

# Distribution-Free Monitoring of Univariate Processes

Peihua Qiu<sup>1</sup> and Zhonghua Li<sup>1,2</sup>

<sup>1</sup>School of Statistics, University of Minnesota, USA

<sup>2</sup>LPMC and Department of Statistics, Nankai University, China

## Abstract

We consider statistical process control (SPC) of univariate processes when observed data are not normally distributed. Most existing SPC procedures are based on the normality assumption. In the literature, it has been demonstrated that their performance is unreliable in cases when they are used for monitoring non-normal processes. To overcome this limitation, we propose two SPC control charts for applications when the process data are not normal, and compare them with the traditional CUSUM chart and two recent distribution-free control charts. Some empirical guidelines are provided for practitioners to choose a proper control chart for a specific application with non-normal data.

*Key Words:* Distribution-free; Non-Gaussian data; Nonparametric procedures; Transformation, Wilcoxon signed-rank test.

## 1 Introduction

Statistical process control (SPC) charts are widely used in industry for monitoring stability of certain sequential processes (e.g., manufacturing processes, health care systems, internet traffic flow, and so forth). The early stage of process monitoring, in which the process needs to be repeatedly adjusted for stable performance, is often called the Phase I analysis, and the afterwards online process monitoring is often called the Phase II SPC. Performance of a Phase II SPC procedure is usually measured by the average run length (ARL), which is the average number of observations needed for the procedure to signal a change in the measurement distribution. The ARL value of the procedure when it is in control (IC), denoted as  $ARL_0$ , is often controlled at some specific level. Then, the procedure performs better if its out-of-control (OC) ARL, denoted as  $ARL_1$ , is shorter, when detecting a given distributional change. See, e.g., Hawkins and Olwell (1998), Montgomery

(2009), Woodall (2000), and Yeh et al. (2004) for related discussion. This paper focuses on Phase II monitoring of univariate processes in cases when process observations are not normally distributed.

In the literature, many Phase II SPC charts have been proposed, including different versions of the Shewhart chart, the cumulative sum (CUSUM) chart, the exponentially weighted moving average (EWMA) chart, and the chart based on change-point detection (cf., e.g., Hawkins and Olwell 1998, Hawkins et al. 2003, Montgomery 2009). All these control charts are based on the assumption that observations of a related process follow a normal distribution. In practice, however, process observations may not follow normal distributions. In such cases, it has been well demonstrated that results from the charts mentioned above would be unreliable (cf., Amin et al. 1995, Hackl and Ledolter 1992, Lucas and Crosier 1982). As a demonstration, Figure 1 shows the actual IC ARL values of the conventional CUSUM chart (Page 1954) based on the assumption that the IC response distribution is  $N(0, 1)$ , in cases when the allowance constant of the chart (see introduction in Section 2) is 0.5, the assumed IC ARL value equals 500 and the true response distribution is the standardized version (with mean 0 and variance 1) of the chi-square (plot (a)) or  $t$  (plot (b)) distribution with degrees of freedom (df) changing from 1 to 60 in plot (a) and from 3 to 60 in plot (b). From the plots, it can be seen that the actual IC ARL values of the conventional CUSUM chart are much smaller than the nominal IC ARL value when the df is small, which implies that the related process would be stopped too often by the control chart and consequently a considerable amount of time and resource would be wasted in such cases.

From Figure 1, it seems necessary to develop appropriate control charts that do not require the normal distribution assumption in cases when the process distribution is actually non-normal. To this end, a number of distribution-free or nonparametric control charts have been developed. See, for instance, Amin and Widmaier (1999), Bakir(2006), Bakir and Reynolds (1979), Chakraborti et al. (2009), and Chakraborti and Eryilmaz (2007). Chakraborti et al. (2001) gives a thorough overview on existing research in the area of univariate distribution-free SPC. In multivariate cases, see Qiu and Hawkins (2001, 2003) and Qiu (2008) for related discussion.

Most existing distribution-free SPC charts mentioned above are based on ordering or ranking information of the observations obtained at the same or different time points. Some of them require multiple observations at each time point (i.e., cases with batch or grouped data). Intuitively, it

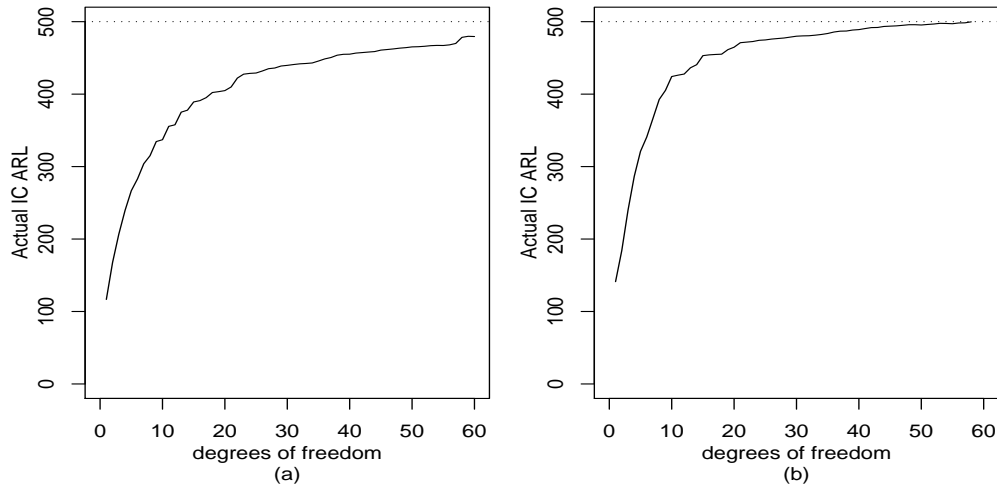


Figure 1: Actual IC ARL (i.e.,  $ARL_0$ ) values of the conventional CUSUM chart in cases when the nominal IC ARL value is 500, and the true response distribution is the standardized version (with mean 0 and variance 1) of the chi-square (plot (a)) or  $t$  (plot (b)) distribution with degrees of freedom changing from 1 to 60 in plot (a) and from 3 to 60 in plot (b).

would lose much information if we only use the ordering or ranking information in the observed data for process monitoring. An alternative approach would first transform the non-normal data properly so that the distribution of the transformed data is close to normal, and then a traditional control chart is applied to the transformed data. To this end, two such control charts are proposed in this paper. They are compared with the traditional CUSUM chart and two recent nonparametric control charts. Some practical guidelines are provided for users to choose a proper control chart for a specific application with non-normal data.

The rest part of the paper is organized as follows. Our proposed control charts are described in Section 2. A numerical study to evaluate their performance in comparison with several existing control charts is presented in Section 3. An application is discussed in Section 4 to demonstrate the use of the proposed method in a real world setting. Some remarks conclude the article in Section 5.

## 2 Proposed Control Charts

Like some existing distribution-free control charts (e.g., Chakraborti et al. 2009), we do not assume that the IC measurement cumulative distribution function (cdf)  $F_0$  is known. Instead, we assume

that an IC dataset is available for us to estimate certain IC parameters.

Let  $\mathbf{X}(n) = (X_1(n), X_2(n), \dots, X_m(n))'$  be  $m$  repeated observations obtained at the  $n$ th time point during Phase II process monitoring. In the literature, such data are called batch data, and  $m$  is the batch size. When  $m = 1$ , the data are sometimes called single-observation data. The traditional CUSUM chart is a standard tool for monitoring univariate processes in practice. Its charting statistics of the two-sided version are defined by

$$\begin{aligned} u_{n,N}^+ &= \max\left(0, u_{n-1,N}^+ + \bar{X}(n) - k_N\right), \\ u_{n,N}^- &= \min\left(0, u_{n-1,N}^- + \bar{X}(n) + k_N\right), \text{ for } n \geq 1, \end{aligned}$$

where  $u_{0,N}^+ = u_{0,N}^- = 0$ ,  $k_N$  is an allowance constant, and  $\bar{X}(n) = \frac{1}{m} \sum_{j=1}^m X_j(n)$ . Then, a mean shift in  $\mathbf{X}(n)$  is signaled if

$$u_{n,N}^+ > h_N \quad \text{or} \quad u_{n,N}^- < -h_N \quad (1)$$

where the control limit  $h_N > 0$  is chosen to achieve a given  $ARL_0$  level under the assumption that all observations are normally distributed. This chart is called N-CUSUM chart in this paper.

When process observations are not normally distributed, as demonstrated in Figure 1, the N-CUSUM chart may not be appropriate to use for process monitoring. In such cases, Chou et al. (1998) suggested an algorithm to transform a non-normal dataset to a standard normal dataset, by using the Slifker and Shapiro's (1980) method for distribution estimation under the Johnson's (1949) system of distributions. More specifically, Johnson (1949) considered three distribution families, labels as  $S_B$ ,  $S_L$ , and  $S_U$ . For each distribution family, Johnson found a transformation to transform the distributions in the family to the standard normal distribution. For a given distribution  $F$ , Slifker and Shapiro (1980) developed a criterion to classify  $F$  to one of the  $S_B$ ,  $S_L$ , and  $S_U$  families. Let

$$QR = \frac{(\tau_4 - \tau_3)(\tau_2 - \tau_1)}{(\tau_3 - \tau_2)^2}$$

where  $\tau_j$  are the  $q_j$ th quantile of  $F$ , for  $j = 1, 2, 3, 4$ ,  $q_1 = \Phi(-3z)$ ,  $q_2 = \Phi(-z)$ ,  $q_3 = \Phi(z)$ ,  $q_4 = \Phi(3z)$ ,  $z$  is a given number between 0.25 and 1.25, and  $\Phi$  is the cdf of the standard normal distribution. Then, Slifker and Shapiro's criterion is as follows.

- If  $QR < 1$ , then  $F$  belongs to the family  $S_B$ ,

- If  $QR = 1$ , then  $F$  belongs to the family  $S_L$ , and
- If  $QR > 1$ , then  $F$  belongs to the family  $S_U$ .

Based on this criterion and Johnson's method to transform distributions in  $S_B$ ,  $S_L$ , and  $S_U$  to the standard normal distribution, Chou et al. (1998) developed a numerical algorithm to transform any non-normal dataset to a standard normal dataset.

At the  $n$ th time point during Phase II monitoring, let  $\mathbf{Z}(n) = (Z_1(n), Z_2(n), \dots, Z_m(n))'$  be the transformed observations by Chou et al.'s algorithm from the original observations  $\mathbf{X}(n)$ . Then,  $\mathbf{X}^*(n) = (X_1^*(n), X_2^*(n), \dots, X_m^*(n))'$  with  $X_j^*(n) = Z_j(n)s_n + \bar{X}(n)$ , for  $j = 1, 2, \dots, m$ , are the transformed observations having the same sample mean and sample standard deviation as  $\mathbf{X}(n)$ , where  $\bar{X}(n)$  and  $s_n$  denote the sample mean and sample standard deviation of  $\mathbf{X}(n)$ . So,  $\mathbf{X}^*(n)$  should be roughly normal if the transformation works well. We then apply the conventional N-CUSUM chart (1) to  $\{\mathbf{X}^*(n), n \geq 1\}$ . The resulting control chart is called T-CUSUM chart, reflecting the fact that it is applied to the transformed data.

When the IC process distribution is unknown but an IC dataset is available, another natural idea is to find a transformation from the IC data such that the distribution of the transformed IC data is close to normal, and then apply the N-CUSUM chart (1) to the transformed Phase II data for process monitoring. In the literature (cf., Section 13.1.4, Cook and Weisberg 1999), a commonly used approach for finding such a transformation is to consider the Box-Cox transformation family

$$BC_\alpha(x) = \begin{cases} (x^\alpha - 1)/\alpha, & \text{if } \alpha \neq 0 \\ \log(x), & \text{otherwise,} \end{cases}$$

where  $\alpha$  is a parameter. The value of  $\alpha$  can be determined by maximizing the Shapiro-Wilk normality test statistic (cf., Shapiro and Wilk 1965) which is included in most statistical software packages, such as SAS and R. The resulting CUSUM chart is called B-CUSUM chart.

### 3 Numerical Study

In this section, we present some numerical examples to evaluate the performance of the charts T-CUSUM and B-CUSUM, in comparison with certain representative existing control charts. The

existing control charts considered here are the traditional N-CUSUM chart, and the two recent distribution-free control charts by Chakraborti and Eryilmaz (2007) and Chakraborti et al. (2009) described below.

The chart by Chakraborti et al. (2009) is a Shewhart-type chart, and it monitors the median of  $\mathbf{X}(n)$  over time and a shift is signaled if the median is out of the upper or lower bound determined by a reference sample (i.e., the IC data) to achieve a given  $ARL_0$  level. This chart was called PRECEDENCE chart in Chakraborti et al. (2009), where several different versions of the PRECEDENCE chart were presented. Here, we use the version of 2-of-2 KL, as suggested by Chakraborti et al. (2009) for detecting either up or down shifts. By this version of the chart, a signal of mean shift is delivered when two consecutive medians are both on or above the upper bound or both on or below the lower bound. The chart by Chakraborti and Eryilmaz (2007) is also a Shewhart-type chart and it is constructed based on the statistic

$$\psi(n) = 2W_n^+ - \frac{m(m+1)}{2}, \text{ for } n \geq 1,$$

where  $W_n^+$  is the Wilcoxon signed-rank statistic of  $\mathbf{X}(n)$ , defined to be the sum of the ranks of  $\{|X_j(n) - \theta_0|, j = 1, 2, \dots, m\}$  over all positive components of  $\{X_j(n) - \theta_0, j = 1, 2, \dots, m\}$ , and  $\theta_0$  is the IC median of the process distribution which can be estimated from the IC data. As well demonstrated in the literature (e.g., Hawkins and Olwell 1998), CUSUM charts are more favorable for detecting persistent shifts, compared to Shewhart-type charts. For that reason, we construct a CUSUM chart based on  $\psi_n$  as follows. Let  $u_{0,S}^+ = u_{0,S}^- = 0$ , and

$$\begin{aligned} u_{n,S}^+ &= \max\left(0, u_{n-1,S}^+ + (\psi(n) - \psi_0) - k_S\right), \\ u_{n,S}^- &= \min\left(0, u_{n-1,S}^- + (\psi(n) - \psi_0) + k_S\right), \text{ for } n \geq 1, \end{aligned}$$

where  $k_S$  is an allowance constant, and  $\psi_0$  is the IC mean of  $\psi(n)$  that can be estimated from the IC data. Then, the CUSUM chart signals a mean shift in  $\mathbf{X}(n)$  if

$$u_{n,S}^+ > h_S \quad \text{or} \quad u_{n,S}^- < -h_S, \tag{2}$$

where the control limit  $h_S$  is chosen to achieve a given  $ARL_0$  level. The chart (2) is called S-CUSUM chart hereafter. As a side note, we do not know how to modify the PRECEDENCE chart properly to make it a CUSUM chart at this moment. So, the original PRECEDENCE chart is included in our numerical study.

Table 1: The actual IC ARL values and their standard errors (in parentheses) of the five control charts when the nominal IC ARL values are fixed at 500 and the actual IC process distribution is the standardized version of  $N(0, 1)$ ,  $t(4)$ ,  $\chi^2(1)$  and  $\chi^2(4)$ .

	$N(0, 1)$	$t(4)$	$\chi^2(1)$	$\chi^2(4)$
N-CUSUM	499.7 (5.25)	396.3 (4.05)	348.7 (3.72)	443.0 (4.73)
T-CUSUM	495.4 (7.36)	341.2 (3.83)	325.6 (5.04)	347.9 (4.91)
B-CUSUM	522.5 (6.19)	500.3 (5.44)	508.9 (5.73)	523.5 (5.97)
S-CUSUM	494.7 (5.79)	515.0 (5.82)	537.0 (6.22)	517.8 (6.00)
PRECEDENCE	515.9 (5.82)	531.4 (5.90)	526.1 (5.88)	522.8 (5.96)

We then compare the two control charts T-CUSUM and B-CUSUM that are both based on transformations with the charts N-CUSUM, S-CUSUM, and PRECEDENCE. In the comparison, the IC distribution is chosen to be the standardized version with mean 0 and standard deviation 1 of one of the following four distributions:  $N(0, 1)$ ,  $t(4)$ ,  $\chi^2(1)$  and  $\chi^2(4)$ . Among these distributions,  $t(4)$  represents symmetric distributions with heavy tails, and  $\chi^2(1)$  and  $\chi^2(4)$  represent skewed distributions with different skewness. It is also assumed that the pre-specified  $ARL_0$  value is 500, and the batch size of Phase II observations at each time point is  $m = 5$ .

First, we compute the actual  $ARL_0$  values of the five control charts, based on 10,000 replicated simulations, in cases with the four actual IC process distributions described above. In computing the control limits of the CUSUM charts, the N-CUSUM chart is based on the assumption that the original process observations are normally distributed, the T-CUSUM chart is based on the assumption that the transformed data by the numerical algorithm developed by Chou et al. (1998) are normally distributed, and the B-CUSUM, S-CUSUM, and PRECEDENCE charts are based on 500 IC observations. For all control charts, their allowance constants are chosen to be 0.2. The results are shown in Table 1. From the table, it can be seen that the actual  $ARL_0$  values of the charts N-CUSUM and T-CUSUM are quite far away from 500 in cases when the actual IC distribution is not normal. As a comparison, the B-CUSUM, S-CUSUM, and PRECEDENCE charts seem quite robust to the IC distribution, because their actual  $ARL_0$  values are all quite close to 500 in all cases considered.

Next, we compare the OC performance of the related control charts in cases when the IC sample size  $M = 200$  or 500. In order to make the comparison more meaningful, we intentionally adjust the control limits of all charts so that their actual  $ARL_0$  values equal 500 in all cases considered. In

this study, 10 mean shifts ranging from -1.0 to 1.0 with step 0.2 are considered, representing small, medium and large shifts. Due to the facts that different control charts have different parameters and that the performance of different charts may not be comparable if their parameters are pre-specified, we use the following two approaches to set up their parameters. One is that all their parameters are chosen to be the optimal ones for detecting a given shift (e.g., the one of size 0.6), by minimizing the OC ARL values of the charts for detecting that shift while their  $ARL_0$  values are all fixed at 500, and the chosen parameters are used in all other cases as well. The second approach is to compare the optimal performance of all the charts when detecting each shift, by selecting their parameters to minimize the  $ARL_1$  values for detecting each individual shift, while their  $ARL_0$  values are all fixed at 500. The second approach for comparing different control charts has been used in the literature. See, for instance, Qiu and Hawkins (2001). Results based on 10,000 replications, when  $M = 200$  or  $500$  and when one of the two approaches described above is used for choosing the parameters, are presented in Figures 2–5, respectively. When reading the plots in these figures, readers are reminded that the scale on the  $y$ -axis is in natural logarithm, to better demonstrate the difference among different control charts when detecting relatively large shifts.

From Figure 2(a), we can see that when the actual process distribution is normal, the N-CUSUM and B-CUSUM charts perform the best, as expected. In such a case, the T-CUSUM and S-CUSUM charts perform slightly worse than the N-CUSUM and B-CUSUM charts, and the PRECEDENCE chart performs the worst because it loses much information when considering the ordering information in the data only and because it is a Shewhart chart. In the case of Figure 2(b), the IC process distribution is  $t(4)$ , which is still symmetric but has heavier tails, compared to the normal distribution. In such cases, the N-CUSUM chart is not the best any more. Instead, the B-CUSUM chart performs well, especially when the mean shift is relatively large. The S-CUSUM chart also performs well, especially when the mean shift is relatively small. The T-CUSUM and PRECEDENCE charts perform relatively worse. In cases of Figure 2(c)-(d), the IC process distribution is either  $\chi^2(1)$  or  $\chi^2(4)$ , which is skewed to the right. In such cases, when the skewness is small (cf., plot (d)), the B-CUSUM chart performs the best for detecting most shifts, the N-CUSUM chart also performs well, the S-CUSUM chart performs well when the mean shift is small, the PRECEDENCE chart performs well only when the mean shift is in the direction of the shorter tail of the IC distribution, and the T-CUSUM chart does not perform well



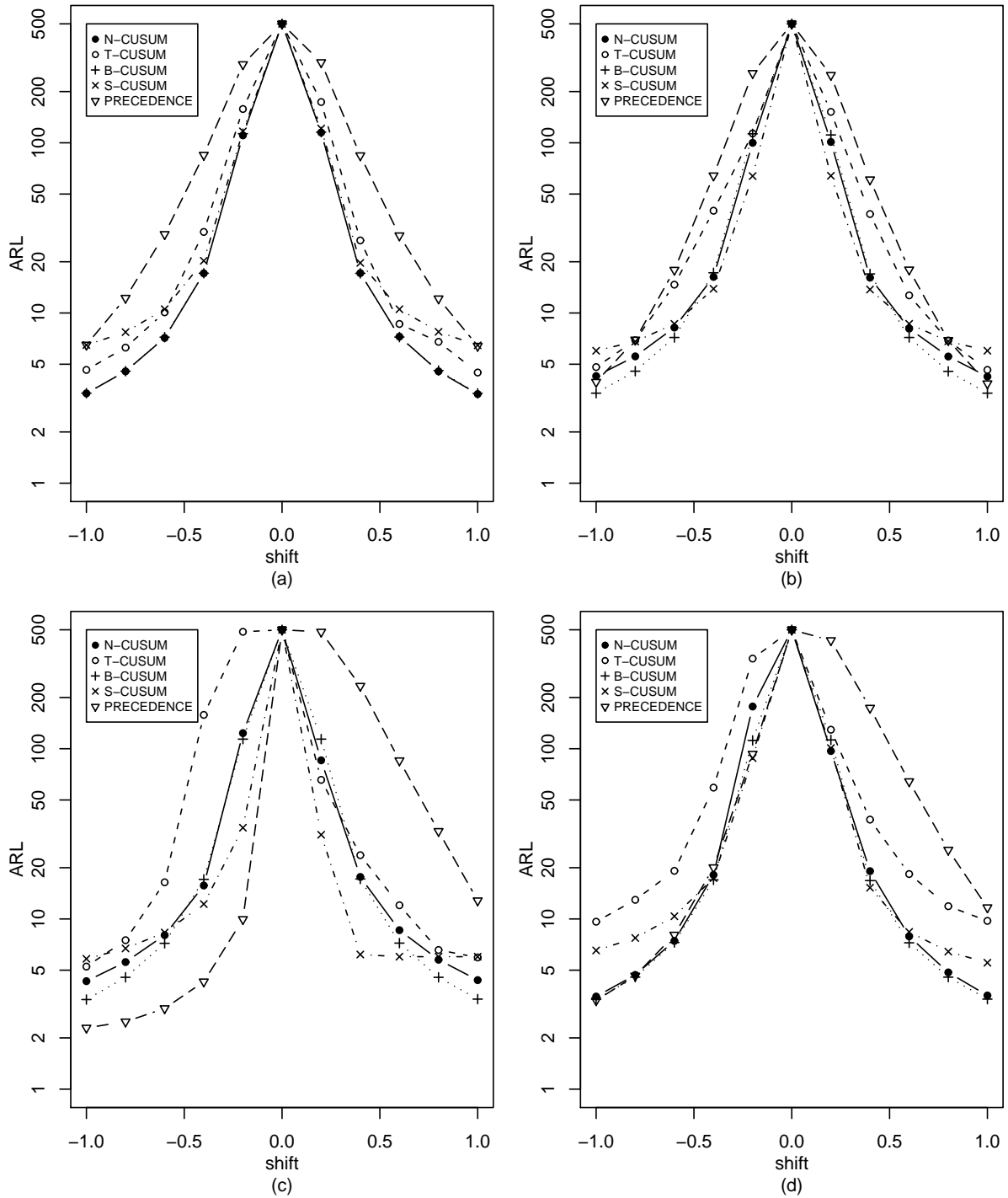


Figure 2: OC ARL values of five control charts when  $ARL_0 = 500$ ,  $M = 200$ ,  $m = 5$ , and the actual IC process distribution is the standardized version of  $N(0, 1)$  (plot (a)),  $t(4)$  (plot (b)),  $\chi^2(1)$  (plot (c)), and  $\chi^2(4)$  (plot (d)). Procedure parameters of the control charts are chosen to be the ones that minimize their OC ARL values when detecting the shift of 0.6. Scale on the  $y$ -axis is in natural logarithm.

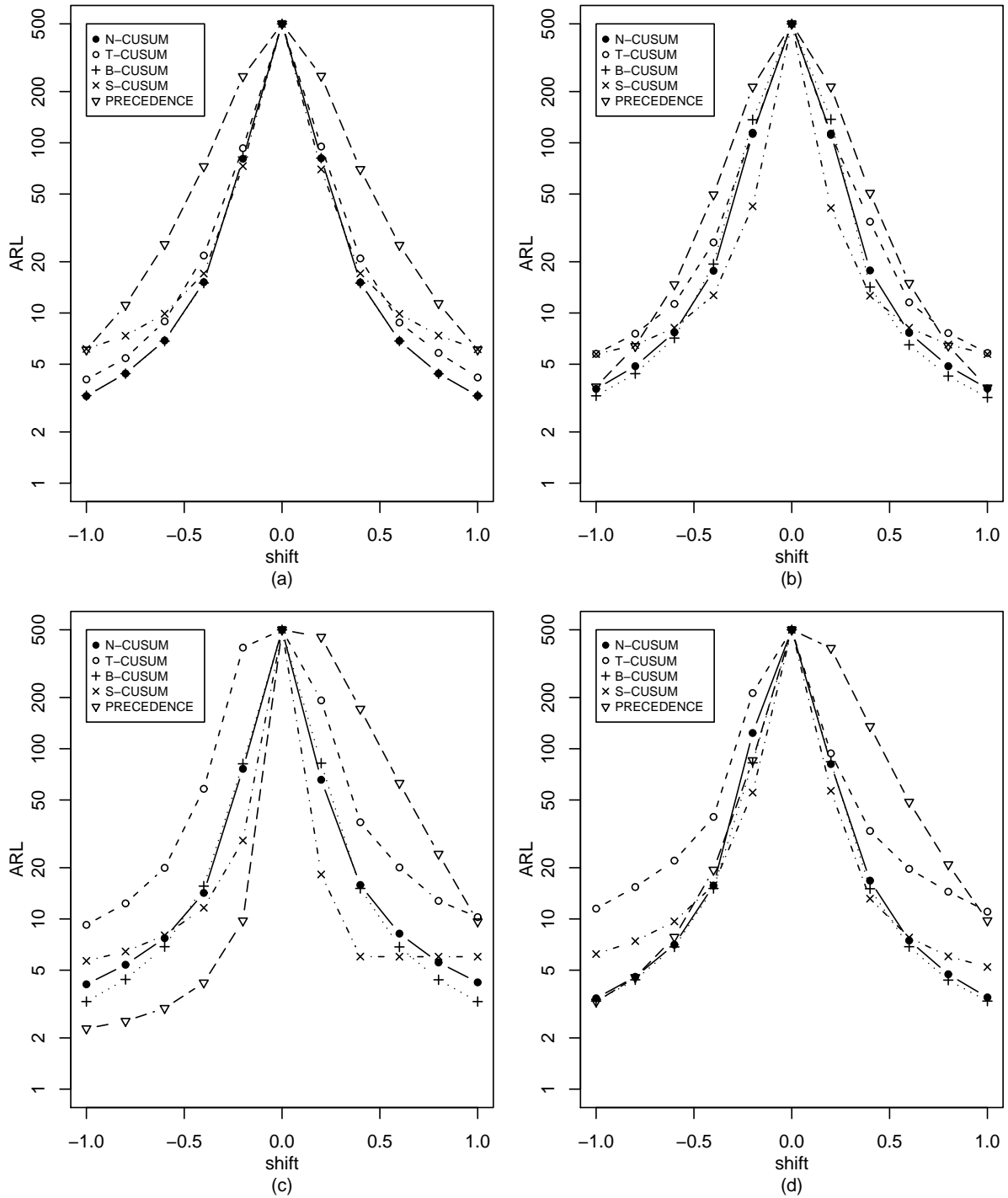


Figure 3: OC ARL values of five control charts when  $ARL_0 = 500$ ,  $M = 500$ ,  $m = 5$ , and the actual IC process distribution is the standardized version of  $N(0, 1)$  (plot (a)),  $t(4)$  (plot (b)),  $\chi^2(1)$  (plot (c)), and  $\chi^2(4)$  (plot (d)). Procedure parameters of the control charts are chosen to be the ones that minimize their OC ARL values when detecting the shift of 0.6. Scale on the  $y$ -axis is in natural logarithm.

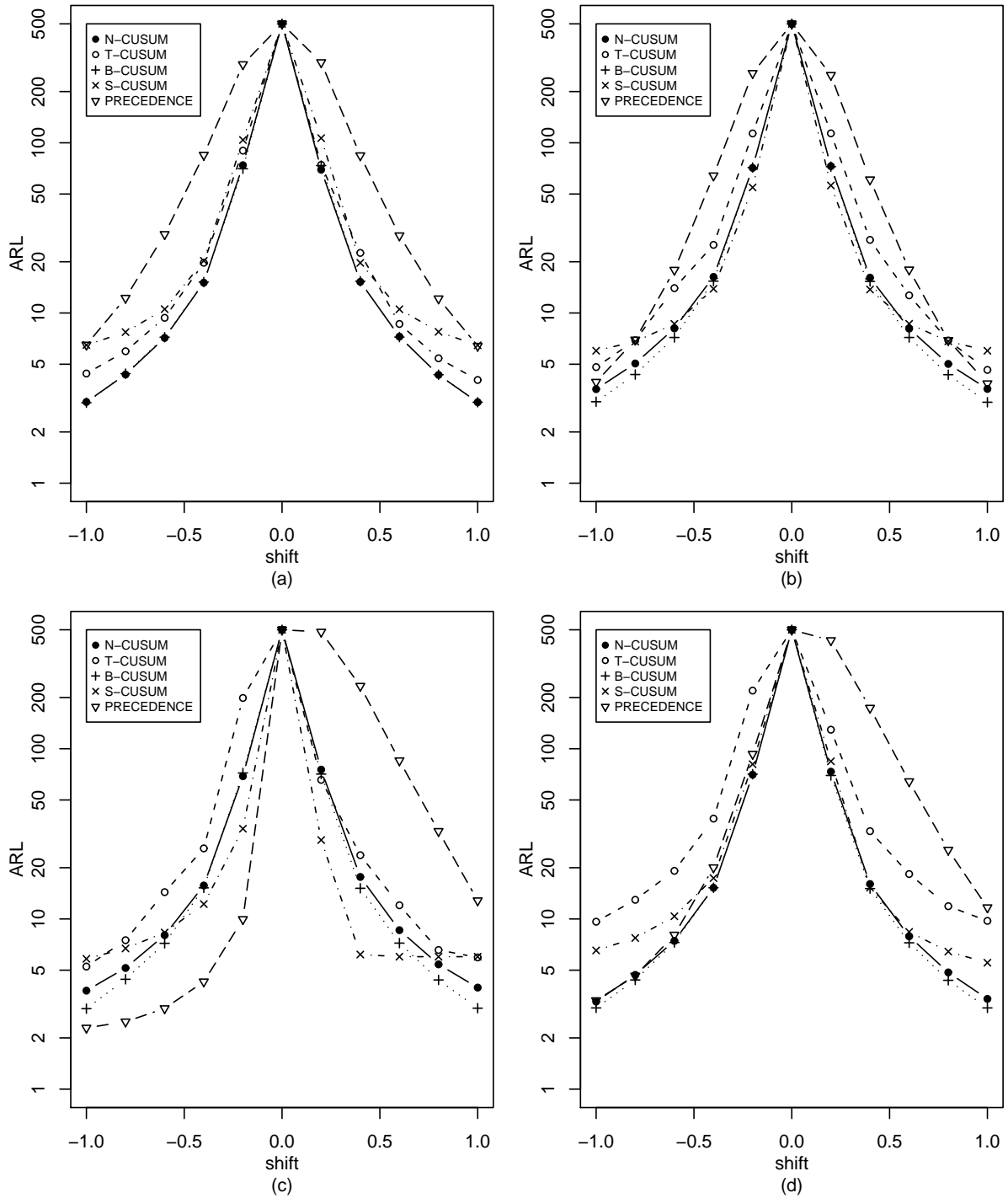


Figure 4: OC ARL values of five control charts when  $ARL_0 = 500$ ,  $M = 200$ ,  $m = 5$ , and the actual IC process distribution is the standardized version of  $N(0,1)$  (plot (a)),  $t(4)$  (plot (b)),  $\chi^2(1)$  (plot (c)), and  $\chi^2(4)$  (plot (d)). Procedure parameters of the control charts are chosen to be the ones that minimize their OC ARL values when detecting each individual shift. Scale on the  $y$ -axis is in natural logarithm.

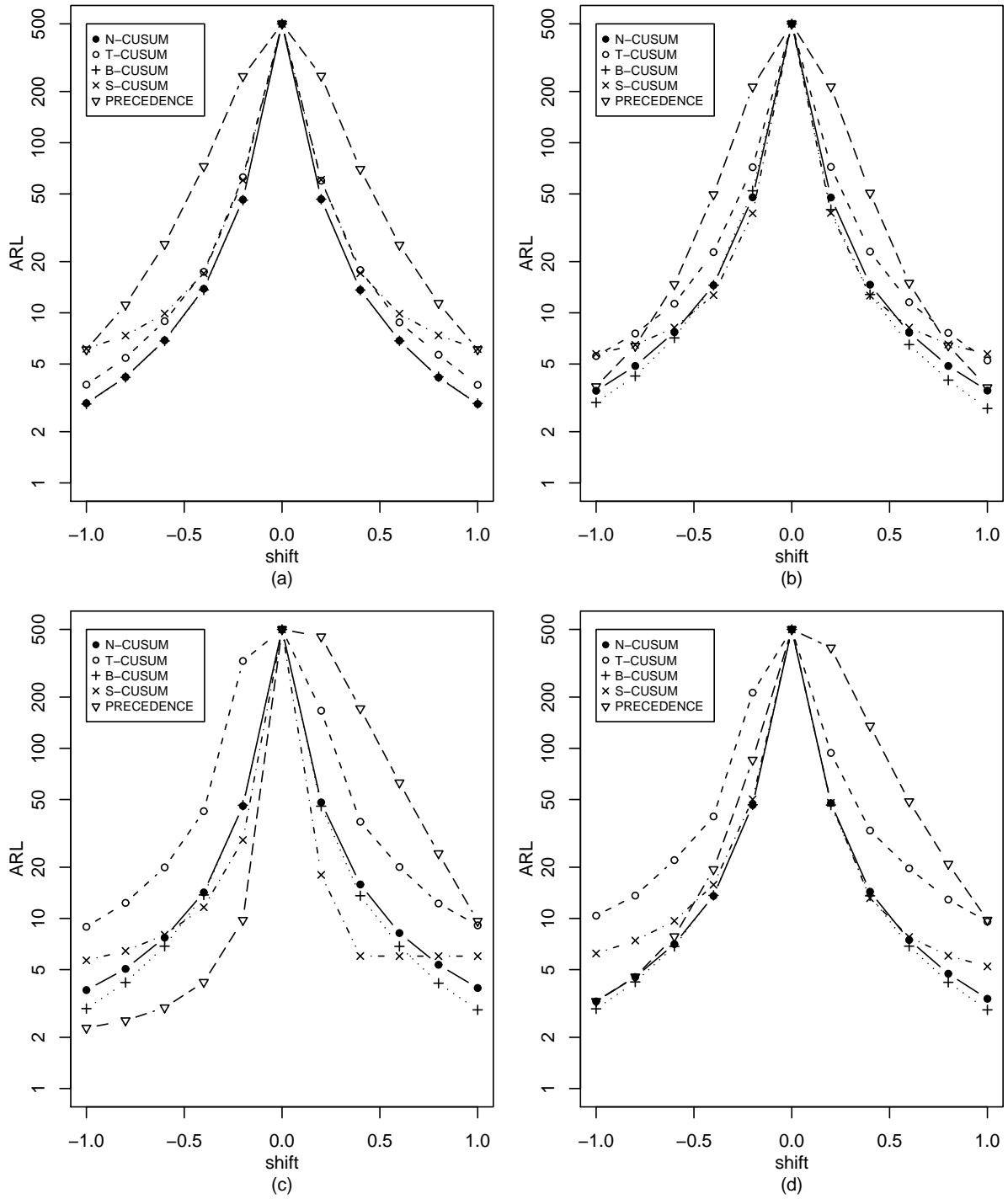


Figure 5: OC ARL values of five control charts when  $ARL_0 = 500$ ,  $M = 500$ ,  $m = 5$ , and the actual IC process distribution is the standardized version of  $N(0,1)$  (plot (a)),  $t(4)$  (plot (b)),  $\chi^2(1)$  (plot (c)), and  $\chi^2(4)$  (plot (d)). Procedure parameters of the control charts are chosen to be the ones that minimize their OC ARL values when detecting each individual shift. Scale on the y-axis is in natural logarithm.

in all cases. When the skewness is large (cf., plot (c)), we notice three major differences from the results in plot (d): (i) when the shift is in the direction of the shorter tail of the IC distribution, the PRECEDENCE chart performs extremely well, (ii) when the shift is in the direction of the longer tail of the IC distribution, the S-CUSUM chart performs well, especially when the shift is small to moderate, and (iii) it seems that the B-CUSUM chart is more effective for detecting the mean shift than the N-CUSUM chart in such cases, especially when the shift is large.

Based on the above IC and OC results, we may have the following conclusions about the five control charts. (i) When the IC distribution is normal, the N-CUSUM chart is the one to use. In such cases, the B-CUSUM and T-CUSUM also performs well in both the IC and OC situations. (ii) When the IC distribution is non-normal, the N-CUSUM and T-CUSUM charts may not be reliable because their actual  $ARL_0$  values could be quite different from the nominal  $ARL_0$  value. In such cases, the B-CUSUM, S-CUSUM, and PRECEDENCE charts are reliable; but, the S-CUSUM chart is efficient only when the mean shift is quite small, and the PRECEDENCE chart is efficient only when the IC distribution is skewed and the mean shift is in the direction of its shorter tail. (iii) The B-CUSUM chart has a reasonably good performance in all cases considered. Therefore, in a given application, if we are not sure whether the IC distribution is normal, the B-CUSUM chart might be the one to consider.

## 4 An Application

We illustrate the proposed method using a real-data example about daily exchange rates between Korean Won and US currency Dollar between March 28, 1997 and December 02, 1997. During this period of time, the daily exchange rates were quite stable early on and became unstable starting from early August. This can be seen from Figure 6(a) in which 162 daily exchange rates (Won/Dollar) observed in that period are shown. As a side note, the world financial market experienced a serious crisis in the winter of 2007, and Korean Won was seriously affected by the crisis.

Like many other Phase II SPC procedures, our proposed procedure assumes that observations at different time points are independent of each other. However, for the exchange rate data, we found that observations at different time points are substantially correlated. Following the suggestions in Qiu and Hawkins (2001), we first pre-whiten the data using the following auto-regression model

that can be accomplished by the R function `ar.yw()`:

$$\begin{aligned}
 x(i) - 919.01 &= 0.86(x(i-1) - 919.01) + 0.05(x(i-2) - 919.01) - 0.09(x(i-3) - 919.01) + \\
 & 0.12(x(i-4) - 919.01) - 0.06(x(i-5) - 919.01) + 0.02(x(i-6) - 919.01) + \\
 & 0.25(x(i-7) - 919.01) - 0.40(x(i-8) - 919.01) + 0.16(x(i-9) - 919.01) + \epsilon(i), \\
 & \text{for } i = 10, 11, \dots, 162,
 \end{aligned}$$

where  $x(i)$  denotes the  $i$ -th observation of the exchange rate,  $\epsilon(i)$  is a white noise process with zero mean, and the order of the model is determined by the default Akaike's Information Criterion (AIC) in `ar.yw()`. The pre-whitened data are shown in Figure 6(b).

We then try to apply the related control charts considered in the previous section to the pre-whitened data. To this end, the first 96 residuals are used as an IC data, which correspond to the first 105 original observations, and the remaining residuals are used for testing. In Figure 6(a)-(b), the training and testing data are separated by a dashed vertical line. To take a closer look at the IC data and at the first several testing observations as well, the first 121 residuals are presented in Figure 6(c) again, in which the solid horizontal line denotes the sample mean of the IC data and the dashed vertical line separates the IC and testing data. From plot (c), it can be seen that there is an upward mean shift starting from the very beginning of the test data. The Shapiro test for checking the normality of the IC data gives a  $p$ -value of  $1.323 \times 10^{-4}$ , which implies that the IC data are significantly non-normal. To demonstrate this, the density histogram of the IC data is shown in Figure 6(d), along with its estimated density curve (solid) and the density curve of a normal distribution (dashed) with the same mean and standard deviation.

Because the T-CUSUM, S-CUSUM and PRECEDENCE charts can only be used for batch data and the current exchange rate data have a single observation at each time point, they are not considered here. Also, the N-CUSUM chart is not appropriate here because the IC distribution has been demonstrated to be non-normal. Therefore, only the B-CUSUM chart is used here. When the nominal IC ARL value is fixed at 200 and the allowance constant of the B-CUSUM chart is chosen to be 0.2, its charting statistics are shown in Figure 7, in which the dashed horizontal lines denote its control limits. From the figure, it can be seen that the B-CUSUM chart gives a signal of mean shift at  $i = 113$ .

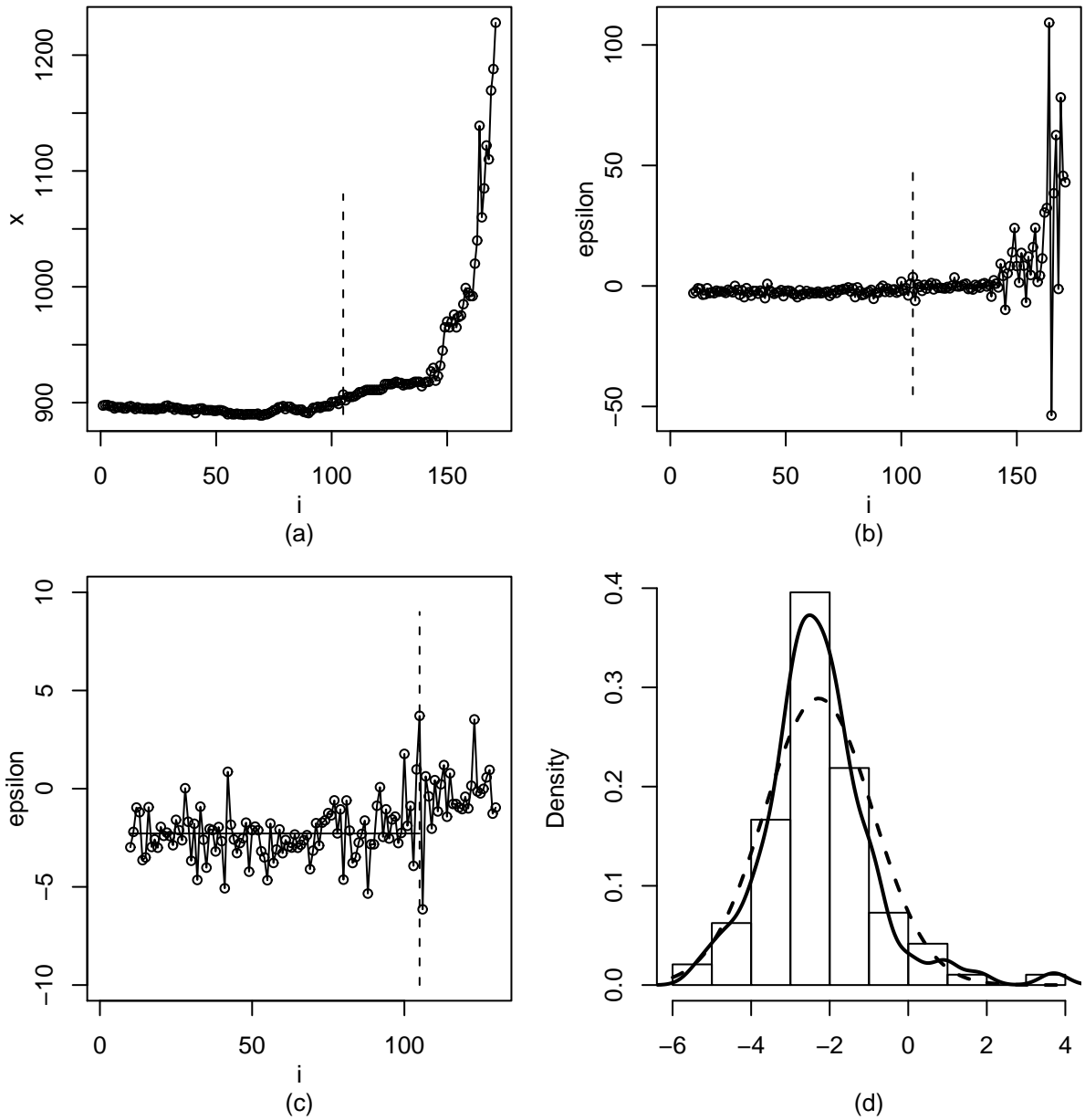


Figure 6: (a) Original observations of the exchange rates between Korean currency Won and US currency Dollar between March 28, 1997 and December 02, 1997. (b) Pre-whitened values of the original observations. (c) The first 121 pre-whitened values. (d) Density histogram, estimated density curve (solid) of the first 96 pre-whitened values (i.e., IC data), and the density curve of a normal distribution (dashed) with the same mean and variance as those of the IC data. In plots (a)–(c), the dashed vertical line separates the IC and testing data. In plot (c), the solid horizontal line denotes the sample mean of the IC data.

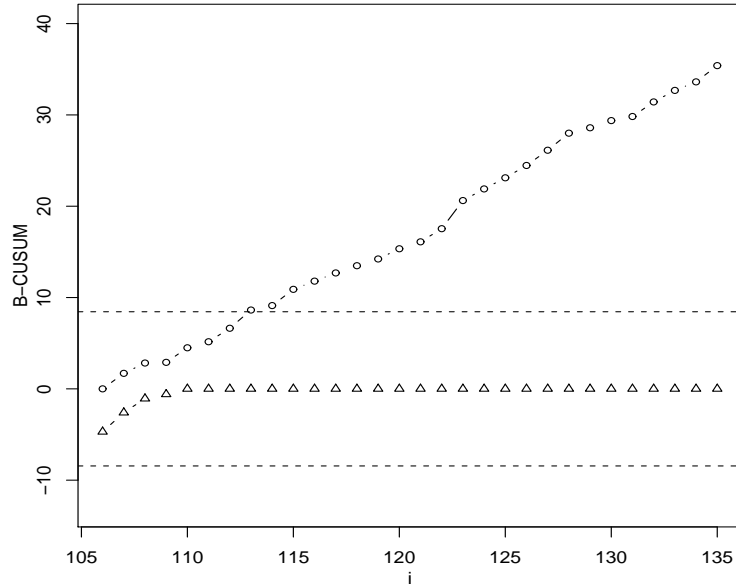


Figure 7: The B-CUSUM chart is applied to the exchange rate data. In the plot, the horizontal dashed lines denote its control limits, and the little circles and little triangles denote its upward and downward charting statistics, respectively.

## 5 Concluding Remarks

We have proposed two control charts B-CUSUM and T-CUSUM in this paper for applications in which the IC process distribution is not normal. These two charts are then compared with the traditional N-CUSUM chart and two recent distribution-free control charts S-CUSUM and PRECEDENCE in various cases. Based on the comparative study, we conclude that, when the IC process distribution is not normal, the N-CUSUM and T-CUSUM charts are not reliable, the S-CUSUM and PRECEDENCE charts are reliable but they are efficient for detecting the mean shift only in certain limited situations, and the B-CUSUM chart has a reasonably good performance in all cases considered. The comparative study presented in this paper is empirical. Much future research is required to confirm our conclusions by mathematically more rigorous arguments. In this paper, we focus on monitoring the process mean. In some applications, simultaneous monitoring of the process mean and variance would be of our interest, which also requires much future research.



## References

- Amin, R., Reynolds, M. R. Jr., and Bakir, S. T. (1995), “Nonparametric quality control charts based on the sign statistic,” *Communications in Statistics-Theory and Methods*, **24**, 1597–1623.
- Amin, R.W., and Widmaier, O. (1999), “Sign control charts with variable sampling intervals,” *Communications in Statistics: Theory and Methods*, **28**, 1961–1985.
- Bakir, S. T. (2006), “Distribution-free quality control charts based on signed-rank-like statistics,” *Communications in Statistics-Theory and Methods*, **35**, 743–757.
- Bakir, S.T., and Reynolds, M.R. Jr. (1979), “A nonparametric procedure for process control based on within group ranking,” *Technometrics*, **21**, 175–183.
- Chakraborti, S., and Eryilmaz, S. (2007), “A nonparametric Shewhart-type signed-rank control chart based on runs,” *Communications in Statistics-Simulation and Computation*, **36**, 335–356.
- Chakraborti, S., Eryilmaz, S., and Human, S. W. (2009), “A phase II nonparametric control chart based on precedence statistics with runs-type signaling rules,” *Computational Statistics and Data Analysis*, **53**, 1054–1065.
- Chakraborti, S., van der Laan, P., and Bakir, S. T. (2001), “Nonparametric control charts: an overview and some results,” *Journal of Quality Technology*, **33**, 304–315.
- Chou, Y.M., Polansky, A.M., and Mason, R.L. (1998), “Transforming non-normal data to normality in statistical process control,” *Journal of Quality Technology*, **30**, 133–141.
- Cook, R.D., and Weisberg, S. (1999), *Applied Regression Including Computing and Graphics*, New York: John Wiley & Sons.
- Hackl, P., and Ledolter, J. (1992), “A new nonparametric quality control technique,” *Communications in Statistics-Simulation and Computation*, **21**, 423–443.
- Hawkins, D.M., and Olwell, D.H. (1998), *Cumulative Sum Charts and Charting for Quality Improvement*, New York: Springer-Verlag.

- Hawkins, D.M., Qiu, P., and Chang Wook Kang (2003), “The changepoint model for statistical process control,” *Journal of Quality Technology*, **35**, 355–366.
- Johnson, N.L. (1949), “Systems of frequency curves generated by methods of translation,” *Biometrika*, **36**, 149–176.
- Lucas, J. M., and Crosier, R. B. (1982), “Robust CUSUM: a robust study for CUSUM quality control schemes,” *Communications in Statistics-Theory and Methods*, **11**, 2669–2687.
- Montgomery, D. C. (2009), *Introduction To Statistical Quality Control (6th edition)*, New York: John Wiley & Sons.
- Page, E.S. (1954), “Continuous inspection schemes,” *Biometrika*, **41**, 100-114.
- Qiu, P. (2008), “Distribution-free multivariate process control based on log-linear modeling,” *IIE Transactions*, **40**, 664–677.
- Qiu, P., and Hawkins, D. M. (2001), “A rank based multivariate CUSUM procedure,” *Technometrics*, **43**, 120–132.
- Qiu, P., and Hawkins, D. M. (2003), “A nonparametric multivariate cumulative sum procedure for detecting shifts in all directions,” *Journal of Royal Statistical Society (Series D) - The Statistician*, **52**, 151–164.
- Shapiro, S.S., and Wilk, M.B. (1965), “An analysis of variance test for normality (complete samples),” *Biometrika*, **52**, 591–611.
- Slifker, J.F., and Shapiro, S.S. (1980), “The Johnson system: selection and parameter estimation,” *Technometrics*, **22**, 239–246.
- Woodall, W.H. (2000), “Controversies and contradictions in statistical process control,” *Journal of Quality Technology*, **32**, 341–350.
- Yeh, A.B., Lin, D.K.-J., and Venkataramani, C., (2004), “Unified CUSUM control charts for monitoring process mean and variability,” *Quality Technology and Quantitative Management*, **1**, 65–85.