Recent Adv. in Stat. and Prob., pp. 155-171
J. Pérez Vilaplana and M. L. Puri (Eds)

Process Tracking of Time Series with Change Points

S. ZACKS[1] and R.S. KENETT[2]
[1] Department of Mathematical Sciences, SUNY–Binghamton, P.O. Box 6000, Binghamton, NY 13902-6000, U.S.A.
[2] Department of Mathematical Sciences, SUNY–Binghamton, P.O. Box 6000, Binghamton, NY 13902-6000, U.S.A.

Abstract – The problem of industrial process control has been tackled in several ways by both theoreticians and practitioners. The most common approach was developed in the twenties by Shewhart in the form of "control charts". Major developments since the control chart include Cumulative Sum Procedures, Exponentially Weighted Moving Average charts and Fast Initial Response Startups.

We apply here a Bayesian framework for tracking a process over time. We use a Bayesian model that explicitly models the effect of assignable causes as a shift in the mean. We derive, after each observation, the posterior distribution of the process mean, given the observations collected in the past.

The fractiles of the posterior distribution of the process mean and the posterior probabilities of a shift at any given time constitute powerful control and diagnostic tools. The sensitivity of the procedure to the prior parameters and the properties of detection criteria are studied by simulation experiments.

1. INTRODUCTION

How do we keep a process under control? Obviously the decision tool to use depends on how we interpret the process control problem.

The most common interpretation is attributed to WShewhart [1931]. In the Shewhart framework one first identifies process variability due to common causes. Then, control limits based on estimates of the process mean and standard deviation permit us to detect significant assignable causes, thereby triggering corrective actions. The control limits are set to give fixed and preset probabilities for undercontrol and overcontrol activities. Shewhart control charts have been adapted to wear out or other known changes in distribution parameters. For a modern exposition of the subject see the book by Stephens et al (1986). In all cases, control charts carry the flavor of hypothesis testing procedures, with control limits and action rules designed to provide a known significance level, α, such as .27%, 1% or 5%. Using three sigma limits ($\alpha = 0.0027$), for instance, provides a test of stability for statistical variations over time. This approach has proven efficient and practical and gained wide range use and acceptability throughout industry.

A drawback of using control charts for process control is the nature of repeated testing of significance over time. The significance level α is determined from an individual, one time, test perspective. A long term perspective on the performance of process monitoring tools demeans the role of α. In the long run the statistically based control scheme will always trigger an alarm. Overall significance levels are therefore one, and so is the power to detect assignable causes. Any change or non-change in the process triggers action, in the long run. This has led to investigations of Average Run Length (ARL) characteristics of control charts. In order to increase the sensitivity for detecting changes and get improved ARLs, Cumulative Sum procedures (CUSUM) were developed (Page (1954)). Further enhancements to CUSUMs including a combined Shewhart-CUSUM control scheme and Fast Initial Response (FIR) start ups, were proposed by Lucas (1982) and Lucas and Crosier (1982).

Numerical algorithms and analytical determination of the Run Length probability distribution of CUSUM procedures is presented in Zacks (1980, 1981). An extension of CUSUMs to cases were processes cannot be reset is given by Kenett and Pollak (1983). Pollak investigated optimality of an alternative sequential procedure (Pollak (1985)) labeled after Shiryayev and Roberts. Such a Shiryayev-Roberts control scheme was applied to track reliability growth in a software development effort (Kenett and Pollak (1986)).

Another approach was taken by BHoadley and his colleagues at AT&T (Hoadley 1981) resulting in the Quality Measurement Plan (QMP). The QMP mechanism dynamically estimates a quality index distribution using hierarchical Bayes methods. The index is a ratio of observed to expected defects. Posterior fractiles of the quality index are presented in a box and whiskers plot format. An alert is triggered by having the 95th percentile of the index posterior distribution above the standard level of one. When the 99th percentile is above one a "Below Normal" condition is declared. A scheme which is more sensitive to drifts in the mean was developed by Phadke (1981).

Like Hoadley, we use a Bayesian framework for tracking the process over time. The control problem we tackle is that of a continuous variable with a normal distribution. The basic reference for our study is the work of Chernoff and Zacks (1964) originally motivated by a radar tracking problem.

The uncertainty in the process is reflected by a probabilistic structure for the observations, which includes random errors around a mean drawn from a normal distribution.

We assume no initial bias in process startups so that, on the average, the mean is on target. At some unknown point in time a perturbation in the mean occurs. That time instance can be in the past, present or future. Our goal is to control fractiles of the posterior distribution of the current mean and of predictive distributions of future observations, given the past behavior of the process.

In Section 2 we present a Bayesian tracking model for the case of at most one change. The posterior distribution of the current mean, μ_n, is derived in Section 3. As shown, the posterior distribution of μ_n is a mixture of normal distribution, with mixing posterior probabilities which are nonlinear functions of the data. The central moments of the posterior distribution of μ_n are derived, and a method is given for the computation of the posterior fractiles.

Numerical illustrations of the tracking procedure, and an associated detection procedure are given in Section 4. This detection procedure is based on the posterior probability, after n observations, that the change has not yet occurred. An alert is issued when that posterior probability drops below a prescribed threshold, Π^*. At the time of an alert, an estimate of the point of change is computed. Simulation results generated from a fractional factorial computer experiment demonstrate the effects of the prior parameters on the procedure's performance. Section 5 concludes the paper with a discussion of our results.

2. THE BAYESIAN TRACKING MODEL

Observations from a continuous distribution characteristic of a process are recorded at discrete times $i = 1, 2, \ldots$ The observations, X_i, are determined by the value of a base average level, μ_0, and an unbiased error, ϵ_i. Possible system deviations at start-up are modeled by having μ_0 randomly determined according to a normal distribution with known mean μ_T (a specified target value) and a known variance σ^2. At some unknown point in time a disturbance might change the process average level to a new value, $\mu_0 + Z$. The probability of having a disturbance between any two observations, taken at fixed operational time intervals, is considered a constant, p. Accordingly we apply a model of at most one change point at an unknown location. The corresponding statistical model, for at most one change point, within n consecutive observations is as follows:

$$X_i = \mu_i + \epsilon_i, \quad i = 1, 2, \ldots, n. \tag{2.1}$$

Where μ_i is the observation's mean and ϵ_i the measurement error, the at-most-one-change (AMOC) point model is:

$$\mu_i = \begin{cases} \mu_0, & i = 1, \ldots, j-1 \\ \mu_0 + Z, & i = j, \ldots, n \end{cases} \tag{2.2}$$

Where j indicates that the disturbance occurred in the time interval between the $(j-1)$st and j-th observations. $j = 1$ indicates a disturbance before X_1, $j = n+1$ indicates no disturbance among the first n observations.

The Bayesian framework is based on the following distributional assumptions:

$$\mu_0 \sim N(\mu_T, \sigma^2) \tag{2.3}$$

$$Z \sim N(\delta, \tau^2), \tag{2.4}$$

$$\epsilon_i \sim N(0, 1) \quad i = 1, 2, \ldots \tag{2.5}$$

The parameters μ_T, σ^2, τ^2 are assumed to be known and $N(\cdot, \cdot)$ designates a random variable having a normal distribution law. Assumption (2.5) of the model requires that the raw data be scaled, to have conditional variance equal to 1.

For n observations, let J_n denote the random variable corresponding to the index of the time interval at which a change occurred. It is assumed that the prior distribution of J_n is the truncated geometric, i.e.,

$$\begin{aligned} \Pr(J_n = j \mid p) &= p(1-p)^{j-1}, \quad j = 1, 2, \ldots, n; \qquad 0 < p < 1 \\ &= (1-p)^n, \qquad j = n+1; \qquad 0 < p < 1 \end{aligned} \tag{2.6}$$

The assumption underlining (2.6) is that disturbances occur at random according to a homogeneous Poisson process. Finally, we assume that μ_0, Z, J_n and ϵ are independent, for each n.

3. THE POSTERIOR DISTRIBUTION OF THE CURRENT MEAN

In the present section we develop the formula of the posterior distribution of μ_n, given $\mathbf{X}^{(n)} = (X_1, \ldots, X_n)$, and compute its moments and fractiles.

3.1. General Derivation

For $j = 1, 2, \ldots, n+1$, let

$$\bar{X}_{j-1} = \frac{1}{j-1}\sum_{i=1}^{j-1} X_i \quad \text{and} \quad \bar{X}^*_{n-j+1} = \frac{1}{n-j+1}\sum_{i=j}^{n} X_i$$

where $\bar{X}_0 \equiv 0$ and $\bar{X}^*_0 \equiv 0$. The current mean μ_n can be written, in the AMOC model, as

$$\mu_n = \mu_0 + I\{J_n \leq n\}Z, \tag{3.1}$$

where $I\{\cdot\}$ is an indicator variable, assuming the value 1 if the relation in brackets is true, and the value 0 otherwise. Thus, given $\{J_n = j, \mathbf{X}^{(n)}\}$, the Bayesian model of Section 2 implies that the posterior distribution of μ_n is normal $\mathcal{N}\big(E_{j,n}(\mathbf{X}^{(n)}),\ D^2_{j,n}\big)$, where $E_{j,n}(\mathbf{X}^{(n)})$ is the posterior mean of μ_n, given $\{J_n = j,\ \mathbf{X}^{(n)}\}$, and $D^2_{j,n}$ is the corresponding posterior variance. Formulae for $E_{j,n}(\mathbf{X}^{(n)})$ and $D^2_{j,n}$, $j = 1, 2, \ldots, n+1$, will be derived below.
Let $\Pi(J_n = j \mid \mathbf{X}^{(n)})$ denote the posterior probability function of $\{J_n = j\}$, given $\mathbf{X}^{(n)}$. Then, the posterior distribution law of μ_n, given $\mathbf{X}^{(n)}$, $\mathcal{L}(\mu_n \mid \mathbf{X}^{(n)})$, is the mixture of the normals $\mathcal{N}\big(E_{j,n}(\mathbf{X}^{(n)}),\ D^2_{j,n}\big)$, with mixing probabilities $\Pi_{j,n}(\mathbf{X}^{(n)}) = \Pi(J_n = j \mid \mathbf{X}^{(n)})$, i.e.,

$$\mathcal{L}(\mu_n \mid \mathbf{X}^{(n)}) = \sum_{j=1}^{n+1} \Pi_{j,n}(\mathbf{X}^{(n)})\ \mathcal{N}\big(E_{j,n}(\mathbf{X}^{(n)}), D^2_{j,n}\big). \tag{3.2}$$

To determine $E_{j,n}(\mathbf{X}^{(n)})$ and $D^2_{j,n}$, for $j = 1, \ldots, n+1$, notice that the joint distribution of $(\mu_n, \bar{X}_{j-1}, \bar{X}^*_{n-j+1})$, given $\{J_n = j\}$, $j = 2, \ldots, n$, is multi-normal with mean $(\mu_T + \delta, \mu_T, \mu_T + \delta)$ and covariance matrix

$$\Sigma_{\mid j} = \begin{bmatrix} \sigma^2+\tau^2 & \ldots & \sigma^2 & \sigma^2+\tau^2 \\ \vdots & \ddots & \vdots & \vdots \\ \sigma^2 & \ldots & \sigma^2 + \dfrac{1}{j-1} & \sigma^2 \\ \sigma^2+\tau^2 & \ldots & \sigma^2 & \sigma^2+\tau^2+\dfrac{1}{n-j+1} \end{bmatrix}. \tag{3.3}$$

From the common formulae of regression theory we obtain, for $j = 2, \ldots, n$.

$$E_{j,n}(\mathbf{X}^{(n)}) = \mu_0 + \delta + \tag{3.4}$$

$$+ (\sigma^2, \sigma^2+\tau^2)\begin{bmatrix} \sigma^2 + \dfrac{1}{j-1} & \sigma^2 \\ \sigma^2 & \sigma^2+\tau^2+\dfrac{1}{n-j+1} \end{bmatrix}^{-1} \begin{pmatrix} \bar{X}_{j-1} - \mu_T \\ \bar{X}^*_{n-j+1} - \mu_T - \delta \end{pmatrix},$$

and

$$D^2_{j,n} = \sigma^2+\tau^2-(\sigma^2,\sigma^2+\tau^2)\begin{bmatrix} \sigma^2 + \dfrac{1}{j-1} & \sigma^2 \\ \sigma^2 & \sigma^2+\tau^2+\dfrac{1}{n-j+1} \end{bmatrix}^{-1} \begin{pmatrix} \sigma^2 \\ \sigma^2+\tau^2 \end{pmatrix}. \tag{3.5}$$

The cases of $j = 1$ or $j = n + 1$ are similar. Thus, after some algebraic manipulations we obtain the following formulae:

$$E_{j,n}(\mathbf{X}^{(n)}) = \begin{cases} \dfrac{\mu_T + \delta}{1 + n(\sigma^2 + \tau^2)} + \dfrac{n(\sigma^2 + \tau^2)}{1 + n(\sigma^2 + \tau^2)} \bar{X}_n , & j = 1 \\ W_{j,n}^{(1)}(\mu_T + \delta + (j-1)\sigma^2) + W_{j,n}^{(2)} \bar{X}_{j-1} + W_{j,n}^{(3)} \bar{X}^*_{n-j+1}, & j = 2, \dots, n \\ \dfrac{\mu_T}{1 + n\sigma^2} + \dfrac{n\sigma^2}{1 + n\sigma^2} \bar{X}_n, & j = n + 1 , \end{cases} \quad (3.6)$$

where

$$\begin{aligned} W_{j,n}^{(1)} &= [1 + n\sigma^2 + \tau^2(n - j + 1)(1 + \sigma^2(j-1))]^{-1} \\ W_{j,n}^{(2)} &= \sigma^2(j-1) W_{j,n}^{(1)} , \qquad j = 2, \dots, n \\ W_{j,n}^{(3)} &= 1 - W_{j,n}^{(1)} - W_{j,n}^{(2)} \end{aligned} \quad (3.7)$$

Furthermore,

$$D_{j,n}^2 = \begin{cases} \left(\sigma^2 + \tau^2 + \sigma^2\tau^2(j-1)\right) W_{n,j}^{(1)} , & j = 1, \dots, n \\ \dfrac{n\sigma^2}{1 + n\sigma^2} , & j = n + 1 . \end{cases} \quad (3.8)$$

We derive now the formulae of the posterior probabilities $\Pi_{j,n}(\mathbf{X}^{(n)})$.

3.2. The Likelihood Functions and Posterior Probabilities

Let $L(j, \mu_0, Z; \mathbf{X}^{(n)})$ denote the likelihood function of (j, μ_0, Z) given the first n observations $\mathbf{X}^{(n)}$. Let $L^*(j; \mathbf{X}^{(n)})$ denote the marginal (predictive) likelihood of j, given $\mathbf{X}^{(n)}$; which is the expected value of $L(j, \mu_0, Z; \mathbf{X}^{(\mathbf{n})})$ with respect to the prior distribution of (μ_0, Z). By Bayes theorem, for all $j = 1, 2, \dots$

$$\Pi_{j,n}(\mathbf{X}^{(n)}) = \frac{I\{j \le n\} p(1-p)^{j-1} L^*(j; \mathbf{X}^{(n)}) + I\{j \ge n+1\}(1-p)^n L^*(n+1; \mathbf{X}^{(n)})}{p \sum_{j=1}^{n} (1-p)^{j-1} L^*(j; \mathbf{X}^{(n)}) + (1-p)^n L^*(n+1; \mathbf{X}^{(n)})} \quad (3.9)$$

We further develop this formula. From the normal distribution of ϵ_i $(i = 1, \dots, n)$,

$$\begin{aligned} L(j, \mu_0, Z; \mathbf{X}^{(n)}) &= \frac{1}{(2\pi)^{n/2}} \exp\left\{ -\frac{1}{2}\sum_{l=1}^{j-1}\left(X_l - \mu_0\right)^2 - \frac{1}{2}\sum_{l=j}^{n}\left(X_l - \mu_0 - Z\right)^2 \right\} = \quad (3.10) \\ &= \frac{1}{(2\pi)^{n/2}} \exp\left\{ -\frac{1}{2} Q_{j-1} - \frac{1}{2} Q^*_{n-j+1} - \frac{j-1}{2}\left(\bar{X}_{j-1} - \mu_0\right)^2 - \right. \\ &\qquad \left. - \frac{n-j+1}{2}\left(\bar{X}^*_{n-j+1} - \mu_0 - Z\right)^2 \right\}, \qquad j = 2, \dots, n \end{aligned}$$

where $Q_{j-1} = \sum_{l=1}^{j-1} (X_l - \bar{X}_{j-1})^2$ and $Q^*_{n-j+1} = \sum_{l=j}^{n} (X_l - \bar{X}^*_{n-j+1})^2$. When $j = 1$ or $j = n + 1$ we substitute in (3.10) $Q_0 \equiv 0$ or $Q^*_0 \equiv 0$, respectively. Let $Q_n = \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$. Simple algebraic manipulations yield

$$-\frac{1}{2}Q_{j-1}-\frac{1}{2}Q^*_{n-j+1}=-\frac{1}{2}\left[Q_n-2(j-1)\left(1-\frac{j-1}{n}\right)(\bar{X}_{j-1}-\bar{X}^*_{n-j+1})^2\right], \quad (3.11)$$

$j=1,\dots,n+1$. Substituting (3.11) in (3.10) we obtain

$$L(j,\mu_0,Z;\mathbf{X}^{(n)})= \quad (3.12)$$
$$=\frac{1}{(2\pi)^{n/2}}\exp\left\{-\frac{1}{2}Q_n\right\}\exp\left\{(j-1)\left(1-\frac{j-1}{n}\right)\left(\bar{X}_{j-1}-\bar{X}^*_{n-j+1}\right)^2-\right.$$
$$\left.-\frac{j-1}{2}\left(\bar{X}_{j-1}-\mu_0\right)^2-\frac{n-j+1}{2}\left(\bar{X}^*_{n-j+1}-\mu_0-Z\right)^2\right\}.$$

Moreover, for $j=1,\dots,n+1$,

$$L^*(j;\mathbf{X}^{(n)}) = \quad (3.13)$$
$$=\frac{1}{(2\pi)^{n/2}}\exp\left\{-\frac{1}{2}Q_n\right\}\exp\left\{(j-1)\left(1-\frac{j-1}{n}\right)\left(\bar{X}_{j-1}-\bar{X}^*_{n-j+1}\right)^2\right\}\cdot$$
$$\cdot E\left\{\exp\left\{-\frac{j-1}{2}\left(\bar{X}_{j-1}-\mu_0\right)^2-\frac{n-j+1}{2}\left(\bar{X}^*_{n-j+1}-\mu_0-Z\right)^2\right\}\right\},$$

where the expectation in (3.13) is with respect to the prior bivariate normal distribution of (μ_0, Z). Thus,

$$L^*(j;\mathbf{X}^{(n)}) = \frac{1}{(2\pi)^{\frac{n}{2}-1}}\exp\left\{-\frac{1}{2}Q_n\right\}\cdot \quad (3.14)$$
$$\cdot\left\{(j-1)\left(1-\frac{j-1}{n}\right)\left(\bar{X}_{j-1}-\bar{X}^*_{n-j+1}\right)^2\right\}L^{**}(j;\mathbf{X}^{(n)}),$$

$j=1,\dots,n+1$, where

$$L^{**}(1;\mathbf{X}^{(n)})=\frac{1}{\left(1+n(\sigma^2+\tau^2)\right)^{1/2}}\cdot\exp\left\{-\frac{n}{2\left(1+n(\sigma^2+\tau^2)\right)}(\bar{X}^2_n-\mu_T-\delta)^2\right\}, \quad (3.15)$$

and for $j=2,\dots,n$,

$$L^{**}(j;\mathbf{X}^{(n)})=\frac{1}{\left(1+n\sigma^2+(n-j+1)\tau^2(1+(j-1)\sigma^2)\right)^{1/2}}\cdot \quad (3.16)$$
$$\cdot\exp\left\{-\frac{1}{2\left(1+n\sigma^2+(n-j+1)\tau^2(1+(j-1)\sigma^2)\right)}\cdot\right.$$
$$\cdot\left[\frac{1+(n-j+1)(\sigma^2+\tau^2)}{n-j+1}\cdot\right.$$
$$\cdot\left((\bar{X}_{j-1}-\mu_t)-2\sigma^2(\bar{X}_{j-1}-\mu_T)(\bar{X}^*_{n-j+1}-\mu_T-\delta)\right)+$$
$$\left.\left.+\frac{1+(j-1)\sigma^2}{j-1}\left(\bar{X}^*_{n-j+1}-\mu_T-\delta\right)^2\right]\right\},$$

while

$$L^{**}(n+1;\mathbf{X}^{(n)})=\frac{1}{\sqrt{1+n\sigma^2}}\exp\left\{-\frac{n}{2(1+n\sigma^2)}(\bar{X}_n-\mu_T)^2\right\}. \quad (3.17)$$

3.3. Moments of the Posterior Distribution of μ_n

Let $m_r^*(\mathbf{X}^{(n)})$ denote the r-th central moment of the posterior distribution of μ_n, given $\mathbf{X}^{(n)}$; i.e.,

$$m_r^*(\mathbf{X}^{(n)}) = E\left\{(\mu_n - \hat{\mu}_n)^r \mid \mathbf{X}^{(n)}\right\}, \tag{3.18}$$

where $\hat{\mu}_n = \sum_{j=1}^{n+1} \Pi_{j,n}(\mathbf{X}^{(n)})\, E_{j,n}(\mathbf{X}^{(n)})$ is the mean of the posterior distribution (the Bayes estimator of μ_n for a squared-error loss). According to the mixture (3.2), the central posterior moments of μ_n can be expressed in terms of the corresponding moments of $\mathcal{N}\left(E_{j,n}(\mathbf{X}^{(n)}),\, D_{j,n}^2\right)$ and the moments of $E_{j,n}(\mathbf{X}^{(n)})$ around $\hat{\mu}_n$. Let $M^*(t)$ designate the posterior moment generating function of $(\mu_n - \hat{\mu}_n)$, i.e.,

$$\begin{aligned} M^*(t) &= E\left\{e^{t(\mu_n-\hat{\mu}_n)} \mid \mathbf{X}^{(n)}\right\} = \\ &= \sum_{j=1}^{n+1} \Pi_{j,n}(\mathbf{X}^{(n)}) e^{t\left(E_{j,n}(\mathbf{X}^{(n)})-\hat{\mu}_n\right)}\, E\left\{e^{t\left(\mu_n - E_{j,n}(\mathbf{X}^{(n)})\right)} \mid J = j, \mathbf{X}^{(n)}\right\}. \end{aligned} \tag{3.19}$$

Since the conditional posterior distribution of μ_n, given $\{J = j, \mathbf{X}^{(n)}\}$ is $\mathcal{N}\left(E_{j,n}(\mathbf{X}^{(n)}),\right.$ $\left.D_{j,n}^2\right)$, the conditional posterior moment generating function is

$$M_j^*(t) = E\left\{e^{t\left(\mu_n - E_{j,n}(\mathbf{X}^{(n)})\right)} \mid J = j, \mathbf{X}^{(n)}\right\} = \exp\left\{\frac{t^2}{2} D_{j,n}^2\right\}, \quad j = 1, \ldots, n+1. \tag{3.20}$$

Thus,

$$M^*(t) = \sum_{j=1}^{n+1} \Pi_{j,n}(\mathbf{X}^{(n)}) \exp\left\{t\left(E_{j,n}(\mathbf{X}^{(n)}) - \hat{\mu}_n\right) + \frac{t^2}{2} D_{j,n}^2\right\}. \tag{3.21}$$

From (3.21) one can obtain the central posterior moments of μ_n, given $\mathbf{X}^{(n)}$. In particular,

$$\mathrm{Var}\left\{\mu_n \mid \mathbf{X}^{(n)}\right\} = \sum_{j=1}^{n+1} \Pi_{j,n}\left(\mathbf{X}^{(n)}\right) D_{j,n}^2 + \sum_{j=1}^{n+1} \Pi_{j,n}\left(\mathbf{X}^{(n)}\right)\left(E_{j,n}\left(\mathbf{X}^{(n)}\right) - \hat{\mu}_n\right)^2, \tag{3.22}$$

the third central moment is

$$m_3^*\left(\mathbf{X}^{(n)}\right) = \sum_{j=1}^{n+1} \Pi_{j,n}\left(\mathbf{X}^{(n)}\right)\left[E_{j,n}\left(\mathbf{X}^{(n)}\right) - \hat{\mu}_n\right]^3 + 3\sum_{j=1}^{n+1} \Pi_{j,n}\left(\mathbf{X}^{(n)}\right)\left[E_{j,n}\left(\mathbf{X}^{(n)}\right) - \hat{\mu}_n\right] D_{j,n}^2 \tag{3.23}$$

and the fourth central moment is

$$\begin{aligned} m_4^*\left(\mathbf{X}^{(n)}\right) &= \sum_{j=1}^{n+1} \Pi_{j,n}\left(\mathbf{X}^{(n)}\right)\left[E_{j,n}\left(\mathbf{X}^{(n)}\right) - \hat{\mu}_n\right]^4 + \\ &\quad + 6\sum_{j=1}^{n+1} \Pi_{j,n}\left(\mathbf{X}^{(n)}\right)\left[E_{j,n}\left(\mathbf{X}^{(n)}\right) - \hat{\mu}_n\right]^2 D_{j,n}^2 + 3\sum_{j=1}^{n+1} \Pi_{j,n}\left(\mathbf{X}^{(n)}\right) D_{j,n}^4. \end{aligned} \tag{3.24}$$

3.4. Posterior Fractiles of μ_n

According to (3.2), the posterior c.d.fóf μ_n, at a point x is

$$F(x \mid \mathbf{X}^{(n)}) = \sum_{j=1}^{n+1} \Pi_{j,n}(\mathbf{X}^{(n)}) \Phi\left(\frac{x - E_{j,n}(\mathbf{X}^{(n)})}{D_{j,n}}\right), \tag{3.25}$$

where $\Phi(Z)$ is the standard normal integral. Accordingly, the p-th fractile of μ_n is the root $\mu_n(p)$ of the equation

$$\sum_{j=1}^{n+1} \Pi_{j,n}(\mathbf{X}^{(n)})\Phi\left(\frac{\mu_n(p) - E_{j,n}(\mathbf{X}^{(n)})}{D_{j,n}}\right) \equiv p, \tag{3.26}$$

for $0 < p < 1$. The root $\mu_n(p)$ of the non-linear equation (3.26) can be obtained by the Newton-Raphson iterative formula:

$$\mu_n^{(i+1)}(p) = \mu_n^{(i)}(p) - \frac{\sum_{j=1}^{n+1}\Pi_{j,n}\left(\mathbf{X}^{(n)}\right)\Phi\left(\frac{\mu_n^{(i)}(p) - E_{j,n}\left(\mathbf{X}^{(n)}\right)}{D_{j,n}}\right) - p}{\sum_{j=1}^{n+1}\frac{\Pi_{j,n}\left(\mathbf{X}^{(\mathbf{n})}\right)}{\mathbf{D_{j,n}}}\psi\left(\frac{\mu_n^{(i)}(p) - E_{j,n}\left(\mathbf{X}^{(n)}\right)}{D_{j,n}}\right)} \tag{3.27}$$

$i = 1, 2, \ldots$, where $\psi(Z)$ is the standard normal p.d.f. One can start the iterations with the initial approximation

$$\mu_n^{(1)}(p) = \sum_{j=1}^{n} \Pi_{j,n}\left(\mathbf{X}^{(n)}\right) E_{j,n}\left(\mathbf{X}^{(n)}\right) + Z_p \sum_{j=1}^{n+1} \Pi_{j,n}\left(\mathbf{X}^{(n)}\right) D_{j,n}, \tag{3.28}$$

in which $Z_p = \Phi^{-1}(p)$.
Numerical examples show that generally (3.27) requires only a few iterations to obtain convergence within four or five significant figures.

4. NUMERICAL EXAMPLES AND SENSITIVITY ANALYSIS

4.1. Numerical Examples

In Tables 4.1 and 4.2 we present portions of two simulated sequences (SIM1 and SIM2) of 15 observations, with a change point at $J = 10$. In both sequences the prior parameters are $\mu_T = 0$, $\delta = 1.5$, $\sigma^2 = \tau^2 = 1$. The random values of μ_i $(i = 1, \ldots, 15)$ for the two sequences are:

$$\mu_i^{(1)} = \begin{cases} -1.91523, & i \leq 10 \\ 0.53421, & i \geq 11 \end{cases}$$

and

$$\mu_i^{(2)} = \begin{cases} 1.32238, & i \leq 10 \\ 0.72655, & i \geq 11 \end{cases}$$

In Tables 4.1 and 4.2 we also present the posterior probabilities of $\{J = j\}$, computed sequentially on the basis of $\mathbf{X}^{(i)}$, $i = 7, \ldots, 13$. We see in Table 4.1 (SIM1) that for $i \leq 10$, $\Pi_{i+1,i}(\mathbf{X}^{(i)})$ assumes the maximal value, which is always greater than 0.8; while immediately after the change, for $i \geq 11$, $\Pi_{11,i}(\mathbf{X}^{(i)})$ assumes the maximum (close to 0.9), while $\Pi_{i+1,i}(\mathbf{X}^{(i)})$ is close to zero. This is a case in which a detection of a change point based on $\Pi_{ji}(\mathbf{X}^{(i)})$ can be easily attained. This is not the case in Table 4.2. The simulated sequence SIM2 does not lead to a close detection of the change point. The value of the prior parameter δ is small relative to $(\tau^2 + \sigma^2)$.

Table 4.1.

Posterior Probabilities $\Pi_{j,i}(\mathbf{X}^{(i)})$, and Posterior Fractiles, Means and Standard Deviations of μ_i ($i = 7, 8, \ldots, 13$), Based on the Simulated Sequence SIM1. Change Point at $i = 11$.

i	7	8	9	10	11	12	13
X_i \ j	-2.081	-2.837	-0.731	-1.683	1.289	0.565	0.256
1	0.0027	0.0028	0.0023	0.0025	0.	0.	0.
2	0.0057	0.0066	0.0037	0.0037	0.	0.	0.
3	0.0022	0.0023	0.0016	0.0016	0.	0.	0.
4	0.0072	0.0085	0.0029	0.0026	0.	0.	0.
5	0.0050	0.0054	0.0019	0.0017	0.	0.	0.
6	0.0047	0.0043	0.0017	0.0016	0.	0.	0.
7	0.0194	0.0039	0.0051	0.0055	0.0004	0.0001	0.
8	<u>0.9606</u>	0.0079	0.0078	0.0074	0.0011	0.0003	0.0002
9		<u>0.9583</u>	0.1583	0.0643	0.0964	0.0794	0.0927
10			<u>0.8149</u>	0.0139	0.0155	0.0149	0.0184
11				<u>0.8952</u>	<u>0.8833</u>	<u>0.9052</u>	<u>0.8886</u>
12					0.0032	0.0001	0.0001
13						0.	0.
14							0.
$\mu_{.01}$	-2.9375	-2.9796	-2.7464	-2.6971	-1.7890	-1.0448	-0.9110
$\mu_{.25}$	-2.3287	-2.4012	-2.1939	-2.1819	-0.5875	-0.1011	-0.0948
$\mu_{.50}$	-2.0855	-2.1720	-1.9473	-1.9664	-0.1136	0.2923	0.2431
$\mu_{.75}$	-1.8399	-1.9419	-1.6270	-1.7358	0.3684	0.6907	0.5844
$\mu_{.99}$	-0.8312	-1.2333	0.8723	-0.2637	1.5813	1.6683	1.4259
$\hat{\mu}_i$	-2.0694	-2.1617	-1.7315	-1.9184	-0.1085	0.2968	0.2463
$v_i^{1/2}$	0.3830	0.3652	0.7782	0.4364	0.8368	0.5848	0.5281

It is of practical and theoretical interest to study the properties of the sampling distributions of various statistics of interest. From a Bayesian point of view, the Bayes estimators of the current mean, μ_n, are optimal as long as the Bayesian model is valid. The question is, how good and effective are the estimators when the Bayesian model is not necessarily valid. Answers to various such questions can be obtained by appropriate simulation experiments.

For example, if we consider a detection procedure which flags a warning on the basis of the stopping time

$$N_{\Pi^*} = \text{ first } \ n \geq 2 \ \text{ for which } \ \Pi_{n+1,n}(\mathbf{X}^{(n)}) \leq \Pi^*, \tag{4.1}$$

it would be of interest to estimate the probability distribution of $(N_{\Pi^*} - J)$, where J is the actual (random) epoch of change. This probability distribution depends on the parameters of the process and on the constants we use in the detection procedure.

Table 4.2.
Posterior Probabilities $\Pi_{ji}(\mathbf{X}^{(i)})$ and Posterior Fractiles, Means and Standard Deviations of μ_i ($i = 7, \ldots, 13$), Based on the Simulated Sequence SIM2. Change Point at $i = 11$.

i	7	8	9	10	11	12	13
X_i \ j	1.483	0.959	1.057	1.419	2.614	0.046	1.626
1	0.0197	0.0187	0.0181	0.0183	0.0203	0.0185	0.0191
2	0.0286	0.0229	0.0200	0.0200	0.0268	0.0192	0.0203
3	0.0206	0.0144	0.0177	0.0115	0.0168	0.0106	0.0112
4	0.0069	0.0049	0.0041	0.0040	0.0054	0.0036	0.0038
5	0.0030	0.0025	0.0023	0.0022	0.0022	0.0020	0.0020
6	0.0083	0.0041	0.0030	0.0027	0.0039	0.0023	0.0024
7	0.0059	0.0031	0.0024	0.0021	0.0022	0.0018	0.0017
8	<u>0.9070</u>	0.0048	0.0030	0.0023	0.0023	0.0018	0.0017
9		<u>0.9246</u>	0.0047	0.0029	0.0039	0.0019	0.0018
10			<u>0.9310</u>	0.0055	0.0098	0.0024	0.0022
11				<u>0.9286</u>	0.0341	0.0030	0.0025
12					<u>0.8724</u>	0.0081	0.0030
13						<u>0.9249</u>	0.0065
14							<u>0.9219</u>
$\mu_{.01}$	0.5112	0.5142	0.5290	0.5760	0.6537	0.6240	0.6822
$\mu_{.25}$	1.1147	1.0802	1.0654	1.0886	1.2105	1.1094	1.1416
$\mu_{.50}$	1.3620	1.3107	1.2831	1.2962	1.4162	1.3007	1.3258
$\mu_{.75}$	1.6154	1.5442	1.5029	1.5059	1.6270	1.4933	1.5118
$\mu_{.99}$	2.3380	2.1641	2.0771	2.0607	2.4834	2.0004	2.0148
$\hat{\mu}_i$	1.3710	1.3149	1.2861	1.2995	1.4293	1.3024	1.3291
$v_i^{1/2}$	0.3829	0.3502	0.3296	0.3161	0.3544	0.2931	0.2839

In Section 4.2 we present results of simulation analysis, in which the parameters of the detection procedure coincide with those of the process.
In Section 4.3 we present the results of a sensitivity analysis based on simulation, in which the parameters of the procedure are different from those of the process.

4.2. The Joint Empirical Distribution of (N_{Π^*}, N_A) and Related Statistics.

We present in the Tables 4.3 to 4.5 the results of a simulation experiment, designed to study the sampling distribution of N_{Π^*} and of the mode, N_A, of the posterior distribution of J_n at the stopping time N_{Π^*}. The mode N_A provides an estimate of the location of the shift. To avoid a high proportion of false alarms, in which $N_{\Pi^*} \leq J - 1$, we used the constant $\Pi^* = .5$ for the stopping time (4.1). In the present experiment we consider the

Table 4.3.
Frequency Distribution of (N, NA) Obtained by 100 Simulation Runs, with Parameters $\mu_T = 0$, $\sigma^2 = \tau^2 = 1$, $\delta = 1.5$, $p = .01$, $\Pi^* = .5$.

N_{Π^*}	*NA*																					SUM
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
1																						
2																						
3																						
4			1																			1
5					1	1																2
6		1																				1
7							2															2
8								1														1
9							1		1													2
10										1	3											4
11									1	4	28	1										34
12									1	1	8	7										17
13											2	3	4	2								11
14										1				1								2
15												1	1			3						5
16																						
17																		1				1
18																			1			1
19																				1		1
20																					15	15
SUM		1	1		1	1	3	1	3	7	41	12	5	3		3		1	1	1	15	100

Table 4.4.
Frequency Distribution of (N_{Π^*}, NA) Obtained by 100 Simulation Runs, with Parameters $\mu_T = 0$, $\sigma^2 = \tau^2 = 1$, $\delta = 2.0$, $p = .01$, $\Pi^* = .5$.

N_{Π^*}	NA																					SUM
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
1																						
2		2																				2
3																						
4			1	2																		3
5		2			2																	4
6					1																	1
7						1	1															2
8							1	2	1													4
9									1													1
10										1												1
11								1	1	2	22	4										30
12					1					2	8	5	2									18
13										1	1	2	8	1								13
14											1		1	1	2							5
15											2		1									3
16															1		1					2
17																1						1
18																1			1			2
19																				1		1
20																					7	7
SUM		4	1	2	4	1	2	3	3	6	34	11	12	2	3	2	1		1	1	7	100

Table 4.5.
Frequency Distribution of (N_{Π^*}, NA) Obtained by 100 Simulation Runs,
With Parameters $\mu_T = 0$, $\sigma^2 = \tau^2 = 1$, $\delta = 2.5$, $p = .01$, $\Pi^* = .5$.

N_{Π^*}	NA																					SUM
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
1																						
2		1																				1
3																						
4	1		1	1																		3
5					2	1																3
6					1																	1
7					1																	1
8								2	1													3
9									1													1
10							1			3												4
11							2			1	49	2										54
12											9	6										15
13											2		1	2								5
14											1			1								2
15															1							1
16										1							1					1
17														1								1
18																		1				1
19																						
20																					3	3
SUM	1	1	1	1	4	1	3	2	2	5	61	8	1	4	1			1			3	100

case in which one knows the parameters of the process μ_T, σ^2, τ^2 and δ. We applied the values $\mu_T = 0$, $\sigma^2 = \tau^2 = 1$ and $\delta = 1.5$, 2.0 and 2.5. The joint frequency distributions of (N_{Π^*}, N_A), obtained from 100 independent simulation runs (for each case) are presented in Tables 4.3-4.5. In these simulations the epoch of change is fixed at $J = 10$. Each run is terminated after $n = 20$ observations, if it has not stopped earlier. The proportion of runs in which $N_{\Pi^*} \leq 10$ is an estimator of the "false alarms".

The statistic

$$CED = \sum_{n=11}^{20} (n - J) f_n / \sum_{n=11}^{20} f_n, \tag{4.2}$$

in which f_n ($n = 1, \ldots, 20$) are the marginal frequencies of N_{Π^*}, is an estimator of the conditional expected delay, i.e., $E\{N_{\Pi^*} - J \mid N_{\Pi^*} \geq J\}$. Notice that this empirical estimator is biased downwards because of the truncation at $n = 20$. The conditional median delay (CMD) is, however, unaffected by the truncation. Notice that if $N_{\Pi^*} = J$ there is no delay in the warning signal and no false alarm.

Let PFA denote an estimate of the probability of false alarms; PES denote an estimate of the probability of exact stopping and CED the estimator (4.2). From Tables 4.3–4.5 we obtain the following estimates.

Table 4.6.
Estimates PFA, PES and CED for
$\mu_T = 0$, $\sigma^2 = \tau^2 = 1$, $p = 0.01$, $\Pi^* = 0.5$.

δ	1.5	2.0	2.5
PFA	0.13	0.18	0.17
PES	0.28	0.22	0.49
CED	2.540	2.098	0.964
CMD	1.00	1.00	0.00

As anticipated, there is no significant difference between the PFA values of the three δ's. The critical decision value Π^* plays the important role. The values of PES do depend on δ. There is no significant difference between the estimates of PES when $\delta = 1.5$ or $\delta = 2.0$. But, when $\delta = 2.5$ the value of PES increases significantly (about doubles). A similar picture is obtained on the CED. It is interesting to know how small could Π^* be without substantially effecting the CED. In additional simulation runs we kept $\delta = 2.0$ and decreased the value of Π^*. The results are presented in Table 4.7. 100 independent simulation runs were performed for each value of Π^*.

Table 4.7.
Estimates PFA, PES and CED for
$\mu_T = 0$, $\sigma^2 = \tau^2 = 1$, $p = 0.01$, $\delta = 2.0$.

Π^*	0.4	0.3	0.2
PFA	0.17	0.11	0.11
PES	0.34	0.35	0.25
CED	1.639	1.461	2.966

We see in Table 4.7 that, decreasing Π^* from 0.5 to 0.4 or 0.3 we obtain significantly higher PES values, the PFA values tend to decrease and so are the CED values. However, further decrease to $\Pi^* = 0.22$ increases CED. Thus, we recommend the use of $\Pi^* = 0.3$ in practical applications. Observe in Tables 4.3–4.5 that, there are cases in which $N_A = N_{\Pi^*} + 1$ when $N_{\Pi^*} < 20$. These are cases of stopping without a clear indication that a change took place. In addition, there are runs in which there was no stopping before $N = 20$. If we denote by PNS the proportion of runs without stopping ($N_{\Pi^*} \geq 20$) and by PFS the proportion of false stopping (stopping without an indication of a change) then we obtain from Tables 4.3–4.5 the estimates which are presented in Table 4.8.

Table 4.8.
Estimates of PNS and PFS for $\mu_T = 0$,
$\sigma^2 = \tau^2 = 1$, $p = 0.01$, $\Pi^* = 0.50$.

δ	PNS	PFS
1.5	0.15	0.13
2.0	0.07	0.13
2.5	0.03	0.07

4.3. Robustness Study of the AMOC Model with Respect to Choice of Parameters

Obviously, if the Bayesian model is valid, and one knows the prior parameters of the process, then the Bayesian estimators of the current mean, μ_n, and other detection indices are optimal. However, in reality, even if the assumed prior distributions are the correct ones, the prior parameters may not be exactly known. It is therefore interesting to study the effects of using wrong values of the parameters.

In the present section we study the sensitivity of the tracking procedure to deviations of the assumed values of p, σ^2, τ^2 and δ from the actual parameters of the process. For this purpose we performed a simulation experiment in which the assumed values of the prior parameters were varied around the true values of the process. These simulations were planned according to a 2^4 factorial experiment, in which ten independent replicas (runs) were performed of each treatment combination. The observations were simulated from sequences of $NS = 15$ random variables, with a fixed change point at $J = 10$. The process parameters used for simulating the random variables were: $\mu_T = 0$, $\sigma^2 = \tau^2 = 1$, $\delta = 2.5$. The levels of the four factors substituted in the detection formulae were: $p = 0.067, 0.100$; $\sigma^2 = 0.8, 1.2$; $\tau^2 = 0.8, 1.2$ and $\delta = 1.5, 3.5$.

If the probability of a shift in the future, $\Pi_{n,n+1}$ drops below $\Pi^* = 0.3$ we consider it as sufficient evidence to declare that a shift already occurred and "sound an alarm". We define the experiment's response variable in each one of the 16 treatment combinations as an angular transformation of the proportion of false alarms among the 10 replicas, i.e.,

$$Y_\nu = 2 \sin^{-1} \sqrt{\hat{p}_\nu},$$

where $\hat{p}_\nu$ is the observed proportion of false alarms among the 10 replicas of the ν-th treatment combination. The other experiment's response variable measures consistency in process level estimates and is the average squared distance of $\hat{\mu}_{nl}$ from μ_n over the 10 replicas, i.e.,

$$ASD_\nu = \frac{1}{10NS} \sum_{l=1}^{10} \sum_{n=1}^{NS} (\hat{\mu}_{nl} - \mu_n)^2,$$

where $\hat{\mu}_{nl}$ is the Bayes estimate of the current mean, μ_n, at the l-th replication. The values of Y_ν and ASD_ν are given in Table 4.9. Y_ν is approximately normal with mean $\eta_\nu = 2 \sin^{-1} \sqrt{p_\nu}$ and constant variance 0.1, where p_ν is the true proportion of false alarms. In all the 160 runs, an alarm was flagged when $\Pi_{n,n+1} < \Pi^* = 0.3$.

Table 4.9.
The observed values of Y_ν and ASD_ν for the 16 treatment combinations of the four factors.

p	σ^2	τ^2	δ	Y_ν	ASD_ν
0.067	0.8	0.8	1.5	1.1593	0.0099
0.100	0.8	0.8	1.5	1.1593	0.0091
0.067	1.2	0.8	1.5	1.3694	0.0159
0.100	1.2	0.8	1.5	1.5708	0.0183
0.067	0.8	1.2	1.5	0.9273	0.0162
0.100	0.8	1.2	1.5	1.1593	0.0124
0.067	1.2	1.2	1.5	1.3694	0.0111
0.100	1.2	1.2	1.5	1.1593	0.0163
0.067	0.8	0.8	3.5	0.9273	0.0136
0.100	0.8	0.8	3.5	0.6435	0.0200
0.067	1.2	0.8	3.5	1.1593	0.0115
0.100	1.2	0.8	3.5	0.9273	0.0066
0.067	0.8	1.2	3.5	0.9273	0.0085
0.100	0.8	1.2	3.5	0.9273	0.0124
0.067	1.2	1.2	3.5	1.1593	0.0190
0.100	1.2	1.2	3.5	0.9273	0.0129

Routine analysis of the experimental results yields that the main effects of σ^2 and δ are the largest effects on y_ν, and that the third order interaction, $\sigma^2 \times \tau^2 \times \delta$, has the largest effect on ASD$_\nu$. This implies that, deflating σ^2 and inflating δ will reduce the proportion of false alarms. The significant third order interaction of σ^2, τ^2 and δ can be exploited to produce Bayes estimates of the current mean with small ASD$_\nu$ values. For example, choosing an inflated (high) value of τ^2 yields ASD$_\nu \approx .012$, whatever the values of δ and σ^2 are. In all cases the procedure was found to be robust with respect to the choice of p.

5. DISCUSSION

There is a fundamental difference between sequential detection procedures and filtering procedures. The sequential detection procedures, like the CUSUM (non-Bayesian) or the Shiryayev-Roberts (Bayesian), are designed to monitor a process and flag as soon as sufficient evidence have accumulated that a change has occurred. A filtering procedure, on the other hand, is designed to estimate a signal (process mean level) in the presence of noise. The tracking algorithm developed and illustrated in the present paper is a filtering procedure. Unlike the Kalman filter it is a non-linear algorithm designed to detect sudden changes in the mean level of a stochastic process. The tracking algorithm provides, in addition to the fractiles (or other parameters) of the posterior distribution of the current mean, also the posterior probabilities of the change epoch, for the time interval under consideration. These posterior probabilities can be used, as illustrated in Section 4, for monitoring objectives. They provide estimates of the location of the change points.

Although our tracking algorithm is based on a model of at most one change, it is a filtering algorithm, and as such it can be applied with appropriate modification, also in the more general case of frequent or even continuous changes in the mean level. The examples presented in Section 4 fall in the class of processes in which the mean level is a

step function. The estimates provided by the tracking algorithm fluctuate around these constant levels, not completely at random due to the interdependence (autocorrelations) between the estimates $\mu_{.5}$, for different n values (Tables 4.1 and 4.2).
Numerical examples have shown that the estimates of μ_n obtained by the tracking algorithm are sometimes sensitive to the choice of the parameters μ_T, δ and p. In Section 4 we study with simulation experiments how to choose these parameters. We propose that, before implementing the tracking algorithm, one should try it on historical data and "fine tune" it by means of choosing the parameters which yield robust and reasonable estimates. Guidelines for such fine tuning are provided in the robustness study presented in Section 4.3. The chosen values of the parameters can then be used for tracking the future movements of the process.

REFERENCES

[1] CHERNOFF, H. AND ZACKS, S. (1964). Estimating the Current Mean of a Normal Distribution which Is Subjected to Changes in Time. *Ann. Math. Statist*, **35**, 999–1018.

[2] HOADLEY, B. (1981). The Quality Measurement Plan. *Bell System Technical Journal*, **60**, 215–271.

[3] KENETT, R. AND POLLAK, M. (1983). On Sequential Detection of a Shift in the Probability of a Rare Event. *J. Amer. Statist. Assoc.*, **78**, 389–395.

[4] KENETT, R. AND POLLAK, M. (1986). A Semi-Parametric Approach to Testing for Reliability Growth, With Application to Software Systems. *IEEE Transactions on Reliability*, **R-35**, 304–311.

[5] LUCAS, J.M. (1982). Combined Shewhart CUSUM Quality Control Schemes. *J. of Quality Technology*, **14**, 51–59.

[6] LUCAS, J.M. AND CROSIER, R.B. (1982). Fast Initial Response of CUSUM Quality-Control Schemes: Give Your CUSUM a Head Start. *Technometrics*, *24*, 199–205.

[7] PAGE, E.S. (1954). Continuous Inspection Schemes. *Biometrika*, **41**, 100–114.

[8] PHADKE, M.S. (1981). Quality Audit Using Adaptive Kalman Filtering. *ASCQ Quality Congress Trans.*, 1045–1052.

[9] POLLAK, M. (1985). Optimal Detection of a Change in Distribution. *Ann. Statistics*, **13**, 206–227.

[10] SHEWHART, W.A. (1931). *Economic Control of Quality of Manufactured Product.* D. Van Nostrand, Princeton, NJ.

[11] STEPHENS, K.S., WADSWORTH, H.M. AND GODFREY, A.B. (1986). *Modern Methods for Quality Control and Improvement.* John Wiley, New York.

[12] ZACKS, S. (1980). Distribution of Stopping Variables In Sequential Procedures For The Detection of Shifts In Distributions of Discrete Random Variables. *Communications In Statistics, B*, **9**, 1–18.

[13] ZACKS, S. (1981). The Probability Distribution and the Expected Value of a Stopping Variable Associated with One-Sided CUSUM Procedures for Non-Negative Integer Valued Random Variables, *Communications in Statistics, A*, **10**, 2245–2258.