APPLIED STOCHASTIC MODELS IN BUSINESS AND INDUSTRY
Appl. Stochastic Models Bus. Ind., 2004; **20**:185–200 (DOI: 10.1002/asmb.521)

Control charts: a cost-optimization approach for processes with random shifts

András Zempléni[1,*,†], Miklós Véber[1], Belmiro Duarte[2] and Pedro Saraiva[3]

[1] *Department of Probability Theory and Statistics, Eötvös Loránd University, Budapest, Hungary*
[2] *Instituto Superior de Engenharia de Coimbra, Portugal*
[3] *Department of Chemical Engineering, University of Coimbra, Portugal*

SUMMARY

In this paper we describe an approach for establishing control limits and sampling times which derives from economic performance criteria and a model for random shifts. The total cost related to both production and control is calculated, based on cost estimates for false alarms, for not identifying a true out of control situation, and for obtaining a data record through sampling. We describe the complete process for applying the method and compare with conventional procedures to real data from a Portuguese pulp and paper industrial plant. It turns out that substantial cost-reductions may be obtained. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: cost function; CUSUM charts; process control; shift in process mean

1. INTRODUCTION

Control charts are one of the most widely used tools in industrial practice for achieving process control and improvement. Montgomery [1] provides a good review, where economic aspects are mentioned. Does *et al*. [2] describe possible pitfalls in real-life applications of statistical process control—with one of the critical issues—associated with the correct implementation of such a tool being the proper definition of control limits and sampling frequencies. Very often these decisions are not supported by sound statistical or economic decision-making criteria, leading to a suboptimal use and results derived from such applications.

In this paper we describe an approach for establishing control limits and sampling times which derives from economic performance criteria that do consider the total cost related to both production and control, based upon cost estimates for false alarms, for not identifying a true out

*Correspondance to: András Zempléni, Department of Probability Theory and Statistics, Eötvös Loránd University, Budapest H-1117, Hungary.
† E-mail: zempleni@ludeus.elte.hu

Contract/grant sponsor: European Commission; contract/grant number: G6RT-CT-2001-05059

of control situation, and for obtaining a data record through sampling. Similar approaches can be found in the works of Duncan [3], with only the simplest case of fixed shift size and the X-chart treated, and a slightly different cost function used. A recent review of statistical tools in quality control is Arnold and Göb [4], where economic evaluations are also considered. We give optimal control charts for a general problem characterized by random (upward) shift occurrences in the monitored variable. The approach is applied to X-charts with upward shift occurrences, but problems with downward shifts (like the real-life example in Section 4) are easily handled through the same methodology.

Previous historical data are used to estimate the intensity and mean of process mean shifts. A similar approach was used for the case of acceptance sampling in Véber and Zempléni [5] and in Klaassen [6]. For the one-sided problem under consideration, maximum-likelihood methods are applied to estimate the variance components (i.e. the process' inherent variation and the variation induced by the shift). Markov chain models are used in order to find the optimal values for control limits and sampling frequencies. It turns out that CUSUM charts have favourable properties for the simpler case of constant shift sizes, but due to the heavy computational burden related to them, in the general framework that we propose for dealing with random shift size disturbances, a simple X-control chart is proposed and optimized.

As the paper describes the complete process for applying the suggested method, we were able to compare the new results obtained from its application to the conventional procedures to real sets of data gathered from a Portuguese pulp and paper industrial plant and it turned out that a substantial cost-reduction is associated with the optimal procedure found.

The detection of process shifts in retrospective discrete time data series can be treated as a classical change-point problem [7–9]. The basic goal underpinning this paradigm is to find the point(s) in time where a certain statistic did suffer pronounced behaviour changes through enumerative procedures. Some of the most commonly used statistics in parametric detection are the Geometrical Mean Average [10], CUSUM [11], or the Generalized Likelihood Ratio [12], while Mann-Whitney [9] and CUSUM [13] values have been employed for non-parametric change detections. However, our approach is different from the classical change-point approach by the used cost functions and the optimization methods.

The remaining parts of this article are organized as follows: in Section 2 the model and solution for fixed shift sizes is developed for X and CUSUM charts. It turns out that CUSUM charts are slightly favourable, but due to the much more complicated optimization algorithm that is needed, it is beyond the scope of this paper to find the optimal CUSUM chart to the general case of random shift sizes. In Section 3 we extend the results to randomly exponentially distributed shift sizes (an exponential distribution was chosen due to its simplicity, but the method can be easily extended to other non-negative probability distributions). Section 4 comprises the presentation of a motivating industrial example and a maximum likelihood estimation procedure to find the parameters required by the optimization algorithm. Its application is preferred to heuristic calculations, especially for random shift size cases. We also investigate the robustness of the procedure with respect to process constraints. We conclude our paper with an analysis of the strength and implementation difficulties of such an approach to typical industrial process monitoring problems. Technical details are given in the Appendix.

2. SIMPLEST MODEL

2.1. Definitions and notation

Let us consider a normally distributed process with mean μ and standard deviation σ under statistical control, which is sampled at a certain frequency. We assume the presence of a random effect (occurring at random and unknown time points), which leads to upward shifts in the process mean. Our aim is to detect the occurrence of such shifts, through an economically designed control chart.

Next, we introduce the main cost parameters used throughout this paper.

c_s: cost of sampling (per unit sample element), which involves the cost of labour, transportation, investigation, material and data processing.

c_f: cost of a false alarm, including the cost of labour required and possible time delays.

c_w: cost (per unit time interval) of a non-recognized shift including its effects on product waste, reliability reduction and possible contractual consequences.

The above cost elements have to be estimated by previous experience-based knowledge, gathered from process operators and managers at different levels. We believe that these considerations are important not only as part of the process investigated in this paper, but to the everyday planning of statistical process control, since they help one to properly assign the levels of risk to the relevant factors and thus lead to a more convenient and economic operation mode.

The following are the parameters needed for statistical investigations:

d: expected number of shifts in a unit time interval, assumed to be constant over time. Thus the probability of a shift occurring is (approximately) given by the product of d and the interval length for small time intervals. This amounts to assuming the time interval between shifts is exponentially distributed with expectation $1/d$.

s: shift size (considered as fixed in this section, but modelled by an exponential distribution in Section 3).

t: time period between consecutive observations.

c: UCL (Upper Control Limit).

2.2. Markov-model for fixed shifts

To derive a Markov model for fixed shifts we consider the simplest case of alarm detection based on a single sample element and extend it to allow for the option of using the last n observations to signal an alarm if all of them are higher than the threshold previously set. In our current work we investigate situations with $n < 5$.

The process states and actions are presented in Table I, where states 1 and 2 correspond to a process in its normal state (with a false alarm or not), and states 3 and 4 correspond to a process with a shift occurrence (either detected or not).

Table I. States of the process and possible decisions.

States	1	2	3	4
Process	No shift	No shift	Shifted	Shifted
Action	No alarm	Alarm	No alarm	Alarm

The process behaviour at consecutive instants (captured through sampling) is characterized by the above states. The assumptions regarding a shift occurrence imply that the probabilities of future transitions are only dependent on the current state. This is the so-called Markov property, and the model under consideration is called a Markov chain. Mathematical theory (see Karlin and Taylor [14], for example) ensure that there exist a stationary distribution of this chain, which may be interpreted as the long run frequency of being in each of the states.

The states of the chain are described by pairs comprising the process mean (which may represent the in-control or shifted process) and the number m of consecutive observations beyond the UCL ($0 \leqslant m \leqslant n$).

The probabilities of transitions from one state to another (i.e. the transition matrix) can be calculated easily. For example, if the process now is in its normal state and there are no out-of-control observations, the probability of remaining in such a state is $(1 - F(t))\Phi(c)$, where we denote the monitored process variable distribution by $\Phi(.)$ and the probability that during a time interval of length t no shift occurs by $1 - F(t)$, with $F(.)$ standing for the exponential distribution with expectation $1/d$.

For the case $n = 1$ the complete transition matrix A is given in Equation (1). The rows correspond to the states above, with transition probabilities to state j in the jth column. Rows 1, 2 and 4 are the same, since we model the process as immediately turning back to its original controlled mean, with the same transition probabilities as before for an immediate shift to happen after an alarm situation.

$$\begin{array}{c} \text{state 1} \\ \text{state 2} \\ \text{state 3} \\ \text{state 4} \end{array} \begin{bmatrix} (1-F(t))\Phi(c) & (1-F(t))(1-\Phi(c)) & F(t)\Phi(c-s) & F(t)(1-\Phi(c-s)) \\ (1-F(t))\Phi(c) & (1-F(t))(1-\Phi(c)) & F(t)\Phi(c-s) & F(t)(1-\Phi(c-s)) \\ 0 & 0 & \Phi(c-s) & 1-\Phi(c-s) \\ (1-F(t))\Phi(c) & (1-F(t))(1-\Phi(c)) & F(t)\Phi(c-s) & F(t)(1-\Phi(c-s)) \end{bmatrix} \tag{1}$$

However, states 1 and 2 play a completely different role in the cost calculation, since the total process control cost C, based on the stationary distribution is given as

$$C = (c_s + p_2 c_f)\frac{1}{t} + p_3 c_w + p_4 c_w \beta \tag{2}$$

where p_i is the stationary probability of state i and β the fraction of time between consecutive observations where the shift occurred and remained undetected. Similar transition matrices can be calculated for cases where $n > 1$, with the resulting Markov chain having $2(n+1)$ different states.

Now we are in a position to find the control chart with the minimal total cost C defined by an optimal choice of c and t. The first step in our optimization approach is to determine the stationary distribution $\pi = (\pi_1, \pi_2, \ldots, \pi_{2(n+1)})$ of the chain, which can be obtained by solving the system of linear equations $\pi A = \pi$. In the simplest case of $n = 1$ this can easily be done, resulting in the following solution: $\pi_2 = (1 - \Phi(c))\pi_1/\Phi(c)$, $\pi_3 = 1 - \pi_1/[(1 - F(t))\Phi(c)]$, $\pi_4 = F(t)\pi_1/[(1 - F(t))\Phi(c)]$ and π_1 can be determined from the constraint $\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$. The total cost C can be computed by using Equation (2) and therefore the combination (c, t) leading to the smallest cost can be obtained through the use of standard optimization procedures available within the scope of the *R* statistical package. This was carried out separately for $n = 1, 2, 3, 4$ with the results for the case $d = 0.2$ (i.e. expected time to shift equal to 5) summarized in Figure 1.

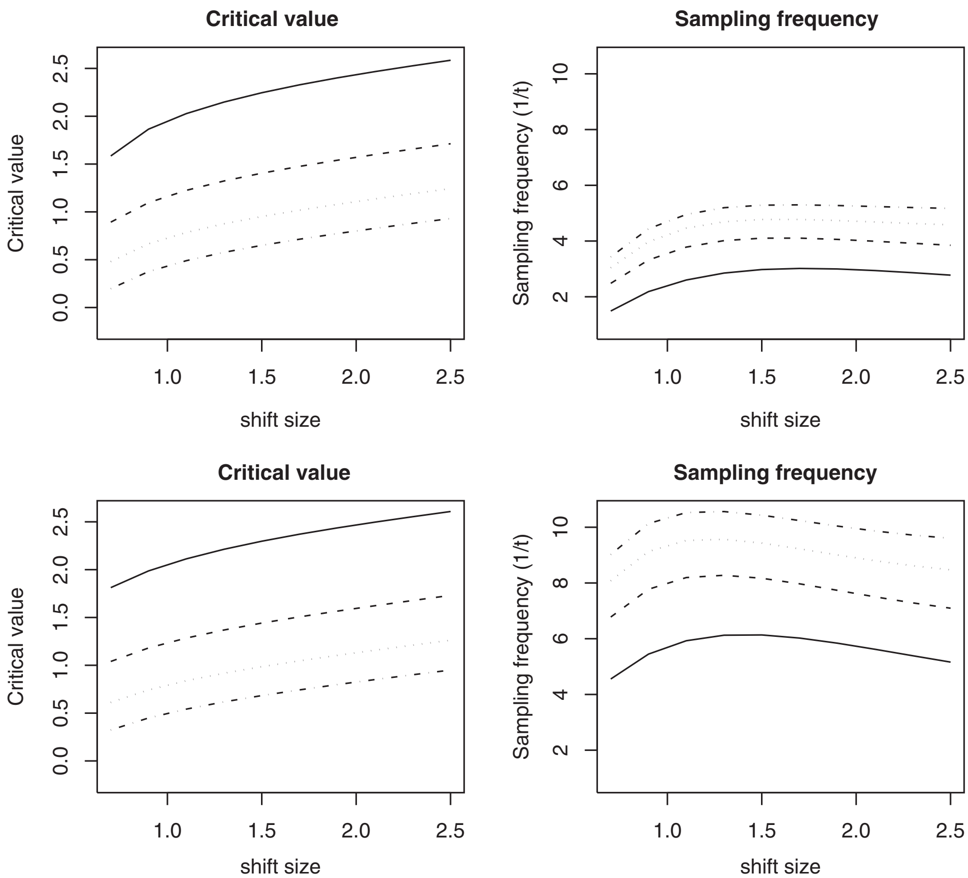

Figure 1. Optimal parameter values as a function of shift size, for $d = 0.2$. Top: $(c_{\mathrm{f}}, c_{\mathrm{w}}) = (150, 50)$; bottom: $(c_{\mathrm{f}}, c_{\mathrm{w}}) = (150, 150)$; solid line: $n = 1$, broken: $n = 2$, dotted: $n = 3$, dots-and-lines: $n = 4$.

The procedure presented is based on the assumption that the process variable under statistical control is normally distributed, which can be transformed to the standard normal distribution we used throughout this and the next section. It can be observed that the parameters fulfil the most obvious properties, with the optimal sampling frequency being higher for higher non-detection costs and critical values increasing with shift size. When the decision rule for alarm signal is based on n ($n > 1$) consecutive observations higher than a threshold, to obtain a similar performance such a value is lower than the value used for $n = 1$.

Figure 2 shows the results for the case $d = 0.05$ (i.e. expected time to shift here equals 20). The critical values are not given, since they behave very much like in the case $d = 0.2$. It is worth noting that the sampling frequencies shown in Figure 2 are much lower than those for $d = 0.2$ case, with the lower risk of shift needing not as many checks. The average cost obtained with the optimal approach shows a significant downward trend for increasing shift sizes, which agrees with the dependence of statistical control cost with respect to shift size. The reason for such a trend has to do with the increasing difficulty in detecting smaller shifts, which leads one to

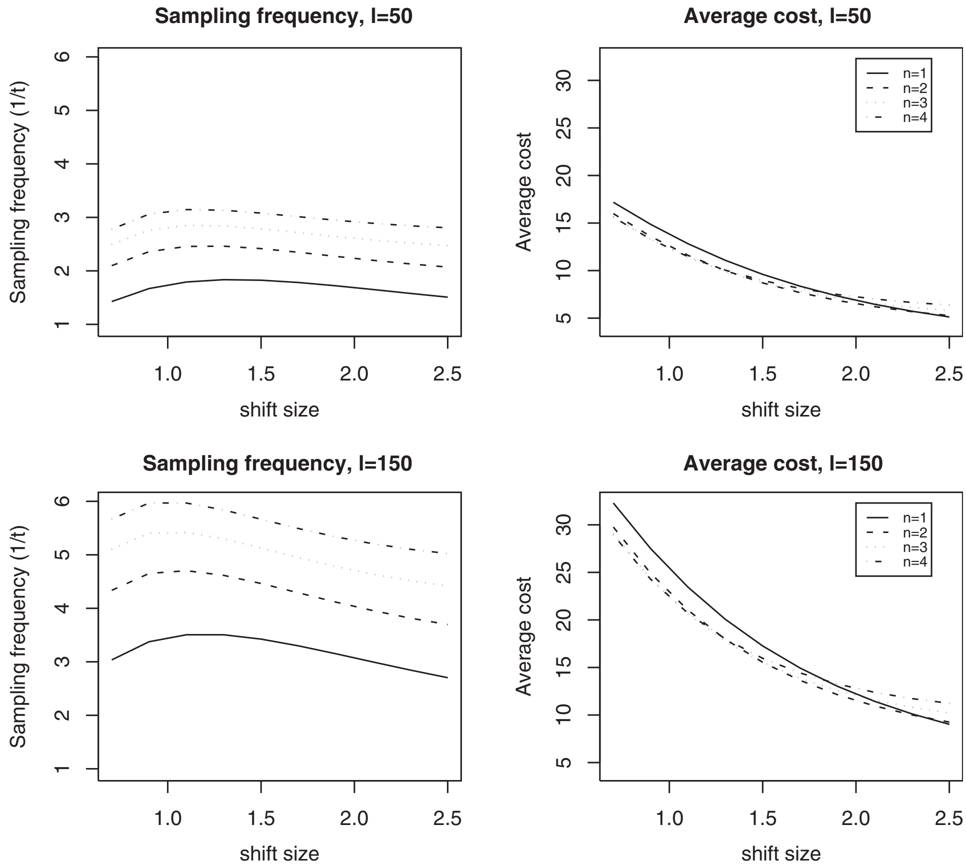

Figure 2. Sampling frequency and average cost as a function of shift size, for $d = 0.05$. Top: $(c_f, c_w) = (150, 50)$; bottom: $(c_f, c_w) = (150, 150)$; solid line: $n = 1$, broken: $n = 2$, dotted: $n = 3$, dots-and-lines: $n = 4$.

expect more incorrect decisions. Furthermore, the cost is weakly dependent upon n. In the next section we focus our attention on the detection based upon single samples ($n = 1$).

2.3. *CUSUM charts*

CUSUM charts are based on cumulative discrepancies from the target value [1, 15]. The cumulative measurement of deviations across time provides the main reason for the robustness of this method in detecting shifts. In our case on the chart the points $(tn, \sum_{i=1}^{n} X_i$ are shown (emphasizing also the dependence on the sampling interval, where n is the index of the current observation and $\sum_{i=1}^{n} X_i$ is the cumulative sum of deviations for samples 1 to n). The control limits are drawn projecting horizontally the last observation at distance a ahead. Next, a line at an angle α to the horizontal axis and with the origin located at the point resulting from previous projection (vertex) is marked. An alarm is signaled when this slanted line intersects one or more earlier observations. The procedure can be demonstrated graphically (see

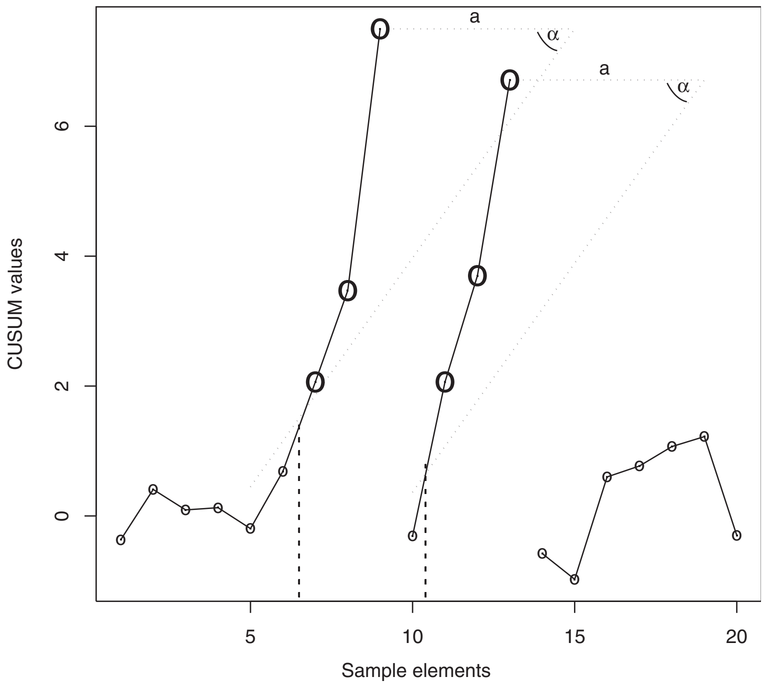

Figure 3. CUSUM chart with half-masks, $t = 1$. Larger circles represent CUSUM values based on out-of-control sample elements. Those slanted lines, where an alarm is signaled, are also shown, together with the parameter values. Broken vertical lines correspond to time points, where a shift occurred.

Figure 3). Each time the alarm is signaled, the CUSUM operator is updated and it restarts from the beginning.

The cost function is similar to Equation (2), with the optimal parameters involved in the procedure being the angle of the mask (α), the distance of the vertex from the last point on the chart (a), and time interval between consecutive observations (t).

This optimization was also based on the Markov-chain methodology, where states represent the distance of the nearest point from the line, which, coupled with the new observation, provides enough information to determine the distance for the next step. Such a formulation leads to a continuous-state chain, whose stationary distribution was computed through a discretization scheme with the distances represented by m classes of values (details can be found in the Appendix).

The transitions between shifted and regular operation states are similar to the case of X-charts presented in Section 2.2. As before, the stationary distribution is found by solving a system of linear equations. We have chosen $m = 100$, which turned out to be accurate enough and allowed for a relatively quick optimization.

The optimal cost-results obtained for CUSUM charts are shown in Figure 4. One may see that as the magnitude of shifts is reduced, greater becomes the distance a, and lower the angle α, due to the requirement of detecting smaller changes. The trends of sampling frequency and average cost are similar to those obtained for the X-chart. The results are encouraging, since

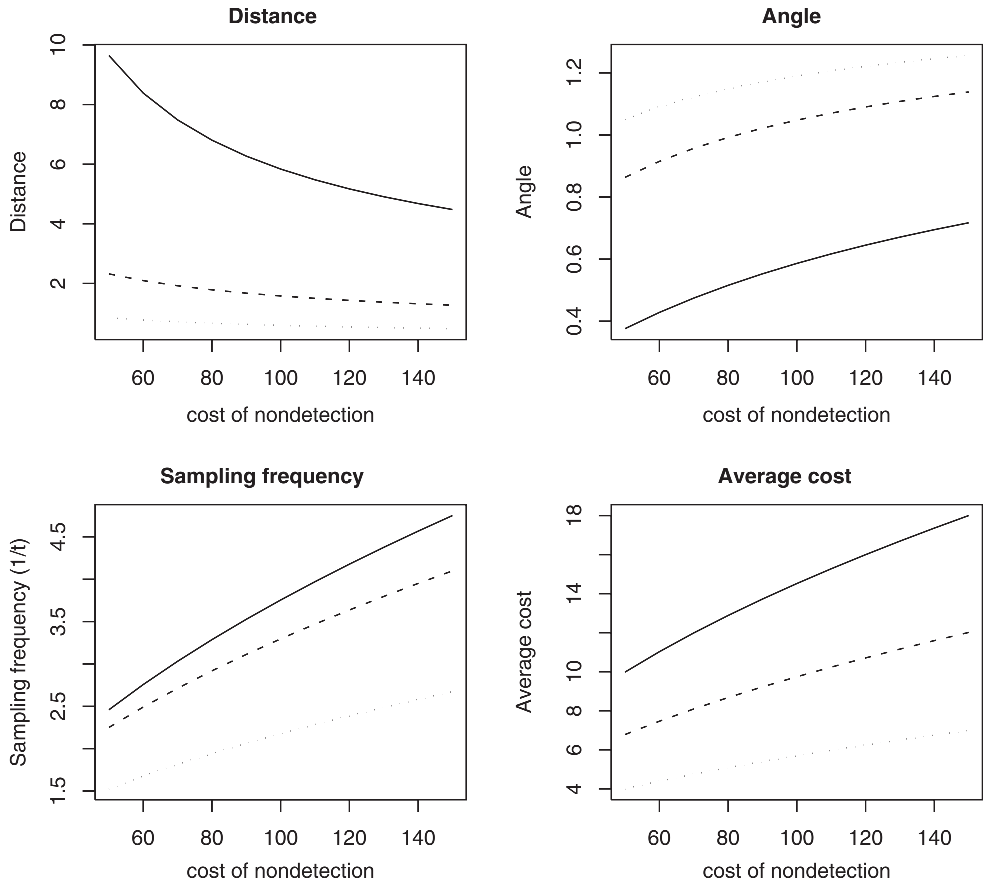

Figure 4. Parameters for the optimal CUSUM chart; $d = 0.05$, $f = 50$. Solid line: $s = 0.35$, broken: $s = 0.65$, dotted: $s = 1.25$.

CUSUM charts allow one to reduce the total cost by 5–25%, when compared to the best results obtained in the previous subsection (Figure 5). The decrease was more prominent for small values of d and higher costs of false alarm.

3. RANDOM SHIFT-SIZE

The main difference of the approach introduced in this section when compared with the methodologies previously described, arises from considering a more realistic situation, where random size shifts may occur. The shifts' distribution is considered to follow an exponential probability distribution with expected value s, but the strategy can easily be extended to other probability distributions.

The formulation presented in this paper is focused on the X-chart type with $n = 1$, i.e. our decision is based on a single observation, which is compared to a threshold (critical value). A

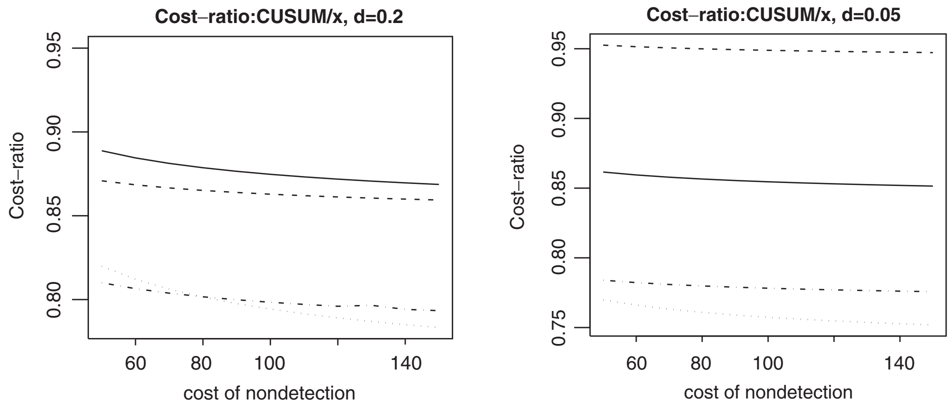

Figure 5. Cost-ratio of the optimal CUSUM versus the optimal X-chart. Solid: $c_f = 50$, $s = 0.35$, broken: $c_f = 50$, $s = 0.65$, dotted: $c_f = 150$, $s = 0.35$, dots-and-lines: $c_f = 150$, $s = 0.65$.

discretization scheme for continuous-state Markov-chains was used to find the optimal parameter vector, where the actual state of the chain is given by the current mean value. The actual states are not observable for the experimenter, thus leading to a hidden structure, and the expected cost is evaluated based on the stationary distribution of the chain.

The formulation of the transition matrix needs a vector of probabilities to represent the shift size distribution q during a sampling interval. Our assumptions ensure that the shift-times form a homogenous Poisson process, while the shift size is exponentially distributed, regardless of any previous events. Besides the capacity of the exponential distribution in describing real-world shift-phenomena, its choice was motivated by the advantage it provides regarding tractability of its convolution powers as a gamma distribution. This leads to time- and location-free evaluation schemes of this distribution, as follows:

$$q(i) = \sum_k P(N = k)P(i\Delta < Y_k < (i+1)\Delta) \tag{3}$$

Equation (3) also holds for the case $i = 0$ with Y_0 set equal to 0, where N is a Poisson distributed random variable, representing the number of events (shifts) during the sampling interval; Y_k is modelled by a $\Gamma(k, 1/s)$ distribution (the sum of k independent exponential variates each with mean s) and Δ stands for the length of an interval, taken as one state of the Markov chain (the derivation of the transition matrix is presented in the Appendix).

The discretization scheme derived for evaluating the continuous Markov chain involves $m = 300$ classes, which turned out to be accurate enough and yet provided a quick solution to the optimization problem.

In this random-shift case we used a Taguchi-type loss function ($c_w x^2$) for the non-detection part of the total cost, where x is just the difference of the current mean from the target value, leading to:

$$C = (c_s + p_f c_f)\tfrac{1}{t} + c_w E(S^2) + p_a c_w [E(S)]^2 \tag{4}$$

where S is the shift distribution of the process; p_f and p_a denote the false and correct alarm probabilities, respectively, and E is the expected value. The last term in (4) is just an approximation for the non-detection cost for those part-intervals where the shift was not yet detected. The use of a smaller coefficient here is required to avoid an overestimation of the amount of time and cost of the process spent out of control. In Section 4 we investigate the accurateness of the Markov-chain approach, including the choice of this loss function.

One may adopt any other loss function for non-detection cost by suitable transformation of the term $E(S^2)$ (cf. Section 4). The objective function involves the same parameters introduced in Section 2. For the chosen value of m the CPU time required to perform the minimization is about 1–2 min per case, so we have been able to investigate a wide range of different parameter values. The initial values of the numeric optimization do not seem to play an important role, since the same optimal values were found regardless of the starting point of the algorithm.

The results obtained with our approach are shown in Figure 6. One may observe the weak dependence of the critical value on c_{w} (cost of non-detection), as well as its strong dependence on s, similarly to what happened for the deterministic shift case. The sensitivity of sampling frequency on c_{w} now depends on the expected shift size, since higher probabilities of large

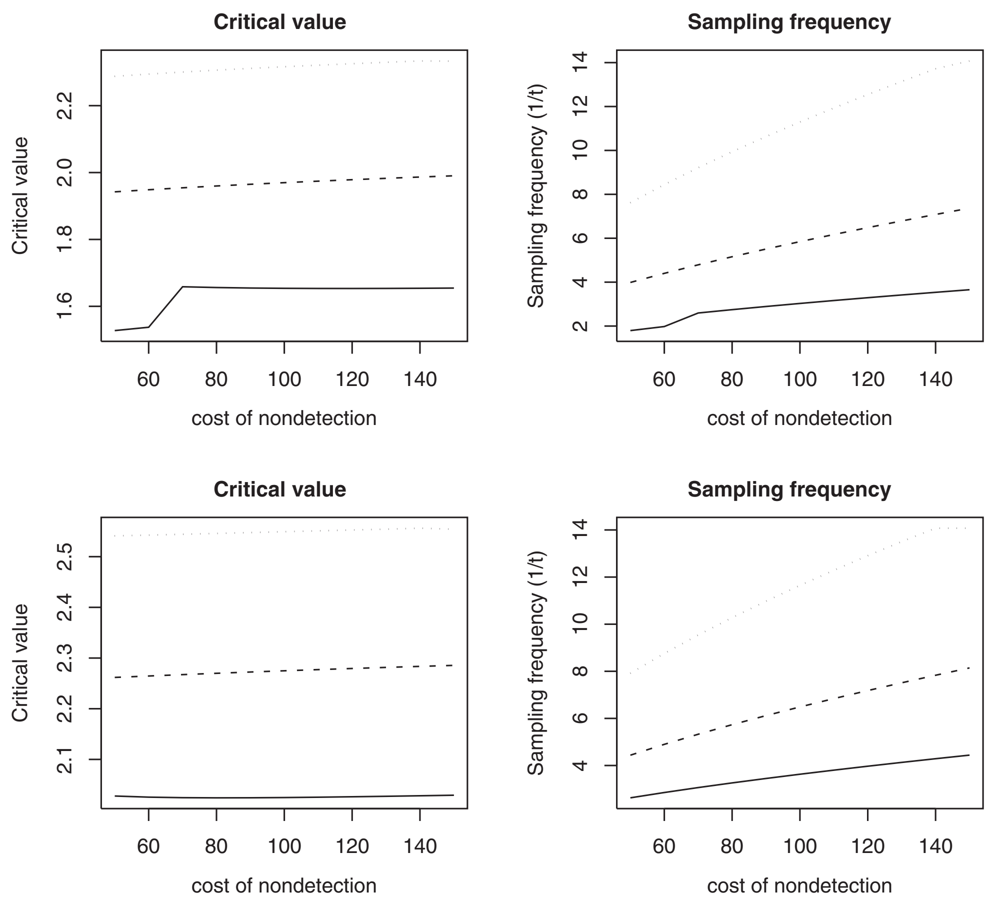

Figure 6. Parameters of the optimal X-chart, for random shift sizes, $d = 0.2$. Top row: $c_{\mathrm{f}} = 50$, bottom row: $c_{\mathrm{f}} = 150$; solid line: $s = 1.4$, broken: $s = 1.3$, dotted: $s = 2.5$.

shifts increase the potential loss due to the form of Taguchian loss function used. For larger values of c_f we may observe a slight increase in the critical values, and correspondingly in sampling frequency.

Figure 7 shows results for the case $d = 0.05$, and their comparison with the results in Figure 6 allow us to observe a substantially smaller sample size coupled with slightly higher critical values.

Finally, Figure 8 shows the average costs, related to the optimal X-chart in this random shift-size set-up. One may see that the higher costs result from the higher average shift (Taguchi-effect) and non-detection costs have a slightly more prominent effect on the total cost than false alarms. The substantial cost reduction for the case of $d = 0.05$ in comparison to $d = 0.2$ is worth noting. This is not surprising since the average frequency of the shift has been reduced by a factor of four, leading to cost reductions of about one half. The relation between $1/d$ and C is non-linear since reductions in sampling frequency and shift occurrence give rise to an increase of the cost for non-detection per unit of time.

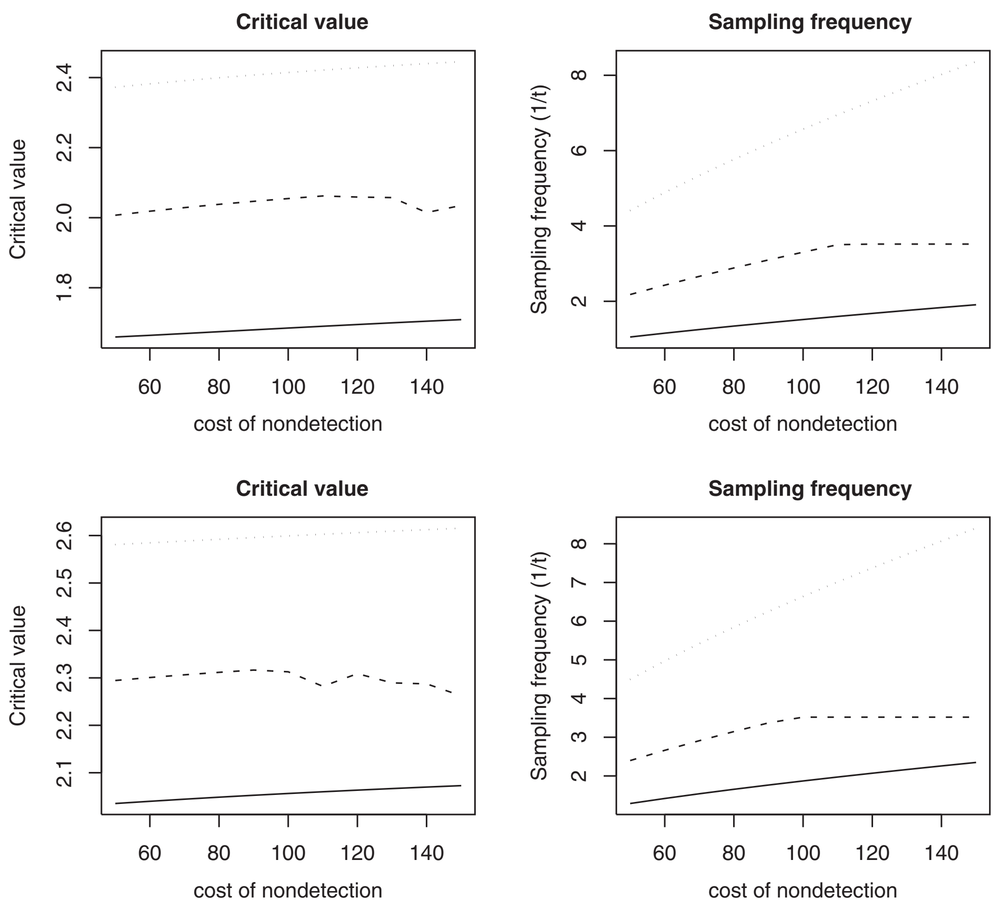

Figure 7. Parameters of the optimal X-chart, for random shift sizes, $d = 0.05$. Top panels: $f = 50$, bottom panels: $f = 150$; solid line: $s = 0.7$, broken: $s = 1.3$, dotted: $s = 2.5$.

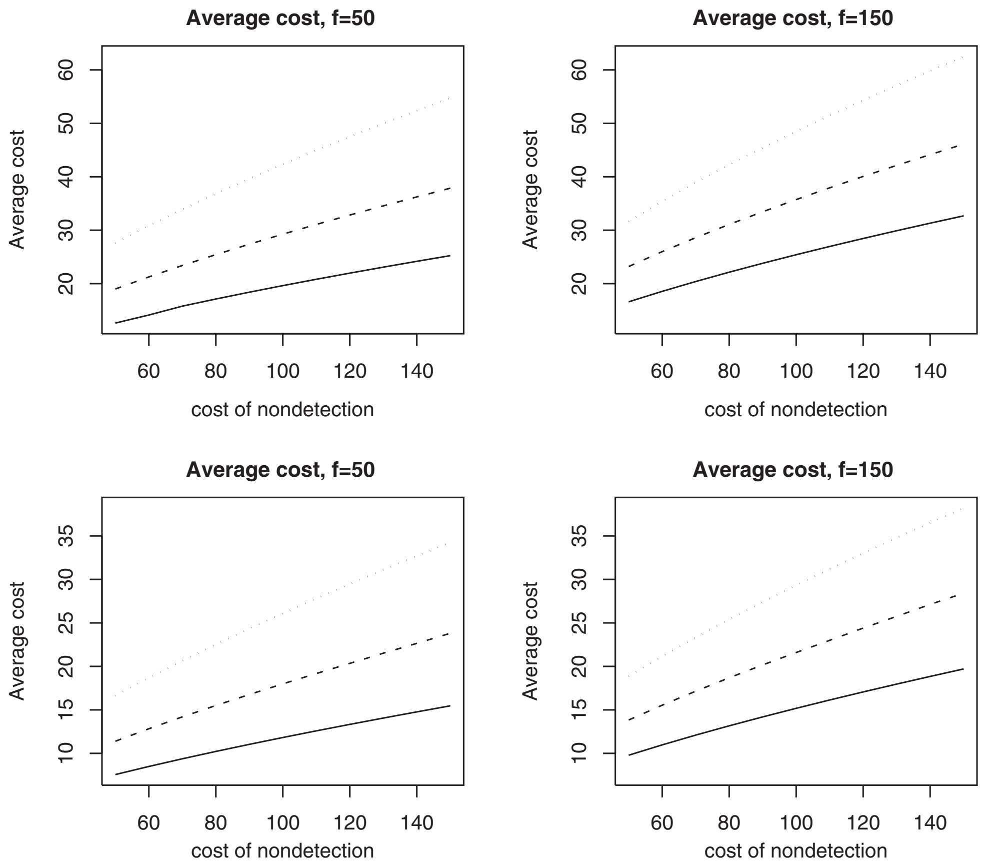

Figure 8. Average cost of the optimal X-charts, for random shift sizes. Upper panels: $d = 0.2$, bottom panels: $d = 0.05$; solid lines: $s = 0.7$, broken: $s = 1.3$, dotted: $s = 2.5$.

4. REAL LIFE APPLICATIONS

4.1. Formulation of the problem

We have applied the above approach to a set of data collected from a Portuguese pulp plant (Companhia de Celulose do Caima, S.A.). This paper mill produces pulp, and one of the quality variables of the final product that needs to be controlled is the pulp brightness. Statistical control of such a variable leads to a one-sided problem, with the target value being 88.5 °ISO and LSL = 87.5 °ISO, where LSL stands for the lower specification level. The production below LSL is re-processed or sold by a small price due to the problems it causes regarding client contracts.

The first step in the optimal chart tuning was to estimate *a priori* process parameters. This resulted in the following estimators:

- $c_{\mathrm{f}} = 90$
- $c_{\mathrm{w}} = 4500$, which represents the maximal loss. At 87.5 °ISO a value $c_{\mathrm{w}} = 90$ was set, thus leading to a two-parameter quadratic form for loss function.

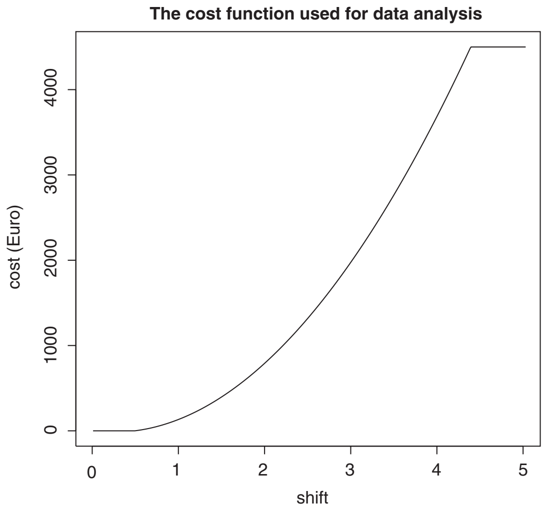

Figure 9. Modified Taguchi-type loss function, used for the real-life data analysis.

Taking into account this information, we have modified the Taguchi-type loss function, as illustrated in Figure 9.

The next step of shift-parameter estimation involves distinguishing between random deviations of the original process and small shifts which take place as a result of special causes of variation. This rather complex task was in this case achieved through a maximum likelihood estimation procedure (explained in the Appendix).

4.2. Data analysis

The process' estimated parameters were $s = 2.5$, $d = 0.33$ and the optimal parameters obtained after standardization were $c = 1.56$ and $1/t = 3.60$. The corresponding minimal value of the cost function is 56.8 monetary units per hour. The currently applied SPC chart is based on five element-samples, corresponding to our set-up with five-times higher sampling cost, and expected shift $s\sqrt{5}$. The optimal solution for such a scenario gives $c = 1.40$, $1/t = 1.43$ and the associated cost function value is 89.9. The current procedure applies the usual 3σ-rule as control limit, which gives rise to a cost of 93.1 Euros. Thus we have shown that at least 40% of the total cost related to sampling can be spared by the suggested approach.

4.3. Sensitivity analysis for the optimal chart

Table II shows the effect when the optimal chart was used for different parameter values (in each simulation a time horizon of 1000 units was used, being replicated 100 times), with the first row corresponding to the situation analysed in Section 4.2. This simulation study was aimed at checking whether the stationary distribution is useful in cost estimation. The results are in good agreement with the cost values shown above, so that the approximation in (A2) does seem to be adequate.

Table II. Average cost of the optimal procedure, for different processes.

d	s	Cost (std. dev.)
0.33	1.25	53.78 (2.05)
0.5	1.5	68.94 (3.44)

The results in the last row of Table II show that even for larger and more frequent shifts the actual costs do not increase dramatically, showing the robustness of the suggested methods.

5. CONCLUSION

We presented an economic approach for tuning statistical control charts based on Markov chains. We firstly derived the methodology for fixed shifts and then extended it for random, exponentially distributed shifts. The performance of CUSUM charts was compared with X-charts, and the approach was extended to accommodate random shift size distributions. The latter cases lead to continuous-state Markov chains that were treated by discretizing the domain of shift size into classes. Finally, the approach was applied to an industrial problem respecting quality control in the pulp and paper industry. Such an application involved the formulation of cost function as a Taguchi type loss function, with the process parameters estimated from data using MLE procedures. The comparison of optimally tuned control X-charts with the actual set-up enables to achieve a substantial cost reduction.

It is important that managers and engineers be aware of both the costs of applying SPC, together with the possible gains resulting from its optimal use. We hope that approaches like the one presented here will become more and more accepted, as the financial gains obtained through a statistically driven six-sigma projects become more and more of common knowledge.

APPENDIX A

A.1. Markov chain method for CUSUM charts

The calculations are based on the following observation. If the distance of the nearest value of the CUSUM chart to the line after n observations is h_n, then

$$h_{n+1} = \min\left(h_n - x_{n+1}\cos\alpha + \frac{\sin\alpha}{j}, a\sin\alpha\right) \tag{A1}$$

For a process under statistical control, this results in a continuous-state Markov chain, where $h_n < a\sin\alpha$. $h_n < 0$ corresponds to the alarm as the line intersects the CUSUM chart. In order to simplify the problem, we approximated it by a discrete Markov chain, dividing the interval $[0, a\sin\alpha]$ into m equal parts, and considering the $m+1$ cutpoints as the possible states of this chain. Now we can formulate the transition from state i to state l, as this is equivalent to

$$l\frac{a\sin\alpha}{m} - \frac{a\sin\alpha}{m} \leqslant i\frac{a\sin\alpha}{m} - x_{n+1}\cos\alpha + \frac{\sin\alpha}{j} \leqslant l\frac{a\sin\alpha}{m} - \frac{a\sin\alpha}{m} \tag{A2}$$

which can be rearranged as

$$v - (2(l-i)+1)w \leqslant x_{n+1} \leqslant v - (2(l-i)-1)w \tag{A3}$$

where

$$v = t\tan\alpha, \quad w = \frac{a\tan\alpha}{2m}$$

As there is the possibility of a (fixed) shift occurrence, we have to introduce another component of the states: 0 corresponds to correct operation, and 1 to the shifted case.

Assuming the lack of shift, the (i, l)th element of the transition matrix is the following:

$$(1-d)[\Phi(v-(2(l-i)-1)w) - \Phi(v-(2(l-i)+1)w)]$$

by (A3); the cases with shift are analogous.

A.2. Transition matrix for random shift sizes

The upper-left corner of the transition matrix is the following (with analogous continuation):

$$\begin{array}{l} \text{state } 1 \\ \text{state } 2 \\ \text{state } 3 \\ \text{state } 4 \end{array} \begin{bmatrix} 1-\gamma_1 & q(1)\Phi(c-\Delta) & q(2)\Phi(c-2\Delta) & q(3)\Phi(c-3\Delta) \\ 1-\gamma_2 & q(0)\Phi(c-\Delta) & q(1)\Phi(c-2\Delta) & q(2)\Phi(c-3\Delta) \\ 1-\gamma_3 & 0 & q(0)\Phi(c-2\Delta) & q(1)\Phi(c-3\Delta) \\ 1-\gamma_4 & 0 & 0 & q(0)\Phi(c-3\Delta) \end{bmatrix} \tag{A4}$$

where γ_i is the sum of the transition probabilities in the ith row from state $\max(i, 2)$ to m, the total number of states.

A.3. Maximum likelihood estimator for shift intensity and size

We present maximum likelihood estimators for d, s and σ. This is preferred to heuristic calculations, especially for the random shift size, since it is not easy to determine the time point of a relatively small shift.

The inference is based on the differences $X_{i+1} - X_i$, these being assumed independent, identically distributed, with X_i denoting the ith observation. The mean value of this difference is ds and its variance consists of two components: a part from the normal operation and an additional one, based on the shift as

$$X_{i+1} - X_i = Z_i + W_i \tag{A5}$$

where Z_i is the random term of normal operation and W_i represents the shift. We used the first two moments of the shift-distribution to formulate the log-likelihood function:

$$l(x) = -n\ln(2\pi(2\sigma^2 + 2d^2s^2))/2 - \sum \frac{(x_i - \mathrm{d}s)^2}{2(2\sigma^2 + 2\mathrm{d}s^2)} \tag{A6}$$

The fact that both ds and ds^2 appeared in the formula for the log-likelihood function (A6) allowed for estimating the parameters separately. A small simulation study was carried out in order to find out the properties of the estimator, with a set of 100 replicates being generated from sample sizes of 100 and 500. The observed results are summarized in Table A1. The parameters to be estimated were $(\sigma, d, s) = (1.0, 0.1, 1.0)$. It might be interesting to investigate the

Table A1. Properties of the maximum likelihood estimator.

n	Mean (std. dev.) for σ	Mean (std. dev.) for d	Mean (std. dev.) for s
100	0.981 (0.084)	0.098 (0.044)	0.997 (0.091)
500	1.001 (0.047)	0.096 (0.020)	1.012 (0.045)

estimator's robustness against discrepancies from the assumed model, but we do not tackle this problem here.

ACKNOWLEDGEMENTS

We acknowledge improvements suggestions made by the anonymous referees. The research was supported by funding under the European Commission's Fifth Framework 'Growth Programme' via Thematic Network "Pro-ENBIS", contract reference G6RT-CT-2001-05059. The author (i.e. Pro-ENBIS) is solely responsible for the content and it does not represent the opinion of the Community, the Community is not responsible for any use that might be made of data therein.

REFERENCES

1. Montgomery DC. *Introduction to Statistical Quality Control*. Wiley: New York, 1985.
2. Does RJMM, Roes KCB, Trip A. *Statistical Process Control in Industry*. Kluwer Publishers: Dordrecht, 1996.
3. Duncan AJ. *Quality Control and Industrial Statistics* (4th edn). Homewood: Illinois, 1974.
4. Arnold BF, Göb R. *Knowledge for Sustainable Development—An Insight into the Encyclopedia of Life Support Systems*, vol. 1. EOLSS Publishers: Oxford, UK, 2002.
5. Véber M, Zempléni A. Optimization of acceptance sampling methods. In: *Proceedings of the 1st Conference of ENBIS*, Oslo, 2001.
6. Klaassen CAJ. Credit in acceptance sampling on attributes. *Technometrics* 2001; **43**(2):211–222.
7. Basseville M, Nikiforov IV. *Detection of Abrupt Changes*: *Theory and Applications*. Prentice Hall: New Jersey, 1993.
8. Chen J, Gupta AK. *Parametric Statistical Change Point Analysis*. Birkhauser: New York, 2000.
9. Darkhovsky BS, Brodsky BE. *Nonparametric Methods in Change Point Problems*. Kluwer Academic Publishers: Amsterdam, 1993.
10. Hines, WGS. A simple monitor of a system with sudden parameter changes. *IEEE Transactions on Information Theory* 1976; **22**(2):210–216.
11. Page ES. Estimating the point of change of a continuous process. *Biometrika* 1957; **44**:248–252.
12. Lorden G. Procedures to react to a change in distribution. *Annals Mathematical Statistics* 1971; **42**:1897–1908.
13. Csörgő M, Horváth L. *Limit Theorems in Change Point Problems*. Wiley: New York, 1997.
14. Karlin S, Taylor HM. *A First Course in Stochastic Processes*. Academic Press: New York, 1975.
15. Woodall WH, Adams BM. The statistical design of CUSUM charts. Q*uality Engineering* 1993; **5**(4):559–566.

