A non parametric CUSUM control chart based on the Mann–Whitney statistic

Dabuxilatu Wang\textsuperscript{a,b}, Liang Zhang\textsuperscript{c}, and Qiang Xiong\textsuperscript{a}

\textsuperscript{a}Department of Statistics, School of Economics and Statistics, Guangzhou, China; \textsuperscript{b}Lingnan Research Center for Statistical Science, Guangzhou University, Higher Education Mega Center, Guangzhou, China; \textsuperscript{c}School of Applied Mathematics, Guangdong University of Technology, Tianhe District, Guangzhou, China

\section*{ABSTRACT}
We consider a novel univariate non parametric cumulative sum (CUSUM) control chart for detecting the small shifts in the mean of a process, where the nominal value of the mean is unknown but some historical data are available. This chart is established based on the Mann–Whitney statistic as well as the change-point model, where any assumption for the underlying distribution of the process is not required. The performance comparisons based on simulations show that the proposed control chart is slightly more effective than some other related non parametric control charts.

\section*{1. Introduction}
Statistical process control (SPC) has been widely used to monitor various industrial processes (e.g., manufacturing processes, health care monitoring, credit card and financial fraud detection, internet traffic flow, and so forth) (cf., Montgomery, 2004; Qiu, 2014). In practice, SPC is roughly divided into two phases. In phase I stage, a collection of sample from process is analyzed and the process is repeatedly adjusted for stable performance. Once the process is confirmed being in condition of a stable performance, then one can utilize the sample taken from the process to estimate the unknown parameters as well as the control limits. In phase II stage, the in-control (IC) distribution of the process obtained from the phase I stage is applied for monitoring the process operation, i.e., to detect changes in the response distribution after an unknown time point. The performance of the phase II is usually measured by the average run length (ARL), which is the expected number of observations needed for the procedure to signal a change in the response distribution. Most of research in SPC focuses on the charting techniques. Many control charts such as Shewhart chart, EWMA (exponential weighted moving average) chart and CUSUM (cumulative sum) chart, p-chart, c-chart, and chart based on change-point model have been proposed (cf., Montgomery, 2004; Qiu, 2014). All these control charts are based on the assumption that observations of a related process follow a normal distribution or other parametric distribution. The parameters of a underlying distribution may be known or unknown. In the later case, problems related to the estimation of the unknown parameters and the affection of the estimators on the performances of the designed control charts as well as the detection of the changes in the concerned unknown parameters have
been discussed in detail by Zhou et al. (2009). However, it is well-recognized that in many applications, the underlying process distribution does not satisfy the assumption of normality (or any other parametric distribution), so that the decision results obtained from using the charts mentioned above would be unreliable (cf., Qiu and Li, 2011a; Qiu and Li, 2011b; Qiu, 2014). In such cases, it is desirable to develop appropriate control charts that do not require an assumption of any specific parametric distribution. The distribution-free control charts or non parametric control charts were proposed for this purpose in the literature. An extensive overview on non parametric control charts was presented by Qiu (2014) and Chakraborti et al. (2001). It is well-known that the Shewhart chart is effective for detection of large shifts in the process mean, whereas both EWMA and CUSUM charts are effective for detection of small mean shifts. The non parametric versions of the EWMA chart and CUSUM chart have been paid much attention. Most of them are based on the ordering or ranking information of the observations obtained at the same time or different time points. For instance, based on the within-group Wilcoxon signed-rank statistic, Bakir and Reynolds (1979) proposed a non parametric CUSUM chart to track the shift of a location parameter \( \mu \) from an IC known value \( \mu_0 \). McDonald (1990) established some non parametric CUSUM chart for detecting the process mean shifts by using sequential rank test. Using cross-sectional antiranks of the measurement and the order information of the sample as well as the log-linear modeling, Qiu and Hawkins (2001), Qiu and Hawkins (2003), and Qiu (2008) designed some non parametric multivariate CUSUM control charts for detecting process variability and process mean shifts, respectively. Recently, Zhou et al. (2009) established a non parametric EWMA control chart using the Mann–Whitney statistic for detecting small shifts of the unknown mean in the case where only some historical IC data are available. This control chart has the advantage of distribution robustness and can be used in start-up or short-run situation, and some simulated results show that it has a good performance across the range of possible shifts. Chakraborti et al. (2008) also proposed a non parametric average control chart–MW chart based on the Mann–Whitney statistic under the condition that the available reference sample is taken from an IC process by a phase I analysis, which is highly sensitive in detection of the large shifts in the mean when the process was distribution free.

Some other novel non parametric methods for establishing CUSUM control charts can be mentioned as follows. Chatterjee and Qiu (2009) proposed a non parametric CUSUM control chart using bootstrap-based control limits. Qiu and Li (2011b) established a T-CUSUM chart using data transformation method. Qiu and Li (2011a) proposed P-CUSUM, L-CUSUM, and K-CUSUM charts by means of categorical data analysis based on categorizing observed data. Qiu and Zhang (2015) also consider several CUSUM charts using data transformation method. Yang and Cheng (2010) proposed a non parametric CUSUM mean chart based on the total number of univariate data exceeding the IC mean, which is already known or estimated by the available IC reference sample.

It is well-known that the Mann–Whitney statistic is equivalent to the Wilcoxon rank-sum statistic. Though the rank-based CUSUM charts have been proposed, there is no report on a non parametric CUSUM chart based on the Mann–Whitney statistic. In this article, we design such a CUSUM chart under the same condition of the EWMA chart proposed by Zhou et al. (2009) for detecting the shifts of mean.

The rest of the article is organized as follows. In Section 2, the concept of change-point and the Mann–Whitney statistic have been recalled. In Section 3, the test statistic of a non parametric CUSUM chart based on the Mann–Whitney statistic is proposed, and the relevant control limits and control rule are designed. In Section 4, some performance comparisons of the designed chart with other charts are given.
2. A non parametric control chart based on the Mann–Whitney statistic

In this section, some basic concepts on change-point and Mann–Whitney statistic as well as some related existing control charts are briefly introduced, and then the design of a new non parametric CUSUM control chart is considered.

2.1. Change-point and the Mann–Whitney statistic

The traditional change-point model can be illustrated as follows. Let \( \{X_i, i = 1, 2, \ldots\} \) be a sequence of independent random variables, and

\[
X_i \sim \begin{cases} 
F(x; \mu_0, \sigma_0^2), & i = 1, 2, \ldots, \tau; \\
F(x; \mu_1, \sigma_1^2), & i = \tau + 1, \tau + 2, \ldots 
\end{cases}
\]

where \( F \) stands for a continuous distribution function with unknown types, \( \mu_0, \mu_1 \) stand for the process mean, and \( \sigma_0^2, \sigma_1^2 \) the process variance. If it holds \( \mu_0 \neq \mu_1 \) or \( \sigma_0^2 \neq \sigma_1^2 \), then \( \tau \) is said to be a change point. Detecting and finding out a change point for random process is an important issue with which we can predict properly the potential change in the observed process. The detection of the mean shifts or variance change in process control is similar to the detection of change point in a process. Like Zhou et al. (2009), we consider a change-point model based on sequence of finite independent random variables \( \{X_i\} \), where \( i = 1, 2, \ldots, l \).

In the non parametric statistical analysis, a sort of change-point detection method is the well-known Mann–Whitney two-sample test (cf., Mann and Whitney, 1947), which is applied for detecting differences existed in the distributions of two populations through two independent group samples. For any \( 1 \leq t < l \), the Mann–Whitney statistic is defined as

\[
MW_{t,l} = \sum_{i=1}^{t} \sum_{j=t+1}^{l} I(x_j < x_i) = \sum_{j=t+1}^{l} [I(x_j < x_1) + I(x_j < x_2) + \cdots + I(x_j < x_t)],
\]

where

\[
I(x_j < x_i) = \begin{cases} 
1, & x_j < x_i, \\
0, & x_j \geq x_i.
\end{cases}
\]

If \( \mu_0 = \mu_1, \sigma_0^2 = \sigma_1^2 \), then it is said that the process is in IC state. In this article, we only consider the variation on the location parameter (the mean of the process) and let \( \sigma_0^2 = \sigma_1^2 \). It is verified that (cf., Mann and Whitney, 1947; Zhou et al., 2009), under IC state, the expectation and variance of \( MW_{t,l} \) can be obtained as

\[
E_0(MW_{t,l}) = \frac{t(l - t)}{2}, \ Var_0(MW_{t,l}) = \frac{t(l - t)(l + 1)}{12}.
\]

In practice, when there are ties in the data, the usual correction to the variance of \( MW_{t,l} \) can be made by multiplying the factor

\[
1 - \sum_{i=1}^{r} g_i(g_i^2 - 1)l^{-1}(l^2 - 1)^{-1},
\]
where \( r \) is the number of distinct values in the \( l \) observations and the \( i \)th value occurs with frequency \( g_i (\sum_{i=1}^{r} = l) \). In this situation, the IC variance of MW\(_{t,l}\) is
\[
\text{Var}_0(\text{MW}_{t,l}) = \frac{t(l-t)(l+1)}{12} \left( 1 - \sum_{i=1}^{r} g_i (g_i^2 - 1) l^{-1} (l^2 - 1)^{-1} \right).
\]
The standardized Mann–Whitney statistic \( \text{SMW}_{t,l} \) is defined by
\[
\text{SMW}_{t,l} = \frac{\text{MW}_{t,l} - E_0(\text{MW}_{t,l})}{\sqrt{\text{Var}_0(\text{MW}_{t,l})}}.
\]
Based on the conclusions given by Mann–Whitney (cf., Mann and Whitney, 1947), when the process is in IC state, the distribution of \( \text{SMW}_{t,l} \) is symmetric about zero for each \( t \), and large values of \( \text{SMW}_{t,l} \) indicate a negative mean shift, whereas small values indicate a positive shift (Zhou et al., 2009). As explained in Zhou et al. (2009), a test statistic for detection of change point about the mean (i.e., the hypothesis \( H_0 : \mu_0 = \mu_1 \)) is proposed by Pettitt (1979) as the maximal standardized Mann–Whitney statistic
\[
T_l = \max_{1 \leq t \leq l-1} |\text{SMW}_{t,l}|.
\]
If \( T_l \) exceeds some critical value \( h_l \), then we conclude that there is a shift in the mean. Otherwise, we conclude that there is no sufficient evidence of a shift. For finding a suitable critical values \( h_l \), we can use the limiting distribution of \( T_l \) given by Pettitt (1979) to make an approximation.

Note that for each \( t \), \( \text{SMW}_{t,l} \) can be modeled approximately by standard normal distribution \( N(0, 1) \) when \( l \) is large.

### 2.2. A non parametric CUSUM chart

For detecting a small mean shift occurred at the change point of the process as soon as possible, the CUSUM and EWMA control charts are usually recommended. Assuming that the occurred change point indicates an upward or downward mean shift, then at the change point, the expectation of the Mann–Whitney statistic becomes larger and the sample mean around the change point might occur a small shift. In such case, it is hard to find the shift of mean quickly using the statistic \( T_l \), because it only depends on the finite individual observations without considering all historical sample information. Therefore, we attempt to construct a non parametric CUSUM control chart based on the standardized Mann–Whitney statistic \( \text{SMW}_{t,l} \).

Zhou et al. (2009) did some modification on the maximal standardized Mann–Whitny statistic \( T_l \) so as to make it suitable to the case where the total \( m \) IC historical individual observations and the \( n \) future observations are available, and they proposed the SMW chart, which has a test statistic
\[
T_{m,n} = \max_{m-m_0 \leq t < m+n} |\text{SMW}_{t,(m+n)}|.
\]
Furthermore, for detecting the change point in the process more accurately, they established an EWMA chart based on \( \text{SMW}_{t,(m+n)} \) using charting statistic
\[
Y_j(m, n) = \lambda \cdot \text{SMW}_{j,(m+n)} + (1 - \lambda) \cdot Y_{j-1}(m, n),
\]
where \( m \) is the number of IC historical individual observations, \( n \) is the number of the future observations, and \( m + n \) is the number of the total observations and \( m - m_0 \leq t < m + n \),
CUSUM charts are more effective for detection of smaller shifts of mean than the EWMA chart. It indicates that the mean (cf., Lucas and Saccucci, 1990), for instance, when the smoothing constant smaller (larger) smoothing constant leads to quicker detection of smaller (larger) shifts of mean (cf., Lucas and Saccucci, 1990), CUSUM charts are usually slightly more sensitive than EWMA charts when the IC ARL becomes large (cf., Srivastava and Wu, 1993). It is easy to see that an EWMA chart can be viewed as a CUSUM chart approximately when the smoothing constant \( \lambda \) of the EWMA chart tends to zero. The EWMA charts are easier to understand and implement. On the other hand, CUSUM charts have certain theoretical optimality properties, and the corresponding theory for the EWMA charts is still lacking (Qiu, 2014; Qiu and Zhang, 2015).

Under the same condition of the SMW chart, we propose the test statistic for our CUSUM chart as

\[
S^+_n(m,n) = \max\{0, S^+_m(m,n) + \text{SMW}_{j(m+n)} - k\}, \\
S^-_n(m,n) = \max\{0, S^-_{m-1}(m,n) + \text{SMW}_{j(m+n)} + k\},
\]

where \( j = m - m_0, m - m_0 + 1, \ldots, m - m_0 + n - 1 \). \( 0 \leq m_0 \leq m \), \( S^+_m(m,n) = S^-_{m-1}(m,n) = 0 \), \( k \) is the reference value, in this article we assume \( k = 0.5 \). Set

\[
S^+_{\text{max}}(m,n) = \max_{m-m_0 \leq j \leq m+n-1} S^+_n(m,n), S^-_{\text{min}}(m,n) = \min_{m-m_0 \leq j \leq m+n-1} S^-_n(m,n).
\]

Our control rule is as follows,

1. After the \( n \)th future sample is monitored, compute \( S^+_{\text{max}}(m,n) \) or \( S^-_{\text{min}}(m,n) \).
2. Let \( h_{m,n} \) be the decision value, which is chosen to obtain the given IC ARL. If \( S^+_{\text{max}}(m,n) \leq h_{m,n} \) or \( S^-_{\text{min}}(m,n) \geq -h_{m,n} \), then we conclude that there is no evidence of a shift and continue to monitor the \( (n+1) \)st future sample. If \( S^+_{\text{max}}(m,n) > h_{m,n} \) or \( S^-_{\text{min}}(m,n) < -h_{m,n} \), then an out-of-control (OC) signal is triggered.

Note that the difference between the EWMA chart proposed by Zhou et al. (2009) and the CUSUM chart is that after the \( (m+n) \)th sample is monitored, the EWMA chart is to calculate the maximum value of exponentially moving average of \( \text{SMW}_{t(m+n)} \); however, the CUSUM chart is to calculate the maximum value of the cumulative sum of \( \text{SMW}_{t(m+n)} - k \) (if the “sum” is less than zero, then the “sum” takes value of zero) or minimum value of the cumulative sum of \( \text{SWM}_{t(m+n)} + k \) (if the “sum” is larger than zero, then the “sum” takes value of zero), for each \( t, m - m_0 \leq t < m + n \).

The smoothing constant \( \lambda \) of the EWMA chart by Zhou et al. is taken to be 0.2. In general, smaller (larger) smoothing constant leads to quicker detection of smaller (larger) shifts of mean (cf., Lucas and Saccucci, 1990), for instance, when the smoothing constant \( \lambda < 0.01 \), the EWMA chart may be approximately considered as a CUSUM chart. It indicates that the CUSUM chart is more effective for detection of smaller shifts of mean than the EWMA chart. For an easy comparison, we assume that \( \lambda = 0.2 \), and the proper value of \( m_0 \) can be chosen at the range of (Qiu and Hawkins, 2001; Hawkins et al., 2003) for \( m \geq 10 \), and \( m_0 = 4 \) for our proposed CUSUM chart. For the given type one error \( \alpha \), the decision value \( h_{m,n} \) can be obtained by solving the following equations

\[
\Pr(S^+_{\text{max}}(m,n) > h_{m,n}(\alpha) | S^+_{\text{max}}(m,i) \leq h_{m,i}(\alpha), 1 \leq i < n) = \alpha, n > 1,
\]

or

\[
\Pr(S^-_{\text{min}}(m,n) < -h_{m,n}(\alpha) | S^-_{\text{min}}(m,i) \leq h_{m,i}(\alpha), 1 \leq i < n) = \alpha, n > 1,
\]

\[
\Pr(S^-_{\text{min}}(m,1) < -h_{m,1}(\alpha)) = \alpha.
\]
Table 1. The decision values $h(m, n)(\alpha)$ of CUSUM control chart.

| n | $m = 10$ | $m = 50$ | ARL(0) |
|---|---------|---------|--------|
| 100 | 1.120 | 1.254 | 1.877 |
| 200 | 1.329 | 1.502 | 2.038 |
| 370 | 1.715 | 2.150 | 2.316 |
| 500 | 2.127 | 2.365 | 2.551 |

Due to the intricacy of this conditional probability, it seems to be impossible to solve it analytically. Therefore, similar to Zhou et al. (2009), we use one million sequences of length 500 which come from the standard normal distribution to estimate them. The historical sample size is assumed to be larger than 10. Table 1 shows the control limit of CUSUM chart for $\alpha$ values of 0.01, 0.005, 0.0027, and 0.002, corresponding to IC ARLs of 100, 200, 370, 500, for $m = 10$ and $m = 50$, $m_0 = 4$, and $n$ values in the range 1–490. As shown in Table 1, $h_{m,n}(\alpha)$ increase initially, but then stabilizes. We can obtain the optimal decision value using such approach of estimation. The missing values in Table 1 can be approximated by the last entries in the same column. Compared with the decision values $h_{m,n}(\alpha)$ shown in Table 1 of Zhou et al. (2009) for their EWMA chart, the decision values of CUSUM chart seem slightly smaller. A number of simulations by generating independent sequences of observations show that the control limits perform well. The advantage of the proposed CUSUM chart is same as the EWMA chart of Zhou et al. (2009), that is, it is completely distribution free, which means that the ARLs for different distributions are the same.

The empirical distribution of the run length (RL) for the CUSUM chart can be considered under the choice of decision values $h_{m,n}(\alpha)$ for different $m$. From Table 1, we observe that the difference of control limits for $m = 10$ and $m = 50$ is very small. The reason is that the distribution of $S_+^{m_1}(m_1, n)(S^-_{m_2}(m_2, n))$ is approximately the same as that of $S_+^{m_1}(m_1, n)(S^-_{m_2}(m_2, n))$ when $m_1, m_2$ are large enough. Thus, the control limits of the CUSUM chart for different $m$ ought to be close. From Table 1, we can choose the control limits for $m = 10$ and $m = 50$. We also hope that those control limits would be appropriate for other values of $m$. Same as the case of the EWMA chart proposed by Zhou et al. (2009), the run length distribution of CUSUM charts for $\alpha = 0.005, m = 25$ using the control limits of $h_{10,1}$ given in Table 1 are
very close to the geometric distribution. Simulations also demonstrate the same behavior of the RL distribution for $m = 100, 300, 500$ using the decision values $h_{50,t}(0.005)$ from Table 1. It seems reasonable using $h_{50,t}(\alpha)$ as the limits for the cases $m > 50$ when the requirement of IC behavior of RL is not very strict.

2.3. The diagnostic support and implementation

In the practical applications of quality control, there are usually two tasks that need to be carried out. One is to detect if the process is in control, the other is to point out the position of the shift if the observed process has shifted. Confirming the process change point would help engineers to identify the special cause quicker. Similar to Zhou et al. (2009), based on the Mann–Whitney statistic, we may obtain an estimator of the change point to assist our two-sided CUSUM chart. We assume that an OC signal is given at $m + n$ observation using the CUSUM chart, i.e., an upward shift or downward shift has occurred after $\tau$th future sample, $m \leq \tau < m + n$. Considering the properties of the two test statistic $S_{+}^{\tau}(m, n), S_{-}^{\tau}(m, n)$, we propose the estimator of the change-point $\tau$ of a step shift as

$$\hat{\tau} = \arg_{m \leq t < m+n} \max |SMW_{t,(m+n)}|,$$

this estimator is a non parametric one, which is the same estimator of change point proposed by Zhou et al. (2009). Also this estimator is less effective than some estimators of change point obtained by parametric likelihood methods, but it is still an accurate and useful estimator of the position of change point (cf., Zhou et al., 2009).

The equivalence between the Wilcoxon rank-sum statistic $W_{t,l}$ and the Mann–Whitney statistic $MW_{t,l}$ is shown with the equality $MW_{t,l} = W_{t,l} - \frac{(t+1)}{2}$, where $W_{t,l} = \sum_{i=1}^{t} R_i$, and $R_i$ denotes the rank of the $i$th observation $x_i$ in the total $l$ observations. We may use this equivalence to reduce the computational complexity of the Mann–Whitney statistic.

In the following, we present an example of simulation for detecting change point of the process with fixed times of IC observations and several future observations using our proposed CUSUM chart.

Example 1. For comparison with the EWMA chart proposed by Zhou et al. (2009), here we use the same dataset $\{x_i\}$ shown in Table 2 of Zhou et al. (2009). Assume that the underlying distribution of the observed process is Chi-square $\chi^2(4)$ with degree of freedom 4. There are $m = 15$ IC historical observations, which are the first 15 rows in Table 2. Suppose the mean has increased 0.75 standard deviation after the 5th future observation. The control limits of the CUSUM chart $h_{10,n}(0.005)$ for $m = 10$ are used which yield IC ARL $\approx 200$. The $h_{10,n}(0.005)$ for $n = 1, 2, \ldots, 8$ and the statistics $S_{+}^{\max}(15, n), S_{-}^{\min}(15, n)$ are shown in Table 2. It is clear that our CUSUM chart signals after three OC observations, which means that our CUSUM chart signals five OC observations earlier than the EWMA chart of Zhou et al. (2009). The reason of the signal delay of the EWMA chart proposed by Zhou et al. (2009) may be the inertia properties of EWMA chart. The maximum of $|SMW_{t,m+n}|$ for $n = 1, 2, \ldots, 8$ is the $|SMW_{20,23}| = 1.735$, which indicates accurately that the estimator of the location of the shift (the change point) is $\hat{\tau} = 20$.

3. Performance comparisons

In this section, we present performance comparisons between our CUSUM chart and the EWMA chart based on the available data from Zhou et al. (2009). Also we consider some comparisons of the characteristics of our test statistic to other non parametric CUSUM statistic.
Table 2. The values of the CUSUM statistics and the control limits based on process data with a shift after 20th sample.

| l  | \( x_i \) | \( S_{\text{max}}^{+} (15, n) \) | \( S_{\text{min}}^{-} (15, n) \) | \( h_{10, l-15} (0.005) \) |
|----|------------|-----------------|-----------------|-----------------|
| 1  | 2.061      |                 |                 |                 |
| 2  | 1.113      |                 |                 |                 |
| 3  | 4.298      |                 |                 |                 |
| 4  | 6.972      |                 |                 |                 |
| 5  | 1.675      |                 |                 |                 |
| 6  | 3.614      |                 |                 |                 |
| 7  | 3.446      |                 |                 |                 |
| 8  | 8.057      |                 |                 |                 |
| 9  | 4.702      |                 |                 |                 |
| 10 | 2.827      |                 |                 |                 |
| 11 | 1.637      |                 |                 |                 |
| 12 | 4.925      |                 |                 |                 |
| 13 | 7.506      |                 |                 |                 |
| 14 | 1.170      |                 |                 |                 |
| 15 | 2.308      |                 |                 |                 |
| 16 | 3.606      | 0.644           | 0               | 1.185           |
| 17 | 7.425      | 0               | -1.119          | 1.276           |
| 18 | 0.277      | 1.138           | 0               | 1.329           |
| 19 | 5.455      | 0.134           | -0.413          | 1.461           |
| 20 | 3.597      | 0.0943          | -0.130          | 1.589           |
| 21 | 6.068      | 0               | -1.092          | 1.661           |
| 22 | 4.618      | 0               | -1.601          | 1.752           |
| 23 | 8.384      | 0               | -7.729          | 1.812           |

3.1. A comparison with the EWMA chart

The non parametric CUSUM control chart proposed in this article is compared with the EWMA chart proposed by Zhou et al. (2009) under the assumption of normal distribution. As we know that a standard univariate CUSUM control chart is established using probability sequential ratio test based on normality assumption of underlying process distribution. Without loss of generality, the underlying IC distribution is assumed to be the standard normal distribution. The ARL comparison between CUSUM and EWMA control charts for \( N(0,1) \) data and \( m = 10, \alpha = 0.005 \) and different values of \( \tau \) are shown. In Table 3 (50,000 simulations), where \( \delta \) denotes the coefficient of standard deviation for measuring the shift size, \( \tau \) denotes the change point, whose values are chosen to be 10, 50, 100, and 250 for a representative illustration. The reference value \( k = 0.5 \) and the smoothing constant \( \lambda = 0.2 \).

Table 3. The ARL comparisons between CUSUM chart and EWMA chart for \( N(0,1) \) data and \( m = 10, \alpha = 0.005 \).

| \( \delta \) | \( \tau = 10 \) | \( \tau = 50 \) | \( \tau = 100 \) | \( \tau = 250 \) |
|-----------|---------------|---------------|---------------|---------------|
|           | CUSUM | EWMA | CUSUM | EWMA | CUSUM | EWMA | CUSUM | EWMA |
| 0.00      | 200.0 | 200.0 | 200.0 | 200.0 | 200.0 | 200.0 | 200.0 | 200.0 |
| 0.025     | 173.1 | 179.7 | 123.4 | 129.7 | 96.4 | 101.9 | 70.8 | 76.3 |
| 0.050     | 134.8 | 140.2 | 39.9 | 44.9 | 26.8 | 30.3 | 21.9 | 25.4 |
| 0.75      | 92.1 | 95.3 | 13.9 | 17.5 | 11.9 | 14.9 | 10.9 | 13.9 |
| 1.00      | 52.8 | 55.3 | 8.7 | 10.5 | 8.2 | 9.8 | 9.2 | 9.5 |
| 1.25      | 26.6 | 29.1 | 6.9 | 7.8 | 6.7 | 7.5 | 7.9 | 7.3 |
| 1.50      | 14.2 | 15.4 | 6.3 | 6.3 | 5.9 | 6.2 | 5.8 | 6.1 |
| 1.75      | 8.2 | 9.1 | 5.2 | 5.5 | 5.3 | 5.4 | 5.5 | 5.3 |
| 2.00      | 5.8 | 6.3 | 4.7 | 4.9 | 4.8 | 4.9 | 4.9 | 4.8 |
| 2.25      | 5.4 | 5.2 | 4.5 | 4.6 | 4.4 | 4.5 | 4.7 | 4.5 |
| 2.50      | 4.8 | 4.6 | 4.4 | 4.3 | 4.3 | 4.3 | 4.5 | 4.3 |
| 2.75      | 4.3 | 4.2 | 4.3 | 4.2 | 4.2 | 4.1 | 4.3 | 4.1 |
| 3.00      | 4.1 | 4.0 | 4.2 | 4.1 | 4.0 | 4.0 | 4.2 | 4.0 |
From Table 3, we can find that
(1) As the future observed IC data increases, both charts become more sensitive to the
shifts as the new observation updates the information already known.
(2) For detecting a relatively large shift in mean, the performances of the CUSUM chart
and the EWMA chart seem to be similar; sometimes, the EWMA chart is a little more
sensitive than the CUSUM chart. Both CUSUM and EWMA charts are less sensitive
to the relatively large mean shifts.
(3) Both CUSUM and EWMA control charts are relatively sensitive to small mean shifts
based on the rank information of the observed data. Our CUSUM chart is more sen-
titive than the EWMA chart for mean shifts of smaller size, for instance, when the mean
shift size is about \( \delta = 0.025 \) standard deviation, and \( \tau = 100 \), the ARL of the CUSUM
chart is 96.4, which is about 94.6% of ARL of EWMA chart.

As we know that the univariate CUSUM control charts for detecting small mean shifts of
the processes somewhat have responses to different types of the underlying process distri-
butions. When the underlying distribution of the process is not symmetric one, the control
limits of the two-sided CUSUM charts are usually recommended to be taken as \( h_1 > 0 \) and
\( h_2 < 0 \) and \( |h_1| \neq |h_2| \). However, the EWMA control charts have not so strong sensitivity to
distribution types of the underlying processes, and it can be used for all kinds of independent
observations from various distributions with almost same effectiveness. Therefore, similar to
Zhou et al. (2009), we also consider the cases where the observed populations are not only
skewed, but also heavy tailed, for which the IC ARL of the CUSUM chart may be different
from the case of normality. In start-up or short-run situation where we usually do not have
knowledge of the underlying distribution, a non parametric or distribution-free scheme such
as our proposed CUSUM chart or the EWMA chart is suitable to be used. We chose the under-
lying distributions of process like the \( \chi^2(4) \) distribution, \( t(4) \) distribution, and lognormal (0, 1)
distribution because they can represent a wide variety of shapes such as symmetric, skewed,
and heavy-tailed distributions. The shifts in the mean of \( \delta = 0.0(0.25)3 \) times standard devi-
ation are considered. The performance results under simulation for \( m = 10, \alpha = 0.005, \) and
\( \tau = 10, 50, 100, 250 \) are summarized in Tables 4, 5, and 6. Data from these tables show that
(1) For all distributions, both CUSUM and EWMA charts have almost the same perfor-
mances. They are sensitive to moderate and small mean shifts, but the CUSUM chart

| \( \delta \) | CUSUM | EWMA | CUSUM | EWMA | CUSUM | EWMA | CUSUM | EWMA | CUSUM | EWMA |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.00 | 200.0 | 200.0 | 200.0 | 200.0 | 200.0 | 200.0 | 200.0 | 200.0 | 200.0 | 200.0 |
| 0.025 | 192.1 | 199.5 | 131.6 | 139.2 | 93.7 | 99.7 | 58.7 | 64.3 |
| 0.050 | 161.9 | 170.32 | 36.7 | 41.9 | 23.1 | 24.8 | 16.1 | 20.2 |
| 0.75 | 110.3 | 124.0 | 14.1 | 15.5 | 11.9 | 12.5 | 8.9 | 11.5 |
| 1.00 | 72.5 | 81.1 | 8.7 | 9.5 | 8.1 | 8.6 | 6.8 | 8.2 |
| 1.25 | 41.2 | 49.0 | 6.3 | 7.2 | 6.3 | 6.8 | 6.2 | 6.6 |
| 1.50 | 23.9 | 29.1 | 5.6 | 6.1 | 5.5 | 5.8 | 5.6 | 5.7 |
| 1.75 | 17.3 | 17.7 | 5.2 | 5.4 | 5.0 | 5.2 | 5.0 | 5.1 |
| 2.00 | 11.8 | 11.6 | 4.8 | 4.9 | 4.7 | 4.8 | 4.6 | 4.7 |
| 2.25 | 8.2 | 8.1 | 4.5 | 4.6 | 4.5 | 4.5 | 4.4 | 4.5 |
| 2.50 | 6.7 | 6.5 | 4.4 | 4.4 | 4.2 | 4.3 | 4.5 | 4.3 |
| 2.75 | 5.8 | 5.5 | 4.3 | 4.3 | 4.3 | 4.2 | 4.3 | 4.2 |
| 3.00 | 5.3 | 5.0 | 4.3 | 4.2 | 4.2 | 4.1 | 4.2 | 4.1 |
Table 5. The ARL comparisons between CUSUM chart and EWMA chart for \( t(4) \) data and \( m = 10, \alpha = 0.005 \).

| \( \delta \) | \( \tau = 10 \) | \( \tau = 50 \) | \( \tau = 100 \) | \( \tau = 250 \) |
|------------|-------------|-------------|-------------|-------------|
|            | CUSUM       | EWMA        | CUSUM       | EWMA        | CUSUM       | EWMA        | CUSUM       | EWMA        |
| 0.00       | 200.0       | 200.0       | 200.0       | 200.0       | 200.0       | 200.0       | 200.0       | 200.0       |
| 0.025      | 172.3       | 177.4       | 108.2       | 113.9       | 77.8        | 81.2        | 52.3        | 57.7        |
| 0.050      | 127.6       | 132.3       | 28.6        | 31.1        | 20.1        | 22.0        | 15.2        | 19.0        |
| 0.75       | 75.7        | 81.6        | 10.9        | 13.1        | 9.1         | 11.6        | 9.7         | 10.9        |
| 1.00       | 39.1        | 44.3        | 7.1         | 8.6         | 7.6         | 8.1         | 7.1         | 7.8         |
| 1.25       | 19.9        | 23.5        | 6.2         | 6.7         | 5.9         | 6.4         | 6.1         | 6.3         |
| 1.50       | 12.5        | 13.2        | 5.3         | 5.7         | 5.2         | 5.5         | 5.2         | 5.4         |
| 1.75       | 7.9         | 8.6         | 4.7         | 5.1         | 4.8         | 5.0         | 4.8         | 4.9         |
| 2.00       | 6.1         | 6.5         | 4.4         | 4.7         | 4.3         | 4.6         | 4.5         | 4.6         |
| 2.25       | 5.2         | 5.4         | 4.1         | 4.5         | 4.1         | 4.4         | 4.3         | 4.4         |
| 2.50       | 4.6         | 4.9         | 4.0         | 4.3         | 4.0         | 4.3         | 4.1         | 4.3         |
| 2.75       | 4.3         | 4.5         | 3.9         | 4.2         | 4.1         | 4.1         | 3.9         | 4.1         |
| 3.00       | 4.1         | 4.3         | 3.9         | 4.1         | 4.0         | 4.1         | 3.8         | 4.1         |

seems slightly quicker than the EWMA chart in detecting smaller mean shifts for all cases. Thus, our proposed CUSUM chart has some sense of advantage.

(2) For detection of large shifts, the performances of the EWMA chart under these distributions are similar to the normal case. However, the effectiveness of the CUSUM chart sometimes are slightly affected by the non symmetry of the underlying process distribution (as our control limits are derived from the symmetric case), so that the performances of the CUSUM chart for detecting large shifts of mean are not similar to the performances in the case of normality, it signals a little slower to the large mean shifts than that of the EWMA chart.

3.2. Some simple comparisons with several related non parametric CUSUM charts

In the following, we also mention some simple comparisons between our CUSUM chart and some non parametric CUSUM charts proposed by Bakir and Reynolds (1979), McDonald (1990), and Yang and Cheng (2010) based on the characteristics of the test statistics.

Bakir and Reynolds (1979) consider a sequence of observations \( \{x_{ij}, i = 1, 2, \ldots; j = 1, \ldots, g\} \), each of size \( g = 4 \) or \( 5 \), and define a within group Wilcoxon signed rank-sum \( SR_i = \)

Table 6. The ARL comparisons between CUSUM chart and EWMA chart for lognormal \((0, 1)\) data and \( m = 10, \alpha = 0.005 \).

| \( \delta \) | \( \tau = 10 \) | \( \tau = 50 \) | \( \tau = 100 \) | \( \tau = 250 \) |
|------------|-------------|-------------|-------------|-------------|
|            | CUSUM       | EWMA        | CUSUM       | EWMA        | CUSUM       | EWMA        | CUSUM       | EWMA        |
| 0.00       | 200.0       | 200.0       | 200.0       | 200.0       | 200.0       | 200.0       | 200.0       | 200.0       |
| 0.025      | 184.8       | 193.8       | 134.9       | 142.3       | 97.8        | 105.6       | 67.2        | 71.7        |
| 0.050      | 156.2       | 161.1       | 43.2        | 49.6        | 23.9        | 29.1        | 18.9        | 22.8        |
| 0.75       | 121.5       | 127.1       | 17.9        | 20.1        | 12.2        | 14.8        | 11.1        | 13.5        |
| 1.00       | 92.20       | 97.2        | 10.2        | 12.0        | 9.6         | 10.4        | 8.3         | 9.9         |
| 1.25       | 70.8        | 74.3        | 7.9         | 9.1         | 7.2         | 8.4         | 7.6         | 8.1         |
| 1.50       | 49.8        | 56.4        | 6.5         | 7.6         | 6.6         | 7.2         | 6.8         | 7.0         |
| 1.75       | 37.6        | 43.8        | 6.1         | 6.7         | 6.1         | 6.4         | 6.1         | 6.3         |
| 2.00       | 30.3        | 34.0        | 5.6         | 6.1         | 5.6         | 5.9         | 5.6         | 5.8         |
| 2.25       | 24.7        | 27.3        | 5.3         | 5.7         | 5.3         | 5.5         | 5.3         | 5.5         |
| 2.50       | 19.3        | 22.0        | 5.1         | 5.4         | 5.1         | 5.2         | 5.1         | 5.2         |
| 2.75       | 16.1        | 18.2        | 4.9         | 5.1         | 4.9         | 5.0         | 5.0         | 5.0         |
| 3.00       | 14.6        | 15.4        | 4.8         | 5.0         | 4.9         | 4.9         | 4.9         | 4.8         |
\[\sum_{j=1}^{g} \text{sign}(x_{ij})R_{ij}\] for each observation, where \(R_{ij}\) denotes the rank of \(|x_{ij}|\) in \([|x_{i1}|, \ldots, |x_{ig}|]\), and based on which they propose their CUSUM statistic as
\[
\sum_{i=1}^{n} (SR_i - k) - \min_{0 \leq m \leq n} \sum_{i=1}^{m} (SR_i - k),
\]
or
\[
\max_{0 \leq m \leq n} \sum_{i=1}^{m} (SR_i + k) - \sum_{i=1}^{n} (SR_i + k).
\]

From the structure of the statistic, we are not able to compare it straightforwardly with our CUSUM statistic because the rank in Mann–Whitney is based on the comparison of the total samples, whereas the rank in former statistic is only based on comparison of within-group sample with its absolute values. If we restrict a sequence of group sample of size \(g\) to finite \(l\) times observations, then we may view them as a finite sequence of \(l \times g\) independent random variables, and we may assume the former \(m\) variables of the sequence are IC state, so that we can obtain our standardized Mann–Whitney statistic with a reference IC data. Noting that the rank involved in our CUSUM statistic is for total samples and the rank in Bakir and Reynolds’s CUSUM statistic (Bakir and Reynolds, 1979) is only the within-group signed rank, the sum of the former ranks is obviously larger than the sum of the later ranks. Therefore, for a fixed reference value \(k\) and a decision value \(h\), our CUSUM chart is clearly more sensitive than Bakir and Reynolds’s chart for small mean shifts.

For comparing our CUSUM chart with McDonald’s CUSUM chart (1990), we note that their sequential rank \(R_i\) is defined as
\[R_i = 1 + \sum_{j=1}^{i-1} I(x_j < x_i)\] for the observation \(\{x_i, i = 2, \ldots\}\), and the CUSUM statistic \(T_j = \max\{0, T_{j-1} + U_j - k\}\), where \(U_i = \frac{R_{i+1}}{i+1}\). The Mann–Whitney statistic \(MW_{i,n}\) can be written as
\[
MW_{i,n} = (R_{i+1} - 1) + \sum_{s=i+2}^{n} (I[x_1 < x_s] + I[x_2 < x_s] + \cdots + I[x_i < x_s]),
\]
i = 2, \ldots, n – 1. Therefore, \(MW_{i,n}\) is obviously larger than \(R_n\), which means that the cumulative sample information of our CUSUM statistic is more abundant than the cumulative sample information of \(T_j\). It roughly demonstrates that our CUSUM chart based on \(MW_{i,n}\) may be more sensitive than the CUSUM chart based on \(R_i\) proposed by McDonald (1990).

We now consider a simple comparison of our CUSUM chart with Yang and Cheng’s CUSUM chart (Yang and Cheng, 2010). Noting that their chart is based on statistic \(M_t = \sum_{j=1}^{t} I(x_j > \mu)\) for the observations \(\{x_i, i = 1, 2, \ldots\}\), where \(\mu\) is the in control mean of the process. Obviously, the Mann–Whitney statistic \(MW_{i,n}\) implies more information than \(M_t\), especially in the case where the unknown IC mean need to be estimated by the available reference IC samples. They only utilize the average information like \(\bar{x}\) of the available reference IC samples, whereas in our CUSUM case, the \(m\) available reference IC samples are fully utilized with an individual comparison. So, our CUSUM chart is also relatively more sensitive than Yang and Cheng’s CUSUM chart.

**Conclusion**

The rank based statistical method is an important well-known non parametric approach for the case where the underlying distribution of the process is completely unknown, in which the Mann–Whitney statistic is the popular and a powerful one. We establish a sort of non
parametric CUSUM chart based on the standardized Mann–Whitney statistic for quickly
detection of the small location shifts under the same condition of the EWMA chart proposed
by Zhou et al. (2009). Some comparisons results (under $\lambda = 0.2, k = 0.5$) show that the
proposed CUSUM chart is slightly more effective than the non parametric charts proposed
by Zhou et al. (2009), Bakir and Reynolds (1979), McDonalds (1990), and Yang and Cheng
(2010) in the detection of the small mean shifts. However, this CUSUM chart also has some
drawbacks of the EWMA chart proposed by Zhou et al. (2009), i.e., it is less efficient than the
parametric methods for detection of the large shift. It is also to be pointed out that the non
parametric control charts rest on the theoretical research only, of which the development of
the practical applications is highly desired.

**Acknowledgments**

The authors thank the editor and the reviewer for their valuable comments and suggestions on the
manuscript, which led to a significant improvement of the article.

**Funding**

The research of Dabuxilatu Wang and Qiang Xiong is supported by the NNSF of China under grant
number 11271096. The authors are very appreciate for the financial aid.

**References**

Bakir, S.T., Reynolds, M.R.I.R. (1979). A non parametric procedure for process control based on within
group ranking. *Technometrics* 21:175–183.

Chakraborti, S., Van Der Laan, P., Bakir, S. T., (2001). Nonparametric control charts: An overview and
some results. *J. Qual. Technol.* 33(3):304–315.

Chakraborti, S., Van De Wiel Mark, A. (2008). A non parametric control chart based on the Mann–
Whitney statistic, IMS collections, beyond parametrics in interdisciplinary research: festschrift in
honor of professor Pranab K. Sen, Vol 1:156–172.

Chatterjee, S., Qiu, P. (2009). Distribution-free cumulative sum control charts using bootstrap-based
control limits. *Ann. Appl. Stat.* 3:349–369.

Hawkins, D.M., Qiu, P., Kang, C.W. (2003). A changepoint model for statistical process control. *J. Qual.
Technol.* 35:355–366.

Lucas, J.M., Saccucci, M.S. (1990). Exponentially weighted moving average control charts schemes:
properties and enhancements. *Technometrics* 32:1–30.

Mann, H.B., Whitney, D.R. (1947). On the test whether one of two random variables is stochastically
larger than other. *Ann. Math. Stat.* 18:50–60.

McDonald, D. (1990). A CUSUM procedure based on sequential ranks. *Naval Res. Logist.* 37:627–646.

Montgomery, D.C. (2004). *Introduction to Statistical Quality Control*, 5th edition. New York: Wiley.

Pettitt, A.N. (1979). A non-parametric approach to the change-point problem. *Appl. Stat.* 28:126–135.

Qiu, P., Hawkins, D.M. (2001). A rank based multivariate CUSUM procedure. *Technometrics* 43:120–
132.

Qiu, P., Hawkins, D.M. (2003). A non parametric multivariate CUSUM procedure for detecting shifts
in all directions. *Statistician* 52:152–164.

Qiu, P. (2014). *Introduction to Statistical Processes Control*. Boca Raton, FL: Chapman & Hall/CRC.

Qiu, P., Zhang, J. (2015). On phase II SPC in cases when normality is invalid. *Qual. Reliab. Eng. Int.*
31:27–35.

Qiu, P., Li, Z. (2011a). On non parametric statistical process control of univariate processes. *Technometrics*
81:390–405.

Qiu, P., Li, Z. (2011b). Distribution-free monitoring of univariate processes. *Stat. Probab. Lett.* 81:1833–
1840.
Qiu, P. (2008). Distribution-free multivariate process control based on log-linear modeling. *IIE Trans.* 40:664–677.

Srivastava, M.S., Wu, Y. (1993). Comparison of EWMA, CUSUM and Shiryayev-Roberts procedures for detecting a shift in the mean. *Ann. Stat.* 21(2):645–670.

Yang, S., Cheng, S.W. (2010). A new non parametric CUSUM mean chart. *Qual. Reliab. Eng. Int.* doi:10.1002/qre.1171.

Zhou, C., Zou, C., Zhang, Y., Wang, Z. (2009). Non parametric control chart based on change-point model. *Stat. Pap.* 50:13–28.