



HAL
open science

Evaluation of CUSUM Charts for Finite-Horizon Processes

George Nenes, George Tagaras

► **To cite this version:**

George Nenes, George Tagaras. Evaluation of CUSUM Charts for Finite-Horizon Processes. Communications in Statistics - Simulation and Computation, 2010, 39 (03), pp.578-597. 10.1080/03610910903528319 . hal-00567041

HAL Id: hal-00567041

<https://hal.science/hal-00567041>

Submitted on 18 Feb 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Evaluation of CUSUM Charts for Finite-Horizon Processes

Journal:	<i>Communications in Statistics - Simulation and Computation</i>
Manuscript ID:	LSSP-2009-0172.R1
Manuscript Type:	Original Paper
Date Submitted by the Author:	30-Nov-2009
Complete List of Authors:	Nenes, George; University of Western Macedonia, Mechanical Engineering Tagaras, George; Aristotle University of Thessaloniki, Mechanical Engineering
Keywords:	CUSUM chart, Short Runs, Statistical Process Control
Abstract:	This paper analyses and evaluates the properties of a CUSUM chart designed for monitoring the process mean in short production runs. Several statistical measures of performance that are appropriate when the process operates for a finite-time horizon are proposed. The methodology developed in this paper can be used to evaluate the performance of the CUSUM scheme for any given set of chart parameters from both an economic and a statistical point of view, and thus, allows comparisons with various other charts.



Evaluation of CUSUM Charts for Finite-Horizon Processes

George Nenes^{1,*} (gnesenes@uowm.gr) and George Tagaras² (tagaras@auth.gr)

¹ University of Western Macedonia
Department of Mechanical Engineering
Bakola & Sialvera, 50100 Kozani, Greece
tel.: +30 2461 056600, fax.: +30 2461 056601

² Aristotle University of Thessaloniki
Department of Mechanical Engineering
54124 Thessaloniki, Greece
tel.: +30 231 0996017, fax.: +30 231 0996018

Abstract:

This paper analyses and evaluates the properties of a CUSUM chart designed for monitoring the process mean in short production runs. Several statistical measures of performance that are appropriate when the process operates for a finite-time horizon are proposed. The methodology developed in this paper can be used to evaluate the performance of the CUSUM scheme for any given set of chart parameters from both an economic and a statistical point of view, and thus, allows comparisons with various other charts.

Key words: CUSUM chart; Short Runs; Statistical Process Control

Revised December 2009

* Corresponding author

1. Introduction

Control charts are widely used for monitoring critical quality characteristics of products and processes. One of the most popular control charts in industry is the Cumulative Sum (CUSUM) chart.

CUSUM charts, which were introduced by Page (1954, 1961), can be used both when the quality characteristic is a continuous random variable (for monitoring the mean and the variance) and when it is a discrete attribute. This paper focuses on CUSUM charts used for monitoring the mean value μ of a continuous quality characteristic X . There are two ways to implement a CUSUM chart for monitoring the mean; the V-mask and the tabular (algorithmic) approach. The V-mask was originally proposed by Barnard (1959) and is applied to successive values of the CUSUM statistic:

$$C_t = \sum_{i=1}^t \frac{\bar{x}_i - \mu_0}{\sigma / \sqrt{n}} = \sum_{i=1}^t z_i$$

where \bar{x}_i is the mean of sample i , μ_0 the target mean, σ the standard deviation of X , n the sample size and C_t the chart's statistic. Statistical properties of CUSUM charts using the V-mask approach have been studied by Johnson (1961) and Goldsmith and Whitfield (1961).

Over the years, the tabular approach has prevailed, mainly because of its simplicity as well as because of the rapid development of computers that facilitated the use of this approach compared to the V-mask. When using the tabular approach to identify upward shifts in the mean, at each sample t after the initial setup the CUSUM statistic is evaluated by:

$$C_t = \max\{0, C_{t-1} + z_t - k\}, \quad C_0 \geq 0 \quad (1)$$

where z_t is the standardized observation and k is the reference value. The formulation for downward shifts is analogous while in case of two-sided charts both CUSUM statistics are used. The CUSUM chart issues an alarm if and when C_t exceeds the control limit H . When the

1
2
3 process setup may be imperfect, the CUSUM scheme is often used with Fast Initial Response
4 (FIR); the CUSUM statistic is given a head-start by setting $C_0 > 0$ rather than $C_0 = 0$, so as to
5 identify the possible initial deviation from μ_0 more effectively. For the same reason, if the setup
6 of the process after a correct signal may be imperfect, after each attempt to restore the process
7 to the in-control condition, C_t is reset to the nonzero value C_0 . However, if the alarm at sample
8 t is proved to be false, C_t is set equal to zero.

9
10
11
12
13
14
15
16
17
18 The run length until the chart triggers a signal is the key measure of the performance of a
19 CUSUM procedure. The mean value of the run length, ARL (Average Run Length), is often
20 used in practice to select the most appropriate CUSUM scheme. There are several papers that
21 deal with the effectiveness of the CUSUM chart. Vance (1986) provides a computer program
22 for evaluating ARL , and Hawkins (1992) gives a relatively simple yet very accurate
23 approximating equation for the evaluation of the ARL . Several Markov chain approaches have
24 been used for the computation of the ARL , like the ones by Ewan and Kemp (1960), Brook and
25 Evans (1972) and Fu et al. (2002). Statistical comparisons between CUSUM and Shewhart
26 charts have been presented by Hawkins and Olwell (1998), Reynolds and Stoumbos (2004) and
27 others.

28
29
30
31
32
33
34
35
36
37
38
39
40
41
42 It has been commonly argued that when the magnitude of the process disturbance due to
43 the occurrence of assignable causes is small, CUSUM charts are more effective than Shewhart
44 charts in monitoring the process (both from an economic and from a statistical point of view),
45 because they can detect small disturbances more rapidly on average. However, in a very recent
46 paper, Nenes and Tagaras (2008) show that the differences between the two charts are
47 negligible in infinite runs, as far as their economic performance is concerned: namely, they
48 produce very similar economic results, even when the shift is very small.

49
50
51
52
53
54
55
56
57
58
59 The analysis of CUSUM charts is typically based on the assumption that the process which
60 is monitored will be operating continuously and indefinitely. In practice though, most

1
2
3 production processes are periodically set up to produce a specific quantity over a specified time
4
5 period, e.g., an 8-hour shift. In such cases, the limited duration of the production run has to be
6
7 taken into account in the design of the control chart in order for this chart to be maximally
8
9 effective.
10

11
12 There is an increasing number of publications that acknowledge the need to design process
13
14 control schemes specifically for short runs; see, for example, Del Castillo and Montgomery
15
16 (1996), Kuo (2006) and Makis (2008). However, the research community has not directed its
17
18 efforts into designing CUSUM schemes specifically for short runs yet, not only because of the
19
20 additional complexity but also because it is not at all obvious that CUSUM charts will be
21
22 effective in short runs, as the limited number of samples does not fit well with the
23
24 accumulative nature of the scheme. The work of Nenes and Tagaras (2005, 2006) is the only
25
26 one, at least to our knowledge, that deals with CUSUM schemes designed for short runs.
27
28 Specifically, Nenes and Tagaras (2005) propose *statistical* measures of performance for control
29
30 charts that are appropriate for short runs while Nenes and Tagaras (2006) develop a model for
31
32 the *economic* optimization of CUSUM schemes in finite-horizon processes. The current paper
33
34 starts with these models as a basis but extends these earlier works in two directions with
35
36 respective objectives as follows:
37
38
39
40
41
42

- 43
44 ➤ it revisits the problem of short runs by deriving properties of the *statistical* measures of
45
46 performance for CUSUM charts and by developing a model for the *statistical* evaluation
47
48 of a CUSUM scheme in short runs;
49
50
51
52 ➤ it compares the behavior of CUSUM and Shewhart charts designed for short runs, using
53
54 both economic and statistical criteria.
55
56

57 The next section describes the problem setting and assumptions. The stochastic model that
58
59 expresses the operation of the CUSUM scheme is presented in section 3. The measures of
60

1
2
3 performance are developed and presented in section 4. Section 5 presents the comparison
4
5 between CUSUM and Shewhart charts in short runs through numerical examples. The final
6
7 section summarizes the main points and findings of the paper.
8
9

10 11 **2. Problem setting and assumptions**

12
13 A production process is set up for processing a specific batch of items over a limited time
14
15 interval (short run). The key measure of the process quality is a continuous random variable X ,
16
17 which is normally distributed with target value μ_0 and constant variance σ^2 . The setup operation
18
19 may not always be perfect in the sense that although in the beginning of the process the mean
20
21 of X is supposed to be set equal to μ_0 , there is a probability that the process starts its operation
22
23 with a mean different from μ_0 .
24
25
26
27

28
29 The process may be affected by the unobservable occurrence of an assignable cause at
30
31 some random time. The effect of the assignable cause is a shift in the mean of X from μ_0 to $\mu_1 =$
32
33 $\mu_0 + \delta\sigma$. The process remains in that undesirable out-of-control state, until the occurrence of the
34
35 assignable cause is detected and its effect removed or until the end of the run if the problem is
36
37 not detected.
38
39

40
41 The process is monitored by means of an one-sided CUSUM chart for detecting a possible
42
43 upward shift in the mean using the CUSUM statistic (1). The total number of samples that will
44
45 be taken till the end of the production run is N . If the control chart indicates a possible out-of-
46
47 control condition, that is, if $C_t > H$, the process is stopped for investigation; if the investigation
48
49 reveals that the assignable cause has indeed occurred, then there is an intervention to restore
50
51 the process to its in-control condition and operation resumes. This intervention may be
52
53 imperfect and despite the detection of the cause and the attempt to remove it, the process may
54
55 continue operating out of control.
56
57
58
59
60

There are five parameters that affect the operation and effectiveness of the CUSUM chart during the run:

- the total number of samples N , (equivalently the sampling interval $h = T / (N + 1)$).
- the sample size n ,
- the reference value k ,
- the control limit H and
- the initial value of the CUSUM statistic C_0 at the beginning of the process ($t = 0$) and after each true alarm.

The values of N , n , k , H , C_0 obviously affect both the statistical and economic performance of the monitoring scheme. Therefore, it is necessary to develop a model for evaluating the appropriate measures of performance for any combination of these five parameters, so as to be able to select, at a second stage, the set of parameter values that satisfy any given requirements.

3. Stochastic model of the process and the control chart

In this section a discrete-time stochastic model for the process and its monitoring scheme is developed, based on the value of the CUSUM statistic C_t for $t = 0, 1, \dots, N$. Although C_t is theoretically a continuous random variable, for practical computational reasons it is discretized into $m+1$ values following the approach of Brook and Evans (1972). Specifically, the interval from 0 to H is partitioned into m sub-intervals. Let w be the width of sub-intervals 1 to $m-1$:

$$w = \frac{2H}{2m-1} \Leftrightarrow m = \frac{H}{w} + \frac{1}{2}. \quad (2)$$

Then, the real-valued C_t is transformed to an integer between 0 and m in the following manner:

$$\text{for } C_t < \frac{w}{2} \rightarrow C_t = 0$$

$$\text{for } \left(i - \frac{1}{2}\right)w \leq C_t < \left(i + \frac{1}{2}\right)w \rightarrow C_t = i \quad i = 1, 2, \dots, m-1$$

$$\text{for } \left(m - \frac{1}{2}\right)w = H \leq C_t \rightarrow C_t = m. \quad (3)$$

If the process is actually in statistical control at sampling instance t ($\mu = \mu_0$), then the standardized variable $z_t = \frac{x_t - \mu_0}{\sigma / \sqrt{n}}$ follows the standard normal distribution $z_t \sim N(0,1)$ while if the process operates in the out-of-control condition ($\mu = \mu_1 = \mu_0 + \delta\sigma$) then z_t follows a normal distribution with mean $\delta\sqrt{n}$ and variance equal to 1: $z_t \sim N(\delta\sqrt{n}, 1)$.

Similar to Nenes and Tagaras (2005), the probabilities p_{ij} of moving from $C_{t-1} = i$ to $C_t = j$ may be computed from:

$$p_{ij} = \begin{cases} \int_{-\infty}^{\left(\frac{1-i}{2}\right)w+k-\delta\sqrt{n}} \varphi(z) dz & i = 0, 1, \dots, m-1 & j = 0 \\ \int_{\left(j-i+\frac{1}{2}\right)w+k-\delta\sqrt{n}}^{\left(j-i-\frac{1}{2}\right)w+k-\delta\sqrt{n}} \varphi(z) dz & i = 0, 1, \dots, m-1 & j = 1, 2, \dots, m-1 \\ \int_{H-iw+k-\delta\sqrt{n}}^{\infty} \varphi(z) dz & i = 0, 1, \dots, m-1 & j = m \\ 0 & i = m & j = 0, 1, \dots, m-1 \\ 1 & i = m & j = m \end{cases} \quad (4)$$

where $\varphi(z)$ is the density function of the standard normal distribution, $\delta = 0$ if the process is under statistical control and $\delta = (\mu_1 - \mu_0)/\sigma$ if the process operates under the effect of the assignable cause. In the latter case the transition probabilities p_{ij} are denoted by \tilde{p}_{ij} . To

simplify notation, when the indices i, j of the transition probabilities are equal to 0 or m , they will be denoted 0 or m , e.g., $p_{0m} \equiv p_{ij}$ for $i = 0$ and $j = m$.

The CUSUM statistic C_t evolves as a Markov chain with $m+1$ states and transition probabilities p_{ij} if the process is in statistical control. The transition probability matrix \mathbf{P} may be written as:

$$\mathbf{P} = \left[\begin{array}{cccc|c} p_{00} & p_{01} & \cdots & p_{0m-1} & p_{0m} \\ p_{10} & p_{11} & \cdots & p_{1m-1} & p_{1m} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{m-10} & p_{m-11} & \cdots & p_{m-1m-1} & p_{m-1m} \\ \hline 0 & 0 & \cdots & 0 & 1 \end{array} \right] = \left[\begin{array}{c|c} \mathbf{A}_{(m \times m)} & \mathbf{B}_{(m \times 1)} \\ \hline \mathbf{0}_{(1 \times m)} & \mathbf{1}_{(1 \times 1)} \end{array} \right]. \quad (5)$$

If the process operates in the out-of-control state, then we obtain in a similar way the matrix $\tilde{\mathbf{P}}$, which is the exact analogue of \mathbf{P} with transition probabilities \tilde{p}_{ij} .

As Fu et al. (2002) have shown multiplication of \mathbf{P} by itself yields the following form of \mathbf{P}^t (the form for $\tilde{\mathbf{P}}^t$ is analogous):

$$\mathbf{P}^t = \left[\begin{array}{c|c} \mathbf{A}^t_{(m \times m)} & \mathbf{W}_t \mathbf{B}_{(m \times 1)} \\ \hline \mathbf{0}_{(1 \times m)} & \mathbf{1}_{(1 \times 1)} \end{array} \right] \quad t = 1, 2, \dots, N \quad (6)$$

where

$$\mathbf{W}_t = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \dots + \mathbf{A}^{t-1} = \sum_{x=0}^{t-1} \mathbf{A}^x. \quad (7)$$

4. Statistical measures of performance

4.1 Truncated ARL, Truncated ATS and Average Number of False Alarms

The most common statistical measure of a control chart's performance is the Average Run Length (ARL), i.e., the expected number of samples until the chart triggers a signal, given that the process remains in the same condition, either in-control (ARL_0), or out-of-control (ARL_δ). In

the case of a short run, where the total number of samples is limited and fixed (N), the run may end without the chart having issued any out-of-control signal. Therefore, let the *Truncated ARL*, denoted by $TARL_0$ or $TARL_\delta$, be the mean number of samples until a signal or until the completion of the process, whichever occurs first. If the run ends without any signal in the N samples, then the number of samples is assigned the value $N+1$. By using $TARL_0$ and $TARL_\delta$, the *Truncated ATS* (Average Time to Signal), denoted by $TATS_0$ and $TATS_\delta$, can be equivalently defined as the mean time until a signal or until the completion of the run, whichever occurs first. If the run ends without any signal in the N samples, then the time to signal is assigned the value T , which is the full duration of the run.

If the process is monitored with a Shewhart \bar{X} -chart, these performance measures are computed as follows:

$$TARL_0 = \sum_{t=1}^N t(1-\alpha)^{t-1}\alpha + (N+1)(1-\alpha)^N = \frac{1-(1-\alpha)^{N+1}}{\alpha} \quad (8)$$

$$TARL_\delta = \sum_{t=1}^N t\beta^{t-1}(1-\beta) + (N+1)\beta^N = \frac{1-\beta^{N+1}}{1-\beta} \quad (9)$$

$$TATS_0 = TARL_0 \cdot h \quad (10)$$

$$TATS_\delta = TARL_\delta \cdot h \quad (11)$$

where $\alpha = \Phi(-k_s)$ is the probability of a type I error, $\beta = \Phi(k_s - \delta\sqrt{n})$ the probability of a type II error at each sample and k_s is the control limit coefficient of the Shewhart chart. $\Phi(\cdot)$ is the cumulative density function of the standard normal distribution and as already mentioned, h is the sampling interval which, in case of N total samples till the end of the run, takes the value $h = T / (N + 1)$; for $N = 0$, $h = T$.

If the short run is monitored with the CUSUM chart described in the previous sections, then $TARL_0$ is the expected value of the following random variable U :

$$U = \begin{cases} \min \{t : C_t \geq H, t = 1, 2, \dots, N\} & \text{if } C_t \geq H \text{ for some } t \in \{1, 2, \dots, N\} \\ N + 1 & \text{if } C_t < H \text{ for all } t \in \{1, 2, \dots, N\} \end{cases} \quad (12)$$

The following equation is obtained by using an approach similar to Fu et al. (2002), for $C_0 = 0$:

$$TARL_0 = [1 \ 0 \ \dots \ 0] \cdot W_{N+1} \cdot \mathbf{1} = 1 + \sum_{t=1}^N \sum_{j=0}^{m-1} p_{0j}^{(t)} \quad (13)$$

where $\mathbf{1} = [1 \ 1 \ \dots \ 1]^T$ is the $m \times 1$ column, every element of which is assigned the value 1, $[1 \ 0 \ \dots \ 0]$ is the $1 \times m$ row, every element of which equals 0, except for the first one which equals 1 and $p_{ij}^{(t)}$ is the probability of the chart's statistic to move from i to j in exactly t steps (sampling instances).

In a similar way, $TARL_0$ for a CUSUM chart with Fast Initial Response (FIR), i.e., starting with $C_0 > 0$, can be easily evaluated. Note that the value of $TARL_0$ in (13) is essentially the first element of the $m \times 1$ vector $W_{N+1} \cdot \mathbf{1}$; in the case of a FIR-CUSUM with $C_0 > 0$, $TARL_0$ is the element located in the C_0 -th row of this column vector. When operating under the effect of the assignable cause, $TARL_\delta$ is computed in the same way but using the elements of matrix \tilde{P} .

The measures $TATS_0$ and $TATS_\delta$ of CUSUM charts are computed from (10) and (11) using the Truncated ARL 's of the CUSUM chart.

In addition to the Truncated ARL and ATS a potentially useful characteristic of a control chart in short runs, which has already been used in practice by Nikolaidis et al. (2007), is the average number of false alarms during the production run, F . For the Shewhart chart, this measure is simply

$$F = N\alpha \quad (14)$$

In order to compute F for the CUSUM chart, the formulation of the transition probabilities in (4) must be slightly modified. In particular, since each false alarm (state m) is followed by the restoration of the chart statistic's value to zero (state 0), the values of p_{mj} must be set equal to the respective values of p_{0j} , $j = 0, 1, \dots, m$ so that state m stops being absorbing. Equivalently, we may say that the bottom row of matrix \mathbf{P} in (5) must be modified and made identical to the top row. Then, for $C_0 = i$, F is computed by:

$$F = \sum_{t=1}^N p_{im}^{(t)} \quad (15)$$

using the modified \mathbf{P} . Note that for CUSUM charts the expected percentage of false alarms in N samples, F/N , may be viewed as the analogue of the type I error (false alarm probability) of the Shewhart chart.

4.2 Properties of Truncated ARL and Truncated ATS

The following propositions provide useful insights about the effect of the CUSUM design parameters on $TARL_0$ and $TARL_\delta$ (Proposition 1) and on $TATS_0$ and $TATS_\delta$ (Proposition 2). Their proofs are given in the Appendix for discretized values of C_0 , k and H and integer n and N . To facilitate the exposition of the proofs, three lemmas concerning the relationships between the elements of the transition probability matrix \mathbf{P} of (5) and \mathbf{P}' are also given and proved in the Appendix.

Proposition 1: $TARL_0$ and $TARL_\delta$ are decreasing in C_0 and increasing in k , H and N . Also, $TARL_0$ is independent of n while $TARL_\delta$ is decreasing in n .

The relationships between $TARL_0$, $TARL_\delta$ and the CUSUM chart parameters are clearly intuitive except probably for the effect of N , which is worth commenting separately. When the total number of samples (N) increases, then the random variable U in (12) can take larger

values, especially when a signal from the chart is not very probable, while if N is small U is bounded from above by a smaller maximum value.

Proposition 2: $TATS_0$ and $TATS_\delta$ are decreasing in C_0 and N and increasing in k and H . $TATS_\delta$ is decreasing in n while $TATS_0$ is independent of n .

4.3 An Application

To illustrate the model and the measures of performance developed in the previous sections, consider the application presented in Nikolaidis et al. (2007), which refers to the monitoring of the tile formation process at a major tile manufacturer in Greece. A critical stage of tile formation is the press operation, which is controlled indirectly through the homogeneity of tile penetrability after stamping. The measurements are done with an instrument that records the penetrability of a pin in which a certain force is exercised.

The penetrability is measured at three points on each of the four sides of a square tile and then the sums of the three measurements of each side are calculated. The monitored quality characteristic X is the difference between the largest and the smallest of these four sums. According to the analysis of Nikolaidis et al. (2007), X is normally distributed with $\mu_0 = 0.0308$ mm and $\sigma = 0.0092$ mm. An assignable cause may shift the process mean to $\mu_1 = \mu_0 + \delta\sigma = 0.0446$ mm ($\delta = 1.5$). The occurrence rate of the cause is $\lambda = 0.0007$ per hour, the sampling cost per item is $c = \text{€}0.56$, the cost of operating in the out-of-control state is $M = \text{€}52.80$ per hour, the false alarm cost is $L_0 = \text{€}2.87$ and the restoration cost is $L_1 = \text{€}16.84$. The optimum expected cost of a Shewhart chart for net operating time $T = 72$ hours (nine shifts) is $\text{€}25.86$ for a sampling interval $h = 6.55$ ($N = 10$), sample size $n = 1$ and control limit $k_s = 0.82$. The respective optimum expected cost of a CUSUM chart is $\text{€}25.80$ for a sampling interval $h = 6.00$ ($N = 11$) sample size $n = 1$, reference value $k = 0.28$, initial value $C_0 = 0$, and control limit $H = 0.8$. The statistical measures of performance for the optimum Shewhart chart are computed

from equations (8) to (11) and (14): $TARL_0 = 4.24$, $TARL_\delta = 1.33$, $TATS_0 = 27.78$, $TATS_\delta = 8.71$ and $F = 2.06$. The respective values for the optimum CUSUM chart are obtained from equations (10), (11), (13) and (15): $TARL_0 = 5.16$, $TARL_\delta = 1.42$, $TATS_0 = 30.94$, $TATS_\delta = 8.51$ and $F = 1.91$. Note that since the sampling interval is different, $TATS_0$ and $TATS_\delta$ are more meaningful here than $TARL_0$ and $TARL_\delta$. Hence, although the economic performance of the optimal Shewhart and CUSUM charts is very similar in this case, the statistical performance of the CUSUM chart is clearly superior.

5. Numerical investigation – Comparisons with Shewhart charts

The purpose of this numerical investigation is to evaluate systematically the economic and statistical performance of the CUSUM scheme in short runs by comparing them against those of the Shewhart scheme. The comparisons are presented in two subsections. First, in 5.1 we present and discuss the economic comparison of the two charts and we also compare the statistical measures of performance of the two charts using the optimal parameters of the economically optimized schemes. Then, in 5.2, we compare the two charts from a purely statistical point of view. Specifically, we compare the behavior of $TARL_\delta$ and $TATS_\delta$ for CUSUM and Shewhart charts that have exactly the same $TARL_0$ and $TATS_0$.

5.1 Comparisons based on the economically optimized schemes

The comparisons here are made through 96 numerical examples. Specifically, the duration of the run T , the magnitude of the shift δ , the occurrence rate of the assignable cause λ , the fixed sampling cost b and the cost per time unit of operating in the out-of-control state M are examined at two levels ($2^5 = 32$ combinations) and each of those 32 combinations is examined for three different sets of cost of false alarm L_0 and restoration cost L_1 ; low values for both L_0 and L_1 , low L_0 and high L_1 , and high values for L_0 and L_1 costs. The sampling cost per item is kept constant: $c = 1$. These parameter choices serve two purposes; on the one hand they cover a

1
2
3 broad range of values, while on the other hand they keep the size of the experiment at a
4 manageable level. The specific values of the parameters that generate the 96 cases are
5 presented in Table 1.
6
7
8
9

10 [Insert Table 1 about here]
11

12
13 Each of the 96 examples is optimized for both Shewhart and CUSUM charts using a cost-
14 minimization criterion. Specifically, optimization of the CUSUM chart is based on the model
15 of Nenes and Tagaras (2006), while optimization of the Shewhart chart is performed using the
16 cost model of Tagaras (1996). Note that no statistical optimization of any kind, for any of the
17 two charts is being attempted at this point. The two charts are first optimized and compared in
18 economic terms. Then, the various statistical measures of performance are calculated by
19 applying the analysis of the present paper on the economically optimized schemes. After that,
20 these statistical measures are compared against each other for the two charts. The exact same
21 comparisons are performed for the case where the sample size is restricted to be equal to one.
22 The special case $n = 1$ is worth examining separately because it is quite common in many
23 industrial applications, especially in the process industries. Table 2 summarizes the cost
24 improvement of the CUSUM over the Shewhart chart and the effect of each parameter on the
25 improvement for both unrestricted n and $n = 1$. Note that for the pair L_0, L_1 there are three
26 average percentages since this combination is examined at three levels.
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

46 [Insert Table 2 about here]
47
48
49

50 The most important conclusion of this numerical investigation is that the economic
51 performance of the two charts is almost identical when there are no sample size restrictions
52 (the CUSUM chart reduces the cost only by 0.14% on average and by 0.72% at most). On the
53 other hand, when the sample size is by necessity equal to one, the average cost reduction in the
54 96 cases that have been examined is 6.73% while it has been recorded as high as 45.84% in one
55 of those cases. These findings coincide with the results of Nenes and Tagaras (2008), who also
56
57
58
59
60

1
2
3 observe negligible differences between the two charts as far as their economic performance is
4
5 concerned for unrestricted sample size, in the infinite-run version of the problem.
6
7

8 The examination of the effect of the parameters shows that the cost reduction is higher for
9
10 high values of T , M and L_0 and L_1 and low values of b , λ and δ . These findings are even more
11
12 pronounced for $n = 1$ since, as pointed out, it is exactly in these cases that large cost
13
14 improvements are observed.
15
16

17
18 Tables 3 and 4 present the average improvements of $TATS_0$ and $TATS_\delta$ respectively
19
20 (average increase of $TATS_0$ and average reduction of $TATS_\delta$) when using a CUSUM instead of
21
22 a Shewhart chart to monitor the process, again for both cases: unrestricted n and $n = 1$. The
23
24 effects of each parameter on the improvement are presented in these two tables exactly as in
25
26 Table 2. Note that the comparisons refer to the values of $TATS_0$ and $TATS_\delta$ (rather than $TARL_0$
27
28 and $TARL_\delta$) since the sampling interval may be different in the optimum design of the two
29
30 charts and consequently the actual time till the detection of the cause (or the false alarm) is
31
32 more interesting and also allows a fair comparison.
33
34
35

36
37 [Insert Tables 3 and 4 about here]
38
39

40 The differences between the two charts are again negligible when n is unrestricted. In
41
42 particular, the CUSUM chart seems to be slightly faster in detecting the assignable cause
43
44 (0.97% improvement of $TATS_\delta$) but it issues type I errors more frequently (0.15% deterioration
45
46 of $TATS_0$). The differences, though, are not large enough to draw solid conclusions.
47
48
49

50 When the sample size is restricted to $n = 1$, the differences in the statistical performance of
51
52 the two charts are very evident, similarly to the economic comparisons. In particular, both the
53
54 average time to the detection of the assignable cause ($TATS_\delta$) and the average time to false
55
56 alarms ($TATS_0$) improve when using the CUSUM instead of the Shewhart chart; the average
57
58 $TATS_0$ increase is 21.81% while the average $TATS_\delta$ decrease is 11.94%.
59
60

1
2
3 As for the effect of the parameters when $n = 1$, $TATS_0$ improves more at high values of T ,
4 M and λ , low values of b and δ while the values of L_0 and L_1 seem to have little effect on $TATS_0$
5 improvement. On the other hand, $TATS_\delta$ improves for high values of T and L_0 and L_1 , low
6 values of b , δ , λ while M seems to have little effect on $TATS_\delta$ reduction.
7
8
9

10
11
12 Table 5 presents the average reduction in F when moving from the Shewhart to the
13 CUSUM chart, exactly as in Tables 2, 3 and 4. The differences between the two charts are
14 again very small for unrestricted n ; the average reduction in F is only 0.76%. However, when
15 the sample size is necessarily $n = 1$, CUSUM charts issue 30.02% fewer type I errors during
16 the production run. In particular CUSUM charts tend to issue fewer type I errors for high
17 values of T , M , L_0 and L_1 and for low values of b .
18
19
20
21
22
23
24
25
26

27 [Insert Table 5 about here]
28
29

30 Overall, substantial differences between the two charts are observed only when the
31 samples are unitary by necessity. In such cases, CUSUM charts are far better than their
32 Shewhart counterparts, both in economic and in statistical terms. Specifically, the differences
33 are more pronounced for high values of T , M , L_0 and L_1 . This means that if a process is set to
34 operate for a production run that is not too short, then it is worth monitoring it with a more
35 sophisticated chart of cumulative nature like the CUSUM. Also, if out-of-control operation is
36 very costly and it is also expensive to interrupt the process and remove the assignable cause,
37 then again the CUSUM chart seems to lead to higher percentage cost reductions and to better
38 statistical results. This is worth underscoring because it shows that CUSUM charts are most
39 effective in monitoring expensive production processes; since the percentage cost reductions
40 are higher, the total savings are also very high in those cases. On the other hand, as the fixed
41 cost per sample b increases CUSUM charts tend to have similar behavior to Shewhart charts.
42 This is explained by the fact that CUSUM charts typically exploit their accumulative feature
43 through more frequent sampling but when the fixed sampling cost increases, frequent sampling
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

is discouraged and consequently their effectiveness is reduced. Finally, the superiority of the CUSUM charts is far more evident when the magnitude of the shift is small because for small values of δ Shewhart charts cannot easily detect the assignable causes.

5.2 Comparative statistical effectiveness of assignable cause detection

In the preceding subsection the various statistical measures of performance were calculated using the economically optimized schemes. In this paragraph we present a purely statistical comparison between the CUSUM and Shewhart charts assuming that a production process is set up to operate for an 8-hour shift ($T = 8$) and samples are scheduled to be taken every 10 minutes ($h = 1/6$ hrs), i.e., $N = 47$ samples within 8 hours. The scope of this comparison is to show how the two charts can be compared from a pure statistical point of view, when no economic optimization has been undertaken.

One possibility is to use a Shewhart chart with $k_s = 3$. Using (8) we obtain $TARL_0 = 46.51$ (equivalently $TATS_0 = 7.75$ hrs).

The alternative is to use a CUSUM chart. For $T = 8$ and $N = 47$, there are numerous combinations of k and H that lead to the same $TARL_0$ and $TATS_0$. The combinations that are used in this investigation, which all lead to $TARL_0 = 46.51$ and $TATS_0 = 7.75$ hrs, are shown in Table 6 along with the values of $TARL_\delta$ for sample sizes $n = \{1, 5, 10, 20\}$ and magnitudes of the shift $\delta = \{0.5, 1.0, 1.5, 2.0, 3.0\}$. Table 6 also contains the $TARL_\delta$ values for the Shewhart chart with $k_s = 3$ for the same n and δ . The shadowed cells in the table highlight the lowest $TARL_\delta$ values given δ and n . Figure 1 presents the values of $TARL_\delta$ for the Shewhart chart and the minimum (shadowed) $TARL_\delta$ of the CUSUM charts.

[Insert Table 6 and Figure 1 about here]

It is obvious from Table 6 and Figure 1 that the differences between the two charts are substantial for $n = 1$ but become progressively negligible as the sample size increases. It is

1
2
3 worth noting that the advantage of the CUSUM is largest for moderate values of shift
4
5 magnitudes δ . If δ is large (e.g., $\delta = 3$) then $TARL_\delta$ is not very different because Shewhart
6
7 charts are almost equally effective in detecting large shifts. If δ is too small ($\delta = 0.5$) then the
8
9 difference in $TARL_\delta$ between Shewhart and CUSUM is not very large because in that case even
10
11 the CUSUM scheme has difficulty in detecting the small shift.
12
13

14 15 16 **6. Summary and conclusions**

17
18 A methodology has been presented for the statistical evaluation of a CUSUM scheme
19
20 designed for finite runs. The economic and statistical characteristics of CUSUM schemes have
21
22 been compared numerically against those of Shewhart charts. The general conclusion of this
23
24 paper is that the economic and statistical performance of the CUSUM chart is superior to that
25
26 of the Shewhart scheme in short runs only if the sample size is restricted to be unitary or very
27
28 small. In these cases, using economic criteria the CUSUM chart is far superior to the Shewhart
29
30 one when production processes are longer, out-of-control operation cost and costs of
31
32 investigating the process and removing the assignable cause are high, fixed sampling cost is
33
34 low and the magnitude of the shift is small. From a purely statistical point of view, the
35
36 performance of the CUSUM chart is also similar to that of the Shewhart chart, unless the
37
38 sample size is unitary and the magnitude of the shift is not large; it is only in those cases that
39
40 the CUSUM scheme clearly outperforms the simple Shewhart chart.
41
42
43
44
45
46

47 48 **Appendix**

49
50 All lemmas and propositions are proved for discretized values of C_0 , k and H . The three
51
52 lemmas are proved by induction.
53
54
55
56
57
58
59
60

Lemma 1

$$\sum_{j=0}^{m-1} p_{aj}^{(t)} < \sum_{j=0}^{m-1} p_{bj}^{(t)} \text{ and } p_{am}^{(t)} > p_{bm}^{(t)} \text{ for any } t \text{ and } \mathbf{a} > \mathbf{b}.$$

Proof of Lemma 1

Since $p_{am} > p_{bm}$ for $t = 1$, it follows that: $\sum_{j=0}^{m-1} p_{aj} < \sum_{j=0}^{m-1} p_{bj}$. Assume that for t : $\sum_{j=0}^{m-1} p_{aj}^{(t)} < \sum_{j=0}^{m-1} p_{bj}^{(t)}$.

It is then sufficient to show that the relationship holds for $t + 1$, namely that $\sum_{j=0}^{m-1} p_{aj}^{(t+1)} < \sum_{j=0}^{m-1} p_{bj}^{(t+1)}$.

From the form of p_{ij} in (4) it follows that $p_{b0} = \sum_{x=0}^{a-b} p_{ax}$. Consequently:

$$\sum_{j=0}^{m-1} p_{aj}^{(t+1)} = \sum_{j=0}^{m-1} \sum_{x=0}^m p_{ax} p_{xj}^{(t)} = \sum_{x=0}^m p_{ax} \sum_{j=0}^{m-1} p_{xj}^{(t)} \text{ and}$$

$$\sum_{j=0}^{m-1} p_{bj}^{(t+1)} = \sum_{x=0}^m p_{bx} \sum_{j=0}^{m-1} p_{xj}^{(t)} = p_{b0} \sum_{j=0}^{m-1} p_{0j}^{(t)} + \sum_{x=1}^m p_{bx} \sum_{j=0}^{m-1} p_{xj}^{(t)} = \sum_{x=0}^{a-b} p_{ax} \sum_{j=0}^{m-1} p_{0j}^{(t)} + \sum_{x=1}^m p_{bx} \sum_{j=0}^{m-1} p_{xj}^{(t)}$$

Therefore:

$$\sum_{j=0}^{m-1} p_{bj}^{(t+1)} - \sum_{j=0}^{m-1} p_{aj}^{(t+1)} = \sum_{x=0}^{a-b} p_{ax} \sum_{j=0}^{m-1} p_{0j}^{(t)} + \sum_{x=1}^m p_{bx} \sum_{j=0}^{m-1} p_{xj}^{(t)} - p_{a0} \sum_{j=0}^{m-1} p_{0j}^{(t)} - \sum_{x=1}^m p_{ax} \sum_{j=0}^{m-1} p_{xj}^{(t)} =$$

$$\sum_{x=1}^{a-b} p_{ax} \left(\sum_{j=0}^{m-1} p_{0j}^{(t)} - \sum_{j=0}^{m-1} p_{xj}^{(t)} \right) - \sum_{x=a-b+1}^m p_{ax} \sum_{j=0}^{m-1} p_{xj}^{(t)} + \sum_{x=1}^m p_{bx} \sum_{j=0}^{m-1} p_{xj}^{(t)} >$$

$$\sum_{x=1}^{a-b} p_{ax} \left(\sum_{j=0}^{m-1} p_{0j}^{(t)} - \sum_{j=0}^{m-1} p_{xj}^{(t)} \right) - \sum_{x=1}^{m-a+b} p_{bx} \sum_{j=0}^{m-1} p_{xj}^{(t)} + \sum_{x=1}^m p_{bx} \sum_{j=0}^{m-1} p_{xj}^{(t)}$$

The inequality holds because $\sum_{x=a-b+1}^m p_{ax} \sum_{j=0}^{m-1} p_{xj}^{(t)} < \sum_{x=1}^{m-a+b} p_{bx} \sum_{j=0}^{m-1} p_{xj}^{(t)}$, since from (4) it follows that

$$p_{ax} = p_{bx} \text{ and by assumption } \sum_{j=0}^{m-1} p_{xj}^{(t)} < \sum_{j=0}^{m-1} p_{\dot{x}j}^{(t)} \text{ for } \dot{x} = x - \mathbf{a} + \mathbf{b} < x. \text{ Consequently:}$$

$$\sum_{j=0}^{m-1} p_{bj}^{(t+1)} - \sum_{j=0}^{m-1} p_{aj}^{(t+1)} >$$

$$\sum_{x=1}^{a-b} p_{ax} \left(\sum_{j=0}^{m-1} p_{0j}^{(t)} - \sum_{j=0}^{m-1} p_{xj}^{(t)} \right) - \sum_{x=1}^{m-a+b} p_{bx} \sum_{j=0}^{m-1} p_{xj}^{(t)} + \sum_{x=1}^{m-a+b} p_{bx} \sum_{j=0}^{m-1} p_{xj}^{(t)} + \sum_{x=m+1-a+b}^m p_{bx} \sum_{j=0}^{m-1} p_{xj}^{(t)} =$$

$$\sum_{x=1}^{a-b} p_{ax} \left(\sum_{j=0}^{m-1} p_{0j}^{(t)} - \sum_{j=0}^{m-1} p_{xj}^{(t)} \right) + \sum_{x=m+1-a+b}^m p_{bx} \sum_{j=0}^{m-1} p_{xj}^{(t)}$$

All the terms above are nonnegative because $\sum_{j=0}^{m-1} p_{xj}^{(t)} < \sum_{j=0}^{m-1} p_{0j}^{(t)}$ for $x > 0$ by assumption.

Therefore: $\sum_{j=0}^{m-1} p_{aj}^{(t)} < \sum_{j=0}^{m-1} p_{bj}^{(t)}$ and consequently $p_{am}^{(t)} > p_{bm}^{(t)}$ for any t and $a > b$. ■

Lemma 2

Let p_{im} be the elements of \mathbf{P} for a CUSUM with some parameter k and let \dot{p}_{im} be the elements of $\mathbf{P}(\tilde{k})$ for a CUSUM differing only in the value, \tilde{k} , of that parameter. Then: $\dot{p}_{im}^{(t)} < p_{im}^{(t)}$ for all t, i and $\tilde{k} > k$.

Proof of Lemma 2

From the form of p_{im} in (4) it is clear that $\dot{p}_{im} < p_{im}$ ($t = 1$). Assume that for some t $\dot{p}_{im}^{(t)} < p_{im}^{(t)}$ for all $i < m$. It is then sufficient to show that $\dot{p}_{im}^{(t+1)} < p_{im}^{(t+1)}$, or $\dot{p}_{im}^{(t+1)} - p_{im}^{(t+1)} < 0$.

Let $\tilde{k} = k + w > k$ and let $i = 0$. Then:

$$\dot{p}_{0m}^{(t+1)} - p_{0m}^{(t+1)} = \dot{p}_{00} \cdot \dot{p}_{0m}^{(t)} + \sum_{x=1}^{m-1} \left(\dot{p}_{0x} \cdot \dot{p}_{xm}^{(t)} \right) + \dot{p}_{0m} \cdot \dot{p}_{mm}^{(t)} - p_{00} \cdot p_{0m}^{(t)} - \sum_{x=1}^{m-1} \left(p_{0x} \cdot p_{xm}^{(t)} \right) - p_{0m} \cdot p_{mm}^{(t)}$$

$\dot{p}_{mm}^{(t)} = p_{mm}^{(t)} = 1$, while from (4) it follows that $\dot{p}_{00} = p_{00} + p_{01}$, $p_{0m} = \dot{p}_{1m}$ and $\dot{p}_{0x} = p_{0(x+1)}$ for

$\tilde{k} = k + w$ and $1 \leq x \leq m - 2$. Therefore: $\dot{p}_{0m}^{(t+1)} - p_{0m}^{(t+1)} =$

$$(p_{00} + p_{01}) \cdot \dot{p}_{0m}^{(t)} + \sum_{x=1}^{m-2} (p_{0(x+1)} \cdot \dot{p}_{xm}^{(t)}) + \dot{p}_{0(m-1)} \cdot \dot{p}_{(m-1)m}^{(t)} + \dot{p}_{0m} - p_{00} \cdot p_{0m}^{(t)} - \sum_{x=1}^{m-1} (p_{0x} \cdot p_{xm}^{(t)}) - p_{0m} =$$

$$p_{00} \cdot (\dot{p}_{0m}^{(t)} - p_{0m}^{(t)}) + \sum_{x=1}^{m-1} [p_{0x} \cdot (\dot{p}_{(x-1)m}^{(t)} - p_{xm}^{(t)})] + \dot{p}_{0(m-1)} \cdot \dot{p}_{(m-1)m}^{(t)} + \dot{p}_{0m} - \dot{p}_{1m} <$$

$$p_{00} \cdot (\dot{p}_{0m}^{(t)} - p_{0m}^{(t)}) + \sum_{x=1}^{m-1} [p_{0x} \cdot (\dot{p}_{(x-1)m}^{(t)} - p_{xm}^{(t)})] + \dot{p}_{0(m-1)} + \dot{p}_{0m} - \dot{p}_{1m}.$$

Noting that (4) implies $\dot{p}_{0(m-1)} + \dot{p}_{0m} = \dot{p}_{1m}$, for $\dot{k} = k + w$ we get:

$$\dot{p}_{0m}^{(t+1)} - p_{0m}^{(t+1)} < p_{00} \cdot (\dot{p}_{0m}^{(t)} - p_{0m}^{(t)}) + \sum_{x=1}^{m-1} [p_{0x} \cdot (\dot{p}_{(x-1)m}^{(t)} - p_{xm}^{(t)})].$$

The right-hand-side of the above inequality is negative because $\dot{p}_{0m}^{(t)} < p_{0m}^{(t)}$ by assumption,

$\dot{p}_{(x-1)m}^{(t)} < \dot{p}_{xm}^{(t)}$ by Lemma 1 and $\dot{p}_{xm}^{(t)} < p_{xm}^{(t)}$ by assumption. Therefore, the left-hand-side of the

above inequality is also negative and consequently the lemma has been proved for $i = 0$.

By similar reasoning we prove that $\dot{p}_{im}^{(t)} < p_{im}^{(t)}$ for any $i < m$ and $\dot{k} > k$. Additionally, since:

$$p_{im}^{(t)} = 1 - \sum_{j=0}^{m-1} p_{ij}^{(t)}, \text{ it follows that } \sum_{j=0}^{m-1} \dot{p}_{ij}^{(t)} > \sum_{j=0}^{m-1} p_{ij}^{(t)} \text{ for any } i \text{ and } \dot{k} > k. \quad \blacksquare$$

Lemma 3

Let p_{ij} be the elements of \mathbf{P} for a CUSUM with some parameter H and let \dot{p}_{ij} be the elements

of $\mathbf{P}(\tilde{\mathbf{P}})$ for a CUSUM differing only in the value, \dot{H} , of that parameter. Then: $\dot{p}_{ij}^{(t)} \geq p_{ij}^{(t)}$ for

any $t, i < m, j < m$ and $\dot{H} = (\dot{m} - 1/2)w > H = (m - 1/2)w$.

Proof of Lemma 3

For $t = 1$: $p_{ij} = \dot{p}_{ij}$ for $i < m, j < m$.

Assume that $\dot{p}_{ij}^{(t)} > p_{ij}^{(t)}$. Then, it suffices to show that: $\dot{p}_{ij}^{(t+1)} - p_{ij}^{(t+1)} \geq 0$ for $i < m, j < m$.

$$\begin{aligned} \dot{p}_{ij}^{(t+1)} - p_{ij}^{(t+1)} &= \sum_{x=0}^{\dot{m}} \dot{p}_{ix}^{(t)} \dot{p}_{xj} - \sum_{x=0}^m p_{ix}^{(t)} p_{xj} = \sum_{x=0}^{\dot{m}-1} \dot{p}_{ix}^{(t)} \dot{p}_{xj} + \dot{p}_{im}^{(t)} \dot{p}_{mj} - \sum_{x=0}^{m-1} p_{ix}^{(t)} p_{xj} - p_{im}^{(t)} p_{mj} \\ &= \sum_{x=0}^{m-1} \dot{p}_{ix}^{(t)} \dot{p}_{xj} + \sum_{x=m}^{\dot{m}-1} \dot{p}_{ix}^{(t)} \dot{p}_{xj} - \sum_{x=0}^{m-1} p_{ix}^{(t)} p_{xj} = \sum_{x=0}^{m-1} \left[\left(\dot{p}_{ix}^{(t)} - p_{ix}^{(t)} \right) p_{xj} \right] + \sum_{x=m}^{\dot{m}-1} \dot{p}_{ix}^{(t)} \dot{p}_{xj} \end{aligned}$$

Since $\dot{p}_{ix}^{(t)} \geq p_{ix}^{(t)}$ by assumption for all $x < m$, it follows that $\dot{p}_{ij}^{(t)} \geq p_{ij}^{(t)}$ for $i < m, j < m$. ■

Proof of Proposition 1

➤ Let \dot{TARL} (\dot{TARL}_0 and \dot{TARL}_δ) be the *Truncated ARL* for $C_0 = \mathbf{a}$ and let $TARL$ ($TARL_0$ and $TARL_\delta$) be the *Truncated ARL* for $C_0 = \mathbf{b}$ where $\mathbf{a} > \mathbf{b}$.

Since $TARL = 1 + \sum_{t=1}^N \sum_{j=0}^{m-1} p_{bj}^{(t)}$ and $\dot{TARL} = 1 + \sum_{t=1}^N \sum_{j=0}^{m-1} p_{aj}^{(t)}$, it follows immediately from Lemma 1

that $TARL_0 > \dot{TARL}_0$ and $TARL_\delta > \dot{TARL}_\delta$, i.e., $TARL_0$ and $TARL_\delta$ are decreasing in C_0 .

➤ Let \dot{TARL} (\dot{TARL}_0 and \dot{TARL}_δ) be the *Truncated ARL* for some reference value \dot{k} and let $TARL$ ($TARL_0$ and $TARL_\delta$) be the *Truncated ARL* for k where $k < \dot{k}$. Then, for $C_0 = i$:

$TARL = 1 + \sum_{t=1}^N \sum_{j=0}^{m-1} p_{ij}^{(t)}$ and $\dot{TARL} = 1 + \sum_{t=1}^N \sum_{j=0}^{m-1} \dot{p}_{ij}^{(t)}$. By Lemma 2 it follows that $TARL_0 > \dot{TARL}_0$

and $TARL_\delta > \dot{TARL}_\delta$, i.e., $TARL_0$ and $TARL_\delta$ are increasing in k .

➤ Let \dot{TARL} (\dot{TARL}_0 and \dot{TARL}_δ) be the *Truncated ARL* for some control limit \dot{H} and let $TARL$ ($TARL_0$ and $TARL_\delta$) be the *Truncated ARL* for H where $H = (m-1/2)w <$

$\dot{H} = (\dot{m}-1/2)w$. By Lemma 3 for $C_0 = i$ it follows that:

$$TARL = 1 + \sum_{t=1}^N \sum_{j=0}^{m-1} \dot{p}_{ij}^{(t)} = 1 + \sum_{t=1}^N \left(\sum_{j=0}^{m-1} \dot{p}_{ij}^{(t)} + \sum_{j=m}^{m-1} \dot{p}_{ij}^{(t)} \right) = 1 + \sum_{t=1}^N \sum_{j=0}^{m-1} \dot{p}_{ij}^{(t)} + \sum_{t=1}^N \sum_{j=m}^{m-1} \dot{p}_{ij}^{(t)} >$$

$$1 + \sum_{t=1}^N \sum_{j=0}^{m-1} p_{ij}^{(t)} + \sum_{t=1}^N \sum_{j=m}^{m-1} \dot{p}_{ij}^{(t)} = TARL + \sum_{t=1}^N \sum_{j=m}^{m-1} \dot{p}_{ij}^{(t)} > TARL.$$

Thus, $TARL_0 > \dot{TARL}_0$ and $TARL_\delta > \dot{TARL}_\delta$, i.e., $TARL_0$ and $TARL_\delta$ are increasing in H .

➤ Let \dot{TARL} (\dot{TARL}_0 and \dot{TARL}_δ) be the *Truncated ARL* for some \dot{N} (total number of samples) and let $TARL$ ($TARL_0$ and $TARL_\delta$) be the *Truncated ARL* for N where $N < \dot{N}$. Then,

$$\text{for } C_0 = i: TARL = 1 + \sum_{t=1}^N \sum_{j=0}^{m-1} p_{ij}^{(t)} \quad \text{and} \quad \dot{TARL} = 1 + \sum_{t=1}^{\dot{N}} \sum_{j=0}^{m-1} p_{ij}^{(t)} = 1 + \sum_{t=1}^N \sum_{j=0}^{m-1} p_{ij}^{(t)} + \sum_{t=N+1}^{\dot{N}} \sum_{j=0}^{m-1} p_{ij}^{(t)}$$

$$= TARL + \sum_{t=N+1}^{\dot{N}} \sum_{j=0}^{m-1} p_{ij}^{(t)} > TARL. \text{ Thus, } TARL_0 \text{ and } TARL_\delta \text{ are increasing in } N.$$

➤ From (4) it follows that the transition probabilities p_{ij} are independent of n for in-control operation ($\delta = 0$) and thus, $TARL_0$ is also independent of n . On the other hand, for out-of-control operation, we see from (4) that an increase in n is equivalent to a decrease in k . We have already shown that $TARL_\delta$ is increasing in k , and thus $TARL_\delta$ is decreasing in n . ■

Proof of Proposition 2

The proof regarding the relationship of $TATS_0$ and $TATS_\delta$ with C_0 , k , H and n is omitted since it comes as a direct result of Proposition 1 and equations (10) and (11). Regarding N , its effect on $TATS$ is not immediately obvious because although $TARL$ is increasing in N , $TATS$ is decreasing since the sampling interval h decreases as N increases. The proof is as follows.

Let \dot{TATS} (\dot{TATS}_0 and \dot{TATS}_δ) be the *Truncated ATS* for some \dot{N} and let $TATS$ ($TATS_0$ and $TATS_\delta$) be the *Truncated ATS* for N where $N < \dot{N}$. In the same way we denote two

alternative sampling intervals: $\dot{h} = \frac{T}{\dot{N}+1} < h = \frac{T}{N+1}$. Then, for $C_0 = i$, it suffices to prove that

$T\dot{A}TS < TATS$ or equivalently that $TATS - T\dot{A}TS > 0$:

$$TATS - T\dot{A}TS = h \cdot TARL - \dot{h} \cdot TARL = h \cdot \left(1 + \sum_{t=1}^N (1 - p_{im}^{(t)}) \right) - \dot{h} \cdot \left(1 + \sum_{t=1}^{\dot{N}} (1 - p_{im}^{(t)}) \right) =$$

$$h - \dot{h} + h \cdot \sum_{t=1}^N (1 - p_{im}^{(t)}) - \dot{h} \cdot \sum_{t=1}^{\dot{N}} (1 - p_{im}^{(t)}) = h - \dot{h} + h \cdot \sum_{t=1}^N (1 - p_{im}^{(t)}) - \dot{h} \cdot \left(\sum_{t=1}^N (1 - p_{im}^{(t)}) + \sum_{t=N+1}^{\dot{N}} (1 - p_{im}^{(t)}) \right) =$$

$$T \cdot \left[\frac{1}{N+1} - \frac{1}{\dot{N}+1} + \frac{1}{N+1} \cdot \sum_{t=1}^N (1 - p_{im}^{(t)}) - \frac{1}{\dot{N}+1} \cdot \sum_{t=1}^N (1 - p_{im}^{(t)}) - \frac{1}{\dot{N}+1} \cdot \sum_{t=N+1}^{\dot{N}} (1 - p_{im}^{(t)}) \right] =$$

$$T \cdot \left[\frac{1}{\dot{N}+1} \cdot \sum_{t=1}^{\dot{N}} p_{im}^{(t)} - \frac{1}{N+1} \cdot \sum_{t=1}^N p_{im}^{(t)} \right] = \frac{T}{(\dot{N}+1) \cdot (N+1)} \cdot \left[(N+1) \cdot \sum_{t=1}^{\dot{N}} p_{im}^{(t)} - (\dot{N}+1) \cdot \sum_{t=1}^N p_{im}^{(t)} \right].$$

Consequently, it suffices to show that:

$$(N+1) \cdot \sum_{t=1}^{\dot{N}} p_{im}^{(t)} - (\dot{N}+1) \cdot \sum_{t=1}^N p_{im}^{(t)} > 0.$$

The left-hand-side of the above inequality can be written as:

$$(N+1) \cdot \left(\sum_{t=1}^N p_{im}^{(t)} + \sum_{t=N+1}^{\dot{N}} p_{im}^{(t)} \right) - (N+1 + \dot{N} - N) \cdot \sum_{t=1}^N p_{im}^{(t)} =$$

$$N \cdot \sum_{t=N+1}^{\dot{N}} p_{im}^{(t)} + \sum_{t=N+1}^{\dot{N}} p_{im}^{(t)} - \dot{N} \cdot \sum_{t=1}^N p_{im}^{(t)} + N \cdot \sum_{t=1}^N p_{im}^{(t)} = N \cdot \sum_{t=1}^{\dot{N}} p_{im}^{(t)} + \sum_{t=N+1}^{\dot{N}} p_{im}^{(t)} - \dot{N} \cdot \sum_{t=1}^N p_{im}^{(t)}.$$

We will prove that the right-hand-side of the above is nonnegative for $\dot{N} = N+1$. Then, by

similar reasoning it can be proved for any $\dot{N} > N$. For $\dot{N} = N+1$ that expression becomes:

$$N \cdot \sum_{t=1}^{N+1} p_{im}^{(t)} + \sum_{t=N+1}^{N+1} p_{im}^{(t)} - (N+1) \cdot \sum_{t=1}^N p_{im}^{(t)} = N \cdot \sum_{t=1}^N p_{im}^{(t)} + N \cdot p_{im}^{(N+1)} + p_{im}^{(N+1)} - N \cdot \sum_{t=1}^N p_{im}^{(t)} - \sum_{t=1}^N p_{im}^{(t)} =$$

$$(N+1) \cdot p_{im}^{(N+1)} - \sum_{t=1}^N p_{im}^{(t)}.$$

Since $p_{im}^{(i)} > p_{im}^{(t)} \quad \forall i > t$ (state m is absorbing), $(N+1) \cdot p_{im}^{(N+1)} - \sum_{t=1}^N p_{im}^{(t)}$ is always nonnegative

because $(N+1) \cdot p_{im}^{(N+1)} > N \cdot p_{im}^{(N)} > \sum_{t=1}^N p_{im}^{(t)}$. Thus, $TATS_0$ and $TATS_\delta$ are decreasing in N . ■

References

- Barnard, G.A. (1959). "Control Charts and Stochastic Processes". *Journal of the Royal Statistical Society*, (B), **21**, 239-271.
- Brook, D. and Evans, D.A. (1972). "An approach to the probability distribution of cusum run length". *Biometrika*, **59**, 539-549.
- Del Castillo, E. and Montgomery, D.C. (1996). "A general model for the optimal economic design of \bar{X} charts used to control short or long run processes". *IIE Transactions*, **28**, 193-201.
- Ewan, W.D. and Kemp, K.W. (1960). "Sampling inspection of continuous processes with no autocorrelation between successive results". *Biometrika*, **47**, 363-380.
- Fu, J.C., Spiring, F.A. and Xie Hansheng (2002). "On the average run lengths of quality control schemes using a Markov chain approach". *Statistics & Probability Letters*, **56**, 369-380.
- Goldsmith, P.L. and Whitfield, H. (1961). "Average run lengths in cumulative chart quality control schemes". *Technometrics*, **3**, 11-20.
- Hawkins, D.M. (1992). "A fast accurate approximation of average run lengths of cusum control charts". *Journal of Quality Technology*, **24**, 37-42.
- Hawkins, D.M. and Olwell, D.H. (1998). *Cumulative Sum Charts and Charting for Quality Improvement*, Springer, New York.

- 1
2
3 Johnson, N.L. (1961). "A simple theoretical approach to cumulative sum control charts".
4
5 *Journal of the American Statistical Association*, **54**, 835-840.
6
7
8 Kuo, Y. (2006). "Optimal adaptive control policy for joint machine maintenance and product
9
10 quality control". *European Journal of Operational Research*, **171**, 586-597.
11
12
13 Makis, V. (2008). "Multivariate Bayesian process control for a finite production run".
14
15 *European Journal of Operational Research*, doi:10.1016/j.ejor.2008.01.001.
16
17
18 Nenes, G. and Tagaras, G. (2005). "The CUSUM Chart for Monitoring Short Production
19
20 Runs". *5th International Conference on Analysis of Manufacturing Systems -*
21
22 *Production Management*, Zakynthos, 43-50.
23
24
25 Nenes, G. and Tagaras, G. (2006). "The Economically Designed CUSUM Chart for
26
27 Monitoring Short Production Runs". *International Journal of Production Research*, **44**,
28
29 1569–1587.
30
31
32 Nenes, G. and Tagaras, G. (2008). "An Economic Comparison of CUSUM and Shewhart
33
34 Charts". *IIE Transactions*, **40**, 133-146.
35
36
37 Nikolaidis, Y., Rigas, G. and Tagaras, G. (2007). "Using Economically Designed Shewhart and
38
39 Adaptive \bar{X} - Charts for Monitoring the Quality of Tiles". *Quality and Reliability*
40
41 *Engineering International*, **23**, 233-245.
42
43
44 Page, E.S. (1954). "Continuous inspection schemes". *Biometrika*, **41**, 100–115.
45
46
47 Page, E.S. (1961). "Cumulative sum charts". *Technometrics*, **3**, 1–9.
48
49
50 Reynolds, M.R., Jr. and Stoumbos, Z.G. (2004). "Should Observations Be Grouped for
51
52 Effective Process Monitoring?". *Journal of Quality Technology*, **36**, 343-366.
53
54
55 Tagaras, G. (1996). "Dynamic Control Charts for Finite Production Runs". *European Journal*
56
57 *of Operational Research*, **91**, 38-55.
58
59
60 Vance, L.C. (1986). "Average run lengths of cumulative sum control charts for controlling
normal means". *Journal of Quality Technology*, **18**, 189–193.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Tables and Figures

Table 1: Parameter values used in the numerical investigation

Parameter	Values
T	8 or 40
δ	1 or 2
λ	0.01 or 0.1
b	0 or 5
M	100 or 1000
L_0, L_1	50, 50 or 50, 500 or 500, 500

Table 2: Average cost improvement of CUSUM against Shewhart chart for small (-) and large (+) values of the experiment parameters

	cost improvement					
	unrestricted n		$n = 1$			
	-	+	-	+		
T	0.10%	0.19%	4.91%	8.55%		
b	0.25%	0.04%	10.97%	2.49%		
M	0.05%	0.23%	3.18%	10.28%		
λ	0.15%	0.14%	7.77%	5.69%		
δ	0.16%	0.12%	8.23%	5.23%		
L_0, L_1	0.20%	0.08%	0.16%	5.47%	2.73%	11.99%
overall	0.14%		6.73%			

Table 3: Average $TATS_0$ improvement of CUSUM against Shewhart chart for small (-) and large (+) values of the experiment parameters

$TATS_0$ improvement						
unrestricted n				$n = 1$		
	-	+		-	+	
T	0.13%	-0.43%		7.88%	35.73%	
b	-0.21%	-0.10%		35.05%	8.57%	
M	-0.02%	-0.29%		12.29%	31.33%	
λ	0.02%	-0.32%		16.61%	27.00%	
δ	0.73%	-1.03%		29.80%	13.82%	
L_0, L_1	-0.58%	-0.18%	0.30%	23.76%	22.12%	19.54%
overall	-0.15%			21.81%		

Table 4: Average $TATS_\delta$ improvement of CUSUM against Shewhart chart for small (-) and large (+) values of the experiment parameters

$TATS_\delta$ improvement						
unrestricted n				$n = 1$		
	-	+		-	+	
T	0.23%	1.71%		8.37%	15.50%	
b	1.98%	-0.03%		17.92%	5.95%	
M	0.34%	1.60%		12.29%	11.58%	
λ	0.71%	1.23%		13.49%	10.38%	
δ	0.52%	1.43%		14.47%	9.41%	
L_0, L_1	0.94%	1.92%	0.06%	6.20%	6.46%	23.15%
overall	0.97%			11.94%		

Table 5: Average F improvement of CUSUM against Shewhart chart for small (-) and large (+) values of the experiment parameters

	F improvement					
	unrestricted n			$n = 1$		
	-	+		-	+	
T	2.98%	-1.46%		20.17%	39.88%	
b	-2.25%	3.77%		45.50%	14.55%	
M	0.67%	0.85%		21.33%	38.71%	
λ	1.10%	0.42%		30.36%	29.68%	
δ	2.43%	-0.91%		28.68%	31.36%	
L_0, L_1	0.66%	-1.25%	2.87%	25.43%	21.62%	43.01%
overall	0.76%			30.02%		

Table 6: $TARL_\delta$ comparisons between CUSUM and Shewhart charts

δ	n	CUSUM						Shewhart
		H k	1.0 2.062	1.5 1.63	2.0 1.265	2.5 1.016	3.0 0.839	$k_s=3$
0.5	1		41.24	40.40	37.90	35.34	32.97	41.62
	5		23.19	18.76	13.22	10.35	8.99	25.65
	10		10.08	7.44	5.57	4.90	4.70	12.56
	20		3.63	3.05	2.75	2.72	2.81	4.50
1	1		27.36	23.31	17.16	13.39	11.36	29.39
	5		3.63	3.05	2.75	2.72	2.81	4.50
	10		1.64	1.60	1.62	1.72	1.86	1.77
	20		1.07	1.08	1.11	1.16	1.25	1.08
1.5	1		11.76	8.65	6.31	5.44	5.14	14.43
	5		1.48	1.47	1.50	1.60	1.74	1.57
	10		1.04	1.05	1.06	1.10	1.17	1.04
	20		1.00	1.00	1.00	1.00	1.00	1.00
2	1		4.98	3.97	3.37	3.24	3.27	6.30
	5		1.08	1.08	1.11	1.16	1.25	1.08
	10		1.00	1.00	1.00	1.00	1.01	1.00
	20		1.00	1.00	1.00	1.00	1.00	1.00
3	1		1.81	1.74	1.74	1.84	1.97	2.00
	5		1.00	1.00	1.00	1.00	1.00	1.00
	10		1.00	1.00	1.00	1.00	1.00	1.00
	20		1.00	1.00	1.00	1.00	1.00	1.00

Figure 1: $TARL_\delta$ comparisons between CUSUM and Shewhart charts

