A Phase II EWMA Chart for the Random Effects Model [1]

Y. Yao and S. Chakraborti
Department of Information Systems, Statistics, and Management Science
The University of Alabama

Abstract: Statistical process control (SPC) and monitoring techniques are often used in applications where multiple sources of variation are present, as for example in a variance components model. The Shewhart $\overline{X}$ control charts for these situations have been studied by Roes and Does (1995), Woodall and Thomas (1995) and other researchers, but Montgomery (2009) argued that the Shewhart $\overline{X}$ control charts are less useful in Phase II and recommended the exponentially weighted moving average (EWMA) chart which is more sensitive to small process mean shifts. We study refinements and adaptations of these charts for the variance components model taking proper account of the effects of parameter estimation while designing and implementing the control charts. In the sequel, we derive and calculate the corrected (adjusted) control limits for Phase II applications. Two types of corrected limits are provided, following the recent literature, one based on the unconditional perspective and the other on the conditional perspective and the exceedance probability criterion (EPC). Tabulations of the charting constants and illustrations with data are provided. An R package is provided to help deployment of the new methodology in practice.

Keywords: Statistical process control; Random effects model; Analysis of variance; Phase II monitoring; Parameter estimation; Robustness.

[1] 9/30/2019

1. Introduction

Statistical Process Control (SPC) and monitoring methods are widely used in various industries. Many famous companies such as IBM, Phillips, Mercedes, GM and GE have successfully taken advantage of the technology and have revolutionized their product lines. Design of Experiments (DOE) is an important component of SPC where various mathematical models are used for the phenomenon under study and the resulting output is monitored with control charts. So control charts need to be developed for the type of model one uses, the parameter(s) of interest, along with the assumptions on the error term. Over twenty years ago, Roes and Does (1995), Woodall and Thomas (1995) among others, considered various control charts including the Shewhart $\overline{X}$ charts (the simplest and the most popular control charts) for monitoring processes that follow some more general linear models, with multiple sources of variation, that are commonly used in DOE. However, Montgomery (2009) considered a major disadvantage of Shewhart $\overline{X}$ charts is that focuses on only the information from the last sample observation, so it is insensitive to detect small process shift. This feature makes these charts less useful in Phase II. One of alternatives is the exponentially weighted moving average (EWMA) control chart. The analogy EWMA charts were developed and studied by Park (1998) and other researchers, but their idea has been silent for years and recently we have some developments in this topic. There is a wealth of knowledge about the effects of parameter estimation, which is most common while setting up the control charts in practice, how it affects chart performance and how to adjust the control limits in order to achieve nominal in-control and reasonable out-of-control chart performance. In this paper, we consider these issues and provide correct Phase II control limits for the EWMA charts for a model more general than the basic Shewhart model. Under this model there are two sources of common cause variation, the variation between and the variation within the batches. In the first section we describe the models, starting from the basic Shewhart model to the model under consideration, and provide the background for our methodology.

1.1 The Basic Shewhart Model

The basic Shewhart model (Montgomery (2009)) is given by

$$X_{ij} = \mu + \epsilon_{ij} \tag{1}$$

where $X_{ij}, i = 1,2,\dots,m,m+1,m+2,\dots, j = 1,2,\dots,n$ is the i-th observation of the j-th batch, μ is the true mean of measurements, and ϵ_{ij} is the i-th normal random error of the j-th batch with mean 0 and variance σ^2.

For $i = 1,2,\dots,m$ the process is in Phase I, so the observations are Phase I data, which are monitored with Phase I control charts constructed using a performance metric (such as the false alarm probability) until process control is established and parameter estimates are obtained. On the other hand, when $i = m+1, m+2,\dots$, the process is in Phase II, and prospective process monitoring starts based on new control limits, derived using some performance metric (such as the in-control average run-length) and using parameter estimates found at the end of Phase I analysis from the reference data. The basic Shewhart model is the simplest and the most commonly studied model in SPC.

The Phase II EWMA plotting statistic for the for this basic model is given by

$$Z_h = \lambda \bar{X}_{h.} + (1-\lambda) Z_{h-1}, Z_0 = \bar{\bar{X}}_{..}, \tag{2}$$

where $h = m+1, m+2, \dots$, $\overline{X}_{h.} = 1/n \sum_{j=1}^{n} X_{hj}$ is h-th Phase II sample (batch mean), $\overline{X}_{..} = 1/m \sum_{i=1}^{m} \overline{X}_{i.}$ is the grand mean of the Phase I data and $\lambda \in (0,1]$ is the smoothing constant. Note that when $\lambda = 1$, the EWMA process is as same as the Shewhart process. The error variance σ_e^2 is

typically estimated by $S_p^2 = 1/m \sum_{i=1}^{m} S_i^2$, the pooled variance of the m Phase I reference sample variances, the $S_i^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_{ij} - \overline{X}_{i.})^2, i = 1,2,\ldots,m.$

Over the last twenty plus years a great deal of knowledge has been acquired on how to monitor both the mean μ and the variance σ^2 of the basic model. We will expand on some of these advances a little later in the paper. For example, in case UU, when the mean μ is unspecified or is unknown, along with σ^2, the control limits for the Phase II EWMA chart for the mean thickness μ are given by

$$\overline{X}_{..} \pm a\sqrt{\frac{S_p^2}{n}}\left(\sqrt{\frac{\lambda}{2-\lambda}(1-(1-\lambda)^{2i})}\right) \qquad i = m+1, m+2, \ldots$$

for the zero state, and by

$$\overline{X}_{..} \pm a\sqrt{\frac{S_p^2}{n}}\left(\sqrt{\frac{\lambda}{2-\lambda}}\right)$$

The charting constant a is typically determined so that the control chart has a specified nominal in-control average run length.

1.2 The Random Effects Model: A Linear Model with Two Random Components for Common Cause Variability

In some applications, there are sources of variation in addition to the error term and thus the basic Shewhart model is not deemed adequate. Several authors have studied this type of situation in the literature. For example, Woodall and Thomas (1995) (hereafter W&T) summarized and reviewed some related approaches to monitor a relatively homogeneous batch of product with some "inherent batch-to-batch variability" which must be accounted for in the model. To this end, they considered an example and stated that "(in the example) batches correspond to production runs of plastic material. Some between-run variation was to be expected because, in addition to setup fluctuations, there were significant time delays between production runs with resulting variability in raw materials and ambient conditions." A similar justification of an extended model was offered by Roes and Does (1995) (hereafter R&D) after they carefully analyzed the production process of integrated circuits (IC) made by Philips Semiconductors, a leading IC chip manufacturer in Europe. They stated that the "wafers are ground(ed) in batches of 31 on the MPS-R600 grinder". They continued to state that "these wafers are positioned on so-called *frits* on a grinding table and located under the grindstone". Thus, R&D argued that a mixed model with fixed and random effects was more appropriate than a basic model. In the R&D model, the fixed effect represented a position effect associated with positions of silicon wafers on the grinder and the random effect represented the between-batch effect.

For a discussion more consistent with the literature, we follow the notation and the terminology in W&T and first consider a simple mixed model with one additional source of variability due to the batch, the so-called variance components model, which is also called a random effects model. This linear model has two components for the variance of an observation, given by

$$X_{ij} = \mu + \beta_i + \epsilon_{ij} \tag{3}$$

where $X_{ij}, i = 1,2,..,m$ and $j = 1,2,\ldots,n$ is the i-th observation from the j-th batch (sampled from all the batches) in Phase I, μ is the true mean, β_i is the effect of the i-th batch which is assumed to be random, and ϵ_{ij} is the random error corresponding. It is assumed that β_i follows a normal distribution with mean 0 and variance σ_b^2 and ϵ_{ij} is normal random variable with mean 0 and variance σ_e^2. In Phase II the linear model is the same in Equation (3), given by

$$X_{ij} = \mu + \beta_i + \epsilon_{ij} \quad (4)$$

where $X_{ij}, i = m+1, m+2, .., \text{and } j = 1,2, \ldots, n$ is the i-th observation from the j-th batch in Phase II and the definitions of the other terms are the same as for Equation (3). Under the model assumptions, the data are from a $N(\mu, \sigma_e^2 + \sigma_b^2)$ distribution in both phases. The goal is to monitor the Phase II mean with and EWMA chart for the random effects model in Equation (4).

The Phase II EWMA chart is based on the plotting statistic $Z_i = \lambda \bar{X}_{i.} + (1-\lambda) Z_{i-1}, Z_0 = \bar{X}_{..}$ for $i = m+1, m+2, \ldots$ In order to study chart performance, the event when the plotting statistic plots outside the control limit, called a signaling event (or its compliment, the non-signaling event) and its probability is important. To this end, for a chosen estimator $\hat{\sigma}$ (discussed below) and the *zero state* EWMA chart, the non-signaling event is given by

$$\bar{X}_{..} - L\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}(1-(1-\lambda)^{2i})} \le Z_i \le \bar{X}_{..} + L\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}(1-(1-\lambda)^{2i})}\,, \quad (5)$$

Similarly, the Phase II non-signaling event for the *steady state* EWMA chart is given by

$$\bar{X}_{..} - L\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}} \le Z_i \le \bar{X}_{..} + L\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}} \quad (6)$$

Before considering the probabilities of these events, note that the batch mean $\bar{X}_{i.}$ follows a $N(\mu, \sigma^2 = \sigma_e^2/n + \sigma_b^2)$ distribution both in Phase I and II, and the Phase I grand mean $\overline{X}_{..} = \sum_{i=1}^{m} \overline{X}_{i.}/m$ follows a $N(\mu, \sigma^2/m)$ distribution. The ANOVA table for the Phase I data under the random effects model is shown in Table 1 (Montgomery, 2017).

Table 1: Phase I ANOVA table for the random effects model

Source	D.F.	SS	MS	E(MS)
Batch	$m-1$	$SSB = n\sum_{i=1}^{m}(\bar{X}_{i.} - \bar{X}_{..})^2$	$MSB = \frac{SSB}{m-1}$	$n\sigma_b^2 + \sigma_e^2$
Error	$m(n-1)$	$SSE = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(X_{ij} - \bar{X}_{i.}\right)^2$	$MSE = \frac{SSE}{m(n-1)}$	σ_e^2
Total	$mn-1$	$SST = \sum_{i=1}^{m}\sum_{j=1}^{n}(X_{ij} - \bar{X}_{..})^2$		

Among the various estimators of the standard deviation (σ) two are popular, each with its pros and cons. The first is based on the uniformly minimum variance unbiased estimator (UMVUE) of σ^2 based on the within-batch mean-squared error. This estimator, denoted by s_b^2, is defined by

$$s_b^2 = \frac{\sum_{i=1}^{m}(\bar{X}_{i.} - \bar{X}_{..})^2}{m-1} = \frac{MSB}{n}. \quad (7)$$

So the first estimator of σ is given by $\hat{\sigma} = s_b$. Note that this estimate is directly obtained from the ANOVA table.

The second estimator is based on the mean of the subgroup moving ranges of span two, denoted mr_m, defined by

$$mr_m = \frac{1}{m-1}\sum_{i=2}^{m}\left|\bar{X}_{i.} - \bar{X}_{(i-1).}\right|$$

The two corresponding unbiased estimators of σ are given by

$$\hat{\sigma}_1 = s_b/c_4 \text{ and } \hat{\sigma}_2 = mr_m/d_2 \quad (8)$$

respectively, where c_4 and d_2 are the corresponding unbiasing constants, tabulated in many textbooks, such as Montgomery (2009).

Both estimators have their pros and cons (see for example Nelson (1982); Cryer and Ryan (1990)). In this paper, we work with the estimator $\hat{\sigma}_1$based on the MSB first as the requisite theory

is exact under normality and some recent research (see for example Mahmoud et al. 2010) has recommended its use for the basic Shewhart model. Moreover, as shown above, this estimator is obtainable from the ANOVA table and thus has some practical advantage when a routine analysis of the data is performed in the process monitoring context. For convenience of computation, we discuss the Phase II EWMA Chart using the. The moving range estimator can be accommodated in our framework with some adaptations, this is not pursued in this paper.

We provide the correct charting constants for the EWMA chart for the random effects model accounting for parameter estimation. It is now well-known that the parameter estimates obtained in the computation of the charting constants necessary to establish the control limits for process monitoring. The charting constants are affected by the model fitted for the problem at hand, the chosen variance estimator, the perspective (unconditional or conditional) and the metric used for in-control performance (which depends on the particular phase of monitoring). One of the most common ways to obtaining the charting constants is to assume that the in-control mean and variance are known (case KK) so that one can calculate the limits using a geometric distribution for the run-length. The famous "3-sigma" rule is an example of this method which yields a FAR of 0.0027 or an in-control ARL of 370. However, in practice the process parameters, the mean and the standard deviation are mostly unknown (case UU) and must be estimated from reference data obtained from a Phase I analysis. In the last several years, many authors have contributed to this literature starting with Chakraborti (2000), Gandy and Kvaloy (2013), Albers et al. (2005), Epprecht et al. (2015), Goedhart et al. (2017) Jardim et al. (2017) and Jardim et al. (2019). This list is not exhaustive but most of this work is for the Shewhart chart. The present work extends this literature to the case of the EWMA charts and corrected charting constants for case UU in the case of the random effects model.

In Section 2, we discuss determination of the charting constants under two perspectives dominant in the current literature. One, under the unconditional perspective, using the conditioning-unconditioning (CUC) method (see, for example, Chakraborti (2000); Felipe et al. (2019)) which guarantees a nominal average in-control ARL value. Two, under the conditional perspective, and using the exceedance probability criterion (EPC) (see, for example, Jardim et al. (2019)). In Section 3, an illustration is provided using the EWMA chart using the corrected charting constants with a dataset based on a case study presented in Roes and Does (1995). In Section 4, we summarize and discuss the discovery of the proposed method.

2 Methodology

It is well known that in Phase II the EWMA control charts are more effective for detecting small shifts and more robust to the violation of normality than Shewhart $\bar{X}$ control charts. The performance of the chart is measured in terms of its robustness and sensitivity. This will be examined by comparing the in-control and the out-of-control performance of the EWMA charts to those of the Phase II Shewhart $\bar{X}$ control charts in Section 3. Recall that the plotting statistic Z_h of the EWMA control chart for the h-th Phase II sample is defined in Equation (2). Also, in this case, the non-signaling event for the zero state is shown in Equation (4) and that for the steady state is shown in Equation (5). Note that the charting constant L are assumed to be unchanged for the zero state and the steady state to some extent, because the difference of ARL between the two states is trivial examined and stated by Lucas and Saccucci (1990). Thus, we only consider the steady state case in the following derivations.

The run length distribution

Many researchers have shown that the key to solving for the charting constant L with estimated parameters is to obtain the distribution of the conditional in-control average run length, $CARL_0$ and to this end we start with the distribution of the conditional false alarm rate $CFAR$. The "conditional" refers to the fact that both of these chart performance characteristics are random

variables and their distributions depend on the parameter estimators used in calculating the Phase II control limits.

Using Equation (5), the *CFAR* for the steady state case is given by

$$CFAR = P\left(|Z_h - \bar{\bar{X}}_{..}| > L\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}}\right) =$$

$$= 1 - P\left(\bar{\bar{X}}_{..} - L\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}} \le Z_h \le \bar{\bar{X}}_{..} + L\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}}\right) \quad (9)$$

Since the charting constant is calculated via the $CARL_0$, it is more meaningful in our context. However, finding the distribution of $CARL_0$ is not easy since unlike in the case of the Shewhart chart, the conditional run length distribution of an EWMA chart is not geometric. For Case KK, Crowder (1987) obtained the exact probability mass function (p.m.f.) of the run length distribution of the EWMA chart and showed that this is not geometric. So the reciprocal of the $CFAR$ is not equal to $CARL_0$ and the distribution of the one can not be calculated from that of the other, as conveniently as in the case of the Shewhart chart. Lucas and Saccucci (1990) introduced an approximate approach to calculate the average run length of the EWMA chart using the absorbing Markov Chain theory. The computation is inspired by theirs and the detailed derivation is given in the Appendix.

Chakraborti and Graham (2019) used results in Fu and Lou (2003) to find the average of the run-length of the conditional in-control average run length. This is given by

$$CARL_0 = \underline{\xi}(\underline{I} - \underline{Q}(U,V;\lambda,L))^{-1}\underline{1} \quad (10)$$

where $\underline{\xi} = [\xi_k]$ is an initial vector which contains the probabilities that the Markov chain starts in a given state, under the condition that $\sum_{k=-t}^{t}\xi_k = 1$. This initial vector depends on the condition of the Phase II process. If the process is (starts) in-control, the elements of the vector are assigned as $\xi_0 = 1$ and $\xi_k = 0$, $k \neq 0$. . Further, $\underline{I}$ is a $(2t+1)\times(2t+1)$ identity matrix, the $(2t+1)\times(2t+1)$ essential transition probability matrix with its elements is given by

$$\underline{Q}(U,V) = \underline{Q}(U,V;\lambda,L) = [q_{kl}(U,V)] = [q_{kl}(U,V;\lambda,L)] \quad (11)$$

where the elements are given by

$$q_{kl}(U,V;\lambda,L) = \Phi\left(\frac{\Phi^{-1}(U)}{\sqrt{m}} + \left(\frac{2l-(1-\lambda)2k+1}{2t+1}\right)\left(\frac{L}{c_4}\sqrt{\frac{F^{-1}_{\chi^2_{m-1}}(V)}{\lambda(2-\lambda)(m-1)}}\right)\right)$$

$$-\Phi\left(\frac{\Phi^{-1}(U)}{\sqrt{m}} + \left(\frac{2l-(1-\lambda)2k-1}{2t+1}\right)\left(\frac{L}{c_4}\sqrt{\frac{F^{-1}_{\chi^2_{m-1}}(V)}{\lambda(2-\lambda)(m-1)}}\right)\right),$$

and $\underline{1}$ is a $(2t+1)\times 1$ vector of 1's. The notation of $\underline{Q}(U,V;\lambda,L)$ and $q_{kl}(U,V)$ are compressed as $\underline{Q}(U,V)$ and $q_{kl}(U,V)$ until it is necessary. Note that the $CARL_0$ is a random variable as a function of U and V, which is the key to obtaining the charting constants for the Phase II Shewhart $\overline{X}$ control charts for the random effects model. We follow this same idea to derive the charting constants for the EWMA chart. Using Equation (11), the c.d.f. of the $CARL_0$ for the EWMA chart is given by

$$F_{CARL_0}(t) = P(CARL_0 \le t) = P\left(\underline{\xi}(\underline{I} - \underline{Q}(U,V))^{-1}\underline{1} \le t\right), t \ge 1 \quad (12)$$

In the next sections, we derive the charting constants under two different perspectives and provide some side results.

2.1 The Unconditional Perspective

Chakraborti (2000, 2006) introduced the charting constant for the Shewhart $\bar{X}$ control chart with estimated parameters, for the basic Shewhart model, using the CUC (conditioning-unconditioning) method. This is called the *unconditional perspective* in Jardim et al. (2019). We follow this idea and find the EWMA charting constant using the *unconditional perspective* for the random effects model. Under this perspective, the charting constant is found by equating the *unconditional in-control average run length* (denoted ARL_{in}), which is the expectation of the conditional in-control average run length $CARL_0$, to a desired nominal value say ARL_0. The nominal values are taken to be 370 or 500, as in the known parameters case. Thus, the ARL_{in} is set to equal to ARL_0 as follow

$$ARL_{in} = \int_0^1 \int_0^1 CARL_0 dudv = \int_0^1 \int_0^1 \underline{\xi}(\underline{I} - \underline{Q}(u, v; \lambda, L))^{-1} \underline{1} dudv = ARL_0 \tag{13}$$

and the charting constant is the solution of this equation. Some values of the charting constants from this perspective are provided in Table 1.

Table 2: EWMA charting constant L under the unconditional perspective for different smoothing constants and numbers of Phase I samples

	$ARL_0 = 370$			$ARL_0 = 500$		
m	$\lambda = 0.5$	$\lambda = 0.8$	$\lambda = 1$	$\lambda = 0.5$	$\lambda = 0.8$	$\lambda = 1$
20	2.7015	2.6800	2.6666	2.7647	2.7416	2.7281
25	2.7635	2.7451	2.7329	2.8325	2.8121	2.7996
30	2.8041	2.7886	2.7774	2.8771	2.8594	2.8479
50	2.8816	2.8747	2.8669	2.9629	2.9533	2.9450
100	2.9343	2.9376	2.9337	3.0219	3.0224	3.0180
150	2.9501	2.9581	2.9558	3.0399	3.0449	3.0422
200	2.9576	2.9682	2.9669	3.0484	3.0560	3.0543

Table 2 is constructed using the solutions of Equation (13). This shows how the charting constants vary for number of batches, smoothing constants and the nominal *ARL*. It is seen that a) the charting constants are monotonically increasing with increasing values of the nominal *ARL*. b) The charting constants are increasing with increasing numbers of batches. c) For a specific number of batches, the maximum values of charting constants are at the lower endpoints of λ. For example, for $m = 20$, the maximum is at $\lambda = 0.5$.

In summary, there may be no evidence that the charting constants converge to 3 or the values decided by the "3-sigma" rule.

2.1.1 Robustness for Normality under the Unconditional Perspective

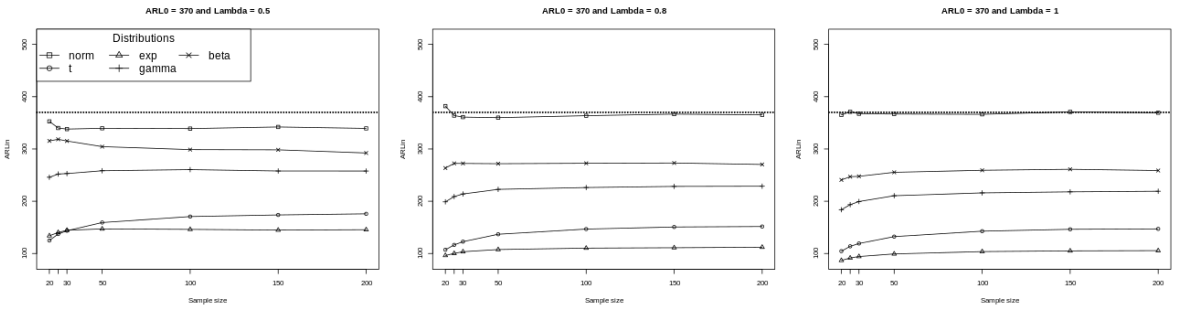
From a practical point of view, it is necessary to examine how the nominal performance is affected when the underlying distribution changes from the assumed normal. In this section we examine this robustness aspect of our adjusted chart, following similar work in Graham et al. (2010) and Capizzi and Masarotto (2013). The investigation is divided into two parts, namely the in-control (IC) and the out-of-control (OOC), respectively. The OOC case is interpreted as the situation where the process distribution is non-normal. The purpose of the IC case is to examine the robustness of the charts in terms of the closeness between the nominal average run length ARL_0 and the simulated

average run length ARL_{in} as a function of the number of batches m and smoothing constants λ, when the data are from a normal distribution. This closeness should be "high". On the other hand, the OOC case is important in order to investigate the robustness of the charts (again, in terms of the closeness between ARL_0 and ARL_{in} as a function of m and λ) but for distributions that are not normal. The various distributions considered are listed and described in Table 3. The IC case is defined by the assumption that all batch means $\bar{X}_{i.}$ are independently and identically distributed, each as a normal random variable with mean μ and variance σ^2. This is Case 1 in Table 3. The OOC cases are defined by Cases 2 through 5 in Table 3.

Table 3: Distributions used for investigating robustness

Case	Process Status	Distribution of the batch mean	Value of Mean μ, Variance σ^2
1	IC	Normal(0,1)	0, 1
2	OOC	t(5)	0, 1.6667
3	OOC	Exponential(1)	1, 1
4	OOC	Gamma(5,1)	5, 5
5	OOC	Beta(8,2)	0.8, 0.0145

In summary, the distribution of the charting statistic is first chosen to be the standard normal distribution. Next, the t distribution with 5 degrees of freedom is used to examine a "small" change in the shape (tails) of the distribution, since this distribution is standard normal like but with heavier tails. This is an OOC situation. Similarly, the last three distributions (Cases 3 through 5) also depict the OOC setting, chosen to investigate the impact of "bigger" changes in the shape, from symmetric to asymmetric distributions, over different types of support. The exponential distribution is asymmetric and extremely right-skewed and the gamma distribution is asymmetric and right-skewed with an infinite support; the gamma distribution begins to look symmetric for larger scale and shape parameters. The beta (8,2) distribution is left-skewed with a finite support. Note that the simulated observations are all "standardized" to eliminate the effects of different means and variances of the underlying distributions. The standardization is performed. The robustness is graphed as follows.



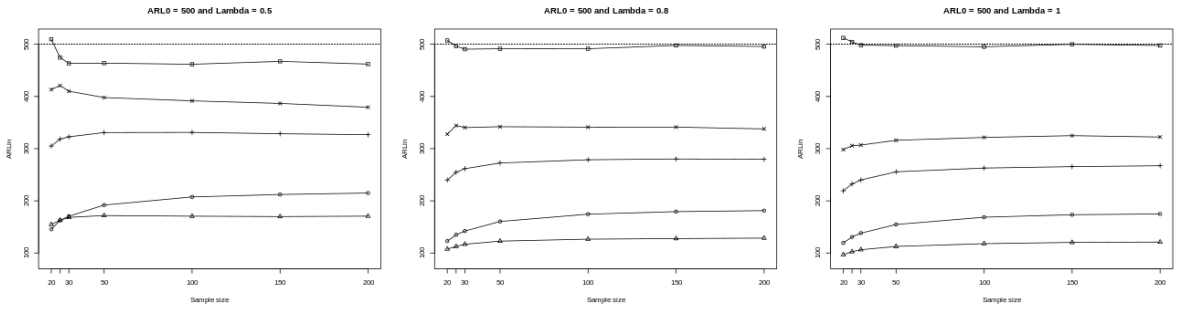


Figure 1: Robustness for $ARL_0 = 370, 500$, $\lambda = 0.5, 0.8$ and 1, and different distributions. Note that the dotted line is the reference line for the different nominal ARL_0's.

In Figure 1, it is seen that a) the IC case is close to the nominal values for high $\lambda = 0.8$ and 1 and the closeness is reduced slightly if $\lambda = 0.5$, and b) the EWMA chart is sensitive to the normality assumption, since the curves of non-normal distributions in cases through 2 to 5 are much far away from the reference line, and c) the EWMA chart is more robust than the Shewhart chart which is $\lambda = 1$, since the curves for non-normal distributions are closer to the ARL_0 for $\lambda < 1$ than those for $\lambda = 1$.

In summary, the EWMA chart is more robust to the violation of normality assumption than the Shewhart chart as expected.

2.2 The Conditional Perspective

Equation (12) shows that the $CARL_0$ is a random variable, so it is quite likely to be different (sometimes significantly) from its expectation, the ARL_0, as specified via the nominal value in the construction of the chart under the unconditional perspective. This may be problematic from a practical point of view. The CUC method guarantees that the average chart performance reaches the target ARL_0 in the long run but it does not consider the variability. A more recent alternative option recognizes the randomness of $CARL_0$ and considers various solutions. To this end, several authors including Albers et al. (2005), Epprecht et al. (2015), Goedhart et al. (2017), and Jardim et al. (2019) have considered setting up the control chart limits so that the $CARL_0$ has a high probability of exceeding a given nominal value (a lower bound) such as 370. This formulation is related to the *exceedance probability criterion* (EPC) proposed by Gandy and Kvaloy (2013) in which one finds an upper bound (that can be tolerated in an application) to the random variable, the conditional in-control false alarm rate $CFAR = CARL_0^{-1}$ with a high probability. The probability of the conditional in-control average run length $CARL_0$ greater than the prediction lower bound using Equation (12) is given by

$$\begin{aligned} P(CARL_0 \geq ARL_b) &= P(CARL_0 \geq (1-\tilde{\varepsilon})ARL_0) \\ &= \int_0^1 \int_0^1 Ind(CARL_0 \geq (1-\tilde{\varepsilon})ARL_0)dudv \\ &= \int_0^1 \int_0^1 Ind(\underline{\xi}(\underline{I} - \underline{Q}(u,v;\lambda,L))^{-1}\underline{1} \geq (1-\tilde{\varepsilon})ARL_0)dudv \\ &= 1 - p_0 \end{aligned} \tag{14}$$

where p_0 $(0 < p_0 < 1)$ is a specified probability, typically a small value, such as 0.1 or 0.05, ε $(0 \leq \varepsilon < 1)$ is a "tolerance factor" allowing the user some flexibility for choosing a nominal conditional false alarm rate, such as 0 or 0.1, $\tilde{\varepsilon}(0 \leq \tilde{\varepsilon} < 1)$ is a "tolerance factor" allowing the user some flexibility for choosing a nominal conditional false alarm rate, $ARL_b = (1-\tilde{\varepsilon})ARL_0$ is a lower

prediction bound which can be seen as the p_0-quantile of the distribution of $CARL_0$, and $Ind(\cdot)$ is an indicator function given by

$$Ind(a > b) = \begin{cases} 1 & \text{if } a > b \\ 0 & \text{if } a \leq b \end{cases}$$

Many researchers (see, for example, Jardim et al., 2019) pointed out that the charting constants under the conditional perspective were different from those from "3-sigma" rules for the Shewhart $\bar{X}$ charts. We verify whether this phenomenon also happens for the EWMA charts.

Table 4: EWMA charting constants under the conditional perspective, for $\lambda = 0.5,0.8,1$, $p_0 = 0.05, 0.1$, $\tilde{\varepsilon} = 0,0.1$ and various values of m

			$ARL_0 = 370$			$ARL_0 = 500$		
p_0	$\tilde{\varepsilon}$	m	$\lambda = 0.5$	$\lambda = 0.8$	$\lambda = 1$	$\lambda = 0.5$	$\lambda = 0.8$	$\lambda = 1$
0.05	0	20	4.2897	4.1898	4.1516	4.4143	4.3169	4.2770
0.05	0	25	4.0904	4.0182	3.9868	4.2127	4.1401	4.1072
0.05	0	30	3.9533	3.9012	3.8742	4.0757	4.0195	3.9912
0.05	0	50	3.6681	3.6519	3.6315	3.7816	3.7627	3.7412
0.05	0	100	3.4178	3.4192	3.4049	3.5255	3.5229	3.5075
0.05	0	150	3.3223	3.3312	3.3155	3.4268	3.4322	3.4155
0.05	0	200	3.2719	3.2770	3.2660	3.3747	3.3764	3.3646
0.05	0.1	20	4.2450	4.1446	4.1070	4.3712	4.2728	4.2335
0.05	0.1	25	4.0452	3.9748	3.9440	4.1707	4.0979	4.0654
0.05	0.1	30	3.9107	3.8590	3.8326	4.0325	3.9784	3.9506
0.05	0.1	50	3.6327	3.6124	3.5925	3.7436	3.7242	3.7031
0.05	0.1	100	3.3793	3.3822	3.3683	3.4882	3.4869	3.4719
0.05	0.1	150	3.2850	3.2952	3.2799	3.3906	3.3972	3.3808
0.05	0.1	200	3.2351	3.2415	3.2309	3.3391	3.3419	3.3304
0.1	0	20	3.9810	3.9225	3.8794	4.0928	4.0415	3.9960
0.1	0	25	3.8289	3.7924	3.7555	3.9484	3.9073	3.8686
0.1	0	30	3.7319	3.7003	3.6672	3.8494	3.8125	3.7779
0.1	0	50	3.5123	3.4971	3.4793	3.6229	3.6017	3.5841
0.1	0	100	3.3274	3.3242	3.3181	3.4321	3.4250	3.4183
0.1	0	150	3.2535	3.2545	3.2517	3.3559	3.3532	3.3494
0.1	0	200	3.2098	3.2162	3.2154	3.3108	3.3138	3.3115
0.1	0.1	20	3.9374	3.8801	3.8379	4.0495	4.0002	3.9553
0.1	0.1	25	3.7865	3.7515	3.7149	3.9067	3.8675	3.8293
0.1	0.1	30	3.6899	3.6603	3.6282	3.8016	3.7735	3.7395
0.1	0.1	50	3.4728	3.4601	3.4423	3.5846	3.5649	3.5477
0.1	0.1	100	3.2901	3.2882	3.2825	3.3958	3.3890	3.3835
0.1	0.1	150	3.2170	3.2195	3.2167	3.3204	3.3189	3.3158
0.1	0.1	200	3.1737	3.1815	3.1808	3.2758	3.2799	3.2779

Table 4 shows that the charting constants under the conditional perspective increase with decreasing p_0, decreasing $\tilde{\varepsilon}$, decreasing m, decrease λ and increasing ARL_0, for small and large sample sizes such as $20 \leq m \leq 100$. For extreme large sample size such as $m > 100$, the maximum of the charting constants is at $\lambda = 0.8$. Also, the charting constants in Table 4 are much greater than those in Table 2, so it is expected that the ARL_{in} for the conditional perspective is much greater than the ARL_0.

Next, we verify the theoretical ARL_{in} in Equation (13).

2.2.1 The In-control Performance under the Conditional Perspective

In this section, we calculate the theoretical ARL_{in} in Equation (13) for measuring the robustness as that in Section 2.1.1. The charting constants are used in Table 4. The ARL_{in} is shown in Table 5.

Table 5: ARL_{in} under the conditional perspective, for $\lambda = 0.5,0.8,1$, $p_0 = 0.05, 0.1$, $\tilde{\varepsilon} = 0,0.1$ and various values of m

			$ARL_0 = 370$			$ARL_0 = 500$		
p_0	$\tilde{\varepsilon}$	m	$\lambda = 0.5$	$\lambda = 0.8$	$\lambda = 1$	$\lambda = 0.5$	$\lambda = 0.8$	$\lambda = 1$
0.05	0	20	2.92E+07	1.08E+07	8.15E+06	1.13E+08	4.20E+07	2.71E+07
0.05	0	25	9.65E+05	6.67E+05	5.71E+05	2.63E+06	1.77E+06	1.47E+06
0.05	0	30	1.67E+05	1.41E+05	1.27E+05	3.97E+05	3.21E+05	2.85E+05
0.05	0	50	1.08E+04	1.11E+04	1.06E+04	1.94E+04	1.99E+04	1.87E+04
0.05	0	100	2.19E+03	2.29E+03	2.22E+03	3.41E+03	3.55E+03	3.41E+03
0.05	0	150	1.37E+03	1.43E+03	1.37E+03	2.04E+03	2.12E+03	2.02E+03
0.05	0	200	1.09E+03	1.10E+03	1.07E+03	1.59E+03	1.61E+03	1.55E+03
0.05	0.1	20	1.70E+07	7.02E+06	5.39E+06	7.22E+07	2.71E+07	1.78E+07
0.05	0.1	25	6.74E+05	4.77E+05	4.11E+05	1.85E+06	1.25E+06	1.06E+06
0.05	0.1	30	1.25E+05	1.06E+05	9.61E+04	2.91E+05	2.40E+05	2.14E+05
0.05	0.1	50	9.01E+03	9.11E+03	8.68E+03	1.59E+04	1.62E+04	1.53E+04
0.05	0.1	100	1.88E+03	1.97E+03	1.91E+03	2.92E+03	3.04E+03	2.93E+03
0.05	0.1	150	1.19E+03	1.25E+03	1.19E+03	1.77E+03	1.85E+03	1.76E+03
0.05	0.1	200	9.53E+02	9.68E+02	9.40E+02	1.39E+03	1.41E+03	1.36E+03
0.1	0	20	1.30E+06	9.63E+05	7.31E+05	3.56E+06	2.73E+06	1.99E+06
0.1	0	25	1.33E+05	1.24E+05	1.04E+05	3.20E+05	2.86E+05	2.34E+05
0.1	0	30	3.88E+04	3.78E+04	3.32E+04	8.28E+04	7.77E+04	6.71E+04
0.1	0	50	5.01E+03	5.16E+03	4.97E+03	8.58E+03	8.63E+03	8.32E+03
0.1	0	100	1.53E+03	1.56E+03	1.56E+03	2.32E+03	2.35E+03	2.35E+03
0.1	0	150	1.06E+03	1.07E+03	1.07E+03	1.55E+03	1.56E+03	1.56E+03
0.1	0	200	8.70E+02	8.83E+02	8.88E+02	1.25E+03	1.27E+03	1.27E+03
0.1	0.1	20	8.93E+05	6.73E+05	5.19E+05	2.40E+06	1.89E+06	1.39E+06
0.1	0.1	25	9.82E+04	9.29E+04	7.84E+04	2.34E+05	2.13E+05	1.76E+05
0.1	0.1	30	2.99E+04	2.95E+04	2.61E+04	6.06E+04	6.03E+04	5.24E+04
0.1	0.1	50	4.16E+03	4.32E+03	4.17E+03	7.10E+03	7.19E+03	6.94E+03
0.1	0.1	100	1.33E+03	1.35E+03	1.35E+03	2.01E+03	2.03E+03	2.03E+03
0.1	0.1	150	9.27E+02	9.37E+02	9.40E+02	1.36E+03	1.36E+03	1.37E+03
0.1	0.1	200	7.66E+02	7.78E+02	7.83E+02	1.10E+03	1.12E+03	1.12E+03

In Table 5, the ARL_{in} under the conditional perspective increase with decreasing p_0, decreasing $\tilde{\varepsilon}$, decreasing m, decrease λ and increasing ARL_0, for small and moderate sample sizes such as $20 \leq \lambda \leq 30$. For moderate and extreme large sample size such as $\lambda > 30$, the maximum of the charting constants is at $\lambda = 0.8$. Also, all ARL_{in} are much greater than the ARL_0 compared to those in section 2.1.1.

In summary, since the ARL_{in} is much larger than the ARL_0, it is sensitive to the range of different constants under the conditional perspective even if the normality is satisfied. Thus, it is unreasonable to continue to study the robustness using ARL_{in}.

Next, we show an application using the charting constants under different perspectives.

3 Application

We illustrate the proposed charts with the data from Roes and Does (1995). The data were collected from the MPS-R600 grinder measuring the wafer thickness, with a target thickness of 244 μm. They were taken from 5 of 31 positions on the grinder. We consider these data as Phase I data number of Phase I batches $m = 30$, each with $n = 5$ observations, with the overall mean and standard deviation equal to 245.1 and 2.0367, respectively, so the correcting constant c_4 is 0.9914 in Montgomery (2009). Following this setting, we simulate 20 batches of Phase II data shown in Table 6.

Table 6: Simulated Phase II sample for the example

Batch	Batch Mean	Batch	Batch Mean	Batch	Batch Mean	Batch	Batch Mean
1	246.303	6	241.365	11	247.312	16	249.229
2	246.558	7	246.395	12	251.285	17	245.730
3	244.875	8	244.533	13	248.312	18	246.870
4	244.168	9	244.516	14	248.620	19	249.853
5	246.345	10	243.211	15	246.009	20	248.165

The data are divided into 2 sets. The first set contains the first 10 observations with the distribution as same as Roes and Does (1995)'s sample, that is a normal (245.1, 2.0367). The last 10 observations make up the second set which are from a normal (247.55, 2.0367) distribution, which has a 1% increase in the mean over that for the first set. The control charts are shown in Figure 4. Because we have two types of charting constants (unconditional and the conditional) that may be used in Phase II, two pairs of control charts (limits) are shown. First, under the unconditional perspective, for a nominal $ARL_0 = 500$ and $\lambda = 0.5$, the value of charting constant is 2.8771 which can be found from Table 2, which leads to the lower and upper control limits, 241.6875 and 248.5125 in the steady state, respectively. Second, given under the EPC perspective, the control limits are obtained as 240.3171 and 249.8829 for $ARL_0 = 500, p_0 = 0.05, \tilde{\varepsilon} = 0.1$ and $\lambda = 0.5$, and the value of charting constant is 4.0325 from Table 4. Also, the reference control limits are shown, 241.4575 and 248.7425, calculated with the charting constant 3.071 in Lucas and Saccucci (1990).

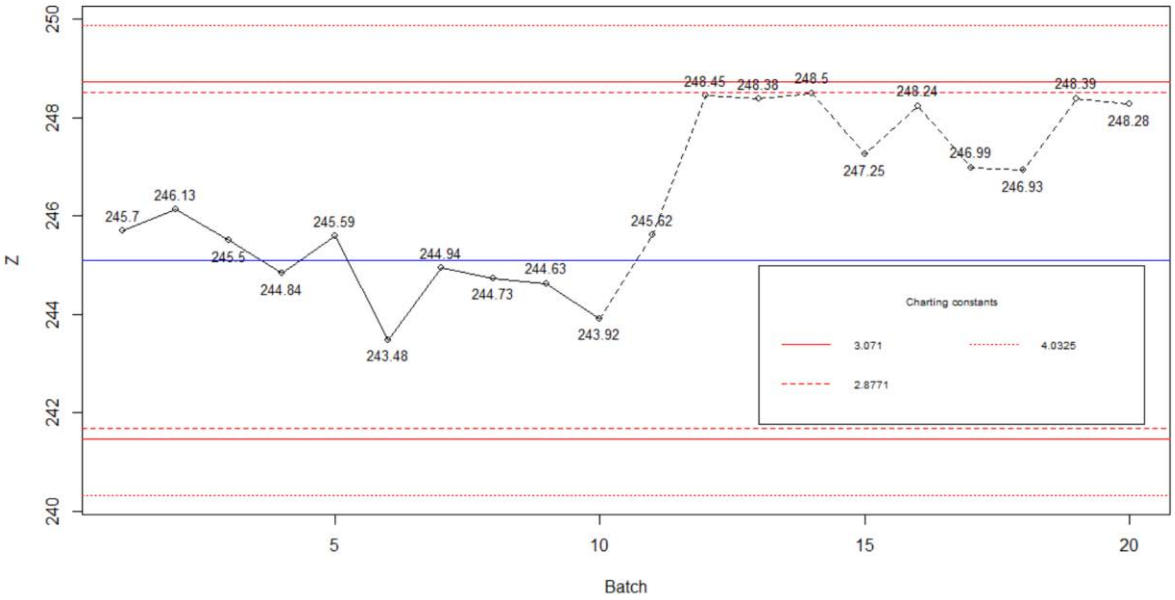


Figure 2: EWMA chart for simulated data

For the value of the charting constant 2.8771 obtained under the unconditional perspective, there is one possible signal, compared to no signal for those from Lucas and Saccuci (1990) and the conditional perspective. Thus, for the unconditional perspective, it can capture signals quicker than other methods. The limits under the conditional perspective are much wider.

4 Conclusion

Based on the results of our study, several conclusions can be made in terms of the Phase II EWMA chart for the random effects model as follows.

a. The charting constants under the unconditional perspective are less than those in Lucas and Saccucci (1990). Thus, for the unconditional perspective is less. Since the ARL_{in} using Table 2 is close to the ARL_0, the ARL_{in} using Lucas and Saccucci (1990) may be greater than the ARL_0 for the random effects model.
b. The performance of the EWMA chart under the unconditional perspective is more robust with decreasing λ.
c. The control limits under the conditional perspective is the widest, so the ARL_{in} is extremely large shown in Table 5. To reach appropriate charting constants with more reasonable ARL_{in}, users may have to increase $\boldsymbol{p_0}$, $\tilde{\varepsilon}$ and m. That may lead to two results. One is there may be no guaranteed to perform well if $\boldsymbol{p_0}$ and $\tilde{\varepsilon}$ are unreasonably high. Another is users may be extremely costly if m is large.

Reference

Albers, Willem, Wilbert CM Kallenberg, and Sri Nurdiati. (2005). "Exceedance Probabilities for Parametric Control Charts." *Statistics* 39 (5). Taylor & Francis: 429–43.

Chakraborti, S. (2000). “Run Length, Average Run Length and False Alarm Rate of Shewhart X-Bar Chart: Exact Derivations by Conditioning.” *Communications in Statistics-Simulation and Computation* 29 (1). Taylor & Francis: 61–81.

Epprecht, Eugenio K, Lorena D Loureiro, and S. Chakraborti. (2015). “Effect of the Amount of Phase I Data on the Phase Ii Performance of S 2 and S Control Charts.” *Journal of Quality Technology* 47 (2). Taylor & Francis: 139–55.

Gandy, Axel, and Jan Terje Kvaløy. (2013). “Guaranteed Conditional Performance of Control Charts via Bootstrap Methods.” *Scandinavian Journal of Statistics* 40 (4). Wiley Online Library: 647–68.

Goedhart, Rob, Marit Schoonhoven, and Ronald JMM Does. (2017). “Guaranteed in-Control Performance for the Shewhart X and X Control Charts.” *Journal of Quality Technology* 49 (2). Taylor & Francis: 155–71.

Jardim, F., S. Chakraborti, and & Epprecht. E. K. (2019). “Two Perspectives for Designing a Phase II Control Chart with Estimated Parameters: The Case of the Shewhart X-Bar Chart.” *Journal of Quality Technology*. Taylor & Francis:: 1-20

Lucas, James M, and Michael S Saccucci. (1990). “Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements.” *Technometrics* 32 (1). Taylor & Francis Group: 1–12.

Mahmoud, Mahmoud A, G Robin Henderson, Eugenio K Epprecht, and William H Woodall. (2010). “Estimating the Standard Deviation in Quality-Control Applications.” *Journal of Quality Technology* 42 (4). Taylor & Francis: 348–57.

Montgomery, Douglas C. (2009). *Introduction to Statistical Quality Control*. John Wiley & Sons (New York).

Park, Changsoon. (1998). “Design of and Ewma Charts in a Variance Components Model.” *Communications in Statistics-Theory and Methods* 27 (3). Taylor & Francis: 659–72.

Roes, Kit CB, and Ronald JMM Does. (1995). “Shewhart-Type Charts in Nonstandard Situations.” *Technometrics* 37 (1). Taylor & Francis: 15–24.

Woodall, William H, and Edward V Thomas. (1995). “Statistical Process Control with Several Components of Common Cause Variability.” *IIE Transactions* 27 (6). Taylor & Francis: 757–64.

Crowder, Stephen V. (1987). "Average run lengths of exponentially weighted moving average control charts." *Journal of Quality Technology* 19 (3). Taylor & Francis: 161-164.

Chakraborti, S., & Graham, M. (2019). *Nonparametric statistical process control*. Wiley.

Gandy, A., & Kvaløy, J. T. (2013). Guaranteed conditional performance of control charts via bootstrap methods. *Scandinavian Journal of Statistics*, 40(4): 647-668.

Cryer, J. D., & Ryan, T. P. (1990). The Estimation of Sigma for an X Chart: or S/c4?. *Journal of Quality Technology*. *22*(3): 187-192

Fu, J. C., & Lou, W. W. (2003). *Distribution theory of runs and patterns and its applications: a finite Markov chain imbedding approach*. World Scientific.

Appendix

Lucas and Saccucci (1990) introduced an approximate approach to calculate the average run length of the EWMA chart using the absorbing Markov Chain theory. Their results are close to those given in Crowder (1987) but more computationally convenient. Chakraborti and Graham (2019) provide a thorough explanation of the Markov Chain approach used in SPC. We describe it briefly here. Let the region between the *LCL* and the *UCL* (where the process is in-control) be divided into an odd number of $v = 2t + 1$ mutually exclusive transient states. The out-of-control state is the region below the *LCL* or above the *UCL* and these two regions together comprise the so-called absorbing state. The $2t + 1$ in-control states can be identified by $k, l = -t, -t + 1, \dots, -1, 0, 1, \dots, t - 1, t$. The $(2t + 2) \times (2t + 2)$ transition probability matrix, $\underline{M}$, is given by

$$\underline{M} = \begin{pmatrix} \underline{Q} & | & \underline{P} \\ \hline \underline{0}^t & | & 1 \end{pmatrix}$$

where $\underline{Q}$ is the $(2t + 1) \times (2t + 1)$ sub-matrix called the essential transition probability matrix containing all the probabilities of going from one transient state to another, $\underline{P}$ is the column vector with $(2t + 1)$ probabilities of going from one transient state to the absorbing state, $\underline{0}$ is the column vector with $(2t + 1)$ zero probabilities of going from the absorbing state to one transient state, and 1 is the scalar value with the probability of going from the absorbing state to the absorbing state. Chakraborti and Graham (2019) showed the computation of the average run length for the EWMA chart using the absorbing Markov chain. We get the conditional average run length using the similar method. The conditional average run length is given by

$$CARL_0 = \underline{\xi}(\underline{I} - \underline{Q}(U, V; \lambda, L))^{-1}\underline{1}$$

where $\underline{\xi} = [\xi_k]$ is an $1 \times (2t + 1)$ initial vector which contains the probabilities that the Markov chain starts in a given state conditioning on $\sum_{k=-t}^{t} \xi_k = 1$, $\underline{I}$ is an $(2t + 1) \times (2t + 1)$ identity matrix, $\underline{Q}(U, V; \lambda, L) = [q_{kl}(U, V; \lambda, L)]$ is the $(2t + 1) \times (2t + 1)$ essential transition probability matrix, and $\underline{1}$ is a $(2t + 1) \times 1$ vector with 1's. If $\lambda > 0$, the $\underline{Q}(U, V; \lambda, L)$ matrix is given by

$$\begin{aligned} q_{kl}(U, V; \lambda, L) &= P(S_l \le \lambda \bar{X}_{h.} + (1 - \lambda)Z_{h-1} \le S_{l+1} | Z_{h-1} = T_k) \\ &= P(T_l - \tau \le \lambda \bar{X}_{h.} + (1 - \lambda)Z_{h-1} \le T_l + \tau | Z_{h-1} = T_k) \\ &= P(T_l - \tau \le \lambda \bar{X}_{h.} + (1 - \lambda)T_k \le T_l + \tau) \\ &= P\left(\frac{T_l - \tau - (1 - \lambda)T_k}{\lambda} \le \bar{X}_{h.} \le \frac{T_l + \tau - (1 - \lambda)T_k}{\lambda}\right) \end{aligned}$$

where $k, l = -t, -t + 1, \dots, t - 1, t$. Also, the half width τ is given by

$$\tau = \tau(\hat{\sigma}; \lambda, L, \theta) = \frac{UCL - LCL}{2(2t + 1)} = \frac{L\hat{\sigma}\sqrt{\frac{\lambda}{2 - \lambda}(1 - (1 - \lambda)^{2h})}}{2t + 1} = \frac{L\hat{\sigma}\sqrt{\frac{\lambda}{2 - \lambda}\theta}}{2t + 1}$$

where $\theta = 1 - (1 - \lambda)^{2h} \in (0, 1]$ is a constant for the zero or steady state. When $\theta = 1$, the Phase II process is in the steady state. When θ is less than 1, the Phase II process is in the zero state. Also, $\tau(\hat{\sigma}; \lambda, L, \theta)$ is written as τ until it is necessary. This midpoint T_k of the k-th state can be derived from the cumulative half width function starting at the lower control limit as follow

$$T_k = LCL + (2(t + k) + 1)\tau$$

where $2(t + k) + 1 = 1, 3, 5, \dots, 4t + 1$ is the number of the cumulative half widths. Also, the lower and upper limits S_k and S_{k+1} of the k-th state are given by

$$S_k = T_k - \tau = LCL + 2(t+k)\tau \text{ and } S_{k+1} = T_k + \tau = LCL + (2(t+k)+2)\tau$$

The midpoint of the k-th state can be derived further as follows

$$T_k = LCL + (2(t+k)+1)\tau = \bar{\bar{X}}_{\cdot\cdot} - L\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}\theta} + (2(t+k)+1)\frac{L\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}\theta}}{2t+1}$$
$$= \bar{\bar{X}}_{\cdot\cdot} + \frac{2k}{2t+1}\left(L\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}\theta}\right) = \bar{\bar{X}}_{\cdot\cdot} + 2k\tau$$

and the difference between T_l and T_k is given by

$$T_l - T_k = \bar{\bar{X}}_{\cdot\cdot} + \frac{2l}{2t+1}\left(L\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}\theta}\right) - \bar{\bar{X}}_{\cdot\cdot} - \frac{2k}{2t+1}\left(L\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}\theta}\right) = \frac{2(l-k)}{2t+1}\left(L\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}\theta}\right)$$
$$= 2(l-k)\tau$$

So the lower limit of the element of $\underline{Q}(U,V;\lambda,L)$ is given by

$$\frac{T_l - \tau - (1-\lambda)T_k}{\lambda} = \frac{1}{\lambda}(T_l - T_k - \tau + \lambda T_k) = \frac{1}{\lambda}(2(l-k)\tau - \tau) + \bar{\bar{X}}_{\cdot\cdot} + 2k\tau$$
$$= \bar{\bar{X}}_{\cdot\cdot} + \frac{1}{\lambda}(2l - (1-\lambda)2k - 1)\tau$$

And the lower limit of the element of $\underline{Q}(U,V;\lambda,L)$ can be transformed further as follow

$$\frac{T_l - \tau - (1-\lambda)T_k}{\lambda\sqrt{\sigma^2}} - \frac{\mu_2}{\sqrt{\sigma^2}} = \frac{\bar{\bar{X}}_{\cdot\cdot}}{\sqrt{\sigma^2}} + \frac{1}{\lambda\sqrt{\sigma^2}}(2l - (1-\lambda)2k - 1)\tau(\hat{\sigma};\lambda,L,\theta)$$
$$= \frac{W}{\sqrt{m}} + \frac{1}{\lambda}(2l - (1-\lambda)2k - 1)\tau\left(\frac{\sqrt{Y}}{c_4\sqrt{m-1}};\lambda,L,\theta\right) - \delta$$

where $W = \frac{\hat{\mu}-\mu}{\sqrt{\sigma^2/m}} \sim N(0,1)$, $Y = \frac{(m-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{m-1}$ and $\delta = \frac{\mu_2}{\sqrt{\sigma^2}} - \frac{\mu}{\sqrt{\sigma^2}}$. Similarly, the upper limit of the element of $\underline{Q}(U,V;\lambda,L)$ for the zero state can be transformed as follow

$$\frac{T_l + \tau - (1-\lambda)T_k}{\lambda\sqrt{\sigma^2}} - \frac{\mu_2}{\sqrt{\sigma^2}} = \frac{W}{\sqrt{m}} + \frac{1}{\lambda}(2l - (1-\lambda)2k + 1)\tau\left(\frac{\sqrt{Y}}{c_4\sqrt{m-1}};\lambda,L,\theta\right) - \delta$$

So the elements of the conditional essential transition probability matrix are given by

$$q_{kl}(W,Y;\lambda,L) = \Phi\left(\frac{W}{\sqrt{m}} + \frac{1}{\lambda}(2l - (1-\lambda)2k - 1)\tau\left(\frac{\sqrt{Y}}{c_4\sqrt{m-1}};\lambda,L,\theta\right) - \delta\right)$$
$$- \Phi\left(\frac{W}{\sqrt{m}} + \frac{1}{\lambda}(2l - (1-\lambda)2k + 1)\tau\left(\frac{\sqrt{Y}}{c_4\sqrt{m-1}};\lambda,L,\theta\right) - \delta\right)$$

For the convenience of computation, let $W = \Phi^{-1}(U)$ and $Y = F^{-1}_{\chi^2_{m-1}}(V)$ where U and V are i.i.d. from uniform distribution with minimum 0 and maximum 1. Φ^{-1} is the quantile function of the

standard normal distribution and $F^{-1}_{\chi^2_{m-1}}$ is the quantile function of the chi-square distribution with degrees of freedom $m-1$. The elements can be rewritten as

$$q_{kl}(U,V;\lambda,L) = \Phi\left(\frac{\Phi^{-1}(U)}{\sqrt{m}} + \frac{1}{\lambda}(2l-(1-\lambda)2k+1)\tau\left(\frac{\sqrt{F^{-1}_{\chi^2_{m-1}}(V)}}{c_4\sqrt{m-1}};\lambda,L,\theta\right) - \delta\right)$$
$$- \Phi\left(\frac{\Phi^{-1}(U)}{\sqrt{m}} + \frac{1}{\lambda}(2l-(1-\lambda)2k-1)\tau\left(\frac{\sqrt{F^{-1}_{\chi^2_{m-1}}(V)}}{c_4\sqrt{m-1}};\lambda,L,\theta\right) - \delta\right)$$

If the process for the steady state is in-control, $\delta = 0$ and as $h \to \infty$, $\theta \to 1$, so the in-control elements are given by

$$q_{kl}(U,V;\lambda,L) = \Phi\left(\frac{\Phi^{-1}(U)}{\sqrt{m}} + \frac{1}{\lambda}(2l-(1-\lambda)2k+1)\tau\left(\frac{\sqrt{F^{-1}_{\chi^2_{m-1}}(V)}}{c_4\sqrt{m-1}};\lambda,L\right)\right)$$
$$- \Phi\left(\frac{\Phi^{-1}(U)}{\sqrt{m}} + \frac{1}{\lambda}(2l-(1-\lambda)2k-1)\tau\left(\frac{\sqrt{F^{-1}_{\chi^2_{m-1}}(V)}}{c_4\sqrt{m-1}};\lambda,L\right)\right)$$

where $\tau\left(\frac{\sqrt{F^{-1}_{\chi^2_{m-1}}(V)}}{c_4\sqrt{m-1}};\lambda,L\right)$ is the special case of $\tau\left(\frac{\sqrt{F^{-1}_{\chi^2_{m-1}}(V)}}{c_4\sqrt{m-1}};\lambda,L,\theta=1\right)$.