EMPIRICAL COMPARISONS OF X-BAR CHARTS WHEN CONTROL LIMITS ARE ESTIMATED

Paper published in

Quality and Reliability Engineering International

Full citation to this publication:

Muñoz-Rosas, J. F., Álvarez-Verdejo, E., Pérez-Aróstegui, M. N., & Gutiérrez-Gutiérrez, L. (2016): "Empirical Comparisons of X-bar Charts when Control Limits are Estimated". *Quality and Reliability Engineering International*, Vol.32, n.2, pp.453-464.

DOI: https://doi.org/10.1002/qre.1763

Thank you for your interest in this publication.

EMPIRICAL COMPARISONS OF X-BAR CHARTS WHEN CONTROL LIMITS ARE ESTIMATED

Muñoz-Rosas, J.F. Álvarez-Verdejo, E., Pérez-Aróstegi, M.N., Gutiérrez-Gutiérrez, L.

University of Granada

Emails: jfmunoz@ugr.es, encarniav@ugr.es , mnperez@ugr.es, leogg@ugr.es

ABSTRACT

A control chart is a very common tool used to monitor the quality of business processes. An estimator of the process variability is generally considered to obtain the control limits of a $\overline{X}$ chart when parameters of the process are unknown. Assuming Monte Carlo simulations, this paper firstly compares the efficiency of the various estimators of the process variability. Two empirical measures used to analyze the performance of control charts are defined. Results derived from various empirical studies reveal the existence of a linear relationship between the performance of the various estimators of the process variability and the performance of $\overline{X}$ charts. The various Monte Carlo simulations are conducted under the assumption that the process is in both situations of in-control and out-of-control.

Key-words: Monte Carlo simulations; average run length; Type I error; mean square error

1. INTRODUCTION

Research on quality involves a range of concerns about definitions, practices and such specific mechanisms as statistical quality control (SQC). These techniques are used in many industries to control the quality of the product by analyzing one or more product characteristics. The most widely used tools in SQC are control charts. A control chart is a powerful tool used to determine if a business or manufacturing process is in a state of statistical control. Shewhart[1] developed this technique and provided a framework for deciding whether the variation in the result is due to assignable causes.

An indispensable assumption for the correct development of control charts is that the parameters related to the quality characteristic of the process are assumed known. However, such parameters are usually unknown in practice. In this situation, control charts are based upon a two-phase procedure. In phase I, m representative samples or subgroups, with size n, are used to estimate the parameters of the process. Control limits are also estimated in phase I and they are used in phase II, in which samples from the process and with size n are selected to study if the process is stable or in-control. It is customary to say that the process is out-of-control when a chart statistics related to a sample from phase II is plotted outside the control limits.

Recent research indicates that the $\overline{X}$ charts are very simple to understand, implement and design, and may be more suitable in many SQC applications (see Montgomery[2], Yang et al.[3]). When the parameters of the process are unknown and they need to be estimated in phase I, the most common statistics used to estimate the variability of the process are based on the sample standard deviations, the sample ranges or the pooled sample standard deviation. As far as the estimation of the mean of the process is concerned, it is common to consider the grand average of the subgroup averages. Therefore, it is very common in practice to use $\overline{X}$ charts based upon the grand average and some of the previously commented methods to estimate the process variability.

The performance of a $\overline{X}$ chart can depend on various aspects, such as the sample sizes or the number of samples used in phase I to obtain the control limits. The method used to estimate the process variability can have also an impact on the performance of the $\overline{X}$ chart. In addition, it is a common practice to use the average run length (*ARL*) as a measure of the performance of control charts (see, for example, Chen[4], Jones et al.[5], Jensen et al.[6], Chakraborti et al.[7]). The main aim of this paper is to analyze the performance of $\overline{X}$ charts under different possible scenarios and assuming various empirical measures defined in this paper. In addition, we also analyze the relationship between the performance of $\overline{X}$ charts and the corresponding estimator used for the process variability. Note that the performance of control charts can be measured by using techniques such as Monte Carlo simulations (see, for example, Li and Pu[8], Mahmoud et al.[9]) and Factorial experiments (Ou et al.[10]).

This paper is organized as follows. In Section 2 and assuming the theoretical definition of *ARL*, we define two empirical measures, which can be used to evaluate the performance of control charts. Assuming Monte Carlo simulations and different sample sizes, various empirical studies and analysis are carried out in this paper. First, we compare numerically, in Section 3, various common estimators of the process variability in terms of bias and mean square error. This is an important aspect to analyze, since it can be seen the empirical performance of the various methods under different situations. Second, we evaluate, in Section 4, the various $\overline{X}$ charts in terms of the suggested empirical measures defined in Section 2, and they are compared to the theoretical value of *ARL*. Results derived from studies of Sections 3 and 4 reveal a possible linear relationship between the performance of the various estimators of the process variability and the performance of $\overline{X}$ charts. Another contribution in this paper is to analyze, in Section 5, the linear relationship between the performance of estimators of the process variability to the performance of $\overline{X}$ charts. Sections 4 and 5 assume that the process is in-control in phase II, whereas studies derived in Section 6 assume that the process is out-of-control in phase II. Conclusions are summarized in Section 7.

2. TWO EMPIRICAL MEASURES TO EVALUATE $\overline{X}$ CHARTS

Let $x \to N(\mu, \sigma)$ be the variable of interest for a process variable, where μ is the true process mean and σ is the true process standard deviation. The process can be monitored by plotting on the $\overline{X}$ chart the sample means

$$\overline{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij}$$

obtained from samples with size *n* and taken in sequence, and where x_{ij} denotes the value of quality characteristics for the *j*th product in the *i*th sample. When parameters μ and σ are known, the $\overline{X}$ chart with 3σ control limits is given by

$$LCL = \mu - 3\frac{\sigma}{\sqrt{n}} \quad ; \quad CL = \mu \quad ; \quad UCL = \mu + 3\frac{\sigma}{\sqrt{n}} \tag{1}$$

We observe that previous control limits depend on the parameters μ and σ, i.e., it is assumed that the standards μ and σ are known. However, the standards are generally unknown in practice, hence they need to be estimated in phase I by using the sample information. Assuming this scenario, control limits are estimated by using m previous samples with size n and taken when the process is believed to be in-control. Expressions for the estimated control limits are:

$$L\hat{C}L = \overline{\overline{x}} - 3\frac{\hat{\sigma}}{\sqrt{n}} \quad ; \quad \hat{C}L = \overline{\overline{x}} \quad ; \quad U\hat{C}L = \overline{\overline{x}} + 3\frac{\hat{\sigma}}{\sqrt{n}} \tag{2}$$

where $\overline{\overline{x}} = m^{-1}\sum_{i=1}^{m}\overline{x}_i$ is the grand average of the m subgroup averages and $\hat{\sigma}$ is an estimator of σ. The most used estimators of $\hat{\sigma}$ are defined in Section 3.

We consider the most popular measure used to analyze the performance of control chart, and this measure is the average run length (*ARL*). The *ARL* indicates the average number of points that will be plotted on a control chart before the process is out-of-control. For example, we can consider that the process is out-of-control when a point falls outside the control limits. This measure is used, for example, by Chen[4], Jones et al.[5], Jensen et al.[6], Abbas[11]. When the process is in-control, the *ARL* can be defined as

$$ARL = \frac{1}{\alpha} = \frac{1}{1 - P\left(LCL < \overline{x}_i < UCL\right)}, \tag{3}$$

where

$$\alpha = P(\text{reject } H_0 | H_0 \text{ is true}) = 1 - P\left(LCL < \overline{x}_i < UCL\right)$$

is the probability of a Type I error associated to the hypothesis test

$$\begin{cases} H_0 : \text{The process is in - control} \\ H_1 : \text{The process is out - of - control} \end{cases}$$

Assuming that the standards are given, control limits are based upon the 3σ criteria and the process is in-control, it is well known that $\alpha = 0.0027$ and $ARL = 370.4$.

It is said that the process is out-of-control when the process is shifted and/or changed to a status with mean μ_0 and standard deviation σ_0. In this situation, the *ARL* is defined as

$$ARL = \frac{1}{1-\beta} = \frac{1}{1 - P\left(LCL < \overline{x}_i < UCL | \mu_0, \sigma_0\right)} \tag{4}$$

where

$$\beta = P(\text{fail to reject } H_0 | H_0 \text{ is false}) = P\left(LCL < \bar{x}_i < UCL | \mu_0, \sigma_0\right) \qquad (5)$$

is the probability of a Type II error.

A simulation study based upon Monte Carlo simulations is a technique generally used to analyze empirically the performance of procedures such as estimators or confidence intervals. Some references that evaluate estimation procedures via Monte Carlo simulations are Li and Pu[8], Mahmoud et al.[9], Rao et al.[12], Silva and Skinner[13] and Muñoz and Rueda[14].

We now define two different empirical measures to analyze the performance of $\bar{X}$ charts. Such measures are based upon a Monte Carlo simulation, which is also defined as follows. First, we assume that the process is in-control, i.e., it has normal distribution with parameters μ and σ. At the first iteration run, control limits are estimated in phase I by selecting *m* samples with size *n* from a normal distribution with mean μ and standard deviation σ. When control limits are estimated, D_{II} samples with size *n* are selected in phase II from the same normal distribution, i.e, it is also assumed that the process is in-control in phase II. Then, it is studied whether such samples fall outside the estimated control limits. At the second iteration run, *m* new samples with size *n* are obtained in order to calculate the new control limits and D_{II} new samples are selected in phase II to analyze if they are out-of-control. This process is repeated D_I times.

Let us assume that the parameters μ and σ of the process are unknown. The first empirical measure (ARL_1) is directly related to definition of *ARL* given by equations (3) and (4). Let B_{ij} be a variable denoting if the *j*th sample, with $j = 1, \ldots D_{II}$, in the *i*th iteration run, with $i = 1, \ldots D_I$, falls inside the estimated control limits, i.e., $B_{ij} = 1$ if $L\hat{C}L_i < \bar{x}_j < U\hat{C}L_i$ and $B_{ij} = 0$ otherwise. Following the definition of *ARL* given by equations (3) and (4), we define the measure

$$ARL_1 = \frac{1}{1 - E[B]}$$

where the empirical expectation of B is given by

$$E[B] = \frac{1}{D_I}\sum_{i=1}^{D_I} E[B_i] = \frac{1}{D_I}\sum_{i=1}^{D_I}\frac{1}{D_{II}}\sum_{j=1}^{D_{II}} B_{ij} = \frac{1}{D_I D_{II}}\sum_{i=1}^{D_I}\sum_{j=1}^{D_{II}} B_{ij}$$

The second measure is directly defined as the empirical average of the total number of run lengths in the Monte Carlo simulation, i.e:

$$ARL_2 = \frac{1}{D_I}\sum_{i=1}^{D_I}\frac{1}{r_i}\sum_{j=1}^{r_i} L_{ij} = \frac{1}{D_I}\sum_{i=1}^{D_I}\frac{1}{r_i} D_{II} = \frac{D_{II}}{D_I}\sum_{i=1}^{D_I}\frac{1}{r_i}$$

where r_i is the number of runs at the ith iteration, and L_{ij} is the length for the jth run and the ith iteration. A run is defined as the number of points that will be plotted on the $\overline{X}$ chart before a point falls outside the control limits. Assuming that the process is in-control, the empirical measures ARL_1 and ARL_2 should be close to 370.4, which is the theoretical value of the measure ARL defined by equation (3).

When assuming that the process is out-of-control, samples from phase II are generated from a normal distribution with mean μ_0 and standard deviation σ_0, and the previous measures ARL_1 and ARL_2 are similarly defined in this situation.

3. MONTE CARLO SIMULATIONS TO COMPARE ESTIMATORS OF THE PROCESS VARIABILITY

Assuming that parameters of the process (μ and σ) are known, expressions of control limits of the corresponding $\overline{X}$ chart are given by equation (1). However, the standards are usually unknown, and the usual solution is to estimate the control limits, such as described by equation (2).

In practice, the most known methods used to estimate σ are based upon the sample standard deviations, the sample ranges and the pooled sample standard deviation, i.e, such estimators are, respectively, given by

$$\hat{\sigma}_S = \frac{\overline{S}}{c_4[n]} \quad ; \quad \hat{\sigma}_R = \frac{\overline{R}}{d_2[n]} \quad ; \quad \hat{\sigma}_p = \frac{S_p}{c_4[1+m(n-1)]} \tag{6}$$

where $\overline{S} = m^{-1}\sum_{i=1}^{m} S_i$ is the average of the sample standard deviations

$$S_i = \left(\frac{1}{n-1} \sum_{j=1}^{n} \left(x_{ij} - \overline{x}_i \right)^2 \right)^{1/2} ,$$

$\overline{R} = m^{-1}\sum_{i=1}^{m} R_i$ is the average of the simple ranges $R_i = \max(x_{ij}) - \min(x_{ij})$ and

$$S_p = \left(\frac{1}{m} \sum_{i=1}^{m} S_i^2 \right)^{1/2}$$

is the pooled sample standard deviation. $c_4[\cdot]$ and $d_2[\cdot]$ are constants based on the sample sizes and which are defined, for example, by Chen[4]. In addition, Appendix VI from Montgomery[2] contains tables for these constants and for various values of *n*.

Assuming different samples sizes, Monte Carlo simulations are now carried out to compare the performance of the various estimators of σ described by (6). This simulation study consists on generating observations from a normal distribution with mean $\mu = 10$ and standard deviation $\sigma = \{0.5, 1, 2\}$. Thus, we assume that the process has normal distribution with parameters μ and σ. Then, the Monte Carlo simulation described in Section 2 is carried out. We considered values of *m* from 10 to 1000 and values of *n* from 3 to 25. We also considered $D_I = 1000$ iterations and $D_{II} = 1000$ samples (with size *n*) for each iteration.

The various estimators of σ are compared in terms of relative bias (*RB*) and relative root mean square error (*RRMSE*), which are defined as

$$RB = 100 \times \frac{E[\hat{\sigma}] - \sigma}{\sigma} \quad ; \quad RRMSE = 100 \times \frac{\sqrt{MSE[\hat{\sigma}]}}{\sigma}$$

where the empirical expectation $E[\cdot]$ and the empirical mean square error $MSE[\cdot]$ are, respectively, defined as

$$E[\hat{\sigma}] = \frac{1}{D_I} \sum_{i=1}^{D_I} \hat{\sigma}_i \quad ; \quad MSE[\hat{\sigma}] = \frac{1}{D_I} \sum_{i=1}^{D_I} \left(\hat{\sigma}_i - \sigma \right)^2$$

Note that the *RB* and the *RRMSE* are common measures used to compare estimation methods. For example, the *RB* and the *RRMSE* have been used by Rao et al.[12], Silva and Skinner[13] and Muñoz and Rueda[14].

FIGURE 1 ABOUT HERE

***Figure 1**. Values of RB and RRMSE of estimators (standard deviations: $\hat{\sigma}_S$; ranges: $\hat{\sigma}_R$; and pooled standard deviation: $\hat{\sigma}_p$) of the true process standard deviation $\sigma = 1$. Values of m between 10 and 50.*

FIGURE 2 ABOUT HERE

***Figure 2**. Values of RB and RRMSE of estimators (standard deviations: $\hat{\sigma}_S$; ranges: $\hat{\sigma}_R$; and pooled standard deviation: $\hat{\sigma}_p$) of the true process standard deviation $\sigma = 1$. Values of m between 100 and 1000.*

From Figures 1 and 2 we compare the various estimators ($\hat{\sigma}_S$, $\hat{\sigma}_R$ and $\hat{\sigma}_p$) of $\sigma = 1$ in terms of *RB* and *RRMSE*. For the various values of *n* and *m*, we observe that the various estimators of σ have a good performance in terms of bias, since the values of *RB* are all less than 0.5% in relative terms. As far as the efficiency in terms of *RRMSE* is concerned, we observe that the various estimators are more efficient as both values of *n* and *m* increase. Estimators $\hat{\sigma}_S$ and $\hat{\sigma}_p$ perform similarly, and they are slightly more efficient than the estimator $\hat{\sigma}_R$, especially as the value of *n* increases. In practice, the estimator $\hat{\sigma}_R$ is the most common estimator used to obtain the control limits. This is due to the fact that this estimator is the simplest estimator. However, it can be seen from this simulation study that the alternative estimators $\hat{\sigma}_S$ and $\hat{\sigma}_p$ can be more efficient. Results derived when $\sigma = \{0.5, 2\}$ are omitted in this paper, since similar conclusions were obtained.

4. MONTE CARLO SIMULATIONS TO ANALYZE $\overline{X}$ CHARTS WHEN THE PROCESS IS IN-CONTROL

Assuming that the process is in-control and different samples sizes, Monte Carlo simulations are now carried out to compare the performance of the various $\overline{X}$ charts in terms of the empirical measures ARL_1 and ARL_2 defined in Section 2. A linear relationship between results derived from Sections 3 and 4 can be observed, and for this reason, we evaluate numerically, in Section 5, such relationship via a linear correlation coefficient.

The Monte Carlo simulation described in Sections 2 and 3 is carried out. In Sections 4 and 5, we assume that the process is in-control in phase II. In Section 6, we assume that the process in out-of-control in phase II.

FIGURE 3 ABOUT HERE

Figure 3. *Values of* ARL_1 *and* ARL_2 *of the various* $\overline{X}$ *charts (using the standard deviations:* $\hat{\sigma}_S$*; using the ranges:* $\hat{\sigma}_R$*; and using the pooled standard deviation:* $\hat{\sigma}_p$*). The true process standard deviation is* $\sigma = 1$*. Values of m between 10 and 50.*

FIGURE 4 ABOUT HERE

Figure 4. *Values of* ARL_1 *and* ARL_2 *of the various* $\overline{X}$ *charts (using the standard deviations:* $\hat{\sigma}_S$*; using the ranges:* $\hat{\sigma}_R$*; and using the pooled standard deviation:* $\hat{\sigma}_p$*). The true process standard deviation is* $\sigma = 1$*. Values of m between 100 and 1000.*

Figures 3 and 4 analyze the performance of the various $\overline{X}$ charts in terms of the empirical measures ARL_1 and ARL_2 defined in Section 2. Note that the values of ARL_1 and ARL_2 should be close to 370.4, which is the theoretical value of the measure *ARL* defined by (3). We observe that the measure ARL_1 performs generally better than the measure ARL_2. First, we observe that the values of ARL_1 are closer to the required value 370.4 as the values of m and n increase, which is a desirable property. However, this property is not satisfied by ARL_2, which generally takes values between 300 and 350. The values of ARL_1 are very close to 370.4 for large values of m. In terms of ARL_1, we also observe that $\overline{X}$ charts based on $\hat{\sigma}_S$ and $\hat{\sigma}_p$ perform better than $\overline{X}$ charts based on $\hat{\sigma}_R$, since values of ARL_1 are closer to 370.4. These results are consistent with the study of the efficiency in the problem of estimating the parameter σ (see Section 3). In other words, from results of Sections 3 and 4 we observe that the values of *RRMSE* (Figures 1 and 2) and the values of ARL_1 (Figures 3 and 4) can have a strong linear relationship. For this reason, we analyze, in Section 5, this relationship by using the linear correlation coefficient between the values of *RRMSE* and the corresponding values of ARL_1.

5. MONTE CARLO SIMULATIONS TO COMPARE VALUES OF *RRMSE* WITH VALUES OF ARL_1.

In this section, we compare numerically the values of RRMSE obtained in Section 3 with the values of ARL_1 obtained in Section 4. From Table 1 we observe a very strong linear relationship for values of m smaller than 50, which is the most common situation in practice (see, for example, Montgomery[2], Chen[4] and Quesenberry[15]). For larger values of m, $m = \{500, 1000\}$, the relationship is smaller, although the performance of $\overline{X}$ charts and estimators of σ is very good in both cases. This smaller linear relationship can be due to the fact that the values of ARL_1 are close to the required 370.4 for the various values of n. However, the values of RRMSE are slightly decreasing according to n.

TABLE 1 ABOUT HERE

6. MONTE CARLO SIMULATIONS WHEN THE PROCESS IS OUT-OF-CONTROL

We now analyze the performance of $\overline{X}$ charts when the process is out-of-control, i.e., it is shifted and/or changed to a status with mean $\mu_0 = \mu + a\sigma$ and standard deviation $\sigma_0 = b\sigma$, where a and b are constants. Note that the process is in-control when $a = 0$ and $b = 1$, and this situation is analyzed in Sections 4 and 5. The process has a shift in the mean when $a \neq 0$ and $b = 1$, it has a shift in the variance when $a = 0$ and $b \neq 1$, and it has a shift in both the mean and the variance otherwise. Table 2 summarizes the various scenarios discussed in this paper.

TABLE 2 ABOUT HERE

When the process is out-of-control, it can be easily seen that the value of β defined by (5) can be expressed as

$$\beta = \phi\left(\frac{3 - a\sqrt{n}}{b}\right) - \phi\left(\frac{-3 - a\sqrt{n}}{b}\right),$$

hence the theoretical value of *ARL* defined (4) can be easily obtained. $\phi(\cdot)$ denotes the distribution function of a $N(0,1)$ random variable.

Figure 5 gives the values of ARL_1 and ARL_2 for various values of *n*, *m*, *a* and *b*. For reasons of clearness, Figure 5 only contains $\overline{X}$ charts based on the estimator $\hat{\sigma}_R$. Note that alternative estimators ($\hat{\sigma}_S$ and $\hat{\sigma}_p$) of σ give $\overline{X}$ charts with similar results, and for this reason they are omitted. When the process has a shift in the mean ($b = 1$), we observe that the measures ARL_1 and ARL_2 are closer to the theoretical value of ARL as the value of *n* and *m* increase. Similar conclusions can be derived when the process has a shift in both the mean and the variance ($a \neq 0$ and $b \neq 1$). Finally, when the process has a shift in the variance ($a = 0$), we observe a clear improvement in the performance of the $\overline{X}$ charts in terms of ARL_1 for values of *m* larger than 10.

FIGURE 5 ABOUT HERE

Figure 5*. Values of ARL_1 (thicker lines) and ARL_2 (thinner lines) of $\overline{X}$ charts based on $\hat{\sigma}_R$. The true process standard deviation is $\sigma = 1$. The solid line represents the theoretical value of ARL obtained by equation (5).*

7. **CONCLUSIONS**

A $\overline{X}$ chart is a very common tool used to monitor the quality of business processes. In practice, the true process mean (μ) and the true process standard deviation (σ) are unknown, and for his reason control limits of $\overline{X}$ charts need to be estimated in this situation. The most known estimators ($\hat{\sigma}_S$, $\hat{\sigma}_R$ and $\hat{\sigma}_p$) of σ used in practice are defined by (6). In this paper, we first compared numerically (in terms of *RB* and *RRMSE*) the performance of the various estimators of σ. Results derived from this simulation study indicate that the various estimators perform well in terms of relative bias. We also observed that estimators $\hat{\sigma}_p$ and $\hat{\sigma}_S$ are more efficient than the estimator $\hat{\sigma}_R$, especially for samples with larger sample sizes. Therefore, $\hat{\sigma}_p$ and $\hat{\sigma}_S$ are the recommended estimators when *n* is large.

We also defined the empirical measures ARL_1 and ARL_2, which are based on the theoretical definition of *ARL*. Monte Carlo simulations indicate that the performance of $\overline{X}$ charts, in terms of ARL_1, is better as the values of *n* and *m* increase. In general, $\overline{X}$ charts based on $\hat{\sigma}_p$ and $\hat{\sigma}_S$ perform better than $\overline{X}$ charts based on $\hat{\sigma}_R$, since values the corresponding values of ARL_1 are closer to the theoretical value of *ARL*.

The aforementioned studies reveal the existence of a possible linear relationship between the values of *RRMSE* and the values of ARL_1. In other word, the performance of $\overline{X}$ charts can have a strong relationship with the efficiency of estimators of σ. We also carried out Monte Carlo simulations to study this relationship, and results derived from this study reveal a very strong linear relationship for values of *m* smaller than 50, which is a very common situation in practice (see Montgomery[2], Chen[4] and Quesenberry[15]). Finally, recall that the various simulation studies in this paper have been conducted under the assumption that the process was in both situations of in-control and out-of-control.

REFERENCES

1. Shewhart, WA. *Economic control of quality of manufactured product*. Milwaukee, ASQC Quality Press, 1931.

2. Montgomery DC. *Statistical quality control. A modern introduction* (6th edn). New York, Wiley, 2009.

3. Yang M, Wu Z, Lee K, Khoo M. The X control chart for monitoring process shifts in mean and variance. *International Journal of Production Research* 2012; **50:**893–907.

4. Chen G. The mean and standard deviation of the run length distribution of X-bar charts when control limits are estimaterd. *Statistica Sinica* 1997; **7:**789–798.

5. Jones LA, Champ CW, Rigdon SE. The performance of exponentially weighted moving average chartswith estimated parameters. *Technometrics* 2001; **43:**156–167.

6. Jensen WA, Jones-Farmer LA, Champ CW, Woodall WH. Effects of parameter estimation on control chart properties: A literature review. *Journal of Quality Technology* 2006; **38:**349–364.

7. Chakraborti S, Human SW, Grahan MA. Phase I Statistical Process Control Charts: An overview and some results. *Quality Engineering* 2008; **21:**52–62.

8. Li Y, Pu X. On the performance of two-sided control charts for short production runs. *Quality Reliability Engineering International* 2012; **28:**215–232.

9. Mahmoud MA, Saleh NA, Madbuly DF. Phase I Analysis of individual observations with missing data. *Quality and Reliability Engineering International* 2013. DOI: 10.1002/qre.1508.

10. Ou Y, Wen D, Wu Z, Knoo MBC. A comparison study on effectiveness and robustness of control charts for monitoring process mean and variance. *Quality and Reliability Engineering International* 2012; **28:**3–17.

11. Abbas N, Riaz M, Does RJMM. Enhancing the performance of EWMA charts. *Quality and Reliability Engineering International* 2012; **27:**821–833.

12. Rao JNK, Kovar JG, Mantel HJ. On estimating distribution functions and quantiles from survey data using auxiliary information. *Biometrika* 1990; **77:**365–375.

13. Silva PLD, Skinner CJ. Estimating distribution functions with auxiliary information using poststratification. *Journal of Official Statistics* 1995; **11:**277–294.

14. Muñoz JF, Rueda M. New imputation methods for missing data using quantiles. *Journal of Computational and Applied Mathematics* 2009; **232:**305–317.

15. Quesenberry CP. The effect of sample size on estimated limits X-bar and X control charts. *Journal of Quality Technology* 1993; **25:**237–247.

	$\sigma = 0.5$			$\sigma = 1.0$			$\sigma = 2.0$		
m	$\hat{\sigma}_S$	$\hat{\sigma}_R$	$\hat{\sigma}_p$	$\hat{\sigma}_S$	$\hat{\sigma}_R$	$\hat{\sigma}_p$	$\hat{\sigma}_S$	$\hat{\sigma}_R$	$\hat{\sigma}_p$
10	0.99	0.98	0.99	0.97	0.99	0.97	0.97	0.99	0.97
20	0.96	0.95	0.96	0.97	0.96	0.97	0.97	0.96	0.97
30	0.96	0.98	0.96	0.96	0.95	0.96	0.96	0.95	0.96
50	0.95	0.92	0.95	0.94	0.90	0.94	0.94	0.90	0.94
100	0.65	0.46	0.58	0.79	0.72	0.75	0.79	0.72	0.75
200	0.72	0.65	0.75	0.71	0.66	0.64	0.71	0.66	0.64
500	0.53	0.54	0.47	0.45	0.41	0.47	0.45	0.41	0.47
1000	0.40	0.47	0.39	0.03	0.03	0.05	0.03	0.03	0.05

Table 1*. Linear correlation coefficients (in absolute terms) between the values of RRMSE of the various estimators (*$\hat{\sigma}_S$*,* $\hat{\sigma}_R$ *and* $\hat{\sigma}_p$*) of* σ *and the values of* ARL_1.

	b		
a	1.0	1.2	1.4
0.0	In control	Shift-variance	Shift-variance
0.3	Shift-mean	Shift-both	Shift-both
0.6	Shift-mean	Shift-both	Shift-both

Table 2. *Different values of the constants a and b and situation of the process for each case.*