

Research Article

A New Nonparametric Multivariate Control Scheme for Simultaneous Monitoring Changes in Location and Scale

Jin Yue ^{1,2} and Liu Liu ^{1,2}

¹College of Mathematics and Physics, Chengdu University of Technology, Chengdu 610059, China

²School of Mathematics and VC & VR Key Lab of Sichuan Province, Sichuan Normal University, Chengdu 610068, China

Correspondence should be addressed to Liu Liu; liuliu@sicnu.edu.cn

Received 7 April 2022; Revised 20 June 2022; Accepted 21 June 2022; Published 4 July 2022

Academic Editor: Po-Hsiang Tsui

Copyright © 2022 Jin Yue and Liu Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Real-time monitoring of the breast cancer index is becoming increasingly important. It can help create advances in the diagnosis and treatment of breast cancer. In today's modern medical processes, simultaneously monitoring changes in observations in terms of location and scale are convenient for the implementation of control schemes but can be challenging. In this paper, we consider a new nonparametric control scheme for monitoring location and scale parameters in multivariate processes. The proposed method is easy to implement, and the performance of the proposed control procedure is discussed. Then, we compare the proposed scheme with some competing methods. Simulation results show that the proposed scheme can efficiently detect a range of shifts. The proposed chart can trigger an alert and timely discover the change of the breast cancer index.

1. Introduction

Control schemes play an important role in biosurveillance studies [1–9]. Control schemes have been frequently used for fault detection in quality control with products and health-care monitoring [10–14]. A process should be monitored using statistical means to determine whether a shift occurs, and action should be taken once the process is considered out-of-control (OC) [15–18]. Many researchers have discussed and proposed many useful charts, such as Shewhart charts [19, 20], cumulative sum (CUSUM) charts [21–30], and exponentially weighted moving average (EWMA) charts [31–38], to detect whether there is a change in quality characteristics in a process. These proposed control schemes can be used for data analysis, including control and forecasting, which are useful for fault diagnosis in practice. Most charts require that these observations be univariate and typically assume that these observations follow a normal distribution. Unfortunately, the assumption of multivariate normality is unrealistic in most cases and would lead to a poor performance if underlying assumptions are invalid.

Nonparametric control charts are important in manufacturing and service sectors when samples of observations are nonnormal. Some control schemes are used to monitor high-dimensional processes when we know little about the underlying distribution [39–42]. Most control schemes are designed to monitor location parameters. For example, Liu and Singh [43] introduced several multivariate rank tests based on data depth. Liu [44] used the concept of data depth to propose several new control charts to monitor multivariate process. Data depth provides an efficient metric of the process' performance without using parametric assumptions. In addition, Zou et al. [45] provided a multivariate spatial rank for monitoring high-dimensional processes with unknown parameters. For detecting the location changes in nonparametric multivariate processes, we also recommend the discussions by [46, 47]. To detect the changes in the location and scale of observations simultaneously, several monitoring methods are proposed in the literature, including Mukherjee and Chakraborti [48] and Chowdhury et al. [49]. Recently, Mukherjee and Marozzi [50] consider the sum of the squares of standardized Wilcoxon and the Bradley statistics for monitoring high-

dimensional processes with unknown parameters which is advantageous in simultaneous monitoring of multiple aspects.

Recently, some schemes have been proposed to monitor the changes in location and scale simultaneously using a single chart. Performance advantages of these charts have been clearly established [51]. Lepage [52] discussed a nonparametric two-sample test for location and dispersion. Based on Lepage [52], Mukherjee and Marozzi [51] introduced new circular-grid charts for simultaneous monitoring of process location and process scale based on Lepage-type statistics. Meanwhile, Mukherjee and Marozzi [53] investigated a new single distribution-free Phase-II CUSUM procedure based on the Cucconi statistic for simultaneously monitoring changes in location and scale parameters of a process. In addition, Mukherjee and Sen [54] discussed a distribution-free (nonparametric) Shewhart-Lepage scheme for simultaneous monitoring of location and scale parameters using an adaptive strategy. Li et al. [55] and Shi et al. [56] provided powerful control schemes aimed at simultaneously monitoring the location and the scale parameters of any continuous process. Moreover, Zafar et al. [57] proposed a new parametric memory-type charting structure based on progressive mean under max statistic for the joint monitoring of location and dispersion parameters. Song et al. [58] introduced distribution-free adaptive Shewhart-Lepage-type schemes for simultaneous monitoring of location and scale parameters using information about symmetry and tail weights of the process distribution. Huang et al. [59] proposed a new statistical process monitoring scheme with a double-sampling plan for simultaneously monitoring location and scale shifts. Bai and Li [60] considered monitoring ordinal categorical factors for monitoring which considers shifts in the location or scale parameters of latent variables. For multivariate processes, Cheng and Shiau [61] proposed a distribution-free phase I monitoring scheme for both location and scale parameters based on the multisample Lepage statistic.

Although these literatures contain many control schemes for monitoring location and scale parameters simultaneously, much less focus has been placed on control strategies that simultaneously monitor location and scale parameters in multivariate processes. In this study, we propose a useful and easy-to-implement control scheme for simultaneously monitoring location and scale parameters, which is based on nonparametric location and scale hypothesis testing. Reference samples are denoted as phase I data streams, and test samples are denoted as phase II data streams. One problem is that the size of phase II increases with the number of data streams. Considering this issue, we performed hypothesis testing repeatedly with each new data stream. Thus, the amount of phase II data became a constant for each acquisition time.

The remainder of this paper is organized as follows: In Section 2, we review nonparametric hypothesis testing in detail. In Section 3, we propose a new scheme based on a hypothesis testing statistic for monitoring location and scale parameters. Then, we discuss the proposed method's performance and validity. In Section 4, we perform a simulation-

based comparison to compare the proposed chart with other existing charts. In Section 5, breast cancer data are investigated to describe the performance of the proposed chart. Lastly, we briefly draw conclusions in Section 6.

2. Review of Nonparametric Hypothesis Testing

Hypothesis testing is a form of statistical inference that uses data from a sample to draw conclusions about a population parameter or a population probability distribution, considering reference sample $\{X_{1,t}, X_{2,t}, \dots, X_{m,t}\}$ of size m and test sample $\{Y_{1,t}, Y_{2,t}, \dots, Y_{n,t}\}$ of size n . Thus, null hypothesis $H_0 : \mu_1 = \mu_2, \sigma_1^2 = \sigma_2^2$ versus alternative hypothesis $H_1 : \mu_1 \neq \mu_2$ or $\sigma_1^2 \neq \sigma_2^2$, where μ_1 is the location parameter of reference sample; μ_2 is the location parameter of test sample; σ_1^2 and σ_2^2 are the scale parameters of the reference and test samples, respectively. We can use a reasonable statistical decision procedure to reject the null hypothesis H_0 . In real situations, it is difficult for us to identify the exact distribution of data streams. Therefore, nonparametric hypothesis testing is also introduced, which does not consider the distribution of the original data. For hypothesis testing about the location parameter, Mood [62] proposed the median test, which is based on the rank of each datum. Considering the interaction between the reference and test samples, Wilcoxon [63] and Mann and Whitney [64] introduced the Mann-Whitney-Wilcoxon statistic. In addition, rank-based nonparametric hypothesis testing of scale parameter is used in the literature [65–67].

2.1. Methods for Location Detection. In general, people often check whether there is a change for a given location parameter in a process. We often use the t -statistic under the assumption that the distribution is normal. However, there is a risk in using the t -statistic with unknown population distributions. Thus, some distribution-free statistics have been developed. Brown-Mood median testing is a useful nonparametric method. However, the bilateral test does not yield satisfactory results when $m \neq n$. To use more information about the relative size of the reference sample and test sample, the Wilcoxon rank-sum test was developed. We assume that a reference sample of size m and test sample of size n are given, and we let $N = m + n$. Considering the pooled sample $\{X_{1,t}, X_{2,t}, \dots, X_{m,t}, Y_{1,t}, Y_{2,t}, \dots, Y_{n,t}\}$ at time t , Mann and Whitney [64] developed the Mann-Whitney statistic as follows:

$$W_{1,t} = \sum_{j=1}^n \sum_{i=1}^m I\{X_{i,t} < Y_{j,t}\}. \quad (1)$$

Therefore, the Wilcoxon rank-sum statistic is

$$W_{2,t} = W_{1,t} + \frac{n(n+1)}{2}, \quad (2)$$

where $W_{2,t} = \sum_{i=1}^n R_{i,t}$, and $R_{i,t}$ is the rank of $Y_{i,t}$ in the pooled sample $\{X_{1,t}, X_{2,t}, \dots, X_{m,t}, Y_{1,t}, Y_{2,t}, \dots, Y_{n,t}\}$. $E(W_{2,t}|H_0) = n(n+1)/2$.

$N + 1)/2$. It can be seen that [68]

$$\begin{aligned} E(W_{1,t}|H_0) &= \frac{mn}{2}, \\ \text{Var}(W_{1,t}|H_0) &= \frac{mn(N+1)}{12}. \end{aligned} \quad (3)$$

Under the null hypothesis, we also calculate the approximate normal statistic when the sample N is sufficiently large.

2.2. Methods for Scale Detecting. A location parameter typically describes the position of a distribution, and a scale parameter is also an important characteristic that describes a distribution. When the distribution of observations is unknown, some distribution-free methods are typically used. Given a two-phase independent sample $\{X_{1,t}, X_{2,t}, \dots, X_{m,t}\} \sim F(\mu_1, \sigma_1^2)$ and $\{Y_{1,t}, Y_{2,t}, \dots, Y_{n,t}\} \sim F(\mu_2, \sigma_2^2)$. We assume that the location parameters of the two samples are equal ($\mu_1 = \mu_2$). Based on the Mann-Whitney statistic, Siegel and Tukey [65] proposed the Siegel-Tukey statistic. The implementation design of this statistic consists of the following steps: (1) mix the two samples $\{X_{1,t}, X_{2,t}, \dots, X_{m,t}, Y_{1,t}, Y_{2,t}, \dots, Y_{n,t}\}$ in ascending order, $Q_{(1),t}, Q_{(2),t}, \dots, Q_{(m+n),t}$; (2) assign the rank $R'_{i,t}$ of $Q_{(1),t}, Q_{(2),t}, \dots, Q_{(m+n),t}$ as shown in Table 1; and (3) calculate the $S_t = \sum_{i=1}^n (R'_{i,t} - n(n+1)/2)$; $R'_{i,t}$ represents the rank of $Y_{i,t}$.

Mood [62] also provided a useful test statistic for scale parameters. As before, we consider two sequences of $\{X_{1,t}, X_{2,t}, \dots, X_{m,t}\} \sim G(\mu_1, \sigma_1^2)$ and $\{Y_{1,t}, Y_{2,t}, \dots, Y_{n,t}\} \sim G(\mu_2, \sigma_2^2)$, where $\mu_1 = \mu_2$. The Mood statistic can be described as follows:

$$\text{MD}_t = \sum_{i=1}^n (R_{i,t} - E(R_{i,t}))^2, \quad (4)$$

where $R_{i,t}$ is the rank of $Y_{i,t}$, $i = 1, 2, \dots, n$, in sample $\{X_{1,t}, X_{2,t}, \dots, X_{m,t}, Y_{1,t}, Y_{2,t}, \dots, Y_{n,t}\}$ of size $N (= m + n)$. For $m, n \rightarrow +\infty$ and $m/N \rightarrow \text{constant } C$. Additionally [68],

$$\begin{aligned} E(\text{MD}_t|H_0) &= \frac{n(N^2 - 1)}{12}, \\ \text{Var}(\text{MD}_t|H_0) &= \frac{mn(N+1)(N^2 - 4)}{180}. \end{aligned} \quad (5)$$

Filgner and Killeen [69] also introduced a test statistic for scale parameters that is based on the absolute rank. The statistic is defined as

$$F_t = \sum_{i=1}^n R_{i,t}. \quad (6)$$

$R_{i,t}$ is the rank of $V_{i,t}^R$ in pooled sample $\{V_{1,t}^R, V_{2,t}^R, \dots, V_{m,t}^R, V_{1,t}^T, V_{2,t}^T, \dots, V_{n,t}^T\}$, where $V_{i,t}^R = |X_{i,t} - M|$, $V_{i,t}^T = |Y_{i,t} - M|$. M represents the median of the sample $\{X_{1,t}, X_{2,t}, \dots, X_{m,t}, Y_{1,t}, Y_{2,t}, \dots, Y_{n,t}\}$. F_t has the distribution of Wilcoxon's

rank-sum statistic under the null hypothesis. Therefore,

$$\begin{aligned} E(F_t|H_0) &= \frac{mn}{2}, \\ \text{Var}(F_t|H_0) &= \frac{mn(N+1)}{12}. \end{aligned} \quad (7)$$

3. Proposed Monitoring Strategy

We assume that there are m -independent observations from an unknown multivariate continuous distribution with dimensionality p . We assume that independent observations, \mathbf{X}_i , follow the model below:

$$\mathbf{X}_i \sim \begin{cases} \mathbf{G}_p(\mu_0, \Sigma_0), & \text{if } i = 1, 2, \dots, \tau, \\ \mathbf{G}_p(\mu_1, \Sigma_1), & \text{if } i = \tau + 1, \tau + 2, \dots, \end{cases} \quad (8)$$

where μ_0 and μ_1 are the in-control (IC) location vector and the OC location vector, respectively; Σ_0 and Σ_1 represent the IC covariance matrix and the OC covariance matrix, respectively, where $(\mu_0, \Sigma_0) \neq (\mu_1, \Sigma_1)$; τ represents an unknown change point; and $\mathbf{G}_p(\cdot)$ is an unknown continuous distribution function. In phase I, we assume that the IC sample of size m is given at time t , $R = \{\mathbf{X}_{1,t}, \mathbf{X}_{2,t}, \dots, \mathbf{X}_{i,t}, \dots, \mathbf{X}_{m,t}\}$ where $\mathbf{X}_{i,t} = \{X_{1,i,t}, X_{2,i,t}, \dots, X_{p,i,t}\}'$, $i = 1, 2, \dots, m$. In phase II, $T = \{\mathbf{Y}_{1,t}, \mathbf{Y}_{2,t}, \dots, \mathbf{Y}_{n,t}\}$ of size n is obtained. After the phase I sample R is analyzed, the phase II sample T is monitored.

Inspired by Mukherjee and Marozzi [50] for multivariate processes, we consider the p -dimension statistic of the Euclidean distance of new observations and the mean vector of phase I data, $\mathbf{X}_{i,t}$, $i = 1, 2, \dots, m$. That is, $D_{i,t}^R = \|\mathbf{X}_{i,t} - \bar{\mathbf{X}}\|$ and $D_{i,t}^T = \|\mathbf{Y}_{i,t} - \bar{\mathbf{X}}\|$, where $\bar{\mathbf{X}} = (1/m) \sum_{i=1}^m \mathbf{X}_{i,t}$. Now, a univariate phase II sequence is obtained, $\{D_{1,t}^R, D_{2,t}^R, \dots, D_{m,t}^R, D_{1,t}^T, D_{2,t}^T, \dots, D_{n,t}^T\}$. Then, a Shewhart-type chart for monitoring location changes that is based on the Wilcoxon rank-sum statistic (i.e., S-W chart) can be constructed. The statistic of the S-W chart is $Z_{W,t} = (W_{1,t} - mn/2) / \sqrt{mn(N+1)/12}$ with upper control limit (UCL)

$$\text{UCL} = E(Z_{W,t}|H_0) + L\sqrt{\text{Var}(Z_{W,t}|H_0)}, \quad (9)$$

and lower control limit (LCL)

$$\text{LCL} = E(Z_{W,t}|H_0) - L\sqrt{\text{Var}(Z_{W,t}|H_0)}, \quad (10)$$

where L is an unknown constant. The Shewhart-type chart can be constructed based on three other types of hypothesis statistics for the scale parameter. The S-ST chart (i.e., the Shewhart-type chart based on the Siegel-Tukey statistic) is calculated using $Z_{ST,t} = (S_t - E(S_t|H_0)) / \sqrt{\text{Var}(S_t|H_0)}$ with $\text{UCL} = E(Z_{ST,t}|H_0) + L\sqrt{\text{Var}(Z_{ST,t}|H_0)}$ and $\text{LCL} = E(Z_{ST,t}|H_0) - L\sqrt{\text{Var}(Z_{ST,t}|H_0)}$. The S-MD chart (i.e., the Shewhart-type chart based on the mood statistic) is given

TABLE 1: Rank of $Q_{(1),t}, Q_{(2),t}, \dots, Q_{(m+n),t}$.

Data	$Q_{(1),t}$	$Q_{(2),t}$	$Q_{(3),t}$	$Q_{(4),t}$	\dots	$Q_{(m+n-3),t}$	$Q_{(m+n-2),t}$	$Q_{(m+n-1),t}$	$Q_{(m+n),t}$
Rank	1	4	5	8	\dots	7	6	3	2

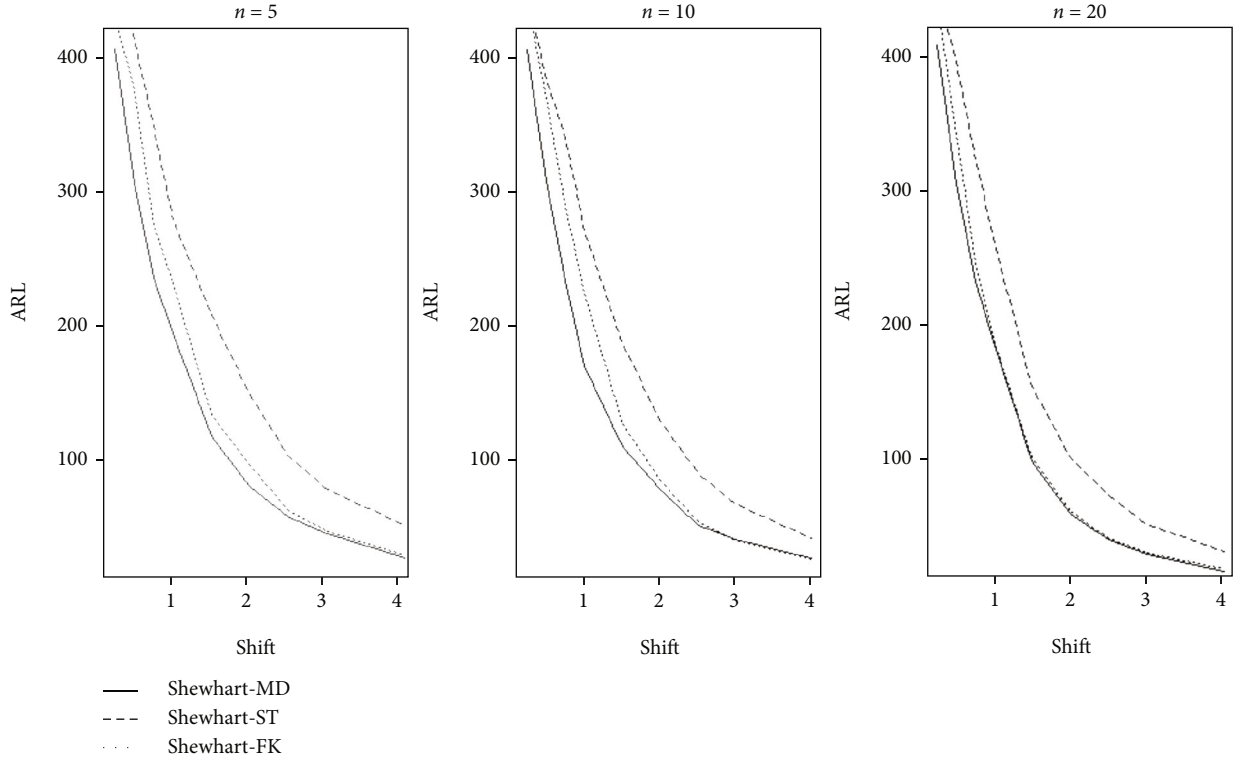


FIGURE 1: Comparison of the three Shewhart-type schemes when detecting changes in scale.

as follows: $Z_{MD,t} = (MD_t - E(MD_t|H_0)) / \sqrt{\text{Var}(MD_t|H_0)}$ with $UCL = E(Z_{MD,t}|H_0) + L\sqrt{\text{Var}(Z_{MD,t}|H_0)}$ and $LCL = E(Z_{MD,t}|H_0) - L\sqrt{\text{Var}(Z_{MD,t}|H_0)}$. The S-FK chart (i.e., the Shewhart-type chart based on the Filgner-Killeen statistic) is given by $Z_{FK,t} = (F_t - E(F_t|H_0)) / \sqrt{\text{Var}(F_t|H_0)}$ with $UCL = E(Z_{FK,t}|H_0) + L\sqrt{\text{Var}(Z_{FK,t}|H_0)}$, and $LCL = E(Z_{FK,t}|H_0) - L\sqrt{\text{Var}(Z_{FK,t}|H_0)}$.

We then use the average run length (ARL) to evaluate the performance of these methods. ARL is the number of points that, on average, will be plotted on a control chart before an OC condition occurs. If the process is IC, $ARL_0 = 1/\alpha$; otherwise, $ARL_1 = 1/(1 - \beta)$ when the process is OC. In addition, α is the probability of a type I error occurring, and β is the probability of a type II error occurring. Therefore, we typically fix IC ARL, which is denoted as ARL_0 , and compare the OC ARL, which is denoted as ARL_1 . A small ARL_1 is considered better. Figure 1 shows the OC ARL of the S-ST, S-MD, and S-FK charts. We let $m = 50$, $n = \{5, 10, 20\}$, and $p = 4$ under the multivariate Gaussian distribution with expectations μ_0 and the variance matrix, Σ_0 . For a fair comparison, we set $ARL_0 = 500$ for all control schemes. Figure 1 shows the OC ARL of the three Shewhart-type schemes when detecting scale parameters.

Figure 1 shows that the S-MD chart's performance is better than the other charts when detecting a range of scale shifts.

When calculating the Mahalanobis distance, the sample population must exceed the sample dimension; otherwise, the inverse matrix of the population sample covariance matrix obtained does not exist. Thus, the Mahalanobis distance sometimes fails to meet practical requirements. It is also not appropriate to simply use the Euclidean distance to reduce the dimensionality of high-dimensional data, because this process would equate the differences between different data attributes (i.e., the dimensions of each index or variable). The standardized Euclidean distance is an improvement strategy that can overcome the shortcoming of the simple Euclidean distance. Since the distribution of each dimension component of the data is different, the first to "standardize" each component to the associated mean and variance are equal.

Mukherjee and Marozzi [50] consider the sum of the squares of standardized Wilcoxon and Bradley statistics for monitoring high-dimensional processes with unknown parameters. Inspired by Mukherjee and Marozzi [50], we combine the idea of control schemes and hypothesis testing to propose an effective control scheme that simultaneously monitors expectation and variance. Based on this analysis, we propose an alternative control scheme, whose statistic is

TABLE 2: OC ARL values of these charts for various m and n when zero-state $ARL_0 = 500$ with the IC distribution $N(\mu_0, \Sigma_0)$.

m	Shifts	S-W	S-MD	Proposed	S-W	S-MD	Proposed	S-W	S-MD	Proposed
		$n = 5$			$n = 10$			$n = 20$		
50	$\delta = 0.5, \sigma = 2$	85.5	161.2	69.1	63.7	169.6	48	48.6	251	42.8
	$\delta = 0.5, \sigma = 4$	25.1	58.1	16.6	15.1	62.3	11.8	10.8	129.7	9.9
	$\delta = 1, \sigma = 2$	26.3	71.7	25.3	16	67.8	15.8	16.9	180.4	16.5
	$\delta = 1, \sigma = 4$	9.8	28.9	9.6	6.5	36.1	6.4	5.5	90	5.3
		$n = 5$			$n = 20$			$n = 50$		
100	$\delta = 0.5, \sigma = 2$	69.1	12.6	57.6	22.9	110.7	21.2	17.7	180.7	16.2
	$\delta = 0.5, \sigma = 4$	16.9	34.4	12.2	5.6	27.8	4.6	4.8	72.1	4.1
	$\delta = 1, \sigma = 2$	20.7	42.8	18.5	7.3	40.5	6.9	8	114.8	7.3
	$\delta = 1, \sigma = 4$	7.9	14.6	7.5	3.1	11.7	3	2.5	47.4	2.5
		$n = 5$			$n = 20$			$n = 50$		
200	$\delta = 0.5, \sigma = 2$	58.3	96.7	49.8	16.7	62.7	14.2	7.8	10.6	7.3
	$\delta = 0.5, \sigma = 4$	14.9	23.3	9.8	4.1	15	3.3	2.5	10.6	2.5
	$\delta = 1, \sigma = 2$	19	28.3	18.7	4.4	15	4.1	3.5	20.3	3.3
	$\delta = 1, \sigma = 4$	6.7	10.5	6.1	2.3	4.3	2.3	2	4.6	2

as follows:

$$Z_t = Z_{W,t} + Z_{MD,t}, \quad (11)$$

with

$$\begin{aligned} \text{UCL} &= E(Z_t|H_0) + L\sqrt{\text{Var}(Z_t|H_0)}, \\ \text{LCL} &= E(Z_t|H_0) - L\sqrt{\text{Var}(Z_t|H_0)}. \end{aligned} \quad (12)$$

The term asymptotic distribution is used in the sense of convergence in law when $m \rightarrow \infty$ and $n \rightarrow \infty$ with the ratio m/N constant [52]. Under H_0 , the statistics $Z_{W,t}$ and $Z_{MD,t}$ are uncorrelated for all m and n . Since, for all m and n ,

$$\begin{aligned} E(W_{2,t}MD_t|H_0) &= E\left[\sum_{i=1}^n \left(R_{i,t} - \frac{N+1}{2}\right)^2 \sum_{j=1}^n R_{j,t}|H_0\right], \\ E\left(\sum_{j=1}^n R_{j,t}|H_0\right) &= \frac{n(N+1)}{2}, \\ E\left(\sum_{i=1}^n \sum_{j=1}^n R_{i,t}R_{j,t}|H_0\right) &= \frac{n^2(N+1)(3N+2)}{12}, \\ E\left(\sum_{i=1}^n \sum_{j=1}^n R_{i,t}^2 R_{j,t}|H_0\right) &= \frac{n^2N(N+1)^2}{6}. \end{aligned} \quad (13)$$

Thus, we have

$$E(W_{2,t}MD_t|H_0) = \frac{n^2(N-1)(N+1)^2}{24}. \quad (14)$$

Equality (14) is the product of $E(W_{2,t}|H_0)$ and $E(MD_t|H_0)$. Therefore,

$$E(W_{1,t}MD_t|H_0) = E(W_{1,t}|H_0)E(MD_t|H_0). \quad (15)$$

It is obvious that

$$E(Z_{W,t}Z_{MD,t}|H_0) = E(Z_{W,t}|H_0)E(Z_{MD,t}|H_0). \quad (16)$$

Under H_0 , $Z_{W,t} = (W_{1,t} - E(W_{1,t}|H_0))/\sqrt{\text{Var}(W_{1,t}|H_0)} \rightarrow N(0, 1)$ and $Z_{MD,t} = (MD_t - E(MD_t|H_0))/\sqrt{\text{Var}(MD_t|H_0)} \rightarrow N(0, 1)$ with $m \rightarrow \infty$, $n \rightarrow \infty$, and the ratio m/N constant.

4. Performance Evaluation

In this section, we compare the performances of these charts with different reference sample sizes m and test sample sizes n when shifts occur. We assume that the t th future observation, \mathbf{X}_t , is collected over time using the following multivariate model:

$$\mathbf{X}_t \sim \begin{cases} \mathbf{N}_p(\mu_0, \Sigma_0), & \text{for } t = 1, 2, \dots, \tau, \\ \mathbf{N}_p(\mu_1, \Sigma_1), & \text{for } t = \tau + 1, \tau + 2, \dots, \end{cases} \quad (17)$$

where $\mu_0 = (0, 0, 0, 0)$, $\mu_1 = (0, 0, \delta, \delta)$, and Σ_0 represents the 4×4 identity matrix. We let $\tau = 50$ and dimensionality $p = 4$. Table 2 shows the OC ARL of these charts. Table 3 presents the OC ARL of these charts when there is a correlation between variables:

$$\mathbf{X}_t \sim \begin{cases} \mathbf{N}_p(\mu_0, \Sigma_2), & \text{for } t = 1, 2, \dots, \tau, \\ \mathbf{N}_p(\mu_1, \Sigma_3), & \text{for } t = \tau + 1, \tau + 2, \dots, \end{cases} \quad (18)$$

TABLE 3: OC ARL values of these charts for various m and n when zero-state $ARL_0 = 500$ with the IC distribution $N(\mu_0, \Sigma_2)$.

m	Shifts	S-W	S-MD	Proposed	S-W	S-MD	Proposed	S-W	S-MD	Proposed
		$n = 5$			$n = 10$			$n = 20$		
50	$\delta = 0.5, \sigma = 2$	96	182.6	70.7	63.5	189.9	53.3	53	274	49.9
	$\delta = 0.5, \sigma = 4$	24.7	68.4	19.3	15.9	72.7	12.6	11.2	147	10.9
	$\delta = 1, \sigma = 2$	31.5	85.7	30.2	19.3	101.5	18.8	19	193.3	18.8
	$\delta = 1, \sigma = 4$	11.2	30.8	11	7.1	40.6	6.8	6.1	103.1	6.1
		$n = 5$			$n = 20$			$n = 50$		
100	$\delta = 0.5, \sigma = 2$	73	140.9	67.7	23.9	33.2	22.8	22.5	92.2	22.3
	$\delta = 0.5, \sigma = 4$	18.4	39.5	14	5.6	33.2	5.1	3.9	92.2	3.7
	$\delta = 1, \sigma = 2$	27.8	47	24.1	5.9	13.8	5.8	8.2	143.2	8.2
	$\delta = 1, \sigma = 4$	17.7	16.5	22.1	4.8	14.4	4.8	4.5	56.6	4.2
		$n = 5$			$n = 20$			$n = 50$		
200	$\delta = 0.5, \sigma = 2$	65.9	120.3	57.9	16.2	79	16.2	8.8	77.8	8.4
	$\delta = 0.5, \sigma = 4$	15.5	27.8	11.3	4	12.2	3.4	2.5	13.4	2.5
	$\delta = 1, \sigma = 2$	18.9	34.8	17.6	4.7	16.6	4.3	2.9	23.8	2.7
	$\delta = 1, \sigma = 4$	7.5	12.6	7.5	2.4	4.7	2.4	2	5.5	2

TABLE 4: OC ARL values of these charts for various m and n when zero-state $ARL_0 = 500$ with the IC distribution Weibull(1, 1).

m	δ	S-W	S-MD	Proposed	S-W	S-MD	Proposed	S-W	S-MD	Proposed
		$n = 5$			$n = 10$			$n = 20$		
50	2	5.5	10.3	4	3.7	6.3	2.8	2.6	5.2	2.3
	4	2.3	2.5	2	2.1	2.1	2	2	2.1	2
	6	2.1	2.1	2	2	2	2	2	2	2
		$n = 5$			$n = 20$			$n = 50$		
100	2	5	7.2	3.8	2.3	2.8	2.1	2	2.4	2
	4	2.2	2.2	2	2	2	2	2	2	2
	6	2	2	2	2	2	2	2	2	2
		$n = 5$			$n = 20$			$n = 50$		
200	2	4.6	5.8	3.4	2.2	2.3	2.2	2	2	2
	4	2.2	2.1	2	2	2	2	2	2	2
	6	2	2	2	2	2	2	2	2	2

where

$$\Sigma_2 = \begin{pmatrix} 1 & 0.2 & 0.2 & 0.2 \\ 0.2 & 1 & 0.2 & 0.2 \\ 0.2 & 0.2 & 1 & 0.2 \\ 0.2 & 0.2 & 0.2 & 1 \end{pmatrix}, \tag{19}$$

$$\Sigma_3 = \begin{pmatrix} 1 + \sigma & 0.2 & 0.2 & 0.2 \\ 0.2 & 1 & 0.2 & 0.2 \\ 0.2 & 0.2 & 1 & 0.2 \\ 0.2 & 0.2 & 0.2 & 1 \end{pmatrix}.$$

The Weibull type of distributional changes for detecting general distributional changes is shown in Table 4, where Weibull(θ_1, θ_2) represents the Weibull distribution with the shape parameter θ_1 and the scale parameter θ_2 . The IC distribution is Weibull(1, 1), and the OC distribution is Weibull(1, 1 + δ). We also consider the three types of general changes (multivariate t with 3 df , multivariate exponential, and multivariate gamma distributions) in Table 5. Tables 2–5 show that the proposed method performs well for detecting a range of shifts.

5. Illustration

5.1. Data Source. To describe the proposed method, we analyze a real clinical case. Samples arrive periodically as Dr.

TABLE 5: OC ARL values of these charts for various n when $m = 100$ and zero-state $ARL_0 = 500$ under other types of distribution.

τ	Type	S-W	S-MD	Proposed	S-W	S-MD	Proposed	S-W	S-MD	Proposed
50		$n = 5$			$n = 20$			$n = 50$		
	1	43.1	126.1	37.5	10.2	103.2	9.4	4.5	252.1	6.9
	2	13.3	64.4	8.1	4.7	5.5	2.8	3.6	10.7	2.4
	3	11.8	59	7.7	4.7	5.5	2.8	3.5	10.3	2.4
	1	43	129.6	41.1	10.1	104.8	9.42	4.4	238.9	6.7
	2	12.7	65.4	7.7	4.8	5.8	2.7	3.4	10.5	2.4
100		$n = 5$			$n = 20$			$n = 50$		
	1	42.9	132.9	39.1	10.7	109.7	9.2	4.7	243.3	6.9
	2	13.4	61.5	8.1	4.8	5.7	2.8	3.4	10.2	2.3
	3	12	58.8	8	4.7	5.7	2.8	3.5	10.6	2.4
	1	42.9	132.9	39.1	10.7	109.7	9.2	4.7	243.3	6.9
	2	13.4	61.5	8.1	4.8	5.7	2.8	3.4	10.2	2.3
200		$n = 5$			$n = 20$			$n = 50$		
	1	42.9	132.9	39.1	10.7	109.7	9.2	4.7	243.3	6.9
	2	13.4	61.5	8.1	4.8	5.7	2.8	3.4	10.2	2.3
	3	12	58.8	8	4.7	5.7	2.8	3.5	10.6	2.4
	1	42.9	132.9	39.1	10.7	109.7	9.2	4.7	243.3	6.9
	2	13.4	61.5	8.1	4.8	5.7	2.8	3.4	10.2	2.3

1: multivariate t with 3 df distribution; 2: multivariate gamma distribution; 3: multivariate exponential distribution.

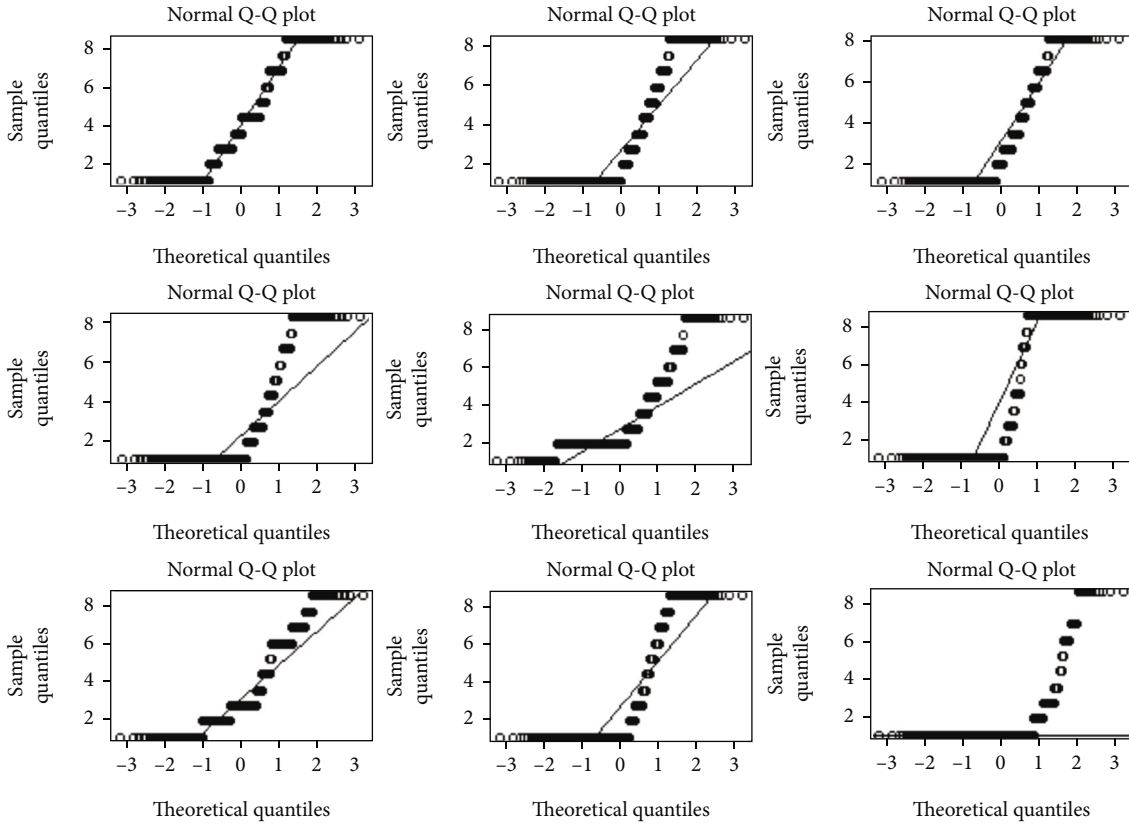


FIGURE 2: Corresponding normal Q-Q plot of the breast cancer data.

Wolberg reports in his clinical cases. The database therefore reflects this chronological grouping of the data. For each of the 599 clinical cases, several clinical features were observed or measured. Quantitative attributes including clump thickness, uniformity of cell size, uniformity of cell shape, marginal adhesion, single epithelial cell size, bare nuclei, bland chromatin, normal nucleoli, and mitoses. The datasets are publicly available in the “Breast Cancer Wisconsin (Original) Data Set” of the UCI Machine Learning Repos-

itory and can be downloaded from the website <http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Original%29>. Breast cancer screening is an important strategy to allow for early detection and ensure a greater probability of having a good outcome in treatment. More details about these datasets can be related to [70–73]. In this work, we aim to monitor the Breast Cancer Wisconsin Data Set and identify whether there is a shift in a process.

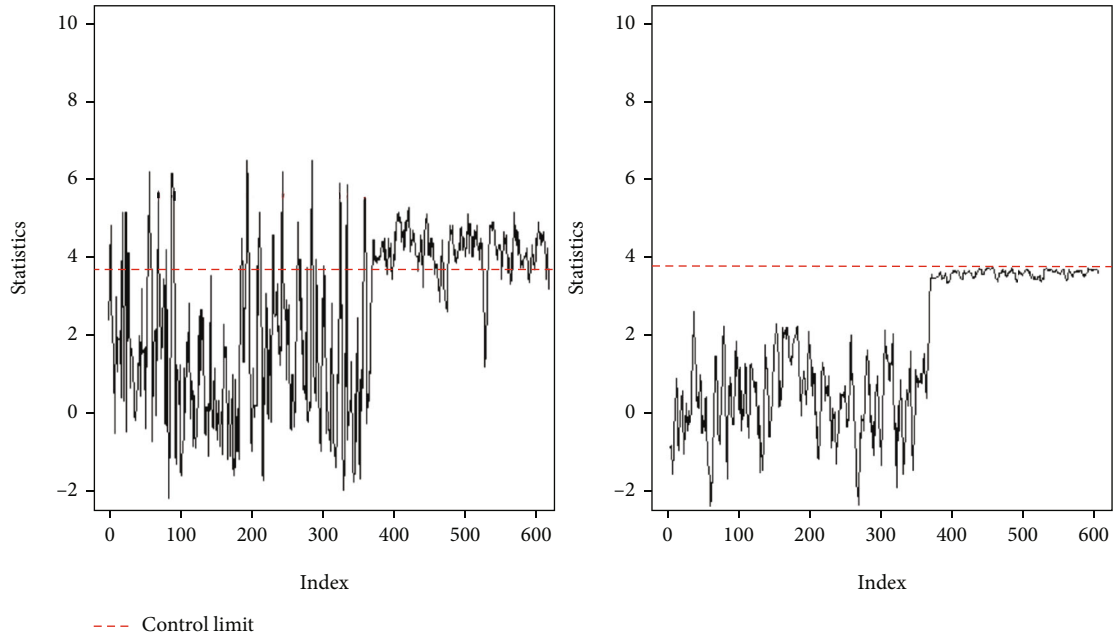


FIGURE 3: (a) S-W chart for monitoring breast cancer data. (b) S-MD chart for monitoring breast cancer data.

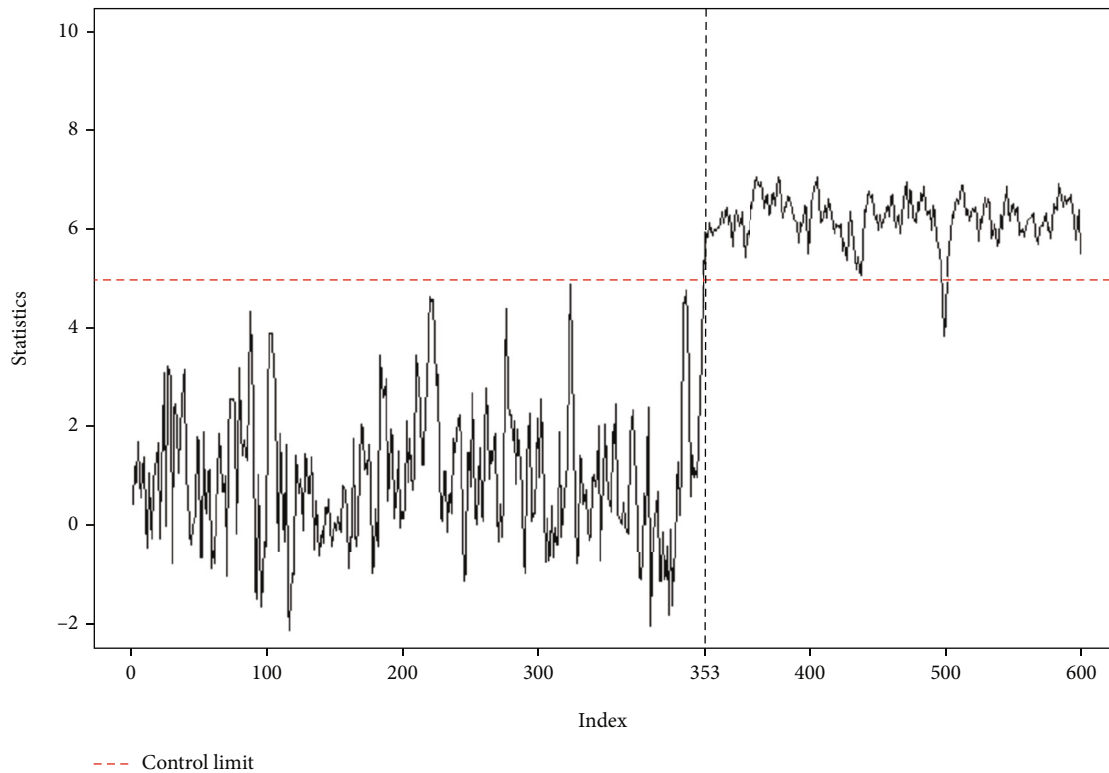


FIGURE 4: The proposed chart for monitoring breast cancer data.

5.2. *Data Analysis.* A quantile-quantile (Q-Q) plot of each index, including 599 historical observations, is shown in Figure 2, which highlights that the normality assumption is invalid, which leads us to reject the null hypothesis that the data are normally distributed. Thus, we use the proposed

distribution-free control scheme to monitor the breast cancer data.

We let $m = 100$ and $n = 5$. We use the 1–350 IC data to find the control limits of the S-W chart, S-MD chart, and proposed chart. For a fair comparison, the IC ARL of all

control charts is set equal to 400, and the remaining 249 breast cancer data are monitored. The curves of the S-W and S-MD charts of the monitored banknote authentication data are shown in Figure 3, which indicates that the S-W chart produces a false alarm when the process is IC; conversely, the S-MD chart produces no OC signal when the process is OC. Figure 4 shows the proposed chart for monitoring breast cancer data and shows that the statistic of the proposed chart falls out of the control limits after 353 observations. Compared with the S-W and S-MD charts, the proposed chart can detect a shift more accurately and earlier than the other charts.

6. Conclusions and Discussion

This paper provided a new control scheme for detecting location and scale changes. Inspired by Mukherjee and Marozzi [50], we proposed an effective control chart that simultaneously monitors changes in both location and scale. In this paper, Breast Cancer Wisconsin Data Sets are provided by using the proposed method. Spectral analysis is also reviewed and conducted to investigate the periodicities of shorter time series, and then, nonlinear least squares fitting is used for fitting analysis. The real-data example shows that the proposed scheme performed well for detecting process changes. In this study, we mainly considered the standard Euclidean distance to reduce the dimensionality of high-dimensional data; the other methods of dimensionality reduction still need to be investigated in more detail.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest statement.

Authors' Contributions

Liu Liu and Jin Yue designed the study and performed the research, Jin Yue discussed the experiment and the related issues in the data analysis parts, and Jin Yue wrote the manuscript. Liu Liu and Jin Yue reviewed the manuscript.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 12075162), the Sichuan Sciences and Technology Program (No. 2020YJ0357), and the VC & VR Key Lab of Sichuan Province.

References

- [1] R. Sarfstein, K. Nagaraj, D. LeRoith, and H. Werner, "Differential effects of insulin and IGF1 receptors on ERK and AKT subcellular distribution in breast cancer cells," *Cell*, vol. 8, no. 12, p. 1499, 2019.
- [2] M. Dettin, E. Sieni, A. Zamuner et al., "A novel 3D scaffold for cell growth to assess electroporation efficacy," *Cell*, vol. 8, no. 11, p. 1470, 2019.
- [3] F. Schochter, T. W. P. Friedl, A. deGregorio et al., "Are circulating tumor cells (CTCs) ready for clinical use in breast cancer? An overview of completed and ongoing trials using CTCs for clinical treatment decisions," *Cell*, vol. 8, no. 11, p. 1412, 2019.
- [4] S. H. Steiner and M. Jones, "Risk-adjusted survival time monitoring with an updating exponentially weighted moving average (EWMA) control chart," *Statistics in Medicine*, vol. 29, no. 4, pp. 444–454, 2010.
- [5] O. A. Grigg and V. T. Farewell, "A risk-adjusted sets method for monitoring adverse medical outcomes," *Statistics in Medicine*, vol. 23, no. 10, pp. 1593–1602, 2004.
- [6] O. A. Grigg and D. J. Spiegelhalter, "A simple risk-adjusted exponentially weighted moving average," *Journal of the American Statistical Association*, vol. 102, no. 477, pp. 140–152, 2007.
- [7] S. Das, P. Duggal, R. Roy et al., "Identification of hot and cold spots in genome of Mycobacterium tuberculosis using Shewhart control charts," *Scientific Reports*, vol. 2, no. 1, p. 297, 2012.
- [8] L. Liu, B. Chen, J. Zhang, and X. Zi, "Adaptive phase II non-parametric EWMA control chart with variable sampling interval," *Quality and Reliability Engineering International*, vol. 31, no. 1, pp. 15–26, 2015.
- [9] E. Monmasson, L. Idkhajine, M. Cirstea, I. Bahri, A. Tisan, and M. W. Naouar, "FPGAs in industrial control applications," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 2, pp. 224–243, 2011.
- [10] S. W. Roberts, "Control chart tests based on geometric moving averages," *Technometrics*, vol. 1, no. 3, pp. 239–250, 1959.
- [11] D. Spiegelhalter, C. S. Johnson, M. Bardsley, I. Blunt, C. Wood, and O. Grigg, "Statistical methods for healthcare regulation: rating, screening and surveillance," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 175, pp. 1–47, 2011.
- [12] K. H. Ong, C. M. Harvey, R. L. Shehab, J. D. Dechert, and A. Darisipudi, "The effects of three statistical control charts on task performance," *Production Planning and Control*, vol. 15, no. 3, pp. 313–323, 2004.
- [13] L. S. Nelson, "The Shewhart control chart—tests for special causes," *Journal of Quality Technology*, vol. 16, no. 4, pp. 237–239, 1984.
- [14] D. A. Cook, S. H. Steiner, R. J. Cook, V. Farewell, and A. P. Morton, "Monitoring the evolutionary process of quality: risk-adjusted charting to track outcomes in intensive care," *Critical Care Medicine*, vol. 31, no. 6, pp. 1676–1682, 2003.
- [15] W. Tian, H. You, K. Gu, C. Zhang, and X. Jia, "Two-level nested control chart for batch process in the semiconductor manufacturing," *IEEE Transactions on Semiconductor Manufacturing*, vol. 29, no. 4, pp. 399–410, 2016.
- [16] N. A. Adegoke, M. Riaz, K. O. Ganiyu, and S. A. Abbasi, "One-sided and two one-sided multivariate homogeneously weighted moving charts for monitoring process mean," *IEEE Access*, vol. 9, pp. 80388–80404, 2021.
- [17] Y. Wu, S. Younas, K. Abbas, A. Ali, and S. A. Khan, "Monitoring reliability for three-parameter Frechet distribution using control charts," *IEEE Access*, vol. 8, pp. 71245–71253, 2020.
- [18] M. Aslam, R. Bhattacharya, and M. S. Aldosari, "Design of control chart in presence of hybrid censoring scheme," *IEEE Access*, vol. 6, pp. 14895–14907, 2018.

- [19] A. Faraz, E. M. Saniga, and C. Heuchenne, "Shewhart control charts for monitoring reliability with Weibull lifetimes," *Quality and Reliability Engineering International*, vol. 31, no. 8, pp. 1565–1574, 2015.
- [20] Z. Khan, M. Gulistan, W. Chammam, S. Kadry, and Y. Nam, "A new dispersion control chart for handling the neutrosophic data," *IEEE Access*, vol. 8, pp. 96006–96015, 2020.
- [21] L. Liu and J. Zhang, "Adaptive nonparametric CUSUM scheme for detecting unknown shifts in location," *International Journal of Production Research*, vol. 52, no. 6, pp. 1592–1606, 2014.
- [22] M. A. Graham, S. Chakraborti, and A. Mukherjee, "Design and implementation of CUSUM exceedance control charts for unknown location," *International Journal of Production Research*, vol. 52, no. 18, pp. 5546–5564, 2014.
- [23] W. Peerajit, W. Areepong, and S. Sukparungsee, "Explicit analytical solutions for ARL of CUSUM chart for a long-memory SARFIMA model," *Communications in Statistics: Simulation and Computation*, vol. 48, no. 4, pp. 1176–1190, 2019.
- [24] S. A. Abbasi, M. Riaz, and A. Miller, "Enhancing the performance of CUSUM scale chart," *Computers and Industrial Engineering*, vol. 63, no. 2, pp. 400–409, 2012.
- [25] R. S. Lashkari and M. A. Rahim, "An economic design of cumulative sum charts to control non-normal process means," *Computers & Industrial Engineering*, vol. 6, no. 1, pp. 1–18, 1982.
- [26] T. C. Chang and F. F. Gan, "A cumulative sum control chart for monitoring process variance," *Journal of Quality Technology*, vol. 27, no. 2, pp. 109–119, 1995.
- [27] R. M. Dickinson, D. A. O. Roberts, A. R. Driscoll, W. H. Woodall, and G. G. Vining, "CUSUM charts for monitoring the characteristic life of censored Weibull lifetimes," *Journal of Quality Technology*, vol. 46, no. 4, pp. 340–358, 2014.
- [28] Z. Wu and Q. N. Wang, "A single CUSUM chart using a single observation to monitor a variable," *International Journal of Production Research*, vol. 45, no. 3, pp. 719–741, 2007.
- [29] C. W. Lu and M. R. Reynolds, "CUSUM charts for monitoring an autocorrelated process," *Journal of Quality Technology*, vol. 33, no. 3, pp. 316–334, 2001.
- [30] R. Ali and A. Haq, "New GWMA-CUSUM control chart for monitoring the process dispersion," *Quality and Reliability Engineering International*, vol. 34, no. 6, pp. 997–1028, 2018.
- [31] J. S. Hunter, "The exponentially weighted moving average," *Journal of Quality Technology*, vol. 18, no. 4, pp. 203–210, 1986.
- [32] S. H. Steiner and R. J. Mac Kay, "Monitoring processes with data censored owing to competing risks by using exponentially weighted moving average control charts," *Journal of the Royal Statistical Society: Series C: Applied Statistics*, vol. 50, no. 3, pp. 293–302, 2001.
- [33] A. K. Patel and J. Divecha, "Modified MEWMA control scheme for an analytical process data," *Global Journal of Computer Science and Technology Software & Data Engineering*, vol. 13, pp. 23–33, 2013.
- [34] S. L. Lu, "An extended nonparametric exponentially weighted moving average sign control chart," *Quality and Reliability Engineering International*, vol. 31, no. 1, pp. 3–13, 2015.
- [35] J. Brown and E. Moltchanova, "New exponentially weighted moving average control charts for monitoring process mean and process dispersion," *Quality and Reliability Engineering International*, vol. 31, no. 5, pp. 877–901, 2015.
- [36] S. V. Crowder and M. C. Hamilton, "An EWMA for monitoring a process standard deviation," *Journal of Quality Technology*, vol. 24, no. 1, pp. 12–21, 1992.
- [37] C. L. Zou and F. Tsung, "A multivariate sign EWMA control chart," *Technometrics*, vol. 53, no. 1, pp. 84–97, 2011.
- [38] J. M. Lucas and M. S. Saccucci, "Exponentially weighted moving average control schemes: properties and enhancements," *Technometrics*, vol. 32, no. 1, pp. 1–12, 1990.
- [39] S. Chakraborti, D. Van, and S. T. Bakir, "Nonparametric control charts: an overview and some results," *Journal of Quality Technology*, vol. 33, no. 3, pp. 304–315, 2001.
- [40] R. Taboran, S. Sukparungsee, and Y. Areepong, "A new non-parametric Tukey MA-EWMA control charts for detecting mean shifts," *IEEE Access*, vol. 8, pp. 207249–207259, 2020.
- [41] A. Shafqat, Z. Huang, M. Aslam, and M. S. Nawaz, "A non-parametric repetitive sampling DEWMA control chart based on linear prediction," *IEEE Access*, vol. 8, pp. 74977–74990, 2020.
- [42] Z. G. Stoumbos and J. H. Sullivan, "Robustness to non-normality of the multivariate EWMA control chart," *Journal of Quality Technology*, vol. 34, no. 3, pp. 260–276, 2002.
- [43] R. Y. Liu and K. Singh, "A quality index based on data depth and multivariate rank tests," *Journal of the American Statistical Association*, vol. 88, pp. 252–310, 1993.
- [44] R. Y. Liu, "Control charts for multivariate processes," *Journal of the American Statistical Association*, vol. 90, no. 432, pp. 1380–1387, 1995.
- [45] C. L. Zou, Z. J. Wang, and F. Tsung, "A spatial rank-based multivariate EWMA control chart," *Naval Research Logistics*, vol. 59, no. 2, pp. 91–110, 2012.
- [46] C. L. Zou, Z. J. Wang, X. M. Zi, and W. Jiang, "An efficient online monitoring method for high-dimensional data streams," *Technometrics*, vol. 57, no. 3, pp. 374–387, 2015.
- [47] W. Y. Hwang, "A new rank-based multivariate cusum approach for monitoring the process mean," *Quality and Reliability Engineering International*, vol. 32, no. 3, pp. 1167–1178, 2016.
- [48] A. Mukherjee and S. Chakraborti, "A distribution-free control chart for the joint monitoring of location and scale," *Quality and Reliability Engineering International*, vol. 28, pp. 335–352, 2012.
- [49] S. Chowdhury, A. Mukherjee, and S. Chakraborti, "Distribution-free phase ii CUSUM control chart for joint monitoring of location and scale," *Quality and Reliability Engineering International*, vol. 31, no. 1, pp. 135–151, 2015.
- [50] A. Mukherjee and M. Marozzi, "Nonparametric phase-ii control charts for monitoring high-dimensional processes with unknown parameters," *Journal of Quality Technology*, vol. 54, no. 1, pp. 44–64, 2022.
- [51] A. Mukherjee and M. Marozzi, "Distribution-free Lepage type circular-grid charts for joint monitoring of location and scale parameters of a process," *Quality and Reliability Engineering International*, vol. 33, no. 2, pp. 241–274, 2017.
- [52] Y. Lepage, "A combination of Wilcoxon's and Ansari-Bradley's statistics," *Biometrika*, vol. 58, no. 1, pp. 213–217, 1971.
- [53] A. Mukherjee and M. Marozzi, "A distribution-free phase-II CUSUM procedure for monitoring service quality," *Total Quality Management & Business Excellence*, vol. 28, no. 11–12, pp. 1227–1263, 2017.
- [54] A. Mukherjee and R. Sen, "Optimal design of Shewhart-Lepage type schemes and its application in monitoring service

- quality,” *European Journal of Operational Research*, vol. 266, no. 1, pp. 147–167, 2018.
- [55] Z. Li, M. Xie, and M. Zhou, “Rank-based EWMA procedure for sequentially detecting changes of process location and variability,” *Quality Technology & Quantitative Management*, vol. 15, no. 3, pp. 354–373, 2018.
- [56] L. X. Shi, L. Gong, and D. K. J. Lin, “CRPS chart: simultaneously monitoring location and scale under data-rich environment,” *Quality and Reliability Engineering International*, vol. 34, no. 4, pp. 681–697, 2018.
- [57] R. F. Zafar, T. Mahmood, N. Abbas, M. Riaz, and Z. Hussain, “A progressive approach to joint monitoring of process parameters,” *Computers & Industrial Engineering*, vol. 115, pp. 253–268, 2018.
- [58] Z. Song, A. Mukherjee, Y. Liu, and J. Zhang, “Optimizing joint location-scale monitoring - an adaptive distribution-free approach with minimal loss of information,” *European Journal of Operational Research*, vol. 274, no. 3, pp. 1019–1036, 2019.
- [59] S. Huang, J. Yang, and M. Xie, “A double-sampling SPM scheme for simultaneously monitoring of location and scale shifts and its joint design with maintenance strategies,” *Journal of Manufacturing Systems*, vol. 54, pp. 94–102, 2020.
- [60] K. Z. Bai and J. Li, “Location-scale monitoring of ordinal categorical processes,” *Naval Research Logistics*, vol. 68, no. 7, pp. 937–950, 2021.
- [61] C.-R. Cheng and J.-J. H. Shiau, “A distribution-free multivariate control chart for phase I applications,” *Quality and Reliability Engineering International*, vol. 31, no. 1, pp. 97–111, 2015.
- [62] A. M. Mood, “On the asymptotic efficiency of certain nonparametric two-sample tests,” *The Annals of Mathematical Statistics*, vol. 25, no. 3, pp. 514–522, 1954.
- [63] F. Wilcoxon, “Individual comparisons by ranking methods,” *Biometrics*, vol. 1, no. 6, pp. 80–83, 1945.
- [64] H. B. Mann and D. R. Whitney, “On a test of whether one of two random variables is stochastically larger than the other,” *The Annals of Mathematical Statistics*, vol. 18, no. 1, pp. 50–60, 1947.
- [65] S. Siegel and J. W. Tukey, “A nonparametric sum of ranks procedure for relative spread in unpaired samples,” *Journal of the American Statistical Association*, vol. 55, no. 291, pp. 429–445, 1960.
- [66] A. R. Ansari and R. A. Bradley, “Rank-sum tests for dispersions,” *The Annals of Mathematical Statistics*, vol. 31, no. 4, pp. 1174–1189, 1960.
- [67] H. W. Liliefors, “On the Kolmogorov-Smirnov test for normality with mean and variance unknown,” *Journal of the American Statistical Association*, vol. 62, no. 318, pp. 399–402, 1967.
- [68] A. Mukherjee, W. Kössler, and H. Murakami, “Two new distribution-free two-sample tests for versatile alternative,” *Statistics*, vol. 55, no. 5, pp. 1123–1153, 2021.
- [69] M. A. Fligner and T. J. Killeen, “Distribution-free two-sample tests for scale,” *Journal of the American Statistical Association*, vol. 71, no. 353, pp. 210–213, 1976.
- [70] O. L. Mangasarian and W. H. Wolberg, “Cancer diagnosis via linear programming,” *SIAM News*, vol. 23, pp. 1–18, 1990.
- [71] W. H. Wolberg and O. L. Mangasarian, “Multisurface method of pattern separation for medical diagnosis applied to breast cytology,” *Proceedings of the National Academy of Sciences USA*, vol. 87, no. 23, pp. 9193–9196, 1990.
- [72] O. L. Mangasarian, R. Setiono, and W. H. Wolberg, “Pattern recognition via linear programming: theory and application to medical diagnosis,” in *Large-Scale Numerical Optimization*, T. F. Coleman and Y. Y. Li, Eds., pp. 22–30, PA: SIAM Publications, Philadelphia, 1990.
- [73] K. P. Bennett and O. L. Mangasarian, *Robust Linear Programming Discrimination of Two Linearly Inseparable Sets*, Gordon & Breach Science Publishers, 1992.