Conditional design of the Shewhart X chart with unknown process parameters based on median run length

Xuelong Hu, Philippe Castagliola, Anan Tang, Xiaojian Zhou

Conditional design of the Shewhart $\bar{X}$ chart with unknown process parameters based on median run length

XueLong Hu*

School of Management,
Nanjing University of Posts and Telecommunications,
Nanjing, China
E-mail: hxl0419@hotmail.com
*Corresponding author

Philippe Castagliola

Université de Nantes,
LS2N UMR CNRS, 6004,
Nantes, France
E-mail: philippe.castagliola@univ-nantes.fr

AnAn Tang

School of Management,
Nanjing University of Posts and Telecommunications,
Nanjing, China
E-mail: tanganan2@sina.com

XiaoJian Zhou

School of Management,
Nanjing University of Posts and Telecommunications,
Nanjing, China
E-mail: xjzhou8888@163.com

Abstract: Numerous researches have been done on the Shewhart $\bar{X}$ chart based on the average run length (ARL) metric. Since the shape of the run length (RL) distribution changes with the mean shift size, the median run length (MRL) is argued to be a better criterion for evaluating the performance of the Shewhart $\bar{X}$ chart. Moreover, when the process parameters are unknown, the Phase II properties of the Shewhart $\bar{X}$ chart are conditioned on the parameter estimates arising from different practitioners in Phase I. This variability among the estimated process parameters is usually called as the between-practitioners variability. In order to investigate this variability in the conditional MRL values, both the average MRL ($AMRL$) and the standard deviation of MRL ($SDMRL$) will be used together in our article. The performance analyses of the MRL based

Shewhart $\bar{X}$ chart are provided. To prevent too many lower in-control MRL values than the desired one, an appropriate bootstrap approach is adopted to adjust the control limits, and to further balance the in- and out-of-control MRL values of the Shewhart $\bar{X}$ chart.

Keywords: Shewhart $\bar{X}$ chart; median run length; estimated parameters; average median run length; Standard deviation of median run length.

Reference to this paper should be made as follows: XueLong, H. Castagliola, P. Anan, T. and XiaoJian Z. (2022) 'Conditional design of the Shewhart $\bar{X}$ chart with unknown process parameters based on median run length', *European Journal of Industrial Engineering*, Vol. x, No. x, pp.xxx–xxx.

Biographical notes: XueLong Hu is currently an Associate Professor in the School of Management, Nanjing University of Posts and Telecommunications. His main areas of interest are in new statistical quality control techniques.

Philippe Castagliola graduated (PhD 1991) from the Université de Technologie de Compiégne (UTC), France. He is currently a professor at the Université de Nantes, Institut Universitaire de Technologie de Nantes, France, and he is also a member of the Institut de Recherche en Communications et Cybernétique de Nantes (IRCCyN), UMR CNRS 6597. He is an associate editor for the Journal of Quality Technology and Quantitative Management and for the International Journal of Reliability, Quality and Safety Engineering. His research activity includes developments of new SPC techniques.

AnAn Tang is currently a lecture in the School of Management, Nanjing University of Posts and Telecommunications. His main areas of interest are in non-parametric statistical quality control techniques.

XiaoJian Zhou is currently an Associate Professor in the School of Management, Nanjing University of Posts and Telecommunications. His main areas of interest are in robust parameter design and statistical quality control.

1 Introduction

Control charts are one of the most common used tools in Statistical Process Control (SPC). When monitoring the mean of a quality characteristic, such as a dimension, weight or volume, the Shewhart $\bar{X}$ chart is usually used to detect assignable cause(s) in the quality characteristic. This chart is especially useful to monitor relatively large shifts in the process mean (see Montgomery (2009)).

In practice, there are generally two phases (Phases I and II) in the implementation of control charts. In Phase I, an in-control reference set of observations is collected by practitioners to estimate the unknown process parameters, which are then used to set the control limits of a Phase II control chart. In Phase II, samples from the process are prospectively collected, and are plotted on the control chart for detecting assignable cause(s) in the process. A Phase II control chart's performance is usually evaluated by

assuming known process parameters but, in practice, the process parameters are often unknown and are estimated from an in-control reference set of observations. Under this case, a inherent variability in these estimated process parameters is encountered because different practitioners will use different Phase I samples (or different Phase I observations). Thus the control chart's performance will vary among practitioners. This leads to the *between-practitioners variability* effect in the properties of control charts (see Saleh, Mahmoud, Keefe and Woodall (2015)).

Control charts' performance are usually assessed through various properties of the run length (RL) distribution, where RL is defined as the number of chart statistics plotted until the chart signals. One of the most commonly used metrics of a Phase II chart is the average run length (ARL), which is defined as the average number of plotted statistics until the chart signals. As pointed out in Montgomery (2009), the shape of the run length distribution changes for different mean shift sizes, i.e. from highly right skewed when the process is in-control to nearly symmetric when the shift size is large. Therefore, the interpretation only based on the ARL corresponding to a highly skewed distribution is certainly different from a nearly symmetric distribution. For example, the Shewhart $\bar{X}$ chart is usually designed to have an in-control $ARL_0 = 370$ by choosing suitable parameters. Using the same parameters, the in-control MRL_0 is 257, which means 50% of all the RLs are less than 257. As a comparison, about 63% of all the run lengths are less than 370. In recent years, the use of the ARL only as a measure of control charts' performance has been criticized in some researches, see Jones et al. (2004), Jensen et al. (2006) and Chakraborti (2007), to name a few. They all pointed out that the interpretation based on the ARL is sometimes misleading as the shape of the run length distribution changes with the magnitude of the mean shift. On the other hand, the MRL does not have the same interpretation problem and, because of its meaningful interpretation, Gan (1993), Gan (1994), Khoo et al. (2011), Khoo et al. (2012), Teoh, Chong, Khoo, Castagliola and Yeong (2016), Lee and Khoo (2017) and Tang et al. (2018) have done some researches on the area of MRL based control charts with known parameters. As stated above, a Phase II control chart's properties need to be investigated with the process parameters estimated from a Phase I dataset. Considering this issue, when the process parameters are estimated, some researches have been conducted on the MRL based control charts: readers may refer to Chakraborti (2007), Teoh et al. (2015), Teoh, Khoo, Castagliola and Lee (2016) and You et al. (2016). All these researches focused on the marginal MRL performance of control charts with estimated parameters. Therefore, this article suggests the MRL as a design measure of the Shewhart $\bar{X}$ chart. A detailed motivation of the use of the MRL instead of the ARL in the design of the Shewhart $\bar{X}$ chart can be found in Khoo (2004), Chakraborti (2007) and Teoh, Khoo, Castagliola and Lee (2016).

When the process parameters are estimated, the between-practitioners variability generates a inherent variability in the in-control ARL values. Using just the ARL averaged over the distribution of the parameter estimates does not give a complete picture of the performance of control charts. In addition to this, much researches recommended the standard deviation of the ARL ($SDARL$) as a new metric, to measure the between-practitioners variability in control charts, see Saleh, Mahmoud, Keefe and Woodall (2015), Saleh, Mahmoud, Jones-Farmer, Zwetsloot and Woodall (2015), Faraz et al. (2015), Hu and Castagliola (2017), Keefe et al. (2015), Aly et al. (2015), Aly et al. (2016), Geodhart, da Silvab, Schoonhoven, Epprecht, Chakraborti, Does and Veiga (2017) and Geodhart,

Schoonhoven and Does (2017). These authors use the $SDARL$ metric together with $AARL$ metric to make recommendations for the amount of Phase I samples. Using the MRL as a measure, and considering the *between-practitioners variability* effect in the MRL values of the Shewhart $\bar{X}$ chart, the MRL averaged over the distribution of the parameter estimates ($AMRL$) will also not give a complete picture of the performance of the Shewhart $\bar{X}$ chart with unknown parameters. This is similar to the ARL based Shewhart $\bar{X}$ chart with unknown parameters (Saleh, Mahmoud, Keefe and Woodall (2015)). To measure this variability, the standard deviation of the MRL ($SDMRL$) will be used as an alternative to evaluate the performance of the Shewhart $\bar{X}$ chart. Thus, the MRL based Shewhart $\bar{X}$ chart with unknown process parameters will be investigated in this article by using both the $AMRL$ and $SDMRL$.

The rest of this article is organized as follows: First, in Section 2, we introduce the run length properties of the Shewhart $\bar{X}$ chart, where the properties for the known and unknown process parameters cases are provided in Sections 2.1 and 2.2, respectively. We then introduce in Section 3 a brief background of the MRL based Shewhart $\bar{X}$ chart. Based on the average, the standard deviation, and several percentiles of the MRL distribution, we investigate in Section 4 the MRL performance of the Shewhart $\bar{X}$ chart with estimated parameters. In Section 5, the adjustments of control limits, and the corresponding performance of the Shewhart $\bar{X}$ chart are discussed. In Section 6, an example is given to illustrate the use of the MRL based Shewhart $\bar{X}$ chart. Finally, some conclusions and recommendations are made in Section 7.

2 The run length properties of the Shewhart $\bar{X}$ chart

2.1 The Shewhart $\bar{X}$ chart with known parameters

Suppose that the quality characteristic Y is a normal random variable, $Y \sim N(\mu_0 + \delta\sigma_0, \sigma_0)$, where μ_0 and σ_0 are the in-control mean and standard deviation, respectively, and δ is the magnitude of the standardized mean shift. The process is considered to be in-control when $\delta = 0$. Otherwise ($\delta \neq 0$), the process is out-of-control. A Phase II sample $\{Y_{i,1}, Y_{i,2}, \ldots, Y_{i,n}\}$ of size n are then collected successively at sampling time i, and the sample mean $\bar{Y}_i = \frac{1}{n}\sum_{j=1}^{n} Y_{i,j}$ is used for process monitoring, where $\bar{Y}_i \sim N(\mu_0 + \delta\sigma_0, \sigma_0/\sqrt{n})$.

For the Shewhart $\bar{X}$ chart, an out-of-control signal is obtained when a sample mean $\bar{Y}_i$ fall above the upper control limit (UCL) or below the lower control limit (LCL), where

$$UCL = \mu_0 + K\sigma_0$$
$$LCL = \mu_0 - K\sigma_0$$

where K is a constant that determines the distance, in terms of standard deviation units, of the control limits from the centre line of the control chart and, it is selected in order to obtain a certain in-control performance.

When the parameters μ_0 and σ_0 are known, the control limits UCL and LCL are constants, and the probability that a sample mean $\bar{Y}_i$ falls inside the control limits is,

$$\begin{aligned} p &= P(LCL \leq Y_i \leq UCL) \\ &= F_N(K\sqrt{n} - \delta\sqrt{n}) - F_N(-K\sqrt{n} - \delta\sqrt{n}) \end{aligned}$$

where $F_N(\cdot)$ is the cumulative distribution function (c.d.f.) of the standard normal distribution. Since the run length (RL) of the Shewhart $\bar{X}$ chart follows a geometric distribution(Montgomery, 2009), the probability distribution function (p.d.f.) $f_{RL}(\ell)$ and c.d.f. $F_{RL}(\ell)$ of the RL are defined for $\ell = \{1, 2, 3, \dots\}$ as follows:

$$f_{RL}(\ell) = P(RL = \ell) = p^{\ell-1}(1-p) \tag{1}$$

$$F_{RL}(\ell) = P(RL \leq \ell) = 1 - p^{\ell} \tag{2}$$

Moreover, the 100ξ $(0 < \xi < 1)$ percentage point RL_ξ of the run length distribution is determined as follows (Gan (1994)):

$$F_{RL}(RL_\xi - 1) \leq \xi \qquad \text{and} \qquad F_{RL}(RL_\xi) > \xi. \tag{3}$$

For the geometric distribution in (2), by solving $F_{RL}(RL_\xi - 1) = 1 - p^{RL_\xi - 1} \leq \xi$ and $F_{RL}(RL_\xi) = 1 - p^{RL_\xi} > \xi$, we can get

$$\frac{\ln(1-\xi)}{\ln(p)} < RL_\xi \leq \frac{\ln(1-\xi)}{\ln(p)} + 1 \tag{4}$$

Since the RL_ξ of a control chart is an integer, we can further obtained that

$$RL_\xi = \begin{cases} \lceil \frac{\ln(1-\xi)}{\ln(p)} \rceil, & \text{if } \frac{\ln(1-\xi)}{\ln(p)} \text{ is not an integer} \\ \frac{\ln(1-\xi)}{\ln(p)} + 1, & \text{if } \frac{\ln(1-\xi)}{\ln(p)} \text{ is an integer} \end{cases} \tag{5}$$

where $\lceil \cdot \rceil$ is the rounding up operator. Using these equations enable the computation of any percentage point of the run length distribution. In particular, if we set $\xi = 0.5$, we can obtain the MRL of the Shewhart $\bar{X}$ chart.

2.2 The Shewhart $\bar{X}$ chart with unknown process parameters

Suppose that $\{X_{i,1}, \dots, X_{i,n}\}$, $i = 1, 2, \dots, m$ are m in-control samples, each with size n, from Phase I. Let us assume X is a normal random variable, i.e. $X_{i,j} \sim N(\mu_0, \sigma_0)$, the estimators $\hat{\mu}_0$ and $\hat{\sigma}_0$ used for μ_0 and σ_0 are respectively (see Jensen et al. (2006)):

$$\hat{\mu}_0 = \frac{1}{m}\sum_{i=1}^{m} \bar{X}_i, \tag{6}$$

$$\hat{\sigma}_0 = \sqrt{\frac{1}{m(n-1)}\sum_{i=1}^{m}\sum_{j=1}^{n}(X_{i,j} - \bar{X}_i)^2}, \tag{7}$$

with $\bar{X}_i = \frac{1}{n}\sum_{j=1}^{n} X_{i,j}$. The control limits of the Shewhart $\bar{X}$ chart become

$$\widehat{UCL} = \hat{\mu}_0 + K\hat{\sigma}_0, \tag{8}$$

$$\widehat{LCL} = \hat{\mu}_0 - K\hat{\sigma}_0. \tag{9}$$

For fixed values of $\hat{\mu}_0$ and $\hat{\sigma}_0$, the conditional probability $\hat{p}$ of a sample falling inside the control limits of the Shewhart $\bar{X}$ chart is $\hat{p} = P(\widehat{LCL} \leq \bar{Y}_i \leq \widehat{UCL}|\hat{\mu}_0, \hat{\sigma}_0)$. Therefore, by doing some simple mathematical manipulations, we can obtain the following results,

$$\hat{p} = F_N\left((\hat{\mu}_0 + K\hat{\sigma}_0 - \mu_0 - \delta\sigma_0)\frac{\sqrt{n}}{\sigma_0}\right) - F_N\left((\hat{\mu}_0 - K\hat{\sigma}_0 - \mu_0 - \delta\sigma_0)\frac{\sqrt{n}}{\sigma_0}\right) \tag{10}$$

For fixed values of $\hat{\mu}_0$ and $\hat{\sigma}_0$, the conditional probability mass function (p.m.f.) $\hat{f}_{RL}(\ell)$ and c.d.f. $\hat{F}_{RL}(\ell)$ of the RL for the Shewhart $\bar{X}$ chart are presented in (1) and (2) by replacing p in (1) with $\hat{p}$ in (10). Similarly, the conditional MRL ($\widehat{MRL}$) of the Shewhart $\bar{X}$ chart with estimated parameters can be computed using (3) by setting $\xi = 0.5$ and replacing p with the corresponding estimate $\hat{p}$.

If we define the random variables $U = (\hat{\mu}_0 - \mu_0)\frac{\sqrt{n}}{\sigma_0}$ and $V = \frac{\hat{\sigma}_0}{\sigma_0}\sqrt{n}$, the conditional probability $\hat{p}$ simplifies to,

$$\hat{p} = F_N\left(U + K \times V - \delta\sqrt{n}\right) - F_N\left(U - K \times V - \delta\sqrt{n}\right). \tag{11}$$

Concerning the distributions of the random variables U and V, they are respectively given as follows (see Zhang et al. (2011)):

$$f_U(u|m) = f_N\left(u\left|0, \frac{1}{\sqrt{m}}\right.\right), \tag{12}$$

$$f_V(v|m,n) = 2v f_\gamma\left(v^2\left|\frac{m(n-1)}{2},\ \frac{2n}{m(n-1)}\right.\right), \tag{13}$$

where $f_N(\cdot)$ is the p.d.f. of a normal distribution with mean 0 and standard deviation $\frac{1}{\sqrt{m}}$ and, $f_\gamma(\cdot)$ is the p.d.f. of the gamma distribution with parameters $\frac{m(n-1)}{2}$ and $\frac{2n}{m(n-1)}$.

Since U and V are independent random variables, then the unconditional RL distribution is the conditional $f_{RL}(\ell)$ and $F_{RL}(\ell)$ in (1) and (2) integrated over the the distributions of U and V, which are obtained as follows:

$$f_{RL}(\ell) = \int_{-\infty}^{+\infty}\int_{0}^{+\infty} \hat{p}^{\ell-1}(1-\hat{p}) f_U(u|m) f_V(v|m,n)\mathrm{d}v\mathrm{d}u, \tag{14}$$

$$F_{RL}(\ell) = \int_{-\infty}^{+\infty}\int_{0}^{+\infty} (1-\hat{p}^{\ell}) f_U(u|m) f_V(v|m,n)\mathrm{d}v\mathrm{d}u, \tag{15}$$

Moreover, the unconditional MRL, denoted as $AMRL$, and the standard deviation of MRL, denoted as $SDMRL$, of the Shewhart $\bar{X}$ chart with estimated parameters can be obtained as follows:

$$AMRL = E(\widehat{MRL}) = \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \widehat{MRL} \times f_U(u|m) f_V(v|m,n) \mathrm{d}v \mathrm{d}u, \quad (16)$$

$$SDMRL = \sqrt{E(\widehat{MRL}^2) - (E(\widehat{MRL}))^2}, \quad (17)$$

where

$$E(\widehat{MRL}^2) = \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \widehat{MRL}^2 \times f_U(u|m) f_V(v|m,n) \mathrm{d}v \mathrm{d}u. \quad (18)$$

In order to measure the average and variability of the conditional $\widehat{MRL}$ performance of the Shewhart $\bar{X}$ chart with unknown process parameters, the $AMRL$, $SDMRL$ and some percentiles of $\widehat{MRL}$ can also be obtained using Monte Carlo simulations.

3 Brief background of the interpretation of the MRL based Shewhart $\bar{X}$ chart

A traditional measure of a control chart's property is the ARL. However, the in-control RL distribution of the Shewhart $\bar{X}$ chart is highly right-skewed, the mean of the RL is greater than its median, and thus it is usually not a fair representation of a "typical" or the "center" measure of the RL. For this reason, some other percentiles of the RL distribution may provide important information about the distribution and, consequently about the control chart's properties. For example, Khoo (2004) and Chakraborti (2007) pointed out that using the MRL to design a chart leads to a more accurate picture of a chart's performance, especially for highly skewed RL distribution. An MRL of τ indicates that 50% of all of the RLs are less than τ.

In order to draw attention to the usage of the MRL based Shewhart $\bar{X}$ chart, some percentiles of the RL and the ARLs are presented in Table 1, for different values of K and δ when $n = 5$. From Table 1, we can note that the difference between the ARL and MRL (50% percentile of RL) is large when the process is in-control ($\delta = 0$). This difference decreases when δ increases. For example, when $K = 1.2553$ and $\delta = 0$, the MRL and ARL values of the Shewhart $\bar{X}$ chart are 139 and 200, respectively. While for $\delta = 1$, the MRL and ARL are 3 and 3.52, respectively. This fact shows that the shape of the RL changes with the mean shift size δ. Since the RL distribution is highly skewed, especially when δ is small, a typical measure of the RL distribution is the MRL. The information from the usage of MRL acts as a good practical guidance to practitioners in studying the property of the Shewhart $\bar{X}$ chart. For more discussion about this topic, readers may refer to Khoo (2004), Chakraborti (2007), Teoh, Chong, Khoo, Castagliola and Yeong (2016), Teoh, Khoo, Castagliola and Lee (2016) and Teoh et al. (2015).

(Please Insert Table 1 Here)

4 The MRL performance of the Shewhart $\bar{X}$ chart with unknown process parameters

The MRL based control charts with unknown process parameters are mostly evaluated based on the 50% percentile of the unconditional c.d.f. $F_{RL}(\ell)$ of the RL as in (15) (Teoh et al. (2015), You et al. (2016) and Teoh, Khoo, Castagliola and Lee (2016)). The 50% percentile obtained from (15) doesn't reflect the between-practitioners variability in control charts and, unfortunately, it doesn't account for the variability in the MRL values. For a larger number of the Shewhart $\bar{X}$ chart, the corresponding MRL values can be widely dispersed and thus, a practitioner can't be sure of his/her specific chart's MRL performance.

The conditional in-control MRL properties of the Shewhart $\bar{X}$ chart are first investigated by considering different values of m varying from 20 to 500 and the desired in-control $MRL \in \{69, 139, 257, 347\}$ (denoted as MRL_0) when $n = 5$. Based on 10^5 simulated Shewhart $\bar{X}$ chart, the corresponding $AMRL$, $SDMRL$ and several percentiles of these in-control MRL values are presented in Table 2. The value of parameter K is selected to produce the desired MRL_0. From Table 2, we may note a large variation in the in-control MRL values, especially for small values of m. For example, when $MRL_0 = 257$ in Table 2, 5% of the in-control MRL values are smaller than 106 and, 5% of the in-control MRL values are larger than 543 and the $SDARL = 148.41$ if we use $m = 50$ Phase I samples to estimate the process parameters. The $SDMRL$ value is too large compared to the desired $MRL_0 = 257$. This fact is generally not acceptable in practice, because the specific practitioner may has her/his own $\bar{X}$ chart's MRL value far away from the desired MRL_0. For other values of the desired MRL_0, a similar trend in the variation also appears in the in-control MRL values. In order to display this variation in the in-control MRL values more clearly, the histograms of the in-control MRL values based on 10^5 simulated Shewhart $\bar{X}$ charts are also presented in Figure 1 for different number of Phase I samples $m \in \{50, 100, 200, 400\}$ when $MRL_0 = 257$ and $n = 5$. A large variation of nearly $SDMRL = 45.35$ can also be noted from Figure 1 (d) even with $m = 400$ Phase I samples.

Moreover, from Table 2, it can also be noted that with the increase in the amount m of Phase I samples, the lower percentiles increases and, the upper percentiles and the $SDMRL$ values decrease, thus illustrating that the variability in the in-control MRL distribution decreases. For example, when $MRL_0 = 257$ and $n = 5$, if we increases the number m of Phase I samples from 20 to 500, the 10% percentiles increases from 78 up to 209, the 90% percentiles decreases from 588 up to 310, and the $SDMRL$ value decreases from 312.65 up to 40.51.

(Please Insert Table 2 Here)
(Please Insert Figure 1 Here)

The results presented in Table 2 show a large variation in the in-control MRL properties of the Shewhart $\bar{X}$ chart, especially with a small number of Phase I samples. This variation is the so-called between-practitioners variability. Considering this variability, numerous researches have been done on the conditional properties of different type control charts,

for example, to name a few, Faraz et al. (2015), Geodhart, Schoonhoven and Does (2017); Saleh, Mahmoud, Keefe and Woodall (2015), Geodhart, da Silvab, Schoonhoven, Epprecht, Chakraborti, Does and Veiga (2017) and Hu and Castagliola (2017) all focused on the Shewhart type chart, Saleh et al. (2016) focused on the CUSUM type chart and, Saleh, Mahmoud, Jones-Farmer, Zwetsloot and Woodall (2015) and Hu et al. (2018) focused on the EWMA type chart. These researches were all focused on the ARL based control charts with unknown process parameters, and they also recommended the usage of the $SDARL$ to account for the between-practitioners variability. For the MRL based Shewhart $\bar{X}$ chart, considering this variability, a direct measure to evaluate its conditional properties would be the $SDMRL$.

In order to measure the variation in the in-control MRL values, we also focused on the $AMRL$ and $SDMRL$ values of the Shewhart $\bar{X}$ chart with unknown process parameters. Some researchers considered a $SDARL$ value about 10% of the desired ARL_0 as small enough (see Zhang et al. (2014) and Saleh, Mahmoud, Jones-Farmer, Zwetsloot and Woodall (2015)). Similar to these researches on the ARL based control charts, we consider $SDMRL$ values about 10% of the desired MRL_0 as being small enough in this article. Under this case, the conditional MRL performance of the specific Shewhart $\bar{X}$ chart would be close to the desired MRL_0. Now, the question is "How about the number of Phase I samples for the Shewhart $\bar{X}$ chart to have an $SDMRL$ value about 10% of the desired MRL_0?" To answer this question, Tables 3 to 6 present the $AMRL$ and $SDMRL$ values for different combinations of m ranging from 20 to 5000, and the desired $MRL_0 \in \{69, 139, 257, 347\}$ when $n \in \{3, 5, 7, 9\}$. The $SDMRL$ values that are about 10% of the desired MRL_0 are bolded in each table. The last row ($m = \infty$) corresponds to the cases for which the process parameters are known. The desired MRL_0 values are chosen so that the corresponding Shewhart $\bar{X}$ chart have ARL_0 values equal to $ARL_0 \in \{100, 200, 370.4, 500\}$, respectively. These ARL_0 values are usually used to evaluate the properties of different ARL based control charts.

From Table 3, if one only considers the $AMRL$ values to determine the required number of Phase I samples, we may note that about $m \simeq 100$ Phase I samples are needed to have an in-control relative difference between the $AMRL$ and the desired MRL_0 smaller than 5% when $MRL_0 = 69$ and $n = 3$. The relative difference is defined as $\Delta = (|AMRL - MRL_0|)/MRL_0 \times 100\%$, and in the case studied here $\Delta = (71.84 - 69)/69 \times 100\% = 4.12\%$. We consider the relative difference smaller than 5% is sufficiently enough. Under this case, the average MRL properties of the Shewhart $\bar{X}$ chart is similar to the desired MRL_0. Although $m \simeq 100$ Phase I samples provide an $AMRL$ value close to the desired MRL_0, the $SDMRL = 28.65$ is still large compared with the desired $MRL_0 = 69$. Considering the variability in the in-control MRL values, about $m \simeq 1500$ Phase I samples are needed for the Shewhart $\bar{X}$ chart to have an in-control $SDMRL$ to be about 10%, i.e. 69×10%=6.9, of the desired MRL_0 (see the bolded entry in column 3). We may see that the recommended $m \simeq 100$ Phase I samples based on the relative difference between $AMRL$ and MRL_0 is far smaller than 1500. Practitioners should realize the importance of taking into account both $AMRL$ and $SDMRL$ when choosing the number m of Phase I samples.

Moreover, with the increase in MRL_0, much more Phase I samples are needed to have a $SDMRL$ value about 10% of the desired MRL_0. For example, when $MRL_0 = 347$,

the number of Phase I samples increases up to $m \simeq 2800$ to have a $SDMRL = 34.34$ (see Table 3). For smaller sample sizes, like $n \in \{3, 5\}$, the relative difference Δ increases with an increase in the desired value of MRL_0 for fixed value of m. For example, in Table 3, when $m = 30$, the relative difference Δ for $MRL_0 = 69$ and $MRL_0 = 139$ are $(79.28 - 69)/69 \times 100\% = 14.89\%$ and $(171.33 - 139)/139 \times 100\% = 23.25\%$, respectively. Moreover, in Table 6, even when $n = 9$, we may note that $m \simeq 20$ Phase I samples is large enough for the Shewhart $\bar{X}$ chart to have a Δ smaller than 5%, but at least about $m \simeq 400$ Phase I samples to have a $SDMRL$ value about 10% of the desired $MRL_0 = 69$ (see the bolded entry in column 3). Through the results presented in Tables 3 to 6, more than $m \simeq 400$ Phase I samples are needed to get a relatively small variability in the in-control MRL values of the Shewhart $\bar{X}$ chart.

Furthermore, considering the variability in the in-control MRL values, it can also be noted that, in order to estimate the process parameters accurately, the overall number of observations ($m \times n$) required for the case $n = 3$ is larger than that the one required for the case $n = 9$. For example, the Shewhart $\bar{X}$ chart requires about 4500 observations ($m = 1500$ samples each with size $n = 3$) to get a $SDMRL$ value to be about 10% of the desired $MRL_0 = 69$ (see Table 3). While it requires about 3600 observations ($m = 400$ samples each with size $n = 9$) to get a $SDMRL$ value to be about 10% of the desired $MRL_0 = 69$ (see Table 6). This result is similar to the ARL based Shewhart $\bar{X}$ chart presented in Saleh, Mahmoud, Keefe and Woodall (2015).

(Please Insert Tables 3 to 6 Here)

Based on the results presented above, at least $m \simeq 400$ samples are needed for the Shewhart $\bar{X}$ chart to obtain a smaller in-control $SDMRL$ value. However, this large number of Phase I samples is sometimes hard to collect and too costly in practice. With a realistic amount of Phase I samples, a significant variation exists in the estimated process parameters. Thus, the practitioner has no confidence that the in-control MRL value of its specific Shewhart $\bar{X}$ chart will be close to the desired MRL_0. For this reason, in the following section, a bootstrap approach will be adopted to re-design the MRL based Shewhart $\bar{X}$ chart with unknown process parameters.

5 Adjustments of the control limits of the MRL based Shewhart $\bar{X}$ chart

To avoid too many in-control MRL values smaller than the desired MRL_0, one would have to adjust the control limits parameter K to account for the difference between the estimated and true process parameters. During the last few years, the *bootstrap approach* used in Gandy and Kvaloy (2013) became very popular, which can be adopted to guarantee the conditional in-control MRL values larger than the desired MRL_0 with a certain high probability $\alpha*$. Compared to the unadjusted parameter cases, where many MRL values are below the desired MRL_0, the bootstrap approach are expected to guarantee that this would happen only in $100(1 - \alpha*)\%$ of all the Shewhart $\bar{X}$ charts. In the literature, the designs of control charts for a guaranteed conditional performance have only been discussed by some researchers for the ARL based charts (see Saleh, Mahmoud, Jones-Farmer, Zwetsloot and Woodall (2015); Saleh et al. (2016), Faraz et al. (2015), Hu and Castagliola (2017), Aly et al.

(2015), Zhao and Driscoll (2016)). Therefore, the bootstrap approach is used in this article to re-design the parameters of the MRL based Shewhart $\bar{X}$ chart with unknown process parameters. The steps for the control limit adjustment based on the bootstrap method are enumerated below:

(1) Generate a Phase I data set $X_{i,j}$ composed of $i = 1, 2, \ldots, m$ samples $\{X_{i,1}, X_{i,2}, \ldots, X_{i,n}\}$ of size n from the (assumed) true in-control distribution $P = P(\mu_0, \sigma_0)$ of $X_{i,j}$, and compute the estimated parameters $\hat{\mu}_0$ and $\hat{\sigma}_0$ as in (6) and (7).

(2) Compute the quantity $K(\hat{P}, \hat{\theta})$ that produces a desired in-control MRL_0, where $\hat{P} = P(\hat{\mu}_0, \hat{\sigma}_0)$ is the estimated in-control distribution and, $\hat{\theta} = (\hat{\mu}_0, \hat{\sigma}_0)$ represents the estimated process parameters used to set control limits.

(3) Generate B samples from the distribution $\hat{P} = P(\hat{\mu}_0, \hat{\sigma}_0)$ and, compute these estimated process parameters $\hat{\theta}_i^* = (\hat{\mu}_i^*, \hat{\sigma}_i^*)$, $i = 1, 2, \ldots, B$.

(4) Compute the quantities $K(\hat{P}_i^*, \hat{\theta}_i^*)$ and $K(\hat{P}, \hat{\theta}_i^*)$, $i = 1, 2, \ldots, B$ to produce the desired in-control MRL_0, where $\hat{P}_i^* = (\hat{\mu}_i^*, \hat{\sigma}_i^*)$, $i = 1, 2, \ldots, B$ and $\hat{P}$ are the in-control distributions and, $\hat{\theta}_i^*$, $i = 1, 2, \ldots, B$ are used to set the control limits, respectively.

(5) Obtain the $100\alpha^*\%$ percentile $K_{\alpha^*}^*$ of the bootstrap distribution of $K(\hat{P}^*, \hat{\theta}^*) - K(\hat{P}, \hat{\theta}^*)$ and take $K(\hat{P}, \hat{\theta}) - K_{\alpha^*}^*$ as the adjusted control limit for the Shewhart $\bar{X}$ chart.

Following these bootstrap steps, it is expected that $100\alpha*\%$ of the in-control MRL values will be larger than the desired MRL_0 value. In our simulations, we assume that the true in-control distribution P is a standard normal $N(0, 1)$ distribution and the desired in-control MRL_0 is set to be 257. $m = 50$ Phase I samples each with size $n = 5$ are considered as realistic and are used in simulations to estimate the process parameters. The specified probability that the conditional in-control MRL values are larger than the desired MRL_0 is set as 90%, i.e. $\alpha^* = 0.9$, and $B = 1000$ bootstrap samples are generated from $\hat{P}$ to construct the adjusted control limits. By applying the bootstrap method, we adjust the control limit parameter K to guarantee the conditional in-control MRL performance for the Shewhart $\bar{X}$ chart to be larger than the desired $MRL_0 = 257$ with a high probability $\alpha^* = 0.9$.

In the bootstrap approach, we need to compute different parameters K from different settings of the Shewhart $\bar{X}$ chart: $K(\hat{P}, \hat{\theta})$, $K(\hat{P}_i^*, \hat{\theta}_i^*)$, $i = 1, 2, \ldots, B$ and $K(\hat{P}, \hat{\theta}_i^*)$, $i = 1, 2, \ldots, B$. Since MRL is an integer, it is possible to have more than one value of K for the desired MRL_0. All possible values of K satisfying MRL_0 are found and without loss of generality, the smallest one is used here. The computations are summarised as follows:

- For the quantity $K = K(\hat{P}, \hat{\theta})$, since the in-control distribution $\hat{P}$ is defined with the estimated parameters $\hat{\theta} = (\hat{\mu}_0, \hat{\sigma}_0)$, and the estimated parameters $\hat{\theta} = (\hat{\mu}_0, \hat{\sigma}_0)$ are also used to set the control limits, the computation of $K = K(\hat{P}, \hat{\theta})$ is similar to the case of computation of $K(P, \theta)$, where $K(P, \theta)$ is the parameter that produces the desired MRL_0 for the Shewhart $\bar{X}$ chart with *known* process parameters.
- Similar to the computation of quantity $K = K(\hat{P}, \hat{\theta})$, the quantities $K(\hat{P}_i^*, \hat{\theta}_i^*)$, $i = 1, 2, \ldots, B$ are all equal to $K = K(\hat{P}, \hat{\theta})$.

- For the quantities $K(\hat{P}, \hat{\theta}_i^*)$, the estimated parameters $\hat{\theta} = (\hat{\mu}_0, \hat{\sigma}_0)$ in the in-control distribution $\hat{P} = N(\hat{\mu}_0, \hat{\sigma}_0)$ are different from the ones $\hat{\theta}_i^* = (\hat{\mu}_i^*, \hat{\sigma}_i^*), i = 1, 2, \ldots, B$ used in the set of the Shewhart $\bar{X}$ charts. In this case, a binary search algorithm is used to compute the value K that satisfies the desired MRL_0.

Because the value of $K(\hat{P}, \hat{\theta})$ is equal to $K(\hat{P}_i^*, \hat{\theta}_i^*)$, the adjusted control limit can be simplified to get $100\alpha^*\%$ percentile of $K(\hat{P}, \hat{\theta}_i^*)$, $i = 1, 2, \cdots, B$. Once the (adjusted) control limits are determined, the conditional in- and out-of-control MRLs of the Shewhart $\bar{X}$ chart can be calculated using the method introduced in Section 2.2.

Figures 2 and 3 show the boxplots of the in- and out-of-control MRL distributions of the Shewhart $\bar{X}$ charts with the adjusted and unadjusted limits. If the limits are computed using the bootstrap adjustment approach, they are denoted as "Adjusted Limits". Otherwise, they are denoted as "Unadjusted Limits". The boxplots are constructed from 1000 MRL values.

(Please Insert Figures 2 to 3 Here)

From Figure 2, we may note that many in-control MRL values in the "Adjusted Limits" cases are larger than in the "Unadjusted Limits" cases. This is the result of the adjustment of the parameter K, which widen the distance of the control limits from the center line of the Shewhart $\bar{X}$ chart. Actually, as expected, the "Adjusted Limits" guarantee about 90% of the Shewhart $\bar{X}$ charts have in-control MRL values larger than the desired $MRL_0 = 257$. While for the "Unadjusted Limits" case, about 90% of the Shewhart $\bar{X}$ charts have in-control MRL values larger than 126, which means many of the "Unadjusted Limits" Shewhart $\bar{X}$ charts have a higher false alarm rate than the ones with the "Adjusted Limits".

In practice, control charts with only large in-control MRL performances are not enough for implementation, because we are also interested in the detection ability of control charts. Thus, the out-of-control MRL values of the Shewhart $\bar{X}$ charts with the "Adjusted Limits" and "Unadjusted Limits" are presented in Figure 3 for $\delta \in \{0.3, 0.5, 0.7, 1.0\}$, respectively. It can be noted that for smaller shift size, the out-of-control MRL values of the Shewhart $\bar{X}$ charts with "Adjusted Limits" are obviously larger than the ones of the Shewhart $\bar{X}$ charts with "Unadjusted Limits". For example, in Figure 3 (a), when $\delta = 0.3$, the median of the MRL values for the Shewhart $\bar{X}$ charts with "Adjusted Limits" and "Unadjusted Limits" are about 131 and 69, respectively. For moderate to large shift sizes, the Shewhart $\bar{X}$ charts with the "Adjusted Limits" only causes a small loss in the out-of-control MRL performance compared with using the "Unadjusted Limits". For example, in Figure 3 (c), when $\delta = 0.7$, the median of the out-of-control MRLs of the Shewhart $\bar{X}$ charts are about 14.5 and 9.0 for the "Adjusted Limits" and "Unadjusted Limits" cases, respectively. This small loss in the out-of-control MRL performances is generally acceptable compared with the gain in the in-control MRL performances. For these moderate to large shift sizes, the bootstrap approach can give a good balance between the in- and out-of-control MRL performance for the Shewhart $\bar{X}$ charts with estimated parameters. Moreover, it can also be seen from Figures 2 and 3 that the MRL distributions of the Shewhart $\bar{X}$ charts with the "Adjusted Limits" are loosen compared to the "Unadjusted Limits" cases.

When repeating Step (1) in the bootstrap approach, the estimated parameters $\hat{\mu}_0$ and $\hat{\sigma}_0$ would be different from each other, which means that the "Adjusted Limits" will not be the same each time. Based on many simulations, the results only show a small difference in these adjusted limits. In order to provide practitioners with more guidance on the adjusted limits of the MRL based Shewhart $\bar{X}$ chart, Table 7 presents the average adjusted limits $\bar{K}$, which guarantees about 90% of the in-control MRL values larger than the desired MRL_0. The $\bar{K}$ values are based on 1000 runs of the bootstrap approach for $m \leq 100$, and 100 runs for other m values larger than 100. The practitioners can refer to this table for the adjusted control limits, or they can either follow Steps (1)-(4) for obtaining the final adjusted limits for the MRL based Shewhart $\bar{X}$ chart.

(Please Insert Table 7 Here)

Since both in- and out-of-control MRL performances of control charts are concerned by practitioners and the bootstrap approach only guarantees the in-control MRL performance, we should not ignore the out-of-control MRL performance of the Shewhart $\bar{X}$ chart, If the detection ability of control chart is deemed to be insufficient, two alternative options are possibly adopted:

- Increase the amount of Phase I observations (increase m or n).
- Be more lenient on the desired in-control MRL_0.

For the first option, using more observations to estimate the process parameters $\hat{\theta} = (\hat{\mu}_0, \hat{\sigma}_0)$ would be more accurate and cause less uncertainty to account for in the unknown process parameters. So the adjustment of the control limit K will be smaller, which will result in smaller increase in the in- and out-of-control MRL values compared to the cases with a large adjustment of K. However, only a limited number of Phase I observations may be available in practice. Under this case, the second option has to be adopted. The practitioners could choose a less smaller MRL_0 than the predetermined desired one, or use a smaller value of α^*, which guarantees a smaller probability to have MRL values above the desired in-control MRL_0.

6 An illustrative example

In this section, we illustrate the proposed MRL based Shewhart $\bar{X}$ chart with unknown parameters by applying it to a production process of 500ml milk bottles. Similar to Castagliola et al. (2016), we are interested in the weight Y of each bottle. The in-control data are generated from a normal distribution with mean 500 and standard deviation 1. The Phase I data set consists of $m = 20$ samples, each of size $n = 5$ (see Table 8). From these Phase I samples, we can obtain the estimated process parameters $\hat{\mu}_0$=500.1349 and $\hat{\sigma}_0$=0.8241. Considering the between-practitioners variability, we can get the adjusted parameters $\bar{K} = 1.533$ of the Shewhart $\bar{X}$ chart from Table 7. Then, the control limits of the Phase II Shewhart $\bar{X}$ chart are set as follows:

$$UCL = 500.1349 + 1.533 \times 0.8241 = 501.34$$
$$LCL = 500.1349 - 1.533 \times 0.8241 = 498.87$$

The Phase II data set consisting of $m = 20$ samples, each of size $n = 5$, is shown Table 9. These samples are also presented in Figure 4 (a). The first 10 samples are supposed to be in-control, while the following 10 samples come from an out-of-control process. All the mean values $\bar{Y}_i$ of these Phase II samples are presented in the last column of Table 9, and the corresponding Shewhart $\bar{X}$ chart is shown in Figure 4 (b). From Figure 4 (b), we may note the mean value of samples #11, #15 and #20 (in bold in Table 9) are below the low control limit LCL, which indicates a possible assignable cause in the process.

(Please Insert Tables 8 and 9 Here)
(Please Insert Figure 4 Here)

7 Conclusions and recommendations

Since the shape of the run length distribution for the Shewhart $\bar{X}$ chart is highly skewed, a performance measure, such as the ARL, does not give a complete understanding of the chart's performance. Thus, the MRL is used to evaluate the properties of Shewhart $\bar{X}$ chart. Furthermore, the process parameters are often unknown, and need to be estimated from Phase I samples. Different practitioners will use different Phase I samples to estimate the process parameters, thus causing the between-practitioners variability in control charts' performance. Considering this variability, the MRL based Shewhart $\bar{X}$ chart with unknown process parameters is investigated in this article.

Several tables are presented for the practitioners to assess the between-practitioners variability of the MRL based Shewhart $\bar{X}$ chart, for different amount of Phase I samples. The results show that at least $m \simeq 400$ Phase I samples are needed to reduce the between-practitioners variability in the MRL based Shewhart $\bar{X}$ chart up to an acceptable level. Since taking a large amount of Phase I samples is impractical in practice, a bootstrap approach is adopted to adjust the control limits of the MRL based Shewhart $\bar{X}$ chart. The simulations results show that, for moderate to larger shift sizes, the adjusted control limits can give a good balance between the in- and out-of-control MRL performances. For different number of Phase I samples, sample sizes, and the desired MRL_0, the average adjusted control limits are given as a reference for the practitioners. As this work is conducted for the Shewhart $\bar{X}$ chart, future research works can be conducted for EWMA or CUSUM type charts.

Acknowledgements

This work is supported by National Natural Science Foundation of China No. 71802110, 72101123, 71872088, Natural Science Foundation of JiangSu Province No. BK20200750, Key Research Base of Philosophy and Social Sciences in Jiangsu-Information Industry Integration Innovation and Emergency Management Research Center, the Excellent Innovation Teams of Philosophy and Social Science in Jiangsu Province No. 2017ZSTD022, the Project of Philosophy and Social Science Research in Colleges and Universities in Jiangsu Province No. 2020SJA0090.

References

Aly, A., Mahmoud, M. and Hamed, R. (2016), 'The Performance of the Multivariate Adaptive Exponentially Weighted Moving Average Control Chart with Estimated Parameters', *Quality and Reliability Engineering International* **32**, 957–967.

Aly, A., Saleh, N., Mahmoud, M. and Woodall, W. (2015), 'A Re-evaluation of the Adaptive Exponentially Weighted Moving Average Control Chart When Parameters are Estimated', *Quality and Reliability Engineering International* **31**, 1611–1622.

Castagliola, P., Maravelakis, P. and Figueiredo, F. (2016), 'The EWMA Median Chart with Estimated Parameters', *IIE Transactions* **48**(1), 66–74.

Chakraborti, S. (2007), 'Run Length Distribution and Percentiles: The Shewhart $\bar{X}$ Chart with Unknown Parameters', *Quality Engineering* **19**(2), 119–127.

Faraz, A., Woodall, W. and Heuchenne, C. (2015), 'Guaranteed Conditional Performance of the S^2 Control Chart with Estimated Parameters', *International Journal of Production Research* **53**(14), 4405–4413.

Gan, F. (1993), 'An Optimal Design of EWMA Control Charts Based on Run Length', *Journal of Statistical Computation and Simulation* **45**(3), 169–184.

Gan, F. (1994), 'An Optimal Design of Cumulative Sum Control Chart Based on Median Run Length', *Communications in Statistics–Simulation and Computation* **23**(2), 485–503.

Gandy, A. and Kvaloy, J. (2013), 'Guaranteed Conditional Performance of Control Charts via Bootstrap Methods', *Scandinavian Journal of Statistics* **40**, 647–668.

Geodhart, R., da Silvab, M. M., Schoonhoven, M., Epprecht, E., Chakraborti, S., Does, R. M. and Veiga, A. (2017), 'Shewhart Control Charts for Dispersion Adjusted for Parameter Estimation', *IISE Transactions* **49**(8), 838–848.

Geodhart, R., Schoonhoven, M. and Does, R. (2017), 'Guaranteed In-Control Performance for the Shewhart X and $\bar{X}$ Control Charts', *Journal of Quality Technology* **49**(2), 155–171.

Hu, X. and Castagliola, P. (2017), 'Guaranteed Conditional Design of the Median Chart with Estimated Parameters', *Quality and Reliability Engineering International* **33**(8), 1873–1884.

Hu, X., Castagliola, P., Zhou, X. and Tang, A. (2018), 'Conditional Design of the EWMA Median Chart with Estimated Parameters', *Communications in Statistics-Theory and Methods* .

Jensen, W., Jones, L., Champ, C. and Woodall, W. (2006), 'Effects of Parameter Estimation on Control Chart Properties: A Literature Review', *Journal of Quality Technology* **38**(4), 349–364.

Jones, L., Champ, C. and Rigdon, S. (2004), 'The Run Length Distribution of the CUSUM with Estimated Parameters', *Journal of Quality Technology* **36**(1), 95–108.

Keefe, M., Woodall, W. and Jones-Farmer, L. (2015), 'The Conditional In-Control Performance of Self-Starting Control Charts', *Quality Engineering* **27**, 488–499.

Khoo, M. B. C. (2004), 'Performance Measures for the Shewhart $\bar{X}$ Control Chart', *Quality Engineering* **16**(4), 585–590.

Khoo, M. B. C., Wong, V. H., Wu, Z. and Castagliola, P. (2011), 'Optimal Designs of the Multivariate Synthetic Chart for Monitoring the Process Mean Vector Based on Median Run Length', *Quality and Reliability Engineering International* **27**, 981–997.

Khoo, M. B. C., Wong, V. H., Wu, Z. and Castagliola, P. (2012), 'Optimal Design of the Synthetic Chart for the Process Mean Based on Median Run Length', *IIE Transactions* **44**, 765–779.

Lee, M. and Khoo, M. B. (2017), 'Optimal Design of Synthetic np Control Chart Based on Median Run Length', *Communications in Statistics - Theory and Methods* **46**(17), 8544–8556.

Montgomery, D. C. (2009), *Introduction to Statistical Quality Control*, 6th edn, John Wiley & Sons, Inc.

Saleh, N., Mahmoud, M., Jones-Farmer, L., Zwetsloot, I. and Woodall, W. (2015), 'Another Look at the EWMA Control Charts with Estimated Parameters', *Journal of Qualtiy Technology* **47**(4), 363–382.

Saleh, N., Mahmoud, M., Keefe, M. and Woodall, W. (2015), 'The Diffculty in Designing Shewhart $\bar{X}$ and X Conrol Charts with Estimated Parameters', *Journal of Qualtiy Technology* **47**(2), 127–138.

Saleh, N., Zwetsloot, I., Mahmoud, M. and Woodall, W. (2016), 'CUSUM Charts with Controlled Conditional Performance under Estimated Parameters', *Quality Engineering* **28**(4), 402–415.

Tang, A., Castagliola, P., Sun, J. and Hu, X. (2018), 'Optimal Design of the Aadaptive EWMA Chart for the Mean Based on Median Run Length and Expected Median Run Length', *Quality Technology & Quantitative Management* .

Teoh, W., Chong, J., Khoo, M. B., Castagliola, P. and Yeong, W. (2016), 'Optimal Designs of the Variable Sample Size Chart Based on Median Run Length and Expected Median Run Length', *Quality and Reliability Engineering International* **33**(1), 121–134.

Teoh, W., Khoo, M. B., Castagliola, P. and Chakraborti, S. (2015), 'A Median Run Length-Based Double-Sampling $\bar{X}$ Chart with Estimated Parameters for Minimizing the Average Sample Size', *The International Journal of Advanced Manufacturing Technology* **80**(1-4), 411–426.

Teoh, W., Khoo, M. B., Castagliola, P. and Lee, M. H. (2016), 'The Exact Run Length Distribution and Design of the Shewhart Chart with Estimated Parameters Based on Median Run Length', *Communications in Statistics –Simulation and Computation* **45**(6), 2081–2103.

You, H., Khoo, M. B., Castagliola, P. and Qu, L. (2016), 'Optimal Exponentially Weighted Moving Average Charts with Estimated Parameters Based on Median Run Length and Expected Median Run Length', *International Journal of Production Research* **54**(17), 5073–5094.

Zhang, M., Megahed, F. and Woodall, W. (2014), 'Exponential CUSUM Charts with Estimated Control Limits', *Quality and Reliability Engineering International* **30**, 275–286.

Zhang, Y., Castagliola, P., Wu, Z. and Khoo, M. (2011), 'The Synthetic $\bar{X}$ Charts with Estimated Parameters', *IIE Transactions* **43**, 676–687.

Zhao, M. and Driscoll, A. (2016), 'The c-Chart with Bootstrap Adjusted Control Limits to Improve Conditional Performance', *Quality and Reliability Engineering International* **32**(8), 2871–2881.

Table 1 Percentiles and the ARLs of the Shewhart $\bar{X}$ Chart for different values of δ and K when $n = 5$

		Percentiles of the RL distribution										
δ	ARL	5th	10th	20th	30th	40th	50th	60th	70th	80th	90th	95th
						$K = 1.1519$						
0.0	100.00	6	11	23	36	51	69	92	120	161	230	299
0.2	55.88	3	6	13	20	29	39	51	67	90	128	166
0.5	13.78	1	2	3	5	7	10	13	16	22	31	40
0.8	4.64	1	1	1	2	3	3	4	5	7	10	13
1.0	2.72	1	1	1	1	2	2	3	3	4	6	7
1.5	1.28	1	1	1	1	1	1	1	1	2	2	2
2.0	1.03	1	1	1	1	1	1	1	1	1	1	1
						$K = 1.2553$						
0.0	200.00	11	22	45	72	102	139	183	241	322	460	598
0.2	102.98	6	11	23	37	53	72	94	124	165	236	308
0.5	21.90	2	3	5	8	11	15	20	26	35	50	65
0.8	6.48	1	1	2	3	4	5	6	8	10	14	18
1.0	3.52	1	1	1	2	2	3	3	4	5	7	9
1.5	1.41	1	1	1	1	1	1	1	1	2	2	3
2.0	1.05	1	1	1	1	1	1	1	1	1	1	1
						$K = 1.3416$						
0.0	370.40	19	39	83	132	189	257	339	446	596	852	1109
0.2	177.73	10	19	40	64	91	123	163	214	286	409	531
0.5	33.40	2	4	8	12	17	23	31	40	53	76	99
0.8	8.86	1	1	2	3	5	6	8	11	14	20	26
1.0	4.50	1	1	1	2	3	3	4	5	7	10	12
1.5	1.57	1	1	1	1	1	1	1	2	2	3	3
2.0	1.08	1	1	1	1	1	1	1	1	1	1	2
						$K = 1.3820$						
0.0	500.00	26	53	112	179	256	347	458	602	804	1151	1497
0.2	231.99	12	25	52	83	119	161	213	279	373	534	694
0.5	41.14	3	5	10	15	21	29	38	49	66	94	122
0.8	10.36	1	2	3	4	6	7	10	12	16	23	30
1.0	5.09	1	1	2	2	3	4	5	6	8	11	14
1.5	1.66	1	1	1	1	1	1	1	2	2	3	4
2.0	1.09	1	1	1	1	1	1	1	1	1	1	2

Table 2 The distribution of the in-control MRL of the Shewhart $\bar{X}$ chart for different values of m and the desired MRL_0 when $n = 5$

m	5th	10th	20th	30th	40th	50th	60th	70th	80th	90th	95th	$AMRL$	$SDMRL$
						$MRL_0 = 69$ ($K = 1.1519$)							
20	23	28	36	42	50	58	67	79	97	129	165	71.32	51.55
30	29	34	41	48	54	61	69	79	93	117	142	70.46	39.04
50	35	40	47	53	58	64	71	78	89	106	122	69.74	28.23
80	41	46	52	57	61	66	71	77	85	97	109	69.36	21.40
100	44	48	53	58	62	67	71	77	83	94	104	69.36	19.01
200	50	54	58	62	65	68	71	75	80	87	93	69.37	13.12
300	53	56	60	63	66	69	71	74	78	84	88	69.44	10.74
400	56	58	62	64	66	69	71	74	77	82	86	69.40	9.20
500	57	59	62	65	67	69	71	73	76	80	84	69.46	8.24
						$MRL_0 = 139$ ($K = 1.2553$)							
20	38	48	64	78	94	112	135	164	207	288	382	149.90	133.84
30	50	60	75	89	104	120	139	162	197	257	322	144.80	96.40
50	63	74	88	101	113	127	141	160	185	227	270	141.48	68.22
80	75	85	98	109	120	131	143	157	176	207	236	139.89	51.22
100	81	90	102	113	122	132	143	156	173	199	224	139.72	45.38
200	95	103	113	121	128	135	143	152	163	180	195	139.05	31.00
300	102	109	118	124	130	136	143	150	159	172	184	138.95	25.11
400	107	113	121	127	132	137	143	149	156	167	177	139.00	21.68
500	110	115	122	128	133	137	142	148	154	164	173	138.83	19.32
						$MRL_0 = 257$ ($K = 1.3416$)							
20	61	78	107	135	166	202	248	309	403	588	810	292.41	312.65
30	80	99	129	156	185	218	257	307	379	515	666	277.32	222.23
50	106	125	154	179	204	231	263	302	355	446	543	266.18	148.41
80	129	147	173	196	218	240	265	296	336	404	468	262.16	110.70
100	139	157	182	203	223	243	266	292	327	385	442	260.62	96.54
200	168	183	203	220	234	250	266	284	308	345	379	258.37	65.70
300	182	196	213	227	239	252	265	280	299	327	353	257.79	52.78
400	191	203	219	231	242	253	264	277	293	317	338	257.34	45.35
500	197	209	223	234	244	254	264	276	290	310	329	257.39	40.51
						$MRL_0 = 347$ ($K = 1.3820$)							
20	75	98	137	176	220	270	334	422	558	839	1173	407.27	473.54
30	101	126	167	205	246	291	345	417	523	723	950	379.67	320.52
50	136	162	202	237	272	310	355	411	489	623	771	363.94	217.64
80	167	194	230	262	292	324	360	403	461	559	658	356.86	159.90
100	181	206	240	270	298	327	359	398	449	531	615	353.15	139.07
200	221	242	271	294	315	337	360	387	421	474	523	349.78	94.07
300	242	260	284	304	322	340	359	380	408	448	485	348.52	75.45
400	253	270	293	310	326	342	358	376	399	434	466	348.21	65.36
500	262	278	299	315	329	343	357	373	394	424	451	347.83	58.01

Table 3 In-Control $AMRL$ and $SDMRL$ for different combinations of the desired MRL_0 and m when $n = 3$

	$MRL_0 = 69$ $(K = 1.4872)$		$MRL_0 = 139$ $(K = 1.6206)$		$MRL_0 = 257$ $(K = 1.7321)$		$MRL_0 = 347$ $(K = 1.7841)$	
m	$AMRL$	$SDMRL$	$AMRL$	$SDMRL$	$AMRL$	$SDMRL$	$AMRL$	$SDMRL$
20	86.40	111.20	196.53	351.51	419.67	1025.60	613.31	1759.23
30	79.28	70.03	171.33	194.41	345.72	489.98	489.30	774.40
50	74.71	45.49	155.87	116.53	302.55	268.56	419.14	403.75
70	72.97	35.92	150.25	89.33	287.49	199.58	395.05	294.80
100	71.84	28.65	146.52	69.82	277.25	152.65	378.74	222.87
150	70.98	22.54	143.76	54.23	269.93	116.71	367.24	168.94
200	70.62	19.21	142.45	45.81	266.48	97.91	361.81	141.17
250	70.35	17.01	141.66	40.39	264.51	85.97	358.60	123.64
300	70.18	15.41	141.17	36.53	263.13	77.49	356.59	111.29
400	70.00	13.26	140.52	31.25	261.55	66.12	354.01	94.73
500	69.90	11.76	140.15	27.78	260.59	58.59	352.47	83.91
600	69.86	10.72	139.96	25.22	259.97	53.16	351.56	76.07
700	69.75	9.91	139.79	23.26	259.51	49.00	350.85	70.08
1000	69.71	8.27	139.47	19.35	258.74	40.70	349.62	58.15
1100	69.71	7.90	139.41	18.47	258.55	38.72	349.33	55.36
1300	69.59	7.24	139.31	16.93	258.30	35.56	348.94	50.77
1500	69.63	**6.72**	139.24	15.76	258.11	33.03	348.59	47.18
2000	69.56	5.81	139.10	**13.60**	257.85	28.53	348.16	40.74
2500	69.57	5.16	139.07	12.15	257.65	**25.47**	347.79	36.36
2600	69.54	5.10	139.09	11.96	257.62	24.99	347.85	35.64
2700	69.50	5.02	139.00	11.65	257.55	24.49	347.81	34.95
2800	69.49	4.97	139.01	11.48	257.52	24.07	347.76	**34.34**
3000	69.47	4.84	139.02	11.09	257.53	23.22	347.66	33.17
5000	69.50	3.67	138.93	8.56	257.26	17.94	347.25	25.59
∞	69.00	0.00	139.00	0.00	257.00	0.00	347.00	0.00

Table 4 In-Control $AMRL$ and $SDMRL$ for different combinations of the desired MRL_0 and m when $n = 5$

	$MRL_0 = 69$ $(K = 1.1519)$		$MRL_0 = 139$ $(K = 1.2553)$		$MRL_0 = 257$ $(K = 1.3416)$		$MRL_0 = 347$ $(K = 1.3820)$	
m	$AMRL$	$SDMRL$	$AMRL$	$SDMRL$	$AMRL$	$SDMRL$	$AMRL$	$SDMRL$
20	71.52	51.65	149.90	134.83	292.86	318.13	407.45	484.77
30	70.25	38.62	144.53	96.50	276.50	217.01	380.33	321.87
50	69.62	28.09	141.38	68.15	266.55	148.31	363.43	216.10
70	69.44	23.18	140.27	55.55	262.99	119.36	357.47	172.77
100	69.41	19.03	139.67	45.26	260.73	96.44	353.56	138.97
150	69.36	15.29	139.30	36.25	259.24	76.76	350.96	110.15
200	69.41	13.21	139.10	31.06	258.50	65.60	349.74	94.03
250	69.43	11.71	139.02	27.64	258.16	58.22	349.08	83.36
300	69.42	10.69	138.97	25.14	257.93	52.91	348.61	75.69
400	69.46	9.20	138.94	21.69	257.61	45.52	348.10	65.10
500	69.46	8.25	138.90	19.30	257.46	40.59	347.82	57.99
600	69.42	7.49	138.88	17.60	257.36	36.94	347.64	52.78
700	69.45	**6.94**	138.80	16.25	257.23	34.14	347.46	48.78
1000	69.43	5.81	138.83	**13.61**	257.23	28.50	347.26	40.67
1100	69.44	5.56	138.79	12.96	257.14	27.16	347.25	38.71
1300	69.45	5.08	138.84	11.93	257.13	**24.94**	347.11	35.58
1500	69.39	4.81	138.79	11.08	257.07	23.20	347.05	**33.11**
2000	69.46	4.04	138.79	9.63	257.05	20.09	346.98	28.62
2500	69.45	3.68	138.81	8.56	257.02	17.93	346.91	25.58
3000	69.46	3.38	138.79	7.80	256.99	16.38	346.90	23.35
5000	69.46	2.62	138.78	6.07	256.91	12.63	346.86	18.07
∞	69.00	0.00	139.00	0.00	257.00	0.00	347.00	0.00

Table 5 In-Control $AMRL$ and $SDMRL$ for different combinations of the desired MRL_0 and m when $n = 7$

	$MRL_0 = 69$ $(K = 0.9736)$		$MRL_0 = 139$ $(K = 1.0610)$		$MRL_0 = 257$ $(K = 1.1339)$		$MRL_0 = 347$ $(K = 1.1680)$	
m	$AMRL$	$SDMRL$	$AMRL$	$SDMRL$	$AMRL$	$SDMRL$	$AMRL$	$SDMRL$
20	67.49	37.85	138.09	93.97	263.08	210.21	361.07	311.08
30	67.62	29.55	136.99	71.64	258.07	156.06	351.78	227.57
50	68.06	22.15	137.01	52.83	255.88	112.99	347.17	163.14
70	68.35	18.45	137.27	43.79	255.51	93.00	346.06	133.72
100	68.62	15.31	137.48	36.13	255.57	76.31	345.73	109.44
150	68.85	12.39	137.89	29.17	255.75	61.41	345.69	87.88
200	68.99	10.68	138.04	25.08	255.95	52.78	345.83	75.48
250	69.11	9.52	138.16	22.36	256.10	46.97	345.98	67.16
300	69.16	8.69	138.32	20.39	256.21	42.78	346.04	61.13
400	69.21	7.49	138.37	17.61	256.37	36.91	346.19	52.72
500	69.26	**6.71**	138.47	15.73	256.47	32.93	346.28	47.01
600	69.31	6.14	138.52	14.33	256.52	30.03	346.37	42.87
700	69.33	5.67	138.52	**13.26**	256.55	27.76	346.37	39.64
800	69.36	5.25	138.54	12.38	256.59	25.97	346.43	37.04
900	69.34	5.00	138.59	11.66	256.64	**24.45**	346.49	34.91
1000	69.31	4.80	138.59	11.07	256.66	23.21	346.53	**33.09**
1100	69.40	4.54	138.64	10.57	256.63	22.14	346.51	31.52
1300	69.45	4.16	138.65	9.74	256.74	20.31	346.49	29.01
1500	69.37	3.83	138.70	9.01	256.72	18.92	346.58	26.99
2000	69.43	3.38	138.68	7.81	256.78	16.38	346.57	23.34
2500	69.41	3.01	138.74	6.99	256.81	14.62	346.61	20.87
3000	69.42	2.72	138.69	6.39	256.82	13.34	346.67	19.03
5000	69.35	2.04	138.74	4.99	256.84	10.32	346.65	14.75
∞	69.00	0.00	139.00	0.00	257.00	0.00	347.00	0.00

Table 6 In-Control $AMRL$ and $SDMRL$ for different combinations of the desired MRL_0 and m when $n = 9$

	$MRL_0 = 69$ $(K = 0.8586)$		$MRL_0 = 139$ $(K = 0.9357)$		$MRL_0 = 257$ $(K = 1)$		$MRL_0 = 347$ $(K = 1.0301)$	
m	$AMRL$	$SDMRL$	$AMRL$	$SDMRL$	$AMRL$	$SDMRL$	$AMRL$	$SDMRL$
20	65.61	31.39	132.69	76.13	249.88	166.30	340.75	242.86
30	66.38	24.98	133.54	59.70	249.54	128.08	338.72	185.23
50	67.35	18.95	134.89	44.85	250.78	95.14	339.49	136.69
70	67.82	15.91	135.66	37.44	251.88	79.07	340.64	113.30
100	68.21	13.21	136.46	31.01	253.03	65.28	341.88	93.43
150	68.65	10.71	137.13	25.11	254.11	52.75	343.17	75.44
200	68.83	9.23	137.51	21.67	254.73	45.48	343.95	64.96
250	68.97	8.28	137.73	19.32	255.11	40.54	344.47	57.87
300	69.04	7.52	137.92	17.61	255.38	36.92	344.76	52.72
400	69.12	**6.49**	138.10	15.20	255.75	31.92	345.25	45.48
500	69.19	5.81	138.21	**13.58**	255.95	28.48	345.56	40.63
600	69.25	5.26	138.34	12.40	256.15	**25.97**	345.72	37.07
700	69.25	4.94	138.40	11.47	256.21	24.03	345.83	**34.29**
1000	69.34	4.05	138.48	9.61	256.44	20.12	346.06	28.61
1100	69.32	3.86	138.56	9.12	256.45	19.14	346.12	27.28
1300	69.36	3.61	138.59	8.41	256.49	17.59	346.19	25.09
1500	69.37	3.36	138.60	7.82	256.59	16.40	346.31	23.35
2000	69.41	2.92	138.62	6.77	256.64	14.19	346.40	20.20
2500	69.42	2.61	138.68	6.08	256.68	12.64	346.51	18.05
3000	69.39	2.39	138.68	5.52	256.72	11.57	346.55	16.52
5000	69.42	1.85	138.87	4.32	256.84	8.92	346.68	12.79
∞	69.00	0.00	139.00	0.00	257.00	0.00	347.00	0.00

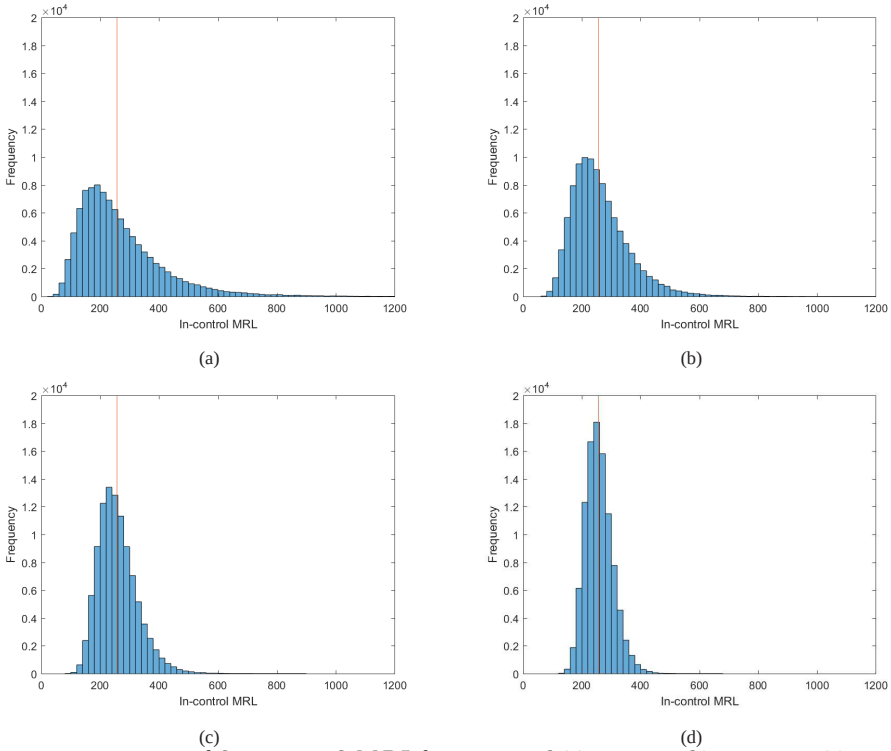

Figure 1 Histograms of the in-control MRL for $n = 5$ and (a) $m = 50$, (b) $m = 100$, (c) $m = 200$, (d) $m = 400$. (The red line corresponds to the $MRL_0 = 257$ on the horizontal axis)

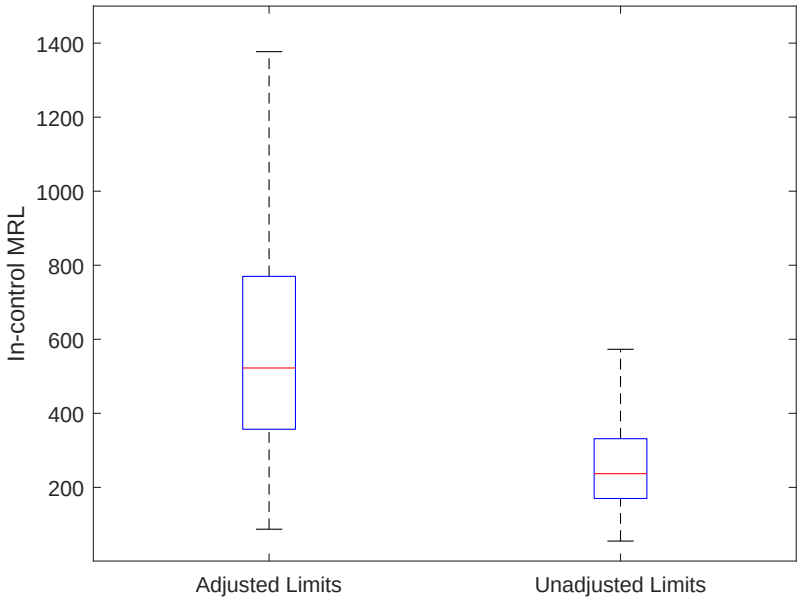

Figure 2 The boxplots of the in-control MRL for the Shewhart $\bar{X}$ chart with the adjusted and unadjusted limits.

Table 7 The average adjusted control limits $\bar{K}$ for the Shewhart $\bar{X}$ chart

	$n=3$				$n=5$			
m	$MRL_0=69$	$MRL_0=139$	$MRL_0=257$	$MRL_0=347$	$MRL_0=69$	$MRL_0=139$	$MRL_0=257$	$MRL_0=347$
20	1.793	1.954	2.074	2.150	1.321	1.434	1.533	1.581
30	1.721	1.871	2.003	2.063	1.279	1.395	1.490	1.539
50	1.651	1.801	1.924	1.988	1.243	1.358	1.449	1.496
80	1.614	1.761	1.882	1.938	1.221	1.333	1.421	1.468
100	1.593	1.742	1.863	1.915	1.213	1.323	1.415	1.455
130	1.580	1.724	1.846	1.898	1.202	1.312	1.403	1.446
150	1.573	1.717	1.835	1.889	1.199	1.310	1.399	1.442
200	1.559	1.702	1.819	1.876	1.191	1.300	1.390	1.433
250	1.553	1.694	1.810	1.863	1.187	1.295	1.384	1.427
300	1.546	1.686	1.803	1.858	1.184	1.292	1.380	1.423
350	1.541	1.680	1.796	1.852	1.180	1.289	1.377	1.419
400	1.537	1.679	1.791	1.845	1.178	1.287	1.375	1.415
	$n=7$				$n=9$			
m	$MRL_0=69$	$MRL_0=139$	$MRL_0=257$	$MRL_0=347$	$MRL_0=69$	$MRL_0=139$	$MRL_0=257$	$MRL_0=347$
20	1.090	1.190	1.272	1.310	0.954	1.042	1.107	1.144
30	1.062	1.161	1.240	1.279	0.929	1.014	1.086	1.114
50	1.038	1.133	1.210	1.249	0.909	0.992	1.061	1.093
80	1.022	1.115	1.191	1.228	0.896	0.977	1.045	1.077
100	1.016	1.108	1.186	1.221	0.890	0.972	1.039	1.071
130	1.009	1.102	1.178	1.212	0.886	0.967	1.033	1.065
150	1.006	1.098	1.174	1.209	0.883	0.965	1.031	1.061
200	1.001	1.094	1.168	1.203	0.880	0.960	1.027	1.057
250	0.997	1.088	1.164	1.198	0.877	0.958	1.022	1.054
300	0.995	1.086	1.161	1.196	0.875	0.955	1.020	1.051
350	0.993	1.083	1.159	1.193	0.873	0.953	1.019	1.050
400	0.991	1.082	1.157	1.192	0.872	0.952	1.018	1.049

Table 8 Phase I Data used to estimate the process parameters

	Phase I data($X_{i,j}$)					X_i	S_i^2
1	500.60	500.44	500.07	499.01	498.81	499.79	0.68
2	500.76	501.76	501.34	500.47	501.88	501.24	0.38
3	499.07	499.95	500.15	500.14	498.95	499.65	0.35
4	499.39	499.55	500.06	500.41	500.79	500.04	0.34
5	500.55	500.88	501.55	502.17	500.96	501.22	0.41
6	499.96	500.97	499.31	501.18	501.59	500.60	0.88
7	500.45	498.82	499.73	501.19	499.78	499.99	0.79
8	500.82	500.37	499.71	499.44	500.87	500.24	0.42
9	500.28	499.30	497.19	497.66	499.30	498.75	1.64
10	499.83	498.41	501.54	501.10	499.86	500.15	1.51
11	500.08	499.13	500.45	499.18	498.99	499.57	0.43
12	500.88	499.89	501.18	500.83	500.72	500.70	0.24
13	499.98	499.03	499.79	499.25	500.52	499.71	0.35
14	500.03	499.97	500.21	499.46	499.26	499.79	0.16
15	499.59	499.20	501.28	500.52	500.30	500.18	0.67
16	500.45	501.96	498.84	501.19	499.43	500.38	1.61
17	499.13	500.82	500.88	499.50	498.91	499.85	0.88
18	500.42	499.83	498.59	500.52	500.54	499.98	0.69
19	500.77	499.86	500.33	500.56	500.64	500.43	0.13
20	499.14	499.75	500.57	501.70	501.05	500.44	1.03

Table 9 Phase II data to be monitored using the Shewhart $\bar{X}$ chart

	Phase II data($Y_{i,j}$)					Y_i
1	498.98	500.12	498.98	499.34	499.68	499.42
2	500.90	499.97	500.74	502.11	499.68	500.68
3	499.97	499.70	499.06	499.61	498.49	499.37
4	500.47	500.34	500.70	499.69	499.27	500.10
5	500.29	500.52	500.32	498.00	499.13	499.65
6	499.21	501.02	500.21	499.51	498.33	499.65
7	500.76	500.29	498.16	499.53	499.75	499.70
8	500.60	498.74	500.05	498.92	500.41	499.75
9	498.46	501.95	500.05	501.28	502.52	500.85
10	497.87	498.63	500.98	500.81	500.82	499.82
11	498.14	498.25	498.02	499.45	499.54	**498.68**
12	500.09	499.68	499.05	499.76	499.11	499.54
13	499.41	500.85	498.63	497.79	499.31	499.20
14	499.03	499.30	499.93	499.17	499.12	499.31
15	498.28	498.71	498.86	497.56	500.58	**498.80**
16	500.20	499.37	500.69	499.45	498.14	499.57
17	497.36	501.54	498.96	500.57	498.77	499.44
18	497.84	499.41	499.84	500.78	497.71	499.12
19	499.90	497.71	498.97	500.56	500.59	499.55
20	497.92	500.21	499.23	498.29	498.39	**498.81**

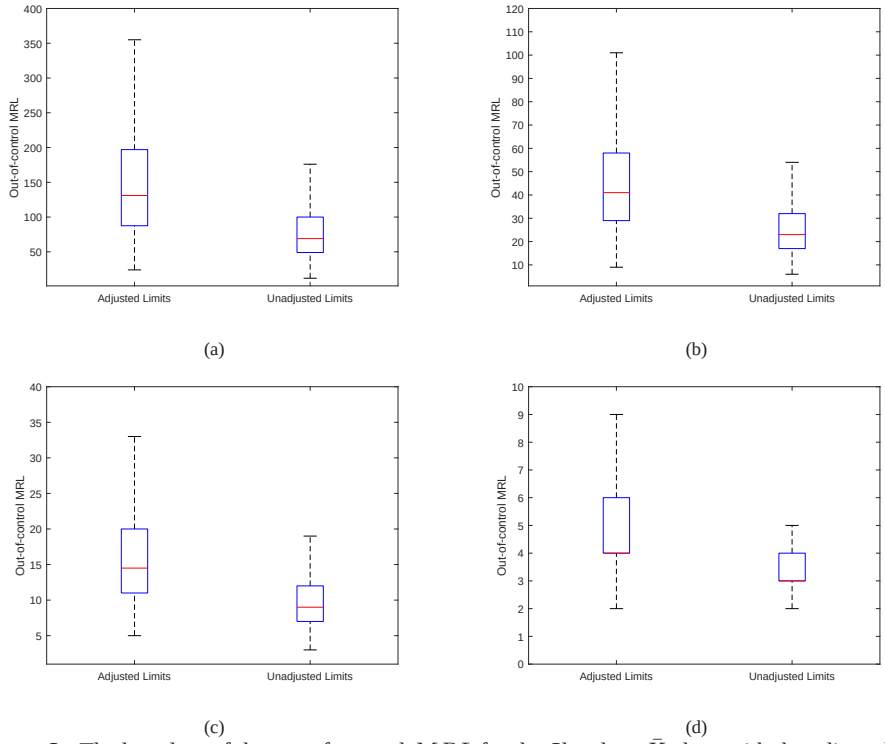

Figure 3 The boxplots of the out-of-control MRL for the Shewhart $\bar{X}$ chart with the adjusted and unadjusted limits, for (a) $\delta = 0.3$, (b) $\delta = 0.5$, (c) $\delta = 0.7$, (d) $\delta = 1.0$

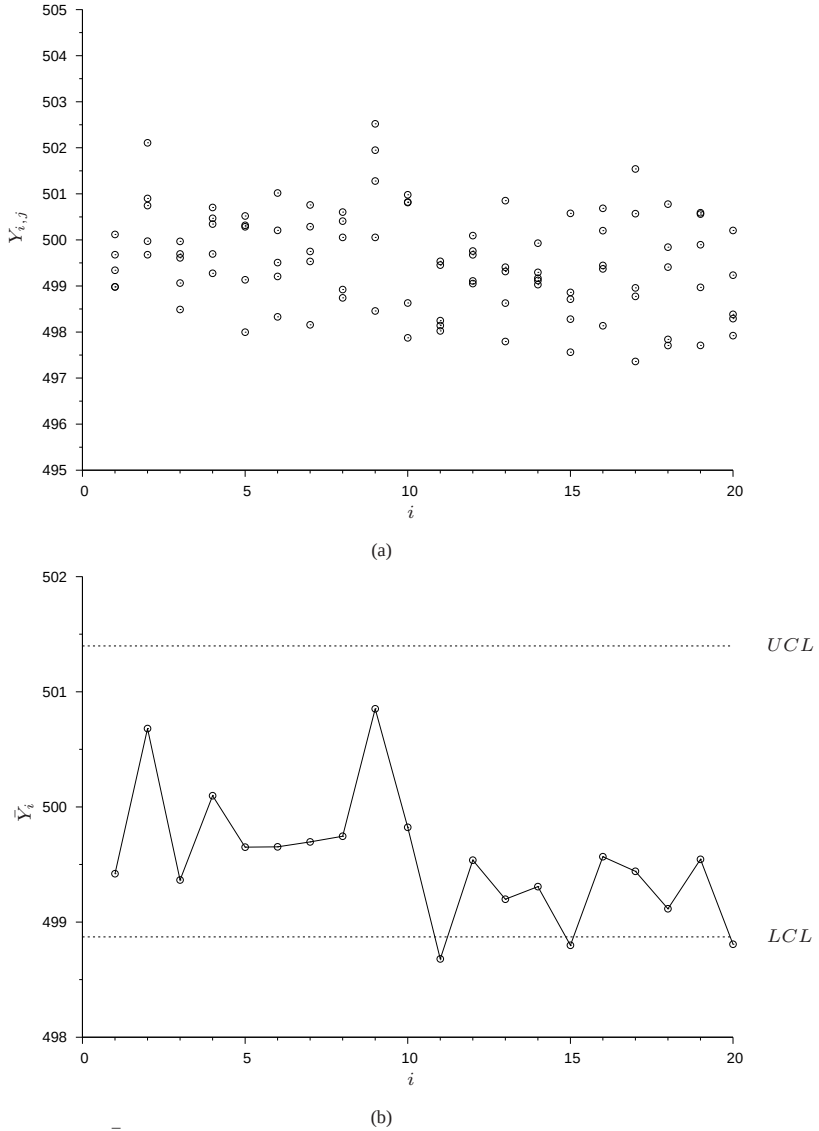

Figure 4 Shewhart $\bar{X}$ chart corresponding to the Phase II data in Table 9