Research Article
Change Point Detection with Robust Control Chart

Ng Kooi Huat[1] and Habshah Midi[1,2]

[1] *Laboratory of Applied and Computational Statistics, Institute for Mathematical Research,*
Universiti Putra Malaysia, Serdang, Selangor 43400, Malaysia
[2] *Department of Mathematics, Faculty of Sciences, Universiti Putra Malaysia, Serdang,*
Selangor 43400, Malaysia

Correspondence should be addressed to Ng Kooi Huat, ngkooihuat@inspem.upm.edu.my

Monitoring a process over time using a control chart allows quick detection of unusual states. In phase I, some historical process data, assumed to come from an in-control process, are used to construct the control limits. In Phase II, the process is monitored for an ongoing basis using control limits from Phase I. In Phase II, observations falling outside the control limits or unusual patterns of observations signal that the process has shifted from in-control process settings. Such signals trigger a search for assignable cause and, if the cause is found, corrective action will be implemented to prevent its recurrence. The purpose of this paper is to introduce a new methodology appropriate for constructing a robust control chart when a nonnormal or a contaminated data that may arise in phase I state. Through extensive Monte Carlo simulations, we examine the behaviors and performances of the proposed MM robust control chart when there is a process shift in mean.

1. Introduction

Statistical process control (SPC) concepts and methods have become very significant in the manufacturing and process industries. Their goal is to monitor the performance of a process over time in order to justify whether or not the process is remaining in a "state of statistical control." This state of control is said to occur if certain process or product variables remain near to their desired values and the only source of variation is "common-cause" variation, that is, the variation which affects the process all the time and is essentially inevitable within the current process. Shewhart charts are used to monitor key product variables in order to detect the occurrence of any event having a "special" or "assignable" cause. By discovering assignable causes, long-term improvements in the process and in product quality can be accomplished by eliminating the causes or improving the process or its operating procedures.

Detecting one or more change points in a batch of observations has attracted substantial investigation in the statistical, engineering, and econometric literature. Assuming that there is an ordered sequence of observations, usually, but not necessarily, taken at equally spaced times, there is a change point between two successive observations if their statistical distributions are different. Between change points, the distributions are usually considered to be identical. In practice, recognizing when a process has changed would simplify the search for the special cause. If the time of the change could be identified, process engineers would have a smaller search window within which to look for the special cause. Consequently, the special cause can be determined more quickly, and necessary actions needed to improve quality can be carried out sooner. In this paper, we will analyze the efficiency of a change point estimator in process mean [1] for each of Shewhart $\overline{X}$, Median, and the proposed MM control chart once issues a signal. The derivation of the change point estimator that is shown in the appendix is in virtue of Hinkley [2]. Hinkley discussed the asymptotic properties of the estimator. Whenever each of Shewhart $\overline{X}$, Median, or the proposed MM chart signals that a special cause is present. The estimator provides practitioners with a useful estimate of the time of the process change. In Section 2, we will introduce a model for a step change in the location of a process. We consider a step change for a process mean occurs when the mean suddenly changes its value and then remains unchanged again until corrective action has been taken. On the basis of this step-change model, we adopt the estimator of the time of the process change when the corresponding chart does signal. In Section 5, we analyze the performances of each chart by means of Monte Carlo simulation.

2. Process Step-Change Model

Suppose that the process is initially in control, with observations coming from a Normal distribution with a known mean of μ_0 and a known standard deviation of σ_0. Even so, after an unknown point in time T (known as the process change point), the process location changes from μ_0 to $\mu_1 = \mu_0 + \delta\sigma_0/\sqrt{n}$, where n is the subgroup size and δ is the unknown magnitude of the change. Assuming also that once this step change in the process location occurs, the process remains at the new level of μ_1, until the special cause has been identified and removed. We let $\overline{X}_T$ be the first subgroup average to exceed a control limit and that this signal is not a false alarm. Hence, $\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_\tau$ are the subgroup averages that come from the in-control process, while $\overline{X}_{\tau+1}, \overline{X}_{\tau+2}, \ldots, \overline{X}_T$ are from the changed process.

3. Definitions

To illustrate, we concentrate on robust estimates for the simple location-scale model, by letting $x_1, \ldots, x_n$ *be n* observations on the real line satisfying

$$x_i = \mu + \sigma\varepsilon_i, \quad i = 1, \ldots, n, \tag{3.1}$$

where ε_i are independent and identically distributed observations with variance equal to 1. We are interested in estimating μ and the scale σ which is a nuisance parameter. We consider

M-location estimates which is proposed by Huber [3]. He defined $\hat{\mu}_n$ as the solution of an estimating equation of the form

$$\sum_{i=1}^{n} \psi\left(\frac{(x_i - \hat{\mu}_n)}{\hat{\sigma}_n}\right) = 0, \tag{3.2}$$

where $\hat{\sigma}_n$ is a robust estimate of the residuals scale, and $\psi : R \rightarrow R$ is a bounded, nondecreasing, and odd real function. We focus on ψ function which is continuous and differentiable influence function given in (3.2) (see [4])

$$\psi_c(u) = \text{sign}(u) \begin{cases} \left|\dfrac{u}{c}\right|, & \text{if } |u| \leq 0.8c, \\ p_4\left(\left|\dfrac{u}{c}\right|\right), & \text{if } 0.8c < |u| \leq c, \\ p_4(1), & \text{if } |u| \geq c, \end{cases} \tag{3.3}$$

where $c > 0$ is a user-chosen tuning constant, and $p_4(u) = 38.4 - 175u + 300u^2 - 225u^3 + 62.5u^4$ (see [5]), for other choices of smooth functions ψ.

The scale estimate $\hat{\sigma}_n$ in (3.2) is an S-estimate of scale (see [6]) which is defined as follows. Let $\rho : R \rightarrow R_+$ be a bounded, continuous, and even function satisfying $\rho(0) = 0$ and let $b \in (0,1)$. The S-scale $\hat{\sigma}_n$ is defined by

$$\hat{\sigma}_n = \inf_{t \in R} s_n(t), \tag{3.4}$$

where, for each $t \in R$, $s_n(t)$ is the solution of

$$\frac{1}{n}\sum_{i=1}^{n} \rho\left(\frac{(x_i - t)}{s_n(t)}\right) = b. \tag{3.5}$$

Indeed, associated with this family are the *S-location* estimates $\tilde{\mu}_n$ is given by

$$\tilde{\mu}_n = \arg \inf_{t \in R} s_n(t). \tag{3.6}$$

Beaton and Tukey [7] proposed a family of functions ρ_d given by

$$\rho_d(u) = \begin{cases} 3\left(\dfrac{u}{d}\right)^2 - 3\left(\dfrac{u}{d}\right)^4 + \left(\dfrac{u}{d}\right)^6, & \text{if } |u| \leq d, \\ 1, & \text{if } |u| > d, \end{cases} \tag{3.7}$$

where the tuning constant d is positive. According to Yohai [8], these M-location estimates obtained with an S-scale estimate are called *MM-location* estimates. Specifically, the estimates $\widehat{\mu}_n$, $\widehat{\sigma}_n$, and $\widetilde{\mu}_n$ solve the following system of equations:

$$
\sum_{i=1}^{n} \psi\left(\frac{(x_i - \widehat{\mu}_n)}{\widehat{\sigma}_n}\right) = 0,
$$

$$
\frac{1}{n} \sum_{i=1}^{n} \left[\rho\left(\frac{(x_i - \widetilde{\mu}_n)}{\widehat{\sigma}_n}\right) - b\right] = 0, \tag{3.8}
$$

$$
\frac{1}{n} \sum_{i=1}^{n} \rho'\left(\frac{(x_i - \widetilde{\mu}_n)}{\widehat{\sigma}_n}\right) = 0.
$$

Let $d = 1.548$ for ρ_d in (3.7), $b = 0.5$ in (3.5), and $c = 1.525$ for ψ_c in (3.3), which yields a location estimate $\widehat{\mu}_n$ with 50% breakdown point and 95% efficiency when the errors have a normal distribution.

3.1. Sample Median

Sample median has been used in early process control charts as it is insensitive to behavior in the tails of the distribution. However, under the normal distribution, the efficiency of the sample median drops off rapidly towards its asymptotic value of 0.64 as sample size increases. For a random sample of size n observations $X_1, X_2, \ldots, X_n$, the sample median, denoted by MD, is defined as follows:

$$
\text{MD} = \begin{cases} X_{((n+1)/2)}, & \text{if } n \text{ is odd,} \\ \dfrac{X_{(n/2)} + X_{((n/2)+1)}}{2}, & \text{if } n \text{ is even.} \end{cases} \tag{3.9}
$$

The interest of using the sample median, MD, is that it is easy to determine, requires only the middle values to calculate, can be used when a distribution is skewed, is not affected by outliers, and has a maximal 50 percent breakdown point. Moreover, its gross-error sensitivity is low and as the sample size n increases, the variance of the MD decreases as $1/n$, but the maximum bias does not change. Hence, the bias is the property of importance for large sample sizes, and MD is the estimator possesses the smallest maximum bias for a given proportion of contamination ε. Huber [3] showed that it minimizes the maximum asymptotic bias over contamination neighborhoods. As opposed to that, the disadvantages for the sample median, MD, are its difficulty to handle in mathematical equations, nonutilizing all available values, and being misleading when the distributions come from a long tail distribution as it might sometimes discard some useful information (see [9, 10]). However, the sample median has become as a good general purpose estimator and is generally considered as an alternative average to the sample mean especially whenever outliers might present in the distribution.

3.2. Median Absolute Deviation from the Sample Median

The median absolute deviation from the sample median, denoted as MAD, is a more robust scale estimator than the standard deviation. The MAD was first introduced by Hampel [11] who attributed it to Gauss. It is simple and easy to compute and mainly used in detecting outliers in a data. The estimate is often used as an initial value for the computation of more efficient robust estimators. Let us denote $X_1, X_2, \ldots, X_n$ as a random sample of size n observations with sample median MD. MAD possesses the following properties:

(i) it has a maximal 50 percent breakdown point which is twice as the IQR;

(ii) in the case of the standard normal distribution, F, the influence function of the MAD estimator, $IF(X; \text{MAD}, \Phi)$, is a step function that takes on two values. This IF is bounded by the sharpest possible bound among all scale estimators. With regard to the optimality properties of MAD, Martin and Zamar [12] established expressions for the maximum asymptotic bias of M-estimates of scale over contamination neighborhood as a function of the fraction of contamination and show that the similar strong results are obtained in terms of maximum asymptotic bias for MAD as with the MD.

3.3. Control Limits

In this paper, we consider the general equations for constructing control limits (see [13]). Thus, with a robust location estimator T and the corresponding scale estimator S, the control limits are given by

$$\text{UCL} = T + \frac{3S}{A\sqrt{n}},$$

$$\text{CL} = T, \tag{3.10}$$

$$\text{LCL} = T - \frac{3S}{A\sqrt{n}}.$$

The constant A in (3.10) is determined in such a way that S/A is an unbiased estimator of the scale parameter. The most commonly used control charts are Shewhart $\overline{X}$ charts using the sample range. For $\overline{X}$ charts using the sample standard deviation, the T in (3.10) is the sample mean $\overline{X}$ and S is the sample standard deviation with $A = c_4$.

Shewhart modified these control limits using rational subgroups (see [14]) in which m rational subgroups with each of size n are taken. According to Shewhart's suggestion, these subgroups are formed in order that the between-groups variability is maximized while the within-group variation is minimized. In this view, then

$$\overline{\overline{X}} = \frac{1}{m} \sum_{i=1}^{m} \overline{X}_i, \qquad \overline{S} = \frac{1}{m} \sum_{i=1}^{m} S_i, \tag{3.11}$$

where $\overline{X}_i$ and S_i are the subgroup mean and standard deviation, respectively. Each of these estimates is an unbiased estimate of the corresponding parameter, then the control limits using rational subgroups are

$$
\begin{aligned}
\text{UCL} &= T + \frac{3\overline{S}}{A\sqrt{n}}, \\
\text{CL} &= T, \\
\text{LCL} &= T - \frac{3\overline{S}}{A\sqrt{n}}.
\end{aligned}
\tag{3.12}
$$

In practice, the control limits are the average of the control limits for the m subgroups. In the case of $\overline{X}$ chart for which we employ the sample range as the scale parameter, it is estimated with the average range computed by averaging over the m subgroups

$$
\begin{aligned}
T &= \overline{\overline{X}} = \frac{1}{m}\sum_{i=1}^{m}\overline{X}_i, \\
S &= \overline{R} = \frac{1}{m}\sum_{i=1}^{m}R_i.
\end{aligned}
\tag{3.13}
$$

Then, the control limits are defined as

$$
\begin{aligned}
\text{UCL} &= \overline{\overline{X}} + \frac{3\overline{R}}{A\sqrt{n}}, \\
\text{CL} &= \overline{\overline{X}}, \\
\text{LCL} &= \overline{\overline{X}} - \frac{3\overline{R}}{A\sqrt{n}}.
\end{aligned}
\tag{3.14}
$$

For constructing the control charts under a normal distribution using the robust estimators, we will determine the appropriate constant A for the desired estimators through computer simulations. To illustrate, a sample of size n was taken from $N(0, 1)$. The constant A was computed by averaging over 100,000 repetitions. Here, for instance, if we consider using S as scale estimator, then over 100,000 repetitions, we expect $E(S/A) = \sigma$ for any σ. Table 1 exhibits the constant A such that $E(scale\ estimator/A) = \sigma$. It can be seen that if scale estimator is the sample range, the simulated values (E_1) agree closely with the standard tabled values. E_2 and E_3 are the corresponding estimates for the Median and the proposed MM charts.

Thus, in the same way, the estimators for location and scale for Median-MAD chart are given by

$$
\begin{aligned}
T &= \text{Median}(x_i), \\
S &= \text{MAD} = \text{Median}\{|x_i - T|\},
\end{aligned}
\tag{3.15}
$$

Table 1: Simulated values of A for different scale estimators.

Estimator	$n = 5$	$n = 10$
E_1	2.3275	3.0563
E_2	0.8255	0.9123
E_3	0.7677	0.8897

while the estimators for the proposed MM chart are given by

$$T = \tilde{\mu}_n = \arg \inf_{t \in R} s_n(t),$$
$$S = \hat{\sigma}_n = \inf_{t \in R} s_n(t). \tag{3.16}$$

In this study, three different estimators under investigation are as follows:

E_1: T = Sample Mean; S = Range,
E_2: T = Median; S = MAD = Median$\{|x_i - T|\}$,
E_3: $T = \tilde{\mu}_n = \arg \inf_{t \in R} s_n(t)$; $S = \hat{\sigma}_n = \inf_{t \in R} s_n(t)$.

3.4. Confidence Regions (Confidence Set)

One of the benchmarks for assessing the performance of a control chart is to construct a confidence region for the time of the process change. The use of confidence region on the change point is that it will suggest practitioners with useful starting points for searching their process log books and records for the special cause. This will provide the practitioners a "search window" for the special cause and aid in quicker identification. Hence, practitioners can then take necessary action for the special cause sooner in order to improve quality as well as to reduce process downtime.

Basically, we will incorporate the likelihood function to obtain a confidence region for the process change point. The confidence region approaches in the statistics literature involve the likelihood function that relies on asymptotic theory (see [15]). In process monitoring, there are relatively small time intervals between the process change point and the time of the control chart signal. Thus, approximations based on asymptotic theory may not be appropriate.

Box and Cox [16] proposed a method involving the log likelihood function for constructing a possibly noncontiguous confidence region (also called a confidence set) on a parameter. Their approach can be used to build a confidence set (CS) for the process change point using the log likelihood function having the form

$$\text{CS} = \{t : \ln L(\hat{\tau}) - \ln L(\tau) < D\}. \tag{3.17}$$

Here, $\ln L(\hat{\tau})$ is the maximum value of the log likelihood function; $\hat{\tau}$ is the MLE of τ (i.e., the value of t that maximizes the log likelihood function), where

$$\ln L(t) = -T \ln\left(\sqrt{2\pi}\sigma_x\right) - \frac{n}{2\sigma_0^2}\left[\sum_{i=1}^{T} X_i^2 - 2\mu_0 \sum_{i=1}^{t} \overline{X}_i + t\mu_0^2 - (T-t)\left(\overline{\overline{X}}_{T,\,t}\right)^2\right] \tag{3.18}$$

is the value of the log likelihood function at t. We let k_1 and k_2 represent constants determined by the subgroup averages. Thus, $\ln L(t)$ can be expressed as

$$\ln L(t) = k_1 - \frac{n}{2\sigma_0^2}\left[k_2 - (T - t)\left(\overline{\overline{X}}_{T,\tau} - \mu_0\right)^2\right]. \tag{3.19}$$

Box and Cox [16] proposed using $D = (1/2)\chi_{1,\alpha}^2$ to obtain a $100(1 - \alpha)\%$ confidence region. Siegmund [17] used asymptotic theory to develop a $100(1 - \alpha)\%$ confidence set for the change point of a normal process mean based on the log likelihood function. He proposed using the value

$$D = -\ln\left[1 - (1 - \alpha)^{1/2}\right]. \tag{3.20}$$

By means of Monte Carlo simulation, we study the confidence sets obtained with nominal confidence coefficient of $1 - \alpha = 0.90$. In accordance with Box and Cox [16] and Siegmund [17], the D values for a 90% confidence set are $D = 1.353$ and $D = 2.97$, respectively. It was observed that $D = 1.353$ value suggested by Box and Cox [16] provides a 90% coverage for a value of δ between 2.0 and 3.0, while Siegmund's [17] of $D = 2.97$ provides at least 90% coverage for $\delta \geq 1.0$. By trial and error, the value of $D = 3.7$ provides at least 90% coverage for $\delta \geq 0.5$.

4. Methodology

In order to analyze the performance of the control charts, we consider using Shewhart $\overline{X}$, Median, and our proposed MM control chart. When a control chart signals that suggest a process change has occurred, the change point estimator (see the appendix) is then applied to the data to estimate the time of the change at which we need to find the value of t in the range $0 \leq t < T$ which maximizes $C_t = (T - \tau)(\overline{\overline{X}}_{T,t} - \mu_0)^2$. The reverse cumulative average, $\overline{\overline{X}}_{T,\tau} = (T - \tau)^{-1}\sum_{i=\tau+1}^{T}\overline{X}_i$, is the overall average of the $T - t$ most recent subgroups for which the value of t maximizing the C_t values is our estimator of the last subgroup from in-control process.

5. Simulation Study

We will now analyze the performance of the change point estimator adopting the three control charts through the simulation study. Assuming that the process is initially in control, with observations coming from a normal distribution with a known mean of μ_0 and a known standard deviation of σ_0. However, after an unknown point in time τ (known as the process change point), the process location changes from μ_0 to $\mu_1 = \mu_0 \pm \delta\sigma_0/\sqrt{n}$, where δ is the unknown magnitude of the change. We also assume that once this step change in the process location occurs, the process remains at the new level of μ_1, until the special cause has been identified and removed.

To illustrate, the data are generated under different settings of distributions. In addition to the normal distribution, two alternative distributional forms are considered.

They are contaminated model (Case 2 of (5.2)) and Slash distribution. Under different types of distributions, for each run, the data consist of $m = 30$ subgroups of size $n = 5$ are used to construct the control limits and summary statistics are calculated. In order to assess the performance of the corresponding chart, observations for 1 to 100 are generated from standard normal distribution. Then, starting from subgroup 101, observations were randomly generated from a normal distribution with mean δ and standard deviation 1 until each of Shewhart $\overline{X}$, Median, and the proposed MM control chart produces a signal. The procedure was repeated a total of 10,000 times for each of the values of magnitude that was studied, namely $\delta = 1.0, 2.0$, and 3.0. For each simulation run, the change point estimate was computed. Subsequently, the average of the estimates of $\overline{\tau}$ for the 10,000 simulation runs was computed along with its standard error, expected length, and coverage probability.

For analyzing the outlier model, we modified the contaminated model by Davis and Adams [18]. Specifically, the in-control conditions with contaminated data values are determined by generating a random number from a Uniform $(0, 1)$ distribution and the corresponding frequency of contaminated data, $0 \leq \beta \leq 1$. A Uniform $(0, 1)$ random probabilistic value p_{ij} is generated for observations j of sample i, x_{ij}. Let $I_{(a,b)}(p)$ represent an indicator function with

$$I_{(a,b)}(p) = \begin{cases} 1, & \text{if } a < p < b, \\ 0, & \text{otherwise.} \end{cases} \tag{5.1}$$

A random observation x_{ij} in the simulated data is described by (5.2). Different changes for each of process states are illustrated in the expression. For the purpose of comparisons, the frequencies of contamination of $\beta = 0.05$ and $\beta = 0.10$ with $C = 9.0$ are considered for the purpose of creating some disturbances in the data

$$x_{ij} \sim N\left(\mu + C, \sigma^2\right) I_{(0,\beta)}(p_{ij}) + N\left(\mu, \sigma^2\right) I_{[\beta,1)}(p_{ij}), \quad \text{for } i = 1, 2, \ldots, 10,000,$$
$$j = 1, 2, \ldots, n. \tag{5.2}$$

Case 1. In-control (no contaminated data): $\mu = 0, \sigma^2 = 1, \beta = 0, C = 0$.

Case 2. In-control (contaminated data): $\mu = 0, \sigma^2 = 1, \beta = 0.05, 0.10, C = 9.0$.

As with the computation of sizes of confidence sets obtained using a specific value of D, for each control chart and magnitude of change studied, a step change in the normal process mean was simulated following $\tau = 100$. The confidence set estimator was applied following a signal from the corresponding control charts considered. The size of the confidence set was recorded as well as whether the confidence set covered the true process change point of $\tau = 100$. This procedure was repeated for a total of $N = 10,000$ simulation runs for each of values δ considered. The proportion of the 10,000 runs that covered the true process change point was also determined. This was reflected by the resulting estimates of the coverage probabilities obtained by specifying the D value to be 2.97, such that it provides at least 90% coverage for $\delta \geq 1.0$, which has been discussed in Section 4. The results are tabulated along with the average sizes of confidence sets. For a given coverage probability, a smaller confidence set is preferred, so that process engineers can more narrowly focus their

search for the special cause. In general, it is presumed that the increase in the magnitude of shift will be followed by the increase in the corresponding coverage probability.

From Tables 2, 3, 4, and 5, we can see that the performances of the three charts in terms of coverage probability are quite similar especially when the process data come from a normal distribution. Generally, the Median chart and the proposed MM chart perform better in most of the cases (particularly for a larger proportion of contamination). For instance, for a change of magnitude $\delta = 3.0$, with $\beta = 0.05$, $C = 9.0$, the coverage probability utilizing $\overline{X}$ chart is 0.588, while the coverage probability for Median chart and the proposed MM chart are 0.626 and 0.632, respectively.

In Tables 2–5, averages of change point estimates $\overline{\hat{\tau}}$ are also tabulated for various sizes of change in the process mean together with its corresponding standard error estimates for a normal case setting. As the actual change point for the simulation was at time 100, the average estimated time of the process change, $\overline{\hat{\tau}}$, should possibly be close to 100. With $\overline{X}$ chart, we see that when the process step change of standardized magnitude $\delta = 1$, the average estimated time of the process change was 100.00, which is fairly close to the actual change point of 100. While for a standardized process location change of size $\delta = 2$, the average estimated time of the change is 99.60. Meanwhile, when $\delta = 3$, the average estimated time of the change is 99.47. Hence, on average, the change point estimate of the time of the process change is considerably close to the actual time of the change, regardless of the magnitude of the change.

By the same taken, with the Median chart, for the process step change of standardized magnitude $\delta = 1$, the average estimated time of the process change was 99.70, which is also close to the actual change point of 100.00. As for a standardized process location change of size $\delta = 2$, the average estimated time of the change is 99.49. And when $\delta = 3$, the average estimated time of the change is 99.57.

Lastly, when the process is monitored with our proposed robust MM chart, by and large, the change point estimate of the time of the process change is fairly close to the actual time of the change, regardless of the magnitude of the change. In the case when the process step change of standardized magnitude $\delta = 1$, the average estimated time of the process change was 99.90. It turns out that for a standardized process location change of size $\delta = 2$, the average estimated time of the change is 99.42. As with $\delta = 3$, the average estimated time of the change is 99.59. Overall, we could say that for all types of charts under study, the change point estimator of the time of the process change is able to detect the change point considerably close to the actual time of the change, irrespective of the magnitude of the change.

Another benchmark of evaluating the control chart is by examining the expected length of the signal. This is the expected time at which the control chart signals a change in the process mean that is supposed to occur at time 100. It is generally perceived that Shewhart $\overline{X}$ control chart might issue a signal of a change in a process mean a considerable amount of time after the change in the process mean actually occurred. Thus, estimating the time of process change with the time when the control chart indeed issues a signal would lead to an unfavorably biased estimate. As a consequence, probably a misleading estimate of the time of the process changes. This bias is in virtue of the potentially large delay in generating a signal from the control chart. Hence, the criterion for evaluating the performance of the control chart is how quick the chart would signal (Expected Length). Tables 2–5 and Figures 1–4 summarize the performance in terms of expected length for the three charts.

In a normal distribution situation, for a step change in the process mean of magnitude $\delta = 1$, it is easy to see from **Figure 1** that the expected length for a Shewhart $\overline{X}$ chart is 57.

The result seems to suggest that $\overline{X}$ chart is the best compared to the Median chart and the proposed MM chart for which each needs 157 and 98, respectively, when the shift is small. The situation improves as the magnitude of shift δ increases to 2 or 3. All types of charts considered here seem to be relatively comparable and perform quite closely.

On the other hand, the Shewhart $\overline{X}$ chart is inferior in an outlier model. With β = 0.10 and C = 9, according to Table 4, both the Median chart and the proposed MM chart appear to be better than Shewhart $\overline{X}$ chart for different magnitudes of shifts. The similar situation arises when a very heavy-tailed distribution (Slash distribution) is considered. Again, the Median chart and the proposed MM chart outperformed Shewhart $\overline{X}$ chart with respect to the expected length. It can be observed from Table 5 that the differences are quite apparent, whereby Shewhart $\overline{X}$ chart requires the expected length of 24 and yet the Median chart and the proposed MM chart both demand about 3 subgroups before detecting the first signal when the magnitude of shift is 1.

We now turn to evaluate the observed frequency in which the estimates of the time of the step were within m observations of the actual time of the change, for $m = 0, 1, 2, \ldots, 10$. The results are tabulated in Tables 6, 7, 8, 9, 10, 11, 12, 13, and 14. This provides an indication of the precision of the estimator by means of the three different charts. The proportion of the 10,000 runs where the estimated time of the change was within $\pm m$ of the actual change is expected to be increase in size as m increases. Referring to Tables 6–14, we observed that the precision increases with the increases of m for each δ value. Let us first focus our attention to a normal setting, when the process step change of magnitude δ = 2. Monitoring the process using traditional Shewhart $\overline{X}$ chart identified correctly the change point in 60.49% of the trials. It was within one observation of the actual change point in 83.23% of the trials, and within two observations of the actual change point in 91.39% of the trials. Turning to the Median chart, which is shown in Table 7, the chart detected correctly the change point in 59.59% of the trials. It was within one observation of the actual change point in 83.46% of the trials, and within two observations of the actual change point in 91.66% of the trials. It then follows that, based on Table 8, our proposed MM chart located accurately the change point in 59.61% of the trials. It was within one observation of the actual change point in 82.88% of the trials, and within two observations of the actual change point in 91.25% of the trials. All types of charts considered here seem to be comparable and performed quite equally.

Next, we observe the situation under outlier model setting, with β = 0.10 and C = 9. Consider again when the process step change of magnitude is δ = 2, for step change of this magnitude, the Shewhart $\overline{X}$ chart estimator exactly identified the time of the change in just 23.65% of the trials and was within one (two) observation of the time of the actual process change in 45.03% (56.69%) of the trials. As for the Median chart, we can notice that, of the 10,000 simulation trials conducted for δ = 2, 24.99% of those simulation trials identified the change point precisely. It was in 46.92% and 58.11% of the trials that the change point was estimated to be within ± 1 and ± 2, respectively, from the actual time of the process change. The results of the study also indicate that 25.25% of those simulation trials identified the change point correctly for the proposed MM chart. In 47.53% of the trials, the estimate was within ± 1 observation, and in 58.61% of the trials, the estimate was within ± 2 observations. Overall, the procedure using the proposed MM chart seems to perform slightly better than the other two charts in this respect.

Finally, we focus on nonnormal data without outliers' situation, the Median chart and the proposed MM chart generally lead to shorter control limits than the traditional Shewhart
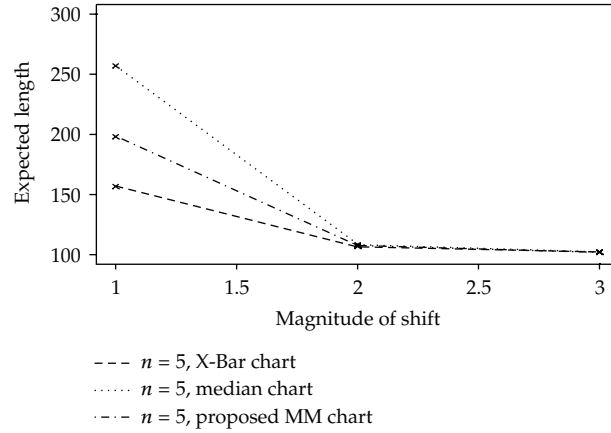
Figure 1: Simulation result: expected length for change point of $\tau = 100$ (normal distribution).

Table 2: Simulation result: estimates of expected length, average change point, standard error, average size of confidence set and coverage probability for change point of $\tau = 100$ (normal distribution).

	X-Bar chart			Median chart			Proposed MM chart		
	Magnitude of shift, δ			Magnitude of shift, δ			Magnitude of shift, δ		
	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$
Expected length, $E(T)$	156.75	106.88	102.05	257.05	108.30	102.16	198.14	107.60	102.14
$\bar{\hat{\tau}}$	100.00	99.60	99.47	99.70	99.49	99.57	99.90	99.42	99.59
Standard error $(\bar{\hat{\tau}})$	3.55	2.04	2.43	4.29	2.60	2.18	3.70	2.76	2.16
Average size, CS	14.00	5.90	4.08	16.20	6.69	4.13	15.80	6.13	4.29
Coverage probability	0.897	0.959	0.984	0.901	0.953	0.979	0.901	0.956	0.981

$\overline{X}$ chart. Here, we just limit ourselves to the study of Slash distribution (see [13]) which is a very heavy-tailed distribution. When the process step change of magnitude $\delta = 2$, monitoring the process using traditional $\overline{X}$ chart identified correctly the change point in 5.31% of the trials. It was within one observation of the actual change point in 14.44% of the trials, and within two observations of the actual change point in 20.64% of the trials. For the Median chart which is shown by Table 13, the chart was able to detect correctly the change point in 5.45% of the trials. It was within one observation of the actual change point in 12.08% of the trials, and within two observations of the actual change point in 17.95% of the trials. It then follows that, based on Table 14, our proposed MM chart discovered accurately the change point in 5.49% of the trials. It was within one observation of the actual change point in 12.28% of the trials, and within two observations of the actual change point in 18.02% of the trials. On the whole, we would say that the Median, and the proposed MM charts perform better and more consistently than the $\overline{X}$ chart in this setting.

6. Conclusions

Control charts are used to detect whether or not a process has changed. When a control chart signals indicate that a process has changed, practitioners must initiate a search for the special cause. However, given a signal from a control chart, practitioners generally do not know
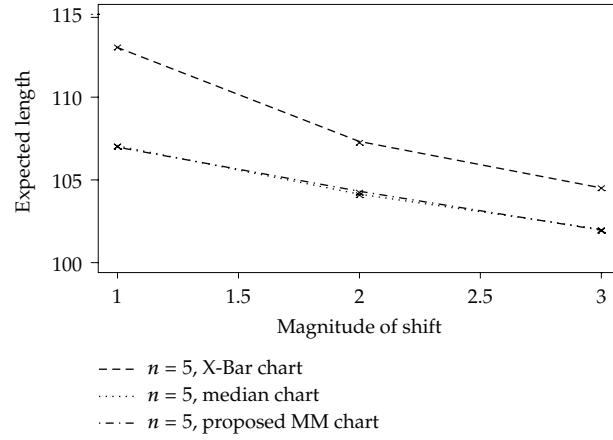
Figure 2: Simulation result: expected length for change point of $\tau = 100$ (contaminated distribution, $\beta = 0.05$, $C = 9$).
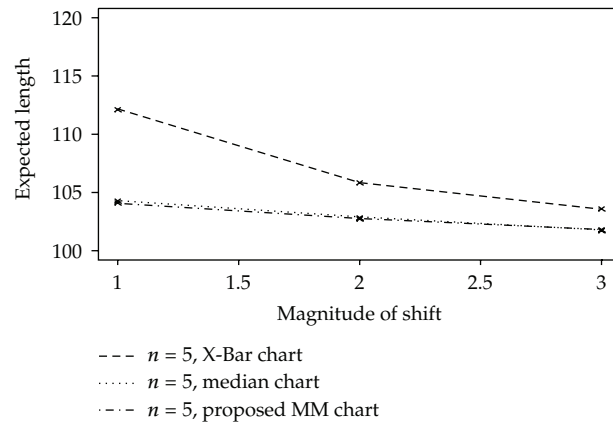

Figure 3: Simulation result: expected length for change point of $\tau = 100$ (contaminated distribution, $\beta = 0.10$, $C = 9$).
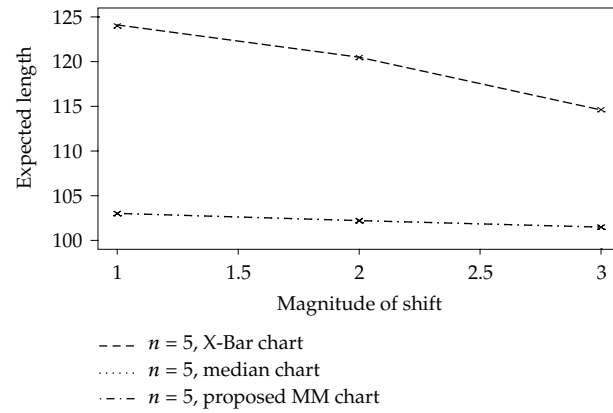

Figure 4: Simulation result: expected length for change point of $\tau = 100$ (slash distribution).

Table 3: Simulation result: estimates of expected length, average change point, standard error, average size of confidence set, and coverage probability for change point of $\tau = 100$ (contaminated distribution, $\beta = 0.05$, $C = 9$).

	X-Bar chart			Median chart			Proposed MM chart		
	Magnitude of shift, δ			Magnitude of shift, δ			Magnitude of shift, δ		
	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$
Expected length, $E(T)$	113.00	107.24	104.50	107.00	104.08	101.92	107.00	104.23	101.97
$\overline{\hat{\tau}}$	105.00	100.11	98.89	96.70	95.17	91.61	97.00	95.93	91.98
Standard error ($\overline{\hat{\tau}}$)	7.14	5.21	4.92	9.54	8.03	9.61	9.48	7.42	9.47
Average size, CS	2.57	2.20	2.06	4.47	4.10	4.54	4.17	3.81	4.24
Coverage probability	0.179	0.438	0.588	0.248	0.476	0.626	0.244	0.481	0.632

Table 4: Simulation result: estimates of expected length, average change point, standard error, average size of confidence set, and coverage probability for change point of $\tau = 100$ (contaminated distribution, $\beta = 0.10$, $C = 9$).

	X-Bar chart			Median chart			Proposed MM chart		
	Magnitude of shift, δ			Magnitude of shift, δ			Magnitude of shift, δ		
	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$
Expected length, $E(T)$	112.11	105.84	103.60	104.11	102.73	101.72	104.27	102.83	101.81
$\overline{\hat{\tau}}$	102.53	99.61	99.72	91.96	92.76	91.89	92.78	93.49	92.44
Standard error ($\overline{\hat{\tau}}$)	8.72	5.24	3.24	10.71	9.12	9.10	10.37	8.80	8.84
Average size, CS	1.82	1.53	1.39	2.96	2.63	2.62	2.79	2.51	2.53
Coverage probability	0.131	0.311	0.470	0.186	0.339	0.465	0.185	0.340	0.473

what caused the process situation to change or when the process has changed. Identifying the time of the process change would simplify the seeking of the special cause. In the event that the practitioners knew when the process changed, the search would simply be reduced for discovering what aspect of the process changed at that time. As a result, practitioners would increase their chances of identifying the special cause more correctly and quickly. Subsequently, This allows them to take the appropriate actions immediately to improve the quality.

In this paper, monitoring processes in the presence of data contamination and under nonnormal setting are of primary concern. We have applied an estimator that is useful for identifying the change point of a step change in normal process mean, nonnormal, and when contamination may exist. We have discussed the performance of the change point estimator and other criteria when they are monitored by Shewhart $\overline{X}$, the Median and the proposed MM control charts. The results show that the proposed MM robust control chart consistently performed well in a range of situations. It provides a useful and much better alternative in using the time of the signal from the conventional $\overline{X}$ control chart. Although the proposed MM chart and the Median chart are comparable under nonnormal and contamination situation; the Median chart becomes worse in the event of normal setting. The performance of our proposed MM control chart has good properties in the aspect of expected length and coverage probability for contamination data and that arise from a heavy-tailed distribution functions for moderate sample sizes. The proposed MM chart compares favorably with traditional Shewhart $\overline{X}$ control chart in normal setting especially when magnitudes of shift are 2 and 3. It is interesting to note that the proposed robust MM chart is more efficient than the Shewhart $\overline{X}$ chart when the process distribution function has a heavy-tailed distribution.

Table 5: Simulation result: estimates of expected length, average change point, standard error, average size of confidence, set and coverage probability for change point of $\tau = 100$ (slash distribution).

	X-Bar chart			Median chart			Proposed MM chart		
	Magnitude of shift, δ			Magnitude of shift, δ			Magnitude of shift, δ		
	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$
Expected length, $E(T)$	124.00	120.46	114.63	103.00	102.17	101.46	103.00	102.24	101.50
$\bar{\hat{\tau}}$	99.40	97.84	92.46	65.40	64.29	64.74	66.00	64.75	64.85
Standard error $(\bar{\hat{\tau}})$	37.40	42.48	28.23	14.90	14.78	14.61	14.90	14.81	14.58
Average size, CS	1.01	1.01	1.02	1.08	1.07	1.07	1.07	1.06	1.07
Coverage probability	0.029	0.056	0.083	0.040	0.055	0.076	0.040	0.056	0.079

Table 6: Precision of the estimator when used with X-Bar chart for different magnitudes of process change (magnitude of shift, δ) based on 10,000 trials. Subgroup size $n = 5$, change point $\tau = 100$ (normal distribution).

	Magnitude of shift, δ		
	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$
$\widehat{P}(\hat{\tau} = \tau)$	0.2660	0.6049	0.8168
$\widehat{P}(|\hat{\tau} - \tau| \leq 1)$	0.4790	0.8323	0.9393
$\widehat{P}(|\hat{\tau} - \tau| \leq 2)$	0.6230	0.9139	0.9734
$\widehat{P}(|\hat{\tau} - \tau| \leq 3)$	0.7080	0.9537	0.9834
$\widehat{P}(|\hat{\tau} - \tau| \leq 4)$	0.7610	0.9720	0.9843
$\widehat{P}(|\hat{\tau} - \tau| \leq 5)$	0.7990	0.9783	0.9893
$\widehat{P}(|\hat{\tau} - \tau| \leq 6)$	0.8260	0.9811	0.9900
$\widehat{P}(|\hat{\tau} - \tau| \leq 7)$	0.8580	0.9831	0.9909
$\widehat{P}(|\hat{\tau} - \tau| \leq 8)$	0.8790	0.9848	0.9909
$\widehat{P}(|\hat{\tau} - \tau| \leq 9)$	0.9100	0.9877	0.9915
$\widehat{P}(|\hat{\tau} - \tau| \leq 10)$	0.9220	0.9899	0.9915

Table 7: Precision of the estimator when used with median chart for different magnitudes of process change (magnitude of shift, δ) based on 10,000 trials. Subgroup size $n = 5$, change point $\tau = 100$ (normal distribution).

	Magnitude of shift, δ		
	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$
$\widehat{P}(\hat{\tau} = \tau)$	0.2520	0.5959	0.8164
$\widehat{P}(|\hat{\tau} - \tau| \leq 1)$	0.4720	0.8346	0.9412
$\widehat{P}(|\hat{\tau} - \tau| \leq 2)$	0.6120	0.9166	0.9769
$\widehat{P}(|\hat{\tau} - \tau| \leq 3)$	0.6960	0.9557	0.9870
$\widehat{P}(|\hat{\tau} - \tau| \leq 4)$	0.7510	0.9722	0.9882
$\widehat{P}(|\hat{\tau} - \tau| \leq 5)$	0.7910	0.9794	0.9923
$\widehat{P}(|\hat{\tau} - \tau| \leq 6)$	0.8200	0.9822	0.9930
$\widehat{P}(|\hat{\tau} - \tau| \leq 7)$	0.8560	0.9833	0.9930
$\widehat{P}(|\hat{\tau} - \tau| \leq 8)$	0.8760	0.9850	0.9930
$\widehat{P}(|\hat{\tau} - \tau| \leq 9)$	0.8880	0.9879	0.9930
$\widehat{P}(|\hat{\tau} - \tau| \leq 10)$	0.9060	0.9888	0.9930

Table 8: Precision of the estimator when used with proposed MM chart for different magnitudes of process change (magnitude of shift, δ) based on 10,000 trials. Subgroup size $n = 5$, change point $\tau = 100$ (normal distribution).

	Magnitude of shift, δ		
	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$
$\widehat{P}(\widehat{\tau} = \tau)$	0.2590	0.5961	0.8163
$\widehat{P}(|\widehat{\tau} - \tau| \leq 1)$	0.4800	0.8288	0.9415
$\widehat{P}(|\widehat{\tau} - \tau| \leq 2)$	0.6240	0.9125	0.9777
$\widehat{P}(|\widehat{\tau} - \tau| \leq 3)$	0.7090	0.9535	0.9878
$\widehat{P}(|\widehat{\tau} - \tau| \leq 4)$	0.7600	0.9714	0.9878
$\widehat{P}(|\widehat{\tau} - \tau| \leq 5)$	0.7980	0.9783	0.9928
$\widehat{P}(|\widehat{\tau} - \tau| \leq 6)$	0.8240	0.9811	0.9935
$\widehat{P}(|\widehat{\tau} - \tau| \leq 7)$	0.8610	0.9822	0.9935
$\widehat{P}(|\widehat{\tau} - \tau| \leq 8)$	0.8830	0.9839	0.9935
$\widehat{P}(|\widehat{\tau} - \tau| \leq 9)$	0.8960	0.9868	0.9935
$\widehat{P}(|\widehat{\tau} - \tau| \leq 10)$	0.9130	0.9877	0.9935

Table 9: Precision of the estimator when used with X-Bar chart for different magnitudes of process change (magnitude of shift, δ) based on 10,000 trials. Subgroup size $n = 5$, change point $\tau = 100$ (contaminated distribution, $\beta = 0.10$, $C = 9$).

	Magnitude of shift, δ		
	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$
$\widehat{P}(\widehat{\tau} = \tau)$	0.0943	0.2365	0.4044
$\widehat{P}(|\widehat{\tau} - \tau| \leq 1)$	0.2295	0.4503	0.6430
$\widehat{P}(|\widehat{\tau} - \tau| \leq 2)$	0.3184	0.5669	0.7574
$\widehat{P}(|\widehat{\tau} - \tau| \leq 3)$	0.3873	0.6547	0.8315
$\widehat{P}(|\widehat{\tau} - \tau| \leq 4)$	0.4595	0.7260	0.8667
$\widehat{P}(|\widehat{\tau} - \tau| \leq 5)$	0.5234	0.7826	0.8986
$\widehat{P}(|\widehat{\tau} - \tau| \leq 6)$	0.5672	0.8331	0.9308
$\widehat{P}(|\widehat{\tau} - \tau| \leq 7)$	0.6074	0.8583	0.9354
$\widehat{P}(|\widehat{\tau} - \tau| \leq 8)$	0.6413	0.8827	0.9501
$\widehat{P}(|\widehat{\tau} - \tau| \leq 9)$	0.6655	0.8991	0.9586
$\widehat{P}(|\widehat{\tau} - \tau| \leq 10)$	0.6928	0.9172	0.9620

The philosophy of the proposed robust MM control chart is more in keeping with the desire to provide robust limits in the face of nonnormal distribution situation or that outlier(s) may arise in data collection. The proposed robust MM charting methodology gives better performance than the traditional Shewhart $\overline{X}$ chart if the underlying distribution of chance cause is nonnormal or contaminated. This is usually desirable feature for any control chart to be applied in the industry.

Appendix

In this appendix, we show the derivation of the maximum likelihood estimator (MLE) of τ, the process location change point. For maximum-likelihood estimation techniques, see [1]. Denoting the MLE of the change point τ as $\widehat{\tau}$, let subgroup averages be $\overline{x}_1, \ldots, \overline{x}_T$, the MLE

Table 10: Precision of the estimator when used with median chart for different magnitudes of process change (magnitude of shift, δ) based on 10,000 trials. Subgroup size $n = 5$, change point $\tau = 100$ (contaminated distribution, $\beta = 0.10$, $C = 9$).

	Magnitude of shift, δ		
	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$
$\widehat{P}(\widehat{\tau} = \tau)$	0.1227	0.2499	0.3714
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 1)$	0.3075	0.4692	0.5843
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 2)$	0.4180	0.5811	0.6783
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 3)$	0.5078	0.6597	0.7309
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 4)$	0.5703	0.7016	0.7499
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 5)$	0.6213	0.7342	0.7662
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 6)$	0.6681	0.7688	0.7843
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 7)$	0.6970	0.7802	0.7893
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 8)$	0.7274	0.7981	0.8041
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 9)$	0.7464	0.8085	0.8087
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 10)$	0.7618	0.8190	0.8127

Table 11: Precision of the estimator when used with proposed MM chart for different magnitudes of process change (magnitude of shift, δ) based on 10,000 trials. Subgroup size $n = 5$, change point $\tau = 100$ (contaminated distribution, $\beta = 0.10$, $C = 9$).

	Magnitude of shift, δ		
	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$
$\widehat{P}(\widehat{\tau} = \tau)$	0.1241	0.2525	0.3788
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 1)$	0.3107	0.4753	0.5946
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 2)$	0.4218	0.5861	0.6901
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 3)$	0.5115	0.6690	0.7437
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 4)$	0.5732	0.7138	0.7631
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 5)$	0.6241	0.7476	0.7812
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 6)$	0.6720	0.7834	0.7995
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 7)$	0.7024	0.7961	0.8045
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 8)$	0.7316	0.8120	0.8169
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 9)$	0.7506	0.8234	0.8215
$\widehat{P}(\|\widehat{\tau} - \tau\| \leq 10)$	0.7684	0.8350	0.8244

of τ is the value of τ that maximizes the likelihood function or, equivalently, its logarithm. It is shown that logarithm of the likelihood function is

$$
\begin{aligned}
\log L(\tau, \mu_1 \mid \overline{x}) &= -\frac{n}{2\sigma_0^2} \left(\sum_{i=1}^{\tau} (\overline{x}_i - \mu_0)^2 + \sum_{i=\tau+1}^{T} (\overline{x}_i - \mu_1)^2 \right) \\
&= -\frac{n}{2\sigma_0^2} \left(\sum_{i=1}^{T} \overline{x}_i^2 - 2\mu_0 \sum_{i=1}^{\tau} \overline{x}_i - \tau\mu_0^2 - 2\mu_1 \sum_{i=\tau+1}^{T} \overline{x}_i + (T - \tau)\mu_1^2 \right).
\end{aligned} \tag{A.1}
$$

Table 12: Precision of the estimator when used with X-Bar chart for different magnitudes of process change (magnitude of shift, δ) based on 10,000 trials. Subgroup size $n = 5$, change point $\tau = 100$ (slash distribution).

	Magnitude of shift, δ		
	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$
$\widehat{P}(\widehat{\tau} = \tau)$	0.0341	0.0531	0.0829
$\widehat{P}(|\widehat{\tau} - \tau| \leq 1)$	0.1057	0.1444	0.1798
$\widehat{P}(|\widehat{\tau} - \tau| \leq 2)$	0.1541	0.2064	0.2547
$\widehat{P}(|\widehat{\tau} - \tau| \leq 3)$	0.1862	0.2542	0.3053
$\widehat{P}(|\widehat{\tau} - \tau| \leq 4)$	0.2225	0.2937	0.3547
$\widehat{P}(|\widehat{\tau} - \tau| \leq 5)$	0.2547	0.3356	0.3892
$\widehat{P}(|\widehat{\tau} - \tau| \leq 6)$	0.2899	0.3659	0.4222
$\widehat{P}(|\widehat{\tau} - \tau| \leq 7)$	0.3038	0.3891	0.4482
$\widehat{P}(|\widehat{\tau} - \tau| \leq 8)$	0.3307	0.4165	0.4690
$\widehat{P}(|\widehat{\tau} - \tau| \leq 9)$	0.3554	0.4358	0.4849
$\widehat{P}(|\widehat{\tau} - \tau| \leq 10)$	0.3787	0.4618	0.5094

Table 13: Precision of the estimator when used with median chart for different magnitudes of process change (magnitude of shift, δ) based on 10,000 trials. Subgroup size $n = 5$, change point $\tau = 100$ (slash distribution).

	Magnitude of shift, δ		
	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$
$\widehat{P}(\widehat{\tau} = \tau)$	0.0437	0.0545	0.0761
$\widehat{P}(|\widehat{\tau} - \tau| \leq 1)$	0.1107	0.1208	0.1493
$\widehat{P}(|\widehat{\tau} - \tau| \leq 2)$	0.1623	0.1795	0.2021
$\widehat{P}(|\widehat{\tau} - \tau| \leq 3)$	0.2004	0.2116	0.2323
$\widehat{P}(|\widehat{\tau} - \tau| \leq 4)$	0.2261	0.2397	0.2567
$\widehat{P}(|\widehat{\tau} - \tau| \leq 5)$	0.2526	0.2549	0.2715
$\widehat{P}(|\widehat{\tau} - \tau| \leq 6)$	0.2845	0.2812	0.2940
$\widehat{P}(|\widehat{\tau} - \tau| \leq 7)$	0.2957	0.2915	0.2995
$\widehat{P}(|\widehat{\tau} - \tau| \leq 8)$	0.3177	0.3072	0.3140
$\widehat{P}(|\widehat{\tau} - \tau| \leq 9)$	0.3310	0.3246	0.3270
$\widehat{P}(|\widehat{\tau} - \tau| \leq 10)$	0.3460	0.3424	0.3396

We can see that there are two unknowns in the log-likelihood function: τ and μ_1. If the change point τ was known, the MLE of μ_1 would be $\widehat{\mu}_1 = \overline{\overline{X}}_{T,\tau} = (T - \tau)^{-1} \sum_{i=\tau+1}^{T} \overline{X}_i$, the average of the $T - \tau$ most recent subgroup averages. Substituting this back into (A.1), we obtain

$$\log L(\tau \mid \overline{x}) = -\frac{n}{2\sigma_0^2} \left(\sum_{i=1}^{T} \overline{x}_i^2 - 2\mu_0 \sum_{i=1}^{\tau} \overline{x}_i + \tau \mu_0^2 - (T - \tau)\overline{\overline{X}}_{T,\tau}^2 \right). \tag{A.2}$$

It is easy to verify that this is equivalent to

$$\log L(\tau \mid \overline{x}) = -\frac{n}{2\sigma_0^2} \left(\sum_{i=1}^{T} \overline{x}_i^2 - 2\mu_0 \sum_{i=1}^{\tau} \overline{x}_i + \tau \mu_0^2 - (T - \tau)\left(\overline{\overline{X}}_{T,\tau} - \mu_0 \right)^2 \right). \tag{A.3}$$

Table 14: Precision of the estimator when used with proposed MM chart for different magnitudes of process change (magnitude of shift, δ) based on 10,000 trials. Subgroup size $n = 5$, change point $\tau = 100$ (slash distribution).

	Magnitude of shift, δ		
	$\delta = 1.0$	$\delta = 2.0$	$\delta = 3.0$
$\widehat{P}(\widehat{\tau} = \tau)$	0.0437	0.0549	0.0780
$\widehat{P}(|\widehat{\tau} - \tau| \leq 1)$	0.1115	0.1228	0.1501
$\widehat{P}(|\widehat{\tau} - \tau| \leq 2)$	0.1631	0.1802	0.2030
$\widehat{P}(|\widehat{\tau} - \tau| \leq 3)$	0.2027	0.2154	0.2329
$\widehat{P}(|\widehat{\tau} - \tau| \leq 4)$	0.2300	0.2439	0.2585
$\widehat{P}(|\widehat{\tau} - \tau| \leq 5)$	0.2565	0.2604	0.2722
$\widehat{P}(|\widehat{\tau} - \tau| \leq 6)$	0.2884	0.2883	0.2947
$\widehat{P}(|\widehat{\tau} - \tau| \leq 7)$	0.2996	0.3002	0.3002
$\widehat{P}(|\widehat{\tau} - \tau| \leq 8)$	0.3228	0.3159	0.3147
$\widehat{P}(|\widehat{\tau} - \tau| \leq 9)$	0.3354	0.3333	0.3277
$\widehat{P}(|\widehat{\tau} - \tau| \leq 10)$	0.3504	0.3505	0.3403

In what follows, the value of τ that maximizes the log-likelihood function is

$$\widehat{\tau} = \arg\max_{t}\left\{ (T - \tau)\left(\overline{\overline{X}}_{T,t} - \mu_0\right)^2\right\}$$
$$= \arg\max_{t}\ \{C_t\},$$

(A.4)

where $C_t = (T-\tau)(\overline{\overline{X}}_{T,t}-\mu_0)^2$, that is, $\widehat{\tau}$ is the value of τ in the range $0 \leq t < T$ which maximizes

$$C_t = (T - \tau)\left(\overline{\overline{X}}_{T,t} - \mu_0\right)^2.$$

(A.5)

Acknowledgment

References

[1] T. R. Samuel, J. J. Pignatiello, and J. A. Calvin, "Identifying the time of a step change with $\overline{X}$ control charts," *Quality Engineering*, vol. 10, no. 3, pp. 521–527, 1998.

[2] D. V. Hinkley, "Inference about the change-point in a sequence of random variables," *Biometrika*, vol. 57, pp. 1–17, 1970.

[3] P. J. Huber, "Robust estimation of a location parameter," *Annals of Mathematical Statistics*, vol. 35, pp. 73–101, 1964.

[4] M. Salibian-Barrera and R. H. Zamar, "Uniform asymptotics for robust location estimates when the scale is unknown," *The Annals of Statistics*, vol. 32, no. 4, pp. 1434–1447, 2004.

[5] R. Fraiman, V. J. Yohai, and R. H. Zamar, "Optimal robust M-estimates of location," *The Annals of Statistics*, vol. 29, no. 1, pp. 194–223, 2001.

[6] P. Rousseeuw and V. Yohai, "Robust regression by means of S-estimators," in *Robust and Nonlinear Time Series Analysis*, vol. 26 of *Lecture Notes in Statistics*, pp. 256–272, Springer, Berlin, Germany, 1984.

[7] A. E. Beaton and J. W. Tukey, "The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data," *Technometrics*, vol. 16, no. 2, pp. 147–185, 1974.

[8] V. J. Yohai, "High breakdown-point and high efficiency robust estimates for regression," *The Annals of Statistics*, vol. 15, no. 2, pp. 642–656, 1987.

[9] G. Betteley, N. Mettrick, E. Sweeney, and D. Wilson, *Using Statistics in Industry: Quality Improvement through Total Process Control*, Prentice-Hall, London, UK, 1994.

[10] A. Francis, *Business Mathematics and Statistics*, DP Publications, London, UK, 1995.

[11] F. R. Hampel, "The influence curve and its role in robust estimation," *Journal of the American Statistical Association*, vol. 69, pp. 383–393, 1974.

[12] R. D. Martin and R. H. Zamar, "Bias robust estimation of scale," *The Annals of Statistics*, vol. 21, no. 2, pp. 991–1017, 1993.

[13] S. Kocherlakota and K. Kocherlakota, "Control chart using robust estimators," *Total Quality Management*, vol. 6, no. 1, pp. 79–90, 1995.

[14] D. C. Montgomery, *Introduction to Statistical Quality Control*, John Wiley and Sons, New York, NY, USA, 2nd edition, 1991.

[15] W. Q. Meeker and L. A. Escobar, "Teaching about approximate confidence regions based on maximum likelihood estimation," *The American Statistician*, no. 49, pp. 48–53, 1985.

[16] G. E. P. Box and D. R. Cox, "An analysis of transformations," *Journal of the Royal Statistical Society B*, vol. 26, pp. 211–243, 1964.

[17] D. Siegmund, "Boundary crossing probabilities and statistical applications," *The Annals of Statistics*, vol. 14, no. 2, pp. 361–404, 1986.

[18] C. M. Davis and B. M. Adams, "Robust monitoring of contaminated data," *Journal of Quality Technology*, vol. 37, no. 2, pp. 163–174, 2005.