Communications for Statistical Applications and Methods
2024, Vol. 31, No. 1, 129–142

https://doi.org/10.29220/CSAM.2024.31.1.129
Print ISSN 2287-7843 / Online ISSN 2383-4757

# Monitoring the asymmetry parameter of a skew-normal distribution

Hyun Jun Kim[a], Jaeheon Lee[1,a]

[a]Department of Applied Statistics, Chung-Ang University, Korea

**Abstract**

In various industries, especially manufacturing and chemical industries, it is often observed that the distribution of a specific process, initially having followed a normal distribution, becomes skewed as a result of unexpected causes. That is, a process deviates from a normal distribution and becomes a skewed distribution. The skew-normal (SN) distribution is one of the most employed models to characterize such processes. The shape of this distribution is determined by the asymmetry parameter. When this parameter is set to zero, the distribution is equal to the normal distribution. Moreover, when there is a shift in the asymmetry parameter, the mean and variance of a SN distribution shift accordingly. In this paper, we propose procedures for monitoring the asymmetry parameter, based on the statistic derived from the noncentral $t$-distribution. After applying the statistic to Shewhart and the exponentially weighted moving average (EWMA) charts, we evaluate the performance of the proposed procedures and compare it with previously studied procedures based on other skewness statistics.

Keywords: asymmetry parameter, EWMA chart, noncentral $t$-distribution, Shewhart chart, skew-normal distribution

## 1. Introduction

In various production processes, product quality characteristics are observed to follow a normal distribution. Examples of such characteristics include the length of bolts or screws, the diameter of bearings, and the electrical resistance of circuit boards. Traditionally, researchers have primarily focused on detecting changes in the mean or variance of these distributions, as investigated by Ryu *et al*. (2010), Shu *et al*. (2010), Guo and Wang (2015), Peng *et al*. (2015), and many others. In addition, considerable attention has been paid to studies of joint monitoring of the mean and variance, as evidenced by the works of Hawkins and Deng (2009), Wu *et al*. (2010), Sheu *et al*. (2012), McCracken *et al*. (2013), Knoth (2015), and Li *et al*. (2016).

It has been generally assumed that only the parameters of the distribution change, while the population distribution itself remains constant. However, several authors have recently pointed out that the normally distributed process may also change its shape due to undesirable causes, resulting in a skewed distribution (see, for example, Zou and Tsung, 2010; Ross and Adams, 2012; Li *et al*., 2013). Real-life examples of this situation have been studied by several authors, such as Vincent and Walsh (1997), Rahman and Hossain (2008), Figueiredo and Gomes (2013), and Mukherjee *et al*. (2013).

To address this issue, the skew-normal (SN) distribution is the most used model. Several authors, such as Azzalini (1985, 1986, 2005), Henze (1986), Chen *et al*. (2004), Arellano-Valle *et al*. (2004),

[1]Corresponding author: Department of Applied Statistics, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul 06974, Korea. E-mail: jaeheon@cau.ac.kr

Published 31 January 2024 / journal homepage: http://csam.or.kr
© 2024 The Korean Statistical Society, and Korean International Statistical Society. All rights reserved.

and Mameli and Musio (2013), have extensively investigated the statistical properties of the SN distribution. Nevertheless, monitoring procedures for the SN distribution have received relatively less attention in the literature. To this end, Li *et al*. (2019) recently performed a comprehensive evaluation of control charts based on different skewness statistics.

In this paper, we propose procedures for monitoring a normally distributed process that may become skewed over time. When a normal distribution becomes skewed, its standard deviation tends to decrease. The proposed statistic in the procedure, which is based on the noncentral $t$-distribution, includes a term that represents variability in the denominator. As a result, it will amplify the effect of the skewness change and increase the detectability. The proposed statistic is applied to Shewhart and the exponentially weighted moving average (EWMA) charts, and the performance is compared with that of the procedure based on previously studied skewness statistics.

The remainder of this paper is organized as follows: Section 2 provides a description of the SN distribution along with its basic properties, and presents previous work related to monitoring the asymmetry parameter of SN distributions. In Section 3, we propose a statistic for monitoring the skewness. Subsequent subsections describe detailed procedures for constructing Shewhart and EWMA charts based on the proposed statistic. A performance comparison is performed and the results are presented in Section 4. Section 5 provides an example of applying the proposed procedure to real process data. Finally, we conclude with some closing remarks in Section 6.

## 2. Skew-normal distribution and previous work

Let a continuous random variable $X$, which represents the quality characteristic of the process being monitored, follow a SN distribution with the location parameter $\xi$, the scale parameter $w$, and the shape (asymmetry) parameter $\lambda$. This can be expressed as

$$X \sim \mathrm{SN}(\xi, w, \lambda)$$

in a standard notation. Then the probability density function (pdf) is of the form:

$$f(x\,;\,\xi, w, \lambda) = \frac{2}{w}\,\phi\left(\frac{x-\xi}{w}\right)\Phi\left(\lambda\,\frac{x-\xi}{w}\right), \quad -\infty < x < \infty, \quad -\infty < \xi < \infty, \quad -\infty < \lambda < \infty, \quad w > 0,$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the pdf and cumulative distribution function (cdf) of a standard normal distribution, respectively. The mean, $\mu$, and variance, $\sigma^2$, of SN$(\xi, w, \lambda)$ are

$$\mu = \xi + \lambda\, w\sqrt{\frac{2}{\pi\,(1+\lambda^2)}} \quad \text{and} \quad \sigma^2 = \left(1 - \frac{2\lambda^2}{\pi\,(1+\lambda^2)}\right) w^2, \tag{2.1}$$

respectively.

The shape parameter $\lambda$ decides the skewness of a SN distribution. It is easy to see that when $\lambda = 0$, the SN distribution becomes equal to the normal distribution with mean $\mu = \xi$ and variance $\sigma^2 = w^2$. When $\lambda > 0$, the distribution is positively skewed, and when $\lambda < 0$, it is negatively skewed. Figure 1 illustrates how the shape of a SN distribution changes with respect to different values of $\lambda$.

Various statistical procedures can be considered that effectively monitor the asymmetry parameter of a SN distribution. In particular, Li *et al*. (2019) proposed a monitoring procedure based on the likelihood ratio test and compared its performance with procedures based on several skewness statistics. They only considered the monitoring procedure for the asymmetry parameter $\lambda$ of a SN distribution

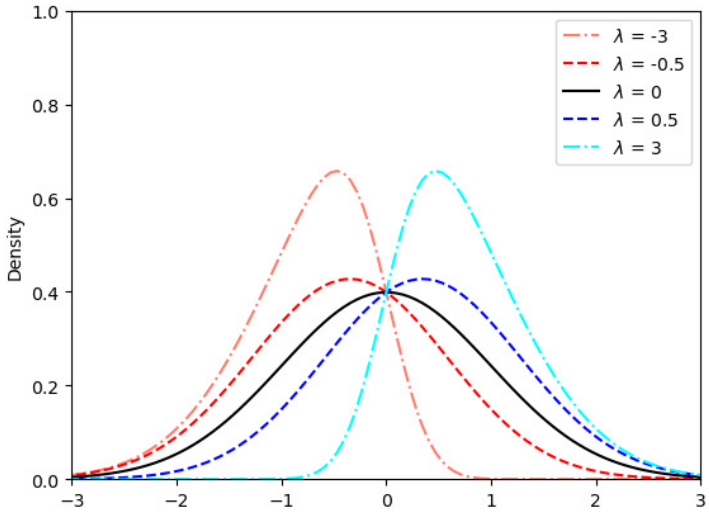


Figure 1: *The pdf of a SN distribution.*

and assumed that the location parameter $\xi$ and the scale parameter $w$ remain constant. In this paper, we will consider the same situation for a fair comparison.

Statistics used for comparisons, considered in Li *et al.* (2019), were the likelihood ratio test statisic, the $L$-skewness statisic, the sample skewness statistic, the sample distance skewness statistic, the median skewness statistic, and the quantile-based statistic. Among them, the performances of the procedures based on the likelihood ratio statisic and the distance skewness statistic were found to be good. However, they noted that the procedure based on the distance skewness statistic operates more competitively, and the procedure based on the likelihood ratio statisic is difficult to use when the sample size is small. Since our study focuses on cases with a sample size of 10 or less, only the procedure based on the distance skewness statistic will be considered for comparison.

Let $\mathbf{X}_k = (X_{k1}, X_{k2}, \ldots, X_{kn})$ represent the sample of size $n$ obtained from SN$(\xi, w, \lambda)$ at sampling time $k$. It is assumed that if the distribution remains SN$(\xi, w, \lambda_0 = 0)$, the process is in control, and if the distribution changes to SN$(\xi, w, \lambda_1 \neq 0)$, the process is out of control.

A brief explanation of the procedure based on the distance skewness statistic is as follows: The distance skewness statistic is a simple and reliable statistic for assessing diagonal symmetry. The sample distance skewness at time $k$ is defined as

$$D_k = 1 - \frac{\sum_{i,j} \left| Z_{ki} - Z_{kj} \right|}{\sum_{i,j} \left| Z_{ki} + Z_{kj} \right|}, \tag{2.2}$$

where $Z$'s represent the standardized observations within a sample. The value of $D_k$ is always between 0 to 1. The case of $D_k = 0$ indicates that the distribution is under symmetry, and the case of $0 < D_k < 1$ indicates a lack of symmetry. Li *et al.* (2019) proposed the Shewhart charting procedure that gives a signal if $D > h_D$, where the control limit $h_D$ is chosen to give specified in-control performance.

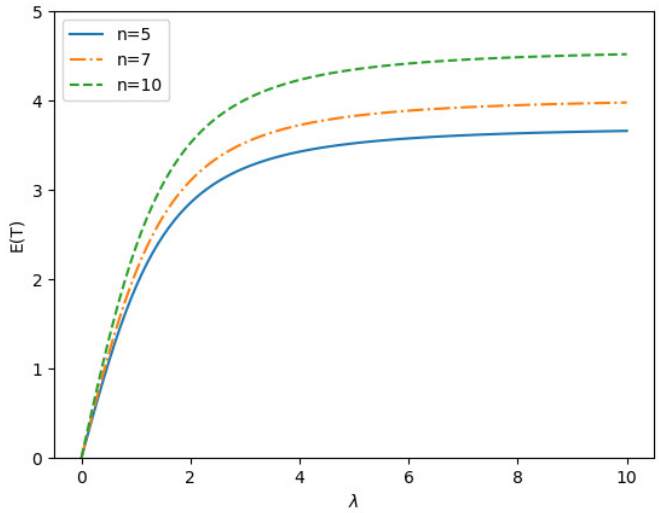


Figure 2: $E(T_k)$ *versus* $\lambda$ *for* $n = 5, 7, 10$ *when* $\xi = 0$ *and* $w = 1$.

## 3. The proposed procedure for monitoring the skewness

### 3.1. The proposed statistic

The objective of this paper is to detect shifts in the asymmetry parameter of a SN distribution during Phase II. We assume that the values of $\xi$ and $w$ are known or have been estimated accurately enough during the Phase I period.

When the process is in control, that is $\lambda = 0$, the in-control mean and variance are $\mu_0 = \xi$ and $\sigma_0^2 = w^2$, respectively. When a shift to some value $\lambda_1$ ($\neq 0$) occurs in the asymmetry parameter $\lambda$, i.e., when the process becomes out-of-control, the corresponding mean and variance also shift accordingly. As can be seen from equation (2.1), the out-of-control mean and variance become

$$\mu_1 = \mu_0 + \lambda_1 \sigma_0 \sqrt{\frac{2}{\pi\left(1+\lambda_1^2\right)}} \quad \text{and} \quad \sigma_1^2 = \left(1 - \frac{2\lambda_1^2}{\pi\left(1+\lambda_1^2\right)}\right)\sigma_0^2, \tag{3.1}$$

respectively.

It is evident that the direction of change of the mean coincides with that of $\lambda$, while the variance consistently decreases. Using these characteristics, we propose the following statistic at time $k$ to detect a shift in $\lambda$.

$$T_k = \frac{\sqrt{n}\,\bar{X}_k}{S_k}, \tag{3.2}$$

where $\bar{X}_k$ and $S_k$ are the sample mean and sample standard deviation of $\mathbf{X}_k$, respectively. It is well known that $T_k$ follows a noncentral $t$-distribution with $n-1$ degrees of freedom and non-centrality parameter $\sqrt{n}\,\mu/\sigma$. When the process is in control, the non-centrality parameter is $\sqrt{n}\,\xi/w$, and when the process is out of control, it becomes $\sqrt{n}\,\mu_1/\sigma_1$.

We expect this statistic to improve the detectability of parameter shifts, as decreasing variance amplifies the impact of shifts. Figure 2 shows values of $E(T_k)$ versus $\lambda$ for $n = 5, 7, 10$ when $\xi = 0$

Table 1: Values of ARL for Shewhart$_D$, EWMA$_D$, $\bar{X} - S$, Shewhart$_T$, and EWMA$_T$ charts when $n = 3$

| $\lambda$ | Shewhart$_D$ | EWMA$_D$ | | | $\bar{X} - S$ |
|---|---|---|---|---|---|
| | | $r = 0.05$ | $r = 0.1$ | $r = 0.2$ | |
| 0 | 374.03 (375.00) | 363.67 (301.20) | 366.46 (336.78) | 370.80 (354.94) | 371.26 (371.16) |
| 0.3 | 319.82 (320.25) | 198.98 (142.41) | 204.19 (175.90) | 220.42 (207.24) | 305.64 (305.23) |
| 0.5 | 263.50 (261.56) | 109.73 (60.21) | 104.27 (77.63) | 116.45 (102.98) | 250.94 (249.92) |
| 1 | 168.10 (166.29) | 48.15 (13.50) | 36.19 (16.24) | 34.23 (22.95) | 182.69 (182.85) |
| 2 | 113.48 (112.64) | 29.66 (4.67) | 19.66 (4.72) | 15.34 (6.04) | 152.18 (152.02) |
| 3 | 100.40 (99.33) | 25.57 (3.12) | 16.56 (3.01) | 12.30 (3.65) | 143.84 (142.84) |
| 5 | 96.17 (95.48) | 23.20 (2.21) | 14.84 (2.08) | 10.72 (2.44) | 136.56 (135.50) |
| 10 | 93.96 (93.62) | 22.03 (1.73) | 14.01 (1.61) | 10.02 (1.88) | 130.75 (129.84) |
| CL | 0.962 | 0.456 | 0.512 | 0.594 | 1.850<br>2.702, 0.026 |
| $\lambda$ | Shewhart$_T$ | EWMA$_T$ | | | |
| | | $r = 0.05$ | $r = 0.1$ | $r = 0.2$ | |
| 0 | 369.93 (366.71) | 370.66 (365.1) | 370.39 (367.87) | 370.18 (372.97) | |
| 0.3 | 318.26 (317.81) | **110.10 (94.69)** | 188.07 (182.08) | 259.44 (257.64) | |
| 0.5 | 261.30 (262.72) | **49.81 (34.27)** | 95.48 (86.99) | 171.30 (168.20) | |
| 1 | 166.70(165.86) | **20.77 (9.59)** | 31.62 (22.10) | 72.53 (67.81) | |
| 2 | 111.83 (111.88) | **13.00 (4.62)** | 16.15 (8.07) | 34.47 (28.90) | |
| 3 | 99.82 (99.45) | **11.37 (3.68)** | 13.47 (5.89) | 27.18 (21.82) | |
| 5 | 94.68 (94.32) | **10.47 (3.17)** | 12.07 (4.81) | 23.81 (18.13) | |
| 10 | 92.50 (91.88) | **10.06 (2.93)** | 11.46 (4.35) | 22.25 (16.52) | |
| CL | 19.206 | 7.280 | 9.181 | 11.96 | |

and $w = 1$. Values of $E(T_k)$ can be calculated using the following equation:

$$E(T_k) = \sqrt{\frac{\nu}{2}}\,\frac{\Gamma\left((\nu-1)/2\right)}{\Gamma(\nu/2)}\,\delta\,,$$

where $\nu$ is the degrees of freedom and $\delta$ is the non-centrality parameter. From Figure 2, it can be seen that on average, $T_k$ reflects the change in $\lambda$ well. However, as the amount of change in $\lambda$ increases, the amount of change in $T_k$ becomes relatively smaller.

## 3.2. Shewhart and EWMA charts based on $T_k$

In this subsection, we propose monitoring procedures using Shewhart and EWMA charts based on the statistic $T_k$. Typically, the performance of control charts is measured by the ARL (average run length), which quantifies the average number of samples required before the chart gives an out-of-control signal. We can choose the control limits of control charts so that the in-control ARL becomes the desired ARL$_0$ value.

First, the Shewhart charting procedure is proposed. It is well known that since the RL (run length) of Shewhart charts follows a geometric distribution, the type I error $\alpha$ and ARL$_0$ have a relationship of $\alpha = 1/\text{ARL}_0$. Here, type I error represents the false alarm rate that will give a signal when the process is in-control. The Shewhart chart (referred to as the Shewhart$_T$ chart) gives a signal at time $k$ when $T_k$ in equation (3.2) falls outside the control limits given by

$$\text{UCL}_T = F_t^{-1}\left(1-\frac{\alpha}{2}\,\middle|\, n-1, \frac{\sqrt{n}\,\xi}{w}\right), \quad \text{LCL}_T = F_t^{-1}\left(\frac{\alpha}{2}\,\middle|\, n-1, \frac{\sqrt{n}\,\xi}{w}\right), \tag{3.3}$$

where UCL$_T$ and LCL$_T$ are the upper and lower control limits, respectively, and $F_t^{-1}(\cdot)$ is the inverse cdf (cumulative distribution function) of the noncentral $t$-distribution with $n-1$ degrees of freedom and non-centrality parameter $\sqrt{n}\,\xi/w$.

Table 2: Values of ARL for Shewhart$_D$, EWMA$_D$, $\bar{X} - S$, Shewhart$_T$, and EWMA$_T$ charts when $n = 5$

| $\lambda$ | Shewhart$_D$ | EWMA$_D$ | | | $\bar{X} - S$ |
|---|---|---|---|---|---|
| | | $r = 0.05$ | $r = 0.1$ | $r = 0.2$ | |
| 0 | 373.53 (373.00) | 364.42 (307.17) | 371.13 (343.79) | 370.81 (358.71) | 370.52 (370.60) |
| 0.3 | 224.74 (243.74) | 129.93 (82.05) | 129.25 (104.28) | 141.39 (128.66) | 243.13 (242.35) |
| 0.5 | 155.04 (154.17) | 62.74 (25.93) | 53.01 (32.78) | 55.26 (44.34) | 160.75 (159.83) |
| 1 | 64.39 (64.00) | 28.05 (6.04) | 19.24 (6.17) | 15.56 (7.53) | 83.80 (83.29) |
| 2 | 31.88 (31.34) | 17.75 (2.32) | 11.53 (2.07) | 8.35 (2.15) | 55.31 (54.97) |
| 3 | 26.30 (25.85) | 15.55 (1.58) | 10.03 (1.38) | 7.14 (1.38) | 48.88 (48.59) |
| 5 | 23.93 (23.34) | 14.31 (1.14) | 9.21 (0.99) | 6.49 (0.97) | 43.59 (42.97) |
| 10 | 23.49 (23.01) | 13.75 (0.93) | 8.83 (0.80) | 6.21 (0.78) | 40.55 (39.95) |
| CL | 0.839 | 0.303 | 0.349 | 0.420 | 1.433<br>2.199, 0.136 |
| $\lambda$ | Shewhart$_T$ | EWMA$_T$ | | | |
| | | $r = 0.05$ | $r = 0.1$ | $r = 0.2$ | |
| 0 | 376.55 (373.70) | 370.54 (359.38) | 370.41 (364.79) | 370.28 (366.41) | |
| 0.3 | 246.44 (244.88) | **32.36 (19.32)** | 40.82 (31.01) | 74.20 (68.16) | |
| 0.5 | 155.39 (154.62) | **17.07 (7.42)** | 18.26 (10.35) | 28.20 (22.27) | |
| 1 | 65.20 (64.57) | 9.03 (2.72) | **8.53 (3.05)** | 9.67 (4.84) | |
| 2 | 32.17 (31.76) | 6.28 (1.45) | **5.70 (1.48)** | **5.70 (1.86)** | |
| 3 | 26.39 (25.96) | 5.64 (1.16) | 5.09 (1.15) | **4.95 (1.36)** | |
| 5 | 24.05 (23.57) | 5.27 (0.99) | 4.74 (0.97) | **4.55 (1.09)** | |
| 10 | 23.73 (23.20) | 5.11 (0.91) | 4.58 (0.87) | **4.37 (0.96)** | |
| CL | 6.651 | 3.575 | 4.059 | 4.694 | |

Next, the EWMA charting procedure is proposed. The EWMA chart has the advantage of being sensitive to small and moderate-sized shifts in process parameters. As a result, the EWMA chart is one of the alternatives to the Shewhart chart when small shifts in parameters need to be detected quickly.

The EWMA chart statistic at time $k$ is defined as

$$V_k = r\,T_k + (1 - r)\,V_{k-1},$$

where $r$ $(0 < r \leq 1)$ is a smoothing constant. It is efficient to use smaller values of $r$ to detect smaller shifts in process parameters. The control limits for the EWMA chart are given by

$$\mathrm{UCL}_{ET} = E_0(T_k) + L\sqrt{\frac{r}{2-r}\,\left[1 - (1-r)^{2k}\right]\,\mathrm{Var}_0(T_k)}\,,$$

$$\mathrm{LCL}_{ET} = E_0(T_k) - L\sqrt{\frac{r}{2-r}\,\left[1 - (1-r)^{2k}\right]\,\mathrm{Var}_0(T_k)}\,,$$

where $E_0(T_k)$ and $\mathrm{Var}_0(T_k)$ are the mean and standard deviation of $T_k$, respectively, when the process is in control, and $L$ is chosen to satisfy the desired $\mathrm{ARL}_0$ value. The EWMA chart (the EWMA$_T$ chart) signals when the statistic $V_k$ exceeds the control limits. Since the SN distribution is $N(\xi, w^2)$ when the process is in control, $E_0(T_k)$ and $\mathrm{Var}_0(T_k)$ are calculated as

$$E_0(T_k) = \frac{\sqrt{n}\,\xi}{c_4 w} \quad \text{and} \quad \mathrm{Var}_0(T_k) = 1 - \frac{n\,\xi^2\left(1 - c_4^2\right)}{c_4^2\,w^2},$$

where $c_4$ is the unbiasing constant which allows $S_k$ to be unbiased. The above control limits converge

Table 3: Values of ARL for Shewhart$_D$, EWMA$_D$, $\bar{X} - S$, Shewhart$_T$, and EWMA$_T$ charts when $n = 7$

| $\lambda$ | Shewhart$_D$ | EWMA$_D$ | | | $\bar{X} - S$ |
|---|---|---|---|---|---|
| | | $r = 0.05$ | $r = 0.1$ | $r = 0.2$ | |
| 0 | 371.56 (371.00) | 365.32 (310.33) | 361.69 (336.87) | 376.58 (364.59) | 371.16 (373.24) |
| 0.3 | 184.47 (183.71) | 94.22 (52.17) | 88.42 (66.72) | 98.99 (88.22) | 198.27 (198.72) |
| 0.5 | 94.58 (94.14) | 44.61 (15.57) | 34.67 (18.09) | 33.9 (24.82) | 109.77 (109.29) |
| 1 | 29.22 (28.76) | 20.25 (3.96) | 13.57 (3.67) | 10.47 (4.11) | 44.55 (44.24) |
| 2 | 11.35 (10.81) | 13.03 (1.56) | 8.52 (1.35) | 6.18 (1.30) | 24.48 (24.04) |
| 3 | 8.28 (7.76) | 11.52 (1.08) | 7.53 (0.92) | 5.41 (0.86) | 20.24 (19.76) |
| 5 | 6.84 (6.33) | 10.70 (0.80) | 7.00 (0.68) | 5.02 (0.64) | 17.46 (16.96) |
| 10 | 6.34 (5.80) | 10.33 (0.66) | 6.76 (0.59) | 4.86 (0.54) | 15.88 (15.38) |
| CL | 0.730 | 0.229 | 0.269 | 0.329 | 1.433<br>1.975, 0.235 |
| $\lambda$ | Shewhart$_T$ | EWMA$_T$ | | | |
| | | $r = 0.05$ | $r = 0.1$ | $r = 0.2$ | |
| 0 | 369.49 (369.24) | 370.46 (356.06) | 370.27 (362.94) | 370.35 (366.84) | |
| 0.3 | 182.17 (181.22) | **22.95 (11.71)** | 25.11 (16.34) | 36.77 (30.71) | |
| 0.5 | 93.34 (93.06) | 12.77 (4.74) | **12.34 (5.680)** | 14.52 (9.09) | |
| 1 | 28.90 (28.38) | 7.11 (1.8) | 6.38 (1.86) | **6.19 (2.28)** | |
| 2 | 11.11 (10.59) | 5.09 (0.98) | 4.49 (0.95) | **4.12 (1.00)** | |
| 3 | 8.09 (7.58) | 4.61 (0.79) | 4.06 (0.75) | **3.70 (0.77)** | |
| 5 | 6.68 (6.20) | 4.33 (0.68) | 3.83 (0.64) | **3.46 (0.64)** | |
| 10 | 6.11 (5.59) | 4.21 (0.63) | 3.72 (0.59) | **3.36 (0.59)** | |
| CL | 4.906 | 3.078 | 3.419 | 3.799 | |

to the asymptotic limits given by

$$\mathrm{UCL}_{ET} = E_0(T_k) + L_T \sqrt{\frac{r}{2-r}\,\mathrm{Var}_0(T_k)}\,, \quad \mathrm{LCL}_{ET} = E_0(T_k) - L_T \sqrt{\frac{r}{2-r}\,\mathrm{Var}_0(T_k)}\,. \tag{3.4}$$

For simplicity, we use the asymptotic control limits defined in equation (3.4).

## 4. Performance evaluation

### 4.1. Procedures being compared

In this paper, we will consider the Shewhart chart and the EWMA chart based on $D$ in equation (2.2) (denoted by the Shewhart$_D$ chart and the EWMA $_D$ chart, respectively) proposed by Li *et al*. (2019), to evaluate the performance of the proposed monitoring procedures. We also consider the traditional $\bar{X}$-$S$ chart, which is used to detect changes in the mean and variance of a process.

The procedure for the Shewhart$_D$ chart is presented in Section 2, and the procedure for the EWMA$_D$ chart is similar to the EWMA$_T$ chart in Section 3. The EWMA$_D$ chart statistic based on $D$ at time $k$ is defined as

$$Y_k = r\,D_k + (1-r)\,Y_{k-1}.$$

Since $D_k$ has only positive values, the EWMA$_D$ chart gives a signal when $Y_k$ is greater than $\mathrm{UCL}_{ED}$, which is defined as follows:

$$\mathrm{UCL}_{ED} = E_0(D_k) + L_D \sqrt{\frac{r}{2-r}\,\mathrm{Var}_0(D_k)}. \tag{4.1}$$

As can be seen in equation (3.1), a change in $\lambda$ can change both the mean and variance of the process, so the use of the traditional $\bar{X}$-$S$ chart will also be considered. Let $\alpha$ be the overall type I

Table 4: Values of ARL for Shewhart$_D$, EWMA$_D$, $\bar{X}-S$, Shewhart$_T$, and EWMA$_T$ charts when $n = 10$

| $\lambda$ | Shewhart$_D$ | EWMA$_D$ | | | $\bar{X}-S$ |
|---|---|---|---|---|---|
| | | $r = 0.05$ | $r = 0.1$ | $r = 0.2$ | |
| 0 | 369.67 (370.00) | 375.80 (320.45) | 369.02 (338.33) | 376.11 (366.88) | 371.66 (371.81) |
| 0.3 | 125.56 (124.80) | 67.79 (32.61) | 58.58 (39.48) | 63.91 (54.34) | 149.65 (149.28) |
| 0.5 | 51.80 (51.34) | 31.83 (9.58) | 23.12 (10.15) | 20.81 (13.15) | 68.94 (68.39) |
| 1 | 12.12 (11.69) | 14.62 (2.64) | 9.74 (2.32) | 7.32 (2.36) | 21.15 (20.66) |
| 2 | 3.85 (3.31) | 9.59 (1.07) | 6.39 (0.91) | 4.70 (0.82) | 9.59 (9.09) |
| 3 | 2.62 (2.05) | 8.58 (0.75) | 5.73 (0.65) | 4.23 (0.55) | 7.44 (6.92) |
| 5 | 2.01 (1.42) | 8.03 (0.57) | 5.36 (0.51) | 4.01 (0.39) | 6.18 (5.65) |
| 10 | 1.75 (1.15) | 7.80 (0.50) | 5.19 (0.41) | 3.93 (0.32) | 5.62 (5.11) |
| CL | 0.602 | 0.169 | 0.199 | 0.250 | 1.014<br>1.792, 0.341 |
| $\lambda$ | Shewhart$_T$ | EWMA$_T$ | | | |
| | | $r = 0.05$ | $r = 0.1$ | $r = 0.2$ | |
| 0 | 371.78 (369.85) | 370.48 (353.61) | 370.60 (362.19) | 370.39 (365.46) | |
| 0.3 | 125.44 (124.70) | **17.48 (7.77)** | 17.50 (9.73) | 21.71 (15.88) | |
| 0.5 | 51.76 (50.99) | 10.12 (3.23) | **9.25 (3.55)** | 9.52 (4.83) | |
| 1 | 12.00 (11.59) | 5.88 (1.27) | 5.15 (1.26) | **4.71 (1.36)** | |
| 2 | 3.80 (3.26) | 4.31 (0.71) | 3.78 (0.67) | **3.37 (0.66)** | |
| 3 | **2.57 (2.00)** | 3.94 (0.58) | 3.46 (0.57) | 3.09 (0.51) | |
| 5 | **2.00 (1.42)** | 3.74 (0.51) | 3.26 (0.49) | 2.94 (0.42) | |
| 10 | **1.74 (1.14)** | 3.65 (0.50) | 3.17 (0.44) | 2.89 (0.38) | |
| CL | 4.095 | 2.836 | 3.118 | 3.395 | |

error on the $\bar{X}$-$S$ chart, and $\alpha_{\bar{X}}$ and $\alpha_S$ be the type I error on the $\bar{X}$ chart and the $S$ chart, respectively. Then, using the independence of $\bar{X}$ and $S$, the following equation can be obtained.

$$1-\alpha = (1-\alpha_{\bar{X}})(1-\alpha_S)\,.$$

Since it is generally assumed that $\alpha_{\bar{X}} = \alpha_S$, let this value be $\alpha^*$. Then if the overall type I error $\alpha$ is given, $\alpha^*$ can be calculated as

$$\alpha^* = 1-(1-\alpha)^{\frac{1}{2}}.$$

Therefore, the control limits for the $\bar{X}$ chart and the $S$ chart are given by

$$\mathrm{UCL}_{\bar{X}} = \mu_0 + z_{\alpha^*/2}\frac{\sigma_0}{\sqrt{n}}\,, \quad \mathrm{LCL}_{\bar{X}} = \mu_0 - z_{\alpha^*/2}\frac{\sigma_0}{\sqrt{n}}\,, \tag{4.2}$$

and

$$\mathrm{UCL}_S = \frac{\sigma_0}{\sqrt{n-1}}\sqrt{\chi^2_{n-1,\alpha^*/2}}\,, \quad \mathrm{LCL}_S = \frac{\sigma_0}{\sqrt{n-1}}\sqrt{\chi^2_{n-1,1-\alpha^*/2}}\,, \tag{4.3}$$

respectively, where $z_{\alpha^*/2}$ is the upper $\alpha^*/2$ percentage point of the standard normal distribution and $\chi^2_{n-1,\alpha^*/2}$ is the upper $\alpha^*/2$ percentage point of the chi-square distribution with $n-1$ degrees of freedom.

## 4.2. Simulation study

A simulation study is performed to compare the performance of the proposed procedure with existing procedures. Simulation study settings are as follows: since $Z = (X-\xi)/w$ follows a $\mathrm{SN}(0,1,\lambda)$ distribution if $X \sim \mathrm{SN}(\xi, w, \lambda)$, we assume $\xi = 0$ and $w = 1$ without loss of generality. The sample

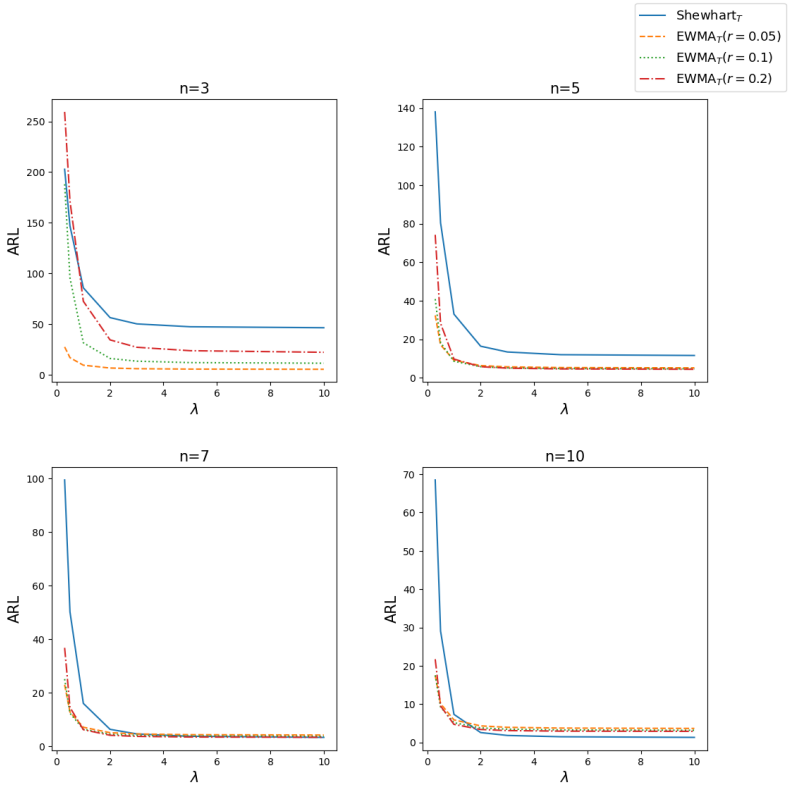


Figure 3: *ARL graphs of the Shewhart$_T$ and EWMA$_T$ charts for various values of $n$ and $\lambda_1$.*

size is set to $n = 3, 5, 7, 10$, and the value of the desired $\text{ARL}_0$ is set to 370.4. The values of $\lambda_1$ range from the relatively small shift of 0.3 up to the relatively large shift of 10. In the case of $\xi = 0$ and $w = 1$, because of symmetry, if the absolute value of $\lambda_1$ is the same, the same ARL value can be obtained. Therefore, only the case of $\lambda_1 > 0$ is considered. The smoothing constant in EWMA$_D$ and EWMA$_T$ charts is set to $r = 0.05, 0.1, 0.2$.

The control limits for Shewhart$_D$, EWMA$_D$, EWMA$_T$ charts ($h_D$, UCL$_{ED}$ in equation (4.1), and UCL$_{ET}$ and LCL$_{ET}$ in equation (3.4), respectively) are obtained so that $\text{ARL}_0$ is approximately 370.4 through simulation. These values can be obtained using the trial and error method. The simulation results shown in this paper are based on 100,000 runs. The control limits for Shewhart$_T$ charts can be obtained using equation (3.3), and the control limits for $\bar{X}$-$S$ charts can be obtained using equations (4.2) and (4.3) to satisfy $\alpha = 0.0027$.

The ARL comparisons of the Shewhart$_D$, EWMA$_D$, $\bar{X}$-$S$, Shewhart$_T$, and EWMA$_T$ charts when $n = 3$, 5, 7, and 10 are shown in Tables 1–4, respectively. In all ARL comparisons, the smallest ARL value appears in bold, and the value in parenthesis represents the standard deviation of the run length. In each table, the row titled CL represents the control limits, $h_D$ for Shewhat$_D$, UCL$_{ED}$ for EWMA$_D$, UCL$_{\bar{X}} = |\text{LCL}_{\bar{X}}|$ for $\bar{X}$, UCL$_S$ and LCL$_S$ for $S$, UCL$_T = |\text{LCL}_T|$ for Shewhart$_T$, and UCL$_{ET} = |\text{LCL}_{ET}|$ for EWMA$_T$ charts. In CL for $\bar{X}$-$S$, the first row represents UCL$_{\bar{X}}$ and the second row represents UCL$_S$ and LCL$_S$.

The results show, as expected, for all the charts we considered, shifts (especially small shifts) are

difficult to detect quickly when $n$ is small, but the performance of detection gradually improves as $n$ increases. Comparing the performance of the charts, it can be seen that for all values of $n$ and all shifts in $\lambda$, the EWMA$_T$ chart based on the proposed statistic demonstrate much better performance compared to the Shewhart$_D$ and EWMA$_D$ charts. Comparing the Shewhart$_D$ chart and the Shewhart$_T$ chart, there is not much difference, but the performance of the Shewhart$_T$ chart appears to be better as $n$ increases. It can be seen that the $\bar{X}$-$S$ chart performs much worse than other control charts in all cases.

When $n$ is small ($n = 3$), the performance of the EWMA$_T$ chart with a small $r$ ($r = 0.05$) is the best. However, as $n$ and the size of shift in $\lambda$ increase, the EWMA$_T$ chart with a larger $r$ performs better. When $n = 10$ and the size of shift is large ($\lambda_1 \geq 2$), the Shewhart$_T$ chart performs best. For a better comparison, ARL graphs of the Shewhart$_T$ and EWMA$_T$ charts are presented in Figure 3.

Li *et al*. (2019) showed that the Shewhart$_D$ chart outperforms the compared procedures when $n$ is not very large ($n \leq 10$). However, when compared to our proposed procedures, the Shewhart$_D$ chart performed relatively poorly in all cases. Therefore, it is confirmed that our proposed procedures are very efficient in detecting a shift in the asymmetry parameter. Based on the overall assessment of the simulation results, it is recommended to use the EWMA$_T$ chart with a small $r$ for cases where $n$ is very small ($n = 3$). For a moderately large $n$ ($n = 5, 7$), it is suggested to use the EWMA$_T$ chart with an adjusted $r$ based on the expected shift size. Additionally, for a large $n$ ($n = 10$) and when the shift size is not small, the use of the Shewhart$_T$ chart is recommended.

## 5. Application to a real process data

In the field of plastics, stamping is a type of cold working that uses hydraulic machinery. It has been observed that the upper vibration data of the hydraulic press initially follows a normal distribution and becomes skewed as the process proceeds. This section aims to compare the performance of the control charts on real production process data.

Stamping is a type of cold working that involves pressing a mold into a workpiece to achieve a desired shape. This technique is commonly used for the mass production of parts made from strong materials, such as automobile parts. Press processing, unlike turning or milling processes, does not generate chips and can produce parts of the same shape on a large scale.

The pressing process is driven by a hydraulic pump, which converts externally supplied mechanical energy into hydraulic energy, that operates the hydraulic system. Due to its structural characteristics, hydraulic pumps generate vibration and noise when operating. As equipment ages, gear wear can cause malfunctions and increase vibration and noise. Reduced hydraulic power can result in insufficient force transfer during the press process and increase the likelihood of defective parts. Additionally, equipment failure due to pump malfunction can disrupt production operations, resulting in up to two to three days of downtime and significant costs.

The data sets used in this study were collected by installing vibration sensors in both the upper and lower sections of the hydraulic pump within the press equipment, as well as current sensors on the power line of the hydraulic pump (The source of the data sets is: Ministry of SMEs and Startups, Korea AI Manufacturing Platform (KAMP), Plastic processing predictive maintenance AI dataset, Smart Manufacturing Innovation Promotion Team, 2022.12.23). Among these data sets, the data obtained from the upper vibration sensor of the hydraulic pump were observed to initially follow a normal distribution under normal operating conditions but then later shift to a SN distribution when the process departs from its normal state. To substantiate these observations, we perform a test for normality using the Shapiro-Wilk normality test and assess the goodness-of-fit to the SN distribution

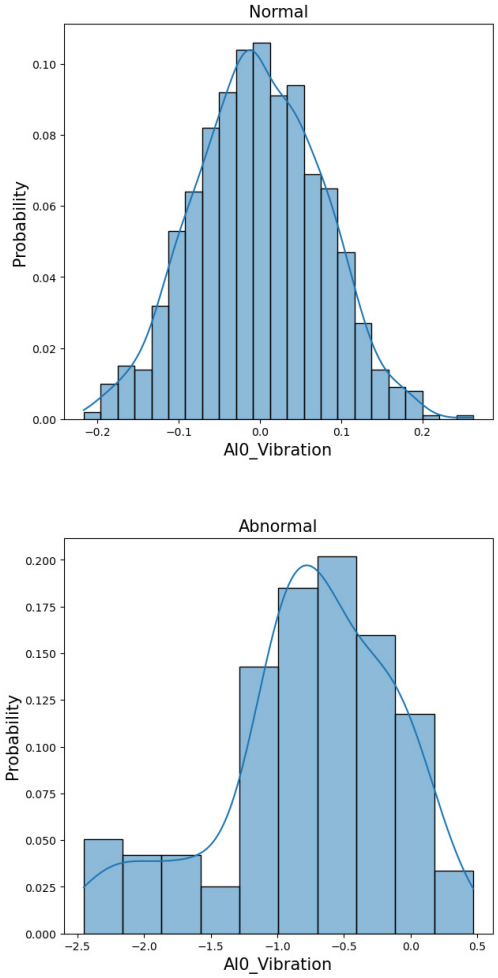


Figure 4: *Histograms of the initial stage of the process (upper) and the later stage of the process (lower).*

using the Kolmogorov-Smirnov (K-S) test.

The Shapiro-Wilk normality test for the normal data set results in a $p$-value of 0.5366, which strongly suggests that the data adheres to the normality assumption during the initial stage of the production process. However, when the process becomes out of control, the $p$-value for the data approaches zero, indicating a departure from the normality assumption. Additionally, the K-S test verifies the suitability of the SN distribution as an appropriate fit to deviant data with a $p$-value exceeding 0.3. As a result, we deduce that the initially normally distributed vibration data becomes skewed. By fitting the data to the SN distribution, the estimated parameter values are $\xi = 0.296$, $w = 0.427$, and $\lambda = -3.528$. As described in the previous sections, the data is normalized using these parameters ($\xi$ and $w$) before calculating the test statistics.

In this context, we applied the Shewhart$_D$, EWMA$_D$, Shewhart$_T$, and EWMA$_T$ charts to evaluate whether these charts are applicable to real data. The out-of-control data set comprises a total of 120 data points, which is then divided into 17 subgroups with each subgroup having 7 data points ($n = 7$).

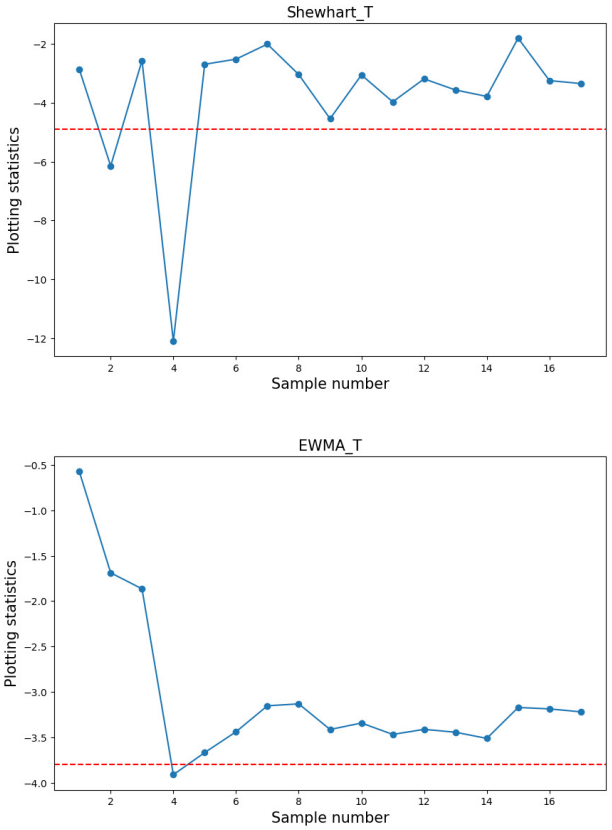


Figure 5: *The Shewhart$_T$ chart (upper) and the EWMA$_T$ chart (lower) for the vibration data.*

A constant value of $r = 0.2$ is used for both the EWMA$_D$ and EWMA$_T$ charts. For $n = 7$, the control limits (CL) for each chart, as detailed in Table 3, are as follows: CL-Shewhart$_D$ = 0.730, CL-EWMA$_D$ = 0.329, CL-Shewhart$_T$ = ±4.906, CL-EWMA$_T$ = ±3.799. The test statistics are calculated for each chart and are shown in Figure 5 along with the corresponding control limits.

The initial signal is detected at the $2^{nd}$ sample for the Shewhart$_T$ chart and at the $4^{th}$ sample for the EWMA$_T$ chart. Both charts based on the distance skewness statistic fail to detect the change until the $17^{th}$ subgroup. As a result, it can be inferred that the proposed procedure shows better sensitivity in detecting skewness change. This is believed to be because, as mentioned earlier, as the data becomes more skewed, the standard deviation of the subsample decreases, which amplifies the effect of the skewness change and increases the possibility of detection.

## 6. Conclusion

To detect shifts in the asymmetry parameter of a SN distribution, this paper proposes a statistic that sensitively reflects shifts, and proposes the Shewhart and EWMA charting procedures based on this statistic. Since the proposed statistic is known to follow the noncentral $t$-distribution, the control limits of Shewhart charts can be easily obtained using this distribution.

To evaluate the performance of the proposed procedure, a simulation study was conducted to com-

pare the Shewhart and EWMA charts based on the distance skewness statistic, which was known to have the best performance in Li *et al*. (2019). Simulation results show that the proposed Shewhart and EWMA charts outperform the Shewhart and EWMA charts of Li *et al*. (2019). Comparing the performance of the proposed Shewhart and EWMA charts, it was found that the EWMA chart shows good performance in most cases, except for cases where the sample size and shift size are very large. Consequently, in practice, one can decide which method to use in accordance with the available sample size and the expected shift size.

In this paper, we considered the situation where only the asymmetry parameter of a SN distribution changes, and the location and scale parameters are constant. However, further studies will be conducted on the procedure to monitor the process when all parameters may change due to undesirable causes.

## References

Arellano-Valle RB, Gomez HW, and Quintana FA (2004). A new class of skew-normal distributions, *Communications in Statistics-Theory and Methods*, **33**, 1465–1480.

Azzalini A (1985). A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics*, **12**, 171–178.

Azzalini A (1986). Further results on a class of distributions which includes the normal ones, *Statistica*, **46**, 199–208.

Azzalini A (2005). The skew-normal distribution and related multivariate families, *Scandinavian Journal of Statistics*, **32**, 159–188.

Chen JT, Gupta AK, and Nguyen TT (2004). The density of the skew normal sample mean and its applications, *Journal of Statistical Computation and Simulation*, **74**, 487–494.

Figueiredo F and Gomes MI (2013). The skew-normal distribution in SPC, *REVSTAT-Statistical Journal*, **11**, 83–104.

Guo B and Wang BX (2015). The design of the ARL-unbiased S2 chart when the in-control variance is estimated, *Quality and Reliability Engineering International*, **31**, 501–511.

Hawkins DM and Deng Q (2009). Combined charts for mean and variance information, *Journal of Quality Technology*, **41**, 415–425.

Henze N (1986). A probabilistic representation of the 'skew-normal' distribution, *Scandinavian Journal of Statistics*, **13**, 271–275.

Knoth S (2015). Run length quantiles of EWMA control charts monitoring normal mean or/and variance, *International Journal of Production Research*, **53**, 4629–4647.

Li C, Mukherjee A, Su Q, and Xie M (2016). Design and implementation of two CUSUM schemes for simultaneously monitoring the process mean and variance with unknown parameters, *Quality and Reliability Engineering International*, **32**, 2961–2975.

Li C, Mukherjee A, Su Q, and Xie M (2019). Some monitoring procedures related to asymmetry parameter of Azzalini's skew-normal model, *REVSTAT-Statistical Journal*, **17**, 1–24.

Li Z, Zou C, Wang Z, and Huwang L (2013). A multivariate sign chart for monitoring process shape parameters, *Journal of Quality Technology*, **45**, 149–165.

Mameli V and Musio M (2013). A generalization of the skew-normal distribution: The beta skew-normal, *Communications in Statistics-Theory and Methods*, **42**, 2229–2244.

McCracken AK, Chakraborti S, and Mukherjee A (2013). Control charts for simultaneous monitoring of unknown mean and variance of normally distributed processes, *Journal of Quality Technology*, **45**, 360–376.

Mukherjee A, Abd-Elfattah AM, and Pukait B (2013). A rule of thumb for testing symmetry about an unknown median against a long right tail, *Journal of Statistical Computation and Simulation*, **84**, 2138–2155.

Peng Y, Xu L, and Reynolds MR Jr (2015). The design of the variable sampling interval generalized likelihood ratio chart for monitoring the process mean, *Quality and Reliability Engineering International*, **31**, 291–296.

Rahman S and Hossain F (2008). A forensic look at groundwater arsenic contamination in Bangladesh, *Environ Forensics*, **9**, 364–374.

Ross GJ and Adams NM (2012). Two nonparametric control charts for detecting arbitrary distribution changes, *Journal of Quality Technology*, **44**, 102–116.

Ryu JH, Wan H, and Kim S (2010). Optimal design of a CUSUM chart for a mean shift of unknown size, *Journal of Quality Technology*, **42**, 311–326.

Sheu SH, Huang CJ, and Hsu TS (2012). Extended maximum generally weighted moving average control chart for monitoring process mean and variability, *Computers and Industrial Engineering*, **62**, 216–225.

Shu L, Yeung HF, and Jiang W (2010). An adaptive CUSUM procedure for signaling process variance changes of unknown sizes, *Journal of Quality Technology*, **42**, 69–85.

Vincent R and Walsh TD (1997). Quantitative measurement of symmetry in CBED patterns, *Ultramicroscopy*, **70**, 83–94.

Wu Z, Yang M, Khoo MBC, and Yu FJ (2010). Optimization designs and performance comparison of two CUSUM schemes for monitoring process shifts in mean and variance, *European Journal of Operational Research*, **205**, 136–150.

Zou C and Tsung F (2010). Likelihood ratio-based distribution-free EWMA control charts, *Journal of Quality Technology*, **42**, 174–196.

*Received July 31, 2023; Revised November 8, 2023; Accepted November 27, 2023*