Robustness of the EWMA control chart for individual observations

S.W. Human
Department of Statistics
University of Pretoria
Lynnwood Road,
Pretoria, 0002
South Africa
schalk.human@up.ac.za

P. Kritzinger
Department of Statistics
University of Pretoria
Lynnwood Road,
Pretoria, 0002
South Africa
pierre.kritzinger@gmail.com

S. Chakraborti
Department of Information Systems, Statistics and Management Science
University of Alabama
Tuscaloosa, AL 35487
U.S.A.
schakrab@cba.ua.edu

ABSTRACT

The traditional exponentially weighted moving average (EWMA) chart is one of the most popular control charts used in practice today. The in-control robustness is the key to the proper design and implementation of any control chart, lack of which can render its out-of-control shift detection capability almost meaningless. To this end, Borror, Montgomery and Runger [5; hereafter BMR] studied the performance of the traditional EWMA chart for the mean for i.i.d. data. We use a more extensive simulation study to further investigate the in-control robustness (to non-normality) of the three different EWMA designs studied by BMR [5]. Our study includes a much wider collection of non-normal distributions including light- and heavy-tailed, symmetric and asymmetric bi-modal as well as the contaminated normal, which is particularly useful to study the effects of outliers. Also, we consider two separate cases: (i) when the process mean and standard deviation are both known and (ii) when they are both unknown and estimated from an in-control Phase I sample. In addition, unlike in BMR [5], the average run-length (ARL) is not used as the sole performance measure in our study, we consider, the standard deviation (SDRL), the median (MDRL), the first and the third quartiles as well as the first and the ninety-ninth percentiles of the in-control run-length distribution for a better overall assessment of the traditional EWMA chart's in-control performance. Our findings sound a cautionary note to the (over) use of the EWMA chart in practice, at least with some types of non-normal data. A summary and recommendations are provided.

Keywords: Average run-length, Boxplot, Distribution-free, Median run-length, Nonparametric, Percentile, Run-length, Simulation.

1. Introduction

The traditional exponentially weighted moving average (EWMA) chart introduced by Roberts [12] is one of the most popular control charts used in practice today in a diverse number of settings and applications. The EWMA chart is simple to implement and has been recognized to be more effective than the Shewhart chart for detecting small shifts. Now, the in-control robustness is the key to the proper design and implementation of any control chart, lack of which can render its out-of-control shift detection capability almost meaningless. This is why we focus on the in-control robustness of the traditional EWMA chart in this paper. BMR [5] showed that the in-control average run-length (ARL) of the Shewhart individuals chart is very sensitive to the normality assumption and the chart's in-control performance is significantly deteriorated otherwise. Consequently, based on their study on a number of gamma and *t*-distributions, they recommended using a properly designed EWMA chart when there is a concern about the assumption of normality. The EWMA chart contains the Shewhart chart as a special case and it has been shown to pick-up smaller shifts better (the EWMA statistic has some optimality properties as a forecast tool for an IMA(1,1) process), and as a result the EWMA chart has been applied in a variety of situations. More recently, Stoumbos *et al.* [16] and Testik *et al.* [17] studied the robustness of the EWMA chart in the multivariate case, Maravelakis *et al.* [9] studied the robustness of EWMA charts for monitoring the variance; and Apley *et al.* [3] studied EWMA charts for auto-correlated data with model uncertainty. The EWMA chart appears to enjoy widespread popularity in practice, even in applications where not much is known about the underlying process distribution, as its robustness (to non-normality) is either assumed or taken for granted based on the findings of BMR [5]. Montgomery [10, p. 413], for example, states that "It (the EWMA) is almost a perfectly nonparametric (distribution-free) procedure."

Against this backdrop, we further investigate the robustness of the EWMA chart in a more thorough study in order to get more insight into its in-control performance for non-normal univariate individuals data. The in-control performance is the key to the correct implementation of a control chart. Unless the in-control performance of a chart is stable and robust, is clearly understood and known, its use and effectiveness can be limited and in fact can be misleading in the realm of detecting shifts. A robust statistical procedure, as described by Balakrishnan *et al*. [4, p. 7299] is a procedure that performs well not only under ideal conditions (under which it is designed and proposed) but also under departures from the ideal. In the same spirit, a control chart is robust if its in-control run-length distribution remains stable (unchanged, or nearly unchanged) when the underlying distributional assumption(s) (normality, for example) are violated (see e.g.,[13]).

BMR [5] used a Markov chain approach (see e.g. [6]) and the steady-state control limits to investigate the in-control ARL of the traditional EWMA chart for the mean assuming that the population mean and variance are both known (so-called Case K) in case of some gamma and *t*-distributions for selected values of the chart's design parameters, λ and L. We investigate the robustness of the traditional EWMA chart in a much larger study involving many more factors than in BMR. First, we use a much larger set of distributions (sixteen; discussed later) that includes the ones used in BMR. Second, we study the performance of the EWMA chart using both steady-state control limits and the exact control limits using intensive computer simulations. The exact time-varying control limits that gradually widen are of interest because they improve the sensitivity of the EWMA chart to mean shifts that occur early. It may be noted that our simulation results match very well with the results in BMR [5] for the subset of distributions considered by them. Third, unlike in BMR where the ARL is used as the sole performance measure, we examine the in-control run-length distribution and because the run-length distribution is skewed (to the right) to make an overall assessment, we also look at other measures such as the standard deviation of the run-length (SDRL) as well as some percentiles, such as the median (denoted MDRL), the two quartiles as well as some (tail) extreme percentiles, which provide valuable information about the performance of a control chart. This has been recommended by several authors (see, for example,[7], [11] and [18]) in the literature. We calculate and examine the ARL, the SDRL, the 1st , 25th, 50th (MDRL), 75th and the 99th percentiles of the in-control run-length distribution using both the steady-state and the exact control limits. Finally, we examine the robustness of the traditional EWMA chart in situations when the mean and the variance of the underlying normal distribution are known (Case K) and when they are unknown (Case U).

The rest of the paper is organised as follows: In the next section we give a detailed description of all the distributions used in our simulation study. In the following section we provide a very basic background on the EWMA control chart. Then, we present our results and finally we close with a summary and some recommendations in the last section.

2. Distributions investigated

The probability density functions (p.d.f.'s) we use as well as expressions for their means (μ_0) and variances (σ_0^2) are shown in Table 1. The graphs of a number of the p.d.f.'s are displayed in Figure 1 to show the variety of distributional shapes considered. The following facts may be noted about the distributions used in our study:

i. The uniform distribution is symmetric with lighter tails than the normal distribution.

ii. The right triangular distribution is asymmetric, positively skewed and bounded below and above (unlike the gamma distribution which is unbounded above).

iii. Student's $t(v)$ distribution is normal-like (i.e. bell-shaped and symmetric) with heavier tails than the normal distribution; this distribution was selected since it can be very similar in appearance to the normal distribution

iv. The gamma distribution is positively skewed and is bounded below by zero; this distribution was chosen to study the effect of skewness and may occur when monitoring quality characteristics such as time until failure or breaking strength.

v. The bi-modal distributions; frequently encountered in industrial practice (see e.g. Schilling and Nelson, 1976) and may occur when several machines pool their output into a common stream. These distributions are constructed by mixing two normal distributions in different proportions. We considered the following:

a) The symmetric bi-modal distribution, formed by mixing equal proportions (50% to 50%) of a $N(0,1)$ and a $N(4,1)$ distribution; this produces a distribution with a mean of 2 and a standard deviation of 2.236.

b) The asymmetric bi-modal distribution, formed by mixing a $N(0,1)$ distribution with a $N(4, \sigma = \frac{1}{3})$ distribution in a 19:1 ratio (95% to 5%); the resulting distribution has a mean of 0.2 and a standard deviation of 1.7156 .

c) The contaminated normal (CN) distribution. This is a mixture distribution widely used in robustness/outlier studies (see e.g. [13]) and can arise in situations where we observe a process that follows a standard normal distribution most of the time but occasionally follows a normal distribution with a larger variance. The CN distribution is widely used in the SPC literature, see for example, the recent papers of Alfaro *et al.* [1] and Sheu *et al.* [14]. In simple terms, the CN model implies that most of the data are "good" but there are occasional "outliers." Our interest is to see how sensitive the EWMA chart is to occasional outliers which may or may not occur due to assignable causes.

We considered two contaminated normal distributions:

- CN1: a 95%-5% mixture of a $N(0,1)$ and a $N(0,\sigma = 5)$ distribution, and
- CN2: a 95%-5% mixture of a $N(0,1)$ and a $N(0,\sigma = 10)$ distribution.

Thus, the CN1 distribution has mean zero and a standard deviation of 1.483 whereas the CN2 distribution has mean zero but a variance of 2.439.

For space limitations we only show the p.d.f.'s of the standard normal (solid line) and the $t(3)$-distributions (dashed line) in Panel (b) of Figure 1; the $t(4)$, $t(5)$ and $t(6)$ distributions are not shown but are all symmetric and heavier-tailed than the $N(0,1)$ distribution. Note that the t distributions become increasingly lighter tailed as the degrees of freedom increases.

Panel (f) in Figure 1 shows the differences in tail behaviours among the standard normal (dotted line), the $t(3)$ (dot-dash line) and the CN2 (solid line) distributions. It is seen that the CN2 has significantly heavier tails than the standard normal whereas the $t(3)$ is in-between the standard normal and the CN2 in terms of tail heaviness. Note that the heavier the tail of the distribution, the higher the chance of observing outliers.

< Table 1 here >

< Figure 1 here >

3. Background on the EWMA control chart

Assume that $X_1, X_2, X_3, \ldots$ denote independent and identically distributed observations with an in-control mean μ_0 and a standard deviation σ_0. First assume that both of these parameters are known; this scenario is referred to as the "standards known" case and is denoted Case K. The exponentially weighted moving average (EWMA) control charting statistic for individual observations in Case K is defined as

$$Z_i = \lambda X_i + (1-\lambda) Z_{i-1} \qquad \text{for} \qquad i = 1,2,\ldots \text{ and } \quad Z_0 = \mu_0 \qquad (1)$$

where X_i is the current observation and the constant $0 < \lambda \leq 1$ is the smoothing parameter.

The exact control limits for the EWMA chart in Case K are

$$UCL / LCL = \mu_0 \pm L\sigma_0 \sqrt{\frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2i}]} \qquad \text{for} \qquad i = 1,2,\ldots \qquad (2)$$

where $L > 0$ determines the width of the control limits. For larger values of i, the exact control limits approach steady-state values. The steady-state EWMA control limits in Case K are

$$UCL / LCL = \mu_0 \pm L\sigma_0 \sqrt{\frac{\lambda}{2-\lambda}} . \qquad (3)$$

When the in-control mean and standard deviation are unknown, they are typically estimated from an in-control Phase I sample (so-called reference or calibration sample) before prospective (i.e. online) monitoring starts in Phase II; this scenario is referred to as the "standards unknown" case and denoted Case U. Note that, the Phase I sample is taken when the process was thought to operate in-control and without any special causes of concern (see e.g. [10], page 199). The point estimates of the unknown in-control mean and standard deviation (denoted by $\hat{\mu}$ and $\hat{\sigma}$, respectively) are used to obtain the starting value Z_0 and substituted in equations (2) and (3) for the respective known parameter values to estimate the Phase II control limits of the traditional EWMA control chart for individual observations. The EWMA control chart statistic for individual observations in Case U is defined as

$$Z_i = \lambda X_i + (1-\lambda) Z_{i-1} \qquad \text{for} \qquad i = 1,2,\ldots \text{ and } \quad Z_0 = \hat{\mu} . \qquad (4)$$

The estimated time-varying and steady-state control limits for Phase II applications in Case U are given by

$$U\hat{C}L\,/\,L\hat{C}L = \hat{\mu} \pm L\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2i}]} \qquad \text{for} \qquad i = 1,2,... \tag{5}$$

and

$$U\hat{C}L\,/\,L\hat{C}L = \hat{\mu} \pm L\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}}\,, \tag{6}$$

respectively.

The EWMA statistic Z_i (defined in (1) and (4) for Case K and Case U, respectively) is plotted on a control chart with the UCL and the LCL given in (2) or (3) (or with the $U\hat{C}L$ and the $L\hat{C}L$ given in (5) or (6)). If a point plots on or outside the control limits, the chart signals and the process is declared out-of-control (OOC) so that a search for assignable causes is started. In this paper we study the performance of the EWMA chart in Case K and in Case U under both sets of limits.

Typically, the EWMA chart is designed for a specified in-control ARL value and a magnitude of the anticipated shift by finding the combination of L and λ that provides the desired in-control ARL and the shortest out-of-control ARL to detect that shift (see e.g. Montgomery [10 p. 411]). However, the determination of the charting parameters (λ, L), although an important consideration, is not the focus of this paper. Our objective is to investigate the in-control robustness of the EWMA chart in Case K and Case U, respectively, for the three pairs of charting parameters studied in BMR [5] (shown in Table 2), under the exact time-varying control limits and the steady-state limits under the sixteen distributions listed in Section 2.

< Table 2 here >

4. Results: Tables, Graphs and Discussion

The results concerning the robustness of the EWMA chart in Case K and Case U are studied and discussed separately in the next two sections. First we focus on Case K.

Standards known: Case K

In Case K the robustness of the EWMA chart are affected and depends on three factors:

i. The design parameters (λ, L). Note that, these pairs of design parameters only guarantees an in-control ARL of 370 if we use steady-state control limits.

ii. The type of control limits i.e. whether we use the exact time-varying limits or the steady-state limits, and

iii. The underlying distribution of the observations.

The run-length distribution of the EWMA chart was obtained via Monte Carlo simulations using 200,000 repetitions. The simulations were conducted according the following steps.

i. One of the 16 distributions (listed Table 1) and a pair of design parameters i.e. (λ, L), listed in Table 2, are chosen.

ii. The mean (μ_0) and the standard deviation (σ_0) of the chosen distribution are calculated according to the formulae provided in Table 1.

iii. The control limits are calculated according to equations (2) or (3) (exact or steady-state limits, respectively) using the μ_0 and σ_0 values from step (ii) and the (λ, L) -values chosen in step (i).

iv. An individual observation, X_i, is generated from the selected distribution and the plotting statistic, Z_i, is calculated according to (1) with $Z_0 = \mu_0$.

v. If $Z_i < UCL$ and $Z_i > LCL$, then a run-length counter is incremented.

vi. Steps (iii) - (v) are repeated until $Z_i \geq UCL$ or $Z_i \leq LCL$; when this occurs, a signal is given and the run-length i.e. the time i, is recorded.

vii. Steps (i) - (vi) are repeated until 200,000 iterations are completed.

The results of our simulation study are summarized in Table 3 and in Figures 2, 3 and 4, respectively. The columns of Table 3 display the numerical values of the in-control run-length distribution characteristics examined: the 1st percentile, the 25th percentile, the MDRL, the 75th percentile, the 99th percentile, the ARL and the SDRL, using both the steady-state and the exact control limits, for each of the 16 distributions used in the study; these distributions are listed in column 2 of Table 3. A couple of quick observations can be made from Table 3.

i. In case of the symmetric bi-modal distribution, the tabulated values are quite (strikingly) different from those for the rest of the distributions; in fact, in all cases these values are much higher compared to the rest. For example, when the steady-state limits and $(\lambda, L) = (0.05, 2.492)$ are used, the in-control ARL is 5,042.5, which is significantly (about 13 times) higher than the next largest in-control ARL value of 392.5 (for the asymmetric bi-modal distribution); this is also true when the exact control limits and/or the other two (λ, L) combinations are used. Although larger in-control ARL values are desirable in general, one as large as 5,042.5 is not useful in practice.

ii. For the contaminated normal distribution, the in-control ARL values are significantly smaller than the rest, which is clearly problematic since they indicate more false alarms. To emphasize this last point, when occasional outliers are present, it is not advisable to use these EWMA designs (i.e. charts) since there can be a larger number of false alarms, much more than what is expected under normality. These two observations demonstrate that the EWMA chart in Case K is not nonparametric, that is, its in-control run-length distribution is not the same or nearly the same for all continuous distributions. The rest of Table 3 is discussed further.

< Table 3 here>

For an alternative and perhaps a more appealing way to explain these results, we show Figures 2, 3 and 4. Each figure shows boxplot-like graphs (see[11]) of the 1st, 25th, 50th, 75th and the 99th percentiles of the in-control run-length distributions of an EWMA chart under each of 11 (out of the total of 16) distributions and a given (λ, L) combination; the symmetric bi-modal distribution is not shown because as we explained earlier, the corresponding results were “way out of line”; the *t*(6), *t*(5), Gamma(4,1), and the Gamma(3,1) are also excluded as the corresponding results are similar to the *t* and the gamma distributions that are shown. Note that, in each “boxplot,” the “whiskers” are extended to the 1st and the 99th percentiles instead of the typical minimum and the maximum. In addition, each boxplot also shows the mean (as a square) and the median (as a circle) inside the box. Also, note that each of the figures displays the results for one particular (λ, L) combination when using the steady-state (in panel (a)) and the exact control limits (in panel (b)), respectively. These figures give a clearer indication of the location, spread and the shape of the in-control run-length distributions.

< Figures 2, 3 and 4 here >

To illustrate, for example, Figure 2 shows that both for the steady-state and exact control limits, the in-control run-length distribution of the EWMA chart with charting constants $(\lambda, L) = (0.05, 2.492)$ is stable around 370 with nearly the same inter-quartile range (IQR) values across all of the underlying distributions except for the contaminated normal distributions. The shapes of the distributions also look very similar. Thus, this pair of charting parameters appears to produce an EWMA chart which is robust for all distributions studied here except for the contaminated normal distributions. In this latter important case, however, the EWMA chart is seen to have a shorter ARL and a shorter MDRL (280.9 and 196, respectively, from Table 3, for the CN1 distribution) and shorter IQR value (302 for CN1 from Table 3), which suggest a much higher number of false alarms. The situation gets worse with increasing variance (CN1 to CN2) with both the ARL and MDRL decreasing along with the IQR.

While Figure 2 makes an encouraging case for the EWMA chart for the combination $(\lambda, L) = (0.05, 2.492)$, Figure 3 paints a remarkably different picture for the pair $(\lambda, L) = (0.1, 2.703)$ and the corresponding EWMA chart. The in-control run-length distributions are quite different (in location, scale and shape) for this chart for the majority of the underlying distributions, and the overall conclusion is that this particular EWMA chart is not robust to non-normality. The same general conclusion can be made for the EWMA chart corresponding to the pair $(\lambda, L) = (0.2, 2.86)$ from Figure 4 where the severity of non-robustness is clearly obvious. An excessive number of false alarms will be expected under this chart for the *t*, gamma and the contaminated normal distributions. Some particular observations with respect to Figures 3 and 4 are:

(i) The ARL and the spread/variability in the run-length decreases for all the t and the gamma distributions and for the asymmetric bi-modal distribution. This means that these EWMA charts will signal more false alarms than what would be typically expected under normality.

(ii) For the right triangular and the uniform distributions the converse is seen to happen, i.e. the ARL and the variability in the run-length increases. This would cause fewer false alarms than expected. Although a higher in-control ARL might seem favourable, it may lead to higher out-of-control ARL so that it would take longer to detect a change when a change occurs.

Finally, it is interesting to note that although the boxplot-like graphs in panel (a) and (b) in each of Figures 2, 3 and 4 suggest that the in-control performance of the EWMA chart with steady-state limits are very similar to that with exact limits, taking a closer look at the actual values given in Table 3, it becomes apparent that there is a difference between the performances of the EWMA chart based on the two types of limits. For example, for all of the three the (λ, L) combinations and for any of the distributions, the values of the in-control run-length characteristics based on the exact limits are slightly smaller than those based on the steady-state limits. This seems reasonable since the exact limits are narrower at start-up and becomes wider and later approach the steady-state limits; so, the EWMA based on exact limits is more likely to signal incorrectly at start-up leading to a shorter overall run-length. Next we discuss our findings for Case U.

Standards unknown: Case U

Because estimates are substituted for the unknown parameter values in Case U, it is of interest to examine the effects of estimation on the Phase II run-length distribution and hence the performance and the robustness of the traditional EWMA chart. We examine this via five factors:

i. The design parameters **(** λ, L **):** We again used the three combinations or pairs used in Case K and listed in Table 2. Note that, these pairs of design parameters only guarantees an in-control ARL of 370 (in Case K) if we use steady-state control limits.

ii. The type of control limits i.e. whether we use the exact time-varying limits or the steady-state limits.

iii. The underlying process distribution. Note that in order to study the in-control robustness of the EWMA control chart in Phase II applications it is important to assume/ensure that the Phase I and Phase II distributions are the same. If this is not the case, we can view the process as having encountered a change and hence to be OOC – this is not the focus of the present investigation. For Phase II robustness study when parameters are estimated, we have focussed on the following nine distributions: N(0,1), *t*(3), Gam(1,1), Right Triangular, Uniform, Asymmetric bi-modal, Symmetric bi-modal, CN1 and CN2. This set of distributions is a representative subset of the 16 distributions considered in Case K and gives an overall picture of the robustness of the EWMA chart in Phase II of Case U.

iv. The size (m) of the in-control Phase I sample which is denoted by $X_1, X_2, ..., X_m$ from which the parameters are estimated. In our study we used $m = 50$ and $m = 100$; note that, we assume that the in-control Phase I observations are independent and identically distributed.

v. The point estimators for the unknown mean and unknown standard deviation. For the mean we used $\overline{X} = \sum_{i=1}^{m} X_i / m$ and for the standard deviation we used two popular estimators $S/c_4 = \sqrt{\sum_{i=1}^{m} (X_i - \overline{X})^2 /(m-1)} / c_4$ and $\overline{MR}/d_2 = \sum_{i=1}^{m} | X_i - X_{i-1} | / d_2(m-1)$, respectively, where c_4 and d_2 are constants that ensure that the point estimates are unbiased in case of the normal distribution. These two standard deviation estimators were used by Cryer and Ryan [8] to investigate and compare the performance of the Shewhart-type X chart for the normal distribution.. Hence, we have two pairs of point estimators to choose from: $(\overline{X}, S/c_4)$ or $(\overline{X}, \overline{MR}/d_2)$ respectively, for $(\hat{\mu}, \hat{\sigma})$ in the equations for the estimated Phase II control limits given in (5) and (6).

It is important to stress that, given a pair of point estimators ($\hat{\mu},\hat{\sigma}$) calculated using an in-control Phase I sample, we obtain the *conditional* run-length distribution and hence we observe the *conditional* performance of the EWMA chart. That is, the observed performance of the chart is based on that specific in-control Phase I sample. Thus, the (conditional) performance and the (conditional) run-length distribution will be different for each practitioner based on his/her own in-control Phase I sample. Therefore the conditional performance does not give us complete insight in to the overall performance of the chart. In order to get an overall picture and a more general idea about the effects of parameter estimation, we study the *unconditional* run-length distribution; this unconditional distribution can be thought of the run-length distribution averaged over all possible values of the parameter estimators.

To obtain one observation from the *conditional* run-length distribution the following simulation algorithm was used:

i. The levels of the five key factors (listed above) are selected i.e. the type of control limits, one of the 9 underlying process distributions, the size (m) of the in-control Phase I sample, one of the 2 pairs of point estimators $(\overline{X}, S/c_4)$ or $(\overline{X}, \overline{MR}/d_2)$ and one of the 3 pairs of design parameters (λ, L).

ii. An in-control Phase I sample of size (m) is simulated from the chosen distribution and then used to calculate the point estimates ($\hat{\mu},\hat{\sigma}$) and estimate the Phase II control limits according to equations (5) and (6).

iii. An individual Phase II observation, X_i, is generated from the selected distribution (which is the same as the Phase I distribution) and the plotting statistic, Z_i, is calculated according to (4) with the starting values taken as $Z_0 = \hat{\mu} = \overline{X}$.

iv. If $Z_i < U\hat{C}L$ and $Z_i > L\hat{C}L$, then a run-length counter is incremented.

v. Steps (iii) and (iv) are repeated until $Z_i \geq U\hat{C}L$ or $Z_i \leq L\hat{C}L$; when this occurs, a signal is given and the run-length i.e. the time i, is recorded. This run-length is one observation from the conditional Phase II run-length as it depends on the in-control Phase I sample is simulated in (ii).

To obtain the *unconditional* run-length distribution and its associated characteristics (such as the 1^{st}, 25^{th}, 50^{th}, 75^{th} and 99^{th} percentiles as well as the ARL and SDRL) steps (ii) - (v) were repeated 200,000 times; these 200,000 independent conditional run-length observations were used to calculate the associated unconditional characteristics of the Phase II run-length distribution. For example, the average of the conditional run-lengths is an estimate of the unconditional average run-length.

These values were calculated for each combination of the levels of the five factors listed above. Due to space limitations however, the results are not shown here but are available from the authors on request. The results in Case U lead to the following general observations:

i. The run-length distribution in Case U, like in Case K, is positively skewed; the ARL is *larger* than the MDRL,

ii. In Case U the in-control SDRL is *larger* than the in-control ARL (in some cases much larger). This is unlike the situation in Case K where the in-control SDRL is smaller than the in-control ARL (see Table 3). This clearly indicates that extra variability is introduced when then mean and the standard deviation are unknown and needs to be estimated.

iii. In general, we can conclude that the unconditional Phase II run-length distribution in Case U is not the same as the run-length distribution in Case K (for the normal distribution).

For a more in-depth (yet concise) analysis of the simulation results of Case U, we focus on the unconditional in-control ARL in the interpretation and the comparison that follow; similar analysis can be done for any of other characteristics of the run-length distribution. These ARL comparisons are shown in Table 4, which shows the percentage difference between the unconditional Phase II in-control ARL in Case U and the nominal in-control ARL of 370. Here are two examples to illustrate the interpretation of the results in Table 4:

Example 1: From Table 4 we observe that if we use (steady-state limits + the CN2 distribution + the pair of design parameters (0.05, 2.492) + $m = 50$ in-control Phase I observations + $\overline{MR}/d_2$ as an estimator for the unknown standard deviation) the unconditional Phase II in-control ARL of the EWMA chart is 80% less (i.e. listed simply as -80 in Table 4) than the nominal in-control ARL value of 370.

Example 2: From Table 4 we observe that if we use (exact limits + the Gam(1,1) distribution + the pair of design parameters (0.1, 2.703) + $m = 100$ in-control Phase I observations + S/c_4 as an estimator for the unknown standard deviation) the unconditional Phase II in-control ARL of the EWMA chart is 26% larger (i.e. listed simply as 26 in Table 4) than the nominal in-control ARL value of 370.

< Table 4 here>

From Table 4, we observe that:

i. For the pair $(\lambda, L) = (0.05, 2.492)$, in general, the in-control ARL in Case U is *less* than the nominal in-control ARL of 370 (i.e. the percentage differences are predominantly negative) except for the $t(3)$ and the Gam(1,1) distributions where we observe the reverse, that is, the in-control ARL in Case U is *larger* than 370 (i.e. the percentage differences are predominantly positive) when S/c_4 is used as an estimator for the standard deviation; this is true for both types of control limits and both values of m. Thus, in general, we would expect *more* false alarms in Case U when using the design parameters $(\lambda, L) = (0.05, 2.492)$, since most of the percentage differences are negative.

ii. For the pair $(\lambda, L) = (0.1, 2.703)$, point i (as listed above) is again observed but we also see that for the N(0,1), the Right Triangular, the Uniform and the Symmetric bi-modal distributions the in-control ARL in Case U is *larger* than the nominal in-control ARL of 370 (i.e. the percentage differences are positive), specifically when using $\overline{MR}/d_2$ as an estimator for the standard deviation.

iii. For the pair $(\lambda, L) = (0.2\,, 2.86)$, in general, the in-control ARL in Case U is *larger* than the nominal in-control ARL of 370. Thus, in general, we expect *less* false alarms in Case U when using $(\lambda, L) = (0.2\,, 2.86)$ since most percentage differences are positive.

iv. For the CN1 and the CN2 distributions, the in-control ARL in Case U is always *less* than the nominal in-control ARL of 370 (i.e. the percentage difference is always negative); this implies *more* false alarms for these two distributions no matter what design parameters, the value of m and the estimator we use for the unknown standard deviation. This is a big problem.

Finally, one may also be interested in the efficacy of the standard deviation estimators and the effects they may have on the performance of the EWMA chart in Case U. This issue was examined in some detail and the results, not presented here, show that (i) S/c_4 is a better estimator of σ and (ii) that as one might expect, as the size (m) of the in-control Phase I sample, increases the performance of the Phase II EWMA chart in Case U becomes more like that of Case K. Details on these results can be obtained from the authors on request.

5. Conclusions: Summary and Recommendations

We have taken a closer look at and done a more thorough examination of the in-control run-length distributions of three traditional exponentially weighted moving average (EWMA) charts considered in BMR [5], both in the standards known and unknown cases. The in-control robustness is crucial in the proper design and implementation of a control chart, without which its shift detection capability becomes questionable at best. Our results show that in general the traditional EWMA charts are neither nonparametric nor completely robust to all nonnormal distributions. In fact, even the best (in terms of in-control robustness) of the three EWMA charts in BMR [5], the one with the smallest $\lambda = 0.05$, performs very poorly for the symmetric bi-modal or the contaminated normal distribution. It is seen that in the bi-modal case the number of false alarms is significantly less than the nominal whereas in the contaminated normal case, which is a model for occasional outliers, there is an excessive number of false alarms (significantly higher than the nominal) and the situation gets progressively worse as the variance of the underlying distribution increases. Parameter estimation (Case U) is also seen to affect the robustness of the traditional EWMA chart applied in Phase II. The larger the size of the in-control Phase I sample, from which the parameter estimates are obtained, the closer the performance of the EWMA Case U chart to that in Case K.

Although it appears that a traditional EWMA chart can be designed (tuned) for a specific application that is fairly robust to a violation of the underlying distributional assumptions, this is bound to be cumbersome and can not be entirely satisfactory. Also, the EWMA chart is seen to be more robust for smaller values of the smoothing parameter λ , but one concern about taking a value of λ too small is the so-called inertia problem (see, for example, Woodall and Mahmoud [19]). These authors showed that "under a worst-case scenario, a sample mean more than 15 standard errors from the target value does not lead to an immediate out-of-control signal" while using the most robust traditional EWMA chart in BMR corresponding to λ = 0.05 and L = 2.492. The point is that the traditional EWMA chart, designed for the normal distribution, no matter how well tuned, may not perform satisfactorily (more false alarms) for all non-normal process distributions, particularly not for the heavier tailed, the contaminated and the bimodal distributions. Our advice is not to discontinue the use of the EWMA chart in practice; instead we suggest using caution against its over-use, particularly in situations where the (shape of the) underlying process distribution is not sufficiently known and occasional outliers are a concern. Since the EWMA chart is well-known to detect smaller shifts more effectively and yet its in-control robustness is somewhat of a concern from a practical standpoint, it may be useful to consider nonparametric adaptations of the EWMA (NPEWMA) charts, such as the ones considered in Amin et al. [2]. This work is currently in-progress and the results will be reported elsewhere. Note that although we consider only univariate EWMA charts in this paper, similar results are expected for the multivariate case

Finally, our findings are relevant in all applications of traditional EWMA charts including the model-based or residual-based control charts (e.g. Apley *et al.* [3]) where robustness is also a concern.

Acknowledgements

The authors would like to thank the two referees for their insightful comments which have led to improvements in the final version. The research of the first two authors was financially supported by STATOMET and the Department of Statistics at the University of Pretoria.

6. References

[1] J.L. Alfaro and J.F. Ortega, *A robust alternative to Hotelling's T^2 control chart using trimmed estimators*, Quality and Reliability Engineering International. 24(5) (2008), pp. 601-611.

[2] R.W. Amin, M.R. Reynolds, Jr. and S.T. Bakir, *Nonparametric quality control charts based on the sign statistic*, Commun. Stat. Theory Meth. 24(6) (1995), pp. 1597-1623.

[3] D.W. Apley and H.C. Lee, *Design of exponentially weighted moving average control charts for autocorrelated processes with model uncertainty*, Technometrics. 45(3) (2003), pp. 187-198.

[4] N. Balakrishnan, C.B. Read and B. Vidakovic, *Encyclopedia of Statistical Sciences*, 2nd edition, John Wiley, New York, NY, 2006.

[5] C.M. Borror, D.C. Montgomery and G.C. Runger, *Robustness of the EWMA control chart to non-normality*, J. Qual. Technol. 31(3) (1999), pp. 309-316.

[6] D. Brook and D.A. Evans, *An approach to the probability distribution of cusum run length*, Biometrika. 59(3) (1972), pp. 539-549.

[7] S. Chakraborti, *Run length distribution and percentiles: The Shewhart $\overline{X}$ chart with unknown parameters*, Quality Engineering. 19(2) (2007), pp. 119-127.

[8] J.D. Cryer and T.P. Ryan, *The estimation of Sigma for an X Chart: $\overline{MR}/d_2$ or S/c_4*, J. Qual. Technol. 22(3) (1990), pp. 187-192.

[9] P.E. Maravelakis, J. Panaretos and S. Psarakis, *An examination of the robustness to non-normality of the EWMA control charts for the dispersion*, Commun. Stat. Simul. Comput. 34(4) (2005), pp. 1069-1079.

[10] D.C. Montgomery, *Introduction to statistical quality control*, 5th edition, John Wiley, New York, NY, 2005.

[11] D. Radson and A.H. Boyd, *Graphical representation of run length distributions*, Quality Engineering. 17(2) (2005), pp. 301-308.

[12] S.W. Roberts, *Control chart tests based on geometric moving averages*, Technometrics. 1(3) (1959), pp. 239-250.

[13] D.M. Rocke, *Robust control charts*, Technometrics. 31(2) (1989), pp. 173-184.

[14] S.H. Sheu and L. Yang, *The generally weighted moving average control chart for monitoring the process median*, Quality Engineering. 18(3) (2006), pp. 333-344.

[15] E.G. Schilling and P.R. Nelson, *The effect of non-normality on the control limits of* $\overline{X}$ *charts*, J. Qual. Technol. 8(4) (1976), pp. 183-188.

[16] Z.G. Stoumbos and J.H. Sullivan, *Robustness to non-normality of the multivariate EWMA control chart*, J. Qual. Technol. 34(3) (2002), pp. 260-276.

[17] M.C. Testik, G.C. Runger and C.M. Borror, *Robustness properties of multivariate EWMA control charts*, Quality and Reliability Engineering International. 19(1) (2003), pp. 31-38.

[18] W.H. Woodall and D.C. Montgomery, *Research issues and ideas in statistical process control*, J. Qual. Technol. 31(4) (1999), pp. 376-386.

[19] W.H. Woodall and M. A. Mahmoud, *The intertial properties of quality control* charts, Technometrics. 47(4) (2005), pp. 321-338.

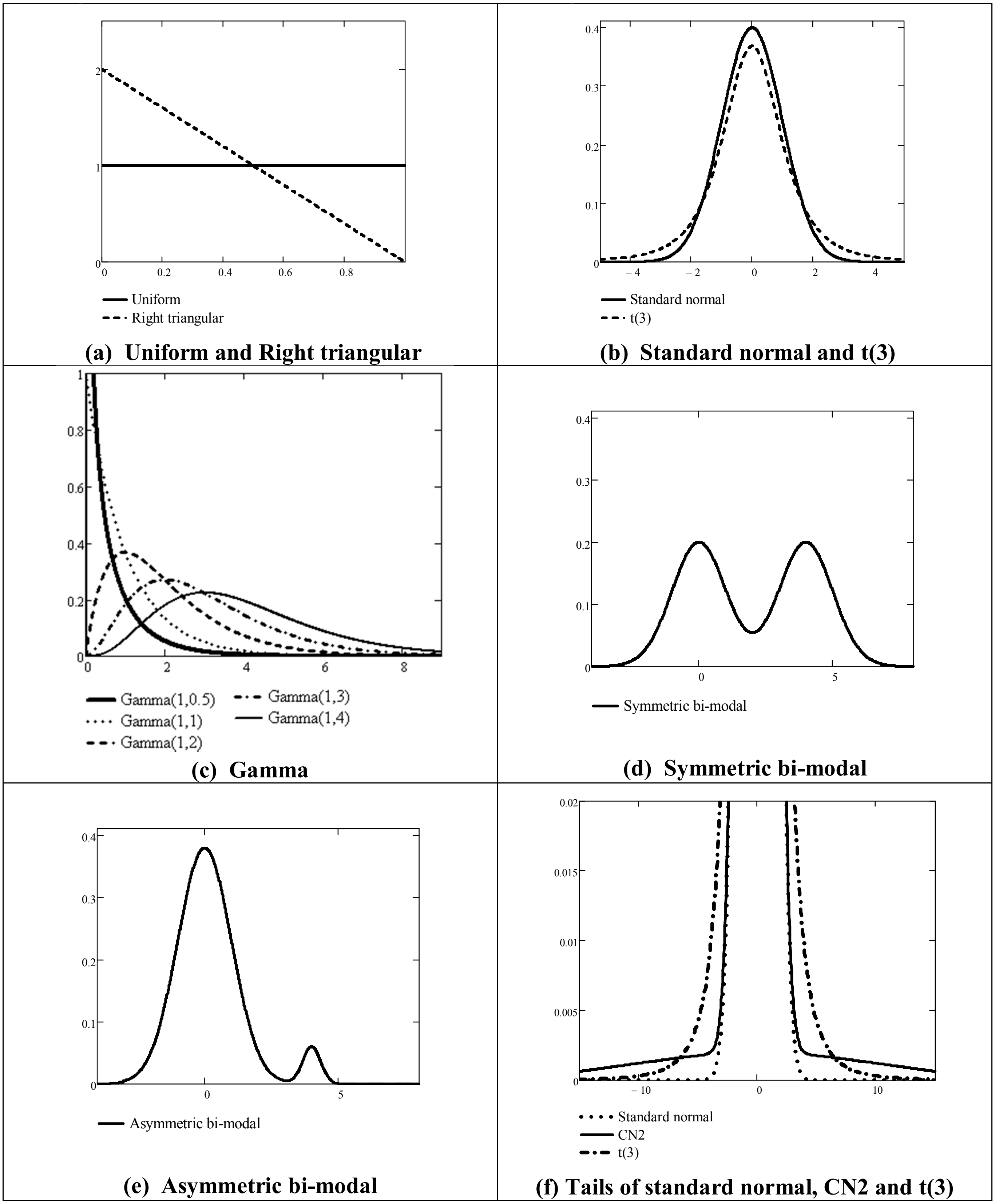


Figure 1: Graphs of the p.d.f.'s of the distributions used in the robustness study

Figure 2: Boxplot-like graphs of the in-control run-length distributions of the EWMA control chart with design parameters $\lambda = 0.05$ and $L = 2.492$

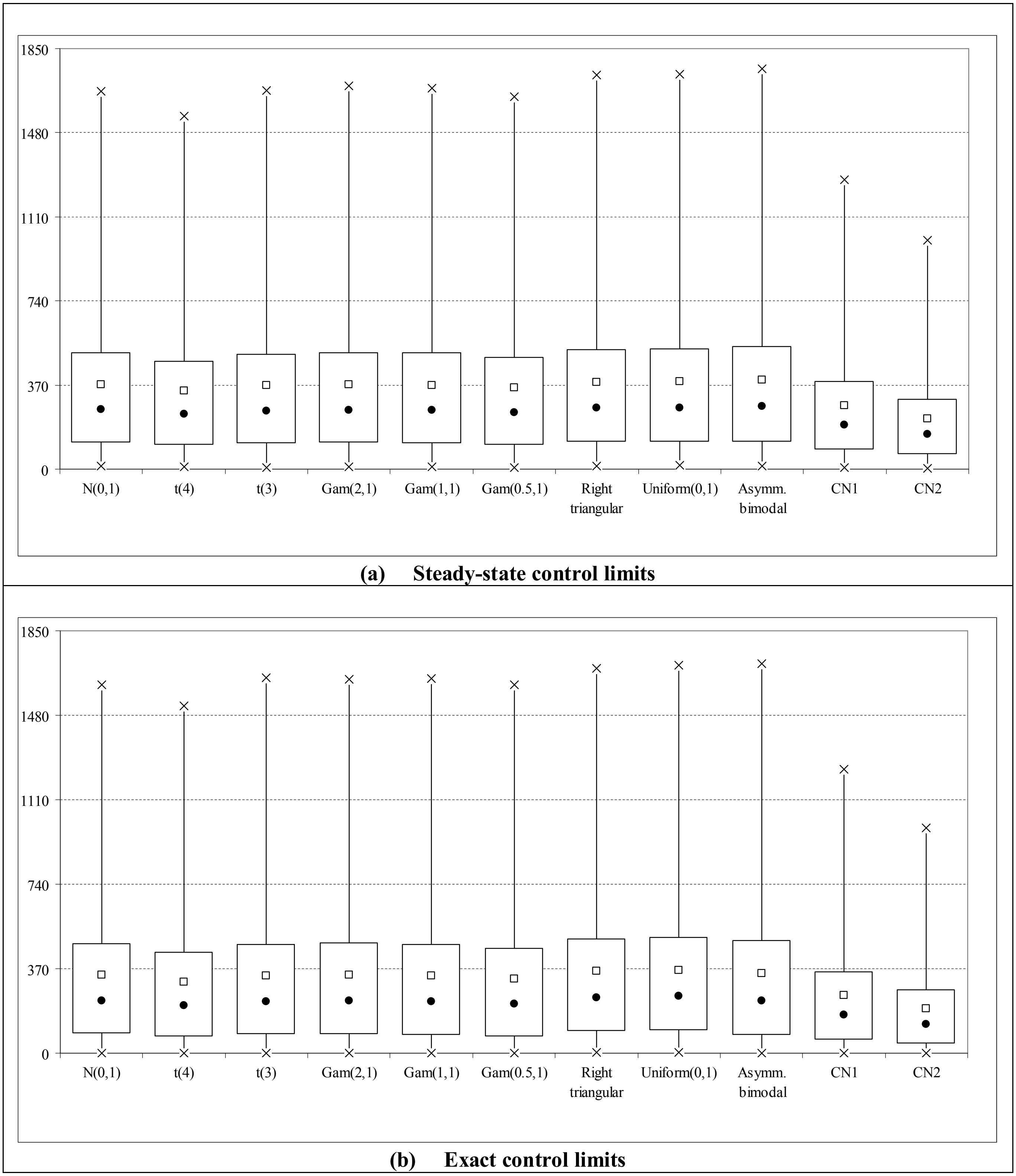
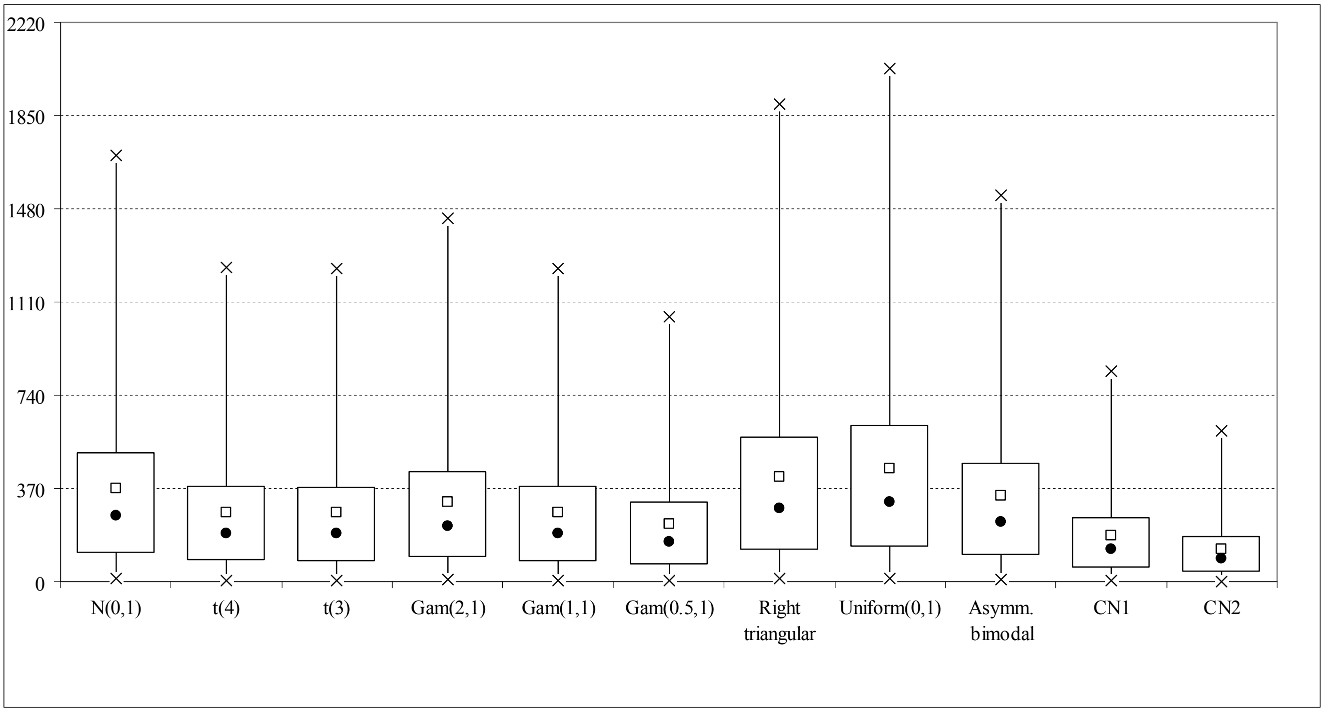
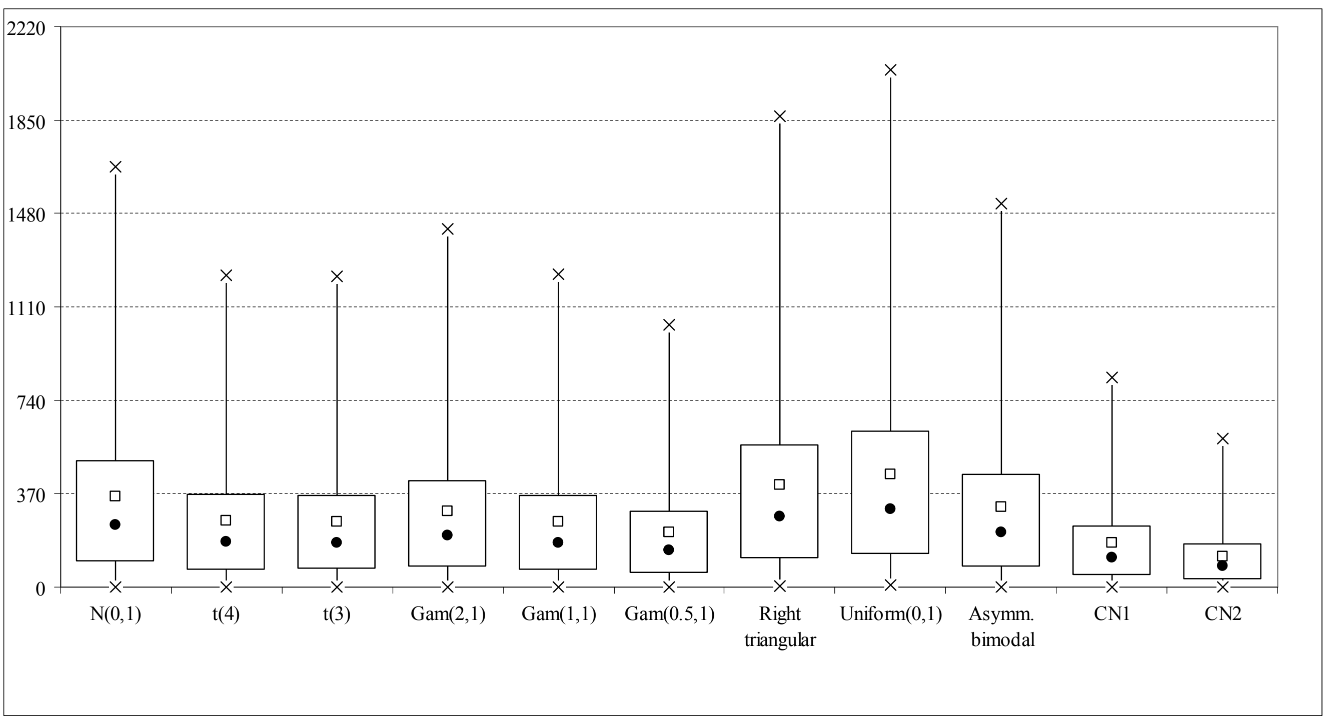


Figure 3: Boxplot-like graphs of the in-control run-length distributions of the EWMA chart with design parameters $\lambda = 0.1$ and $L = 2.703$

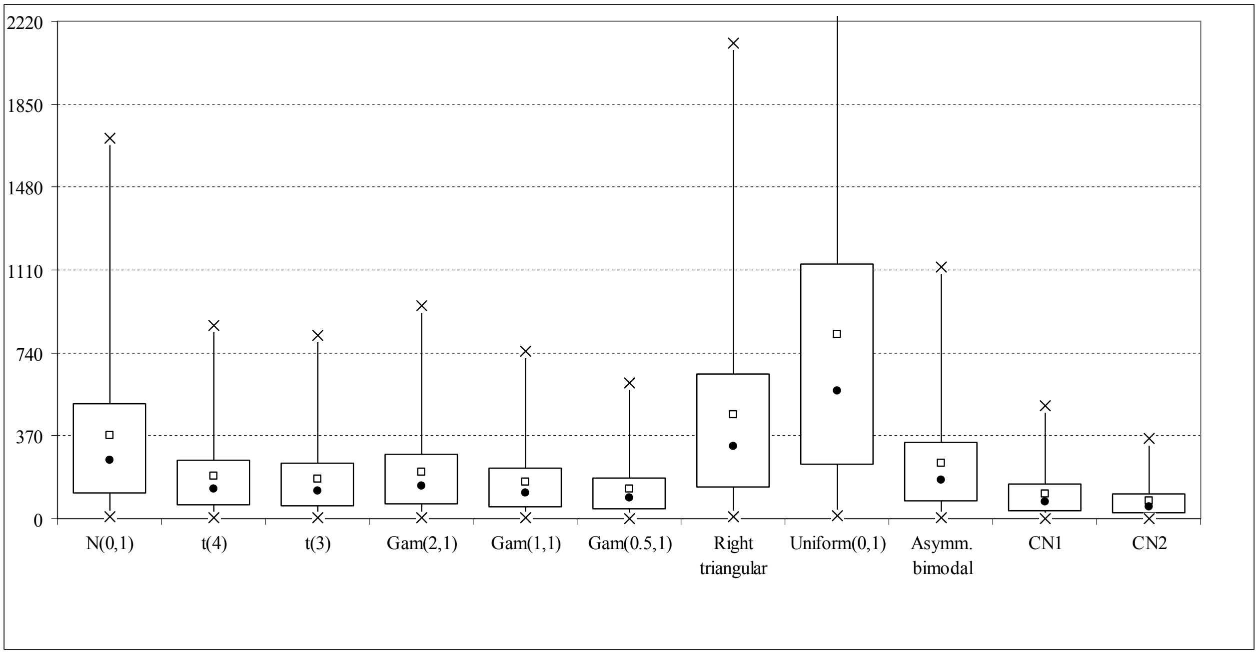


(a) Steady-state control limits

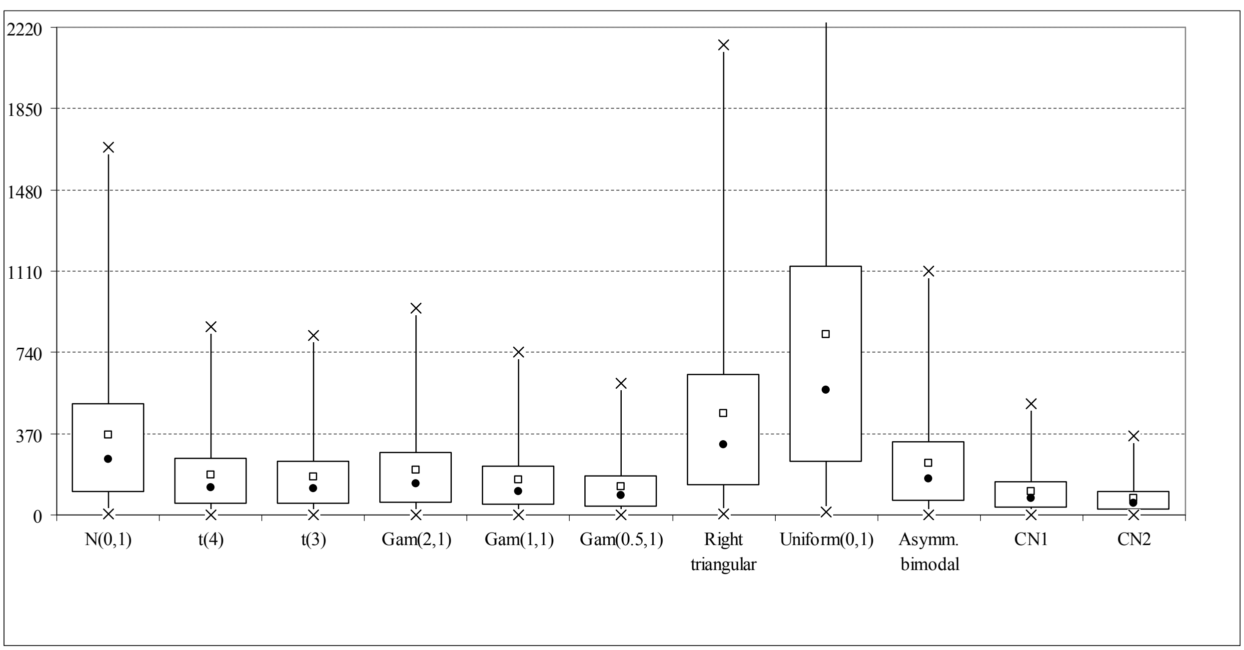


(b) Exact control limits

Figure 4: Boxplot-like graphs of the in-control run-length distributions of the EWMA control chart with design parameters $\lambda = 0.2$ and $L = 2.86$ [1]



(a) Steady-state control limits



(b) Exact control limits

[1]The vertical scales are capped at 2220 to ensure better comparison with other figures; some larger values were observed for the Uniform(0,1) distribution.

Table 1: **Distributions used in the robustness study**

Distribution	p.d.f, mean and variance
Uniform	$f(x)=1 \;\; 0<x<1 \;:\; \mu_0=1/2 \;,\; \sigma_0^2=1/12$
Right triangular	$f(x)=-2x+2 \;\; 0<x<1 \;:\; \mu_0=1/3 \;,\; \sigma_0^2=1/18$
Standard normal	$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \;\; x\in\Re \;:\; \mu_0=0 \;,\; \sigma_0^2=1$
Student's *t*(*v*) with degrees of freedom v = 3, 4, 5 and 6.*	$f(x)=\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\nu/2)\sqrt{\nu\pi}}(1+\frac{x^2}{\nu})^{-\frac{\nu+1}{2}} \quad x\in\Re \;:\; \mu_0=0 \;,\; \sigma_0^2=\frac{v}{v-2},\; v\geq 3$
Gamma with parameters (κ,θ) = (4,1), (3,1), (2,1), (1,1) and (0.5,1).*	$f(x)=\frac{1}{\theta^k\Gamma(\kappa)}x^{\kappa-1}e^{-x/\theta} \quad x>0 \;:\; \mu_0=\kappa\theta,\; \sigma_0^2=\kappa\theta^2$
Symmetric bi-modal	$f(x)=\frac{0.5}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}+\frac{0.5}{\sqrt{2\pi}}e^{-\frac{(x-4)^2}{2}} \quad x\in\Re \;:\; \mu_0=2 \;,\; \sigma_0^2=5$
Asymmetric bi-modal	$f(x)=\frac{0.95}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}+\frac{0.05}{\sqrt{2\pi}/3}e^{-\frac{(x-4)^2}{2(1/3)^2}} \quad x\in\Re \;:\mu_0=0.2 \;,\; \sigma_0^2=1.7156$
Contaminated normal	$f(x)=\frac{0.95}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}+\frac{0.05}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}$ $x\in\Re:\mu_0=0 \;,\; \sigma_0^2=0.95+0.05\sigma^2$

****Note***: The *t*(4), *t*(6) and the gamma distributions were also studied by BMR [5].

Table 2: EWMA design parameters (λ, L) studied in BMR [5]

Pair 1	Pair 2	Pair 3
$(\lambda,L)=(0.05\,,2.492)$	$(\lambda,L)=(0.1\,,2.703)$	$(\lambda,L)=(0.2\,,2.86)$

Table 3: Characteristics of the in-control run-length distribution of the EWMA charts for various distributions

		Pair 1 : $(\lambda, L) = (0.05, 2.492)$							Pair 2 : $(\lambda, L) = (0.1, 2.703)$							Pair 3 : $(\lambda, L) = (0.2, 2.86)$						
	Distribution	1st	25th	MDRL	75th	99th	ARL	SDRL	1st	25th	MDRL	75th	99th	ARL	SDRL	1st	25th	MDRL	75th	99th	ARL	SDRL
Steady-state control limits	N(0,1)	15	117	263	512	1663	372.1	358.3	11	113	260	511	1690	371.2	363.6	8	110	259	512	1696	370.7	366.05
	t(6)	13	110	248	484	1585	352.2	340.1	8	92	214	422	1395	306.3	300.3	5	68	159	316	1050	229.3	227.2
	t(5)	12	108	244	477	1564	347.6	337.5	7	87	204	403	1310	291.2	285.4	4	62	146	290	969	210.7	208.7
	t(4)	11	106	242	475	1553	344.7	335.1	6	83	191	379	1248	275	270.5	3	56	132	260	863	188.5	186.3
	t(3)	8	111	256	505	1667	367.1	359.8	5	80	190	376	1245	272.6	269.6	3	52	124	246	818	178	176.7
	Gam(4,1)	13	116	263	512	1666	372.7	359.6	8	102	238	470	1556	341.6	336.4	5	76	181	360	1184	259.9	257.5
	Gam(3,1)	13	116	262	513	1677	373	361.7	8	100	233	462	1530	334.6	330.3	4	70	165	329	1101	238.1	237.2
	Gam(2,1)	11	115	261	513	1686	372.9	363.2	7	94	219	435	1441	314.9	310.4	4	61	144	287	953	207.8	206.6
	Gam(1,1)	10	112	260	511	1677	370.3	362.5	5	81	191	379	1244	273.7	270	3	48	113	224	747	162.3	161.5
	Gam(0.5,1)	8	105	248	492	1637	356.7	353.7	4	66	158	316	1050	228.8	228.4	2	38	91	181	604	131.3	130.6
	Right triangular	15	120	270	526	1733	383.1	370.8	11	125	290	574	1894	415.9	408.8	9	136	325	645	2123	465.7	461.4
	Uniform(0,1)	16	121	271	528	1736	385.1	372.1	13	137	315	618	2037	449.5	440.1	14	239	570	1136	3785	821	819.1
	Asymm. bi-modal	14	121	275	540	1760	392.5	380.1	9	102	237	469	1532	339.7	333.1	5	73	172	341	1125	246.5	244.1
	Symm. bi-modal	72	1469	3503	6988	23053	5042.5	5011.9	201	5277	12687	25335	83620	18267.9	18221.8	3464	99777	239366	479066	1586839	345245.5	344873.9
	CN1	7	85	196	387	1272	280.9	275.07	3	54	127	253	835	182.9	181	2	32	77	153	504	110.4	109.3
	CN2	4	66	154	306	1008	221.4	217.6	2	39	91	180	599	130.5	129.4	1	23	55	109	360	78.7	78.2

		Pair 1 : $(\lambda, L) = (0.05, 2.492)$							Pair 2 : $(\lambda, L) = (0.1, 2.703)$							Pair 3 : $(\lambda, L) = (0.2, 2.86)$						
	Distribution	1st	25th	MDRL	75th	99th	ARL	SDRL	1st	25th	MDRL	75th	99th	ARL	SDRL	1st	25th	MDRL	75th	99th	ARL	SDRL
Exact control limits	N(0,1)	1	86	231	479	1615	341.3	356.4	2	100	247	500	1667	359.2	364.1	3	104	253	507	1674	364.9	364.5
	t(6)	1	75	213	451	1564	320.2	342.1	1	77	199	407	1377	291.3	299.5	1	62	154	312	1049	224.5	228.3
	t(5)	1	75	212	444	1529	315.6	335	1	73	188	387	1315	276.8	285.6	1	56	140	284	948	203.9	206.8
	t(4)	1	73	209	441	1522	313.1	334	1	68	178	366	1236	261.2	269.8	1	51	126	256	856	183.8	186.7
	t(3)	1	81	227	477	1646	339.2	360.1	1	69	177	363	1233	260.1	268.7	1	48	119	242	817	174.1	177.2
	Gam(4,1)	1	85	232	482	1657	343.6	361.6	1	88	224	459	1537	328.1	336.5	1	70	175	354	1189	254.6	258.3
	Gam(3,1)	1	84	230	480	1634	342.1	360.9	1	84	216	442	1498	316.9	326.8	1	64	160	324	1081	232.2	235.1
	Gam(2,1)	1	82	230	482	1639	342.2	361.8	1	78	204	422	1419	301.1	312.1	1	55	140	283	943	202.7	205.7
	Gam(1,1)	1	79	225	476	1640	337.9	362.1	1	65	175	361	1239	258.6	270.5	1	43	108	220	743	157.8	161.3
	Gam(0.5,1)	1	72	215	460	1612	325.6	352.6	1	53	145	302	1038	215.5	227.4	1	34	87	177	598	127.1	130.7
	Right triangular	2	95	244	500	1687	358.4	369.9	3	114	280	563	1865	405.4	407	5	133	320	640	2142	462.5	463.9
	Uniform(0,1)	4	99	250	508	1699	364	370.8	7	129	309	616	2050	444.5	444	12	239	569	1134	3768	819.3	815.5
	Asymm. bi-modal	1	78	231	495	1706	348.8	375.7	1	80	215	446	1518	317.5	333	1	64	163	331	1109	236.9	242.5
	Symm. bi-modal	53	1448	3482	6962	23148	5028.8	5027.8	195	5320	12666	25316	84351	18295	18272	3421	99588	240877	480231	1602691	346632.9	346679.3
	CN1	1	57	168	358	1244	254.3	273.9	1	44	118	244	832	173.7	181.4	1	29	74	150	505	107.5	109.4
	CN2	1	40	128	279	985	196.7	216.2	1	30	82	172	587	122.4	128.6	1	21	53	107	358	76.5	77.9

Table 4: Percentage difference between the unconditional Phase II in-control ARL in Case U and the nominal in-control ARL of 370

		Pair 1: $(\lambda, L) = (0.05, 2.492)$				Pair 2: $(\lambda, L) = (0.1, 2.703)$				Pair 3: $(\lambda, L) = (0.2, 2.86)$			
		m **= 50**		***m*** **= 100**		***m*** **= 50**		***m*** **= 100**		***m*** **= 50**		***m*** **= 100**	
	Distribution	$\overline{MR}/d_2$	S/c_4	$\overline{MR}/d_2$	S/c_4	$\overline{MR}/d_2$	S/c_4	$\overline{MR}/d_2$	S/c_4	$\overline{MR}/d_2$	S/c_4	$\overline{MR}/d_2$	S/c_4
Steady-state control limits	**N (0,1)**	-18	-29	-18	-23	9	-15	-4	-13	52	7	14	-2
	t(3)	-58	29	-56	59	-63	17	-63	12	-72	-30	-72	-27
	Gam(1,1)	-49	239	-48	10	-27	221	-41	29	-41	13	-60	-34
	Right triangular	-15	-29	-14	-21	42	-4	18	-2	983	98	133	61
	Uniform(0,1)	-7	-32	-5	-23	61	-10	37	-4	2480	105	361	92
	Asymm. bimodal	-34	-22	-35	-18	-19	-4	-32	-10	-10	6	-38	-18
	Symm. bimodal	-10	-32	-9	-24	58	-9	33	-2	1087	116	324	106
	CN1	-63	-8	-62	-21	-71	-35	-71	-45	-79	-63	-79	-67
	CN2	-80	-6	-79	-28	-84	-45	-84	-57	-88	-72	-88	-76
Exact control limits	**N (0,1)**	-24	-34	-24	-29	6	-17	-7	-16	50	5	12	-3
	t(3)	-63	203	-61	37	-66	5	-65	6	-73	102	-73	-18
	Gam(1,1)	-54	61	-53	2	-28	260	-45	26	-42	9	-61	-36
	Right triangular	-20	-34	-19	-27	39	-6	15	-4	932	98	135	62
	Uniform(0,1)	-12	-36	-10	-28	58	-12	35	-6	11928	103	376	91
	Asymm. bimodal	-41	-29	-43	-26	-24	-8	-36	-15	378	3	-40	-20
	Symm. bimodal	-15	-36	-13	-28	64	-10	32	-3	1626	115	319	106
	CN1	-67	-13	-66	-28	-72	-37	-72	-48	-79	-64	-80	-68
	CN2	-82	-10	-82	-34	-85	-46	-85	-59	-88	-72	-88	-76