profile depend on the value of δ^* besides of the values of (m, n) .

Table 2. Optimal designs (λ^* and k^*) and ARL profiles, for $ARL_0 = 200.0$ and $n = 1$.

m	δ^*	Chart	λ^*	k^*	Steady-state ARL_1						
					δ						
					0.5	1.0	1.5	2.0	3.0	4.0	
5	---	residuals GCC	1.000	3.290	154.0	74.8	31.4	13.5	3.6	1.6	
		GEWMA- \bar{d}	0.035	2.715	33.0	14.3	9.2	6.8	4.5	3.5	
	0.5	EWMA- R_t	0.013	1.037	101.9	34.7	16.3	9.6	4.9	3.3	
		MEWMA- S^2	0.037	10.506	33.8	14.7	9.3	6.9	4.6	3.5	
	1.0	GEWMA- \bar{d}	0.111	3.055	40.7	12.8	7.2	5.1	3.3	2.5	
		EWMA- R_t	0.013	1.037	101.9	34.7	16.3	9.6	4.9	3.3	
	2.0	MEWMA- S^2	0.100	12.730	40.2	13.3	7.7	5.4	3.5	2.7	
		GEWMA- \bar{d}	0.294	3.223	67.8	16.9	7.3	4.4	2.6	1.9	
	2.0	EWMA- R_t	0.154	2.399	121.5	40.8	14.9	7.5	3.3	2.2	
		MEWMA- S^2	0.318	14.406	72.1	18.8	8.1	4.8	2.7	1.9	
	10	---	residuals GCC	1.000	3.480	172.2	97.3	40.5	16.1	3.7	1.6
			GEWMA- \bar{d}	0.035	2.978	35.5	15.1	9.5	7.1	4.7	3.6
0.5		EWMA- R_t	0.013	1.027	123.3	45.9	20.5	11.5	5.4	3.4	
		MEWMA- S^2	0.037	18.299	38.2	16.4	10.4	7.7	5.1	3.9	
1.0		GEWMA- \bar{d}	0.088	3.236	41.7	13.3	7.7	5.5	3.6	2.7	
		EWMA- R_t	0.013	1.027	123.3	45.9	20.5	11.5	5.4	3.4	
2.0		MEWMA- S^2	0.113	21.381	49.6	15.3	8.4	5.8	3.7	2.8	
		GEWMA- \bar{d}	0.234	3.410	73.3	16.2	7.3	4.6	2.7	2.0	
2.0		EWMA- R_t	0.131	2.318	137.0	53.6	19.8	9.3	3.7	2.3	
		MEWMA- S^2	0.220	22.624	70.6	18.6	8.5	5.4	3.1	2.3	
20		---	residuals GCC	1.000	3.662	182.9	118.0	52.5	20.8	4.3	1.7
			GEWMA- \bar{d}	0.035	3.215	39.0	16.1	10.1	7.5	5.0	3.8
	0.5	EWMA- R_t	0.013	1.032	144.2	61.2	27.0	14.3	6.0	3.6	
		MEWMA- S^2	0.034	31.470	44.0	19.2	12.2	8.9	5.9	4.5	
	1.0	GEWMA- \bar{d}	0.101	3.476	51.5	14.4	8.0	5.6	3.6	2.7	
		EWMA- R_t	0.013	1.032	144.2	61.2	27.0	14.3	6.0	3.6	
	2.0	MEWMA- S^2	0.089	35.283	56.3	18.2	10.2	7.1	4.5	3.3	
		GEWMA- \bar{d}	0.284	3.620	101.8	21.5	8.2	4.8	2.7	2.0	
	2.0	EWMA- R_t	0.123	2.286	153.4	73.4	27.9	12.1	4.3	2.4	
		MEWMA- S^2	0.267	37.785	98.5	29.2	11.6	6.6	3.5	2.5	

Table 3. Optimal designs (λ^* and k^*) and ARL profiles, for $ARL_0 = 200.0$ and $n = 4$.

m	δ^*	Chart	λ^*	k^*	Steady-state ARL_1						
					δ						
					0.5	1.0	1.5	2.0	3.0	4.0	
5	---	residuals GCC	1.000	3.290	74.8	13.5	3.6	1.6	1.0	1.0	
		GEWMA- \bar{d}	0.111	3.055	12.8	5.1	3.3	2.5	1.8	1.4	
		EWMA- R_t	0.013	1.037	34.7	9.6	4.9	3.3	2.1	1.6	
	0.5	MEWMA- S^2	0.100	12.730	13.3	5.4	3.5	2.7	1.9	1.5	
		GEWMA- \bar{d}	0.294	3.223	16.9	4.4	2.6	1.9	1.3	1.0	
		EWMA- R_t	0.154	2.399	40.8	7.5	3.3	2.2	1.4	1.1	
	1.0	MEWMA- S^2	0.318	14.406	18.8	4.8	2.7	1.9	1.3	1.0	
		GEWMA- \bar{d}	0.766	3.283	48.2	7.6	2.6	1.5	1.0	1.0	
		EWMA- R_t	0.623	2.892	61.3	10.2	3.2	1.7	1.1	1.0	
	2.0	MEWMA- S^2	0.748	14.819	47.2	8.1	2.9	1.6	1.0	1.0	
		residuals GCC	1.000	3.480	97.3	16.1	3.7	1.6	1.0	1.0	
		GEWMA- \bar{d}	0.088	3.236	13.3	5.5	3.6	2.7	1.9	1.5	
10	0.5	EWMA- R_t	0.013	1.027	45.9	11.5	5.4	3.4	2.1	1.6	
		MEWMA- S^2	0.113	21.381	15.3	5.8	3.7	2.8	1.9	1.6	
		GEWMA- \bar{d}	0.234	3.410	16.2	4.6	2.7	2.0	1.4	1.1	
	1.0	EWMA- R_t	0.131	2.318	53.6	9.3	3.7	2.3	1.4	1.1	
		MEWMA- S^2	0.220	22.624	18.6	5.4	3.1	2.3	1.6	1.2	
		GEWMA- \bar{d}	0.766	3.478	63.2	8.5	2.7	1.5	1.0	1.0	
	2.0	EWMA- R_t	0.676	2.873	83.7	13.6	3.7	1.8	1.1	1.0	
		MEWMA- S^2	0.725	23.563	62.3	10.8	3.4	1.9	1.1	1.0	
		residuals GCC	1.000	3.662	118.0	20.8	4.3	1.7	1.0	1.0	
	20	0.5	GEWMA- \bar{d}	0.101	3.476	14.4	5.6	3.6	2.7	1.9	1.5
			EWMA- R_t	0.013	1.031	61.2	14.3	6.0	3.6	2.0	1.5
			MEWMA- S^2	0.089	35.283	18.2	7.1	4.5	3.3	2.3	1.8
1.0		GEWMA- \bar{d}	0.284	3.620	21.5	4.8	2.7	2.0	1.3	1.0	
		EWMA- R_t	0.123	2.285	73.4	12.1	4.3	2.4	1.4	1.1	
		MEWMA- S^2	0.267	37.785	29.2	6.6	3.5	2.5	1.7	1.2	
2.0		GEWMA- \bar{d}	0.761	3.660	80.0	9.8	2.8	1.6	1.0	1.0	
		EWMA- R_t	0.522	2.823	98.6	16.7	4.1	1.9	1.1	1.0	
		MEWMA- S^2	0.657	38.528	75.9	13.8	4.2	2.2	1.2	1.0	

6 Performance analysis

Tables 2 and 3 show also the steady-state ARL_1 profiles (steady-state ARL_1 values for a range of shifts in the mean of one stream) of the GEWMA- \bar{d} chart, the residuals GCC of Epprecht et al. (2011a), the EWMA- R_t chart of Mortell & Runger (1995) and the MEWMA- S^2 chart of Runger et al. (1996), for process with 5, 10 and 20 streams, and for the cases of 1 and 4 observations per stream. The ARLs for the shifts for which the charts were optimized (δ^*) are in bold. Recall that, being a Shewhart-type chart, the residuals GCC with fixed n and ARL_0 constraint does not admit optimization. Tables of ARL profiles of the EWMA charts for $m = 2, 3$ and 15 and also for $ARL_0 = 100$ and 370.4 are available upon request to the first author. For ARL profiles of the residuals GCC, refer to Epprecht et al. (2011a).

It can be seen that the GEWMA- \bar{d} chart outperforms the other EWMA charts practically always: the ARL reduction relative to the MEWMA- S^2 chart is smaller (but consistent), while the reduction relative to the EWMA- R_t chart is substantial. The only exception was with $n = 1$ and $\delta^* = 0.5$. In this case, for shifts of 4 standard deviations, the ARL of the EWMA- R_t chart was between 5% and 6% smaller. Nevertheless, given the two-component nature of the process variance (see Equation 2), we may expect that a shift of 0.5 standard deviations of the individual component (i.e., $0.5\sigma_0$) is likely to be substantially smaller in terms of the total standard deviation and, as a result, not typically relevant, so that in most practical situations the charts should be optimized for shifts larger than $\delta^* = 0.5$.

The residuals GCC was considered just for completeness and for reference, since it was expected that its EWMA counterpart (the GEWMA- \bar{d} chart), and also the other EWMA charts, yielded much lower ARL_1 values. As a matter of fact, the residuals GCC, with $n = 1$, outperforms the GEWMA- \bar{d} chart for $\delta \geq 4$ (and for $\delta \geq 3$ in the case of $\delta^* = 0.5$ with $m = 10$ and 20). With $n = 4$, in the case $\delta^* = 0.5$ or 1.0, the residuals GCC outperforms the GEWMA- \bar{d} chart for $\delta \geq 2$; in the case $\delta^* = 2.0$, the GEWMA- \bar{d} chart is always the most efficient of all charts considered. Note that EWMA charts are devised for quicker detection of small and moderate shifts; for large shifts it is natural that Shewhart-type charts provide shorter ARLs (being, nevertheless, much slower in the detection of small and moderate shifts). So the GEWMA- \bar{d} chart achieves its objective, being, for the detection of small and moderate shifts, the most efficient among the EWMA charts considered. The comparison with the residuals GCC was useful to show in which situations (for which sizes of shift) each one is preferable.

To provide a general picture of the quantitative improvement attained with the new chart, Figure 4 shows the ARL curves of the three charts with $n=1$ for the case of 20 streams, optimized for $\delta^* = 0.5$. For a better visualization, due to the large variation in the magnitude of the ARLs with the mean shifts, the figure was split into two parts with different vertical scales, one for shifts up to 1.5 standard deviations and the other for shifts from 1.5 to 4.0 standard deviations.

Figure 5 is similar to Figure 4, now for the case of subgroups of size $n=4$ per stream. The same pattern can be observed. The GEWMA- \bar{d} chart uniformly outperforms the other EWMA charts: it slightly outperforms the MEWMA- S^2 chart and substantially outperforms the EWMA- R_t chart. They are outperformed by the residuals GCC only for the largest shifts.

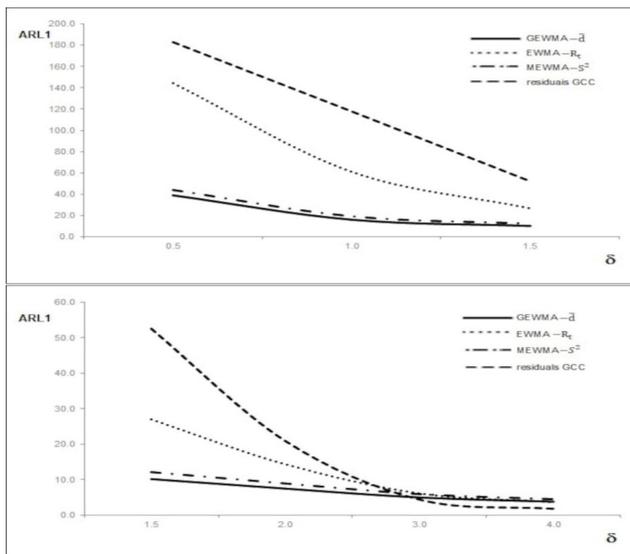


Figure 4. ARL_1 profiles for $n = 1$ ($m = 20$ streams, $ARL_0 = 200.0$, optimized for $\delta = 0.5$).

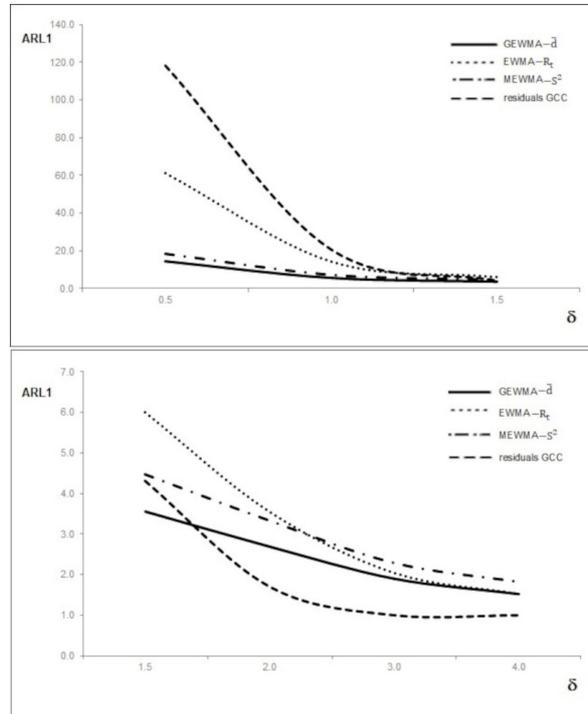


Figure 5. ARL_1 profiles for $n = 4$ ($m = 20$ streams, $ARL_0 = 200.0$, optimized for $\delta = 0.5$).

6 Conclusions

Our objective, as stated in the Introduction section of this paper, was to propose of an EWMA group control chart (GCC) on the smallest and the largest estimates of the individual streams components of a MSP, optimize it and analyze its performance. This chart is the EWMA counterpart of Epprecht et al. (2011a) Shewhart-type *residuals* GCC. For this, we would first need to define the computation of the EWMA statistic to use, the calculation of the chart parameters, and describe its operation. Our motivation was the expectation that such a chart might be faster in detecting small and moderate shifts in the mean of individual streams of the process than the EWMA versions of Mortell & Runger's (1995) R_t chart (EWMA- R_t) and Runger et al. (1996) S^2 chart (MEWMA- S^2). So a second objective was to compare the performances of the proposed chart with the previous ones, in equivalent conditions. This required optimizing all the charts for a variety of situations and for different sizes of shift in the mean of an individual stream, compute their ARLs and compare them side by side. The objective of the optimization was to obtain the design that yields the shortest out-of-control ARL (ARL_1) possible for a given shift (specified), subject to a constraint on the value of ARL_0 .

We succeeded in achieving all these objectives, and named our chart the GEWMA- \bar{d} chart. The comparative performance analysis confirmed our expectations: for the detection of shifts in the mean of the individual component of a single stream, the GEWMA- \bar{d} chart has shown to outperform, in virtually all the cases analyzed, the EWMA- R_t and MEWMA- S^2 charts. Its ARLs were smaller than the ones of the latter and dramatically smaller than the ones of the former.

The only exception observed was when the charts are used with a sample size of 1 and are optimized for a shift of 0.5 standard deviations. In this case, and only for very large shifts (of the order of 4 or more standard deviations), the ARL of the EWMA- R_t chart

was smaller. Nevertheless, given the two-component nature of the process variance, we may expect that a shift of 0.5 standard deviations of the individual component is likely not be typically relevant and that in most practical situations the charts should be optimized for shifts larger than this value. In this case, the proposed chart uniformly outperforms the other two. The differences in performance (i.e., the advantages of the proposed chart) increase with the number of streams.

For signalling larger shifts (of the order of 3 or more standard deviations), the Shewhart-type residuals GCC of Epprecht et al. (2011a) is the fastest chart; this was expected since the EWMA charts are devised for protection against small and moderate shifts and Shewhart charts tend to be faster against larger shifts.

An additional contribution of this work are the optimal designs obtained for the EWMA- R_t and MEWMA- S^2 charts, which had not been given in the previous literature.

As a conclusion, the GEWMA- \bar{d} chart is recommended for the monitoring of small and moderate shifts of the mean of individual streams, being in this case the fastest of the charts considered.

Acknowledgements

The authors have been partly supported by the CNPq (Brazilian Council for the Scientific and Technological Development), project numbers 140933/2006-9, 302326/2008-1 and 308677/2015-3. They are grateful to the anonymous reviewers whose comments contributed to improve the quality of this paper.

References

- Amin, R. W., & Li, K. (2000). The MaxMin EWMA tolerance limits. *International Journal of Quality & Reliability Management*, 17(1), 27-41. <http://dx.doi.org/10.1108/02656710010283351>.
- Boyd, B. F. (1950). Applying the group chart for \bar{d} and R. *Industrial Quality Control*, 3, 22-25.
- Crowder, S. V. (1989). Design of Exponentially Weighted Moving Average Schemes. *Journal of Quality Technology*, 21(3), 155-162. <http://dx.doi.org/10.1080/00224065.1989.11979164>.
- Epprecht, E. K. (2015). Statistical control of multiple-stream processes: a literature review. In S. Knoth & W. Schmid, (Eds.), *Frontiers in statistical quality control 11* (p. 49-64). Springer International Publishing, Switzerland. http://dx.doi.org/10.1007/978-3-319-12355-4_4.
- Epprecht, E. K., Barbosa, L. F. M., & Simões, B. F. T. (2011a). SPC of multiple stream processes – a chart for enhanced detection of shifts in one stream. *Produção*, 21(2), 242-253. <http://dx.doi.org/10.1590/S0103-65132011005000022>.
- Epprecht, E. K., de Luna, M. A., & Aparisi, F. (2011b). Joint EWMA charts for multivariate process control: markov chain and optimal design. *International Journal of Production Research*, 49(23), 7151-7169. <http://dx.doi.org/10.1080/00207543.2010.537383>.
- Kiefer, J. (1953). Sequential minimax search for a maximum. *Proceedings of the American Mathematical Society*, 4(3), 502-506. <http://dx.doi.org/10.1090/S0002-9939-1953-0055639-3>.
- Lanning, J. W., Montgomery, D. C., & Runger, G. C. (2002). Monitoring a multiple stream filling operation using fractional samples. *Quality Engineering*, 15(2), 183-195. <http://dx.doi.org/10.1081/QEN-120015851>.
- Liu, X., Mackay, R. J., & Steiner, S. H. (2008). Monitoring multiple stream processes. *Quality Engineering*, 20(3), 296-308. <http://dx.doi.org/10.1080/08982110802035404>.
- Lowry, C. A., Woodall, W. H., Champ, C. W., & Rigdon, S. E. (1992). A multivariate exponentially weighted moving average control chart. *Technometrics*, 34(1), 46-53. <http://dx.doi.org/10.2307/1269551>.

- Meneças, N. S., Olivera, S. A., Saccone, C. D., & Tessore, J. (2008). Statistical control of multiple-stream processes: a shewhart control chart for each stream. *Quality Engineering*, 20(2), 185-194. <http://dx.doi.org/10.1080/08982110701241608>.
- Mortell, R. R., & Runger, G. C. (1995). Statistical process control of multiple stream processes. *Journal of Quality Technology*, 27(1), 1-12. <http://dx.doi.org/10.1080/00224065.1995.11979554>.
- Nelson, L. S. (1986). Control charts for multiple stream processes. *Journal of Quality Technology*, 18(4), 255-256. <http://dx.doi.org/10.1080/00224065.1986.11979020>.
- Runger, G. C., Alt, F. B., & Montgomery, D. C. (1996). Controlling multiple stream processes with principal components. *International Journal of Production Research*, 34(11), 2991-2999. <http://dx.doi.org/10.1080/00207549608905074>.
- Simões, B. F. T. (2010). *Controle estatístico de processos multicanal* (Tese de doutorado). Pontifícia Universidade Católica do Rio de Janeiro - PUC-Rio, Rio de Janeiro.
- Xiang, L., & Tsung, F. (2008). Statistical monitoring of multi-stage processes based on engineering models. *IIE Transactions*, 40(10), 957-970. <http://dx.doi.org/10.1080/07408170701880845>.