Design of side-sensitive double sampling control schemes for monitoring the location parameter

by

COLLEN MABILUBILU MOTSEPA

Submitted in accordance with the requirements
for the degree of

MASTER OF SCIENCE

in the subject

STATISTICS

at the

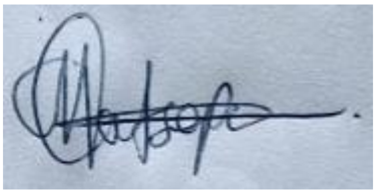
UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: Dr. M J C Malela

December 2019

Declaration

I, Collen Mabilubilu Motsepa, student number 61515426, declare that the thesis entitled: "**Design of side-sensitive double sampling control schemes for monitoring the process location parameter**", which I hereby submit for the degree Master in Statistics at the University of South Africa (UNISA), is my own work that has not previously been submitted by me for a degree at this or any other tertiary institution, and all sources used or quoted in the study have been indicated and acknowledged by means of complete references.

Signature: 

Date: **30 November 2019**

Acknowledgements

I would like to sincerely thank my supervisor, Dr. Jean-Claude Malela-Majika, for his constant support, guidance, efforts and sacrifices for the sake of a successful MSc research project; at University of South Africa's Department of Statistics. I am really grateful for your time, contributions and input in this MSc research project.

My many thanks also go to the University of South Africa's Department of Statistics as well as the Masters & Doctorate (MD) bursary programme. Without their financial assistance, I would not have been able to conduct my research.

My journey to conducting this MSc research could not be easy without the support and advices from my friends and colleagues from the Department of Statistics at the University of South Africa, more specially Mr. Sandile Charles Shongwe. More importantly, I am thankful to Prof Philippe Castagliola from the University of Nantes in France (Université de Nantes, LS2N UMR CNRS 6004, France), who spent his valuable time and gave us constructive comments to improve the content of this project.

With the latter being stated, I would like to express that any theoretical, empirical or grammatical errors are to be ascribed to me and me alone.

Finally, I would like to thank my family and friends who were very supportive throughout the tenure of this research project.

Abstract

Double sampling procedure is adapted from a statistical branch called acceptance sampling. The first Shewhart-type double sampling monitoring scheme was introduced in the statistical process monitoring (SPM) field in 1974. The double sampling monitoring scheme has been proven to effectively decrease the sampling effort and, at the same time, to decrease the time to detect potential out-of-control situations when monitoring the location, variability, joint location and variability using univariate or multivariate techniques. Consequently, an overview is conducted to give a full account of all 76 publications on double sampling monitoring schemes that exist in the SPM literature. Moreover, in the review conducted here, these are categorized and summarized so that any research gaps in the SPM literature can easily be identified. Next, based on the knowledge gained from the literature review about the existing designs for monitoring the process mean, a new type of double sampling design is proposed. The new charting region design lead to a class of a control charts called a side-sensitive double sampling (SSDS) monitoring schemes. In this study, the SSDS scheme is implemented to monitor the process mean when the underlying process parameters are known as well as when they are unknown. A variety of run-length properties (i.e., the 5th, 25th, 50th, 75th, 95th percentiles, the average run-length (ARL), standard deviation of the run-length ($SDRL$), the average sample size (ASS) and the average extra quadratic loss ($AEQL$) metrics) are used to design and implement the new SSDS scheme. Comparisons with other established monitoring schemes (when parameters are known and unknown) indicate that the proposed SSDS scheme has a better overall performance. Illustrative examples are also given to facilitate the real-life implementation of the proposed SSDS schemes. Finally, a list of possible future research ideas is given with hope that this will stimulate more future research on simple as well as complex double sampling schemes (especially using the newly proposed SSDS design) for monitoring a variety of quality characteristics in the future.

Keywords: Double sampling, Monitoring scheme, Statistical process monitoring, Run-length properties, Overall performance measures, Side sensitive, Non-side-sensitive, Estimated parameters, Phase I, Phase II.

List of research outputs

The research outputs associated with this dissertation are listed below. This includes peer-reviewed published article, articles under review and, an upcoming conference where parts of this dissertation will be presented.

- ***Published or Accepted***

1. Malela-Majika J.-C., Motsepa C.M. and Graham M.A. (2019). A new double sampling $\bar{X}$ control chart for monitoring an abrupt change in the process location. *Communications in Statistics – Simulation and Computation*. **DOI: 10.1080/03610918.2019.1577970**.

2. Motsepa C.M., Malela-Majika J.-C., Castagliola P. and Shongwe S.C. (2020). A side-sensitive double sampling $\bar{X}$ monitoring scheme with estimated process parameters. *Communications in Statistics – Simulation and Computation*. **DOI: 10.1080 /03610918. 2020.1722835**

- ***Under Review***

3. Motsepa C.M., Malela-Majika J.-C., Castagliola P. and Shongwe S.C. (2020). Double sampling monitoring schemes: A literature review and some future research ideas. *European Journal of Industrial Engineering.* **Submitted: 15 November 2019** (Manuscript ID: EJIE-33792).

- ***Conference Presentation***

4. Motsepa C.M. (2019). A new double sampling control chart for monitoring an abrupt change in the process location. *The 61st Annual Conference of the South African Statistical Association* (SASA), hosted by Nelson Mandela Metropolitan University (NMMU), Port Elizabeth, 27-29 November 2019.

- ***Departmental Seminar Presentation***

5. Motsepa C.M. (2020). A side-sensitive double sampling $\bar{X}$ monitoring scheme with estimated process parameters. *Departmental Seminar*, Department of Statistics, University of South Africa, Florida Campus. **Presented: 27 February 2020**.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

The quality of products and services is one of the most important factors in the growth and good run of modern organizations. Thus, statistical process monitoring (or statistical process control) has become an important business strategy used to reduce the costs and improve the quality of products (or goods) and services; and therefore, maximize the organization's profits. According to Montgomery (2013), statistical process monitoring (SPM) is a field of quality control that uses statistical tools (or methods) to monitor and control industrial and non-industrial processes. It helps to ensure that a process is stable and operates efficiently to avoid waste and reworks (or scrap). SPM uses seven tools (or techniques) known as magnificent seven: (i) histogram, (ii) flow chart, (iii) scatter diagram, (iv) cause-and-effect diagram, (v) check sheet, (vi) Pareto chart and, (vii) control chart. The latter is the most popular and efficient tool used in SPM. Therefore, this study focuses on the design, implementation and enhancement of control charts to monitor the process location.

1.1 Basic concepts of statistical process monitoring

In this section, we provide a number of background information required to understand the essence of this work. These include, the definition of a monitoring scheme and its originality, important terminologies, different types of monitoring schemes, methods of computing the run-length properties, difference between: univariate and multivariate schemes, parametric and nonparametric schemes, parameters known and unknown, Phase I and Phase II; and finally, define the concept of double sampling (from a general statistical point of view).

1.1.1 Monitoring scheme (or control chart)

The basic theory of SPM was developed in the late 1920's by Dr. W. Shewhart – see for instance the introduction sections of Montgomery (2013), Qiu (2014), Chakraborti and Graham (2019). Montgomery (2013) stated that SPM is an application of a collection of statistical techniques, which allows high quality products to be produced. Among the different tools used in SPM applications, monitoring schemes are the most widely used for identifying changes in processes. One of the main purposes of the control charts is to distinguish between assignable and common causes of variation. The process that works only in the presence of common causes of variation is said to be statistically in-control (IC). When a given process runs in a presence of assignable causes of variation then it is said to be out-of-control (OOC). In this case, an OOC signal is given. This OOC signal can also be classified as a false alarm which is a wrong warning about the presence of assignable causes, when in fact the process is IC. If the issued signal from the process is not a false alarm, corrective action(s) should be implemented to eliminate the assignable cause(s) and, consequently, return the process to the IC state.

A typically monitoring scheme is a two dimensional graphic (or line graph – which can be two- or three-dimensional) consisting of the values of a plotting (i.e. charting) statistic plotted on the vertical axis against time or subgroup number on the horizontal axis along with the associated control limits. That is, the main purpose of a monitoring scheme is to continually monitor a given process by illustrating its behaviour. The charting statistic and the control limits are calculated from the data, which can be individuals or subgroup samples of observations, collected sequentially over time. For instance, a two-sided monitoring scheme showing eleven consecutive sample points with the plotting statistic (indicating some specific quality characteristic of interest, i.e. mean, standard deviation, variance, coefficient of variation, etc.) is shown in Figure 1.1.

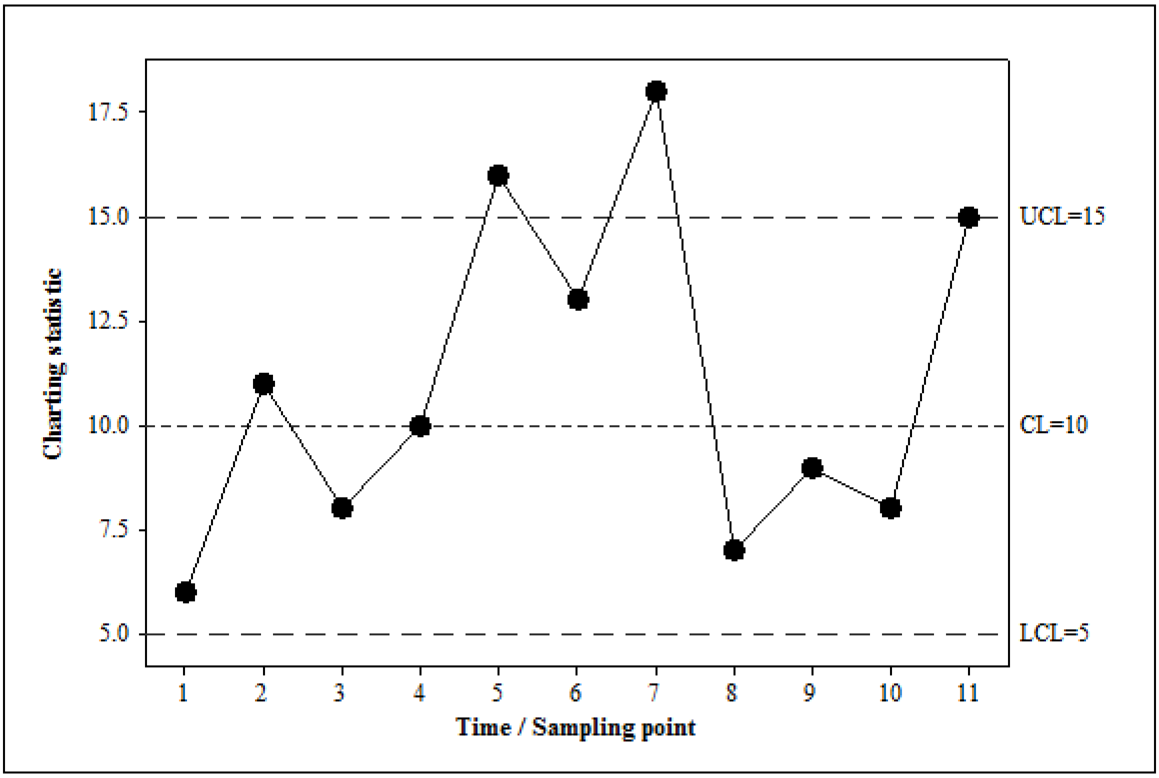


Figure 1.1: A typical Shewhart-type monitoring scheme

From Figure 1.1, it can be seen that a monitoring scheme usually has a center line (*CL*) and two horizontal lines, one on each side of the *CL*. The line above the *CL* is called the upper control limit (*UCL*) whereas the line below the *CL* is called the lower control limit (*LCL*). These three lines are placed on the monitoring scheme to aid the user in making an informed and objective decision whether a process is IC or OOC. For a basic Shewhart-type monitoring scheme, when a charting statistic plots on or outside either of the control limits it is said that a signal has been observed and the process is declared OOC. For instance, assume that

$$\{X_{tj}: t \geq 1; j = 1, 2, \ldots, n\} \tag{1.1}$$

is a sequence of samples from an independent and identically distributed (iid) $N(\mu_0, \sigma_0^2)$ distribution; where μ_0 and σ_0^2 are the specified IC mean and variance, respectively. Let

$$\bar{X}_t = \frac{1}{n}\sum_{j=1}^{n} X_{tj} \tag{1.2}$$

denote the plotting statistic at sampling point *t*. A basic monitoring scheme that is usually used to monitor $\bar{X}_t$ is called the Shewhart $\bar{X}$ scheme and it signals when a single plotting statistic (i.e. sample mean) falls above the *UCL* or below the *LCL* which are given by the following *k*-sigma limits:

$$\begin{aligned} UCL &= \mu_0 + k\frac{\sigma_0}{\sqrt{n}}, \\ CL &= \mu_0, \\ LCL &= \mu_0 - k\frac{\sigma_0}{\sqrt{n}}, \end{aligned} \tag{1.3}$$

where *k* is the distance of the control limits from the *CL*; see for instance Figure 1.1. In Figure 1.1, a process would be thought to be OOC at time points 5, 7 and 15. The corresponding event is called a signalling event. On the contrary, when the charting statistic randomly plots between the *UCL* and *LCL*, the process is thought to be IC and hence, no signal is observed on the monitoring scheme. The corresponding event is called a non-signalling event.

1.1.2 Different types of monitoring schemes

In the literature, there are many different types of monitoring schemes – each depending on what type of data is at hand. Data can be continuous or discrete. Quality characteristics that can be expressed in terms of a numerical measurement are called “variables” and the data collected on variables are called “variables data”, see Montgomery (2013, p. 234). Examples include dimensions such as length or width, temperature, volume, etc. However, quality characteristics that cannot be measured on a numerical scale, for example, the quality of paint on a glass container for a liquid product, are called “attributes” and the corresponding data collected are called “attributes data”; see Montgomery (2013, p. 297). To examine attributes data, we classify them into one of the two categories called conforming and nonconforming, depending on whether the container meets the requirements on one or more quality characteristics. Examples include the number of errors or mistakes made in completing a loan application, the number of medical errors made in a hospital, etc.; see Montgomery (2013, p. 297).

Note that, in general, monitoring schemes differ by the plotting statistic. For example, an $\bar{X}$ scheme is for the mean or average; *p* and *np* schemes are for monitoring the fraction and number of conforming items in a sample; *t* and t_r schemes are for the time between events; etc. and these are based on some assumed distribution, i.e. normal, binomial, exponential or gamma, etc. There are other schemes to monitor special processes, i.e. short-run, profiles, start-up, time series, rare events, etc., – see Qiu (2014). These schemes have different methods to calculate charting statistics and control limits.

Although there are many different monitoring schemes, most of these may be classified under the following two popular types of control charting techniques: the Shewhart schemes and memory-type schemes (e.g., the exponentially weighted moving average (EWMA), Cumulative Sums (CUSUM), etc.). Relative advantages and disadvantages of these schemes are well documented in the literature; see for example, Chapters 6, 7 and 9 in Montgomery (2013). Shewhart-type schemes are the most popular schemes in practice because of their simplicity, ease of application, and the fact that these versatile schemes are quite efficient in detecting moderate to large shifts; in addition, they are closely associated with the theory of hypothesis testing. The CUSUM scheme was developed by Page (1954) using Wald’s sequential testing theory. A comprehensive description of the construction of CUSUM scheme is discussed in Hawkins and Olwell (1998). The EWMA scheme assigns a larger weight to the most recent observations and was developed by Roberts (1959).

1.1.3 Run-length characteristics

Note that the number of rational subgroups to be collected or the number of charting statistics to be plotted on a monitoring scheme before the first OOC signal is observed is the run-length of a scheme, see Montgomery (2013). The run-length is a random variable, denoted usually by *RL*, with a mean and variance. The most widely used monitoring scheme’s performance metric is the mean of the run-length, referred to as the average run-length (*ARL*). However, since the run-length distribution is significantly right-skewed, researchers have advocated using other, more representative, measures for the assessment of chart performance. These include the standard deviation of the run-length (*SDRL*) and percentiles of the run-length (PRL), more specifically, the median run-length (*MRL*), which provides additional and more meaningful information about the IC and OOC performances of a monitoring scheme, not given by the *ARL*, which may be computed using the cumulative distribution function (c.d.f.). Some researchers such as Gan (1994), Chakraborti (2006) and Khoo et al. (2011a) have advocated the use of percentiles (i.e., such as the median, 5th, 25th, 75th and 95th) for assessment of a chart OOC performance. The run-length distribution and the characteristics of the run-length distribution can be obtained using four methods, namely

i. The exact approach
ii. The Markov chain approach
iii. The integral equation approach
iv. The Monte Carlo simulations approach

These approaches are used to evaluate the run-length distribution and the characteristics of the run-length distribution of various types of control charts. In this dissertation, we use the first, third and the fourth methods listed above.

1.1.4 Univariate and multivariate schemes

Since the process in Equation (1.1) generates observations used to monitor a single quality characteristic in Equation (1.2), such monitoring schemes are termed, *univariate*. However, instead of Equation (1.1), when observations are generated by

$$\{\boldsymbol{X}_{tj} = \begin{bmatrix} X_{1tj} \\ X_{2tj} \\ \vdots \\ X_{ptj} \end{bmatrix} : t \geq 1;\, j = 1, 2, \ldots, n\} \tag{1.4}$$

i.e., a p-variate normally distributed random variable with mean $\boldsymbol{\mu_0}$ and variance $\boldsymbol{\Sigma}_0$, see Bersimis et al. (2007); such monitoring schemes are termed *multivariate*. The charting statistic in Equation (1.4) is used to monitor p different quality characteristics. Note that $\boldsymbol{\mu_0}$ and $\boldsymbol{\Sigma}_0$ denote a $(p \times 1)$ vector of sample means and a $(p \times p)$ covariance matrix, respectively. Thus, the corresponding charting statistic is computed as (instead of Equation (1.2), the Hotelling's T^2 statistic),

$$T^2 = n(\overline{\boldsymbol{X}} - \boldsymbol{\mu_0})'\boldsymbol{\Sigma}_0^{-1}(\overline{\boldsymbol{X}} - \boldsymbol{\mu_0}) \tag{1.5}$$

which has a central chi-square distribution with p degrees of freedom, where $\overline{X}$ refers to the $(p \times 1)$ vector of sample means, with each of the p entries computed using Equation (1.2). However, when the process is thought to be OOC, it has non-central chi-square distribution with p degrees of freedom, with the non-central parameter given by $\theta^2 = n(\boldsymbol{\mu_1} - \boldsymbol{\mu_0})'\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\mu_1} - \boldsymbol{\mu_0})$. For instance, after the occurrence of assignable causes, the mean shifts as follows,

- Univariate case: from μ_0 to μ_1 ($\mu_1 = \mu_0 + \delta\sigma_0$), so that $\delta = \frac{\mu_1 - \mu_0}{\sigma_0}$;
- Multivariate case: from $\boldsymbol{\mu_0}$ to $\boldsymbol{\mu_1}$, so that $\boldsymbol{\delta} = \sqrt{(\boldsymbol{\mu_1} - \boldsymbol{\mu_0})'\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\mu_1} - \boldsymbol{\mu_0})}$.

1.1.5 Parametric and nonparametric schemes

The key difference between parametric and nonparametric schemes is that the parametric schemes have some assumed underlying distribution that the process follows. However, if it is different from the assumed distribution, then the corresponding performance may be adversely affected. According to Chakraborti and Graham (2019), a monitoring scheme is called nonparametric if its IC run-length distribution is the same for every continuous distribution. All the publications that are of interest in this dissertation are parametric schemes, except for two publications by Yang and Wu (2017a, b). The main advantage of nonparametric monitoring schemes is the nonparametric statistical methods, which are typically based on order statistics, ranks and various functions of them. That is, the corresponding charting statistics are based on distribution-free test statistics, meaning that their IC *RL* distributions do not depend on any specific underlying distributions. Thus, these monitoring schemes are IC robust. A monitoring scheme is said to be IC robust if the IC characteristics of the run-length distribution are the same over all continuous distributions, see for instance Qiu (2014), Chakraborti and Graham (2019), Mabude et al. (2019), etc. Note though, parametric monitoring schemes are not IC robust. That is, when a gamma distribution is used instead of a normal distribution when the process is IC, the $\bar{X}$ scheme will yield different IC run-length properties under the gamma and normal distributions.

1.1.6 Phase I and Phase II

When the underlying process parameters are unknown, a monitoring scheme needs to be implemented in a two-phase approach, i.e. Phase I and Phase II (see Jensen et al. (2006) and Psarakis et al. (2013) for a review of parameter estimation effect articles). In Phase I, monitoring schemes are implemented retrospectively in order to estimate the distribution parameters and determine the control limits using an IC reference sample. Note that all administrative tasks are planned and executed in Phase I. However, in Phase II, monitoring schemes are implemented prospectively to continuously monitor any departures from an IC state using the parameters estimated in Phase I.

These concepts are further discussed in Chapters 2 and 4, in detail.

1.1.7 Enhancement of monitoring schemes

Different techniques have been proposed in the SPM literature to improve the statistical performance of some basic monitoring schemes. To list just a few, the SPM literature recommends the use of supplementary runs-rules, adaptive approaches (e.g., variable sampling size and interval), the combination of different monitoring schemes (e.g. synthetic schemes), the use of double sampling procedure, etc. The type of techniques is selected depending on the objectives of the investigation. For instance, to improve the traditional basic Shewhart scheme in monitoring small and moderate shifts in the process parameters, the literature recommends the use of double sampling scheme and the combination of the Shewhart and the memory-type scheme such as the EWMA, CUSUM, etc. To improve the memory-type scheme in monitoring large shifts in the process, the literature recommends either the combination of Shewhart and memory-type schemes or the use of supplementary runs-rules. To improve the EWMA scheme in monitoring very small shifts, the extended EWMA schemes such as the double EWMA (DEWMA), Hybrid EWMA (HEWMA), generally weighted moving average (GWMA) and double GWMA (DGWMA) schemes are recommended. Moreover, the side-sensitive designs can also be used to improve the performance of the basic and adaptive monitoring schemes as well as those supplemented with runs-rules in monitoring small to large shifts in the process.

This study will focus on the design of side-sensitive double sampling monitoring scheme in order to improve the sensitivity of the existing double sampling schemes. The concept of double sampling is defined in the next sub-section.

1.1.8 Concept of double sampling

The main focus of this dissertation is on double sampling. Double sampling plan or scheme is not uniquely used in SPM (which is one of the branches of statistics also known as ‘Industrial Statistics’) but is also used in a variety of areas in statistics (see for example the review paper by Rao (2005)). More specifically, double sampling is “borrowed” from the area in statistics called ‘acceptance sampling’. Double sampling design is defined as a procedure in which a master sample is split into two separate homogeneous groups, where the first sample is used in the first stage and the second sample is used in the second stage. Note though, the latter definition has been revised in the literature, and double sampling is now defined in a slightly different manner, and this is the main focus of chapters 2, 3 and 4.

1.2 Problem statement

Researchers and quality practitioners are looking for more sensitive, efficient and accessible monitoring schemes that can detect any change (or shift) in the production or manufacturing process as quickly as possible. However, most of the schemes have the difficulty to detect all range of shifts in the process soon after they have occurred, particularly smaller and moderate shifts. Montgomery (2013) showed that the detection ability of a monitoring scheme decreases when the size of a shift decreases. Montgomery (2013) also showed that the basic Shewhart monitoring schemes are more efficient in unmasking large shifts in the quality process and relatively less efficient in unmasking small and moderate shifts. To increase the sensitivity of monitoring schemes in unmasking small and moderate shifts, the CUSUM and EWMA monitoring schemes were proposed for such purpose. The CUSUM and EWMA monitoring schemes were found to be more efficient in detecting small and moderate changes in the process. However, they also have limitations because they are not as effective as the Shewhart scheme in unmasking large shifts. Hence, considerable efforts have been taken to improve their performance.

Many researchers have hugely contributed to the development of new control schemes in order to increase their sensitivity in detecting small and moderate shifts. Some authors such as Croasdale (1974) and Daudin (1992) have developed double sampling schemes as an attempt to improve the efficiency of the traditional Shewhart scheme in monitoring small and moderate changes in the process mean. Other researchers combined different schemes in order to improve their efficiency, Shewhart charts were combined with CUSUM and EWMA in order to increase their sensitivity for being able to detect moderate and small process mean shifts. Primarily, when a control chart is improved in order to detect a particular type of shifts, the process affects the detection capability for other range of shifts. Therefore, an efficient and most powerful procedure would improve the detection capability irrespective of the magnitude (or size) of the shifts; or improve the detection capability for a specific type of shifts and keep its effectiveness for other shifts.

In this study, a new double sampling scheme for monitoring the process location parameter, namely, the side-sensitive double sampling (SSDS) $\bar{X}$ scheme when the process underlying distribution parameters are known and when they are estimated from some IC historical data. The proposed charts are expected to perform more efficiently than the classical Shewhart scheme and the existing double sampling monitoring schemes, irrespective of the size of the location shifts.

1.3 Research objectives

As mentioned earlier, the traditional Shewhart $\bar{X}$ scheme is less efficient in detecting (or unmasking) small and moderate shifts. Therefore, it is imperative to find alternative schemes to solve this problem. Many schemes have been developed in the past as an attempt to find an alternative of Shewhart $\bar{X}$ scheme. The double sampling $\bar{X}$ scheme is one of the alternatives; therefore, this research is willing to assist in improving the existing double sampling $\bar{X}$ scheme by proposing a new design for the charting regions that will be shown to have a better detection ability than the existing double sampling design. The results from this study can be used as a baseline for developing and designing other schemes for monitoring the median, standard deviation, range, coefficient of variation, etc. In addition, this study is expected to promote the use of SPM approaches in manufacturing, public and private sectors and in all other types of organizations (or institutions) such as academic, engineering, health, education, finance, transportation, etc.

1.4 Scope and limitation of the study

In this dissertation, there are 5 main chapters. Firstly, in this introductory chapter, the main objectives of the whole dissertation are given and the important SPM concepts discussed throughout are defined so that the reader gets a better understanding of the research done in the upcoming chapters. In chapter 2, all the existing SPM literature on double sampling, exactly 76 publications, are reviewed. A new design of double sampling schemes for monitoring abrupt shifts in the process mean when the distributional parameters are known and unknown are proposed in chapters 3 and 4, respectively. In chapter 5, the concluding remarks and possible future research ideas are given. Finally, the Appendices (i.e., A, B, C) explain how the expressions for calculating run-length properties are formulated in MATHCAD®14 software, and how they are derived for the new side-sensitive design.

In this dissertation, the focus is on the univariate parametric double sampling schemes for monitoring the process location. The multivariate perspective of the chart and the monitoring of the variability will be discussed in the future. The scheme is based on some i.i.d. underlying parametric normal distribution. Note though, there are some discussions on other publications that utilized other different distributions and those that are non-i.i.d. within the literature review in Chapter 2.

1.5 Outline of the dissertation

In this section, we give a brief outline of the research done in Chapters 2, 3, 4 and 5.

1.5.1 Research done in Chapter 2

Given the pioneering works by Crosier (1974) and Daudin (1992); double sampling procedures in SPM applications have shown some significant increase in outputs very recently, hence all the publications on this topic are reviewed. To ensure a proper easy to follow discussion, the literature on double sampling is discussed in terms of classes i.e. univariate and multivariate schemes. Based on these two classes, then the discussion is done in terms of double sampling schemes for monitoring the location, variability, as well as both the location and variability simultaneously. Based on some tables constructed to show what has been done in the literature, the key missing gaps are easily predictable.

1.5.2 Research done in Chapter 3

A new design for the double sampling scheme to monitor the process mean when the underlying process distribution (i.e., normal distribution) and parameters are specified is proposed. Using a number of run-length properties described in Section 1.1.3, it is shown that the new design yields a better OOC performance than that of the most used double sampling scheme by Daudin (1992) and a variety of other well-known monitoring schemes to monitor the process mean. Real-life examples are used to illustrate its implementation. The illustration of how the run-length formulae were programmed in MATHCAD® is shown in Appendix A and the derivation of the theoretical run-length formulas are shown in Appendix B.

1.5.3 Research done in Chapter 4

Given that the research done in Chapter 3 deals with the underlying parameters assumed known, then in Chapter 4, the effect of estimating the unknown parameters on the Phase II run-length performance is conducted. More importantly, the run-length properties of the new design for the double sampling scheme are derived and evaluated empirically when the process parameters are assumed unknown. The aim of conducting parameters unknown case is to study the corresponding run-length properties and compare its performance to that of parameters known done in Chapter 3. That is, the intention is to measure the difference in performance when parameters are known versus parameters unknown. As in other different monitoring schemes, researchers like Jones et al. (2004) observed that the run-length properties are significantly different when parameters are known versus parameters unknown. Consequently, the parameters unknown case is more real-life oriented than that in Chapter 3 (i.e. parameters known). Finally, the effect of the Phase I sample size on the performance of Phase II SSDS monitoring scheme in terms of the unconditional *ARL* metric when the parameters known optimal design parameters are used instead of the parameters unknown ones. A real-life example is used to illustrate its implementation. The derivation of the theoretical run-length formulas is shown in Appendix C.

1.5.4 Research done in Chapter 5

Here, the key findings from all the chapters in this dissertation are summarized. That is, the concluding remarks and possible future research ideas are given in this chapter.

Chapter 2. Double sampling schemes: A review and some future research ideas

2.0 Overview

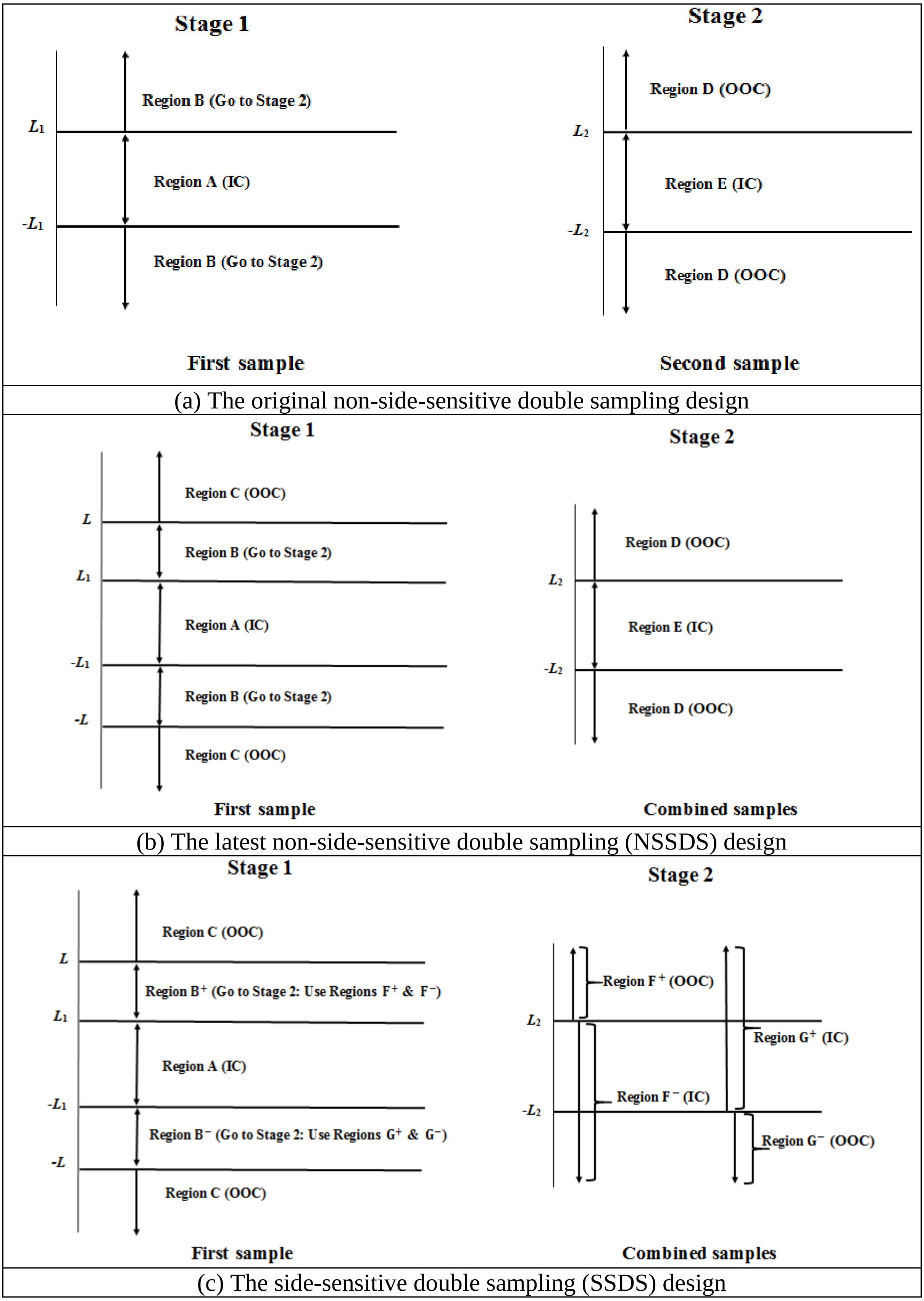
In this chapter, the literature review on double sampling monitoring schemes is discussed. All existing articles on double sampling monitoring schemes were considered without any exclusion criterion. This chapter is divided into six sections organised as follows: in Section 2.1, an introductory discussion presenting a detailed summary of all the publications on double sampling monitoring schemes is given. Section 2.2 discusses some important run-length properties that are used in the SPM literature to evaluate the performance of the double sampling schemes. In Section 2.3, the univariate double sampling monitoring schemes are discussed; whereas the multivariate ones are discussed in Section 2.4. Other different monitoring schemes or procedures that are integrated with the operation of the double sampling scheme are discussed in Section 2.5. Finally, some concluding remarks and future research ideas are given in Section 2.6.

2.1 Introduction

A double sampling monitoring strategy is one of the most powerful tools used in SPM to detect unexpected changes in various types of processes (such as business, health, manufacturing, etc.) as quickly as possible. Double sampling monitoring schemes implement a two-stage monitoring procedure to decide whether the process being monitored is IC or OOC. Croasdale (1974) adapted the idea of double sampling procedure from the acceptance sampling field and implemented its use in the SPM field. Croasdale (1974)'s method entails the use of a sample of size n_1 in stage 1 and of size n_2 in stage 2 to compute the corresponding charting statistics; both sub-samples from the same master sample of size $n\ (= n_1 + n_2)$, where $n_2 > n_1$. Consequently, as an improvement to Croasdale's method, Daudin (1992) showed that the use of the sample size n_1 in stage 1 and both n_1 and n_2 (i.e., $n_1 + n_2$) in stage 2 yields an even more improved performance. Based on the latter, the vast majority of discussions on double sampling schemes done post-1992 were more focused on the method by Daudin (1992) rather than the original version by Croasdale (1974).

There are three main different designs of Shewhart-type double sampling schemes charting regions, which are defined as: (i) original non-side-sensitive, (ii) improved non-side-sensitive and, (iii) side-sensitive. The first non-side-sensitive double sampling scheme is a two-stage scheme based on two *unconnected* samples (i.e. the first sample of size n_1 in stage 1 and the second sample of size n_2 in stage 2). It has only been discussed in 5 research works (see Figure 2.1(a) for its charting regions) – and it was first proposed in Croasdale (1974). The second non-side-sensitive double sampling scheme (by Daudin, 1992) is the most used design by slightly over 90% of publications on this topic; see Figure 2.1(b) for its charting regions – henceforth denoted by NSSDS. Unlike Croasdale's design, the Daudin's design is a two-stage scheme based on two *connected* samples (i.e. the first sample of size n_1 in stage 1 and the second combined sample of size $n_1 + n_2$ in stage 2). The third one is called the side-sensitive double sampling with its charting regions given in Figure 2.1(c) – henceforth denoted by SSDS – this is proposed in Chapter 3 of this dissertation or, in Malela-Majika et al. (2019). The SSDS scheme is based on two *connected* samples. It is important to note from Figure 2.1 that the Croasdale (1974) charting regions imply that a monitoring process never goes to a state of OOC in stage 1, but it only does in stage 2. However, the charting regions in Figures 2.1(b) and (c) do allow for an OOC signal to take place in stage 1, making it more efficient. Note that the SSDS scheme contains work that is fully discussed in Chapters 3 and 4 of this dissertation.

While the majority of the double sampling schemes are focused on the monitoring of the process location parameter, there is a variety of other different parameters that can be monitored by these schemes, e.g. the standard deviation, variance, range, coefficient of variation, etc. All existing publications from 1974 up to February 2020 that we could find in the SPM literature are summarized in Table 2.1. The corresponding journals or conference proceedings that published these ones are outlined in Table 2.2. Next, different authors that have made a contribution of at least two publications in this area of research are listed in Table 2.3 along with their respective affiliations and the number of publications each published. It is observed from Tables 2.2 and 2.3 that *Quality and Reliability Engineering International* journal has the most publications on double sampling schemes, and that Prof M.B.C. Khoo (from Universiti Sains Malaysia) significantly has the most publications than any other author / researcher.



(a) The original non-side-sensitive double sampling design

(b) The latest non-side-sensitive double sampling (NSSDS) design

(c) The side-sensitive double sampling (SSDS) design

Figure 2.1: The charting regions in stages 1 and 2 of the different double sampling designs

Table 2.1: Classification of articles discussing double sampling schemes in SPM (sorted chronologically)

	Quality characteristic				Parameter(s)		Type of Data		Process Type	
Paper	Location	Variability	Location & Variability	Econ and/or Econ-Stat design	Known	Unknown	Univariate	Multivariate	I.I.D.	Serial Dependence
Croasdale (1974)	✓				✓		✓		✓	
Daudin et al. (1990)	✓				✓		✓		✓	
Daudin (1992)	✓				✓		✓		✓	
Irianto and Shinozaki (1998)	✓				✓		✓		✓	
Carot et al. (2002)	✓				✓		✓		✓	
He et al. (2002)	✓				✓		✓		✓	
He and Grigoryan (2002)		✓			✓		✓		✓	
He and Grigoryan (2003)		✓			✓		✓		✓	
Hsu (2004)	✓				✓		✓		✓	
Khoo (2004)		✓			✓		✓		✓	
Grigoryan and He (2005)		✓			✓			✓	✓	
He and Grigoryan (2006)			✓		✓		✓		✓	
Hsu (2007)		✓			✓		✓		✓	
Claro et al. (2008)	✓				✓		✓			✓
Costa and Claro (2008)	✓				✓		✓			✓
Champ and Aparisi (2008)	✓				✓			✓	✓	
Costa and Machado (2008)	✓				✓			✓	✓	
Machado and Costa (2008)		✓			✓			✓	✓	
Torng et al. (2009a)	✓			✓	✓		✓		✓	
Torng et al. (2009b)	✓			✓	✓		✓			✓
Torng and Lee (2009)	✓				✓		✓		✓	
Lee et al. (2009)	✓				✓		✓			✓
Irianto and Juliani (2010)	✓				✓		✓		✓	
Torng et al. (2010)	✓				✓		✓		✓	
Lee et al. (2010)		✓			✓		✓		✓	
Costa and Machado (2011)	✓				✓		✓			✓
De Araújo Rodrigues et al. (2011)	✓				✓		✓		✓	
Khoo et al. (2011b)	✓				✓		✓		✓	
Faraz et al. (2012)	✓			✓	✓		✓		✓	
Lee et al. (2012a)	✓			✓	✓		✓		✓	
Lee et al. (2012b)		✓			✓		✓		✓	

Table 2.1: (continued)

Paper	Quality characteristic			Econ and/or Econ-Stat design	Parameter(s)		Type of Data		Process Type	
	Location	Variability	Location & Variability		Known	Unknown	Univariate	Multivariate	I.I.D.	Serial Dependence
Khoo et al. (2013a)	✓					✓	✓		✓	
Khoo et al. (2013b)	✓					✓	✓		✓	
Khoo et al. (2013c)	✓				✓			✓	✓	
Teoh et al. (2013)	✓					✓	✓		✓	
Lee (2013)			✓		✓		✓		✓	
Chong et al. (2014)	✓				✓		✓		✓	
Teoh et al. (2014a)	✓				✓		✓		✓	
Teoh et al. (2014b)	✓					✓	✓		✓	
Costa and Machado (2015)	✓				✓		✓		✓	
Noorossana et al. (2015)	✓				✓		✓		✓	
Teoh et al. (2015)	✓					✓	✓		✓	
You et al. (2015)	✓					✓	✓		✓	
Khoo et al. (2015)	✓				✓		✓		✓	
Khoo et al. (2016)	✓				✓		✓		✓	
Lee and Khoo (2016)		✓			✓		✓		✓	
Teoh et al. (2016a)	✓					✓	✓		✓	
Teoh et al. (2016b)	✓					✓	✓		✓	
Aghaulor and Ezekwem (2016)	✓				✓		✓		✓	
Costa (2017)		✓			✓		✓		✓	
You (2017)	✓				✓		✓		✓	
Lee and Khoo (2017a)		✓			✓		✓		✓	
Lee and Khoo (2017b)			✓		✓		✓		✓	
Lee and Khoo (2017c)	✓				✓		✓		✓	
Chong et al. (2017)	✓				✓		✓		✓	
Castagliola et al. (2017)		✓				✓	✓		✓	
Yang and Wu (2017a)	✓				✓		✓		✓	
Yang and Wu (2017b)		✓			✓		✓		✓	
Haq and Khoo (2018)	✓				✓		✓			✓
You (2018)	✓					✓	✓		✓	
Ng et al. (2018)			✓		✓		✓		✓	
Saha et al. (2018)	✓				✓		✓		✓	
Lee and Khoo (2018)		✓			✓			✓	✓	
Khatun et al. (2018)	✓				✓			✓	✓	
Chong et al (2018)	✓				✓		✓		✓	
Malela-Majika and Rapoo (2019)	✓				✓		✓		✓	
Malela-Majika et al. (2019)	✓				✓		✓		✓	
Malela-Majika (2019)			✓		✓		✓		✓	
Lee and Khoo (2019a)	✓			✓	✓		✓		✓	
Lee and Khoo (2019b)	✓			✓	✓			✓	✓	
Lee and Khoo (2019c)	✓					✓	✓		✓	
Haq and Khoo (2019)	✓				✓		✓			✓
Rozi et al. (2019)			✓		✓		✓		✓	
Lee et al. (2019)	✓				✓		✓		✓	
Umar et al. (2019)	✓				✓		✓			✓
Katebi and Moghadam (2020)	✓				✓			✓	✓	
Motsepa et al. (2020)	✓					✓	✓		✓	

Table 2.2: Journals / Conference proceedings with double sampling schemes publications

Journal / Conference proceedings title	Number of publications
Quality and Reliability Engineering International	9
Communications in Statistics – Simulation and Computation	9
International Journal of Production Research	7
Communications in Statistics – Theory and Methods	6
International Journal of Production Economics	5
Computers & Industrial Engineering	4
Journal of Applied Statistics	3
South African Journal of Industrial Engineering	2
European Journal of Operational Research	2
International Journal of Advanced Manufacturing Technology	2
Academic Journal of Science	1
COMPUSOFT, An International Journal of Advanced Computer Technology	1
European Journal of Industrial Engineering	1
Expert Systems with Applications	1
IEEE Access	1
IEEE International Conference on Control and Robotics Engineering	1
IEEE International Conference on Industrial Engineering and Engineering Management	1
IIE Transactions	1
International Conference on Smart Sensors and Application	1
International Journal of Applied Engineering Research	1
International Journal of Engineering Research & Technology	1
International Journal of Industrial Engineering – Theory, Applications & Practice	1
International Journal of Production Development	1
International Journal of Pure and Applied Mathematics	1
IOP Conference Series: Materials Science and Engineering	1
ITB Journal of Engineering Science	1
Journal of Probability and Statistics	1
Journal of Quality Measurement and Analysis	1
Journal of Quality Technology	1
Journal of Statistical Computation and Simulation	1
Journal of Testing and Evaluation	1
MATEC Web of Conferences	1
Pesquisa Operacional	1
PLoS ONE	1
Quality Technology and Quantitative Management	1
Revue de Statistique Appliquée	1
Transactions of the Institute of Measurement and Control	1

Table 2.3: Top researchers in SPM with at least two publications on double sampling schemes

Author	Affiliation	Number of publications
Khoo, M.B.C.	Universiti Sains Malaysia; Malaysia	31
Lee, M.H.	Swinburne University of Technology; Malaysia	12
Teoh, W.L.	Heriot-Watt University Malaysia; Malaysia	12
Castagliola, P.	Université de Nantes & LS2N UMR CNRS 6004; France	10
Lee, P.-H.	National Yunlin University of Science and Technology; Taiwan	9
Torng, C.-C.	National Yunlin University of Science and Technology; Taiwan	8
Costa, A.F.B.	Sao Paulo State University; Brazil	7
Yeong, W.C.	University of Malaya; Malaysia	7
Teh, S.Y.	Universiti Sains Malaysia; Malaysia	6
Chong, Z.L.	Universiti Tunku Abdul Rahman; Malaysia	5
He, D.	University of Illinois; USA	5
Grigoryan, A.	University of Illinois; USA	5
Machado, M.A.G.	Sao Paulo State University; Brazil	5
Malela-Majika, J.-C.	University of South Africa; South Africa	4
Irianto, D.	Institute of Technology Bandung; Indonesia	3
Saha, S.	University of Business Agriculture and Technology, Bangladesh	3
Tseng, C.-C.	National Yunlin University of Science and Technology; Taiwan	3
You, H.W.	Universiti Kebangsaan Malaysia; Malaysia	3
Haq, A.	Quaid-i-Azam University, Pakistan	3
Chakraborti, S.	University of Alabama; USA	2
Claro, F.A.E.	Sao Paulo State University; Brazil	2
Daudin, J.J.	UMR MIA 518 AgroParisTech/INRA; France	2
Hsu, L.F.	City University of New York; USA	2
Lee, H.C.	Universiti Sains Malaysia; Malaysia	2
Liao, H.-S.	National Yunlin University of Science and Technology; Taiwan	2
Liao, N.-Y.	National Yunlin University of Science and Technology; Taiwan	2
Motsepa, C.M.	University of South Africa; South Africa	2
Wu, Z.	Nanyang Technological University; Singapore	2
Wu, S.-H.	National Chengchi University; Taiwan	2
Yang, S.-F.	National Chengchi University; Taiwan	2

2.2 Operation and run-length properties of the double sampling schemes

Remark: from Table 2.1, all the publications on double sampling schemes are based on some i.i.d. (independent and identically distributed) underlying parametric distribution (except for Yang and Wu (2017a, b) and the seven publications that are based on serially correlated data).

2.2.1 Operation of the basic double sampling scheme

Assume that Y_{tj} are some i.i.d. observations of some quality characteristic of interest which are from some specified distribution. From these Y_{tj} observations, a first subgroup sample of size n_1 is collected at the t^{th} sampling time (denoted as Y_{1tj}, $t = 1, 2, \ldots$, and $j = 1, 2, \ldots, n_1$). If the standardized charting statistic based on the first sample falls on a region that requires a second stage to make a decision, then a second subgroup sample of size n_2 (where $n_2 \geq n_1$) is also collected at the t^{th} sampling time (denoted as Y_{2tj}, $t = 1, 2, \ldots,$ and $j = 1, 2, \ldots, n_2$). Then any double sampling scheme uses these two separate sub-samples taken from the same master sample to decide whether the process is IC or OOC, and these sub-samples are used to compute the charting statistics of the two stages shown in Figure 2.1. Since the majority of the double sampling schemes in Table 2.1 are based on the univariate sample mean using the standard normal distribution; hence, for illustration purpose of the stages and the operation, we use the NSSDS $\bar{X}$ scheme when parameters are known and unknown.

- **When parameters are known (i.e. Case K), the stages are implemented as follows:**
  - **Stage 1**

Let $\bar{Y}_{1t} = \sum_{j=1}^{n_1} Y_{1tj}/n_1$ be the mean of the first sample of subgroup size n_1 at the t^{th} sampling time. Hence, the standardized statistic for the first sample at the t^{th} sampling time is then given by

$$Z_{1t} = \frac{\bar{Y}_{1t} - \mu_0}{\sigma_0/\sqrt{n_1}} \tag{2.1}$$

where $\bar{Y}_{1t} \sim N(\mu_0 + \delta\sigma_0, \frac{\sigma_0}{\sqrt{n_1}})$ and $\delta = |\mu_1 - \mu_0|/\sigma_0$ represents the magnitude of the standardized mean shift with the OOC mean μ_1 ($\mu_1 = \mu_0 + \delta\sigma_0$), so that $\delta = 0$ means that the process is IC. In this case, Z_{1t} follows a standard normal distribution (i.e. $Z_{1t} \sim N(0,1)$). However, when $\delta \neq 0$, the process is OOC and $Z_{1t} \sim N(\delta, 1)$.

  - **Stage 2**

At the t^{th} sampling time of the second sample, the sample mean, i.e. $\bar{Y}_{2t} = \sum_{j=1}^{n_2} Y_{2tj}/n_2$, and the combined (or pooled) sample mean, i.e. $\bar{Y}_t = (n_1\bar{Y}_{1t} + n_2\bar{Y}_{2t})/(n_1 + n_2)$ are calculated, respectively. Hence, the standardized charting statistic for the combined samples at the t^{th} sampling time is then given by

$$Z_t = \frac{\bar{Y}_t - \mu_0}{\sigma_0/\sqrt{n_1 + n_2}}, \tag{2.2}$$

where $\bar{Y}_t \sim N(\mu_0 + \delta\sigma_0, \frac{\sigma_0}{\sqrt{n_1+n_2}})$. When the process is IC, $Z_{2t} \sim N(0, 1)$ since $\delta = 0$ and when the process is OOC, $Z_{2t} \sim N(\delta, 1)$.

- **When parameters are unknown (i.e. Case U), the stages are implemented as follows:**

Phase I parameter estimation

Since the IC process parameters, μ_0 and σ_0, are usually unknown they have to be estimated from m Phase I subgroup samples, each of size n, i.e. X_{ij}, $i = 1,2, \dots, m$ and $j = 1,2, \dots, n$. The estimated IC process parameters, $\hat{\mu}_0$ and $\hat{\sigma}_0$, are given by

$$\hat{\mu}_0 = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij} \tag{2.3}$$

and

$$\hat{\sigma}_0 = \sqrt{\frac{1}{m(n-1)}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(X_{ij} - \bar{X}_i\right)^2}, \tag{2.4}$$

respectively, where $\bar{X}_i = \sum_{j=1}^{n} X_{ij}/n$.

Phase II charting statistics and operation procedure: Stages 1 and 2

Let Y_{tj} be the Phase II observations from i.i.d. $N(\mu_1\,, \sigma_0)$, where μ_1 is the OOC mean (i.e. $\mu_1 = \mu_0 + \delta\sigma_0$). In Phase II of the NSSDS $\bar{X}$ monitoring scheme, there are two distinct standardized charting statistics in Case U (i.e. $\hat{Z}_{1t}$ and $\hat{Z}_{2t}$, shown below) used during stages 1 and 2, respectively.

Stage 1: Similarly, as in Case K, $\bar{Y}_{1t} = \sum_{j=1}^{n_1} Y_{1tj}/n_1$; so that

$$\hat{Z}_{1t} = \frac{\bar{Y}_{1t} - \hat{\mu}_0}{\hat{\sigma}_0/\sqrt{n_1}}. \tag{2.5}$$

Stage 2: Similarly, $\bar{Y}_{2t} = \sum_{j=1}^{n_2} Y_{2tj}/n_2$ and $\bar{Y}_t = (n_1\bar{Y}_{1t} + n_2\bar{Y}_{2t})/(n_1 + n_2)$, so that

$$\hat{Z}_t = \frac{\bar{Y}_t - \hat{\mu}_0}{\hat{\sigma}_0/\sqrt{n_1 + n_2}}. \tag{2.6}$$

In essence, Equations (2.1) to (2.2) as well as (2.5) to (2.6), imply that there are two distinct standardized charting statistics used during stages 1 and 2 (if needed), respectively. Consequently, based on abovementioned stages, then it follows that the Phase I and Phase II operational procedure of the NSSDS scheme is as summarized in Figure 2.2. Note that in Case K, only the Phase II portion is relevant.

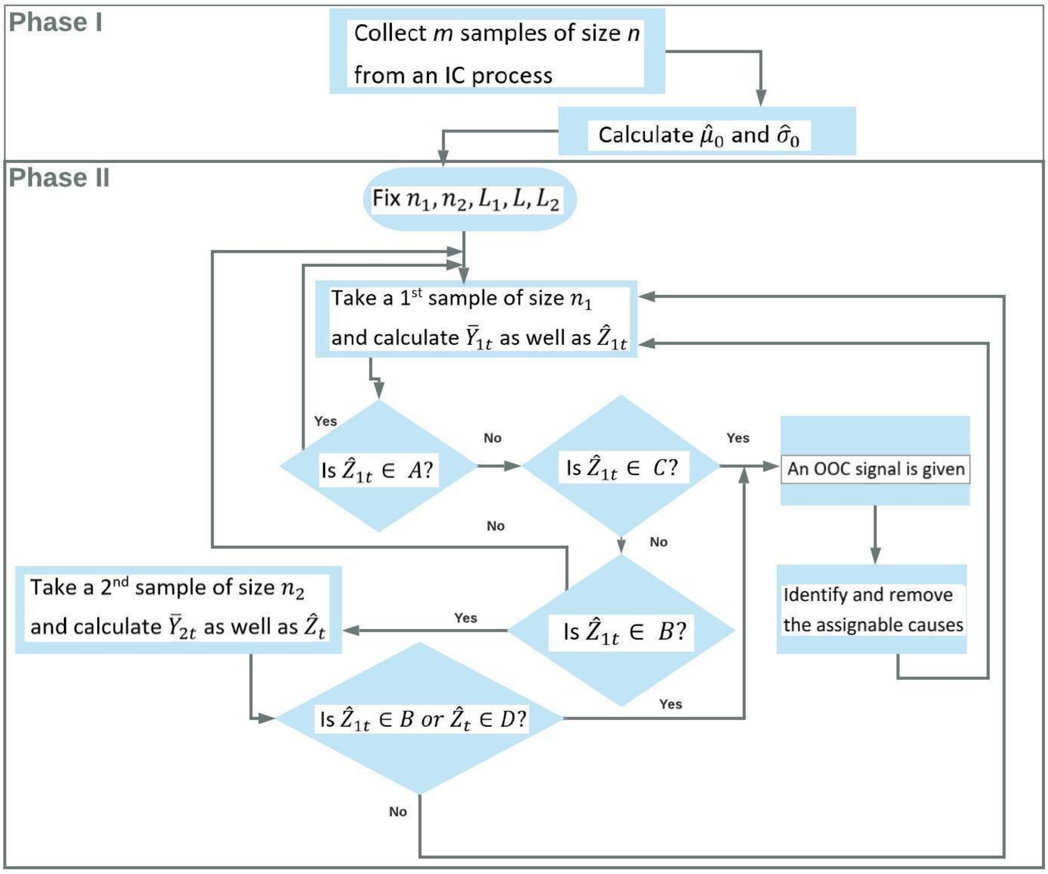


Figure 2.2: Flowchart of the operation of the NSSDS scheme in Phase I and Phase II

Although Figure 2.2 is done for NSSDS $\bar{X}$ scheme, it can easily be modified to account for the different designs outlined in Figure 1 as well as for different charting statistics, i.e. the median, standard deviation, coefficient of variation, etc.

2.2.2 Run-length properties of the NSSDS $\bar{X}$ scheme in Case K

Let P_{0k} represents the probability that the process is regarded as IC at stage k, where k = 1, 2. Then, $P_0 = P_{01} + P_{02}$ is the probability that the process is IC, where:

$$P_{01} = P(Z_{1t} \in \mathrm{A}) = \Phi\left[L_1 + \delta\sqrt{n_1}\right] - \Phi\left[-L_1 + \delta\sqrt{n_1}\right], \tag{2.7}$$

and

$$\begin{aligned} P_{02} &= P[Z_{1t} \in \mathrm{B} \text{ and } Z_{2t} \in \mathrm{E}] \\ &= \int_{Z_{1t} \in \mathrm{I}^{**}} \left\{\Phi\left[cL_2 + rc\delta - z\sqrt{n_1/n_2}\right] - \Phi\left[-cL_2 + rc\delta - z\sqrt{n_1/n_2}\right]\right\} \phi(z)dz \end{aligned} \tag{2.8}$$

where $\Phi(.)$ and $\phi(.)$ are the c.d.f. (cumulative distribution function) and p.d.f. (probability density function) of the standard normal random variable, respectively; $r^2 = n_1 + n_2$, $c = r/\sqrt{n_2}$, and $\mathrm{I}^{**} = [-L + \delta\sqrt{n_1}, -L_1 + \delta\sqrt{n_1}) \cup (L_1 + \delta\sqrt{n_1}, L + \delta\sqrt{n_1}]$. Hence,

$$P_0 = \Phi\left[L_1 + \delta\sqrt{n_1}\right] - \Phi\left[-L_1 + \delta\sqrt{n_1}\right] + \int_{Z_{1t}\in \mathrm{I}^{**}} \left\{\Phi\left[cL_2 + rc\delta - z\sqrt{n_1/n_2}\right] - \Phi\left[-cL_2 + rc\delta - z\sqrt{n_1/n_2}\right]\right\} \phi(z)dz. \tag{2.9}$$

Given that the NSSDS $\bar{X}$ scheme is a Shewhart-type one, its run-length (*RL*) follows a geometric distribution. Therefore, the c.d.f. of the *RL* distribution (denoted $F_{RL}(\ell)$) is obtained as

$$F_{RL}(\ell) = P(RL \leq \ell) = 1 - P_{0,}^{\ell} \tag{2.10}$$

where $\ell \in \{1, 2, 3, ...\}$. Then, the $(100\rho)^{th}$ percentile of the *RL* distribution, ℓ_ρ, is given by

$$P(RL \leq \ell_\rho - 1) \leq \rho \text{ and } P(RL \leq \ell_\rho) > \rho. \tag{2.11}$$

Note that the most used metrics to evaluate the run-length distribution are ρ = 0.05, 0.25, 0.50, 0.75 and 0.95, which denote the 5th, 25th, 50th (or median), 75th and 95th percentiles, respectively. Other well-known *RL* properties are average run-length (*ARL*), standard deviation of the run-length (*SDRL*), average sample size (*ASS*) and average number of observations to signal (*ANOS*) which are given by

$$ARL = \frac{1}{1 - P_0}, \tag{2.12}$$

$$SDRL = \frac{\sqrt{P_0}}{1 - P_0}, \tag{2.13}$$

$$ASS = n_1 + n_2 P_2 \quad \text{and} \quad ANOS = ASS \times ARL, \tag{2.14}$$

respectively, where $P_2 = P(Z_{1t} \in B)$ is the probability of taking the second sample, and it is given by

$$P_2 = \Phi\left(L + \delta\sqrt{n_1}\right) - \Phi\left(L_1 + \delta\sqrt{n_1}\right) + \Phi\left(-L_1 + \delta\sqrt{n_1}\right) - \Phi\left(-L + \delta\sqrt{n_1}\right).$$

Since the *ANOS* depends on the *ASS* and *ARL* values, a larger *ANOS* value implies that either the monitoring scheme is inefficient and/or the cost of inspection is higher. A variety of other *RL* performance measures have been used in the literature, these include the average time to signal (*ATS*), average number of samples to signal (*ANSS*), average number of switches (*ANSW*), standard deviation of time to signal (*SDTS*), standard deviation of number of samples to signal (*SDNSS*), standard deviation of number of switches (*SDNSW*) – these are thoroughly discussed in Noorossana et al. (2015).

2.2.3 Run-length properties of the NSSDS $\bar{X}$ scheme in Case U

In order to calculate the unconditional *RL* properties, we need to first derive the conditional ones, see for instance, You et al. (2015). Hence, the conditional c.d.f. of $\hat{Z}_{1t}$, given $\hat{\mu}_0$ and $\hat{\sigma}_0$ is defined as

$$F_{\hat{Z}_{1t}}(z|\hat{\mu}_0, \hat{\sigma}_0) = \Phi\left(U\sqrt{\frac{n_1}{mn}} + Vz - \delta\sqrt{n_1}\right). \tag{2.15}$$

where $U = (\hat{\mu}_0 - \mu_0)\frac{\sqrt{mn}}{\sigma_0}$ and $V = \hat{\sigma}_0/\sigma_0$. Consequently, the conditional p.d.f. of $\hat{Z}_{1t}$, is given by

$$f_{\hat{Z}_{1t}}(z|\hat{\mu}_0, \hat{\sigma}_0) = V\phi\left(\mathrm{U}\sqrt{\frac{n_1}{mn}} + Vz - \delta\sqrt{n_1}\right). \tag{2.16}$$

Since $\hat{\mu}_0 \sim N(\mu_0, \frac{\sigma_0}{\sqrt{mn}})$, then $U \sim N(0,1)$ so that the p.d.f. of the random variable U is simply,

$$f_U(u) = \phi(u). \tag{2.17}$$

Next, using the fact that $V^2 = (\hat{\sigma}_0/\sigma_0)^2$ has a gamma distribution with parameters $m(n-1)/2$ and $2/[m(n-1)]$, then the p.d.f. of V is defined as

$$f_v(v|m,n) = 2vf_\gamma\left[v^2\middle|\frac{m(n-1)}{2}, \frac{2}{m(n-1)}\right], \tag{2.18}$$

where $f_\gamma(.)$ is the p.d.f. of the gamma distribution with parameters $\frac{m(n-1)}{2}$ and $\frac{2}{m(n-1)}$.

Consequently, to derive the unconditional c.d.f., we need to first derive the unconditional probability of the IC process. Let $\hat{P}_{0k}$ denote the probability that the process with estimated parameters remains IC at the sampling stage k (with $k = \{1, 2\}$). Then, the probability that the process is IC is given by

$$\hat{P}_0 = \hat{P}_{01} + \hat{P}_{02} \tag{2.19}$$

where,

$$\begin{gathered}\hat{P}_{01} = \Phi\left(U\sqrt{\frac{n_1}{mn}} + VL_1 - \delta\sqrt{n_1}\right) - \Phi\left(U\sqrt{\frac{n_1}{mn}} - VL_1 - \delta\sqrt{n_1}\right)\\ \hat{P}_{02} = \int_{Z\in \mathrm{I}^{**}}^{.} \hat{P}_E f_{\hat{Z}_{1t}}(z|\hat{\mu}_0, \hat{\sigma}_0)dz\end{gathered} \tag{2.20}$$

with

$$\hat{P}_E = \Phi\left[U\sqrt{\frac{n_2}{mn}} + V\left(\frac{L_2\sqrt{n_1+n_2} - z\sqrt{n_1}}{\sqrt{n_2}}\right) - \delta\sqrt{n_2}\right] - \Phi\left[U\sqrt{\frac{n_2}{mn}} - V\left(\frac{L_2\sqrt{n_1+n_2} + z\sqrt{n_1}}{\sqrt{n_2}}\right) - \delta\sqrt{n_2}\right].$$

Then, the unconditional c.d.f. of the NSSDS $\bar{X}$ scheme is given by

$$F_{RL}(\ell) = \int_{-\infty}^{+\infty}\int_0^{+\infty}\left(1 - \hat{P}_0^\ell\right)f_U(u)f_V(v)\,dv\,du, \tag{2.21}$$

where $\ell \in \{1, 2, 3, \dots, \}$, $f_U(u)$ and $f_V(v)$ are defined in Equations (2.17) and (2.18), respectively.

Therefore, the unconditional ARL and $SDRL$ of the NSSDS $\bar{X}$ scheme are given by

$$ARL = \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \left(\frac{1}{1 - \hat{P}_0}\right) f_U(u) f_V(v)\, dv\, du \tag{2.22}$$

and

$$SDRL = \left[\int_{-\infty}^{+\infty} \int_{0}^{+\infty} \left(\frac{1 + \hat{P}_0}{1 - \hat{P}_0}\right) f_U(u) f_V(v)\, dv\, du - ARL^2\right]^{1/2}. \tag{2.23}$$

The ASS is given by

$$ASS = \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \left(n_1 + n_2 \hat{P}_2\right) f_U(u) f_V(v)\, dv\, du \tag{2.24}$$

where $\hat{P}_2$ is the probability of taking the second sample, which is given by $\hat{P}_2 = P\left(\hat{Z}_{1t} \in \mathrm{B} \middle| \hat{\mu}_0, \hat{\sigma}_0\right)$, or simply,

$$\begin{aligned} \hat{P}_2 = {} & \Phi\left(U\sqrt{\frac{n_1}{mn}} + VL - \delta\sqrt{n_1}\right) - \Phi\left(U\sqrt{\frac{n_1}{mn}} + VL_1 - \delta\sqrt{n_1}\right) \\ & + \Phi\left(U\sqrt{\frac{n_1}{mn}} - VL_1 - \delta\sqrt{n_1}\right) - \Phi\left(\mathrm{U}\sqrt{\frac{n_1}{mn}} - VL - \delta\sqrt{n_1}\right). \end{aligned} \tag{2.25}$$

Then, the average number of observations to signal ($ANOS$) is given by

$$ANOS = \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \left(n_1 + n_2 \hat{P}_2\right) \left(\frac{1}{1 - \hat{P}_0}\right) f_U(u) f_V(v)\, dv\, du. \tag{2.26}$$

2.2.4 Other run-length properties of the double sampling scheme

Various authors have revealed that if a monitoring scheme is designed based on one specific size of a mean shift, it will perform poorly when the actual size of the shift is significantly different from the assumed size (see You (2017, 2018)). Moreover, since the *ARL* is defined as the average number of samples required before an OOC signal is issued in the process. It is well-known that the *RL* distribution of a monitoring scheme is generally highly right-skewed when parameters are estimated; see for example Jones et al. (2004). Also, the *ARL* is criticized because of its ineffectiveness in assessing the overall performance, especially when the aim of the study is to assess the performance of a monitoring scheme over a range of shifts. Thus, many researchers have recommended the use of a quality loss function (QLF) instead of the *ARL*, *SDRL*, *ANOS*, *ATS*, etc., to assess the performance of a monitoring scheme. A QLF describes the relationship between the shift size and the quality impact. Ryu et al. (2010) observed the uncertainty in δ, and hence, they recommended designing monitoring schemes to rather minimize quality loss, which is measured by a quantity called the expected weighted run-length ($EWRL$) which is given by

$$EWRL = E[w(\delta) \times RL(\delta)] = \int_{\delta_{min}}^{\delta_{max}} \big(w(\delta) \times RL(\delta)\big) \times h(\delta) d\delta \qquad (2.27)$$

where δ follows some p.d.f. with a density function $h(\delta)$ and a range $[\delta_{min}, \delta_{max}]$, where δ_{min} and δ_{max} are the lower and upper bound of the range of δ, $w(\delta)$ is a weight function associated with δ and $RL(\delta)$ is some specific shift run-length metric, e.g., $ARL(\delta)$, $ANOS(\delta)$, $ATS(\delta)$, etc. In the double sampling schemes literature, Equation (2.27) has been utilised by a number of different authors to formulate a bi-objective algorithm to obtain optimal parameter values, see for instance, Chong et al. (2014), Lee and Khoo (2017c), etc. More specifically, You (2017, 2018) used $w(\delta)$=1 to design the NSSDS scheme with the objective of minimizing Equation (2.27) when parameters are known (i.e. $RL(\delta)$ equal to Equation (2.12)) and unknown (i.e. $RL(\delta)$ equal to Equation (2.22)), respectively.

2.3 Univariate double sampling schemes

In this section, the publications discussing basic double sampling schemes for location, variability and, both the mean and variability simultaneously are discussed in sub-sections 2.3.1, 2.3.2 and 2.3.3, respectively. Note that double sampling schemes combined with other monitoring schemes to monitor location, variability, both the mean and variability simultaneously are discussed in Section 2.5.

2.3.1 Location

In an effort to design double sampling schemes for monitoring the mean, many authors have studied the same NSSDS scheme; see for instance Daudin et al. (1990). However, they have designed it by taking into account different design aspects. For example, Irianto and Juliani (2010) outlined the following three design criterions:

(i) Minimize the expected number of sampling and inspections,

(ii) Maximize the OOC detection power (or minimizing the customer risk),

(iii) Minimize the false alarm rate (or minimizing the producer risk).

It worth pointing out that Daudin et al. (1990) and Daudin (1992) method prioritized (i) and (iii), whereas Irianto and Shinozaki (1998) used the charting regions in Figure 2.1(a) as these regions prioritize (ii) without taking into account (i). Irianto and Juliani (2010) formulated a model that takes into account (ii) and (iii), simultaneously. Note though, in Stage 2, Irianto and Shinozaki (1998) as well as Irianto and Juliani (2010) used a sample of size $n_1 + n_2$ instead of just n_2 as done in Croasdale (1974). Next, other publications discussed here, tend to ignore design criterion (iii), by keeping the false alarm rate constant, then prioritizing (i) and (ii), simultaneously.

He et al. (2002) compared the performance of the NSSDS scheme to a triple sampling scheme (i.e. with 3 stages) and they observed that increasing the number of stages improves the detection ability of a monitoring scheme. However, Hsu (2004) raised some valid concerns regarding the manner in which the generic algorithm in He et al. (2002) was designed as it only took into account the ASS only, without using other run-length performance measures.

As outlined in Table 1, seven publications for serially dependent observations using the NSSDS $\bar{X}$ scheme. Costa and Claro (2008) used the autoregressive moving average with order (1,1), whereas Claro et al. (2008) and Costa and Machado (2011) used the first-order autoregressive model; however, Torng et al. (2009a) and Lee et al. (2009) used a correlation model proposed in Yang and Hancock (1990). Finally, Haq and Khoo (2018) designed a NSSDS $\bar{X}$ scheme based on a regression-type estimator of the process mean with an auxiliary variable under some specific conditions of correlation.

Torng and Lee (2009) studied the NSSDS $\bar{X}$ scheme using a variety of *t*- and gamma distributions with different parameters and observed that it is as good as the variable parameter $\bar{X}$ scheme; however, it turns to be much better than the basic $\bar{X}$ scheme in terms of a variety of run-length performance measures.

There have been numerous articles that have investigated the performance of the NSSDS $\bar{X}$ scheme when parameters are unknown, these are: Khoo et al. (2013a, b), Teoh et al. (2013, 2014b, 2015, 2016a, 2016b), You et al. (2015) and You (2018). That is, these latter articles studied the NSSDS scheme in Case U for a variety of design criterion and contexts. The design parameters that are obtained while the process is IC are such that the following performance metrics used in the latter articles (e.g. the unconditional average run-length (ARL), the unconditional median run-length, the unconditional expected ARL ($EARL$), the unconditional average sample size (ASS), the unconditional average number of observations to signal ($ANOS$), etc.) are minimized when the process is in a state of OOC. However, Motsepa et al. (2020) studied SSDS $\bar{X}$ scheme when parameters are unknown (this is part of Chapter 4 of this dissertation), including the effect of Phase I sample size on the Phase II OOC performance.

The economic and economic-statistical design of the NSSDS $\bar{X}$ scheme in Case K have been conducted in an effort to find the optimal set of parameters which minimizes the net sum of all costs involved, so that the scheme can be operated at an economically optimal level by using the classical cost model in Lorenzen and Vance (1986) as well as the sensitivity analysis. The latter was studied by Torng et al. (2009a, b) when observations are serially correlated and i.i.d., respectively.

Until more recently, all the publications on NSSDS $\bar{X}$ schemes have assumed that the observations were obtained using perfect measurements, i.e. without contaminated observations. Note that as discussed in Maleki et al. (2017), this is hardly ever true in real-life applications; hence, Lee et al. (2019) investigated the effect of measurement errors on the NSSDS scheme using a linearly covariate error model to capture the inherent measurement inaccuracy. To reduce the negative effect of measurement errors, Lee et al. (2019) used the multiple measurements sampling strategy (instead of the standard single measurement).

Following a similar operational procedure as that in Figure 2.2, De Araújo Rodrigues et al. (2011) formulated the first double sampling scheme for attribute data called the NSSDS np scheme which monitors the number of nonconforming items in a sample and it was shown to have a significantly better performance than the basic np scheme. More recently, Lee and Khoo (2019c) investigated the performance of the NSSDS np scheme in Phase II when the process parameter is estimated from some IC historical Phase I data.

2.3.2 Variability

The first NSSDS scheme for variability was proposed in He and Grigoryan (2002), where the sample standard deviation is computed by, $S = \frac{1}{n}\sum_{j=1}^{n}\left(X_{tj} - \bar{X}\right)^2$ in each stage, accordingly, by using the operational procedure in Figure 2.2. Lee et al. (2010) illustrated a real-life application of the NSSDS S scheme using a wire bonding process of packaging, where they showed the effectiveness of the scheme in reducing the cost as it requires fewer samples. Next, He and Grigoryan (2003) presented an improved version of the scheme in He and Grigoryan (2002) without the normality assumption for the sample standard deviation. Similar to the manner that Hsu (2004) showed that the sole use of the *ASS* without other run-length measures may, in some cases, yield misleading results; Hsu (2007) showed that He and Grigoryan (2003)'s sole use of the *ASS* is questionable because the conclusion is invalid when using other run-length properties. Next, Khoo (2004) investigated the performance of the NSSDS scheme for monitoring the variability using the S^2 statistic when the underlying parameters are known and later, Castagliola et al. (2017) conducted the same study when the underlying parameters were estimated from a Phase I IC data and they also investigated the effect of Phase I sample size on the Phase II OOC performance.

Contrary, to the above publications that use either the standard deviation or the variance to monitor variability, Costa (2017) proposed an NSSDS scheme based on the sample ranges.

2.3.3 Location and variability

In the review paper by McCracken and Chakraborti (2012), the authors observed that monitoring the process mean alone would imply ignoring the changes in the process standard deviation, despite being well known that the latter can be greatly affected when the mean value gives a poor measure of central tendency. For monitoring both the mean and variability simultaneously, it is assumed that the process is OOC if either the process mean shifts from μ_0 to $\mu_1 = \mu_0 \pm \delta\sigma_0$ (i.e., $|\delta| > 0$) and/or the process standard deviation shifts from σ_0 to $\sigma_1 = \gamma\sigma_0$ (i.e., $\gamma > 1$ for increase in σ_0, or $0 < \gamma < 1$ for decrease in σ_0). The process is IC if $\delta = 0$ and $\gamma = 1$. He and Grigoryan (2006) first proposed the NSSDS scheme to monitor both the mean and standard deviation simultaneously using the NSSDS $\bar{X}$ sub-scheme by Daudin (1992) and the NSSDS S sub-scheme by He and Grigoryan (2002) i.e., with separate schemes for the mean and standard deviation. Later, Lee and Khoo (2017b) proposed the use of the single max-type plotting statistic (see Chen and Cheng (1998)); that is, instead of separately plotting the standardized mean or standard deviation, one needs to plot the maximum value of either the standardized mean or standard deviation at each sampling point (for stage 1, and if needed, for stage 2 also) using the upper one-sided version of the charting regions in Figure 2.1(b).

Since there are cases in SPM application where the process mean and standard deviation may not be constant when the process is in an IC state; however, their corresponding ratios are proportional, then Ng et al. (2018) implemented the SSDS charting regions in Figure 2.1(c) to monitor the coefficient of variation (CV) measuring the run-length performance with the *ANOS*; however, using samples of size n_1 in stage 1, and n_2 *only* in stage 2. Next, Rozi et al. (2019) instead implemented the NSSDS charting regions in Figure 1(b) with samples of size n_1 in stage 1, and (the combined samples) $n_1 + n_2$ in stage 2 to show that it outperforms the Ng et al. (2018) version using the *ANOS*. Considering that Malela-Majika et al. (2019) (i.e. Chapter 3 of this dissertation) shows that the SSDS design has a better OOC performance than the NSSDS design when monitoring the mean, one would expect that it would be the case for the CV statistic too. However, since Ng et al. (2018)'s SSDS scheme did not use the combined samples (i.e. $n_1 + n_2$) in stage 2, instead used a smaller sample of size n_2, then the Rozi et al. (2019)'s NSSDS scheme with a larger sample size in stage 2 ended up outperforming the SSDS scheme.

2.4 Multivariate double sampling schemes

A majority of publications on double sampling control charts are based on univariate monitoring schemes, with just only 8 (out of 76) publications on multivariate schemes - see the outline on Table 2.1. When more than one characteristics (either i.i.d. or correlated) are to be monitored, multivariate charts must be used. If observations are p-variate normal random variable with mean $\boldsymbol{\mu_0}$ and variance $\boldsymbol{\Sigma}_0$, then the sequence of observations is denoted by $\{\boldsymbol{X}_{tj} = \left[X_{1tj} \;\; X_{2tj} \;\; \ldots \;\; X_{ptj}\right]' : t \geq 1; j = 1, 2, \ldots, n\}$. The NSSDS schemes for multivariate data are based on the Hotelling's $T^2 = n(\overline{\boldsymbol{X}} - \boldsymbol{\mu_0})'\boldsymbol{\Sigma}_0^{-1}(\overline{\boldsymbol{X}} - \boldsymbol{\mu_0})$ statistic and the generalized sample variance (or equivalently, the determinant of the sample covariance matrix) $|\mathbf{S}| = \left|\frac{1}{n}\sum_{j=1}^{n}\left(\boldsymbol{X}_{tj} - \overline{\boldsymbol{X}}\right)'\left(\boldsymbol{X}_{tj} - \overline{\boldsymbol{X}}\right)\right|$ which are used to monitor the multivariate sample mean and standard deviation, respectively. The latter were first proposed by Champ and Aparisi (2008) and Grigoryan and He (2005), respectively. Note that, unlike the univariate double sampling schemes, the multivariate ones tend to be designed as one-sided schemes. Faraz et al. (2012) conducted an intensive economic-statistical design for the optimal set of parameters for the NSSDS T^2 scheme and they showed that, in most cases, it even outperforms the well-known multivariate EWMA T^2 scheme.

For the specific bivariate case, Costa and Machado (2008) showed that the one-sided NSSDS T^2 scheme has a better performance than the basic, VSS, VSI T^2 schemes. Moreover, they observed that the one-sided version of Croasdale (1974)'s regions in Figure 2.1(a) are more favourable in terms of implementation it is known that the OOC signal can only take place in stage 2, and in some cases, it yields better OOC performance than the NSSDS T^2 scheme. Next, for the bivariate sample variability, Machado and Costa (2008) proposed a NSSDS scheme based on the VMAX statistic which can be used for monitoring a covariance matrix of a bivariate normal process, i.e. VMAX statistic utilizes the sample variances of two correlated random variables given by VMAX = $\max\{S_x^2, S_y^2\}$, where $S_x^2 = \sum_{j=1}^{n} x_j^2/n$, $S_y^2 = \sum_{j=1}^{n} y_j^2/n$ and the samples are denoted by (x_j, y_j), $j = 1,2, \ldots, n$.

Other multivariate double sampling schemes are discussed under the appropriate subsections of Section 2.5.

2.5 Other monitoring schemes combined with the double sampling scheme

Khoo et al. (2016) and Teoh et al. (2014a) compared the performance of the double sampling $\bar{X}$ scheme against the VSI and VSS $\bar{X}$ schemes, respectively. It was observed that the VSI schemes had a better OOC performance when moderate to large shifts are of interest using the *ATS*; that is, the NSSDS scheme has a better performance for small shifts only. Next, the NSSDS scheme has a better OOC performance than the VSS scheme when using the *ARL* and *SDRL*; however, the converse is true when using the *ASS*.

Because the purpose of integrating different monitoring schemes is to produce an improved scheme that has a better performance than the individual combined schemes, several monitoring schemes have been integrated with the basic double sampling scheme in an effort to improve its performance. Such monitoring schemes that we are aware of, so far, that have been integrated with the basic double sampling schemes are: (i) Variable sampling interval (VSI) scheme, (ii) Variable sample size and interval (VSSI) scheme, (iii) Synthetic scheme, (iv) Group-runs scheme and (v) Exponentially weighted moving average (EWMA) procedure.

2.5.1 VSI and VSSI scheme

For a better understanding of VSI and VSSI schemes, the reader is referred to the literature review by Psarakis (2015). Assume that the possible sample sizes are $n_1 < n_2$ and we define the long and short sampling intervals as d_1 and d_2, respectively, where $d_1 > d_2$. Carot et al. (2002) were the first to combine the NSSDS scheme with the VSI design using the charting regions in Figure 2.1(b). At each sampling point t, in stage 1, the sample size is fixed at n_1; however, the sampling interval is allowed to vary as follows

$$\begin{cases} d_2, & \text{if } Z_{1,t-1} \in \text{Region B or C} \\ d_1, & \text{if } Z_{1,t-1} \in \text{Region A.} \end{cases} \tag{2.28}$$

Later, Torng et al. (2010) studied the corresponding works with the normality assumption relaxed by using various *t*- and gamma distributions with different parameters. Note though, slightly different charting regions were used in stage 1 – see Figure 2.3. That is, Torng et al. (2010) defined the implementation of the sampling intervals at Z_{1t} as follows

$$\begin{cases} d_2, & \text{if } Z_{1,t-1} \in \text{Region B1 or B2 or C} \\ d_1, & \text{if } Z_{1,t-1} \in \text{Region A.} \end{cases} \tag{2.29}$$

Moreover, unlike Carot et al. (2002), the charting procedure moves to stage 2 when Z_{1t} falls in Region B2 in stage 1. As an improvement to Haq and Khoo (2018), Umar et al. (2019) investigated the performance of the NSSDS scheme combined with the VSI design for monitoring the process mean with regression-type estimators under specific conditions of correlation (i.e. with auxiliary based information).

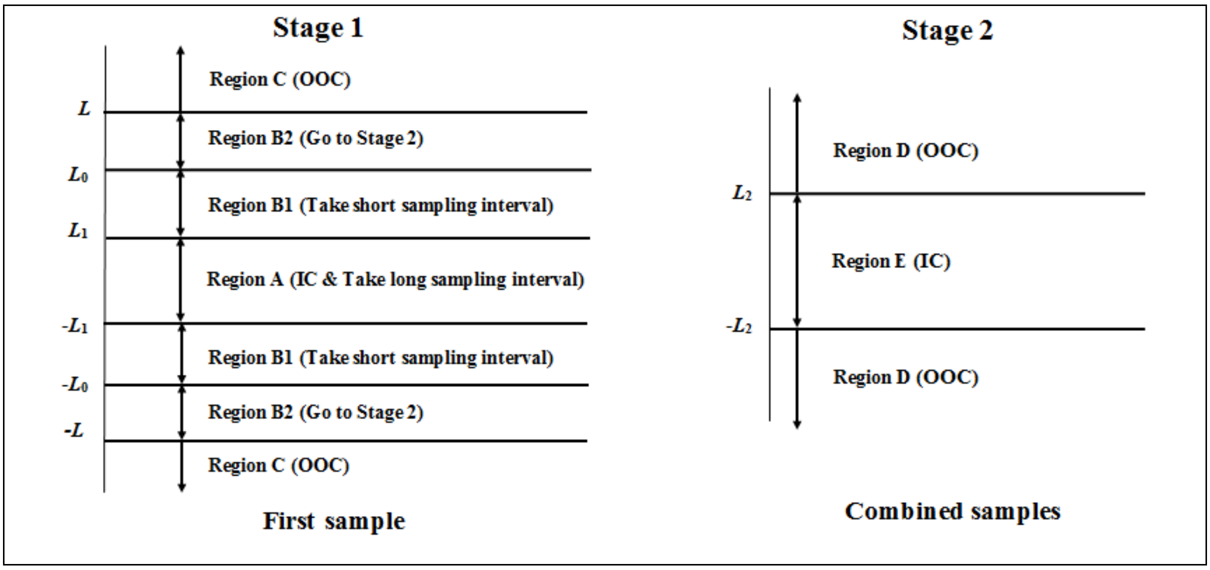


Figure 2.3: The charting regions in stages 1 and 2 of the VSI double sampling scheme

However, Noorossana et al. (2015) combined the NSSDS scheme with the VSSI design using the charting regions in Figure 2.3. While the NSSDS with the VSI design has n_1 and n_2 only, the double sampling with VSSI design has n_1, n_2 and n_3 (with $n_1 < n_2 < n_3$). Consequently, at each sampling point t, the sample size and sampling interval (denoted as (n_i, d_i)) are defined as follows

$$\begin{cases} (n_3, 0) & \text{if } Z_{1,t-1} \in \text{Region B2} \\ (n_2, d_2) & \text{if } Z_{1,t-1} \in \text{Region B1 or C} \\ (n_1, d_1) & \text{if } Z_{1,t-1} \in \text{Region A.} \end{cases} \tag{2.30}$$

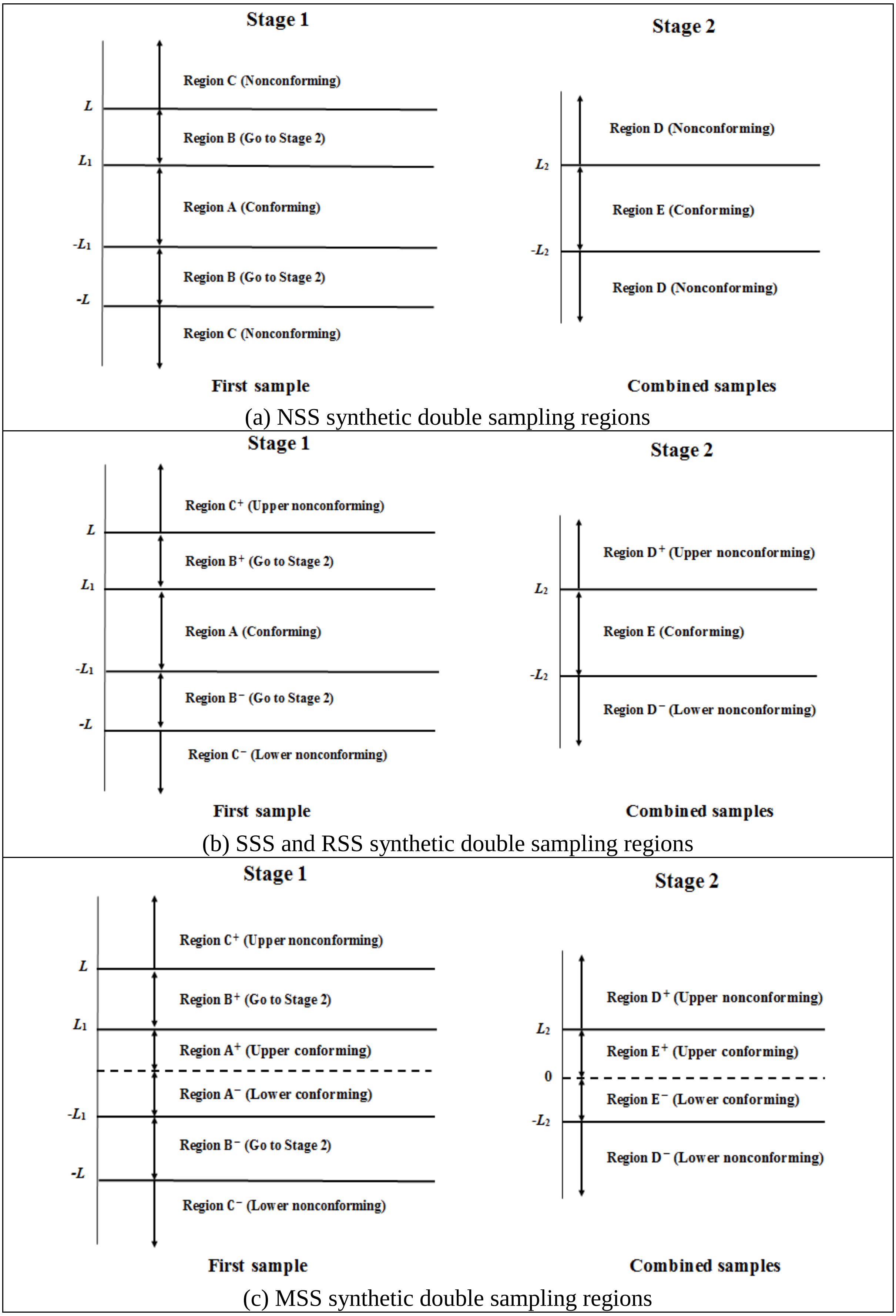
That is, when a plotting statistic falls in Region B2, the charting procedure moves to stage 2 immediately (at that sampling point, i.e. the sampling interval is equal to zero) using a combined sample of size either $(n_1 + n_3)$ or $(n_2 + n_3)$ depending on the previous sample.
Unlike Torng et al. (2010) who used integral equations to evaluate the run-length distribution, Carot et al. (2002) and Noorossana et al. (2015) used the Markov chain approach outlined in Jensen et al. (2008) to obtain the *ATS*, *ANSS* and *ANOS*. Using these run-length properties, Noorossana et al. (2015) showed that the double sampling scheme with the VSSI design has a better performance than the corresponding VSI counterpart. Moreover, it performs better than all the corresponding basic Shewhart VSS, VSI and VSSI $\bar{X}$ schemes. Note that the economic design of the VSI double sampling $\bar{X}$ scheme is studied in Lee et al. (2012a).

The NSSDS S scheme combined with the VSI design for monitoring the standard deviation is discussed in Lee et al. (2012b). The corresponding scheme that jointly monitors the mean and the standard deviation (i.e., which in essence incorporates the VSI design to the He and Grigoryan (2006)'s joint $\bar{X}$ and S NSSDS scheme) was proposed in Lee (2013). Similarly; however, in the case of attributes data, Lee and Khoo (2017c) combined the NSSDS np scheme with the VSI design.

For multivariate data, Khatun et al. (2018) and Katebi and Moghadam (2020) investigated the performance of the NSSDS T^2 scheme combined with the VSI and VSSI designs, respectively. In addition, the VSSI design incorporated into the NSSDS T^2 scheme outperforms that of the VSI design in detecting shifts in the vector of process means. Lee and Khoo (2018) investigated the performance of the NSSDS $|\mathbf{S}|$ scheme combined with the VSI design. In the latter three multivariate articles, the combined schemes were shown to yield much better performance than their individual counterparts.

2.5.2 Synthetic scheme

For a better understanding of synthetic schemes, the reader is referred to the literature review by Rakitzis et al. (2019). The conforming run-length (CRL) is defined as the number of samples observed between two consecutive nonconforming samples, inclusive of the nonconforming sample at the end. The main difference between a basic NSSDS scheme (in Figure 2.1(b)) and a non-side-sensitive (NSS) synthetic double sampling scheme (in Figure 2.4(a)) is that the latter does not issue OOC signal at the first sample point that falls on the nonconforming regions (i.e., the 'OOC regions' in Figure 2.1(b)). That is, the process waits until a second sample point falls on the nonconforming region and, if these two nonconforming samples are relatively close to each other (say, $CRL \leq H$), then an OOC signal is triggered. Note that H is a positive integer greater than 0 and it is defined as a control limit of the CRL scheme.



(a) NSS synthetic double sampling regions

(b) SSS and RSS synthetic double sampling regions

(c) MSS synthetic double sampling regions

Figure 2.4: The charting regions in stages 1 and 2 of the synthetic double sampling scheme

Khoo et al. (2011b) and Khoo et al. (2013) were the first to integrate the operation of a NSS double sampling synthetic scheme with regions given in Figure 2.4(a) using the $\bar{X}$ and T^2 charting statistics, respectively. It was observed that the integrated scheme has a significant improvement over the individual synthetic or NSSDS scheme. Next, the corresponding economic designs were studied in Lee and Khoo (2019a, b) under a variety of constraints for the univariate and multivariate NSS synthetic double sampling schemes, respectively. Aghaulor and Ezekwem (2016) designed the NSS synthetic double sampling scheme in a slightly different manner than Khoo et al. (2011b); that is, they implemented an algorithm such that the samples of sizes n_1 and n_2 used in stages 1 and 2, respectively, are such that: $n_1 < n$, $n_2 < 2n$ and $n_1 + n_2 < n$. Next, Costa and Machado (2015) realized that a side-sensitive version of the Khoo et al. (2011b) scheme yields an improved performance; hence they proposed the standard side-sensitive (SSS) synthetic double sampling scheme using the regions in Figure 2.4(b). Finally, Malela-Majika and Rapoo (2019) proposed the revised and modified side-sensitive (denoted by RSS and MSS, respectively) synthetic double sampling schemes; and they showed that the latter two schemes outperform the other synthetic double sampling schemes.

As an improvement to the NSSDS scheme for monitoring the mean with auxiliary variable by Haq and Khoo (2018), Haq and Khoo (2019) combined the later with the CRL sub-scheme.

After observing that there is no synthetic double sampling scheme dedicated to simultaneously monitoring the mean and standard deviation, Malela-Majika (2019) used the regions in Figure 2.4(c) and the CRL sub-scheme to propose the MSS synthetic double scheme with an OOC performance better than all its Shewhart-type competitors.

Having observed the performance of the basic NSSDS *np* scheme by De Araújo Rodrigues et al. (2011), Chong et al. (2014) investigated the corresponding NSS synthetic double sampling *np* scheme. For the variability case, Lee and Khoo (2017a) extended on He and Grigoryan (2002) work and proposed a NSS synthetic double sampling S scheme.

2.5.3 Group-runs scheme

For a better understanding of group-runs schemes, the reader is referred to Gadre and Rattihalli (2007). Khoo et al. (2015) and Chong et al. (2017) proposed the side-sensitive group-runs (SSGR) double sampling scheme for the process mean and number of nonconforming items in a sample, respectively. Looking at group-runs schemes in a different way, it is a generalized version of the synthetic schemes in Section 2.5.2, i.e. it is similar to the CRL sub-scheme except in the decision making procedure. That is, the group-runs schemes give an OOC signal when the first CRL charting statistic is less or equal to H (i.e., $CRL_1 \leq H$), or any two consecutive CRL charting statistics are both less than or equal to H (i.e., $CRL_i \leq H$ and $CRL_{i+1} \leq H$, i=2,3,…). SSGR double sampling schemes uses the charting regions in Figure 2.4(b) similar to those of the RSS synthetic schemes. The zero- and steady-state OOC performance of these schemes were computed using the ANOS. Chong et al. (2018) studied the run-length in more details by evaluating additional run-length properties, i.e., the median and percentile number of observations to signal. Later, Saha et al. (2018) enhanced the latter scheme by re-defining the CRL sub-scheme so that it has two limits, i.e. a warning limit (denoted by H_1) and a control limit (denoted by H_2), with $H_1 < H_2$, where H_1 and H_2 are positive integers greater than 0. This new scheme was called the modified SSGR double sampling. The modified SSGR double sampling scheme gives an OOC signal when the first CRL charting statistic is less or equal to H_2 (i.e., $CRL_1 \leq H_2$), or any two consecutive $CRL_i \leq H_1$ and $CRL_{i+1} \leq H_2$, i=2,3,…). Using the *ANOS* and *EANOS*, the modified SSGR double sampling has been shown to outperform the SSGR double sampling scheme and a variety of other Shewhart-type competitors.

2.5.4 EWMA procedure

Yang and Wu (2017a) used an asymmetric version of the control limits in Figure 2.1(b) to study the EWMA double sampling scheme based on the nonparametric sign statistic. This latter scheme was studied, and shown to yield a better performance than a variety of parametric and distribution-free schemes under the normal, double exponential, uniform, chi-square and exponential distributions. Similarly, Yang and Wu (2017b) showed that the asymmetric EWMA double sampling scheme for monitoring the variance has a better performance compared with the parametric and distribution-free schemes for monitoring variability.

2.6 Concluding remarks

In this chapter, all 76 existing publications that use the double sampling methodology to monitor the location, variability, both the location and variability simultaneously, etc. using univariate or multivariate techniques are categorized and summarized so that any research gaps can easily be identified. Note that other different monitoring schemes that are integrated with the operation of the double sampling methodology are also included in this chapter.

Based on this review chapter it is apparent that double sampling schemes are one of the most powerful Shewhart-type schemes in SPM literature and yields a better OOC performance than a majority of Shewhart-type schemes and using some run-length metric shows that they yield more competitive performance as compared to memory-type schemes, i.e. CUSUM and EWMA. Moreover, credit to Yang and Wu (2017a, b), the double sampling procedure has been integrated with the EWMA procedure to show that it has an even better OOC performance when integrated with memory-type schemes.

Note that the directions for future research suggestions are given in Chapter 5 of this dissertation.

Chapter 3. A new side-sensitive double sampling $\bar{X}$ scheme for monitoring an abrupt change in the process location

3.1 Introduction

A review of all currently available research works on double sampling schemes in SPM literature has been reviewed in Chapter 2, i.e. from 1974 up to November 2019. To ensure that this chapter is self-contained, there will be some few key concepts and figures that will be reproduced from Chapter 2; however, the rest of the concepts are described in Chapters 1 and 2.

Keep in mind that the double sampling $\bar{X}$ scheme was first presented by Croasdale (1974) as an attempt to improve the standard Shewhart $\bar{X}$ scheme in detecting small and moderate shifts in the process mean. Croasdale's double sampling scheme is a two-stage scheme based on two unconnected samples, where the master sample size is equal to n. Note that 'unconnected samples' imply that the first sample size (denoted as n_1) is used in Stage 1 whereas the second sample size (denoted as n_2) is used in Stage 2; where $n = n_1 + n_2$. Daudin et al. (1990) and Daudin (1992) modified Croasdale's double sampling scheme by connecting the first and the second sample at the second stage; that is, instead of using a sample of size n_2, they used a sample of size n. Since then, many authors have contributed to the design of the double sampling schemes – see Table 2.1 in Chapter 2.

To further increase the sensitivity of monitoring schemes, the SPM literature suggests the use of improved schemes such as the synthetic and runs-rules schemes. These schemes could be classified into two main categories, which are the non-side-sensitive (NSS) and side-sensitive (SS) schemes, respectively. The NSS *w*-*of*-(*w*+*v*) scheme (with integers $w > 1$ and $v \geq 0$) gives an out-of-control (OOC) signal when *w* nonconforming points out of the last *w*+*v* successive points plot outside of the control limits, no matter whether some (or all) of the *w* nonconforming points plot above the *UCL* and others (or all) plot below the *LCL*, which are separated by, at most, *v* conforming points that plot between the *LCL* and the *UCL*. Alternatively, the SS *w*-*of*-(*w*+*v*) scheme gives a signal when *w* nonconforming points out of the last *w*+*v* successive points plot on or above (below) the *UCL* (*LCL*), which are separated by, at most, *v* points that plot below (above) the *UCL* (*LCL*), respectively. Klein (2000) and Shongwe and Graham (2016) showed that the SS schemes not only improves the sensitivity of the basic (i.e. 1-*of*-1) scheme in detecting small shifts, but also outperforms the corresponding NSS scheme.

Motivated by the discussion in the latter paragraph, in this chapter, we propose the side-sensitive double sampling (SSDS) $\bar{X}$ scheme with known process parameters (i.e. Case K) in order to improve the existing non-side-sensitive double sampling (NSSDS) $\bar{X}$ scheme by Daudin (1992) in detecting small ($\delta < 0.75$) and moderate ($0.75 \leq \delta < 1.5$) shifts without affecting its sensitivity in detecting large shifts ($\delta \geq 1.5$).

The remainder of this chapter is organized as follows: In Section 3.2, we present the operation of the proposed SSDS $\bar{X}$ scheme and the exact expressions of the probability of the in-control (IC) process and average run-length (ARL). Section 3.3 presents the measures of the overall performance. In Section 3.4, we evaluate the IC and OOC performances of the proposed monitoring scheme and compare their overall performances with some well-known monitoring schemes. In Section 3.5, we give an illustrative examples using real-life data to demonstrate the implementation and design of the SSDS scheme. Finally, some concluding remarks are given in Section 3.6.

3.2 Operation and design consideration

3.2.1. Operation of the SSDS $\bar{X}$ control scheme

Assume that the observations of the quality characteristic $\{X_{tj}$: t=1,2,3,…; j=1,2,3,…,$n\}$ are independent and identically distributed (iid) from a $N(\mu_0,\ \sigma_0)$ distribution, where μ_0 and σ_0 represent the IC mean and the IC standard deviation, respectively. Let L_1 and L (with $L \geq L_1 > 0$) be the warning and control limits of the first sample at Stage 1, respectively; and L_2 (with $L_2 > 0$) be the control limit of the combined samples at Stage 2. Therefore, the SSDS $\bar{X}$ scheme is divided into eight intervals, i.e. A = $[-L_1, L_1]$, B$^+$ = $(L_1, L]$, B$^-$ = $[-L, -L_1)$, C =$(-\infty, -L) \cup (L, +\infty)$, F$^+$ = $(L_2, +\infty)$, F$^-$ = $(-\infty, L_2]$, G$^-$ = $(-\infty, -L_2)$ and G$^+$ = $[-L_2, +\infty)$.

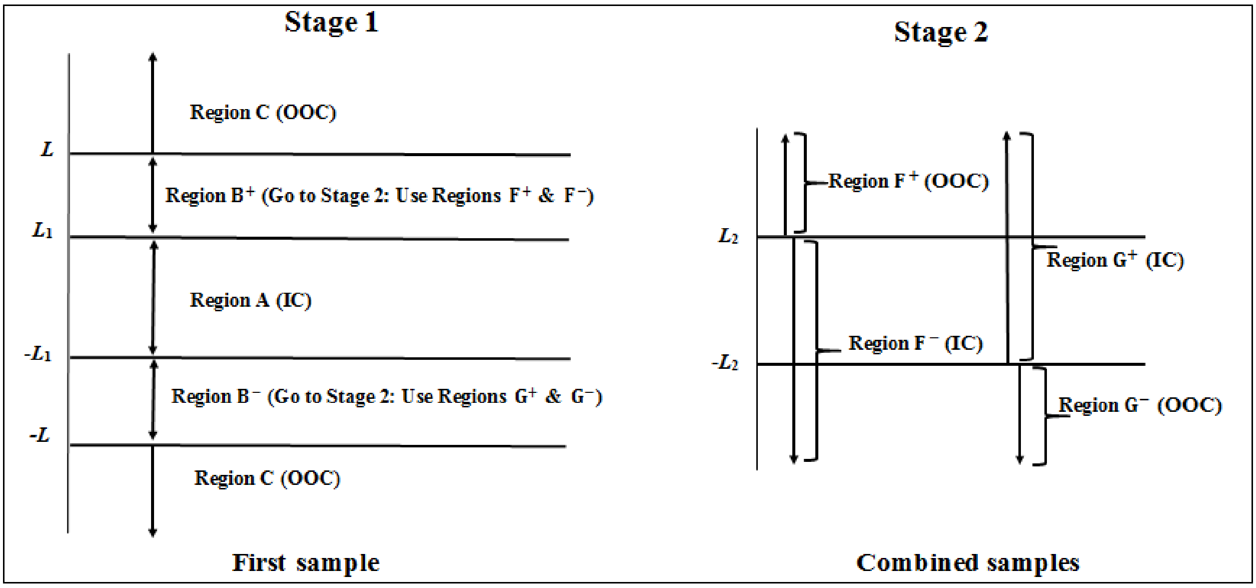


Figure 3.1: The charting regions of the SSDS scheme

From the X_{tj} observations, a first subgroup sample of size n_1 is collected at the t^{th} sampling time (denoted as X_{1tj}, $t = 1, 2, \ldots,$ and $j = 1, 2, \ldots, n_1$). If the standardized charting statistic based on the first sample falls on Region B$^-$ or B$^+$, then a second subgroup sample of size n_2 (where $n_2 \geq n_1$) is also collected at the t^{th} sampling time (denoted as X_{2tj}, $t = 1, 2, \ldots,$ and $j = 1, 2, \ldots, n_2$). Note that in each stage, the charting statistic is as follows.

Stage 1: Let $\bar{X}_{1t} = \sum_{j=1}^{n_1} X_{1tj}/n_1$ be the mean of the first sample of subgroup size n_1 at the t^{th} sampling time. Hence, in Case K, the standardized statistic for the first sample at the t^{th} sampling time is then given by

$$Z_{1t} = \frac{\bar{X}_{1t} - \mu_0}{\sigma_0/\sqrt{n_1}}$$

where $\bar{X}_{1t} \sim N(\mu_0 + \delta\sigma_0, \frac{\sigma_0}{\sqrt{n_1}})$ and $\delta = |\mu_1 - \mu_0|/\sigma_0$ represents the magnitude of the standardized mean shift with the OOC mean μ_1 ($\mu_1 = \mu_0 + \delta\sigma_0$), so that $\delta = 0$ means that the process is IC. In this case, Z_{1t} follows a standard normal distribution (i.e. $Z_{1t} \sim N(0,1)$). However, when $\delta \neq 0$, the process is OOC and $Z_{1t} \sim N(\delta, 1)$.

Stage 2: At the t^{th} sampling time of the second sample, the sample mean, i.e. $\bar{X}_{2t} = \sum_{j=1}^{n_2} X_{2tj}/n_2$, and the combined sample mean, i.e. $\bar{X}_t = (n_1\bar{X}_{1t} + n_2\bar{X}_{2t})/(n_1 + n_2)$ are calculated, respectively. Hence, in Case K, the standardized charting statistic for the combined samples at the t^{th} sampling time is then given by

$$Z_t = \frac{\bar{X}_t - \mu_0}{\sigma_0/\sqrt{n_1 + n_2}}$$

where $\bar{X}_t \sim N(\mu_0 + \delta\sigma_0, \frac{\sigma_0}{\sqrt{n_1+n_2}})$. When the process is IC, then $Z_t \sim N(0, 1)$ as $\delta = 0$ and when the process is OOC, then $Z_t \sim N(\delta, 1)$.

That is, there are two distinct standardized charting statistics (i.e. Z_{1t} and Z_t) used during stages 1 and 2 (if needed), respectively (see Figure 3.1). Thus, the operational procedure of the SSDS $\bar{X}$ scheme is as follows:

1. Take a sample of size n_1 and calculate $\bar{X}_{1t}$ and Z_{1t} at the t^{th} sampling time of the first sample.
2. If $Z_{1t} \in$ A, the process is considered as IC.
3. If $Z_{1t} \in$ C, the process is said to be OOC and then the necessary corrective action must be taken to find and remove the assignable causes.
4. If $Z_{1t} \in$ B$^+$ (or $Z_{1t} \in$ B$^-$), take a second sample of size n_2, with $n_2 \geq n_1$ and calculate $\bar{X}_{2t}$ at the t^{th} sampling time of the second sample.
5. At the t^{th} sampling time, calculate $\bar{X}_t$ and Z_t.
6. Consequently, the process is declared IC at stage 2:

   (a) If $Z_{1t} \in$ B$^+$ and $Z_t \in$ F$^-$, or

   (b) If $Z_{1t} \in$ B$^-$ and $Z_t \in$ G$^+$.

   Otherwise, the process is declared OOC at stage 2:

   (a) If $Z_{1t} \in$ B$^+$ and $Z_t \in$ F$^+$, or

   (b) If $Z_{1t} \in$ B$^-$ and $Z_t \in$ G$^-$.

Remarks:

- If the charting statistic Z_{1t} falls in region B$^+$ at stage 1, then at stage 2, we consider the charting regions F$^+$ = $(L_2, +\infty)$ and F$^-$ = $(-\infty, L_2]$ only (i.e., the upper scheme with control limit L_2).
- However, if the charting statistic Z_{1t} falls in region B$^-$ at Stage 1, then at Stage 2, we consider charting regions G$^-$ = $(-\infty, -L_2)$ and G$^+$ = $[-L_2, +\infty)$ only (i.e., lower scheme with control limit $-L_2$).

Figure 3.2 provides a graphical summary of the operation of the proposed SSDS scheme.

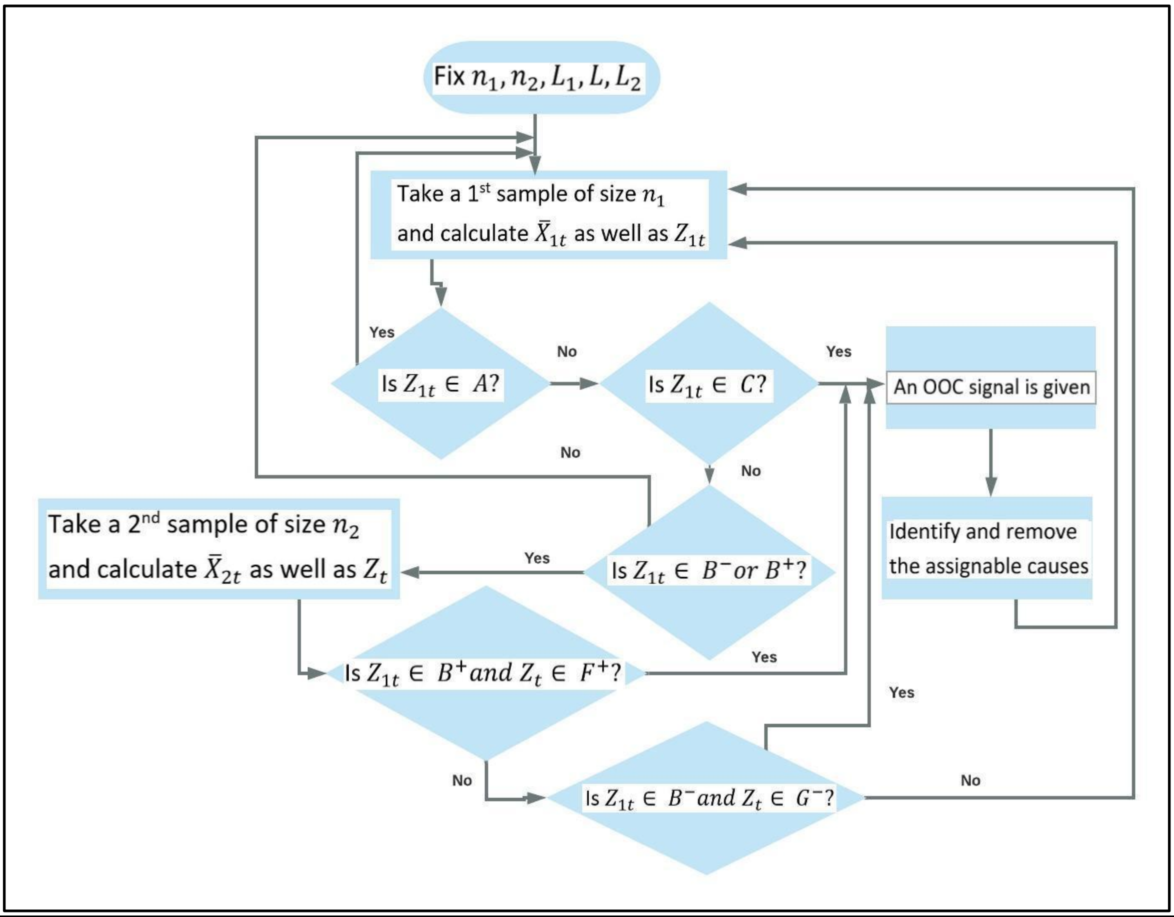


Figure 3.2: Flow chart for the proposed SSDS $\bar{X}$ monitoring scheme

3.2.2. Run-length properties of the SSDS $\bar{X}$ scheme

At stage 1, the SSDS $\bar{X}$ scheme gives an OOC signal if the charting statistic Z_{1t} plots in region C. Unlike the conventional NSSDS scheme by Daudin (1992), the SSDS scheme gives a signal at Stage 2 if both charting statistics Z_{1t} and Z_t plot on one side of the scheme in regions B^+ and F^+ (or B^- and G^-), respectively. In order to properly formulate run-length properties of the SSDS scheme, the following four types of events need to be defined:

(i) the Stage 1 IC event when "$Z_{1t} \in A$",

(ii) the Stage 2 IC event when either "$Z_{1t} \in B^+$ and $Z_t \in F^-$" or "$Z_{1t} \in B^-$ and $Z_t \in G^+$",

(iii) the Stage 1 OOC event when "$Z_{1t} \in C$" and

(iv) the Stage 2 OOC event when either "$Z_{1t} \in B^+$ and $Z_t \in F^+$" or "$Z_{1t} \in B^-$ and $Z_t \in G^-$".

Since μ_0 and σ_0^2 are known (i.e., Case K), let P_{0k} be the probability that the process is regarded as IC at stage k where k = 1, 2. Then, $P_0 = P_{01} + P_{02}$ is the probability that a process in both stages is IC, where:

$$P_{01} = P(Z_{1t} \in \mathrm{A}) = \Phi\left[L_1 + \delta\sqrt{n_1}\right] - \Phi\left[-L_1 + \delta\sqrt{n_1}\right] \tag{3.1}$$

and

$$\begin{aligned} P_{02} &= P[Z_{1t} \in \mathrm{B}^+ \text{ and } Z_t \in \mathrm{F}^-] + P[Z_{1t} \in \mathrm{B}^- \text{ and } Z_t \in \mathrm{G}^+] \\ &= \int_{Z_{1t}\in \mathrm{B}^{++}} \left\{\Phi[cL_2 + rc\delta - z\sqrt{n_1/n_2}]\right\} \phi(z)dz \\ &+ \int_{Z_{1t}\in \mathrm{B}^{--}} \left\{1 - \Phi[-cL_2 + rc\delta - z\sqrt{n_1/n_2}]\right\} \phi(z)dz. \end{aligned} \tag{3.2}$$

Equations (3.1) and (3.2) are based on events (i) and (ii), respectively. Hence,

$$\begin{aligned} P_0 &= \Phi\left[L_1 + \delta\sqrt{n_1}\right] - \Phi\left[-L_1 + \delta\sqrt{n_1}\right] \\ &+ \int_{Z_{1t}\in \mathrm{B}^{++}} \left\{\Phi[cL_2 + rc\delta - z\sqrt{n_1/n_2}]\right\} \phi(z)dz \\ &+ \int_{Z_{1t}\in \mathrm{B}^{--}} \left\{1 - \Phi[-cL_2 + rc\delta - z\sqrt{n_1/n_2}]\right\} \phi(z)dz, \end{aligned} \tag{3.3}$$

where $\Phi(.)$ and $\phi(.)$ are the cumulative distribution function (c.d.f.) and probability density function (p.d.f.) of the standard normal random variable, respectively. In this chapter, $r^2 = n_1 + n_2$, $c = r/\sqrt{n_2}$, $\mathrm{B}^{++} = (L_1 + \delta\sqrt{n_1}, L + \delta\sqrt{n_1}]$ and $\mathrm{B}^{--} = [-L + \delta\sqrt{n_1}, -L_1 + \delta\sqrt{n_1})$. Given that the SSDS $\bar{X}$ scheme is a Shewhart-type one, its run-length (RL) distribution, denoted by $F_{RL}(l)$, is defined by the geometric distribution. Therefore, the c.d.f. of the RL distribution is obtained as

$$F_{RL}(l) = P(RL \leq l) = 1 - P_{o,}^{l} \tag{3.4}$$

where $l \in \{1, 2, 3, \dots\}$.

Then, the $(100\rho)^{th}$ percentile of the RL distribution, l_ρ is given by

$$P\left(RL \leq l_\rho - 1\right) \leq \rho \text{ and } P(RL \leq l_\rho) > \rho \tag{3.5}$$

The ARL, standard deviation of the run-length ($SDRL$) and the average sample size (ASS) at each sampling time are given by

$$ARL = \frac{1}{1 - P_{\mathrm{o}}}, \tag{3.6}$$

$$SDRL = \frac{\sqrt{P_{\mathrm{o}}}}{1 - P_{\mathrm{o}}}, \tag{3.7}$$

and

$$ASS = n_1 + n_2 \times P_2, \tag{3.8}$$

respectively, where $P_2 = P(Z_{1t} \in \mathrm{B}^+ \cup \mathrm{B}^-)$ is the probability of taking the second sample, and it is given by

$$P_2 = \left(\Phi\left(L + \delta\sqrt{n_1}\right) - \Phi\left(L_1 + \delta\sqrt{n_1}\right)\right) + \left(\Phi\left(-L_1 + \delta\sqrt{n_1}\right) - \Phi\left(-L + \delta\sqrt{n_1}\right)\right).$$

Then, the average number of observations to signal (*ANOS*) is given by

$$ANOS = ASS \times ARL. \tag{3.9}$$

There are five parameters (n_1, n_2, L_1, L, L_2) that need to be specified in order to design the SSDS $\bar{X}$ scheme. The efficiency of the proposed SSDS scheme depends on the combinations (n_1, n_2, L_1, L, L_2). We have three steps in the optimal design of the proposed SSDS scheme:

- Firstly, the nominal IC *ARL* $(NARL_0)$ is set to some high recommended values, such as 370.4 or 500.
- Secondly, the (n_1, n_2, L_1, L, L_2) combination, which provides an attained IC ARL (ARL_0), is set much closer to the $NARL_0$ and the smallest OOC ARL (ARL_δ) for a given mean shift δ is considered as an optimal combination.
- Thirdly, the optimization model is presented as follows

  $$\underset{n_1, n_2, L_1, L, L_2}{\text{Min}} ARL_\delta \tag{3.10}$$

  subject to

  $$ARL_0 = NARL_0 \tag{3.11}$$

  and

  $$E[total\ sample\ size | \mu = \mu_0] = n_1 + n_2 \times P_2 = n, \tag{3.12}$$

  where *n* represents the expected IC *ASS* (denoted as ASS_0).

3.3 Measures of the overall performance

Although the ARL value is the most used metric measurement in SPM; numerous authors have advocated against the sole use of the *ARL* as a performance measure (see, for example, Graham et al. (2014) for a recent discussion on this issue). In addition to the arguments made by Graham et al. (2014), a number of authors have shown that if a control scheme is designed based on one specific size of a mean shift, it would perform poorly when the actual size of the shift is significantly different from the assumed size (Reynolds and Lou, 2010; Ryu et al., 2010). This makes the *ARL* deficient in assessing the overall performance of a control scheme. To solve this problem, a number of researchers have suggested the use of quality loss functions (*QLF*s) instead of the ARL to assess the performance of a monitoring scheme (see, for example, Machado and Costa, 2014). A *QLF* describes the relationship between the shift size and the quality impact. Therefore, when the aim of a study is to measure the overall performance of a control scheme over a range of shifts (i.e. $0 < \delta \leq \delta_{max}$), the objective function must be defined in terms of the average extra quadratic loss ($AEQL$) function given as (see Wu et al., 2008; Machado and Costa, 2014; Shongwe and Graham (2019a, b); Shongwe et al. (2019))

$$\underset{n_1,n_2,L_1,L,L_2}{\text{Min}} AEQL(\delta) \tag{3.13}$$

with

$$AEQL(\delta) = \frac{1}{\delta_{max}} \int_0^{\delta_{max}} w(\delta) \times ARL(\delta) \ \ d\delta$$

where δ_{max} is the upper boundary of the range of shifts under consideration and $w(\delta)$ (with $w(\delta) = \delta^2$) represents the weight function associated with δ. It is generally assumed that all location shifts (mean shifts) occur with equal probability. Therefore, a uniform distribution of δ is implied.

The expression of the $AEQL$ given in Equation (3.13) can also be written as follows

$$AEQL(\delta) = \frac{1}{\delta_{max}} \sum_{0}^{\delta_{max}} \delta^2 \times ARL(\delta). \tag{3.14}$$

Note that after using Equations (3.10) to (3.12), the optimal parameters are selected using Equation (3.14), which imply that they yield a minimum $AEQL$ value.

In order to measure the relative effectiveness of two different schemes, Wu et al. (2008) suggests the use of the performance comparison index (PCI), which is given as

$$PCI = \frac{AEQL}{AEQL_{benchmark}}, \tag{3.15}$$

where $AEQL_{benchmark}$ is the $AEQL$ of the benchmark scheme. In this paper, the SSDS $\bar{X}$ scheme is used as the benchmark scheme. In addition to the $AEQL$ and PCI, many authors suggest the use of the average ratio of the ARL (denoted $ARARL$) to measure the overall performance of a benchmark scheme against other competitors; see for instance, Wu et al. (2008). The $ARARL$ is given by

$$ARARL = \frac{1}{\delta_{max}} \sum_{0}^{\delta_{max}} \frac{ARL(\delta)}{ARL(\delta)_{benchmark}}. \tag{3.16}$$

Note that, if the PCI and/or $ARARL$ is larger than one, the competing scheme will produce larger ARL_δ values over the range of shifts under consideration. Therefore, the benchmark scheme outperforms the competing scheme for that specific range; otherwise, the competing scheme is more sensitive than the benchmark scheme. Finally, if the PCI and/or $ARARL$ is equal to one, then the competing scheme and the benchmark scheme have the same performance.

Although it is not the scope of this chapter, it is worth mentioning that the effectiveness of traditional performance measures should be revisited. Even as far back as 1986, Woodall (1986) had started mentioning flaws in the designs of monitoring schemes and, even today, there is still room for improvement. We do not wish to degrade the importance of traditional monitoring scheme performance metrics; however, the key common characteristic of these traditional methods is to design the monitoring scheme for a pre-specified magnitude of shift. Many researchers have now argued that if a monitoring scheme is designed for some pre-specified magnitude of shift, it will perform poorly when the actual shift differs significantly from this pre-specified value (see for example, Ryu et al., 2010; Machado and Costa, 2014). The recommendation is that the overall performance metrics, such as *QLF*s, must be used to supplement the specific shifts metrics. The exploration into the fact that making use of traditional measures can be misrepresentative is currently under investigation and will be reported on in a future research works.

3.4 Performance study of the proposed scheme

3.4.1 Optimal design of the SSDS $\bar{X}$ monitoring scheme

In this section, we investigate the optimal design of the SSDS $\bar{X}$ control scheme in Case K by setting the ARL_0 to some high, acceptable nominal values and minimizing the ARL_δ and / or EQL values with (δ_{min}, δ_{max}) = (0, 2.5) and a step shift of size 0.1. In this study, we set the $NARL_0$ to 370.4 and 500, respectively, with an ASS_0 (i.e. n) ∈ {2, 5, 7, 11}. We used Equations (3.4) to (3.7) in MATHCAD® 14 to compute the IC and OOC characteristics of the run-length distribution, respectively. Moreover, the *ASS*, *ANOS* and *AEQL* values were computed using Equations (3.8), (3.9) and (3.14), respectively. Note that there are three main steps in the search of the optimal design parameters:

(i) For some specific sample sizes (i.e., n_1 and n_2) and shift ($\delta = 0$), find all possible combinations of the design parameters that yield an attained ARL_0 value of 370.4 for a prespecified value of n (i.e., ASS_0);

(ii) For each combination, calculate the *AEQL* value;

(iii) Select the combination that yields the minimum *AEQL* value to be the optimal design parameters.

Table 3.1 presents the optimal design parameters of the proposed scheme when ASS_0 ∈ {2, 5, 7, 11}. For instance, when $(n_1, n_2) =$ (2, 8) with ASS_0 = 5, we compute the optimal design parameters (n_1, n_2, L_1, L, L_2) in order to achieve a specified ARL_0 of 370.4. In this example, we observed that (n_1, n_2, L_1, L, L_2) is such that (2, 8, 0.8856, 3.3526, 3.0085) so that the attained ARL_0 = 370.4 with a minimum EQL value of 33.99. Under the same conditions, and for another choice of design parameters, say (n_1, n_2, L_1, L, L_2) = (2, 8, 0.8815, 3.1026, 3.2731), the SSDS scheme yields an ARL_0 value of 370.4 with an EQL value of 39.31. In this situation, the design parameters that yield the smallest $AEQL$ (i.e. minimum $AEQL$) value are considered to be the winner, and thus are used as optimal design parameters. From Table 3.1 it is observed that the sensitivity of the SSDS scheme is proportional to the first (Stage 1) and second (Stage 2) sample sizes. This means that the larger the samples (i.e. n_1 and n_2), the more sensitive the SSDS scheme is.

Table 3.1: Optimal design parameters and $AEQL$ values when $NARL_0$ ∈{370.4, 500} and ASS_0 ∈{2, 5, 7, 11}

Attained ARL_0			**370.4**		**500**	
ASS_0	n_1	n_2	(L_1, L, L_2)	$AEQL$	(L_1, L, L_2)	$AEQL$
2	2	2	(2.9101, 3.0568, 2.4050)	120.11	(2.9631, 3.1688, 2.5120)	138.68
		5	(2.9001, 3.0073,2.9025)	119.97	(2.9337, 3.1021, 2.9611)	134.89
		8	(2.9751, 3.0009, 2.9805)	129.36	(3.0428, 3.0917, 3.0832)	150.93
		11	(2.9908, 3.0002, 2.9925)	131.56	(3.0732, 3.0906, 3.0926)	155.96
5	2	8	(0.8856, 3.3526, 3.0085)	33.99	(0.8868, 3.6788, 3.0367)	35.52
		11	(1.0941, 3.2339, 3.0101)	32.45	(1.0943, 3.3422, 3.0889)	34.46
		14	(1.2377, 3.1693, 3.0126)	32.01	(1.2373, 3.2522, 3.1180)	34.11
	4	4	(1.1491, 3.4180, 3.0412)	35.64	(1.1423, 3.2402, 3.2926)	42.90
		8	(1.5291, 3.2440, 3.0302)	31.11	(1.5293, 3.2602, 3.2027)	34.48
		11	(1.6821, 3.1435, 3.0617)	30.68	(1.6823, 3.2162, 3.1985)	33.07
		14	(1.7906, 3.0989, 3.0773)	30.61	(1.7943, 3.2302, 3.0846)	31.55
	5	5	(2.9934, 3.0008, 2.9998)	49.54	(2.9655, 3.1102, 3.0016)	52.20
		8	(2.9947, 3.0004, 2.9979)	49.56	(2.9935, 3.0992, 3.0301)	52.87
		11	(2.9961, 3.0002, 2.9993)	49.59	(2.9878, 3.0998, 2.9081)	52.15
7	3	8	(0.6740, 3.5671, 3.0013)	29.84	(0.6738, 3.5308, 3.1203)	32.51
		11	(0.9076, 3.5336, 2.9559)	27.60	(0.9078, 3.6105, 3.0541)	29.22
	5	5	(0.8406, 3.4148, 3.0521)	30.99	(0.8406, 3.4405, 3.1610)	33.88
		8	(1.1496, 3.6358, 2.9798)	27.41	(1.1498, 3.8905, 3.0618)	28.96
		11	(1.3342, 3.5663, 2.9306)	26.01	(1.3338, 3.9704, 3.0019)	27.05
	7	8	(2.9952, 3.0005, 2.9993)	37.72	(3.0008, 3.1025, 3.0014)	39.96
		11	(2.9962, 3.0003, 2.9998)	37.73	(3.0002, 3.0994, 3.0001)	39.82
11	3	8	(0.0001, 3.8993, 3.0020)	29.03	(0.0002, 4.2604, 3.0643)	31.07
		11	(0.3486, 3.8211, 2.9827)	26.48	(0.3482, 4.3104, 3.0643)	27.86
	5	8	(0.3185, 3.8526, 3.0049)	26.90	(0.3182, 4.1103, 3.0904)	28.51
		11	(0.6045, 3.8868, 2.9861)	25.08	(0.6039, 4.1006, 3.0750)	26.26

3.4.2 Performance of the SSDS $\bar{X}$ control scheme

Once the optimal parameters are obtained, the OOC performance of the proposed scheme can then be investigated. Tables 3.2 and 3.3 present the performance of the SSDS scheme for different optimal design parameters with $NARL_0$ = 370.4 when ASS_0 ∈ {2, 5} and {7, 11}, respectively. In terms of the ARL values, it was observed that the performance of the SSDS scheme depends on the ASS_0 value, and the first and second sample sizes (i.e., n_1 and n_2). When the couple (n_1, n_2) is kept constant,

(i) as the ASS_0 (i.e., n) value increases, the performance of the SSDS scheme improves (see Figure 3.3),

(ii) in terms of the cost of inspection, the SSDS scheme is cost effective for small ASS_0 values (see Figure 3.4),

(iii) for $ASS_0 = n_1 \leq 2$, when n_2 increases, the sensitivity of the SSDS scheme decreases regardless of the size of the process mean shift, and

(iv) for $ASS_0 > 2$, when n_2 increases, the sensitivity of the SSDS scheme increases for small and moderate shifts. For large shifts in the process mean, the sensitivity remains the same.

(v) for $ASS_0 = n_1$ and $n_1 > 2$, when n_2 increases, the sensitivity of the SSDS scheme increases regardless of the size of the process mean shift (see Figures 3.3 (a) to (d)). Note that when $ASS_0 = n_1$, the OOC ASS (ASS_δ) values remain closer to the ASS_0 value, which makes the SSDS scheme cost effective (see Figures 3.4 (a) to (d)).

Therefore, to obtain an optimal and cost effective design of the SSDS scheme, we suggest keeping the ASS_0 value and n_1 as small as possible at an acceptable cut-off point (e.g. $ASS_0 = n_1$ = 3 or 4) and increase n_2 in order to get an efficient and economic SSDS $\bar{X}$ scheme. In terms of the percentile of the run-length (PRL) values, for a $NARL_0$ of 370.4 with an ASS_0 value of 2 and (n_1, n_2, L_1, L, L_2) = (2, 5, 2.9001, 3.0073, 2.9025), the results in Table 3.2 reveal that when the process is IC (i.e. δ = 0), there is a 5% chance that the SSDS scheme signals for the first time on the 19th subgroup. However, there is 95% chance that the SSDS scheme signals for the first time on the 1110th subgroup. When there is an abrupt mean shift of size 0.2 (i.e. δ = 0.2), there is 5% chance that the SSDS scheme will signal on the 13th subgroup and 95% chance that it will signal on the 768th subgroup. For large shifts, there is 50% chance that the SSDS scheme signals on the second subgroup and 75% chance that it will signal on the 5th subgroup. When we increase the ASS_0 value, say $ASS_0 = 5$ for (n_1, n_2) = (2, 8), (L_1, L, L_2) = (0.8856, 3.3526, 3.0085) so that the scheme yields a $NARL_0$ of 370.4. In this case, when there is an abrupt mean shift of size 0.2, there is 5% chance that the SSDS scheme will signal on the 7th subgroup and 95% chance that it will signal on the 387th subgroup. For large sample sizes, there is a very high chance that the SSDS scheme signals on the first subgroup. This shows that the sensitivity of the SSDS scheme increases as the sample size increases. Consequently, the cost increases as well. It can also be observed that for small ASS_0 values, for instance when the $ASS_0 = 2$, the IC characteristics of run-length are the same for different optimal combinations regardless of the sample sizes. However, for moderate and large ASS_0 values, the IC characteristics of the run-length are not equal for different optimal combinations regardless of the sample sizes. For more details, see Tables 3.2 and 3.3.

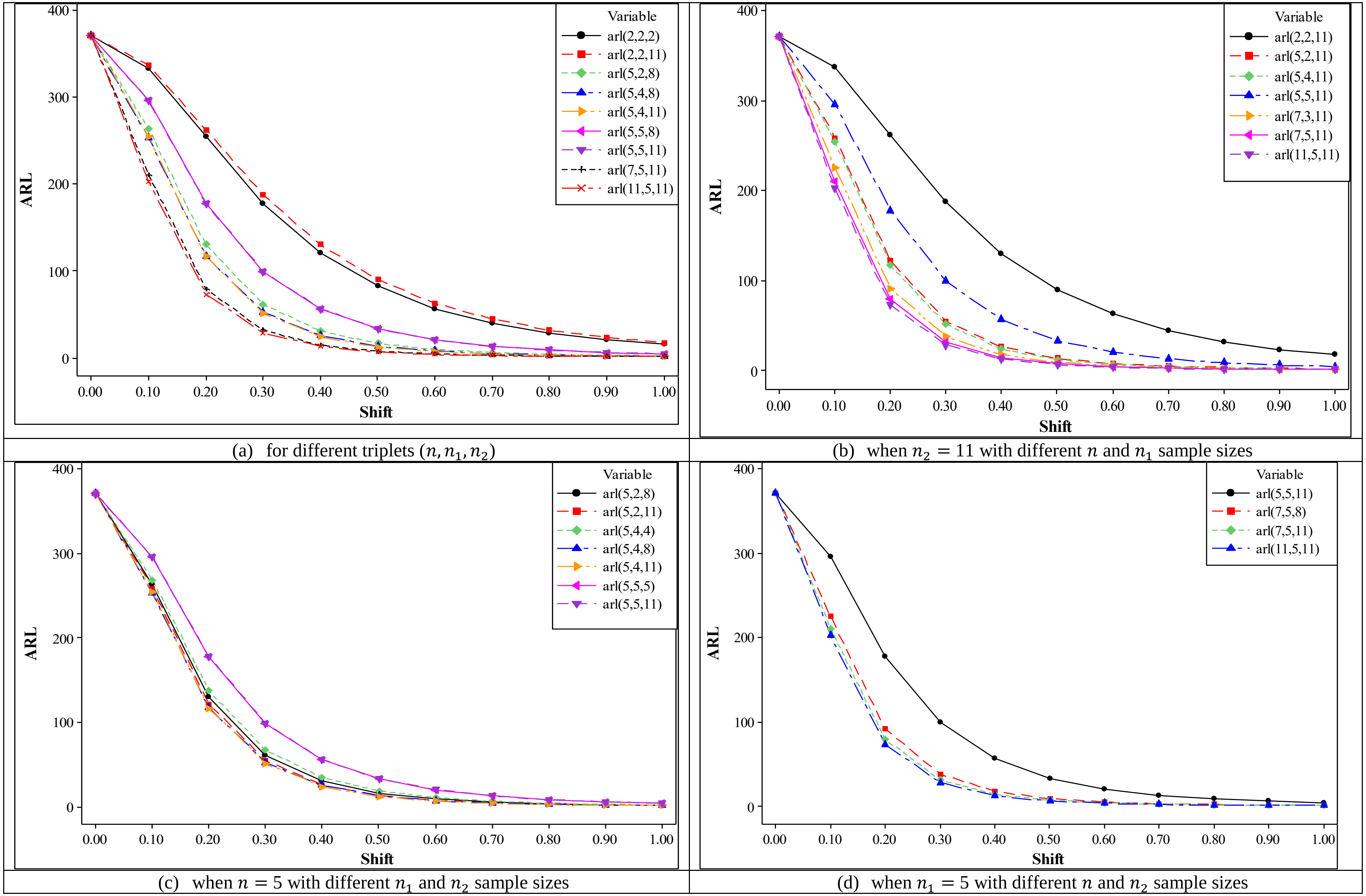
The results in Tables 3.2 and 3.3 also show that the distribution of the ASS_δ is symmetric or skewed or relatively constant about $\delta_{ASS_{max}}$(where $\delta_{ASS_{max}}$is the mean shift that produces the maximum ASS_δ value) depending on the triplet (ASS_0, n_1, n_2). Therefore, the ASS_δ of the proposed scheme may be considered as an increasing and decreasing function of δ in the ranges [0, $\delta_{ASS_{max}}$] and [$\delta_{ASS_{max}}$, δ_{max}], respectively. Figures 3.4 (a) to (d) reveal that, for small and moderate shifts, when n_2 is kept constant, the design of the SSDS scheme is cost effective for small n_1. For large shifts, the design of the SSDS scheme is cost effective for large n_2. When $ASS_0 = n_1$, the SSDS scheme is cost effective regardless of the size of the mean shift.

Table 3.2: Exact *ARL*, *SDRL*, *ASS*, *ANOS*, *AEQL*, percentile values and optimal design parameters of the SSDS scheme when the $NARL_0$ = 370.4, $n = ASS_0$ = 2 and 5 with δ_{max} = 2.5

Shift (δ)	**(*ARL, SDRL, ASS, ANOS*) (P5, P25, P50, P75, P95)**						
0.00	(370.40, 369.90, 2, 742.82) (19, 106, 256, 513, 1110)	(370.42, 369.82, 2, 741.54) (19, 106, 256, 513, 1110)	(370.36, 369.86, 2, 741.07) (19, 106, 256, 513, 1110)	(370.43, 369.93, 5, 1852) (19, 106, 256, 513, 1110)	(370.38, 369.87, 5, 1852) (19, 108, 261, 523, 1118)	(370.40, 369.91, 5, 1887) (19, 108, 261, 522, 1119)	(370.40, 369.90, 5, 1852.0) (19, 106, 256, 513, 1110)
0.20	(257.39, 256.89, 20.1, 516.64) (13, 74, 178, 356, 768)	(261.59, 261.08, 2.01 , 523.83) (13, 74.9, 181, 362, 780)	(262.28, 261.78, 2.02, 524.88) (13, 75, 181, 362, 780)	(130.06, 129.56, 5.15, 669.50) (7, 37, 89, 179, 387)	(122.00, 121.50, 5.20, 634.61) (6, 34, 82, 165, 356)	(117.18, 116.68, 5.23, 625.04) (6, 34, 82, 165, 357)	(177.37, 176.87, 5.01, 886.95) (9, 50, 122, 245, 529)
0.40	(123.33, 122.83, 2.02, 248.25) (6, 35, 85, 170, 367)	(128.97, 128.47, 2.03, 258.53) (7, 37, 89, 178, 385)	(130.08, 129.58, 2.03, 260.44) (7, 37, 89, 179, 385)	(30.63, 30.13, 5.56, 170.37) (2, 9, 21, 42, 90)	(26.23, 25.73, 5.78, 151.49) (1, 7, 17, 35, 76)	(25.51, 25.01, 5.88, 152.66) (1, 7, 17, 35, 76)	(56.33, 55.83, 5.01, 281.74) (3, 16, 38, 77, 167)
0.60	(57.44, 56.94, 2.02, 116.22) (3, 16, 39, 79, 170)	(61.70, 61.20, 2.03, 123.93) (3, 17, 42, 85, 183)	(62.66, 62.15, 2.04, 125.57) (3, 18, 43, 86, 185)	(9.47, 8.95, 6.16, 58.34) (1, 3, 6, 12, 27)	(7.81, 7.30, 6.63, 51.76) (1, 2, 5, 10, 22)	(7.65, 7.13, 6.82, 52.86) (1, 2, 5, 10, 21)	(20.43, 19.92, 5.02, 102.22) (1, 6, 13, 28 , 60)
0.80	(28.46, 27.96, 2.05, 58.09) (1, 8, 19 , 39, 84)	(31.24, 30.73, 2.05, 62.94) (2, 9, 21, 43 , 92)	(31.91, 31.41, 2.04, 64.06) (2, 9, 21, 43, 92)	(3.95, 3.42, 6.85, 27.06) (1, 1, 2, 5, 10)	(3.35, 2.81, 761, 25.50) (1, 1, 2, 4, 8)	(3.22, 2.68, 7.83, 25.44) (1, 1, 2, 4, 8)	(8.79, 8.27, 5.02, 44.01) (1, 2, 6, 11, 25)
1.00	(15.30, 14.79, 2.08, 31.60) (1, 4, 10, 21, 44)	(17.01, 16.50, 2.07, 34.42) (1, 5, 11, 23, 49)	(23.38, 22.87, 2.06, 35.08) (1, 5, 12, 24, 51)	(2.17, 1.60, 7.49, 16.27) (1, 1, 1, 2, 5)	(1.97, 1.38, 8.57, 16.89) (1, 1, 1, 2 , 4)	(1.84, 1.24, 8.60, 15.89 (1, 1, 1, 2,4)	(4.46, 3.93, 5.03, 22.36) (1, 1, 3, 6, 12)
1.20	(8.94, 8.42, 2.10, 18.74) (1, 2, 6, 12, 25)	(9.97, 9.45, 2.10, 20.29) (1, 2, 6, 13, 28)	(10.22, 9.71, 2.08, 20.62) (1, 3, 7, 13, 29)	(1.50, 0.87, 7.98, 11.99) (1, 1, 1, 1, 2)	(1.46, 0.82, 9.34, 13.65) (1, 1, 1 1, 2.)	(1.33, 0.67, 8.87, 11.86) (1, 1, 1, 1, 2)	(2.65, 2.09, 5.03, 13.26) (1, 1, 1, 3, 6)
1.40	(5.64, 5.12, 2.12, 12.04) (1, 1, 4, 7, 15)	(6.26, 5.74, 2.13, 12.83) (1, 2, 4, 8, 17)	(6.41, 5.89, 2.10, 12.97) (1, 2, 4, 8, 18)	(1.23, 0.53, 8.24, 10.10) (1, 1, 1, 1, 2)	(1.25, 0.55, 9.79, 12.20) (1, 1, 1, 1, 1)	(1.13, 0.39, 8.56, 9.71) (1, 1, 1,1, 1)	(1.80, 1.20, 5.05, 9.06) (1, 1, 1, 1, 4)
1.60	(3.83, 3.29, 2.13, 8.302) (1, 1, 2, 5, 10)	(4.20, 3.67, 2.16, 8.67) (1, 1, 3, 5, 11)	(4.29, 3.75, 2.12, 8.71) (1, 1, 2 , 5, 11)	(1.11, 0.34, 8.23, 9.11) (1, 1, 1, 1, 1)	(1.14, 0.40, 9.85, 11.22), (1, 1, 1, 1, 1)	(1.05, 0.24, 7.76, 8.17) (1, 1, 1, 1, 1)	(1.38, 0.73, 5.06, 6.96) (1, 1, 1 , 1, 2)
1.80	(2.77, 2.21, 2.14, 6.082) (1, 1, 2, 3, 7)	(3.00, 2.45, 2.12, 6.22) (1, 1, 2, 3, 7)	(3.05, 2.50, 2.13, 6.21) (1, 1, 2, 4, 8)	(1.05, 0.24, 7.94, 8.36) (1, 1, 1, 1, 1)	(1.08, 0.29, 9.49, 10.25) (1, 1, 1, 1, 1)	(1.02, 0.14, 6.73, 6.87) (1, 1, 1, 1, 1)	(1.18, 0.46, 5.08, 5.90) (1, 1, 1, 1, 2)
2.00	(2.12, 1.54, 2.15, 4.693) (1, 1, 1, 2, 5)	(2.26, 1.69, 2.10 , 4.71) (1, 1, 1, 2, 5)	(2.63, 2.07, 2.11, 4.69) (1, 1, 1, 2, 5)	(1.03, 0.17, 7.39, 7.52) (1, 1, 1, 1, 1)	(1.04, 0.21, 8.78, 9.16) (1, 1, 1, 1, 1)	(1.01, 0.08, 5.74, 5.79) (1, 1, 1, 1, 1)	(1.07, 0.28, 5.09, 5.38) (1, 1, 1, 1, 1)
2.50	(1.36, 0.69, 2.13, 2.957) (1, 1, 1, 1, 2)	(1.40, 0.75, 2.07, 2.91) (1, 1, 1, 1, 2)	(2.30, 1.73, 2.08, 2.88) (1, 1, 1, 1, 2)	(1.00, 0.06, 5.39, 5.41) (1, 1, 1, 1, 1)	(1.01, 0.09, 6.12, 6.16) (1, 1, 1, 1, 1)	(1.00, 0.02, 4.47, 4.32) (1, 1, 1, 1, 1)	(1.00, 0.07, 5.11, 5.02) (1, 1, 1, 1, 1)
AEQL	119.97	129.36	131.56	33.99	32.45	31. 11	49.56
(L_1, L, L_2)	(2.9001, 3.0073, 2.9025)	(2.9751, 3.0009, 2.9805)	(2.9908, 3.0002, 2.9925)	(0.8856, 3.3526, 3.0085)	(1.0941, 3.2339, 3.0101)	(1.5291, 3.2440, 3.0412)	(2.9934, 3.0008, 2.9998)
(n, n_1, n_2)	(2, 2, 5)	(2, 2, 8)	(2, 2, 11)	(5, 2, 8)	(5, 2, 11)	(5, 4, 8)	(5, 5, 5)

Table 3.3: Exact *ARL*, *SDRL*, *ASS*, *ANOS*, *AEQL*, percentile values and the optimal design parameters of the SSDS scheme when the $NARL_0$ = 370.4, $n = ASS_0$ = 7 and 11 with δ_{max} = 2.5

Shift (δ)	**(*ARL, SDRL, ASS, ANOS*)** **(P5, P25, P50, P75, P95)**						
0.00	(370.48, 369.97, 7.00, 2593) (19, 106, 256, 513, 1110)	(370.37, 369.87, 7.00, 197.37) ((19, 106, 256, 513, 1110)	(370.38, 369.88, 7.00, 2593) (19, 106, 256, 513, 1110)	(370.44, 369.94, 7.00, 2593) (19, 106, 256, 513, 1110)	(370.42, 369.92, 11.00, 4075) (19, 106, 256, 513, 1110)	(370.40, 369.90, 11.00, 4074) (19, 106, 256, 513, 1110)	(370.36, 369.86, 11.00, 4074) (19, 106, 256, 513, 1110)
0.20	(108.02, 107.52, 7.20, 777.51) (6, 31, 74, 149, 322)	(115.65, 115.15, 7.22, 99.33) (1, 33, 79, 159, 342)	(91.59, 91.09, 7.37, 674.57) (5, 26, 63, 126, 271)	(79.18, 78.68, 7.46, 591.04) (4, 22, 54, 109, 235)	(84.43, 83.94, 11.17, 942.86) (4, 24, 58, 116, 249)	(88.48, 87.98, 11.18, 989.44) (4, 25, 61, 122, 264)	(73.18, 72.68, 11.42, 835.79) (4, 21, 50, 100, 218)
0.40	(23.11, 22.61, 178.68) (1, 7, 16, 31, 67)	(25.87, 25.37, 7.78, 39.77) (1, 1, 3, 6, 14)	(17.94, 17.43, 8.33, 149.48) (1, 5, 12, 24, 52)	(14.39, 13.88, 8.73, 125,68) (1, 4, 10, 19, 41)	(15.83, 15.32, 11.61, 183,76) (1, 4, 10, 21, 46)	(17.02, 16.51, 11.63, 197.93) (1, 5, 11, 23, 49)	(12.79, 12.28, 12.48, 159.57) (1, 4, 9, 17, 36)
0.60	(6.99, 6.47, 8.44, 58.98) (1, 1, 4, 9, 19)	(7.89, 7.38, 8.44, 20.84) (1, 1, 1, 3, 6)	(5.34, 4.81, 9.57, 51.12) (1, 1, 3, 7, 14)	(4.31, 3.78, 10.43, 44.94) (1, 1, 3, 5, 11)	(4.69, 4.16, 12.18, 57.14) (1, 1, 3, 6, 13)	(5.04, 4.51, 12.11, 60.99) (1, 1, 3, 6, 13)	(3.77, 3.24, 13.69, 51.67) (1, 1, 2, 5, 10)
0.80	(2.97, 2.42, 9.13, 27.16) (1, 1, 2, 3, 7)	(3.30, 2.75, 8.90, 13.83) (1, 1, 1, 1, 3)	(2.36, 1.79, 10.66, 25.13) (1, 1, 1, 3, 5)	(2.02, 1.44, 12.02, 24.29) (1, 1, 1, 2, 4)	(2.12, 1.54, 12.73, 26.95) (1, 1, 1, 2, 5)	(2.23,1.65, 12.42, 27.66) (1, 1, 1, 2, 5)	(1.78, 1.18, 14.59, 25.99) (1, 1, 1, 1, 4)
1.00	(1.71, 1.10, 9.64, 16.50) (1, 1, 1, 2, 3)	(1.84, 1.24, 9.00, 10.79) (1, 1, 1, 1, 2)	(1.45, 0.81, 11.25, 16.32) (1, 1, 1, 1, 3)	(1.34, 0.68, 12.97, 17.43) (1, 1, 1, 1, 2)	(1.35, 0.69, 13.08, 17.73) (1, 1, 1, 1, 2)	(1.38, 0.73, 12.40, 17.17) (1, 1, 1, 1, 2)	(1.22, 0.51, 14.92, 18.14) (1, 1, 1, 1, 2)
1.20	(1.26, 0.57, 9.84, 12.35) (1, 1, 1, 1, 2)	(1.30, 0.62, 8.67, 9.23) (1, 1, 1,1, 1)	(1.14, 0.40, 10.11, 12.71) (1, 1, 1, 1, 1)	(1.12, 0.36, 12.95, 14.47) (1, 1, 1, 1, 1)	(1.11, 0.34, 13.18, 14.57) (1, 1, 1, 1, 1)	(1.11, 0.34, 11.97, 13.24) (1, 1, 1, 1, 1)	(1.05, 0.23, 14.54, 15.27) (1, 1, 1, 1, 1)
1.40	(1.09, 0.31, 9.67, 10.50) (1, 1, 1, 1, 1)	(1.09, 0.32, 8.00, 8.14) (1, 1, 1, 1, 1)	(1.04, 0.20, 10.36, 10.76) (1, 1, 1, 1, 1)	(1.04, 0.20, 11.96, 12.43) (1, 1, 1, 1, 1))	(1.03, 1.930.18, 12.93, 13.31) (1, 1, 1, 1, 1)	(1.02, 0.15, 11.10, 11.35) (1, 1, 1, 1, 1)	(1.01, 0.1, 13.47, 13.60) (1, 1, 1, 1, 1)
1.60	(1.03, 0.17, 9.15, 9.40) (1, 1, 1, 1, 1)	(1.02, 0.16, 7.16, 7.19) (1, 1, 1, 1, 1)	(1.01, 0.1, 9.12, 9.21) (1, 1, 1, 1, 1)	(1.01, 0.11, 10.31, 10.45) (1, 1, 1, 1, 1)	(1.01, 0.09, 12.31, 12.42) (1, 1, 1, 1, 1)	(1.00, 0.06, 9.86, 9.90) (1, 1, 1, 1, 1)	(1.00, 0.04, 11.82, 11.84) (1, 1, 1, 1, 1)
1.80	(1.01, 0.09, 8.32, 8.40) (1, 1, 1, 1, 1)	(1.00, 0.07, 6.35, 6.37) (1, 1, 1, 1, 1)	(1.00, 0.05, 7.77, 7.79) (1, 1, 1, 1, 1)	(1.00, 0.06, 8.52, 8.55) (1, 1, 1, 1, 1)	(1.00, 0.05, 11.32, 11.36) (1, 1, 1, 1, 1)	(1.00, 0.02, 8.45, 8.47) (1, 1, 1, 1, 1)	(1.00, 0.03, 9.89, 9.90) (1, 1, 1, 1, 1)
2.00	(1.00, 0.05, 7.31, 7.32) (1, 1, 1, 1, 1)	(1.00, 0.03, 5.73, 8.73) (1, 1, 1, 1, 1)	(1.00, 0.02, 6.61, 6.61) (1, 1, 1, 1, 1)	(1.00, 0.03, 7.00, 7.00) (1, 1, 1, 1, 1)	(1.00, 0.03, 10.02, 10.03) (1, 1, 1, 1, 1)	(1.00, 0.01, 7.14, 7.14) (1, 1, 1, 1, 1)	(1.00, 0.01, 8.07, 8.07) (1, 1, 1, 1, 1)
2.50	(1.00, 0.01, 4.78, 4.78) (1, 1, 1, 1, 1)	(1.00, 0.00, 5.07, 5.07) (1, 1, 1, 1, 1)	(1.00, 0.00, 5.20, 5.203) (1, 1, 1, 1, 1)	(1.00, 0.00, 5.24, 5.236) (1, 1, 1, 1, 1)	(1.00, 0.01, 6.36, 6.36) (1, 1, 1, 1, 1)	(1.00, 0.00, 5.33, 5.33) (1, 1, 1, 1, 1)	(1.00, 0.00, 5.48, 5.49) (1, 1, 1, 1, 1)
AEQL	29.84	30.99	27.41	26.01	26.48	26.90	25.08
(L_1, L, L_2)	(0.6740, 3.5671, 3.0013)	(0.8406, 3.4148, 3.0521)	(1.1496, 3.6358, 2.9798)	(1.3342, 3.5663, 2.9306)	(0.3486, 3.8211, 2.9827)	(0.3185, 3.8526, 3.0049)	(0.6045, 3.8868, 2.9861)
(n, n_1, n_2)	(7, 3, 8)	(7, 5, 5)	(7, 5, 8)	(7, 5, 11)	(11, 3, 11)	(11, 5, 8)	(11, 5, 11)



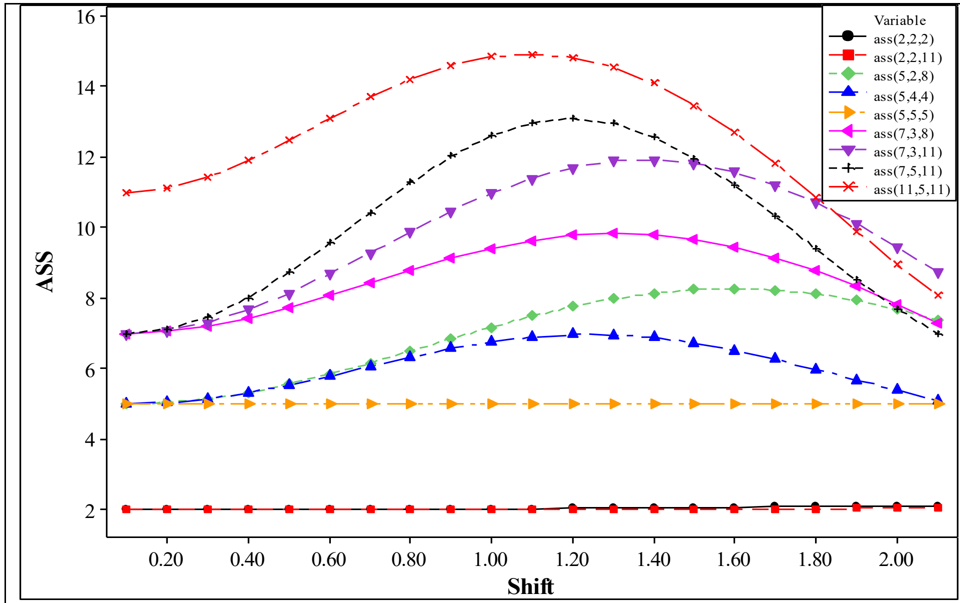
(a) for different triplets (n, n_1, n_2)

(b) when $n_2 = 11$ with different n and n_1 sample sizes

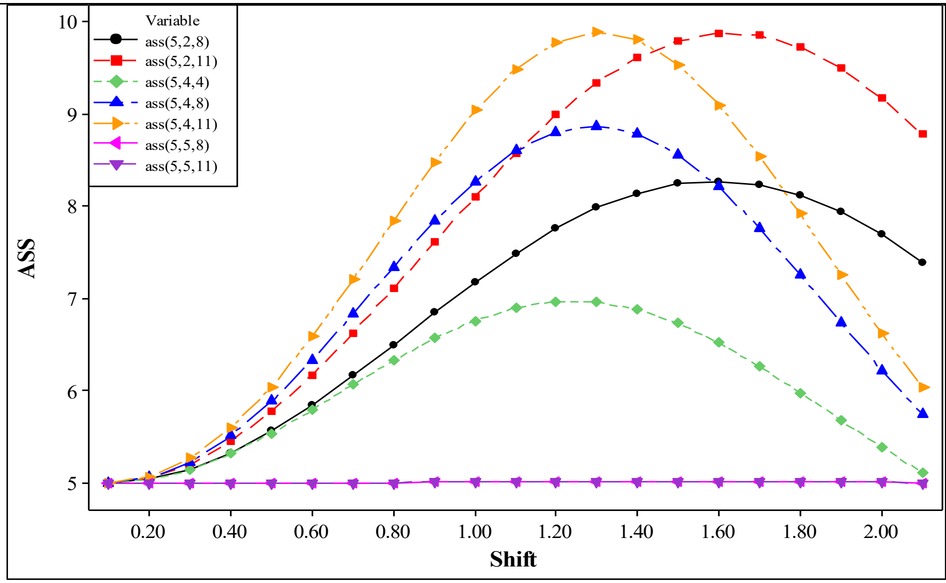
(c) when $n = 5$ with different n_1 and n_2 sample sizes

(d) when $n_1 = 5$ with different n and n_2 sample sizes

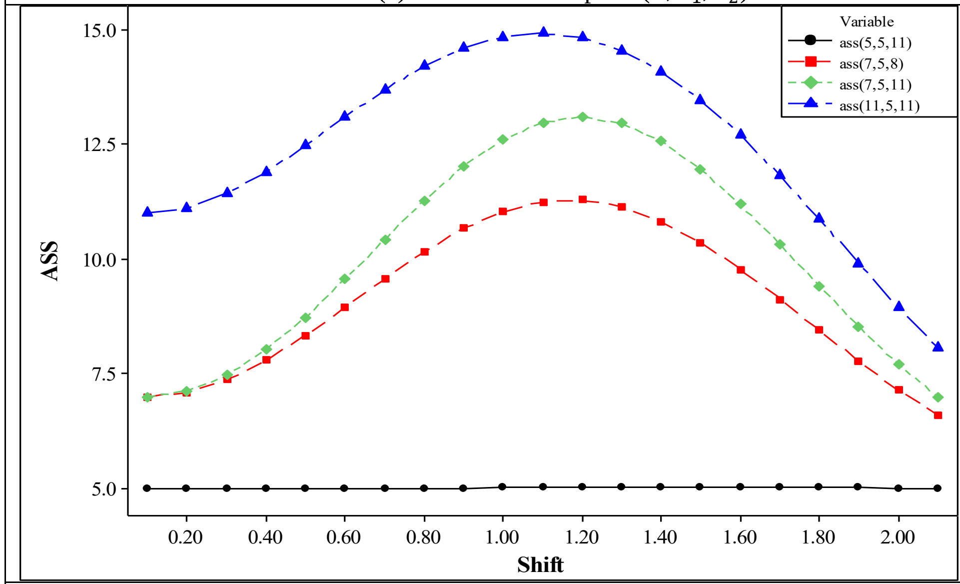
Figure 3.3: *ARL* values of the SSDS $\bar{X}$ scheme for different sample sizes



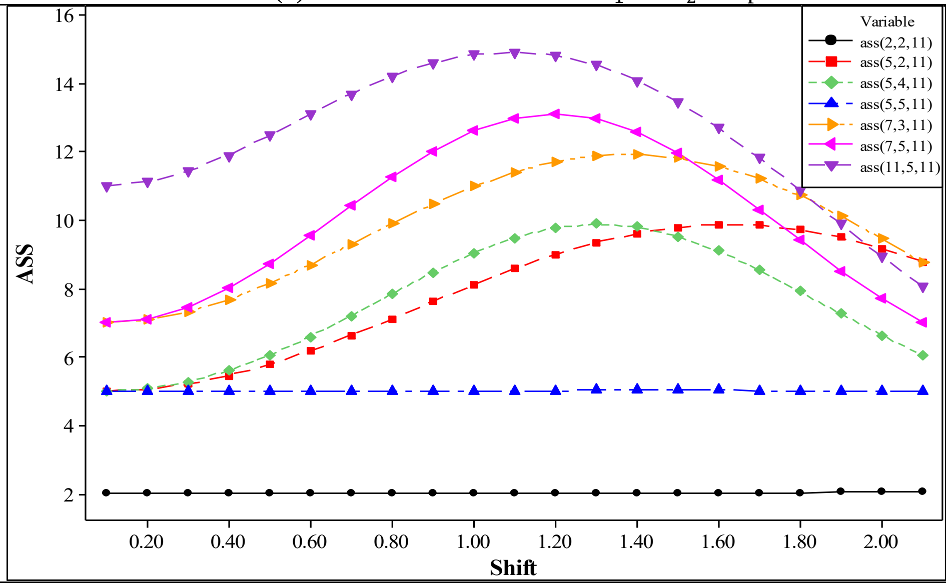
(a) for different triplets (n, n_1, n_2)



(b) when $n = 5$ with different n_1 and n_2 sample sizes

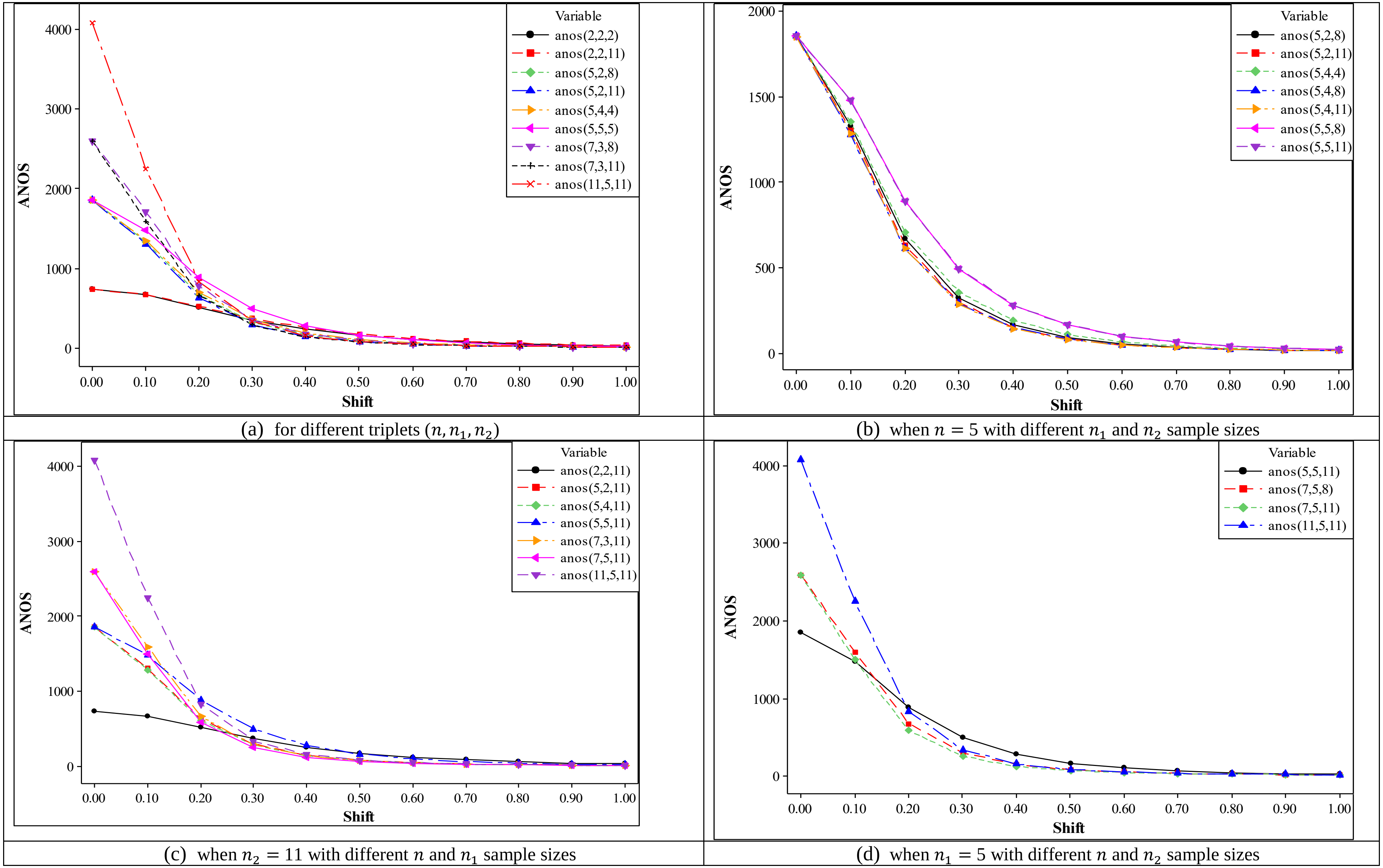


(c) when $n_1 = 5$ with different n and n_2 sample sizes



(d) when $n_2 = 11$ with different n and n_1 sample sizes

Figure 3.4: *ASS* values of the SSDS $\bar{X}$ scheme for different sample sizes



(a) for different triplets (n, n_1, n_2)

(b) when $n = 5$ with different n_1 and n_2 sample sizes

(c) when $n_2 = 11$ with different n and n_1 sample sizes

(d) when $n_1 = 5$ with different n and n_2 sample sizes

Figure 3.5: *ANOS* values of the SSDS $\bar{X}$ scheme for different sample sizes

Figures 3.5 (a) to (d) confirm that for small values of n, the SSDS scheme is cost effective. For instance, in terms of the *ANOS* values, when $\delta = 0.2$, $(n_1, n_2) = (2, 11)$ and n = 2, the SSDS scheme signals for the first time on either the 524th or the 525th observation. When $(n_1, n_2) = (2, 11)$ and n = 5, the SSDS scheme signals for the first time on the either the 634th or the 635th observation. This confirms that the SSDS $\bar{X}$ scheme is cost effective for small values of n.

3.4.3 Performance comparison

In this section, the SSDS $\bar{X}$ scheme is compared to seven well-known monitoring schemes, namely, the traditional $\bar{X}$, NSS synthetic $\bar{X}$, SSS synthetic $\bar{X}$, VSS $\bar{X}$, $\bar{X}$-EWMA (λ), $\bar{X}$-CUSUM and NSSDS $\bar{X}$ schemes in terms of the $AEQL$, $ARARL$ and PCI values. For a fair comparison, these performance measures are computed when $n \in \{4, 7\}$, $n_1 \in \{2, 4\}$, n_2 = 12 and δ_{max} = 2.5 with $NARL_0$ = 370.4 for each scheme. The performance of the $\bar{X}$-EWMA (λ) was investigated for $\lambda = 0.1$ and 0.5. However, the synthetic schemes were investigated such that $H = 1$ where H is a non-zero positive integer representing the control limit of the conforming run-length (CRL) sub-chart. Each competing scheme was optimized by minimizing the ARL_δ values resulting in minimum *AEQL* values.

From Table 3.4, it can be seen that for small ASS_0 values, in terms of the *AEQL*, *ARARL* and PCI values, when the n_1 (i.e. Stage 1 sample size) increases, the proposed SSDS $\bar{X}$ scheme becomes less sensitive than the NSS DS $\bar{X}$ scheme (i.e., its *AEQL* values are greater than the ones of the NSS DS $\bar{X}$ scheme – this is also indicated by the $ARARL$ as well as the PCI values of the NSS DS $\bar{X}$ scheme, which are less than 1) except for large shifts in the process location where the performance of these two schemes are similar. For small Stage 1 sample sizes, the SSDS and NSS DS $\bar{X}$ schemes are almost equivalent. In addition, for small shifts, the $\bar{X}$-EWMA and VSS $\bar{X}$ schemes outperform the SSDS $\bar{X}$ scheme and the $\bar{X}$-CUSUM scheme is as much sensitive as the SSDS $\bar{X}$ scheme. However, for large ASS_0 values, the proposed SSDS $\bar{X}$ scheme outperforms all competing schemes regardless of the size of the shift in the process location (i.e. it has the smallest $AEQL$ value).

Table 3.4: Case K monitoring schemes performance comparison for different shift sizes when $n \in \{4,7\}$, $n_1 \in\{2,4\}$, n_2=12 and δ_{max}=2.5 with $NARL_0$ = 370.4

*Shift	Performance measures	$\bar{X}$	VSS $\bar{X}$	NSS Synthetic $\bar{X}$	SSS Synthetic $\bar{X}$	$\bar{X}$-EWMA(0.1)	$\bar{X}$-EWMA(0.5)	$\bar{X}$-CUSUM	NSS DS $\bar{X}$	SSDS $\bar{X}$	(n_1, n_2)	$ASS_0 = n$
Small	***AEQL***	152.47	53.48	147.82	104.95	55.89	138.16	61.43	**42.35**	61.19	**(4, 12)**	**4**
	ARARL	2.42	0.92	2.19	1.70	0.95	2.09	1.04	**0.75**	1.00		
	PCI	2.49	0.87	2.41	1.72	0.91	2.26	1.00	**0.69**	1.00		
Moderate	***AEQL***	148.45	45.21	151.14	109.52	44.78	79.64	77.53	**18.35**	20.36		
	ARARL	8.89	2.12	8.54	7.50	2.22	3.92	3.88	**0.91**	1.00		
	PCI	7.29	2.22	7.42	5.38	2.20	3.91	3.80	**0.90**	1.00		
Large	***AEQL***	34.83	32.23	106.59	91.86	37.23	46.45	82.87	**17.06**	17.25		
	ARARL	2.11	1.89	9.54	8.77	2.20	2.77	4.86	**0.99**	1.00		
	PCI	2.02	1.87	6.18	5.33	2.16	2.69	4.80	**0.99**	1.00		
Small-to-large	***AEQL***	77.47	73.43	135.95	102.59	81.54	136.10	149.69	**40.66**	47.42		
	ARARL	1.87	1.78	7.18	6.36	1.82	2.96	3.34	**0.88**	1.00		
	PCI	1.63	1.73	2.86	2.16	1.72	2.87	3.16	**0.86**	1.00		
Small	***AEQL***	152.47	49.22	147.82	104.95	55.89	138.16	61.43	38.84	**36.83**	**(2, 12)**	
	ARARL	2.49	1.42	3.34	1.83	1.53	3.59	1.71	1.04	**1.00**		
	PCI	4.14	1.34	4.01	2.85	1.52	3.75	1.67	1.05	**1.00**		
Moderate	***AEQL***	148.45	41.34	151.14	109.52	44.78	79.64	77.53	**14.25**	14.35		
	ARARL	9.10	2.32	9.16	8.17	3.14	5.67	5.38	**0.99**	1.00		
	PCI	10.34	2.88	10.54	7.63	3.12	5.55	5.40	**0.99**	1.00		
Large	***AEQL***	34.83	30.75	106.59	91.86	37.23	46.45	82.87	**17.11**	17.18		
	ARARL	2.23	1.46	10.08	9.11	2.23	2.80	4.90	**1.00**	1.00		
	PCI	2.03	1.79	6.21	5.44	2.17	2.70	4.84	**1.00**	1.00		
Small-to-large	***AEQL***	77.47	70.41	135.95	102.59	81.54	136.10	149.69	37.11	**36.67**		
	ARARL	1.91	1.62	7.65	6.79	2.33	4.04	4.08	1.01	**1.00**		
	PCI	2.11	1.92	3.70	2.80	2.21	3.71	4.08	1.01	**1.00**		
Small	***AEQL***	141.03	49.79	97.19	91.98	55.71	138.36	61.35	28.29	**24.91**	**(4, 12)**	**7**
	ARARL	2.31	1.89	2.70	2.60	2.25	5.47	2.55	12.89	**1.00**		
	PCI	5.66	2.00	3.90	3.69	2.24	5.55	2.46	1.14	**1.00**		
Moderate	***AEQL***	112.44	41.24	103.33	97.06	45.47	79.57	77.58	11.22	**10.09**		
	ARARL	8.62	2.13	8.43	7.98	4.64	8.39	7.75	1.44	**1.00**		
	PCI	11.14	4.09	10.24	9.62	4.51	7.89	7.69	1.11	**1.00**		
Large	***AEQL***	33.56	30.33	91.82	84.99	37.18	46.46	82.78	16.39	**16.15**		
	ARARL	2.09	1.81	6.29	5.59	2.40	3.04	5.28	1.07	**1.00**		
	PCI	2.08	1.88	5.69	5.26	2.30	2.88	5.13	1.01	**1.00**		
Small-to-large	***AEQL***	73.19	69.46	97.93	91.78	81.88	136.13	149.61	31.50	**29.58**		
	ARARL	2.72	2.32	6.52	6.09	3.13	5.64	5.29	1.32	**1.00**		
	PCI	2.47	2.34	3.31	3.11	2.77	4.60	5.08	1.26	**1.00**		
Small	***AEQL***	141.03	46.77	97.19	91.98	55.71	138.36	61.35	42.35	**31.00**	**(2, 12)**	
	ARARL	4.45	1.91	2.57	1.78	1.85	4.44	2.09	1.37	**1.00**		
	PCI	4.54	1.51	3.14	2.97	1.80	4.46	1.98	1.37	**1.00**		
Moderate	***AEQL***	112.44	39.43	103.33	97.06	45.47	79.57	77.58	18.35	**11.45**		
	ARARL	9.44	3.32	8.72	8.42	4.05	7.24	6.79	1.65	**1.00**		
	PCI	9.82	3.44	9.02	8.49	3.97	6.94	6.78	1.60	**1.00**		
Large	***AEQL***	33.56	29.05	91.82	84.99	37.18	46.46	82.78	17.06	**16.41**		
	ARARL	2.01	1.76	8.65	8.01	2.35	2.97	5.17	1.05	**1.00**		
	PCI	2.04	1.77	5.59	5.18	2.67	2.83	5.04	1.04	**1.00**		
Small-to-large	***AEQL***	73.19	63.25	97.93	91.78	81.88	136.13	149.61	40.66	**32.41**		
	ARARL	2.49	2.12	5.99	5.58	2.78	4.90	4.78	1.36	**1.00**		
	PCI	2.26	1.95	3.02	2.83	2.53	4.20	4.62	1.25	**1.00**		

* Small: (0 < $\delta \leq$ 0.7), Moderate: (0.7 < $\delta \leq$ 1.6), Large (1.6 < $\delta \leq$ 2.5) and Small-to-large: (0 < $\delta \leq$ 2.5)

3.5 Illustrative example

3.5.1 Example 1

To illustrate the implementation and application of the proposed SSDS $\bar{X}$ scheme, the well-known dataset from Montgomery (2013) on the inside diameters of piston rings manufactured by a forging process is considered. This data set contains 25 retrospective or Phase I samples, each of size five, that were collected when the process was thought to be IC. A goodness of fit test for normality reveals that the data are normally distributed. For this data, the process parameters, μ_0 and σ_0 are given by 74.001 and 0.008, respectively. This data set also contains 75 Phase II observations (i.e. 15 sub-groups of size 5). Therefore, we consider the SSDS $\bar{X}$ scheme with (n_1, n_2) = (5, 5) and an ASS_0 of 5. The optimal combination (n_1, n_2, L_1, L, L_2) when ASS_0 = 5 is found to be equal to (5, 5, 3.0, 3.0, 3.0) so that the ARL_0 = 370.4 with $AEQL$ = 49.54. For a fair comparison with the Daudin (1992)'s NSSDS $\bar{X}$ scheme, we also considered the NSSDS $\bar{X}$ scheme with (n_1, n_2) = (5, 5) and an ASS_0 of 5. The optimal combination for this scheme is given by (5, 5, 2.51, 3.221, 2.752) so that the ARL_0 = 370.4 with $AEQL$ = 52.46. The standardized statistics for the first sample and combined samples are Z_{1t} and Z_t, respectively. These values are computed using μ_0 and σ_0.

A plot of the charting statistics Z_{1t} and Z_t for both schemes is shown in Figure 3.6. It can be seen that from the first to the 13th sampling time (i.e. t = 1 to 13) of the first sample, the charting statistics Z_{1t} of the SSDS $\bar{X}$ scheme plotted in region A ($L_1 = L$). Therefore, at this stage, the process was IC. At the 14th sampling time, the charting statistic ($Z_{1,14}$) plotted below -L (i.e. below -3). Thus, the SSDS $\bar{X}$ scheme gives a signal for the first time on the 14th sampling time. It was observed that for the NSSDS $\bar{X}$ scheme, at the first and second sampling times of the first sample, the charting statistics ($Z_{1,1}$ and $Z_{1,2}$) plotted in region B^+ and their corresponding stage 2 charting statistics Z_1 and Z_2 plotted in the IC region F^-; thus, the process is IC at this stage. On the twelfth sampling time of the first sample, (i.e., $Z_{1,12}$) plots in region B^- and its corresponding stage 2 charting statistic Z_{12} plots in region G^+, which means that the process is IC. From the fourteenth sampling time of the first sample onwards, all of the charting statistics plotted in the IC regions. Hence, the NSSDS $\bar{X}$ scheme does not signal. Therefore, the SSDS $\bar{X}$ scheme outperforms the NSSDS $\bar{X}$ scheme. These findings confirm the results found in Section 3.4.3. Note that since we needed to go to stage 2 on three different sampling times (using a second sample) at times t =1, 2 and 12, the NSSDS scheme has 22 sampling times, instead of 25.

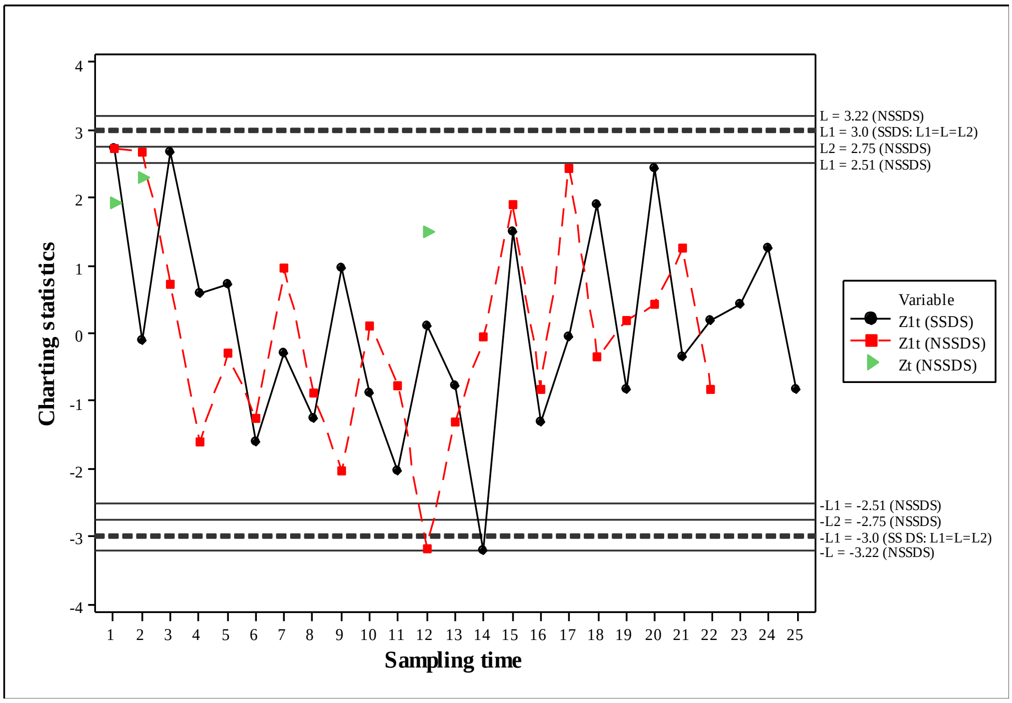


Figure 3.6: The SSDS and NSSDS $\bar{X}$ schemes for the piston ring data

3.5.2 Example 2

In addition to the example in Section 3.5.1, another dataset is used to illustrate the implementation and application of the proposed SSDS $\bar{X}$ scheme. This dataset is based on the hard-bake process which is used in conjunction with photolithography in semiconductor manufacturing – see Chapter 6 of Montgomery (2013). The interval of time between samples or subgroups is modified such that master samples of size 10 are taken every two hours to monitor the flow width of the resist, see Table 3.5. A goodness of fit test for normality reveals that the data are normally distributed. From prior information, the process parameters, μ_0 and σ_0 are known to be 1.5056 microns and 0.1398 microns, respectively.

Table 3.5: Dataset on the hard-bake process

	Stage 1 observations		**Stage 2 observations**							
t	Y_{1t1}	Y_{1t2}	Y_{2t1}	Y_{2t2}	Y_{2t3}	Y_{2t4}	Y_{2t5}	Y_{2t6}	Y_{2t7}	Y_{2t8}
1	1.4483	1.5458	1.4538	1.4303	1.6206	1.5435	1.6899	1.5830	1.3358	1.4187
2	1.5175	1.3446	1.4723	1.6657	1.6661	1.5454	1.0931	1.4072	1.5039	1.5264
3	1.4418	1.5059	1.5124	1.4620	1.6263	1.4301	1.2725	1.5945	1.5397	1.5252
4	1.4981	1.4506	1.6174	1.5837	1.4962	1.3009	1.506	1.6231	1.5831	1.6454
5	1.4132	1.4603	1.5808	1.7111	1.7313	1.3817	1.3135	1.4953	1.4894	1.4596
6	1.5765	1.7014	1.4026	1.2773	1.4541	1.4936	1.4373	1.5139	1.4808	1.5293
7	1.5729	1.6738	1.5048	1.5651	1.7473	1.8089	1.5513	1.8250	1.4389	1.6558
8	1.6236	1.5393	1.6738	1.8698	1.5036	1.412	1.7931	1.7345	1.6391	1.7791
9	1.7372	1.5663	1.4910	1.7809	1.5504	1.5971	1.7394	1.6832	1.6677	1.7974
10	1.4295	1.6536	1.9134	1.7272	1.4370	1.6217	1.822	1.7915	1.6744	1.9404

In this example, we consider the SSDS $\bar{X}$ scheme with (n_1, n_2) = (2, 8) and an ASS_0 of 5. The optimal combination (n_1, n_2, L_1, L, L_2) when ASS_0 = 5 is found to be equal to (2, 8, 0.8856, 3.3526, 3.0085) so that the attained ARL_0 = 370.4 – see Table 3.1.

Table 3.6: Illustration of the operation of the SSDS $\bar{X}$ schemes using the dataset on the hard-bake process

t	$\bar{Y}_{1t}$	Z_{1t}	Region	Take a 2nd Sample	$\bar{Y}_{2t}$	$\bar{Y}_t$	Z_t	Region	Stage 1: Signal	Stage 2: Signal
1	1.4971	-0.0865	A	No					No	
2	1.4311	-0.7542	A	No					No	
3	1.4739	-0.3212	A	No					No	
4	1.4744	-0.3162	A	No					No	
5	1.4368	-0.6965	A	No					No	
6	1.6390	1.3489	B^+	Yes	1.4486	1.4867	-0.4281	F^-	No	No
7	1.6234	1.1911	B^+	Yes	1.6371	1.6344	2.9129	F^-	No	No
8	1.5815	0.7673	A	No					No	
9	1.6518	1.4784	B^+	Yes	1.6634	1.6611	3.5164	F^+	No	**Yes**
10	1.5416	0.3636	A	No					No	

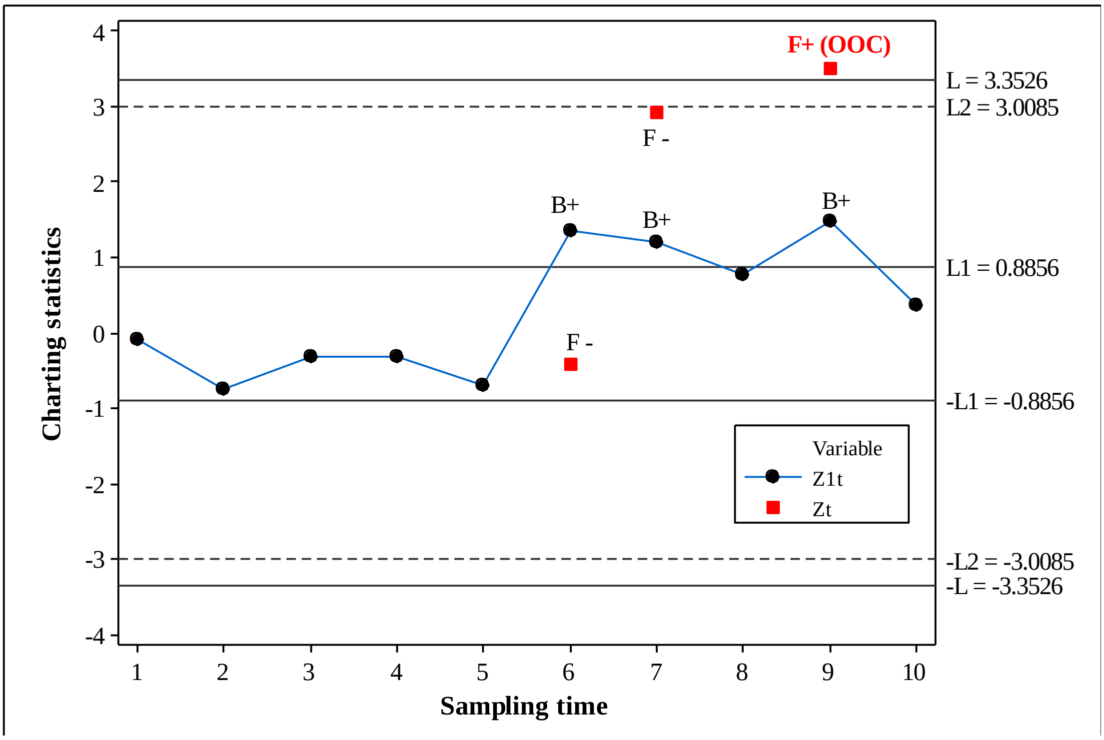


Figure 3.7: The SSDS $\bar{X}$ scheme for the hard-bake process data

Given the illustration in Table 3.6 and Figure 3.7, it is observed that the process does not give an OOC signal at Stage 1. However, the process goes to Stage 2 at the 6th, 7th and 9th sampling time (see the red dots in Figure 3.7); and eventually issue an OOC signal on the 9th subgroup for the first time according to the implementation operation in Figure 3.2.

3.6 Conclusion

In this chapter, a new design of the parametric monitoring scheme based on double sampling is proposed. At the second stage, the proposed SSDS scheme gives a signal if the charting statistic of the first sample and the combined samples plot on one side of the charting regions in each stage (i.e., side-sensitive). The performance of the proposed SSDS scheme with known process parameters was investigated in terms of the characteristics of the run-length distribution, the ASS, the $ANOS$ and the $AEQL$ function. It was observed that for large ASS_0 values, the SSDS $\bar{X}$ scheme outperforms all competing schemes considered in this chapter regardless of the size of the mean shifts. However, for small ASS_0 values and large Stage 1 sample sizes, the NSS DS, EWMA and VSS $\bar{X}$ schemes outperform the proposed SSDS $\bar{X}$ scheme for small shifts in the process location. In terms of the *ASS*, the proposed SSDS scheme is cost effective when compared to the NSSDS and VSS $\bar{X}$ schemes (keep in mind that, in Teoh et al. (2014a), it was concluded that the VSS $\bar{X}$ scheme has a better *ASS* performance than the NSSDS $\bar{X}$ scheme). Note though, the proposed SSDS scheme tends to yield better OOC performance than its competitors in many situations.

However, users / companies that do not face problems involving large sample sizes are advised to use large sample sizes regardless of the shift size of interest, as this guarantees a much better performance. If the sample size is a major concern for a user / company, we recommend the use of small sample sizes for small and moderate shifts and large sample sizes for large shifts. We also suggest that companies use $ASS_0 = n_1 =$ 3, 4 or 5, and increase n_2 considerably in order to reach the desired efficiency.

Chapter 4. A new side-sensitive double sampling $\bar{X}$ scheme for monitoring an abrupt change in the process location with estimated parameters

4.1 Introduction

A literature review of all currently available double sampling schemes in SPM have been reviewed in Chapter 2. To ensure that this chapter is self-contained, there will be some few key concepts and figures that will be reproduced from Chapter 2.

Although there is a lots of research work based on parameters known (i.e. Case K), in practice, the process parameters are generally unknown (i.e. Case U). As stated in Chapter 2, there are a fewer number of articles that investigated the Case U scenario in the context of NSSDS schemes as compared to Case K (see Table 2.1 in Chapter 2). These are: Khoo et al. (2013b), Teoh et al. (2013, 2014, 2015, 2016a, 2016b), Castagliola et al. (2017), You et al. (2015), You (2018) and, Lee and Khoo (2019c). It is worth noting that all the articles on Case U NSSDS schemes mentioned in the previous sentence, used the non-side-sensitive design discussed in Daudin (1992); see these charting regions design in Figure 2.1(b). In Chapter 3 of this dissertation, it is shown that the SSDS $\bar{X}$ scheme has an improved performance over the NSSDS $\bar{X}$ scheme when Case K is assumed. Since the abovementioned articles have thoroughly discussed the design and implementation of the NSSDS $\bar{X}$ scheme. Therefore, in this chapter, the design and implementation of the SSDS $\bar{X}$ scheme in Case U using the (5$^{\text{th}}$, 25$^{\text{th}}$, 50$^{\text{th}}$, 75$^{\text{th}}$, 95$^{\text{th}}$) percentiles, ARL, $SDRL$, ASS and the average extra quadratic loss ($AEQL$) metrics are discussed.

The rest of this chapter is structured as follows: Section 4.2 presents the operation, design and run-length properties of the SSDS $\bar{X}$ scheme in Case U; while, Section 4.3 presents the overall performance run-length metrics. The optimization model is given in Section 4.4. Section 4.5 assesses the IC, OOC performance of the SSDS $\bar{X}$ scheme and compares their overall performance with the NSSDS $\bar{X}$ scheme and other established monitoring schemes in Case U. In Section 4.6, a case study is given using real-life data to demonstrate the implementation and design of the Case U SSDS $\bar{X}$ scheme. Finally, some concluding remarks are given in Section 4.7.

4.2 Design of the SSDS $\bar{X}$ monitoring scheme with estimated process parameters

The Case U scenario requires monitoring schemes to be applied in a two-phase approach, i.e. Phase I and Phase II (see Jensen et al. (2006) and Psarakis et al. (2013) for a review of parameter estimation effect publications). In Phase I, monitoring schemes are implemented retrospectively in order to estimate the distribution parameters using an IC reference sample. However, in Phase II, monitoring schemes are implemented prospectively to continuously monitor any departures from an IC state using the parameters estimated in Phase I.

4.2.1 Phase I and Phase II operation of the SSDS $\bar{X}$ monitoring scheme

Phase I parameter estimation

Since the IC process parameters, μ_0 and σ_0, are usually unknown they have to be estimated from *m* Phase I subgroup samples, each of size n, i.e. $\{X_{ij}\}_{j=1,2,\dots,n}^{i=1,2,\dots,m}$. The estimated IC process parameters, $\hat{\mu}_0$ and $\hat{\sigma}_0$, are given by

$$\hat{\mu}_0 = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij} \tag{4.1}$$

and

$$\hat{\sigma}_0 = \sqrt{\frac{1}{m(n-1)}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(X_{ij}-\bar{X}_i\right)^2}, \tag{4.2}$$

respectively, where $\bar{X}_i = \sum_{j=1}^{n} X_{ij}/n$.

Phase II charting statistics and operation procedure: Stages 1 and 2

Let Y_{tj} be the Phase II observations from i.i.d. $N(\mu_1\,, \sigma_0)$, where μ_1 is the OOC mean (i.e. $\mu_1 = \mu_0 + \delta\sigma_0$) with $\delta = |\mu_1 - \mu_0|/\sigma_0$ the magnitude of the standardized mean shift from μ_0 to μ_1. Let L_1 and L (with $L \geq L_1 > 0$) be the warning and control limits of the first sample of size n_1 at Stage 1, respectively; and L_2 (with $L_2 > 0$) be the control limit of the combined samples of size n $(= n_1 + n_2)$ at Stage 2. Therefore, the SSDS $\bar{X}$ control scheme is divided into eight intervals, i.e. A = $[-L_1, L_1]$, B$^+$ = $(L_1, L]$, B$^-$ = $[-L, -L_1)$, C =$(-\infty, -L) \cup (L, +\infty)$, F$^+$ = $(L_2, +\infty)$, F$^-$ = $(-\infty, L_2]$, G$^-$ = $(-\infty, -L_2)$ and G$^+$ = $[-L_2, +\infty)$.

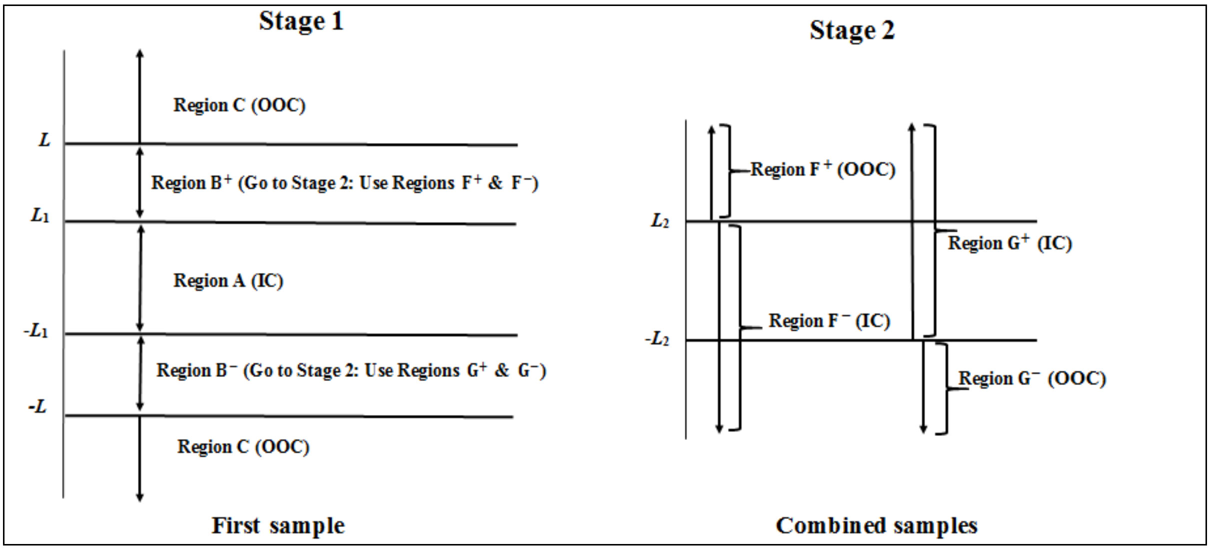


Figure 4.1: The charting regions of the SSDS scheme

In Phase II of the SSDS $\bar{X}$ monitoring scheme, there are two distinct standardized charting statistics in Case U (denoted as $\hat{Z}_{1t}$ and $\hat{Z}_{2t}$, shown below) used during Stage 1 and 2, respectively (see Figure 4.1). From the Y_{tj} observations, a first subgroup sample of size n_1 is collected at the t^{th} sampling time (denoted as Y_{1tj}, $t = 1, 2, \ldots$, and $j = 1, 2, \ldots, n_1$). If the standardized charting statistic based on the first sample falls on Region B$^-$ or B$^+$, then a second subgroup sample of size n_2 (where $n_2 \geq n_1$) is also collected at the t^{th} sampling time (denoted as Y_{2tj}, $t = 1, 2, \ldots$, and $j = 1, 2, \ldots, n_2$). Then the SSDS $\bar{X}$ scheme uses these two stages to decide whether the process is IC or OOC, and each stage's charting statistic is as follows.

Stage 1: Let $\bar{Y}_{1t} = \sum_{j=1}^{n_1} Y_{1tj}/n_1$ be the mean of the first sample of subgroup size n_1 at the t^{th} sampling time. Hence, the standardized statistic for the first sample at the t^{th} sampling time in Case U is given by

$$\hat{Z}_{1t} = \frac{\bar{Y}_{1t} - \hat{\mu}_0}{\hat{\sigma}_0/\sqrt{n_1}}. \tag{4.3}$$

where $\bar{Y}_{1t} \sim N(\mu_0 + \delta\sigma_0, \frac{\sigma_0}{\sqrt{n_1}})$, so that $\delta = 0$ means that the process is IC. In this case, Z_{1t} follows a standard normal distribution (i.e. $Z_{1t} \sim N(0,1)$). However, when $\delta \neq 0$, the process is OOC and $Z_{1t} \sim N(\delta, 1)$.

Stage 2: At the t^{th} sampling time of the second sample, the sample mean is given by $\bar{Y}_{2t} = \sum_{j=1}^{n_2} Y_{2tj}/n_2$, so that the combined sample mean is given by $\bar{Y}_t = (n_1\bar{Y}_{1t} + n_2\bar{Y}_{2t})/(n_1 + n_2)$. Thus, the standardized charting statistic in Case U at the t^{th} sampling time is given by

$$\hat{Z}_t = \frac{\bar{Y}_t - \hat{\mu}_0}{\hat{\sigma}_0/\sqrt{n_1 + n_2}}. \tag{4.4}$$

where $\bar{Y}_t \sim N(\mu_0 + \delta\sigma_0, \frac{\sigma_0}{\sqrt{n_1+n_2}})$. When the process is IC, $\hat{Z}_{2t} \sim N(0,1)$ since $\delta = 0$ and when the process is OOC, $\hat{Z}_t \sim N(\delta, 1)$.

Thus, based on the description above, the operational procedure of the Case U SSDS $\bar{X}$ scheme is given as follows:

1. From the IC retrospective data with m samples, estimate the IC mean and standard deviation of the process using Equations (4.1) and (4.2), respectively.
2. In the prospective phase, take a sample of size n_1 and calculate the sample mean $\bar{Y}_{1t}$ at the t^{th} sampling time at Stage 1.
3. If $\hat{Z}_{1t} \in$ A, the process is considered as IC.
4. If $\hat{Z}_{1t} \in$ C, the process is said to be OOC and then the necessary corrective action must be taken to find and remove the assignable causes.
5. If $\hat{Z}_{1t} \in \mathrm{B}^- \cup \mathrm{B}^+$, take a second sample of size $n_2 (n_2 \geq n_1)$ and calculate the sample mean $\bar{Y}_{2t}$ at the t^{th} sampling time of the second sample.
6. At the t^{th} sampling time, calculate $\bar{Y}_t$ and then $\hat{Z}_t$.
7. The process is declared IC if:
   (a) If $\hat{Z}_{1t} \in \mathrm{B}^+$ and $\hat{Z}_t \in \mathrm{F}^-$, or
   (b) If $\hat{Z}_{1t} \in \mathrm{B}^-$ and $\hat{Z}_t \in \mathrm{G}^+$.

   However, the process is declared OOC:
   (c) If $\hat{Z}_{1t} \in \mathrm{B}^+$ and $\hat{Z}_t \in \mathrm{F}^+$, or
   (d) If $\hat{Z}_{1t} \in \mathrm{B}^-$ and $\hat{Z}_t \in \mathrm{G}^-$.

In essence, if the plotting statistic falls in region B^+ (region B^-) in Stage 1, then it can only fall in regions F^- or F^+ (regions G^+ or G^-) only, in Stage 2, respectively. Conversely, if in Stage 1, $\hat{Z}_{1t} \in \mathrm{B}^+$, then in Stage 2, we have $\hat{Z}_t \notin \{\mathrm{G}^+, \mathrm{G}^-\}$. Similarly, if in Stage 1, $\hat{Z}_{1t} \in \mathrm{B}^-$, then in Stage 2, we have $\hat{Z}_t \notin \{\mathrm{F}^-, \mathrm{F}^+\}$. The flow chart illustrating the steps involved in the operation of the Case U SSDS $\bar{X}$ monitoring scheme is shown in Figure 4.2.

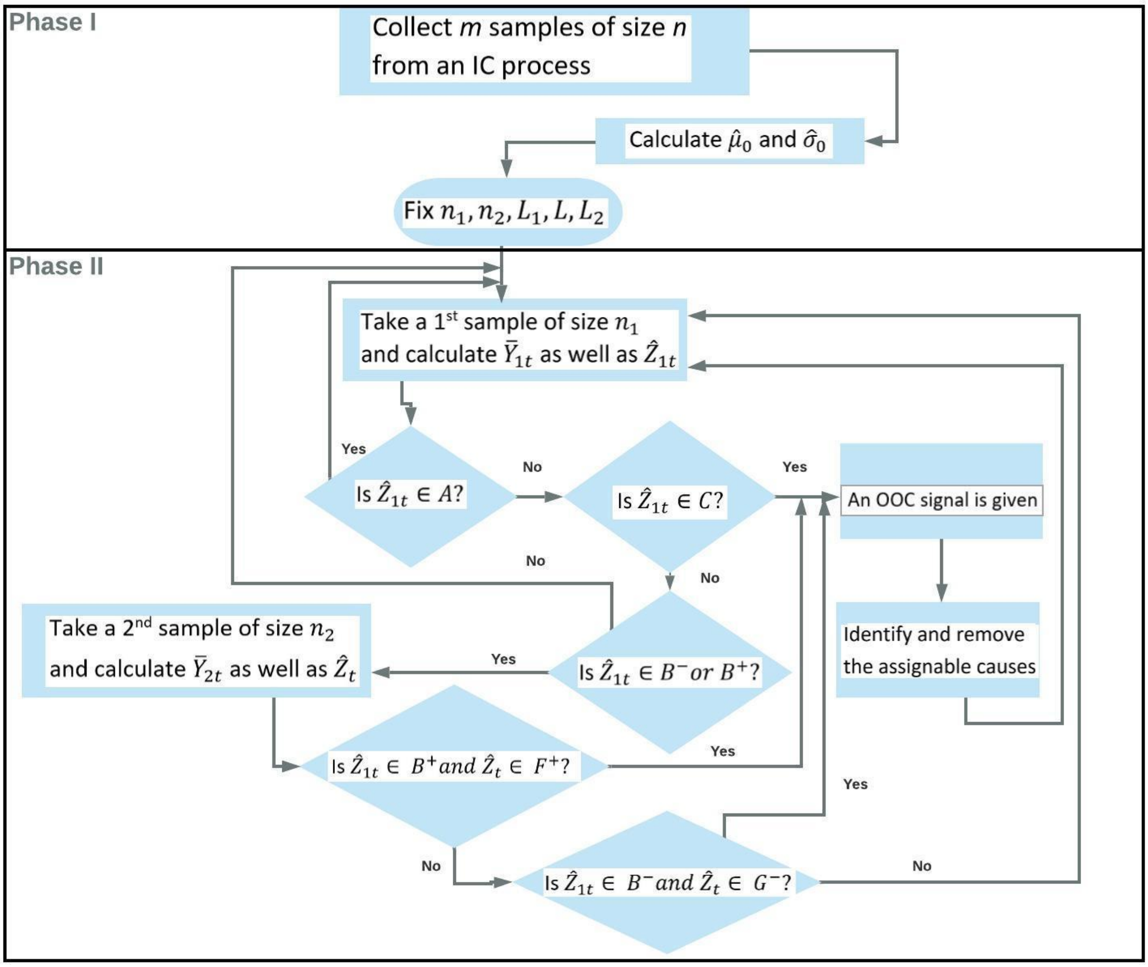


Figure 4.2: Flow chart for the proposed SSDS $\bar{X}$ control scheme in Case U

4.2.2 Unconditional run-length properties of the SSDS $\bar{\mathbf{X}}$ scheme

In order to calculate the unconditional run-length (*RL*) properties, we need to first derive the conditional ones, see Jensen et al. (2006). Hence, the conditional c.d.f. of $\hat{Z}_{1t}$, given $\hat{\mu}_0$ and $\hat{\sigma}_0$ is defined as

$$F_{\hat{Z}_{1t}}(z|\hat{\mu}_0, \hat{\sigma}_0) = \Phi\left(U\sqrt{\frac{n_1}{mn}} + Vz - \delta\sqrt{n_1}\right). \tag{4.5}$$

where $U = (\hat{\mu}_0 - \mu_0)\frac{\sqrt{mn}}{\sigma_0}$ and $V = \hat{\sigma}_0/\sigma_0$. Consequently, the conditional p.d.f. of $\hat{Z}_{1t}$, given $\hat{\mu}_0$ and $\hat{\sigma}_0$ is given by

$$f_{\hat{Z}_{1t}}(z|\hat{\mu}_0, \hat{\sigma}_0) = V\phi\left(U\sqrt{\frac{n_1}{mn}} + Vz - \delta\sqrt{n_1}\right). \tag{4.6}$$

Since $\hat{\mu}_0 \sim N(\mu_0, \frac{\sigma_0^2}{mn})$, then $U \sim N(0,1)$ so that the p.d.f. of the random variable U is simply,

$$f_U(u) = \phi(u). \tag{4.7}$$

Zhang et al. (2011) used the fact that $V^2 = (\hat{\sigma}_0/\sigma_0)^2$ has a gamma distribution with parameters $m(n-1)/2$ and $2/[m(n-1)]$ to show that the p.d.f. of V is defined as

$$f_v(v|m,n) = 2vf_\gamma\left[v^2\middle|\frac{m(n-1)}{2},\frac{2}{m(n-1)}\right], \tag{4.8}$$

where $f_\gamma(.)$ is the p.d.f. of the gamma distribution with parameters $\frac{m(n-1)}{2}$ and $\frac{2}{m(n-1)}$.

Next, to derive the unconditional c.d.f. of the RL of the proposed SSDS scheme, we need to first derive the unconditional probability of the IC process. Let $\hat{P}_{0k}$ denote the probability that the process with estimated parameters remains IC at the sampling stage k (with $k = \{1, 2\}$). Then, the probability that the process is IC is given by

$$\hat{P}_0 = \hat{P}_{01} + \hat{P}_{02} \tag{4.9}$$

where,

$$\begin{aligned}\hat{P}_{01} &= \Phi\left(U\sqrt{\frac{n_1}{mn}} + VL_1 - \delta\sqrt{n_1}\right) - \Phi\left(U\sqrt{\frac{n_1}{mn}} - VL_1 - \delta\sqrt{n_1}\right)\\ \hat{P}_{02} &= \int_{Z\in \mathrm{B}^{++}}^{\cdot} \hat{P}_{\mathrm{F}^-} f_{\hat{Z}_{1t}}(z|\hat{\mu}_0,\hat{\sigma}_0)dz + \int_{Z\in \mathrm{B}^{--}}^{\cdot} \hat{P}_{\mathrm{G}^+} f_{\hat{Z}_{1t}}(z|\hat{\mu}_0,\hat{\sigma}_0)dz\end{aligned} \tag{4.10}$$

with

$$\hat{P}_{\mathrm{F}^-} = \Phi\left[U\sqrt{\frac{n_2}{mn}} + V\left(\frac{L_2\sqrt{n_1+n_2} - z\sqrt{n_1}}{\sqrt{n_2}}\right) - \delta\sqrt{n_2}\right]$$

and

$$\hat{P}_{\mathrm{G}^+} = 1 - \Phi\left[U\sqrt{\frac{n_2}{mn}} - V\left(\frac{L_2\sqrt{n_1+n_2} - z\sqrt{n_1}}{\sqrt{n_2}}\right) - \delta\sqrt{n_2}\right].$$

Then, the unconditional c.d.f. of the SSDS $\bar{X}$ monitoring scheme for Case U is given by

$$F_{RL}(\ell) = \int_{-\infty}^{+\infty}\int_{0}^{+\infty}\left(1 - \hat{P}_0^{\ell}\right)f_U(u)f_V(v)\,dv\,du, \tag{4.11}$$

where $\ell \in \{1,2,3,\dots,\}$, $f_U(u)$ and $f_V(v)$ are defined in Equations (4.7) and (4.8), respectively.

Therefore, the unconditional ARL and $SDRL$ of the proposed SSDS $\bar{X}$ monitoring scheme with estimated process parameters are given by

$$ARL = \int_{-\infty}^{+\infty}\int_{0}^{+\infty}\left(\frac{1}{1-\hat{P}_0}\right)f_U(u)f_V(v)\,dv\,du \tag{4.12}$$

and

$$SDRL = \left[\int_{-\infty}^{+\infty}\int_{0}^{+\infty}\left(\frac{1+\hat{P}_0}{1-\hat{P}_0}\right)f_U(u)f_V(v)\,dv\,du - ARL^2\right]^{1/2}. \tag{4.13}$$

The Case U average sample size (ASS) is given by

$$ASS = \int_{-\infty}^{+\infty}\int_{0}^{+\infty}\left(n_1 + n_2\hat{P}_2\right)f_U(u)f_V(v)\,dv\,du \tag{4.14}$$

where $\hat{P}_2$ is the probability of taking the second sample, which is given by $\hat{P}_2 = P\left(\hat{Z}_{1t} \in B^- \cup B^+ \middle| \hat{\mu}_0, \hat{\sigma}_0\right)$, or simply,

$$\begin{aligned}\hat{P}_2 = \Phi\left(U\sqrt{\frac{n_1}{mn}} + VL - \delta\sqrt{n_1}\right) - \Phi\left(U\sqrt{\frac{n_1}{mn}} + VL_1 - \delta\sqrt{n_1}\right) \\ + \Phi\left(U\sqrt{\frac{n_1}{mn}} - VL_1 - \delta\sqrt{n_1}\right) - \Phi\left(\mathrm{U}\sqrt{\frac{n_1}{mn}} - VL - \delta\sqrt{n_1}\right).\end{aligned} \tag{4.15}$$

Then, the average number of observations to signal ($ANOS$) is given by

$$ANOS = \int_{-\infty}^{+\infty}\int_{0}^{+\infty}\left(n_1 + n_2\hat{P}_2\right)\left(\frac{1}{1-\hat{P}_0}\right)f_U(u)f_V(v)\,dv\,du. \tag{4.16}$$

Since the $ANOS$ depends on the ASS and ARL values, a larger $ANOS$ value implies that either the monitoring scheme is inefficient and/or the cost of inspection is higher.

4.3 Measures of the overall performance

The *ARL* (see Equation (4.12)) is defined as the average number of samples required before an OOC signal is issued in the process. It is well-known that the *RL* distribution of a monitoring scheme is generally highly right-skewed in Case U; see for example Jones et al. (2004). As a result, many researchers prefer to use more meaningful performance measures (such as the percentiles of the *RL* which includes the median run-length (*MRL*)) to better evaluate the performance of the schemes. Furthermore, the *ARL* has been widely criticized by many authors, see for example, Wu et al. (2008) and Machado and Costa (2014). The *ARL* was simply criticized because of its ineffectiveness in assessing the overall performance, especially when the aim of the study is to assess the performance of a monitoring scheme over a range of shifts. Several authors have revealed that if a monitoring scheme is designed based on one specific size of a mean shift, it will perform poorly when the actual size of the shift is significantly different from the assumed size (see Reynolds and Lou (2010), Ryu, Wan and Kim (2010), Machado and Costa (2014) and Shongwe, Malela-Majika and Rapoo (2019)). Therefore, many researchers have recommended the use of a quality loss function (QLF) instead of the *ARL* to assess the performance of a monitoring scheme. A QLF describes the relationship between the shift size and the quality impact. The average extra quadratic loss (*AEQL*) is an alternative measure of the *ARL* used to assess the overall performance of a monitoring scheme for a range of shifts. Therefore, when the aim of a study is to measure the overall performance of a scheme over a range of shifts (say, $0 \leq \delta \leq 2.5$), the objective function can be defined in terms of the $AEQL$ given by

$$AEQL = \int_{\delta_{min}}^{\delta_{max}} \int_{-\infty}^{+\infty} \int_{0}^{+\infty} W(\delta) \left(\frac{1}{1 - \hat{P}_0} \right) f(\delta) f_U(u) f_V(v) \, dv \, du \, d\delta \tag{4.17}$$

where δ_{min} and δ_{max} are the lower and upper boundary of the range of shifts under consideration and $W(\delta)$ (with $W(\delta) = \delta^2$) represents the weight function associated with δ. Since it is generally assumed that all location shifts (mean shifts) occur with equal probability; hence, a uniform distribution of δ is implied, i.e. $f(\delta) = 1/(\delta_{max} - \delta_{min})$.

In order to measure the relative effectiveness of two different schemes, Wu et al. (2008) suggested the use of the *PCI*, which is the ratio between the *AEQL* of a competing monitoring scheme and the *AEQL* of the benchmark scheme under the same settings. In this chapter, the proposed SSDS scheme is used as the benchmark. The PCI is then defined by

$$PCI = \frac{AEQL}{AEQL^*} \tag{4.18}$$

where $AEQL^*$ is the $AEQL$ of the benchmark SSDS scheme. In addition to the $AEQL$ and the PCI, many authors also suggested the use of the $ARARL$ to measure the overall performance of a benchmark SSDS scheme against other competing schemes; see Wu et al. (2008). The $ARARL$ is given by

$$ARARL = \frac{1}{\delta_{max} - \delta_{min}} \sum_{\delta = \delta_{min}}^{\delta_{max}} \frac{ARL(\delta)}{ARL^*(\delta)} \tag{4.19}$$

where ARL^* is the *ARL* of the benchmark SSDS scheme. Note that, if the PCI and/or $ARARL$ is larger than one, the competing scheme will produce larger *ARL*s over the range of shifts under consideration, which means that the SSDS scheme outperforms the competing scheme for that particular range; otherwise, if *PCI* is less than one, then the competing scheme is more sensitive than the SSDS scheme.

4.4 Bi-objective model of the proposed SSDS $\bar{X}$ monitoring scheme

There are three control limits L_1, L and L_2 and two sample sizes n_1 and n_2 that need to be specified for a specified nominal *ASS* value (denoted by ASS_0) in order to design the SSDS $\bar{X}$ monitoring scheme. The efficiency of the proposed SSDS $\bar{X}$ scheme depends on the combination $(m, n_1, n_2, L_1, L, L_2)$. There are three main steps in the optimal design of the SSDS scheme: Firstly, the nominal IC ARL (ARL_0) is set to a high desired value, such as 370.4 or 500. Secondly, the combination that yields an ARL_0 as close as possible to the nominal ARL_0

value and the smallest OOC *ARL* (ARL_δ) for a given mean shift δ and a minimum $AEQL$ value is considered to be the optimal combination. Thirdly, the optimization model is given by:

$$(L_1, L, L_2) = \underset{m,n_1,n_2,L_1,L,L_2}{\operatorname{argmin}} (ARL_1, AEQL) \tag{4.20}$$

subjects to

$$ARL_0 = \tau \tag{4.21}$$

and

$$ASS_0 = \xi, \tag{4.22}$$

where ξ is the prespecified ASS_0 value and τ represents the nominal ARL_0 value. Note that the ASS_0 and OOC ASS (ASS_δ) are used because the sample size is not fixed in advance (it can be n_1 or $n_1 + n_2$). This plays an important role in the estimation of the cost of inspection.

The search of the optimal parameters can be summarized in *three* main steps given as follows:

1. Fix m and for some specific sample sizes (i.e., n_1 and n_2) and mean shift ($\delta = 0$), find all possible combinations of the design parameters that yield an attained ARL_0 value of 370.4 for a prespecified value of the ASS_0. These combination of parameters (L_1, L, L_2) are called local design parameters;
2. For each combination the local design parameters, compute the ARL_δ (where $\delta = 0.1$ to 2.5 with a step shift of 0.1) and then calculate the corresponding $AEQL$ value;
3. Select the combination that yields the minimum $AEQL$ value to be the combination of the optimal design parameters (L_1^*, L^*, L_2^*).

4.5 Performance study of the proposed monitoring scheme

4.5.1 Performance analysis of the SSDS $\bar{\mathbf{X}}$ monitoring scheme

In this section, the performance of the SSDS $\bar{X}$ scheme is investigated in Case U by setting the nominal ARL_0 value to 370.4 with a maximum mean shift of 2.5 (i.e. $\delta_{max} = 2.5$) and ASS_0 values of 5 and 8, see Tables 4.1 to 4.4; where, for illustration purpose, *m* is set at 50 and 100 for Case U SSDS $\bar{X}$ scheme, and *m* is assumed to approach infinity (∞) for Case K. The first row of each cell gives the ARL, $SDRL$, ASS and $ANOS$ values and the second row gives the 5th, 25th, 50th, 75th, 95th percentiles (denoted by (P5, P25, P50, P75, P95)) of the *RL* distribution of the Case U SSDS $\bar{X}$ scheme. Note that the Case K *RL* properties discussed in Chapter 3 are given in the last column. Equations (4.11) to (4.13) are used to compute the IC and OOC characteristics of the *RL* distribution. Moreover, the *ASS, ANOS* and *AEQL* values are computed using Equations (4.14), (4.16) and (4.17), respectively.

For instance (see the second column of Table 4.1), for a Phase I sample of size 50 (i.e. m = 50), when (n_1,n_2) = (5, 5), (δ_{min}, δ_{max}) = (0, 2.5) and ASS_0 = 5, it is found (using the optimization model in Equations (4.20) to (4.22)) that (L_1, L, L_2) = (2.9093, 3.0111, 2.9309) so that the SSDS $\bar{X}$ scheme satisfies ARL_0 = 370.4 with a minimum $AEQL$ = 70.72. However, when n_2 is increased to 8, for the same values of m, n_1, δ_{min}, δ_{max}, and ASS_0, it is found that (L_1, L, L_2) = (2.9101, 3.0108, 2.6310) so that the SSDS scheme also satisfies ARL_0 = 370.4 with a minimum $AEQL$ = 69.06 (see the third column of Table 4.1).

From Table 4.1, it can be seen that when m = 50 and ASS_0 = 5, if (n_1,n_2) = (5, 5) there is a 5% chance that the Case U SSDS $\bar{X}$ scheme gives a signal for the first time on the 18th subgroup and a 95% chance that it signals on the 1102 subgroup in Phase II when the process is IC (shift = 0). For a small shift of size 0.3, there are 5% and 95% chances that the proposed scheme gives a signal on the 11th and 563th subgroups, respectively. For m = 100 with an ASS_0 of 5, when (n_1,n_2) = (5, 5) (i.e. fourth column of Table 4.1) and with $\delta = 0.3$, there are 5% and 95% chances that the Case U SSDS $\bar{X}$ scheme gives a signal on the 7th and 388th subgroups, respectively. For a moderate $\delta = 0.9$, there are 5% and 95% chances that the Case U SSDS $\bar{X}$ scheme signals on the 1st and 18th subgroups, respectively. These findings confirm that the larger the Phase I sample size, the more sensitive the SSDS $\bar{X}$ scheme. As δ increases, the SSDS $\bar{X}$ scheme becomes more sensitive. When n_1 is kept at 5 and increase n_2 (say n_2 = 8) for m = 50, for a δ = 0.3, there is 95% chance that the proposed scheme gives a signal on the 546th subgroup in the prospective phase. This reveals an improvement in the sensitivity of the proposed scheme when the stage 2 sample size increases. In Case K, when δ = 0.3, (n_1,n_2) = (5, 8) and (3, 10) with ASS_0 = 5, there is 5% chance that the SSDS $\bar{X}$ scheme gives a signal on the 4th sample, see the last column of Tables 4.1 and 4.2, respectively. However, there is 95% chance that the SSDS $\bar{X}$ scheme gives a signal on the 260th and 183rd sample, respectively. This shows that the SSDS $\bar{X}$ scheme performs better in Case K. From Tables 4.3 and 4.4, it can be seen that when the *ASS* increases, the sensitivity of the SSDS $\bar{X}$ scheme increases as well. For small Phase I sample sizes (i.e. m = 25, on the second column of Tables 4.3 and 4.4), the detection ability of the SSDS $\bar{X}$ scheme is poor as compared to m = 50 and 100 on columns 3 to 6, respectively.

In terms of the ARL values, for small and moderate shifts in the process mean, the larger the Phase I sample (i.e. m), the more sensitive the Case U SSDS $\bar{X}$ scheme. However, for large shifts in the process mean, the SSDS $\bar{X}$ scheme performs uniformly better regardless of the Phase I sample size. For small and moderate shifts, the SSDS $\bar{X}$ scheme is less sensitive in Case

U than in Case K. This under-performance is due to the effect of estimation that deteriorates the performance of a monitoring scheme.

In terms of the $SDRL$ values, it can be seen that the practitioner-to-practitioner variability in the performance of the SSDS $\bar{X}$ scheme decreases as the Phase I sample size increases. The OOC $SDRL$ ($SDRL_\delta$) drop rapidly as the Phase I sample size increases. Therefore, the larger the Phase I sample size, the more reliable the results. The larger the ASS_0, the smaller the variability in the performance outputs. In terms of the $ANOS$ values, the larger the Phase I sample, the smaller the OOC $ANOS$. For very small shifts, $0 < \delta < 0.2$, the smaller the ASS_0 and $ANOS$ values. When $\delta \geq 0.2$, the larger the ASS_0 values and the smaller the $ANOS$ values, which means that when the process is OOC, the SSDS $\bar{X}$ scheme is more efficient and cost effective for larger values of ASS_0.

Table 4.1: The exact ARL, $SDRL$, ASS, $ANOS$ (first row), Percentiles (second row), $AEQL$ and optimal design parameters of the SSDS $\bar{X}$ scheme when $m\in\{50, 100\}$ and $m = \infty$ (i.e. Case K), (n_1, n_2) ∈{(5,5); (5, 8)}, ASS_0=5 and δ_{max} = 2.5 for a nominal ARL_0 value of 370.4

Shift(δ)	(***ARL, SDRL, ASS, ANOS***) **(P5, P25, P50, P75, P95)**				
0.00	(370.4, 369.98, 5.00, 1847.56) (18, 107, 254, 510, 1102)	(370.4, 371,36, 5.00, 1853.22) (20, 106, 257, 516, 1093)	(370.4, 375.44, 5.01, 1862.09) (19, 103, 256, 518, 1122)	(370.4, 373.22, 5.02, 1858.68) (19, 107, 254, 513, 1121)	(370.4, 367.15, 5.02, 1859.37) (20, 108, 257, 515, 1121)
0.10	(347.63, 340.25,5.00, 1740.89) (24, 132, 314, 620, 1337)	(340.83, 345.39,5.01, 1608.47) (23, 127, 303, 603, 1329)	(279.97, 278.84, 5.01,1504.20) (20, 109, 257, 522, 1139)	(273.23, 272.54, 5.02,1472.99) (20, 109, 258, 516, 1124)	(292.46, 288.87, 5.02, 1467.67) (16, 89, 205, 405, 870)
0.20	(333.23,329.98, 5.00, 1669.22) (17, 96, 233, 457, 998)	(321.71, 322.17, 5.01,1613.34) (16, 91, 222, 449, 948)	(243.47, 242.17, 5.08, 1221.64) (12, 73, 170, 334, 723)	(241.81, 241.39, 5.03, 1215.90) (13, 71, 169, 333, 714)	(159.38, 159.45, 5.03, 801.43) (8, 46, 110, 223, 474)
0.30	(187.70, 189.80, 5.02, 941.31) (11, 55, 129, 260, 563)	(182.85, 182,73, 5.02, 918.68) (10, 53, 128, 252, 546)	(131.10, 131.65, 5.03, 659.31) (7, 38, 91, 182, 388)	(123.83, 124.38, 5.05, 624.92) (7, 35, 84, 172, 375)	(86.17, 86.73, 5.05, 434.89) (4, 25, 59, 121, 260)
0.40	(102.65, 100.97, 5.02, 515) (6, 30, 72, 143, 307)	(97.23, 97.40, 5.03, 489.93) (5, 28, 68, 133, 292)	(69.75, 69.28, 5.05, 352.05) (4, 20, 48, 97, 208)	(66.91, 66.79, 5.08, 339.64) (4, 20, 46, 93, 202)	(47.67, 47.36, 5.08, 241.97) (3, 14, 33, 66, 143)
0.50	(56.26, 56.21, 5.03, 283.38) (3, 16, 39, 77, 170)	(54.73, 54.16, 5.06, 276.92) (3, 16, 38, 75, 165)	(39.00, 38.92, 5.07, 197.91) (3, 12, 27, 52, 117)	(36.99, 36.08, 5.11, 189.36) (2, 11, 26, 51, 109)	(28.02, 27.18, 5.12, 143.40) (2, 9, 20, 39, 82)
0.60	(32.50, 31.65, 5.05, 164.27) (2, 10, 23, 45, 97)	(30.88, 30.61, 5.09, 157.11) (2, 9, 21, 43, 93)	(22.99, 22.69, 5.11, 117.47) (2, 7, 16, 31, 69)	(21.98, 21.46, 5.18, 113.80) (2, 7, 15, 30, 64)	(17.30, 16.94, 5.18, 89.54) (1, 5, 12, 24, 52)
0.70	(19.63, 19.02, 5.07, 99.65) (1, 6, 14, 27, 58)	(19.16, 18.80, 5.12, 98.15) (1, 6, 13, 27, 55)	(14.20, 13.41, 5.16, 73.21) (1, 5, 10, 19, 41)	(13.46, 13.03, 5.25, 70.66) (1, 4, 9, 18, 40)	(11.26, 10.62, 5.25, 59.13) (1, 4, 8, 15, 33)
0.80	(12.62, 12.11, 5.10, 64.38) (1, 4, 9, 17, 36)	(12.18, 11.60, 5.16, 62.91) (1, 4, 9, 17, 35)	(9.15, 8.52, 5.21, 47.68) (1, 3, 7, 12, 26)	(8.59, 7.96, 5.34, 45.85) (1, 3, 6, 12, 24)	(7.52, 7.05, 5.34, 40.14) (1, 3, 5, 10, 22)
0.90	(8.34, 7.76, 5.13, 42.83) (1, 3, 6, 11, 24)	(7.94, 7.42, 5.21, 41.34) (1, 3, 6, 11, 23)	(6.31, 5.78, 5.27, 33.28) (1, 2, 4, 9, 18)	(6.03, 5.55, 5.43, 32.76) (1, 2, 4, 8, 17)	(5.31, 4.85, 5.43, 28.84) (1, 2, 4, 7, 15)
1.00	(5.83, 5.38, 5.15, 30.08) (1, 2, 4, 8, 16)	(5.66, 5.13, 5.25, 29.73) (1, 2, 4, 8, 16)	(4.46, 3.93, 5.42, 14.24) (1, 2, 3, 6, 13)	(4.28, 3.81, 5.53, 23.69) (1, 2, 3, 6, 12)	(4.03, 3.48, 5.53, 22.30) (1, 2, 3, 5, 11)
1.20	(3.22, 2.72, 5.20, 16.72) (1, 1, 2, 4, 9)	(3.21, 2.62, 5.31, 17.03) (1, 1, 2, 4, 8)	(2.62, 2.07, 5.45, 9.64) (1, 1, 2, 3, 7)	(2.56, 1.98, 5.68, 14.56) (1, 1, 2, 3, 6)	(2.40, 1.81, 5.68, 13.66) (1, 1, 2, 3, 6)
1.40	(2.04, 1.46, 5.20, 10.61) (1, 1, 2, 3, 5)	(2.04, 1.44, 5.32, 10.87) (1, 1, 2, 3, 5)	(1.77, 1.17, 5.39, 9.64) (1, 1, 1, 2, 4)	(1.78, 1.18, 5.72, 10.19) (1, 1, 1, 2, 4)	(1.71, 1.08, 5.71, 9.77) (1, 1, 1, 2, 4)
1.60	(1.50, 0.88, 5.17, 7.77) (1, 1, 1, 2, 3)	(1.51, 0.88, 5.27, 7.95) (1, 1, 1, 2, 3)	(1.39, 0.73, 5.39, 7.47) (1, 1, 1, 2, 3)	(1.37, 0.71, 5.62, 7.68) (1, 1, 1, 2, 3)	(1.34, 0.68, 5.62, 7.51) (1, 1, 1, 1, 3)
1.80	(1.25, 0.55, 5.12, 6.37) (1, 1, 1, 1, 2)	(1.24, 0.54, 5.18, 6.42) (1, 1, 1, 1, 2)	(1.17, 0.45, 5.27, 6.18) (1, 1, 1, 1, 2)	(1.17, 0.45, 5.44, 6.39) (1, 1, 1, 1, 2)	(1.16, 0.43, 5.44, 6.32) (1, 1, 1, 1, 2)
2.00	(1.11, 0.34, 5.06, 5.60) (1, 1, 1, 1, 2)	(1.11, 0.35, 5.10, 5.65) (1, 1, 1, 1, 2)	(1.07, 0.28, 5.16, 5.53) (1, 1, 1, 1, 2)	(1.07, 0.27, 5.61, 5.25) (1, 1, 1, 1, 2)	(1.06, 0.26, 5.25, 5.58) (1, 1, 1, 1, 2)
2.50	(1.01, 0.08, 5.01, 5.04) (1, 1, 1, 1, 1)	(1.01, 0.08, 5.01, 5.05) (1, 1, 1, 1, 1)	(1.00, 0.07, 5.02, 5.04) (1, 1, 1, 1, 1)	(1.00, 0.06, 5.03, 5.04) (1, 1, 1, 1, 1)	(1.00, 0.05, 5.03, 5.04) (1, 1, 1, 1, 1)
AEQL	70.72	69.06	59.59	55.28	49.91
$\boldsymbol{m}$ $(\boldsymbol{n_1}, \boldsymbol{n_2})$	50 (5, 5)	50 (5, 8)	100 (5, 5)	100 (5, 8)	∞ (5, 8)
$(\boldsymbol{L_1}, \boldsymbol{L})$ $\boldsymbol{L_2}$	(2.9093, 3.0111) 2.9309	(2.9101, 3.0108) 2.6310	(2.9096, 3.1354) 2.9103	(2.9098, 3.1361) 2.7101	(2.9099, 3.1354) 1.9002

Table 4.2: The exact ARL, $SDRL$, ASS, $ANOS$ (first row), Percentiles (second row), $AEQL$ and optimal design parameters of the SSDS $\bar{X}$scheme when $m\in$ {50, 100} and $m=\infty$ (i.e. Case K), (n_1, n_2) ∈{(3,5); (3, 10)}, ASS_0=5 and δ_{max} = 2.5 for a nominal ARL_0 value of 370.4

Shift(δ)	(*ARL, SDRL, ASS, ANOS*) (P5, P25, P50, P75, P95)				
0.00	(370.4, 369.67, 5.11, 1898.25) (20, 108, 255, 513, 1103)	(370.4, 372.08, 4.95, 1841.79) (20, 105, 258, 517, 1115)	(370.4, 370.85, 5.08, 1874.94) (20, 108, 258, 500, 1112)	(370.4, 373.06, 5.03, 1850.61) (18, 104, 253, 507, 1115)	(370.4, 369.43, 5.03, 1873.42) (19, 107, 259, 518, 1107)
0.10	(349.72, 344.22, 5.14, 1512.51) (23, 133, 312, 618, 1342)	(357.29, 354.83, 5.06, 1693) (26, 136, 320, 628, 1362)	(280.74, 279.53, 5.12, 1450.46) (19, 109, 263, 533, 1128)	(267.35, 268.21, 5.09, 1370.19) (20, 106, 255, 515, 1102)	(267.48, 262.02, 5.09, 1361.74) (14, 75, 190, 374, 791)
0.20	(329.95, 328.82, 5.23, 1728.44) (17, 93, 226, 460, 1000)	(317.14, 318.04, 5.20, 1650.42) (16, 91, 219, 438, 943)	(222.82, 222.43, 5.22, 1163.72) (12, 65, 154, 309, 662)	(235.37, 234.30, 5.28, 1243.25) (13, 70, 164, 325, 699)	(132.85, 130.81, 5.28, 701.74) (7, 39, 93, 184, 397)
0.30	(184.41, 188.05, 5.39, 994.05) (11, 52, 125, 252, 566)	(154.15, 153.63, 5.51, 848.96) (9, 44, 107, 213, 472)	(110.50, 109.24, 5.38, 594.59) (6, 33, 79, 153, 325)	(122.11, 120.98, 5.59, 682.53) (7, 35, 85, 170, 370)	(61.16, 60.69, 5.59, 341.88) (4, 18, 42, 86, 183)
0.40	(96.62, 95.25, 5.59, 539.68) (6, 29, 68, 133, 287)	(71.17, 70.62, 5.91, 420.53) (4, 21, 49, 100, 212)	(55.06, 54.33, 5.59, 307.52) (3, 16, 39, 76, 164)	(60.09, 59.23, 5.99, 360.35) (4, 17, 42, 84, 178)	(29.18, 27.98, 5.99, 175.01) (2, 9, 21, 41, 84)
0.50	(49.86, 48.63, 5.81, 289.59) (3, 15, 35, 69, 145)	(32.75, 32.34, 6.39, 209.19) (2, 10, 23, 45, 99)	(29.09, 29.01, 5.82, 169.38) (2, 9, 20, 40, 87)	(28.90, 28.79, 6.48, 187.32) (2, 9, 20, 40, 86)	(14.68, 14.39, 6.48, 95.14) (1, 5, 10, 20, 44)
0.60	(27.34, 27.00, 6.04, 165.15) (2, 8, 19, 38, 81)	(16.65, 16.14, 6.92, 115.15) (1, 5, 12, 23, 49)	(16.68, 16.30, 6.07, 101.34) (1, 5, 12, 23, 49)	(15.47, 14.96, 7.01, 108.49) (1, 5, 11, 21, 45)	(8.34, 7.76, 7.01, 58.49) (1, 3, 6, 11, 24)
0.70	(15.95, 15.41, 6.27, 99.91) (1, 5, 11, 22, 47)	(9.15, 8.57, 7.46, 68.30) (1, 3, 7, 13, 27)	(9.92, 9.48, 6.32, 62.74) (1, 3, 7, 14, 28)	(8.53, 8.09, 7.56, 64.56) (1, 3, 6, 12, 25)	(5.03, 4.45, 7.57, 38.08) (1, 2, 4, 7, 14)
0.80	(9.62, 9.07, 6.46, 62.19) (1, 3, 7, 13, 27)	(5.51, 5.03, 7.99, 44.03) (1, 2, 4, 7, 16)	(6.27, 5.75, 6.56, 41.11) (1, 2, 4, 9, 18)	(5.07, 4.56, 8.10, 41.09) (1, 2, 4, 7, 14)	(3.38, 2.81, 8.10, 27.40) (1, 1, 2, 4, 9)
0.90	(6.29, 5.79, 6.62, 41.62) (1, 2, 4, 8, 18)	(3.57, 2.99, 8.47, 30.23) (1, 1, 3, 5, 9)	(2.99, 2.42, 6.76, 28.58) (1, 1, 2, 4, 8)	(3.40, 2.81, 8.59, 29.18) (1, 1, 3, 4, 9)	(2.41, 1.84, 8.59, 20.70) (1, 1, 2, 3, 6)
1.00	(4.18, 3.60, 6.72, 28.09) (1, 2, 3, 6, 11)	(2.57, 2.01, 8.87, 22.82) (1, 1, 2, 3, 7)	(4.23, 3.62, 6.92, 20.66) (1, 2, 3, 6, 11)	(2.46, 1.90, 8.99, 22.08) (1, 1, 2, 3, 6)	(1.91, 1.31, 8.99, 17.18) (1, 1, 1, 2, 5)
1.20	(2.28, 1.72, 6.75, 15.41) (1, 1, 2, 3, 6)	(1.62, 1.01, 9.33, 15.09) (1, 1, 1, 2, 4)	(1.83, 1.24, 7.06, 12.89) (1, 1, 1, 2, 4)	(1.54, 0.91, 9.45, 14.57) (1, 1, 1, 2, 3)	(1.36, 0.69, 9.46, 12.83) (1, 1, 1, 2, 3)
1.40	(1.54, 0.90, 6.50, 9.99) (1, 1, 1, 2, 3)	(1.25, 0.55, 9.24, 11.55) (1, 1, 1, 1, 2)	(1.34, 0.68, 6.96, 9.32) (1, 1, 1, 2, 3)	(1.22, 0.51, 9.37, 11.41) (1, 1, 1, 1, 2)	(1.17, 0.44, 9.37, 10.93) (1, 1, 1, 1, 2)
1.60	(1.20, 0.49, 6.04, 7.25) (1, 1, 1, 1, 2)	(1.11, 0.36, 8.62, 9.61) (1, 1, 1, 1, 2)	(1.12, 0.37, 6.61, 7.42) (1, 1, 1, 1, 2)	(1.09, 0.31, 8.75, 9.56) (1, 1, 1, 1, 2)	(1.07, 0.27, 8.76, 9.39) (1, 1, 1, 1, 2)
1.80	(1.08, 0.29, 5.43, 5.85) (1, 1, 1, 1, 2)	(1.05, 0.23, 7.62, 8.02) (1, 1, 1, 1, 1)	(1.04, 0.21, 6.07, 6.32) (1, 1, 1, 1, 1)	(1.04, 0.20, 7.75, 8.04) (1, 1, 1, 1, 1)	(1.03, 0.19, 7.75, 8.01) (1, 1, 1, 1, 1)
2.00	(1.02, 0.15, 4.78, 4.89) (1, 1, 1, 1, 1)	(1.02, 0.15, 6.46, 6.60) (1, 1, 1, 1, 1)	(1.02, 0.16, 5.42, 5.48) (1, 1, 1, 1, 1)	(1.02, 0.13, 6.57, 6.69) (1, 1, 1, 1, 1)	(1.01, 0.12, 6.57, 6.66) (1, 1, 1, 1, 1)
2.50	(1.00, 0.02, 3.55, 3.55) (1, 1, 1, 1, 1)	(1.00, 0.04, 4.09, 4.10) (1, 1, 1, 1, 1)	(1.00, 0.01, 3.93, 3.93) (1, 1, 1, 1, 1)	(1.00, 0.04, 4.15, 4.16) (1, 1, 1, 1, 1)	(1.00, 0.04, 4.15, 4.16) (1, 1, 1, 1, 1)
AEQL	62.32	50.50	46.82	45.79	35.03
m (n_1, n_2)	50 (3, 5)	50 (3, 10)	100 (3, 5)	100 (3, 10)	∞ (3, 10)
(L_1, L) L_2	(0.8018, 3.1071) 3.2650	(1.2903, 3.1079) 3.2210	(0.8093, 3.4354) 3.1721	(1.2693, 3.1354) 3.5001	(1.2693, 3.1354) 3.1600

Table 4.3: The exact ARL, $SDRL$, ASS, $ANOS$ (first row), Percentiles (second row), $AEQL$ and optimal design parameters of the SSDS $\bar{X}$ scheme when $m\in\{25, 50, 100\}$ and $m = \infty$ (i.e. Case K), $(n_1, n_2) \in\{(5,5); (5, 8)\}$, ASS_0=8 and δ_{max} = 2.5 for a nominal ARL_0 value of 370.4

Shift(δ)	**(*ARL, SDRL, ASS, ANOS*)** **(P5, P25, P50, P75, P95)**					
0.00	(370.4, 373.18, 8.41, 3126.71) (20, 106, 256, 517, 1115)	(370.4, 370.84, 7.99, 2961.50) (20, 108, 257, 508, 1116)	(370.4, 365.99, 8.00, 2933.87) (21, 109, 256, 505, 1086)	(370.4, 378.83, 7.99, 3014.07) (18, 107, 256, 522, 1134)	(370.4, 361.99, 8.00, 2914.36) (19, 106, 251, 512, 1077)	(370.4, 368.56, 8.00, 2956.70) (20, 110, 258, 509, 1085)
0.10	(266.04, 266.36,8.44, 2246.64) (13, 77, 181, 367, 806)	(168.63, 165.29, 8.04,1356.19) (9, 50, 119, 236, 502)	(138.89,137.37, 8.10, 1124.34) (8, 41, 97, 192, 417)	(168.88, 180.20, 8.04,1438.60) (10, 51, 125, 249, 531)	(132.11, 157.06,8.10, 1662.76) (35, 187, 444, 901, 671)	(133.41, 119.10, 8.10, 1608.56) (12, 68, 156, 307, 674)
0.20	(128.55, 128.04, 8.55,1098.77) (7, 39, 89, 179, 383)	(73.81, 74.87, 8.17, 602.87) (4, 21, 50, 102, 219)	(59.36, 57.63, 8.37, 454.94) (3, 16, 38, 75, 163)	(49.41, 79.13, 8.17, 648.65) (5, 23, 55, 110, 238)	(56.80, 88.21, 8.37, 480.60) (21, 115, 75, 143, 230)	(56.76, 55.42, 8.37, 734.41) (5, 25, 61, 120, 227)
0.30	(58.67, 57.42, 8.70, 510.52) (4, 18, 41, 81, 174)	(35.23, 34.95, 8.36, 294.47) (2, 11, 25, 49, 103)	(26.06, 25.35, 8.79, 211.56) (2, 7, 17, 33, 70)	(27.00, 36.72, 8.36, 309.25) (2, 11, 26, 51, 109)	(24.63, 24.04,8.79, 389.23) (8, 43, 103, 203, 437)	(24.35, 23.01, 8.79, 319.64) (2, 11, 25, 51, 108)
0.40	(28.64, 28.41, 8.88, 254.43) (2, 8, 20, 39, 86)	(18.06, 17.63, 8.58, 155.04) (1, 6, 12, 25, 54)	(13.73, 12.26, 9.32, 109.35) (1, 4, 8, 16, 34)	(18.87, 18.22, 8.58, 161.96) (1, 6, 13, 26, 56)	(12.32, 56.60, 9.32, 134.38) (3, 17, 40, 79, 171)	(12.02, 11.36, 9.32, 158.68) (1, 5, 12, 24, 49)
0.50	(15.01, 14.43, 9.07, 136.12) (1, 5, 11, 21, 43)	(9.96, 9.31, 8.82, 87.86) (1, 3, 7, 14, 28)	(7.41, 5.86, 9.91, 63.54) (1, 2, 5, 9, 18)	(10.42, 10.08, 8.82, 91.89) (1, 3, 7, 14, 30)	(6.42, 23.99, 9.91, 101.98) (2, 7, 17, 33, 72)	(6.80, 5.33, 9.91, 87.20) (1, 3, 6, 12, 25)
0.60	(9.60, 8.16, 9.23, 79.42) (1, 3, 6, 12, 25)	(5.92, 5.41, 9.03, 53.43) (1, 2, 4, 8, 17)	(4.95, 4.44, 10.51, 41.49) (1, 1, 3, 5, 11)	(6.25, 5.81, 9.03, 56.44) (1, 2, 4, 8, 18)	(4.18, 11.65, 10.50, 67.92) (1, 4, 9, 17, 36)	(4.06, 3.65, 10.50, 53.19) (1, 2, 4, 7, 14)
0.70	(5.31, 4.89, 9.35, 49.63) (1, 2, 4, 7, 15)	(3.93, 3.42, 9.19, 36.09) (1, 1, 3, 5, 11)	(2.94, 2.03, 11.07, 28.73) (1, 1, 2, 3, 7)	(4.04, 3.57, 9.19, 37.16) (1, 1, 3, 5, 11)	(2.61, 6.04, 11.06, 43.13) (1, 2, 5, 9, 19)	(2.25, 1.74, 11.06, 36.01) (1, 1, 2, 4, 9)
0.80	(3.53, 3.04, 9.40, 33.17) (1, 1, 3, 5, 10)	(2.94, 2.22, 9.29, 25.41) (1, 1, 2, 4, 7)	(1.99, 1.43, 11.56, 22.22) (1, 1, 1, 2, 5)	(2.77, 2.22, 9.29, 25.71) (1, 1, 2, 4, 7)	(2.00, 3.43, 11.55, 26.20) (1, 1, 3, 5, 11)	(1.98, 1.69, 11.55, 26.32) (1, 1, 2, 3, 6)
0.90	(2.91, 1.97, 9.38, 23.57) (1, 1, 2, 3, 6)	(2.25, 1.44, 9.29, 19.07) (1, 1, 2, 3, 5)	(1.80, 0.87, 11.97, 17.92) (1, 1, 1, 2, 3)	(2.10, 1.56, 9.30, 19.57) (1, 1, 2, 3, 5)	(1.70, 2.15, 11.95, 22.30) (1, 1, 2, 4, 7)	(1.74, 1.12, 11.96, 20.82) (1, 1, 1, 2, 4)
1.00	(2.09, 1.31, 9.28, 17.57) (1, 1, 1, 2, 4)	(1.95, 1.03, 9.22, 15.17) (1, 1, 1, 2, 4)	(1.59, 0.61, 12.30, 15.85) (1, 1, 1, 1, 2)	(1.65, 1.05, 9.22, 15.24) (1, 1, 1, 2, 4)	(1.33, 1.31, 12.26, 23.61) (1, 1, 1, 2, 5)	(1.30, 0.74, 12.26, 17.16) (1, 1, 1, 2, 3)
1.20	(1.33, 0.67, 8.82, 11.69) (1, 1, 1, 1, 3)	(1.30, 0.52, 8.80, 10.74) (1, 1, 1, 1, 2)	(1.28, 0.29, 12.71, 13.71) (1, 1, 1, 1, 2)	(1.22, 0.52, 8.80, 10.73) (1, 1, 1, 1, 2)	(1.10, 0.62, 12.60, 16.36) (1, 1, 1, 1, 3)	(1.12, 0.37, 12.60, 14.10) (1, 1, 1, 1, 2)
1.40	(1.19, 0.33, 8.08, 8.91) (1, 1, 1, 1, 2)	(1.09, 0.27, 8.08, 8.61) (1, 1, 1, 1, 2)	(1.09, 0.14, 12.90, 13.14) (1, 1, 1, 1, 1)	(1.07, 0.27, 8.08, 8.61) (1, 1, 1, 1, 2)	(1.08, 0.29, 12.57, 13.56) (1, 1, 1, 1, 2)	(1.03, 0.16, 12.57, 12.89) (1, 1, 1, 1, 1)
1.60	(1.12, 0.15, 7.21, 7.38) (1, 1, 1, 1, 1)	(1.07, 0.13, 7.21, 7.33) (1, 1, 1, 1, 1)	(1.06, 0.06, 12.97, 13.02) (1, 1, 1, 1, 1)	(1.02, 0.13, 7.21, 7.32) (1, 1, 1, 1, 1)	(1.02, 0.14, 12.18, 12.39) (1, 1, 1, 1, 1)	(1.01, 0.08, 12.18, 12.26) (1, 1, 1, 1, 1)
1.80	(1.09, 0.07, 6.39, 6.42) (1, 1, 1, 1, 1)	(1.05, 0.04, 6.39, 6.40) (1, 1, 1, 1, 1)	(1.04, 0.02, 12.99, 12.99) (1, 1, 1, 1, 1)	(1.00, 0.05, 6.39, 6.40) (1, 1, 1, 1, 1)	(1.00, 0.06, 11.39, 11.44) (1, 1, 1, 1, 1)	(1.00, 0.03, 11.39, 11.40) (1, 1, 1, 1, 1)
2.00	(1.07, 0.02, 5.75, 5.75) (1, 1, 1, 1, 1)	(1.02, 0.01, 5.75, 5.75) (1, 1, 1, 1, 1)	(1.01, 0.02, 12.99, 13.00) (1, 1, 1, 1, 1)	(1.00, 0.02, 5.75, 5.75) (1, 1, 1, 1, 1)	(1.00, 0.01, 10.22, 10.23) (1, 1, 1, 1, 1)	(1.00, 0.02, 10.22, 10.23) (1, 1, 1, 1, 1)
2.50	(1.04, 0.00. 5.08, 5.08) (1, 1, 1, 1, 1)	(1.01, 0.00, 5.08, 5.08) (1, 1, 1, 1, 1)	(1.00, 0.00, 12.99, 12.99) (1, 1, 1, 1, 1)	(1.00, 0.00, 5.08, 5.08) (1, 1, 1, 1, 1)	(1.00, 0.00, 6.88, 6.88) (1, 1, 1, 1, 1)	(1.00, 0.00, 6.88, 6.88) (1, 1, 1, 1, 1)
AEQL	39.63	36.96	33.59	32.84	30.19	29.07
m (n_1, n_2)	25 (5, 5)	50 (5, 5)	50 (5, 8)	100 (5, 5)	100 (5, 8)	∞ (5, 8)
(L_1, L) L_2	(0.4093, 3.4354) 3.2221	(0.524, 3.435) 3.243	(0.8870, 10.5920) 3.1850	(2.9835, 3.0083) 2.0027	(0.887, 4.866) 3.375	(0.887, 4.866) 2.975

Table 4.4: The exact ARL, $SDRL$, ASS, $ANOS$ (first row), the Percentiles (second row), $AEQL$ and optimal design parameters of the SSDS $\bar{X}$ scheme when m∈{25, 50, 100} and $m = \infty$ (i.e. Case K), (n_1, n_2) ∈{(3,5); (3, 10)}, ASS_0 = 8 and δ_{max} = 2.5 for a nominal ARL_0 value of 370.4

Shift(δ)	**(*ARL, SDRL, ASS, ANOS*)** **(P5, P25, P50, P75, P95)**					
0.00	(370.4, 377.30, 7.98, 2995.65) (20, 107, 262, 516, 1131)	(370.4, 373.36, 7.98, 2979.38) (19, 108, 257, 516, 1126)	(370.4, 371.12, 8.00, 2947.74) (20, 105, 254, 513, 1111)	(370.4, 367.01, 7.98, 2968.61) (20, 108, 263, 515, 1124)	(370.4, 364.11, 8.00, 2938.58) (19, 106, 257, 507, 1090)	(370.4, 374.68, 8.00, 2982.43) (20, 110, 253, 515, 1126)
0.10	(273.72, 274.99, 7.98, 2183.83) (15, 79, 191, 378, 808)	(270.11, 271.52, 7.98, 2155.02) (15, 80, 187, 376, 800)	(134.70, 135.11, 8.07, 1086.59) (7, 39, 94, 188, 405)	(195.76, 196.16, 7.98, 1561.82) (10, 55, 135, 271, 592)	(131.27, 131.95, 8.07, 1020.31) (8, 43, 105, 209, 446)	(129.99, 121.92, 8.07, 1074.63) (11, 64, 151, 303, 466)
0.20	(150.52, 150.47, 7.97, 1200.79) (8, 43, 105, 207, 455)	(150.22, 148.41, 7.97, 1197.84) (8, 44, 105, 208, 448)	(52.96, 52.31, 8.25, 473.12) (3, 16, 37, 74, 158)	(99.86, 99.42, 7.97, 796.25) (6, 30, 69, 137, 302)	(50.03, 49.70, 8.25, 447.20) (4, 18, 41, 81, 173)	(49.29, 49.09, 8.25, 437.00) (5, 26, 61, 122, 165)
0.30	(79.85, 79.34, 7.96, 635.99) (5, 24, 55, 111, 237)	(77.18, 75.54, 7.96, 614.69) (4, 22, 54, 107, 228)	(23.82, 23.60, 8.55, 203.66) (2, 7, 17, 33, 70)	(50.80, 49.60, 7.95, 404.60) (3, 15, 36, 71, 151)	(21.47, 20.86, 8.55, 226.32) (2, 8, 19, 36, 78)	(20.87, 19.06, 8.55, 323.77) (2, 11, 27, 53, 111)
0.40	(42.89, 43.41, 7.95, 340.97) (3, 12, 30, 59, 127)	(41.25, 39.51, 7.95, 327.96) (3, 12, 29, 58, 122)	(11.71, 11.36, 8.93, 104.60) (1, 4, 8, 16, 35)	(27.53, 26.61, 7.95, 218.88) (2, 8, 19, 38, 80)	(11.96, 11.43, 8.93, 115.75) (1, 4, 9, 18, 38)	(11.84, 10.28, 8.93, 119.39) (1, 5, 13, 24, 52)
0.50	(23.54, 23.11, 7.93, 186.62) (2, 7, 16, 32, 69)	(23.29, 22.92, 7.93, 184.63) (2, 7, 16, 32, 68)	(6.54, 6.14, 9.38, 61.30) (1, 2, 5, 9, 19)	(15.74, 15.24, 7.93, 124.78) (1, 5, 11, 21, 46)	(6.07, 5.51, 9.38, 66.35) (1, 2, 5, 9, 20)	(6.20, 5.87, 9.38, 86.24) (1, 3, 6, 12, 27)
0.60	(13.47, 13.07, 7.89, 106.35) (1, 4, 10, 18, 40)	(13.23, 12.71, 7.89, 104.40) (1, 4, 9, 18, 38)	(3.95, 3.37, 9.86, 38.93) (1, 1, 3, 5, 11)	(9.48, 9.03, 7.89, 74.86) (1, 3, 7, 13, 28)	(3.84, 2.70, 9.86, 37.79) (1, 2, 3, 6, 12)	(3.80, 2.86, 9.86, 36.21) (1, 2, 4, 7, 15)
0.70	(8.32, 7.79, 7.84, 65.22) (1, 3, 6, 11, 24)	(8.25, 7.76, 7.84, 64.67) (1, 3, 6, 11, 24)	(2.69, 2.16, 10.34, 27.85) (1, 1, 2, 3, 7)	(6.02, 5.41, 7.84, 47.22) (1, 2, 4, 8, 17)	(2.62, 1.89, 10.34, 27.20) (1, 1, 2, 4, 7)	(2.61, 1.90, 10.34, 35.28) (1, 1, 2, 4, 9)
0.80	(5.45, 4.89, 7.77, 42.40) (1, 2, 4, 7, 15)	(5.42, 4.94, 7.77, 42.16) (1, 2, 4, 7, 15)	(1.97, 1.40, 10.81, 21.29) (1, 1, 1, 2, 5)	(4.10, 3.63, 7.77, 31.85) (1, 1, 3, 5, 11)	(2.06, 1.48, 10.81, 22.32) (1, 1, 2, 3, 5)	(2.32, 1.77, 10.81, 25.12) (1, 1, 2, 3, 6)
0.90	(3.77, 3.24, 7.68, 28.95) (1, 1, 3, 5, 10)	(3.73, 3.19, 7.68, 28.64) (1, 1, 3, 5, 10)	(1.65, 0.92, 11.25, 17.42) (1, 1, 1, 2, 3)	(2.93, 2.38, 7.68, 22.48) (1, 1, 2, 4, 8)	(1.62, 1.00, 11.25, 18.24) (1, 1, 1, 2, 4)	(1.78, 1.17, 11.25, 19.97) (1, 1, 1, 2, 4)
1.00	(2.27, 2.15, 7.56, 20.58) (1, 1, 2, 4, 7)	(2.74, 2.19, 7.56, 20.73) (1, 1, 2, 4, 7)	(1.42, 0.65, 11.63, 15.40) (1, 1, 1, 1, 3)	(2.22, 1.65, 7.56, 16.81) (1, 1, 2, 3, 6)	(1.36, 0.71, 11.63, 15.83) (1, 1, 1, 2, 3)	(1.47, 0.83, 11.63, 17.13) (1, 1, 1, 2, 3)
1.20	(1.74, 1.14, 7.22, 12.55) (1, 1, 1, 2, 4)	(1.71, 1.10, 7.22, 12.36) (1, 1, 1, 2, 4)	(1.21, 0.34, 12.23, 13.49) (1, 1, 1, 1, 2)	(1.51, 0.88, 7.22, 10.87) (1, 1, 1, 2, 3)	(1.12, 0.37, 12.23, 13.71) (1, 1, 1, 1, 2)	(1.16, 0.44, 12.23, 14.24) (1, 1, 1, 1, 2)
1.40	(1.28, 0.60, 6.73, 8.63) (1, 1, 1, 1, 2)	(1.28, 0.60, 6.73, 8.58) (1, 1, 1, 1, 2)	(1.05, 0.18, 12.60, 13.01) (1, 1, 1, 1, 1)	(1.19, 0.48, 6.73, 8.03) (1, 1, 1, 1, 2)	(1.04, 0.21, 12.60, 13.12) (1, 1, 1, 1, 1)	(1.06, 0.25, 12.60, 13.31) (1, 1, 1, 1, 2)
1.60	(1.10, 0.32, 6.12, 6.71) (1, 1, 1, 1, 2)	(1.10, 0.33, 6.12, 6.74) (1, 1, 1, 1, 2)	(1.03, 0.10, 12.81, 12.94) (1, 1, 1, 1, 1)	(1.07, 0.26, 6.12, 6.53) (1, 1, 1, 1, 2)	(1.01, 0.12, 12.81, 12.99) (1, 1, 1, 1, 1)	(1.02, 0.14, 12.81, 13.07) (1, 1, 1, 1, 1)
1.80	(1.03, 0.17, 5.44, 5.59) (1, 1, 1, 1, 1)	(1.03, 0.18, 5.44, 5.60) (1, 1, 1, 1, 1)	(1.01, 0.07, 12.88, 12.94) (1, 1, 1, 1, 1)	(1.02, 0.13, 5.44, 5.54) (1, 1, 1, 1, 1)	(1.01, 0.07, 12.88, 12.95) (1, 1, 1, 1, 1)	(1.01, 0.08, 12.88, 12.97) (1, 1, 1, 1, 1)
2.00	(1.01, 0.08, 4.77, 4.78) (1, 1, 1, 1, 1)	(1.01, 0.09, 5.10, 4.80) (1, 1, 1, 1, 1)	(1.00, 0.03, 12.85, 12.86) (1, 1, 1, 1, 1)	(1.00, 0.07, 4.77, 4.79) (1, 1, 1, 1, 1)	(1.00, 0.05, 12.85, 12.88) (1, 1, 1, 1, 1)	(1.00, 0.04, 12.85, 12.87) (1, 1, 1, 1, 1)
2.50	(1.00, 0.01, 3.54, 3.54) (1, 1, 1, 1, 1)	(1.00, 0.01, 3.54, 3.54) (1, 1, 1, 1, 1)	(1.00, 0.01, 12.17, 12.17) (1, 1, 1, 1, 1)	(1.00, 0.00, 3.53, 3.54) (1, 1, 1, 1, 1)	(1.00, 0.01, 12.17, 12.17) (1, 1, 1, 1, 1)	(1.00, 0.01, 12.17, 12.17) (1, 1, 1, 1, 1)
AEQL	41.22	40.88	28.10	34.83	27.68	27.21
m **(n_1, n_2)**	25 (3, 5)	50 (3, 5)	50 (3, 10)	100 (3, 5)	100 (3, 10)	∞ (3, 10)
(L_1, L) **L_2**	(0.02527, 3.088) 3.2100	(0.02549, 3.091) 3.2010	(0.6740, 5.7150) 3.3102	(0.02544, 3.088) 3.1503	(0.6736, 5.7150) 2.999	(0.6742, 5.7155) 2.945

In terms of the overall performance, i.e. *AEQL*, the SSDS $\bar{X}$ scheme performs better for large Phase I sample sizes and/or large expected number of samples. As the Phase I sample size increases, the Case U run-length properties converge towards the Case K run-length properties. Therefore, it is very important to study the effect of the Phase I sample size on the performance of the SSDS scheme in order to know the amount of Phase I observations required to reach the Case K performance; this is done in the next sub-section.

4.5.2 *ARL* profiles of the Case U SSDS $\bar{X}$ scheme using Case K parameters

In this sub-section, we investigate the *ARL* profile behaviour of the proposed SSDS $\bar{X}$ scheme when the Case U performance is obtained using the Case K optimal design parameters (as discussed in Chapter 3 of this dissertation) instead of the Case U optimal design parameters. To evaluate the impact of using the Case K optimal design parameters in Case U, the percentage difference ($\%Diff$) between the Case U OOC ARL (denoted as ARL_{δ_U}) and Case K OOC ARL (denoted as ARL_{δ_K}) is calculated as follows:

$$\%Diff = \left(\frac{ARL_{\delta_U} - ARL_{\delta_K}}{ARL_{\delta_K}}\right) \times 100 \tag{4.23}$$

Table 4.5 displays the ARL_{δ_U} and ARL_{δ_K} (last column) values using the Case K optimal design parameters when $n \in \{2, 5\}$, $n_1 \in \{2, 5\}$, $n_2 \in \{2, 5, 8, 11\}$ and nominal ARL_0 of 370.4; where $m = \infty$ denotes the Case K values. From Table 4.5, it is observed that the proposed SSDS $\bar{X}$ scheme yields very large ARL_{δ_U} for small Phase I sample sizes. For instance, when $\delta = 0.5$ and (n, n_1, n_2) = (2,2,2) for a nominal ARL_0 value of 370.4, the SSDS $\bar{X}$ scheme yields *ARL* values of 160.2 and 79.41 when m = 25 and m = ∞, respectively; revealing a 101.7% percentage difference as compared to the Case K *ARL* value. Moreover, the results in Table 4.5 show that, as the Phase I sample size increases, the $\%Diff$ decreases considerably. For a large Phase I sample size (e.g. $m = 400$), the $\%Diff$ is less than 1%, meaning that the Case U SSDS $\bar{X}$ scheme performs as if the optimal design parameters were known. Therefore, it is very important to know the number of Phase I observations for which the proposed scheme performs as if it was in Case K. As we can see from Table 4.5, this will depend on the *ASS* as well as the stages 1 and 2 sample sizes. Moreover, the larger the *ASS*, the higher the $\%Diff$.

Table 4.5: Case U and Case K OOC ARL (first row) and $\%Diff$ (second row) of the SSDS scheme using the Case K optimal design parameters when $n \in \{2, 5\}$, $n_1 \in \{2, 5\}$ and $n_2 \in \{2, 5, 8, 11\}$ when $NARL_0 = 370.4$

$\boldsymbol{n}$	$(\boldsymbol{n_1}, \boldsymbol{n_2})$	Case K Optimal parameters $(\boldsymbol{L_1}, \boldsymbol{L}, \boldsymbol{L_2})$	$\boldsymbol{\delta}$	$\boldsymbol{m}$						
				25	**50**	**100**	**200**	**300**	**400**	∞
2	(2, 2)	(2.910, 3.057, 2.405)	0.5	160.20 101.7%	112.64 41.8%	101.41 27.7%	82.01 3.3%	81.11 2.1%	80.01 0.8%	79.41
			1.0	30.14 95.3%	30.19 95.7%	24.58 59.3%	18.71 21.3	17.14 11.1%	15.48 0.3%	15.43
			1.5	11.16 135.9%	8.19 73.2%	6.37 34.7%	5.54 17.1	5.00 5.7%	4.74 0.2%	4.73
			2.0	5.13 135.3%	3.31 51.8%	3.00 37.6%	2.64 21.1	2.39 9.6%	2.20 0.9%	2.18
			2.5	3.36 140.0%	2.24 60.0%	2.24 60.0%	2.03 45.0	1.88 34.3%	1.41 0.7%	1.40
	(2, 8)	(2.975, 3.005, 2.931)	0.5	164.20 104.4%	95.02 18.3%	87.64 9.1%	83.63 4.1%	82.23 2.4%	81.09 0.9%	80.34
			1.0	32.11 101.1%	26.62 66.7%	23.13 44.8%	19.24 20.5	17.44 9.2%	16.03 0.4%	15.97
			1.5	12.94 171.3%	9.04 89.5%	7.18 50.5%	5.70 19.5	5.02 5.2%	4.82 1.0%	4.77
			2.0	6.07 167.4%	3.56 56.8%	2.98 31.3%	2.66 17.2	2.37 4.4%	2.29 0.9%	2.27
			2.5	4.01 180.4%	2.62 83.2%	2.21 54.5%	2.05 43.4	1.76 23.1%	1.44 0.7%	1.43
	(2, 11)	(2.991, 3.00, 2.998)	0.5	165.18 104.6%	96.21 19.2%	88.07 9.1%	83.61 3.6%	82.00 1.6%	81.18 0.5%	80.74
			1.0	34.05 104.4%	28.19 69.2%	24.13 44.8%	19.31 15.9	18.07 8.5%	16.73 0.4%	16.66
			1.5	13.71 176.4%	12.63 154.6%	7.25 46.2%	5.82 17.3	5.38 8.5%	5.00 0.8%	4.96
			2.0	6.60 188.2%	6.11 166.8%	2.99 30.6%	2.67 16.6	2.49 8.7%	2.31 0.9%	2.29
			2.5	4.26 143.4%	4.38 150.3%	2.29 30.9%	2.24 28.0	2.03 16.0%	1.76 0.6%	1.75

Table 4.5: (continued)

n	(n_1, n_2)	Case K Optimal parameters (L_1, L, L_2)	δ	m: 25	50	100	200	300	400	∞
5	(2, 11)	(1.094, 3.234, 3.010)	0.	116.33 501.2%	30.21 56.1%	22.07 14.1%	22.51 16.3%	21.02 8.6%	19.4 0.3%	19.3
			1.	24.42 921.8%	4.23 77.0%	3.72 55.6%	3.31 38.5%	2.89 20.9	2.44 2.1%	2.39
			1.	9.06 512.2%	2.51 69.6%	2.04 37.8%	1.69 14.2%	1.61 8.8%	1.49 0.7%	1.48
			2.	3.10 176.8%	2.08 85.7%	1.63 45.5%	1.39 24.1%	1.26 12.5	1.13 0.9%	1.12
			2.	2.00 96.1%	1.61 57.8%	1.39 36.3%	1.22 19.6%	1.12 9.8%	1.04 2.0%	1.02
	(5, 5)	(2.993, 3.001, 3.000)	0.	120.76 359.3%	41.47 57.7%	32.04 21.9%	29.11 10.7%	27.77 5.6%	26.4 0.5%	26.2
			1.	22.01 470.2%	9.21 138.6%	6.36 64.8%	4.71 22.0%	4.04 4.7%	3.87 0.3%	3.86
			1.	8.86 523.9%	4.92 246.5%	2.28 60.6%	1.61 13.4%	1.50 5.6%	1.42 0.0%	1.42
			2.	3.97 264.2%	2.73 150.5%	1.94 78.0%	1.40 28.4%	1.26 15.6	1.10 0.9%	1.09
			2.	3.45 235.0%	1.87 81.6%	1.47 42.7%	1.19 15.5%	1.10 6.8%	1.03 0.0%	1.02
	(5, 8)	(2.993, 3.001, 2.998)	0.	104.31 460.5%	26.42 42.0%	21.72 16.7%	20.17 8.4%	19.08 2.5%	18.7 0.5%	18.6
			1.	21.79 942.6%	7.26 247.4%	5.48 162.2%	5.01 139.7	3.04 45.5	2.09 0.0%	2.09
			1.	7.08 475.6%	4.54 269.1%	2.19 78.0%	2.01 63.4%	1.63 32.5	1.24 0.8%	1.22
			2.	3.31 212.3%	2.80 164.2%	1.79 68.9%	1.53 44.3%	1.36 28.3	1.07 0.9%	1.06
			2.	2.48 145.5%	1.74 72.3%	1.41 39.6%	1.34 32.7%	1.20 18.8	1.02 1.0%	1.01
	(5, 11)	(2.996, 3.000, 2.999)	0.	99.17 476.2%	25.45 47.9%	21.39 24.3%	19.15 11.3%	17.82 3.5%	17.2 0.1%	17.2
			1.	22.84 982.5%	7.42 251.7%	5.47 159.2%	4.93 133.6	3.00 42.2	2.11 0.0%	2.11
			1.	6.59 478.1%	4.50 294.7%	2.16 89.5%	1.97 72.8%	1.61 41.2	1.15 0.9%	1.14
			2.	3.22 198.1%	2.77 156.5%	1.77 63.9%	1.49 38.0%	1.33 23.1	1.09 0.9%	1.07
			2.	2.29 126.7%	1.70 68.3%	1.39 37.6%	1.33 31.7%	1.21 19.8	1.02 1.0%	1.00

Therefore, to secure stability and better OOC performance in Phase II for the proposed SSDS $\bar{X}$ scheme, the operator must either use a high desired Phase I sample size or choose the appropriate design parameters as suggested in Tables 4.1 to 4. 4.

4.5.3 Performance comparison

In this section, the proposed Case U SSDS $\bar{X}$ scheme is compared to a number of well-known Case U monitoring schemes including the existing NSSDS $\bar{X}$, NSS and side-sensitive synthetic

Shewhart $\bar{X}$, exponentially weighted moving average $\bar{X}$ (denoted as $\bar{X}$-EWMA (λ) where λ represents the smoothing parameter) with λ = 0.1 and 0.5, Cumulative Sum $\bar{X}$ (denoted as $\bar{X}$-CUSUM) monitoring schemes with estimated process parameters. The competing schemes are compared in terms of the *AEQL*, the *ARARL* and *PCI* values. Note that the monitoring scheme with a small *AEQL* value is considered to be superior in performance for the range of shifts under consideration. In this example, the proposed scheme is considered to be the benchmark scheme. Therefore, for the chosen competing schemes, if its *PCI* and *ARARL* values are less than one; then, that particular competing scheme is declared as more efficient than the proposed Case U SSDS $\bar{X}$ scheme. However, if the *PCI* and *ARARL* values are greater than one, then the competing scheme is declared as less efficient than the proposed SSDS $\bar{X}$ scheme. When the *PCI* and *ARARL* values are equal to one, then the competing scheme and the proposed SSDS $\bar{X}$ scheme are equivalent. For a fair comparison, the performance of the competing schemes are investigated when (δ_{min},δ_{max}) = (0,2.5), $m \in \{50,100\}$, $ASS_0 \in \{5,8\}$ corresponding to $n \in \{5,8\}$, $n_1 \in \{3,5\}$, $n_2 \in \{5,8\}$ and a nominal ARL_0 = 370.4. The shifts sizes are divided into three groups which are "small" (0< $\delta \leq$ 0.7), "small to moderate" (0< $\delta \leq$1.6), and "small to large" (0< $\delta \leq$2.5). In Table 4.6, the proposed scheme is compared to the foregoing monitoring schemes in terms of the overall performance. The results corresponding to the best monitoring scheme are highlighted in bold.

Table 4.6: Case U monitoring schemes performance comparison when $n = ASS_0$ = 5, $n_1 \in \{3, 5\}$, $n_2 \in \{5, 8\}$, δ_{min} = 0 and δ_{max} = 2.5 with a nominal ARL_0 of 370.4

*Shift	Performance measures	Control charts: NSS Synthetic $\bar{X}$	SS Synthetic $\bar{X}$	$\bar{X}$-EWMA(0.1)	$\bar{X}$-EWMA(0.5)	$\bar{X}$-CUSUM	NSS DS $\bar{X}$	SSDS $\bar{X}$	(n_1, n_2)	$ASS_0 = n$	m
Small	***AEQL***	98.21	82.33	**70.56**	119.12	96.01	84.13	76.23	**(3, 5)**	**5**	**50**
	ARARL	1.24	1.10	**0.91**	1.48	1.23	1.13	1.00			
	PCI	1.29	1.08	**0.93**	1.56	1.26	1.10	1.00			
Small to moderate	***AEQL***	104.24	80.32	76.79	120.24	103.72	74.22	**70.76**			
	ARARL	1.49	1.17	1.11	1.64	1.50	1.08	**1.00**			
	PCI	1.47	1.14	1.09	1.70	1.47	1.05	**1.00**			
Small to large	***AEQL***	86.04	70.43	94.18	110.44	100.37	67.99	**62.32**			
	ARARL	1.41	1.11	1.47	1.68	1.56	1.13	**1.00**			
	PCI	1.38	1.13	1.51	1.77	1.61	1.09	**1.00**			
Small	***AEQL***	72.89	68.16	**60.30**	73.05	71.18	68.29	62.46	**(5, 8)**	**8**	
	ARARL	1.20	1.14	**0.95**	1.21	1.14	1.12	1.00			
	PCI	1.17	1.09	**0.97**	1.17	1.14	1.09	1.00			
Small to moderate	***AEQL***	71.51	60.16	61.47	79.09	66.34	59.35	**54.24**			
	ARARL	1.34	1.15	1.14	1.52	1.20	1.12	**1.00**			
	PCI	1.32	1.11	1.13	1.46	1.22	1.09	**1.00**			
Small to large	***AEQL***	45.04	40.37	50.56	61.40	59.48	38.55	**33.59**			
	ARARL	1.36	1.31	1.40	1.72	1.63	1.17	**1.00**			
	PCI	1.34	1.20	1.45	1.71	1.65	1.15	**1.00**			
Small	***AEQL***	75.43	67.68	**54.44**	103.57	86.12	71.07	63.40	**(3, 5)**	**5**	**100**
	ARARL	1.17	1.03	**0.82**	1.57	1.32	1.14	1.00			
	PCI	1.19	1.07	**0.86**	1.63	1.36	1.12	1.00			
Small to moderate	***AEQL***	79.69	61.21	59.85	85.59	75.05	56.10	**51.94**			
	ARARL	1.49	1.23	1.21	1.43	1.36	1.13	**1.00**			
	PCI	1.53	1.18	1.15	1.65	1.44	1.08	**1.00**			
Small to large	***AEQL***	67.65	51.18	71.02	83.51	68.71	53.95	**46.82**			
	ARARL	1.48	1.12	1.41	1.74	1.49	1.19	**1.00**			
	PCI	1.44	1.09	1.52	1.78	1.47	1.15	**1.00**			
Small	***AEQL***	72.43	65.68	**53.51**	99.43	88.09	70.12	61.25	**(5, 8)**	**8**	
	ARARL	1.21	1.10	**0.84**	1.58	1.41	1.17	1.00			
	PCI	1.18	1.07	**0.87**	1.62	1.44	1.14	1.00			
Small to moderate	***AEQL***	75.45	59.06	60.76	83.28	71.26	53.23	**49.59**			
	ARARL	1.46	1.21	1.26	1.56	1.47	1.10	**1.00**			
	PCI	1.52	1.19	1.23	1.68	1.44	1.07	**1.00**			
Small to large	***AEQL***	65.69	59.34	68.13	74.79	66.52	51.48	**47.19**			
	ARARL	1.42	1.19	1.45	1.47	1.36	1.06	**1.00**			
	PCI	1.39	1.26	1.44	1.58	1.41	1.09	**1.00**			

* Small: (0<δ ≤0.7), Small to Moderate: (0<δ ≤1.6) and Small to Large: (0<δ ≤2.5).

From Table 4.6, it can be seen that regardless of the sample sizes, the EWMA (0.1) monitoring scheme outperforms the proposed scheme for "small" shifts (δ_{max} = 0.7) in the process mean. However, for "small to moderate" shifts (i.e. δ_{max} = 1.6) as well as for the "small to large" shifts (i.e. δ_{max} = 2.5), the proposed SSDS monitoring scheme outperforms all the competing schemes considered in this paper. These findings are also confirmed in terms of the *ARARL* and *PCI* values. When comparing the existing DS $\bar{X}$ scheme to the proposed scheme, we can observe the following: for "small" shifts, the SSDS $\bar{X}$ monitoring scheme improves the existing DS $\bar{X}$ scheme between 10% and 17%. From "small to moderate" shifts, the overall improvement is between 5% and 9%. From "small to large" shifts, the overall improvement is between 7% and 15%.

4.6 Illustrative example

In this section, the implementation and application of the proposed SSDS $\bar{X}$ scheme is illustrated using the dataset from Zaman et al. (2017). The data gives the information on the inside diameter of cylinder bores in an engine block and contain thirty-five samples, each of size n = 5. In this implementation example, each sample is considered to be a master sample which is divided into two subgroups of sizes 2 and 3 (i.e. n_1 = 2 and n_2 = 3), in stages 1 and 2, respectively, such that $n = n_1 + n_2 = 5$. The estimated IC process mean and standard deviation (using Equations (4.1) and (4.2)) for the inside diameter of cylinder bores are $\hat{\mu}_0$ = 200.15 and $\hat{\sigma}_0$ = 3.47 millimeters (*mm*), respectively. The shift detection ability of the proposed Case U SSDS $\bar{X}$ scheme is also compared to the one of the existing Case U NSSDS $\bar{X}$ scheme.

For (n_1, n_2) = (2, 3) and ASS_0 = 3, the optimal combinations (L_1, L, L_2) of the Case U SSDS $\bar{X}$ scheme and the Case U NSSDS $\bar{X}$ scheme are found to be equal to (2.212, 2.576, 2.305) and (2.306, 2.614, 2.418), respectively, so that these schemes both satisfies ARL_0 = 370.4. A plot of the charting statistics, Z_{1t} and Z_t (i.e., for stages 1 and 2, respectively) of the two monitoring schemes are shown in Figure 4.3. Table 4.7 illustrates the operation of the Case U's Phase II NSSDS and SSDS $\bar{X}$ schemes using the data set on the inside diameter of cylinder bores.

It is seen that the NSSDS $\bar{X}$ scheme does not give a signal at stage 1. However, at the 16th, 19th and 26th sampling time, there was a need for a second sample and the process moved to stage 2. The plotting statistics of the NSSDS $\bar{X}$ scheme at stage 2, Z_t, at the 16th, 19th and 26th sampling time are equal to -0.425, 1.015 and 3.176, respectively. It can be seen that Z_{16} and Z_{19} plot between -L_2 = -2.418 and L_2 = 2.418, which means that the NSSDS $\bar{X}$ scheme does not signal on the 16th and 19th sampling time. Since Z_{26} plots above L_2, the NSSDS $\bar{X}$ scheme gives a signal at the 26th sampling time (see Figure 4.3(a) and Table 4.7) for the first time in Stage 2.

The proposed SSDS $\bar{X}$ scheme moves for the first time to Stage 2 at the 16th sampling time. At this sampling time, $Z_{2,16}$ is equal to -0.425. Since $Z_{16} \in (-L_2,\ L_2)$ = (-2.305, 2.305), the proposed SSDS $\bar{X}$ scheme does not give a signal on the 16th sampling time. However, on the 19th sampling time, $Z_{1,19}$ equal to 2.5864 plots above L = 2.576. Therefore, the SSDS $\bar{X}$ scheme gives a signal for the first time at the 19th sampling time in Stage 1 (see Figure 4.3(b) and Table 4.7).

This example shows that the proposed SSDS $\bar{X}$ scheme is more sensitive than the existing NSSDS $\bar{X}$ scheme in monitoring Phase II samples when the unknown underlying distribution design parameters are estimated from an IC Phase I sample.

Table 4.7: Illustration of the operation of the NSSDS and SSDS $\bar{X}$ schemes using the dataset on the inside diameter of cylinder bores in an engine block

	NSSDS $\bar{X}$ chart								SSDS $\bar{X}$ chart							
Sample No	$\bar{X}_{1t}$	$\hat{Z}_{1t}$	Take a 2nd Sample	$\bar{X}_{2t}$	$\bar{X}_t$	$\hat{Z}_t$	Stage 1: Signal	Stage 2: Signal	$\bar{X}_{1t}$	$\hat{Z}_{1t}$	Take a 2nd Sample	$\bar{X}_{2t}$	$\bar{X}_t$	$\hat{Z}_t$	Stage 1: Signal	Stage 2: Signal
1	203.5	1.3637	N				N		203.5	1.3637	N				N	
2	200.5	0.1409	N				N		200.5	0.1409	N				N	
3	201.5	0.5485	N				N		201.5	0.5485	N				N	
4	204	1.5674	N				N		204	1.5674	N				N	
5	197.5	-1.0818	N				N		197.5	-1.0818	N				N	
6	200.5	0.1409	N				N		200.5	0.1409	N				N	
7	202	0.7523	N				N		202	0.7523	N				N	
8	196.5	-1.4894	N				N		196.5	-1.4894	N				N	
9	199.5	-0.2667	N				N		199.5	-0.2667	N				N	
10	199	-0.4705	N				N		199	-0.4705	N				N	
11	204.5	1.7712	N				N		204.5	1.7712	N				N	
12	200.5	0.1409	N				N		200.5	0.1409	N				N	
13	200.5	0.1409	N				N		200.5	0.1409	N				N	
14	200.5	0.1409	N				N		200.5	0.1409	N				N	
15	200	-0.0629	N				N		200	-0.0629	N				N	
16	194	-2.5084	**Y**	203.67	199.8	-0.4253	N	N	194	-2.5084	**Y**	203.67	199.8	-0.4253	N	N
17	202	0.7523	N				N		202	0.7523	N				N	
18	199.5	-0.2667	N				N		199.5	-0.2667	N				N	
19	206.5	2.5864	**Y**	197.33	201	1.0152	N	N	206.5	2.5864	N				**Y**	
20	202	0.7523	N				N		202	0.7523	N				N	
21	201.5	0.5485	N				N		201.5	0.5485	N				N	
22	199.5	-0.2667	N				N		199.5	-0.2667	N				N	
23	198	-0.8780	N				N		198	-0.8780	N				N	
24	199	-0.4705	N				N		199	-0.4705	N				N	
25	200	-0.0629	N				N		200	-0.0629	N				N	
26	206	2.3826	**Y**	200.67	202.8	3.1760	N	**Y**	206	2.3826	**Y**	200.67	202.8	3.1760	N	**Y**
27	203.5	1.3637	N				N		203.5	1.3637	N				N	
28	200	-0.0629	N				N		200	-0.0629	N				N	
29	198.5	-0.6743	N				N		198.5	-0.6743	N				N	
30	200	-0.0629	N				N		200	-0.0629	N				N	
31	200	-0.0629	N				N		200	-0.0629	N				N	
32	195.5	-1.8970	N				N		195.5	-1.8970	N				N	
33	200.5	0.1409	N				N		200.5	0.1409	N				N	
34	199	-0.4705	N				N		199	-0.4705	N				N	
35	202	0.7523	N				N		202	0.7523	N				N	

Note: N = No and Y = Yes.

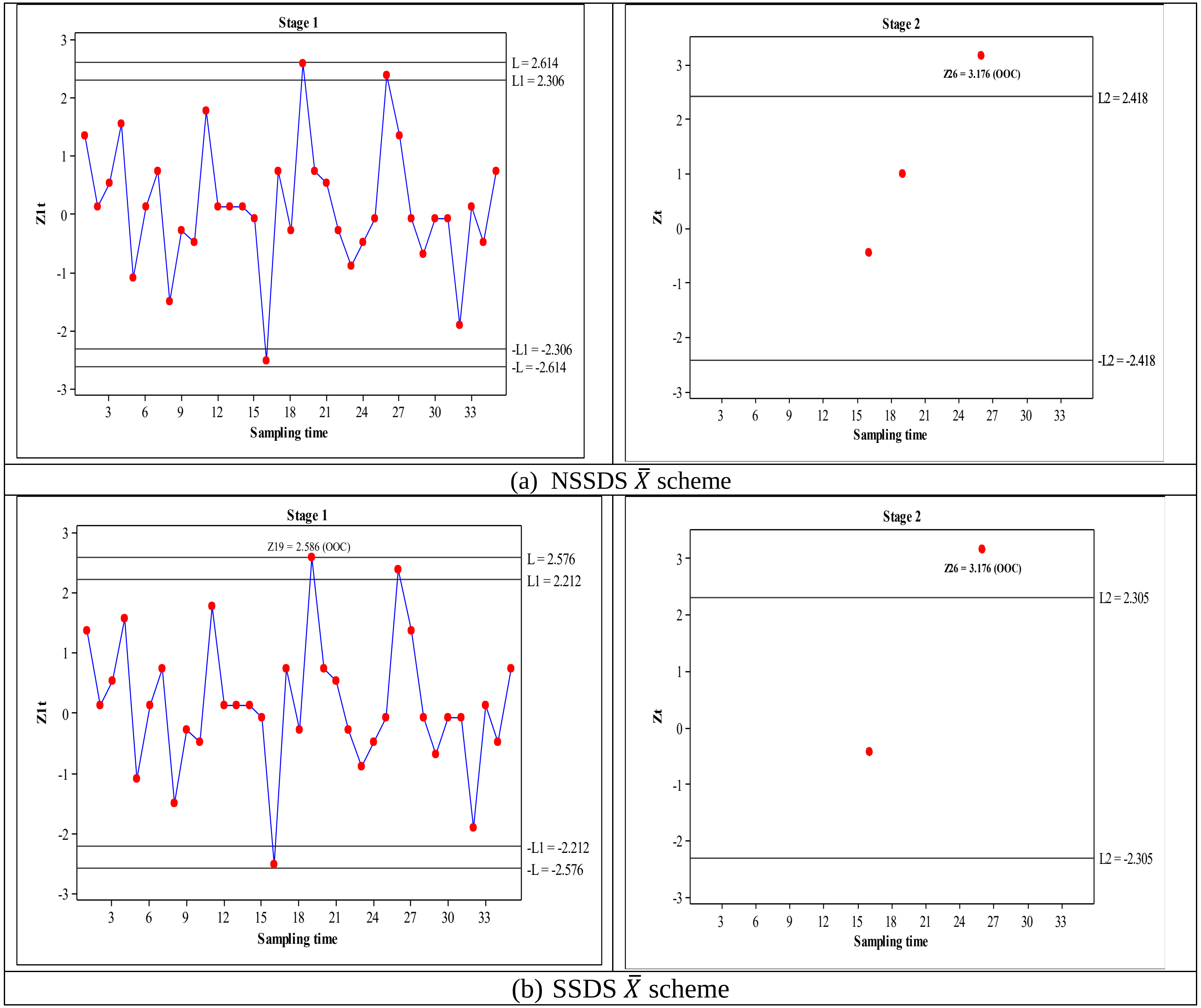


Figure 4.3: The Case U NSSDS and SSDS $\bar{X}$ schemes using the data set on the inside diameter of cylinder bores in an engine block

4.7 Conclusion

In this chapter, a SSDS $\bar{X}$ monitoring scheme is proposed under the assumption of unknown process parameters. The performance of the SSDS scheme is investigated in terms of the different characteristics of the run-length distribution (i.e. *ARL*, *SDRL*, *PRL*) as well as the *ANOS*, *AEQL*, *PCI* and *ARARL* values. The proposed SSDS $\bar{X}$ scheme outperforms the competing schemes considered in this paper in many situations. Moreover, in terms of the *ASS* and *ANOS* values, the proposed SSDS scheme is found to be cost effective and sensitive compared to the competing schemes considered in this chapter. Practitioners in the industrial and non-industrial environments are recommended that, when the underlying process parameters are unknown and need to be estimated, the reference sample size *m* must be sufficiently large (i.e., *m* greater or equal 100) in order to get more accurate estimates and stability in the performance of the proposed SSDS scheme as for such large Phase I sample sizes, the resulting performance is closer to the parameters known scenario.

Chapter 5: Concluding remarks and future research ideas

Shewhart-type double sampling monitoring schemes have been shown to yield better OOC performance than their corresponding basic Shewhart schemes when the process parameters are known (i.e. Case K) and unknown (i.e. Case U) regardless of the nature of the underlying process distribution. This strategy has been proven to effectively decrease the sampling effort and, at the same time, to decrease the time to detect potential out-of-control situations. For these reasons, it has received some attention in the statistical process monitoring. Moreover, from the works of Yang and Wu (2017a, b) it has been shown that this remains valid in the case of the double sampling schemes based on memory-type schemes. Also, the double sampling schemes combined with other designs / schemes (e.g. VSI, VSSI, synthetic, group-runs) have an even better OOC performance than their individual integrated schemes. Thus, this indicates that these monitoring schemes can be more useful in many real-life applications where the traditional Shewhart scheme is currently in use. Implementation tools need to be developed (e.g. using statistical packages like R, Minitab, SAS, Matlab, SPSS, etc.) so that these monitoring schemes can be implemented in real-life monitoring of some online real time applications.

Consequently, while a majority of research works has been dedicated to non-side-sensitive double sampling (i.e., NSSDS) design, very little has been dedicated to the design of side-sensitive double sampling (i.e., SSDS) Shewhart-type scheme. This current dissertation investigates the latter and shows that the integration of side-sensitive design to the traditional Shewhart-type double sampling scheme improves considerably its sensitivity in monitoring unexpected shifts in the location parameters. Moreover, in Case U, the estimation of the process parameters has a negative effect on the performance of the proposed SSDS monitoring scheme. In this case, operators in the industries are advised to use some high desired number of Phase I observations to guarantee stability and better performance. Therefore, the investigation of the design of monitoring schemes in Case U is needed for all types of schemes. In addition, while SPM literature shows that there are a number of estimated parameter(s) research works for double sampling schemes – except the NSSDS S^2 scheme by Castagliola et al. (2017), these are only for the univariate process location, with none dedicated to monitoring both the location and variability simultaneously, as well as the coefficient of variation, etc.

Finally, a list of some possible future research ideas that may be of interest to researchers who are interested in pursuing enhancements of double sampling monitoring schemes are listed below:

1. Majority of double sampling schemes are based on monitoring of the process location parameter for a normally distributed i.i.d. process. There are a number of pitfalls in ignoring or assuming that the corresponding standard deviation is constant or unaffected by changes in location parameter. Thus, future research works need to focus more on monitoring both the location and variability parameters simultaneously in the case of non-normal distributions (e.g. Burr's XII distributions), as well as using for instance, monitoring time between events using, say, the exponential distribution; etc.
2. Double sampling schemes are mostly based on the assumption that the subgroup samples do not have either autocorrelation (within-sample correlation) or cross-correlation (between-samples correlation). However, for sequential observations, there tend to be some inherent underlying correlation within the observations – see for instance, Qiu (2019). Therefore, it is important in the future to focus on double sampling schemes with more focus on the autocorrelated observations as well as on nonparametric or distribution-free monitoring schemes.
3. With the exception of Haq and Khoo (2018, 2019), no other double sampling scheme takes into account auxiliary information. There is only a single research work that takes into account measurement errors, i.e. Lee et al. (2019). Considering the importance of auxiliary information and measurement errors in real-life applications; these important factors require more attention for double sampling schemes.
4. With only eight publications on multivariate schemes, there is a lot of research in double sampling that need to be done based on parametric and nonparametric multivariate double sampling schemes – with Qiu (2014) being the more appropriate starting point. Moreover, for the parametric case, there is a need for a multivariate double sampling schemes based on monitoring both location and variability simultaneously, as well as the coefficient of variation.
5. Since the combined schemes usually perform better than the individual integrated schemes, a fact that has been shown in the case of double samples with synthetic, VSI, VSSI and group-runs schemes. The latter statement needs to be tested whether it holds in the case of memory-type schemes (i.e. exponentially weighted moving average (EWMA), cumulative sum (CUSUM), generally weighted moving average (GWMA),

homogeneously weighted moving average (HWMA)). The only publications that have done this so far in the literature are on distribution-free method using the EWMA double sampling scheme; see Yang and Wu (2017a, b). The latter methodologies need to be adopted for the parametric scenarios and also be extended for other nonparametric scenarios. Moreover, for complex double sampling schemes, research may investigate the possibility of combining the synthetic or group-runs double sampling schemes with the VSI or VSSI designs.

6. Only a few studies on the economic and economic-statistical designs for the process location (those corresponding to the double sampling $\bar{X}$ and T^2 schemes only) have been done in the literature. Hence, more needs to be done especially when there is no assumption of i.i.d. and normality; and more importantly, when parameters are estimated.
7. With only a few studies on attributes data, more investigations are required in the area of double sampling schemes, specifically based on the number of nonconformities as well as high-yield processes, see the review on attributes data by Woodall (1997) as a possible starting point.

In closing, the main objective of this study was first to introduce a new double sampling monitoring scheme under both Case K and Case U. Second, to give a more intensive review as well as more detailed background on this important class Shewhart-type schemes; with hope that this will stimulate more future research on simple as well as complex double sampling schemes (especially using the newly proposed SSDS design) for monitoring a variety of quality characteristics.

Appendices

In this section, there are 3 appendices, i.e. Appendix A, B and C. In Appendix A, some illustrations are given to show how the expressions were implemented in MATHCAD®14 software to calculate the *ARL*, *ASS*, *ANOS*, *SDRL* and *EQL*.

Appendix A: MATHCAD explicit formulas

To illustrate how the empirical values in the main chapter were calculated, the following metrics are used to show how the formulas are entered in MATHCAD®14: *ARL*, *ASS*, *ANOS*, *SDRL* and *EQL*.

A1. *ARL* formula from Equation (3.6)

$$\mathrm{arl}(n1,n2,l1,l,l2,d) := \frac{1}{1-\left[\mathrm{pnorm}\left(l1+d\cdot\sqrt{n1},0,1\right)-\mathrm{pnorm}\left(-l1+d\cdot\sqrt{n1},0,1\right)+\int_{l1+d\cdot\sqrt{n1}}^{l+d\cdot\sqrt{n1}}\left[\mathrm{pnorm}\left[\sqrt{\frac{(n1+n2)}{n2}}\cdot l2-z\cdot\sqrt{\frac{n1}{n2}}+\frac{d\cdot(n1+n2)}{\sqrt{n2}},0,1\right]\right]\cdot\mathrm{dnorm}(z,0,1)\,dz+\int_{-l+d\cdot\sqrt{n1}}^{-l1+d\cdot\sqrt{n1}}\left[1-\mathrm{pnorm}\left[-\sqrt{\frac{(n1+n2)}{n2}}\cdot l2-z\cdot\sqrt{\frac{n1}{n2}}+\frac{d\cdot(n1+n2)}{\sqrt{n2}},0,1\right]\right]\cdot\mathrm{dnorm}(z,0,1)\,dz\right]}$$

$d := 0, 0.1 .. 2.5$

arl(2, 2, 2.9101, 3.0568, 2.4050, d)

370.394
333.364
253.872
177.41
120.604
82.088
56.618
39.754
28.462
20.784
15.48
11.753
9.093
7.165
5.746
...

The above are the resulting values of the *ARL* when δ varies from 0 to 2.5 with an increment of 0.1, where specifically, n_1=2, n_2=2, L_1=2.9101, L_2=2.4050 and L=3.0568.

A2. *ASS* formula from Equation (3.8)

$$\mathrm{ass}(n1,n2,l1,l,l2,d) := n1 + n2\cdot\left(\mathrm{pnorm}\left(l + d\cdot\sqrt{n1},0,1\right) - \mathrm{pnorm}\left(l1 + d\cdot\sqrt{n1},0,1\right) + \mathrm{pnorm}\left(-l1 + d\cdot\sqrt{n1},0,1\right) - \mathrm{pnorm}\left(-l + d\cdot\sqrt{n1},0,1\right)\right)$$

$$d := 0, 0.1 .. 2.5$$

$$\mathrm{ass}(2,2,2.9101,3.0568,2.4050,d)$$

2.003
2.003
2.004
2.005
2.007
2.009
2.012
2.016
2.021
2.027
2.034
2.042
2.051
2.061
2.071
...

The above are the resulting values of the *ASS* when δ varies from 0 to 2.5 with an increment of 0.1, where specifically, n_1=2, n_2=2, L_1=2.9101, L_2=2.4050 and L=3.0568.

A3. *ANOS* formula from Equation (3.9)

anos (n1 , n2 , l1, l , l2, d) := ass (n1 , n2 , l1, l , l2, d)·arl(n1 , n2 , l1, l , l2, d)

d := 0, 0.1.. 2.5

anos (2, 2, 2.9101, 3.0568, 2.405, d)

741.807
667.718
508.668
355.673
241.996
164.909
113.921
80.15
57.525
42.133
31.489
24.003
18.652
14.765
11.899
...

The above are the resulting values of the *ANOS* when δ varies from 0 to 2.5 with an increment of 0.1, where specifically, n_1=2, n_2=2, L_1=2.9101, L_2=2.4050 and L=3.0568.

A4. *SDRL* formula from Equation (3.7)

$$\mathrm{pa}(n1,n2,l1,l,l2,d) := \mathrm{pnorm}\left(l1+d\cdot\sqrt{n1},0,1\right) - \mathrm{pnorm}\left(-l1+d\cdot\sqrt{n1},0,1\right) + \int_{l1+d\cdot\sqrt{n1}}^{l+d\cdot\sqrt{n1}} \left[\mathrm{pnorm}\left[\sqrt{\frac{(n1+n2)}{n2}}\cdot l2 - z\cdot\sqrt{\frac{n1}{n2}} + \frac{d\cdot(n1+n2)}{\sqrt{n2}},0,1\right]\right]\cdot \mathrm{dnorm}(z,0,1)\,dz + \int_{-l+d\cdot\sqrt{n1}}^{-l1+d\cdot\sqrt{n1}} \left[1-\mathrm{pnorm}\left[-\sqrt{\frac{(n1+n2)}{n2}}\cdot l2 - z\cdot\sqrt{\frac{n1}{n2}} + \frac{d\cdot(n1+n2)}{\sqrt{n2}},0,1\right]\right]\cdot \mathrm{dnorm}(z,0,1)\,dz$$

$$\mathrm{sdrl1}(d) := \frac{\sqrt{\mathrm{pa}(2,2,2.9101,3.0568,2.4050,d)}}{1-\mathrm{pa}(2,2,2.9101,3.0568,2.4050,d)}$$

$$d := 0, 0.1 .. 2.5$$

sdrl1(d) =

369.894
332.864
253.372
176.91
120.103
81.586
56.116
39.251
27.957
20.278
14.971
11.242
8.579
6.646
5.222
4.158
3.351
2.73
2.247
1.867
1.563
1.319
...

The above are the resulting values of the *SDRL* when δ varies from 0 to 2.5 with an increment of 0.1, where specifically, n_1=2, n_2=2, L_1=2.9101, L_2=2.4050 and L=3.0568.

A5. *EQL* formula from Equation (3.14)

$c1 := 1$

$c2 := 0.1$

$k := 1, 2 .. 26$

$arlrss(k) := arl[3, 11, 0.3486, 3.8211, 2.9827, [(k - 1) \cdot 0.1]]$

$dsq(k, c1, c2) := [(k - c1) \cdot c2] \cdot [(k - c1) \cdot c2]$

$$eqltot(x) := \sum_{k} \frac{[dsq(k, c1, c2) \cdot arlrss(k) \cdot (k \le x)]}{2.5}$$

$eqltot(25) = 26.482$

The above are the resulting value of the *EQL* when δ varies from δ_{min}=0 to δ_{max}=2.5 with an increment of 0.1, where specifically, n_1=3, n_2=11, L_1=0.3486, L_2=2.9827 and L=3.8211.

Appendix B: Derivations of the SSDS schemes' run-length expressions in Case K

Note that the expressions in Chapter 3 assume that the unstandardized means are $\bar{Y}_{1t} = \sum_{j=1}^{n_1} Y_{1tj}/n_1$, $\bar{Y}_{2t} = \sum_{j=1}^{n_2} Y_{1tj}/n_2$ and $\bar{Y}_t = (n_1\bar{Y}_{1t} + n_2\bar{Y}_{2t})/(n_1 + n_2)$. The corresponding standardized plotting statistics are Z_{1t}, Z_{2t} and Z_t in stages 1 and 2, which are given by

$$Z_{1t} = \frac{\bar{X}_{1t} - \mu_0}{\sigma_0/\sqrt{n_1}}$$
$$Z_{2t} = \frac{\bar{X}_{2t} - \mu_0}{\sigma_0/\sqrt{n_2}} \qquad \text{(B1)}$$
$$Z_t = \frac{\bar{X}_t - \mu_0}{\sigma_0/\sqrt{n_1 + n_2}}$$

respectively.

Based on the latter, the warning and control limits in Figure 3.1 are also standardized. On the contrary, the warning and control limits for unstandardized observations (i.e. $\bar{X}_{1t}$ and $\bar{X}_t$ in stages 1 and 2, respectively) are as given in Figure B1; where the upper & lower control limits in stage 1 are denoted by UCL_1 & LCL_1, the upper & lower warning limits in stage 1 are denoted by UWL & LWL, the upper & lower control limit in stage 2 are denoted by UCL_2 & LCL_2, respectively.

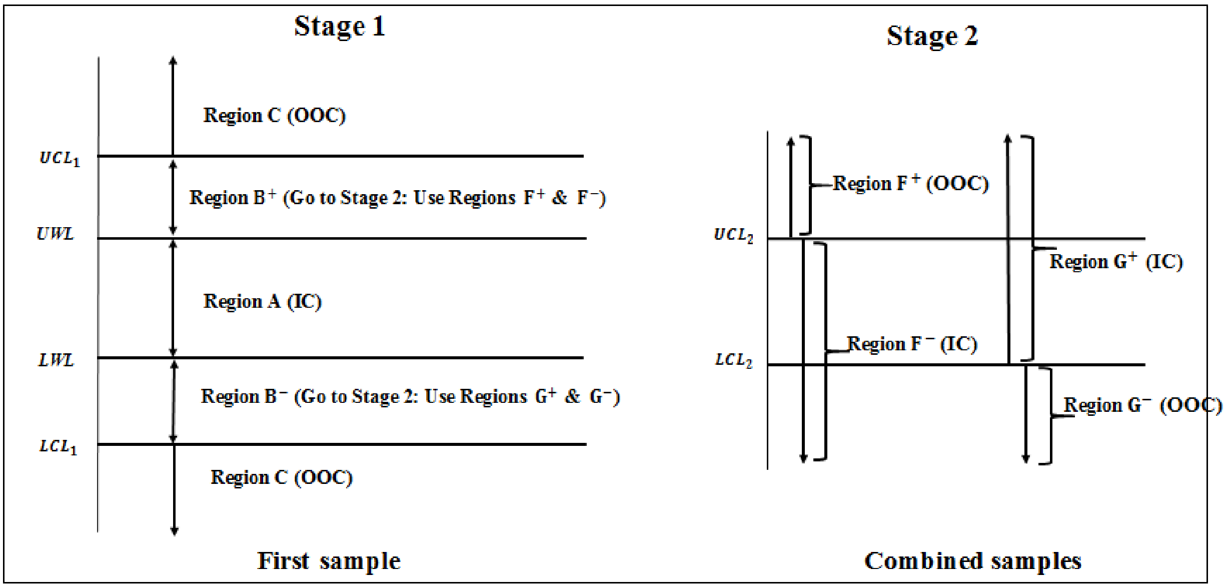


Figure B1: The charting limits for the SSDS scheme for unstandardized observations

The limits in Figure B1for stage 1 of the SSDS monitoring scheme are given by

$$\begin{aligned} UCL_1 &= \mu_0 + L\frac{\sigma_0}{\sqrt{n_1}} \\ UWL &= \mu_0 + L_1\frac{\sigma_0}{\sqrt{n_1}} \\ LWL &= \mu_0 - L_1\frac{\sigma_0}{\sqrt{n_1}} \\ LCL_1 &= \mu_0 - L\frac{\sigma_0}{\sqrt{n_1}} \end{aligned} \qquad \text{(B2)}$$

and those for stage 2 are given by

$$\begin{aligned} UCL_2 &= \mu_0 + L_2\frac{\sigma_0}{\sqrt{n_1 + n_2}} \\ LCL_2 &= \mu_0 - L_2\frac{\sigma_0}{\sqrt{n_1 + n_2}}. \end{aligned} \qquad \text{(B3)}$$

Now that the above unstandardized charting regions have been defined; next, the expressions used in the main chapters are shown how they were derived.

Lemma B1

From Equation (3.1) in Chapter 3, it is given that

$$P_{01} = P[Z_{1t} \in \mathrm{A}] = \Phi\left(L_1 + \delta\sqrt{n_1}\right) - \Phi\left(-L_1 + \delta\sqrt{n_1}\right)$$

i.e. this is the probability that the process is IC in stage 1, hence no need for stage 2.

Proof:

$$\begin{aligned}
P_{01} &= P(Z_{1t} \in \mathrm{A}) \\
&= P(-L_1 < Z_{1t} < L_1) \\
&= P\left(-L_1 < \frac{\bar{X}_{1t} - \mu_0}{\frac{\sigma_0}{\sqrt{n_1}}} < L_1\right) \ \{\text{using Equation (B1)}\} \\
&= P\left(\mu_0 - L_1 \frac{\sigma_0}{\sqrt{n_1}} < \bar{X}_{1t} < \mu_0 + L_1 \frac{\sigma_0}{\sqrt{n_1}}\right) \\
&= P\left(\frac{\left(\mu_0 - L_1 \frac{\sigma_0}{\sqrt{n_1}}\right) - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} < \frac{\bar{X}_{1t} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} \right. \\
&\qquad\qquad \left. < \frac{\left(\mu_0 + L_1 \frac{\sigma_0}{\sqrt{n_1}}\right) - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}}\right) \\
&= \Phi\left(L_1 + \delta\sqrt{n_1}\right) - \Phi\left(-L_1 + \delta\sqrt{n_1}\right); \ \text{ i.e., } \frac{\bar{X}_{1t} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} \sim N(\delta, 1).
\end{aligned}$$

Lemma B2

From Equation (3.2) in Chapter 3, it is given that

$$P_{02} = P[Z_{1t} \in \mathrm{B}^+ \text{ and } Z_t \in \mathrm{F}^-] + P[Z_{1t} \in \mathrm{B}^- \text{ and } Z_t \in \mathrm{G}^+]$$

$$= \int_{Z_{1t} \in \mathrm{B}^{++}} \left\{\Phi\left(cL_2 + rc\delta - z\sqrt{n_1/n_2}\right)\right\} \phi(z)dz$$

$$+ \int_{Z_{1t} \in \mathrm{B}^{--}} \left\{1 - \Phi\left(-cL_2 + rc\delta - z\sqrt{n_1/n_2}\right)\right\} \phi(z)dz.$$

i.e. this is the probability that the process is IC in stage 2, given that in stage 1 it plotted in either region B^+ or B^-; and $\phi(z)$ is the p.d.f. of the standard normal distribution.

Proof:

Since, $P_{02} = P[Z_{1t} \in \mathrm{B}^+ \text{ and } Z_t \in \mathrm{F}^-] + P[Z_{1t} \in \mathrm{B}^- \text{ and } Z_t \in \mathrm{G}^+]$ then we first consider, the first part of the equation, i.e., $P[Z_{1t} \in \mathrm{B}^+ \text{ and } Z_t \in \mathrm{F}^-]$. From the latter, we start with stage 2 component, i.e. $P[Z_t \in \mathrm{F}^-]$.

$$P(Z_t \in \mathrm{F}^-)$$

$$= P(Z_t < L_2)$$

$$= P\left(\frac{\bar{X}_t - \mu_0}{\frac{\sigma_0}{\sqrt{n_1 + n_2}}} < L_2\right)$$

$$= P\left(\bar{X}_t < \mu_0 + L_2 \frac{\sigma_0}{\sqrt{n_1 + n_2}}\right)$$

$$= P\left(\frac{n_1\bar{X}_{1t} + n_2\bar{X}_{2t}}{n_1 + n_2} < \mu_0 + L_2 \frac{\sigma_0}{\sqrt{n_1 + n_2}}\right)$$

$$= P\left(\frac{n_1\big(\bar{X}_{1t} - (\mu_0 - \delta\sigma_0)\big) + n_2\big(\bar{X}_{2t} - (\mu_0 - \delta\sigma_0)\big)}{n_1 + n_2} + (\mu_0 - \delta\sigma_0) \right.$$

$$\left. < \mu_0 + L_2 \frac{\sigma_0}{\sqrt{n_1 + n_2}}\right)$$

$$= P\left(\frac{n_1\left(\frac{\sigma_0}{\sqrt{n_1}} Z_{1t}\right) + n_2\left(\frac{\sigma_0}{\sqrt{n_1}} Z_{2t}\right)}{n_1 + n_2} < \mu_0 + L_2 \frac{\sigma_0}{\sqrt{n_1 + n_2}} - \mu_0 + \delta\sigma_0\right);$$

i.e., $Z_{it} = \frac{\bar{X}_{it} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} \sim N(\delta, 1)$, for $i = 1$ and 2, hence $\frac{\sigma_0}{\sqrt{n_i}} Z_{it} = \bar{X}_{it} - (\mu_0 - \delta\sigma_0)$.

$$= P\left(\frac{\sigma_0}{n_1 + n_2}\left(\sqrt{n_1} Z_{1t} + \sqrt{n_2} Z_{2t}\right) < \sigma_0\left(\frac{L_2}{\sqrt{n_1 + n_2}} + \delta\right)\right)$$

$$= P\left(Z_{2t} < L_2\sqrt{\frac{n_1 + n_2}{n_2}} + \delta\frac{(n_1 + n_2)}{n_2} - Z_{1t}\sqrt{\frac{n_1}{n_2}}\right)$$

Consequently, it follows that

$$P(Z_t < L_2 | Z_{1t} = z) = P\left(Z_{2t} < L_2\sqrt{\frac{n_1+n_2}{n_2}} + \delta\frac{(n_1+n_2)}{n_2} - z\sqrt{\frac{n_1}{n_2}}\right)$$

$$= \Phi\left(L_2\sqrt{\frac{n_1+n_2}{n_2}} + \delta\frac{(n_1+n_2)}{\sqrt{n_2}} - z\sqrt{\frac{n_1}{n_2}}\right)$$

Since $P[Z_{1t} \in \mathrm{B}^+$ and $Z_t \in \mathrm{F}^-]$ is the probability that a sample plots in region F^- in stage 2 given that in stage 1, a sample plotted in region B^+ (i.e. this region is denoted by $\mathrm{B}^{++} = (L_1 + \delta\sqrt{n_1}, L + \delta\sqrt{n_1}]$), and the standardized values are used, then

$$P[Z_{1t} \in \mathrm{B}^+ \text{ and } Z_t \in \mathrm{F}^-] = \int_{Z_{1t}\in \mathrm{B}^{++}} P(Z_{2t} \in \mathrm{F}^- | Z_{1t} = z)\, \phi(z)dz$$

$$= \int_{Z_{1t}\in \mathrm{B}^{++}} \Phi\left(L_2\sqrt{\frac{n_1+n_2}{n_2}} + \delta\frac{(n_1+n_2)}{\sqrt{n_2}} - z\sqrt{\frac{n_1}{n_2}}\right) \phi(z)dz$$

By letting $r^2 = n_1 + n_2$, $c = r/\sqrt{n_2}$, hence it follows that

$$P[Z_{1t} \in \mathrm{B}^+ \text{ and } Z_t \in \mathrm{F}^-] = \int_{Z_{1t}\in \mathrm{B}^{++}} \Phi\left(cL_2 + rc\delta - z\sqrt{n_1/n_2}\right)\, \phi(z)dz. \qquad \text{(B7)}$$

$$P[Z_{1t} \in \mathrm{B}^+ \text{ and } Z_t \in \mathrm{F}^-] = \int_{Z_{1t}\in \mathrm{B}^{++}} \Phi\left(cL_2 + rc\delta - z\sqrt{n_1/n_2}\right)\, \phi(z)dz.$$

Similarly, for $P[Z_{1t} \in \mathrm{B}^-$ and $Z_t \in \mathrm{G}^+]$, we first consider

$$P(\,Z_t \in \mathrm{G}^+) = P\left(\frac{\bar{X}_t - \mu_0}{\frac{\sigma_0}{\sqrt{n_1+n_2}}} > -L_2\right)$$

From the fact that $\bar{X}_t = \frac{n_1\bar{X}_{1t} + n_2\bar{X}_{2t}}{n_1+n_2}$ it follows that

$$= P\left(\frac{n_1\bar{X}_{1t} + n_2\bar{X}_{2t}}{n_1+n_2} > \mu_0 - L_2\frac{\sigma_0}{\sqrt{n_1+n_2}}\right)$$

$$= P\left(\frac{n_1\left(\bar{X}_{1t} - (\mu_0 - \delta\sigma_0)\right) + n_2\left(\bar{X}_{2t} - (\mu_0 - \delta\sigma_0)\right)}{n_1+n_2} + (\mu_0 - \delta\sigma_0) > \mu_0 - L_2\frac{\sigma_0}{\sqrt{n_1+n_2}}\right)$$

$$= P\left(\frac{\sqrt{n_1}Z_1 + \sqrt{n_2}Z_2}{n_1+n_2} > -L_2\frac{1}{\sqrt{n_1+n_2}} + \delta\right)$$

$$= P\left(\sqrt{\frac{n_1}{n_2}}Z_{1t} + Z_{2t} > -L_2\sqrt{\frac{n_1+n_2}{n_2}} + \delta\frac{(n_1+n_2)}{\sqrt{n_2}}\right)$$

$$= 1 - P\left(Z_{2t} < -L_2\sqrt{\frac{n_1+n_2}{n_2}} + \delta\frac{(n_1+n_2)}{\sqrt{n_2}} - Z_{1t}\sqrt{\frac{n_1}{n_2}}\right)$$

Since $P[Z_{1t} \in \mathrm{B}^-$ and $Z_t \in \mathrm{G}^+]$ is the probability that a sample plots in region G^+ in stage 2 given that in stage 1, a sample plotted in region B^- (i.e. this region is denoted by $\mathrm{B}^{--} = (-L + \delta\sqrt{n_1}, -L_1 + \delta\sqrt{n_1}]$), and the standardized values are used, then

$$P[Z_{1t} \in \mathrm{B}^- \text{ and } Z_t \in \mathrm{G}^+] = \int_{Z_{1t}\in\mathrm{B}^{--}} P[Z_t \in \mathrm{G}^+ | Z_{1t} = z]\, \phi(z)dz$$

$$= \int_{Z_{1t}\in\mathrm{B}^{--}} \left\{1 - \Phi\left(L_2\sqrt{\frac{n_1+n_2}{n_2}} + \delta\frac{(n_1+n_2)}{\sqrt{n_2}} - z\sqrt{\frac{n_1}{n_2}}\right)\right\} \phi(z)dz$$

By letting $r^2 = n_1 + n_2$, $c = r/\sqrt{n_2}$, hence it follows that

$$P[Z_{1t} \in \mathrm{B}^- \text{ and } Z_t \in \mathrm{G}^+] = \int_{Z_{1t}\in\mathrm{B}^{--}} \left\{1 - \Phi\left(cL_2 + rc\delta - z\sqrt{\frac{n_1}{n_2}}\right)\right\} \phi(z)dz. \qquad \text{(B8)}$$

Therefore, using Equations (B7) and (B8), then it follows that Lemma 2 is proved.

Lemma B3

From Equation (3.3) in Chapter 3, it is given that

$$P_0 = \Phi\left[L_1 + \delta\sqrt{n_1}\right] - \Phi\left[-L_1 + \delta\sqrt{n_1}\right] + \int_{Z_{1t}\in \mathrm{B}^{++}} \left\{\Phi\left(cL_2 + rc\delta - z\sqrt{n_1/n_2}\right)\right\} \phi(z)dz + \int_{Z_{1t}\in \mathrm{B}^{--}} \left\{1 - \Phi\left(-cL_2 + rc\delta - z\sqrt{n_1/n_2}\right)\right\} \phi(z)dz$$

i.e. this is the probability that the process is IC in both stages 1 and 2, given that in stage 1 it plotted in either region B^{+} or B^{-}.

Proof:

The proof follows directly from Lemmas 1 and 2, since $P_0 = P_{01} + P_{02}$.

Lemma B4

From Equation (3.8) in Chapter 3, it is given that

$$P_2 = \Phi\left(L + \delta\sqrt{n_1}\right) - \Phi\left(L_1 + \delta\sqrt{n_1}\right) + \Phi\left(-L_1 + \delta\sqrt{n_1}\right) - \Phi\left(-L + \delta\sqrt{n_1}\right)$$

i.e. this is the probability of taking the second sample.

Proof:

$$\begin{aligned}
P_2 &= P(Z_{1t} \in \mathrm{B}^+ \cup \mathrm{B}^-) \\
&= P(Z_{1t} \in \mathrm{B}^+) + P(Z_{1t} \in \mathrm{B}^-) \\
&= P(L_1 < Z_{1t} < L) + P(-L < Z_{1t} < -L_1) \\
&= P\left(L_1 < \frac{\bar{X}_{1t} - \mu_0}{\frac{\sigma_0}{\sqrt{n_1}}} < L\right) + P\left(-L < \frac{\bar{X}_{1t} - \mu_0}{\frac{\sigma_0}{\sqrt{n_1}}} < -L_1\right) \\
&= P\left(\mu_0 + L_1\frac{\sigma_0}{\sqrt{n_1}} < \bar{X}_{1t} < \mu_0 + L\frac{\sigma_0}{\sqrt{n_1}}\right) + P\left(\mu_0 - L\frac{\sigma_0}{\sqrt{n_1}} < \bar{X}_{1t} < \mu_0 - L_1\frac{\sigma_0}{\sqrt{n_1}}\right)
\end{aligned}$$

Then using Equation (B5), it follows that

$$\begin{aligned}
P_2 = P&\left(\frac{\mu_0 + L_1\frac{\sigma_0}{\sqrt{n_1}} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} < \frac{\bar{X}_{1t} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} < \frac{\mu_0 + L\frac{\sigma_0}{\sqrt{n_1}} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}}\right) \\
&+ P\left(\frac{\mu_0 - L\frac{\sigma_0}{\sqrt{n_1}} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} < \frac{\bar{X}_{1t} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} < \frac{\mu_0 - L_1\frac{\sigma_0}{\sqrt{n_1}} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}}\right) \\
&= \left(\Phi\left(L + \delta\sqrt{n_1}\right) - \Phi\left(L_1 + \delta\sqrt{n_1}\right)\right) + \left(\Phi\left(-L_1 + \delta\sqrt{n_1}\right) - \Phi\left(-L + \delta\sqrt{n_1}\right)\right).
\end{aligned}$$

Appendix C: Derivations of the SSDS schemes' run-length expressions in Case U

Note that the expressions in Chapter 4 assume that the unstandardized means are $\bar{Y}_{1t}$ and $\bar{Y}_{2t}$ in stages 1 and 2, respectively. The unstandardized combined process means in stages 1 and 2 is obtained by

$$\bar{Y}_t = \frac{n_1\bar{Y}_{1t} + n_2\bar{Y}_{2t}}{n_1 + n_2}. \tag{C1}$$

Consequently, the corresponding standardized process means $\hat{Z}_{1t}$, $\hat{Z}_{2t}$ and $\hat{Z}_t$ are given by

$$\begin{aligned} \hat{Z}_{1t} &= \frac{\bar{Y}_{1t} - \hat{\mu}_0}{\hat{\sigma}_0/\sqrt{n_1}} \\ \hat{Z}_{2t} &= \frac{\bar{Y}_{2t} - \hat{\mu}_0}{\hat{\sigma}_0/\sqrt{n_2}} \\ \hat{Z}_t &= \frac{\bar{Y}_t - \hat{\mu}_0}{\hat{\sigma}_0/\sqrt{n_1 + n_2}}, \end{aligned} \tag{C2}$$

respectively.

The limits in Figure B1for stage 1 of the SSDS monitoring scheme, in Case U, are given by

$$\begin{aligned} U\hat{C}L_1 &= \hat{\mu}_0 + L\frac{\hat{\sigma}_0}{\sqrt{n_1}} \\ U\widehat{W}L &= \hat{\mu}_0 + L_1\frac{\hat{\sigma}_0}{\sqrt{n_1}} \\ L\widehat{W}L &= \hat{\mu}_0 - L_1\frac{\hat{\sigma}_0}{\sqrt{n_1}} \\ L\hat{C}L_1 &= \hat{\mu}_0 - L\frac{\hat{\sigma}_0}{\sqrt{n_1}} \end{aligned} \tag{C3}$$

and those for stage 2, in Case U, are given by

$$\begin{aligned} U\hat{C}L_2 &= \hat{\mu}_0 + L_2\frac{\hat{\sigma}_0}{\sqrt{n_1 + n_2}} \\ L\hat{C}L_2 &= \hat{\mu}_0 - L_2\frac{\hat{\sigma}_0}{\sqrt{n_1 + n_2}}. \end{aligned} \tag{C4}$$

Lemma C1

From Equation (4.3) in Chapter 4, it is given that

$$\hat{P}_{01} = P\big[\hat{Z}_{1t} \in \mathrm{A}\big] = \Phi\left(U\sqrt{\frac{n_1}{mn}} + VL_1 - \delta\sqrt{n_1}\right) - \Phi\left(U\sqrt{\frac{n_1}{mn}} - VL_1 - \delta\sqrt{n_1}\right)$$

i.e. this is the probability that the process is IC in stage 1, hence no need for stage 2.

Proof:

$$\hat{P}_{01} = P\big(\hat{Z}_{1t} \in \mathrm{A}\big|\hat{\mu}_0, \hat{\sigma}_0\big)$$

$$= P\big(-L_1 < \hat{Z}_{1t} < L_1\big|\hat{\mu}_0, \hat{\sigma}_0\big)$$

$$= P\left(-L_1 < \frac{\bar{Y}_{1t} - \hat{\mu}_0}{\frac{\hat{\sigma}_0}{\sqrt{n_1}}} < L_1\middle|\hat{\mu}_0, \hat{\sigma}_0\right) \quad \{\text{using Equation (C1)}\}$$

$$= P\left(\hat{\mu}_0 - L_1\frac{\hat{\sigma}_0}{\sqrt{n_1}} < \bar{Y}_{1t} < \hat{\mu}_0 + L_1\frac{\hat{\sigma}_0}{\sqrt{n_1}}\middle|\hat{\mu}_0, \hat{\sigma}_0\right)$$

$$= P\left(\frac{\left(\hat{\mu}_0 - L_1\frac{\hat{\sigma}_0}{\sqrt{n_1}}\right) - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} < \frac{\bar{Y}_{1t} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} < \frac{\left(\hat{\mu}_0 + L_1\frac{\hat{\sigma}_0}{\sqrt{n_1}}\right) - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}}\middle|\hat{\mu}_0, \hat{\sigma}_0\right)$$

$$= \Phi\left(\frac{\hat{\mu}_0 - \mu_0}{\frac{\sigma_0}{\sqrt{n_1}}} + L_1\frac{\hat{\sigma}_0}{\sigma_0} + \delta\sqrt{n_1}\right)$$

$$- \Phi\left(\frac{\hat{\mu}_0 - \mu_0}{\frac{\sigma_0}{\sqrt{n_1}}} - L_1\frac{\hat{\sigma}_0}{\sigma_0} + \delta\sqrt{n_1}\right), \text{i.e.}\,\frac{\bar{Y}_{1t} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} \sim N(\delta, 1)$$

Let $U = \frac{\hat{\mu}_0 - \mu_0}{\frac{\sigma_0}{\sqrt{mn}}}$ and $V = \frac{\hat{\sigma}_0}{\sigma_0}$. Since $\hat{\mu}_0 \sim N(\mu_0, \frac{\sigma_0}{\sqrt{mn}})$, it follows that $\frac{\hat{\mu}_0 - \mu_0}{\frac{\sigma_0}{\sqrt{n_1}}} = U\sqrt{\frac{n_1}{mn}}$; and therefore,

$$\hat{P}_{01} = \Phi\left(U\sqrt{\frac{n_1}{mn}} + L_1V + \delta\sqrt{n_1}\right) - \Phi\left(U\sqrt{\frac{n_1}{mn}} - L_1V + \delta\sqrt{n_1}\right).$$

Note that when the process is IC, $\hat{P}_{01}$ is given by

$$\hat{P}_{01} = \Phi\left(U\sqrt{\frac{n_1}{mn}} + L_1V\right) - \Phi\left(U\sqrt{\frac{n_1}{mn}} - L_1V\right).$$

Lemma C2

The conditional probability distribution function (p.d.f.) of $\hat{Z}_{1t}$ given $\hat{\mu}_0$ and $\hat{\sigma}_0$,

$$f_{\hat{Z}_{1t}}(z|\hat{\mu}_0,\hat{\sigma}_0) = V\phi\left(U\sqrt{\frac{n_1}{mn}} + Vz - \delta\sqrt{n_1}\right).$$

Proof: Let $F_{\bar{Y}_{1t}}(z|\hat{\mu}_0,\hat{\sigma}_0)$ be the conditional cumulative distribution function (c.d.f.) of $\bar{Y}_{1t}$ given $\hat{\mu}_0$ and $\hat{\sigma}_0$. Since $\bar{Y}_{1t} \sim N(\mu_0 + \delta\sigma_0, \frac{\sigma_0}{\sqrt{n_1}})$

$$\begin{aligned}
F_{\hat{Z}_{1t}}(z|\hat{\mu}_0,\hat{\sigma}_0) &= P\left(\hat{Z}_{1t} < z \middle| \hat{\mu}_0,\hat{\sigma}_0\right) \\
&= P\left(\frac{\bar{Y}_{1t} - \hat{\mu}_0}{\frac{\hat{\sigma}_0}{\sqrt{n_1}}} < z \middle| \hat{\mu}_0,\hat{\sigma}_0\right) \\
&= P\left(\bar{Y}_{1t} < \hat{\mu}_0 + z\frac{\hat{\sigma}_0}{\sqrt{n_1}} \middle| \hat{\mu}_0,\hat{\sigma}_0\right) \\
&= P\left(\frac{\bar{Y}_{1t} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} < \frac{(\hat{\mu}_0 - \mu_0) + z\frac{\hat{\sigma}_0}{\sqrt{n_1}} + \delta\sigma_0}{\frac{\sigma_0}{\sqrt{n_1}}} \middle| \hat{\mu}_0,\hat{\sigma}_0\right) \\
&= P\left(\hat{Z}_{1t} < U\sqrt{\frac{n_1}{mn}} + Vz + \delta\sqrt{n_1} \middle| \hat{\mu}_0,\hat{\sigma}_0\right) \\
&= \Phi\left(U\sqrt{\frac{n_1}{mn}} + Vz + \delta\sqrt{n_1}\right)
\end{aligned}$$

Next, we take the first derivative of $F_{\hat{Z}_{1t}}(z|\hat{\mu}_0,\hat{\sigma}_0)$so that

$$\begin{aligned}
f_{\hat{Z}_{1t}}(z|\hat{\mu}_0,\hat{\sigma}_0) &= \frac{\partial}{\partial z}F_{\hat{Z}_{1t}}(z|\hat{\mu}_0,\hat{\sigma}_0) \\
&= \frac{\partial}{\partial z}\left\{\Phi\left(U\sqrt{\frac{n_1}{mn}} + Vz + \delta\sqrt{n_1}\right)\right\} \\
&= V\phi\left(U\sqrt{\frac{n_1}{mn}} + Vz + \delta\sqrt{n_1}\right).
\end{aligned}$$

Lemma C3

From Equation (4.4) in Chapter 4, it is given that

$$\hat{\hat{P}}_{02} = P\left[\hat{Z}_{1t} \in \mathrm{B}^{+} \text{ and } \hat{Z}_t \in \mathrm{F}^{-}\right] + P\left[\hat{Z}_{1t} \in \mathrm{B}^{-} \text{ and } \hat{Z}_t \in \mathrm{G}^{+}\right]$$

$$= \int_{Z \in \mathrm{B}^{++}} \hat{P}_{\mathrm{F}^-} f_{\hat{Z}_{1t}}(z|\hat{\mu}_0, \hat{\sigma}_0)dz + \int_{Z \in \mathrm{B}^{--}} \hat{P}_{\mathrm{G}^+} f_{\hat{Z}_{1t}}(z|\hat{\mu}_0, \hat{\sigma}_0)dz.$$

with

$$\hat{P}_{\mathrm{F}^-} = \Phi\left[U\sqrt{\frac{n_2}{mn}} + V\left(\frac{L_2\sqrt{n_1+n_2} - z\sqrt{n_1}}{\sqrt{n_2}}\right) - \delta\sqrt{n_2}\right]$$

and

$$\hat{P}_{\mathrm{G}^+} = 1 - \Phi\left[U\sqrt{\frac{n_2}{mn}} - V\left(\frac{L_2\sqrt{n_1+n_2} + z\sqrt{n_1}}{\sqrt{n_2}}\right) - \delta\sqrt{n_2}\right].$$

Proof:

Note that,

$$\hat{P}_{\mathrm{F}^-} = P\left(\hat{Z}_t < L_2 \middle| \hat{Z}_{1t} = z, \hat{\mu}_0, \hat{\sigma}_0\right).$$

Firstly though, since,

$$\hat{Z}_t = \frac{\bar{Y}_t - \hat{\mu}_0}{\left(\frac{\hat{\sigma}_0}{\sqrt{n_1+n_2}}\right)}$$

$$= \frac{\left\{\frac{n_1\bar{Y}_{1t} + n_2\bar{Y}_{2t}}{n_1+n_2}\right\} - \hat{\mu}_0}{\left(\frac{\hat{\sigma}_0}{\sqrt{n_1+n_2}}\right)}, \quad \text{see Equation (C1)}$$

$$= \frac{\left\{\frac{n_1(\bar{Y}_{1t} - \hat{\mu}_0) + n_2(\bar{Y}_{2t} - \hat{\mu}_0)}{n_1+n_2} + \hat{\mu}_0\right\} - \hat{\mu}_0}{\left(\frac{\hat{\sigma}_0}{\sqrt{n_1+n_2}}\right)}$$

$$= \frac{\left\{\frac{n_1\left(\frac{\hat{\sigma}_0}{\sqrt{n_1}}\hat{Z}_{1t}\right) + n_2\left(\frac{\hat{\sigma}_0}{\sqrt{n_2}}\hat{Z}_{2t}\right)}{n_1+n_2} + \hat{\mu}_0\right\} - \hat{\mu}_0}{\left(\frac{\hat{\sigma}_0}{\sqrt{n_1+n_2}}\right)}, \qquad \text{see Equations (C2)}$$

$$= \frac{\frac{n_1\left(\frac{\hat{\sigma}_0}{\sqrt{n_1}}\hat{Z}_{1t}\right) + n_2\left(\frac{\hat{\sigma}_0}{\sqrt{n_2}}\hat{Z}_{2t}\right)}{n_1+n_2}}{\left(\frac{\hat{\sigma}_0}{\sqrt{n_1+n_2}}\right)}$$

$$= \frac{\hat{\sigma}_0\left(\sqrt{n_1}\hat{Z}_{1t} + \sqrt{n_2}\hat{Z}_{2t}\right)}{n_1 + n_2} \times \left(\frac{\sqrt{n_1 + n_2}}{\hat{\sigma}_0}\right)$$

$$= \frac{\sqrt{n_1}\hat{Z}_{1t} + \sqrt{n_2}\hat{Z}_{2t}}{\sqrt{n_1 + n_2}}$$

then,

$$\hat{P}_{\mathrm{F}^-} = P\left(\frac{\sqrt{n_1}\hat{Z}_{1t} + \sqrt{n_2}\hat{Z}_{2t}}{\sqrt{n_1 + n_2}} < L_2 \middle| \hat{Z}_{1t} = z, \hat{\mu}_0, \hat{\sigma}_0\right)$$

$$= P\left(\sqrt{n_2}\hat{Z}_{2t} < L_2\sqrt{n_1 + n_2} - \sqrt{n_1}\hat{Z}_{1t} \middle| \hat{Z}_{1t} = z, \hat{\mu}_0, \hat{\sigma}_0\right)$$

$$= P\left(\hat{Z}_{2t} < L_2\sqrt{\frac{n_1 + n_2}{n_2}} - \sqrt{\frac{n_1}{n_2}}\hat{Z}_{1t} \middle| \hat{Z}_{1t} = z, \hat{\mu}_0, \hat{\sigma}_0\right)$$

$$= P\left(\hat{Z}_{2t} < L_2\sqrt{\frac{n_1 + n_2}{n_2}} - \sqrt{\frac{n_1}{n_2}}z\right).$$

Moreover, since

$$\hat{Z}_{2t} = \frac{\bar{Y}_{2t} - \hat{\mu}_0}{\frac{\hat{\sigma}_0}{\sqrt{n_2}}} \quad \text{with} \quad \bar{Y}_{2t} \sim N(\mu_0 - \delta\sigma_0, \frac{\sigma_0}{\sqrt{n_2}})$$

then,

$$\hat{P}_{\mathrm{F}^-} = P\left(\hat{Z}_{2t} < L_2\sqrt{\frac{n_1 + n_2}{n_2}} - \sqrt{\frac{n_1}{n_2}}z\right)$$

$$= P\left(\frac{\bar{Y}_{2t} - \hat{\mu}_0}{\frac{\hat{\sigma}_0}{\sqrt{n_2}}} < L_2\sqrt{\frac{n_1 + n_2}{n_2}} - \sqrt{\frac{n_1}{n_2}}z\right)$$

$$= P\left(\bar{Y}_{2t} < \hat{\mu}_0 + L_2\sqrt{\frac{n_1 + n_2}{n_2}}\frac{\hat{\sigma}_0}{\sqrt{n_2}} - z\sqrt{\frac{n_1}{n_2}}\frac{\hat{\sigma}_0}{\sqrt{n_2}}\right)$$

$$= P\left(\frac{\bar{Y}_{2t} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_2}}} < \frac{\hat{\mu}_0 + L_2\sqrt{\frac{n_1 + n_2}{n_2}}\frac{\hat{\sigma}_0}{\sqrt{n_2}} - z\sqrt{\frac{n_1}{n_2}}\frac{\hat{\sigma}_0}{\sqrt{n_2}} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_2}}}\right)$$

$$= \Phi\left(\frac{\hat{\mu}_0-\mu_0}{\frac{\sigma_0}{\sqrt{n_2}}}+\frac{L_2\sqrt{\frac{n_1+n_2}{n_2}}\frac{\hat{\sigma}_0}{\sqrt{n_2}}}{\frac{\sigma_0}{\sqrt{n_2}}}-\frac{z\sqrt{\frac{n_1}{n_2}}\frac{\hat{\sigma}_0}{\sqrt{n_2}}}{\frac{\sigma_0}{\sqrt{n_2}}}+\frac{\delta\sigma_0}{\frac{\sigma_0}{\sqrt{n_2}}}\right),$$

$$\text{i.e.}\frac{\bar{Y}_{2t}-(\mu_0-\delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_2}}}\sim N(\delta,1)$$

$$= \Phi\left(U\sqrt{\frac{n_2}{mn}}+VL_2\sqrt{\frac{n_1+n_2}{n_2}}-Vz\sqrt{\frac{n_1}{n_2}}+\delta\sqrt{n_2}\right)$$

$$= \Phi\left(U\sqrt{\frac{n_2}{mn}}+V\left(\frac{L_2\sqrt{n_1+n_2}-z\sqrt{n_1}}{\sqrt{n_2}}\right)-\delta\sqrt{n_2}\right).$$

Since $P\left[\hat{Z}_{1t}\in \mathrm{B}^+ \text{ and } \hat{Z}_t\in \mathrm{F}^-\right]$ is the probability that a sample plots in region F^- in stage 2 given that in stage 1, a sample plotted in region B^+ (i.e. this region is denoted by $\mathrm{B}^{++}=(L_1+\delta\sqrt{n_1}, L+\delta\sqrt{n_1}]$), and the standardized values are used, then (by also invoking Lemma C2)

$$P\left[\hat{Z}_{1t}\in \mathrm{B}^+ \text{ and } \hat{Z}_t\in \mathrm{F}^-\right]=\int_{Z_{1t}\in \mathrm{B}^{++}} P\left[\hat{Z}_t\in \mathrm{F}^-|\hat{Z}_{1t}=z\right] f_{\hat{Z}_{1t}}(z|\hat{\mu}_0,\hat{\sigma}_0)dz$$

$$=\int_{Z_{1t}\in \mathrm{B}^{++}} \Phi\left(U\sqrt{\frac{n_2}{mn}}+V\left(\frac{L_2\sqrt{n_1+n_2}-z\sqrt{n_1}}{\sqrt{n_2}}\right)-\delta\sqrt{n_2}\right)\times V\times\phi\left(U\sqrt{\frac{n_1}{mn}}+Vz+\delta\sqrt{n_1}\right)dz \tag{C5}$$

Next, for $\hat{P}_{\mathrm{G}^+}$, similar derivations as above are done as follows,

$$\hat{P}_{\mathrm{G}^+}=P\left(\hat{Z}_t>-L_2\middle|\hat{Z}_{1t}=z,\hat{\mu}_0,\hat{\sigma}_0\right)$$

$$=P\left(\frac{\sqrt{n_1}\hat{Z}_{1t}+\sqrt{n_2}\hat{Z}_{2t}}{\sqrt{n_1+n_2}}>-L_2\middle|\hat{Z}_{1t}=z,\hat{\mu}_0,\hat{\sigma}_0\right)$$

$$=P\left(\hat{Z}_{2t}>-L_2\sqrt{\frac{n_1+n_2}{n_2}}-\sqrt{\frac{n_1}{n_2}}z\right)$$

$$=1-P\left(\frac{\bar{Y}_{2t}-\hat{\mu}_0}{\frac{\hat{\sigma}_0}{\sqrt{n_2}}}<-L_2\sqrt{\frac{n_1+n_2}{n_2}}-\sqrt{\frac{n_1}{n_2}}z\right)$$

$$=1-P\left(\frac{\bar{Y}_{2t}-(\mu_0-\delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_2}}}<\frac{\hat{\mu}_0-L_2\sqrt{\frac{n_1+n_2}{n_2}}\frac{\hat{\sigma}_0}{\sqrt{n_2}}-z\sqrt{\frac{n_1}{n_2}}\frac{\hat{\sigma}_0}{\sqrt{n_2}}-(\mu_0-\delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_2}}}\right)$$

$$= 1 - \Phi\left(\frac{\hat{\mu}_0 - \mu_0}{\frac{\sigma_0}{\sqrt{n_2}}} - \frac{L_2\sqrt{\frac{n_1+n_2}{n_2}}\frac{\hat{\sigma}_0}{\sqrt{n_2}}}{\frac{\sigma_0}{\sqrt{n_2}}} - \frac{z\sqrt{\frac{n_1}{n_2}}\frac{\hat{\sigma}_0}{\sqrt{n_2}}}{\frac{\sigma_0}{\sqrt{n_2}}} + \frac{\delta\sigma_0}{\frac{\sigma_0}{\sqrt{n_2}}}\right)$$

$$= 1 - \Phi\left(U\sqrt{\frac{n_2}{mn}} - VL_2\sqrt{\frac{n_1+n_2}{n_2}} - Vz\sqrt{\frac{n_1}{n_2}} + \delta\sqrt{n_2}\right)$$

$$= 1 - \Phi\left(U\sqrt{\frac{n_2}{mn}} - V\left(\frac{L_2\sqrt{n_1+n_2} + z\sqrt{n_1}}{\sqrt{n_2}}\right) - \delta\sqrt{n_2}\right).$$

Since $P\left[\hat{Z}_{1t} \in \mathrm{B}^- \text{ and } \hat{Z}_t \in \mathrm{G}^+\right]$ is the probability that a sample plots in region F^- in stage 2 given that in stage 1, a sample plotted in region B^+ (i.e. this region is denoted by $\mathrm{B}^{--} = (-L + \delta\sqrt{n_1}, -L_1 + \delta\sqrt{n_1}]$), and the standardized values are used, then (by also invoking Lemma C2)

$$P\left[\hat{Z}_{1t} \in \mathrm{B}^- \text{ and } \hat{Z}_t \in \mathrm{G}^+\right] = \int_{\hat{Z}_{1t} \in \mathrm{B}^{--}} P\left[\hat{Z}_t \in \mathrm{G}^+ | \hat{Z}_{1t} = z\right] f_{\hat{Z}_{1t}}(z|\hat{\mu}_0, \hat{\sigma}_0)dz$$

$$= \int_{\hat{Z}_{1t} \in \mathrm{B}^{++}} \Phi\left(U\sqrt{\frac{n_2}{mn}} - V\left(\frac{L_2\sqrt{n_1+n_2} + z\sqrt{n_1}}{\sqrt{n_2}}\right) - \delta\sqrt{n_2}\right) \times V \times \phi\left(U\sqrt{\frac{n_1}{mn}} + Vz + \delta\sqrt{n_1}\right)dz \quad \text{(C6)}$$

Therefore, using Equations (C5) and (C6), then it follows that Lemma 3 is proved.

Lemma C4

From Equation (4.15) in Chapter 4, it is given that

$$\begin{aligned}\hat{P}_2 &= P\left(\hat{Z}_{1t} \in \mathrm{B}^- \cup \mathrm{B}^+ \middle| \hat{\mu}_0, \hat{\sigma}_0\right) \\ &= \Phi\left(U\sqrt{\frac{n_1}{mn}} + VL - \delta\sqrt{n_1}\right) - \Phi\left(U\sqrt{\frac{n_1}{mn}} + VL_1 - \delta\sqrt{n_1}\right) \\ &\quad + \Phi\left(U\sqrt{\frac{n_1}{mn}} - VL_1 - \delta\sqrt{n_1}\right) - \Phi\left(\mathrm{U}\sqrt{\frac{n_1}{mn}} - VL - \delta\sqrt{n_1}\right).\end{aligned}$$

i.e., the probability of taking the second sample.

Proof:

$$\begin{aligned}\hat{P}_2 &= P\left(\hat{Z}_{1t} \in \mathrm{B}^+ \cup \mathrm{B}^-\right) \\ &= P\left(\hat{Z}_{1t} \in \mathrm{B}^+\right) + P\left(\hat{Z}_{1t} \in \mathrm{B}^-\right) \\ &= P\left(L_1 < \hat{Z}_{1t} < L\right) + P\left(-L < \hat{Z}_{1t} < -L_1\right) \\ &= P\left(L_1 < \frac{\bar{Y}_{1t} - \hat{\mu}_0}{\frac{\hat{\sigma}_0}{\sqrt{n_1}}} < L\right) + P\left(-L < \frac{\bar{Y}_{1t} - \hat{\mu}_0}{\frac{\hat{\sigma}_0}{\sqrt{n_1}}} < -L_1\right) \\ &= P\left(\frac{\hat{\mu}_0 + L_1\frac{\hat{\sigma}_0}{\sqrt{n_1}} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} < \frac{\bar{Y}_{1t} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} < \frac{\hat{\mu}_0 + L\frac{\hat{\sigma}_0}{\sqrt{n_1}} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}}\right) \\ &\quad + P\left(\frac{\hat{\mu}_0 - L\frac{\hat{\sigma}_0}{\sqrt{n_1}} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} < \frac{\bar{Y}_{1t} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}} < \frac{\hat{\mu}_0 - L_1\frac{\hat{\sigma}_0}{\sqrt{n_1}} - (\mu_0 - \delta\sigma_0)}{\frac{\sigma_0}{\sqrt{n_1}}}\right) \\ &= \left\{\Phi\left(U\sqrt{\frac{n_1}{mn}} + VL - \delta\sqrt{n_1}\right) - \Phi\left(U\sqrt{\frac{n_1}{mn}} + VL_1 - \delta\sqrt{n_1}\right)\right\} \\ &\quad + \left\{\Phi\left(U\sqrt{\frac{n_1}{mn}} - VL_1 - \delta\sqrt{n_1}\right) - \Phi\left(U\sqrt{\frac{n_1}{mn}} - VL - \delta\sqrt{n_1}\right)\right\}.\end{aligned}$$

Bibliography

Aghaulor, C.D. and Ezekwem, C. (2016). An economic design of a modified synthetic double sampling control chart for process monitoring. *International Journal of Engineering Research & Technology,* 5(11), 445-460.

Bersimis, S., Psarakis, S. and Panaretos, J. (2007). Multivariate statistical process control charts: An overview. *Quality and Reliability Engineering International*, 23(5), 517-543.

Carot, V., Jabaloyes, J.M. and Carot, T. (2002). Combined double sampling and variable sampling interval chart, *International Journal of Production Research,* 40(9), 2175-2186.

Castagliola, P., Oprime, P.C. and Khoo, M.B.C. (2017). The double sampling S^2 chart with estimated process variance. *Communications in Statistics – Theory and Methods*, 46(7), 3556-3573.

Chakraborti, S. and Graham, M.A. (2019). *Nonparametric Statistical Process Control*. New York: Wiley, ISBN: 978-1-118-45603-3.

Chakraborti, S. (2006). Parameter estimation and design considerations in prospective applications of the $\bar{X}$ chart. *Journal of Applied Statistics,* 33(4), 439-459.

Champ, C.W. and Aparisi, F. (2008). Double sampling Hotelling's T^2 charts. *Quality and Reliability Engineering International*, 24(2), 153-166.

Chen, G. and Cheng, S.W. (1998). Max chart: Combining $\bar{X}$ chart and s chart. *Statistica Sinica*, 8(1), 263-271.

Chong, N.L., Khoo, M.B.C., Chong, Z.L. and Teoh, W.L. (2018). A study on the run length properties of the side sensitive group runs double sampling (SSGRDS) control chart. MATEC Web of Conferences, DOI: 10.1051/matecconf/201819201005.

Chong, Z.L., Khoo, M.B.C. and Castagliola, P. (2014). Synthetic double sampling np control chart for attributes. *Computers & Industrial Engineering*, 75(1), 157-169.

Chong, Z.L., Khoo, M.B.C., Teoh, W.L., Yeong, W.C. and Teh, S.Y. (2017). Group runs double sampling *np* control chart for attributes. *Journal of Testing and Evaluation*, 45(6), 2267-2282.

Claro, F.A.E., Costa, A.F.B. and Machado, M.A.G. (2008). Double sampling $\bar{X}$ control chart for a first order autoregressive process. *Pesquisa Operacional*, 28(3), 545-562.

Costa, A.F.B. and Claro, F.A.E. (2008). Double sampling X control chart for a first-order autoregressive moving average process model. *International Journal of Advanced Manufacturing Technology*, 39(5-6), 521-542.

Costa, A.F.B. and Machado M.A.G. (2008). Bivariate control charts with double sampling. *Journal of Applied Statistics*, 35(7), 809-822.

Costa, A.F.B. and Machado, M.A.G. (2011). Variable parameter and double sampling $\bar{X}$ charts in the presence of correlation: the Markov chain approach. *International Journal of Production Economics*, 130(2), 224-229.

Costa, A.F.B. (2017). The double sampling range chart. *Quality and Reliability Engineering International*, 33(8), 2739-2745.

Costa, A.F.B. and Machado, M.A.G. (2015). The steady-state behavior of the synthetic and side-sensitive double sampling $\bar{X}$ charts. *Quality and Reliability Engineering International*, 31(2), 297-303.

Croasdale, R. (1974). Control charts for a double-sampling scheme based on average production run lengths. *International Journal of Production Research*, 12(5), 585-592.

Daudin, J.J. (1992). Double sampling $\bar{X}$ charts. *Journal of Quality Technology*, 24(1), 78-87.

Daudin, J.J., Duby, C. and Trecourt, P. (1990). Plans de Controle Double Optimaux (Maitrise des Procedes et Controle de Reception), *Revue de Statistique Appliquée*, 38(4), 45-59.

De Araújo Rodrigues, A.A., Epprecht, E.K. and De Magalhães, M.S. (2011). Double-sampling control charts for attributes. *Journal of Applied Statistics*, 38(1), 87-112.

Faraz, A., Heuchenne, C. and Saniga, E. (2012). Optimal T^2 control chart with a double sampling scheme – an alternative to the MEWMA chart. *Quality and Reliability Engineering International*, 28(7), 751-760.

Gadre, M.P. and Rattihalli, R.N. (2007). A side-sensitive group runs control chart for detecting shifts in the process mean. *Statistical Methods and Applications*, 16(1), 27-37.

Gan, F.F. (1994). An optimal design of cumulative sum control chart based on median run length. *Communications in Statistics – Simulation and Computation*, 23(2), 485-503.

Graham, M.A., Chakraborti, S. and Mukherjee, A. (2014). Design and implementation of CUSUM exceedance control charts for unknown location. *International Journal of Production Research*, 52(18), 5546-5564.

Grigoryan, A. and He, D. (2005). Multivariate double sampling |**S**| charts for controlling process variability. *International Journal of Production Research*, 43(4), 715-730.

Haq, A. and Khoo, M.B.C. (2018). A new double sampling control chart for monitoring process mean using auxiliary information. *Journal of Statistical Computation and Simulation*; 88(5), 869-899.

Haq, A and Khoo, M.B.C. (2019). A synthetic double sampling control chart for process mean using auxiliary information. *Quality and Reliability Engineering International*, 35(6), 1803-1825.

Hawkins, D.M. and Olwell, D.H. (1998). *Cumulative sum charts and charting for quality improvement*. New York, NY: Springer.

He, D. and Grigoryan, A. (2002). Construction of double sampling s-control charts for agile manufacturing. *Quality and Reliability Engineering International*, 18(4), 343-355.

He, D. and Grigoryan, A. (2003). An improved double sampling s chart. *International Journal of Production Research*, 41(12), 2663-2679.

He, D. and Grigoryan, A. (2006). Joint statistical design of double sampling $\bar{X}$ and S charts. *European Journal of Operational Research*, 168(1), 122-142.

He, D., Grigoryan, A. and Sigh, M. (2002). Design of double-and triple-sampling $\bar{X}$ control charts using genetic algorithms, *International Journal of Production Research*, 40 (6), 1387-1404.

Hsu, L.F. (2004). Note on 'Design of double- and triple- sampling X-bar control charts using genetic algorithms'. *International Journal of Production Research*, 42(5), 1043-1047.

Hsu, L.F. (2007). Note on 'Construction of double sampling s-control charts for agile manufacturing'. *Quality and Reliability Engineering International*, 23(2), 269-272.

Irianto, D. and Juliani, A. (2010). A two control limits double sampling control chart by optimizing producer and customer risks. *ITB Journal of Engineering Science*, 42(2), 165-178.

Irianto, D., and Shinozaki, N. (1998). An optimal double sampling $\bar{X}$ control chart. *International Journal of Industrial Engineering – Theory, Applications and Practice*, 5(3), 226-234.

Jensen, W.A., Jones-Farmer, L.A., Champ, C.W. and Woodall, W.H. (2006). Effects of parameter estimation on control chart properties: A literature review. *Journal of Quality Technology*, 38(4), 349-364.

Jones, L.A., Champ, C.W. and Rigdon, S.E. (2004). The run length distribution of the CUSUM with estimated parameters. *Journal of Quality Technology*, 36(1), 95-108.

Katebi, M. and Moghadam, M.B. (2020). A double sampling multivariate T^2 control chart with variable sample size and variable sampling interval. *Communications in Statistics – Simulation and Computation*, DOI: 10.1080/03610918.2020.1716246.

Khatun, M., Khoo, M.B.C., Yeong, W.L., Teoh, W.L. and Chong, Z.L. (2018). Adaptive multivariate double sampling and variable sampling interval Hotelling's T^2 charts. *Quality and Reliability Engineering International*, 34(5), 894-911.

Khoo, M.B.C. (2004). S^2 control chart based on double sampling. *International Journal of Pure and Applied Mathematics*. 13(2), 249-258.

Khoo, M.B.C., Khoo, J.L., Yeong W.C. and Teoh, W.L. (2016). A comparative study between the variable sampling interval $\bar{X}$ and double sampling $\bar{X}$ charts. IEEE International Conference on Control and Robotics Engineering (ICCRE), April 2016, DOI: 10.1109/ICCRE.2016.7476133.

Khoo, M.B.C., Lee, H.C., Wu, Z., Chen, C.H. and Castagliola, P. (2011b). A synthetic double sampling control chart for the process mean. *IIE Transactions*, 43(1), 23-38.

Khoo, M.B.C., Lee, M.H., Teoh, W.L., Liew, J.Y. and Teh, S.Y. (2013a). The effects of parameter estimation on minimizing the in-control average sample size for the double sampling $\bar{X}$ chart, *South African Journal of Industrial Engineering*, 24(3), 58-67.

Khoo, M.B.C., Tan, E.K., Chong, Z.L. and Haridy, S. (2015). Side-sensitive group runs double sampling (SSGRDS) chart for detecting mean shifts. *International Journal of Production Research*, 53(15), 4735-4753.

Khoo, M.B.C., Teoh, W.L., Castagliola, P. and Lee, M.H. (2013b). Optimal designs of the double sampling chart with estimated parameters, *International Journal of Production Economics*, 144(1), 345-357.

Khoo, M.B.C., Wong, V.H., Wu, Z. and Castagliola, P. (2011a). Optimal designs of the multivariate synthetic chart for monitoring the process mean vector based on median run length. *Quality and Reliability Engineering International*, 27(12), 981–997.

Khoo, M.B.C., Wu Z., Castagliola, P. and Lee, H.C. (2013c). A multivariate synthetic double sampling T^2 control chart. *Computers & Industrial Engineering*, 64(1), 179-189.

Klein, M. (2000). Two alternatives to Shewhart $\bar{X}$ control chart. *Journal of Quality Technology*, 32(4), 427-431.

Lee, M.H. and Khoo, M.B.C. (2019a). The economic and economic statistical designs of synthetic double sampling $\bar{X}$ chart. *Communications in Statistics - Simulation and Computation*, 48(8), 2313-2332.

Lee, M.H. and Khoo, M.B.C. (2019b). Economic-statistical design of synthetic double sampling T^2 chart. *Communications in Statistics - Theory and Methods*, 48(23), 5862-5876.

Lee, M.H. and Khoo, M.B.C. (2019c). Double sampling *np* chart with estimated process parameter. *Communications in Statistics - Simulation and Computation*, DOI: 10.1080/03610918.2019.1599017.

Lee, M.H. and Khoo, M.B.C. (2017a). Synthetic double sampling *s* chart. *Communications in Statistics – Theory and Methods*, 46(12), 5914-5931.

Lee, M.H. and Khoo, M.B.C. (2017b). Double sampling max chart. *Communications in Statistics – Simulation and Computation*, 46(10), 7855-7878.

Lee, M.H. and Khoo, M.B.C. (2017c). Combined double sampling and variable sampling interval *np* chart. *Communications in Statistics – Theory and Methods*, 46(23), 11892-11917.

Lee, M.H. and Khoo, M.B.C. (2018). Double sampling |S| control chart with variable sample size and variable sampling interval. *Communications in Statistics – Simulation and Computation*, 47(2), 615-628.

Lee, P.-H, Torng, C.-C., Liao, H.-S. and Tseng, C.-C. (2009). The statistical performance of double sampling $\bar{X}$ control charts for correlation data. 2009 *IEEE International Conference on Industrial Engineering and Engineering Management*, DOI: 10.1109/ieem.2009.5373536.

Lee, P.-H. (2013). Joint statistical design of $\bar{X}$ and s charts with combined double sampling and variable sampling interval. *European Journal of Operational Research*, 225(2), 285-297.

Lee, P.-H., Chang, Y.-C. and Torng, C.-C. (2012b). A design of s control charts with a combined double sampling and variable sampling interval scheme. *Communications in Statistics – Theory and Methods*, 41(1), 153-165.

Lee, P.-H., Torng, C.-C. and Liao, L.F. (2012a). An economic design of combined double sampling and variable sampling interval $\bar{X}$ control chart, *International Journal of Production Economics*, 138(1), 102-106.

Lee, P.-H., Torng, C.-C., Wu, J.-C. and Tseng C.-C. (2010). The effectiveness study of double sampling s charts application on destructive testing process. *International Journal of Product Development*, 12(3-4), 324-335.

Lee, M.H., Si, K.S.K.Y., Chew, X.Y., Lau, E.M.F. and Then, P.H.H. (2019). The effect of measurement errors on the double sampling $\bar{X}$ chart. *COMPUSOFT, An International Journal of Advanced Computer Technology*, 8(9), 3395-3401.

Lorenzen, T.J., and Vance, L.C. (1986). The economic design of control charts: A unified approach. *Technometrics*, 28(1), 3-10.

Machado, M.A.G. and Costa, A.F.B. (2008). The double sampling and the EWMA charts based on the sample variances. *International Journal of Production Economics*, 144(1), 134-148.

Machado, M.A.G. and Costa, A.F.B. (2014). Some comments regarding the synthetic chart. *Communications in Statistics - Theory and Methods*, 43(14), 2897-2906.

Maleki, M.R., Amiri, A. and Castagliola, P. (2017). Measurement errors in statistical process monitoring: A literature review. *Computers & Industrial Engineering*, 103, 316-329.

Malela-Majika, J.-C. and Rapoo, E.M. (2019). Side-sensitive synthetic double sampling $\bar{X}$ control charts. *European Journal of Industrial Engineering*, 13(1), 117-148.

Malela-Majika, J.-C. (2019). Modified side-sensitive synthetic double sampling monitoring scheme for simultaneously monitoring the process mean and variability. *Computers and Industrial Engineering*, 130, 798-814.

Malela-Majika, J.-C., Motsepa, C.M. and Graham, M.A. (2019). A new double sampling $\bar{X}$ control chart for monitoring an abrupt change in the process location. *Communications in Statistics - Simulation and Computation*; DOI: 10.1080/03610918.2019.1577970.

McCracken, A. and Chakraborti, S. (2013). Control charts for joint monitoring of mean and variance: An overview. *Quality Technology and Quantitative Management*, 10(1), 17-36.

Montgomery, D.C. (2013). *Statistical Quality Control: A Modern Introduction*, 7th ed., John Wiley Sons, Singapore Pte. Ltd.

Motsepa C.M., Malela-Majika J.-C., Castagliola P. and Shongwe S.C. (2020). A side-sensitive double sampling $\bar{X}$ monitoring scheme with estimated process parameters. *Communications in Statistics – Simulation and Computation*, DOI: 10.1080/03610918.2020.1722835.

Ng, P.S., Khoo, M.B.C., Saha S. and Yeong, W.C. (2018). Double sampling control charts for monitoring the coefficient of variation. *2nd International Conference on Smart Sensors and Application* (ICSSA), Malaysia, DOI: 10.1109/ICSSA.2018.8535900.

Noorossana, R., Shekary, A.M., and Deheshvar, A. (2015). Combined variable sample size, sampling interval and double sampling (CVSSIDS) adaptive control charts. *Communications in Statistics – Theory and Methods*, 44(6), 1255-1269.

Page, E.S., (1954). Continuous inspection schemes. *Biometrika,* 41(1-2), 100-115.

Psarakis, S. (2015). Adaptive control charts: Recent developments and extensions. *Quality and Reliability Engineering International*, 31(7), 1265-1280.

Psarakis, S., Vyniou, A. and Castagliola, P. (2013). Some recent developments on the effects of parameter estimation on control charts. *Quality and Reliability Engineering International*, 30(8), 1113-1129.

Qiu, P. (2019). Some recent studies in statistical process control. In *Statistical Quality Technologies: Theory and Practice*, Eds.: Lio Y., Ng H.K.T., Tsai T.-R., Chen D.-G. Springer: ICSA Book Series in Statistics.

Qiu, P. (2014). *Introduction to Statistical Process Control*. Chapman & Hall/CRC Press, Taylor & Francis Group, Florida: Baton Rouge.

Rakitzis, A.C., Chakraborti, S., Shongwe, S.C., Graham, M.A. and Khoo, M.B.C. (2019). An overview of synthetic-type control charts: Techniques and Methodology. *Quality and Reliability Engineering International*, 35(7), 2081-2096.

Rao, P.S.R.S. (2005). Double sampling. *Encyclopedia of Biostatistics*, John Wiley & Sons. (pp. 1-7). DOI: 10.1002/0470011815.b2a16020.

Reynolds Jr., M.R., and Lou, J. (2010). An evaluation of a GLR control chart for monitoring the process mean. *Journal of Quality Technology*, 42(3), 287-310.

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

Rozi, F., Pasaribu, U.S., Mukhaiyar, U. and Irianto, D. (2019). Modified double sampling control chart for monitoring the coefficient of variation. *Conference on Industrial and System Engineering*, In *IOP Conference Series: Materials Science and Engineering*, DOI: 10.1088/1757-899X/598/1/012110.

Ryu, J.H., Wan, H. and Kim, S. (2010). Optimal design of a CUSUM chart for a mean shift of unknown size. *Journal of Quality Technology*, 42(3), 311-326.

Saha, S., Khoo, M.B.C., Lee, M.H. and Castagliola, P. (2018). A side-sensitive modified group runs double sampling (SSMGRDS) control chart for detecting mean shifts. *Communications in Statistics - Simulation and Computation*, 47(5), 1353-1369.

Shongwe, S.C. and Graham, M.A. (2016). On the performance of Shewhart-type synthetic and runs-rules charts combined with an $\overline{\mathrm{X}}$ chart. *Quality and Reliability Engineering International*, 32(4), 1357-1379.

Shongwe, S.C. and Graham, M.A. (2019a). Some theoretical comments regarding the run-length properties of the synthetic and runs-rules monitoring schemes – Part 1: Zero-state. *Quality Technology & Quantitative Management*, 16(2), 170-189.

Shongwe, S.C. and Graham, M.A. (2019b). Some theoretical comments regarding the run-length properties of the synthetic and runs-rules monitoring schemes – Part 2: Steady-state. *Quality Technology & Quantitative Management* 16(2), 190-199.

Shongwe, S.C., Malela-Majika, J.-C. and Rapoo, E.M. (2019). One-sided and two-sided *w-of-w* runs-rules schemes: An overall performance perspective and the unified run-length derivations. *Journal of Probability and Statistics,* Article ID: 6187060, 20pages.

Teoh, W.L., Fun, M.S., Teh, S.Y., Khoo, M.B.C. and Yeong, W.C. (2016a). Exact run length distribution of the double sampling $\bar{X}$ chart with estimated process parameters, *South African Journal of Industrial Engineering*, 27(1), 20-31.

Teoh, W.L., Khoo, M.B.C., Castagliola, P. and Chakraborti, S. (2015). A median run length-based double-sampling $\bar{X}$ chart with estimated parameters for minimizing the average sample size. *International Journal of Advanced Manufacturing Technology*, 80(1-4), 411-426.

Teoh, W.L., Khoo, M.B.C., Castagliola, P. and Chakraborti, S. (2014b). Optimal design of the double sampling $\bar{X}$ chart with estimated parameters based on median run length. *Computers & Industrial Engineering*, 67, 104-115.

Teoh, W.L., Khoo, M.B.C. and Teh, S.Y. (2013). Optimal designs of the median run length based double sampling $\bar{X}$ chart for minimizing the average sample size. *PLoS ONE*, 8(7): e68580.

Teoh, W.L., Liow, K.L., Khoo, M.B.C., Yeong, W.C. and Teh, S.Y. (2014a). A comparison between the performance of the double sampling $\bar{X}$ and variable sample size $\bar{X}$ charts. *Journal of Quality Measurement and Analysis*, 10(2), 15-31.

Teoh, W.L., Yeong, W.C., Khoo, M.B.C. and Teh, S.Y. (2016b). The performance of the double sampling $\bar{X}$ chart with estimated parameters for skewed distributions. *Academic Journal of Science*, 5(1), 237-252.

Torng, C.-C. and Lee, P.-H. (2009). The performance of double sampling $\bar{X}$ control charts under non-normality. *Communications in Statistics – Simulation and Computation*, 38(3), 541-557.

Torng, C.-C., Lee, P.-H, Liao, H.-S. and Liao, N.-Y. (2009a). An economic design of double sampling $\bar{X}$ charts for correlated data using genetic algorithms. *Expert Systems with Applications*, 36(10), 12621-12626.

Torng, C.-C., Lee, P.-H. and Liao, N.-Y. (2009b). An economic-statistical design of double sampling $\bar{X}$ control chart. *International Journal of Production Economics*, 120(2), 495-500.

Torng, C.-C., Tseng, C.-C. and Lee, P.-H. (2010). Non-normality and combined double sampling and variable sampling interval control charts. *Journal of Applied Statistics*, 37(6), 955-967.

Umar, A.A., Khoo, M.B.C., Saha, S. and Haq, A. (2019). A combined variable sampling interval and double sampling control chart with auxiliary information for the process mean. *Transactions of the Institute of Measurement and Control*, DOI: 10.1177/0142331219885525.

Woodall, W.H. (1986). Weaknesses of the economical design of control charts. *Technometrics*, 28(4), 408-409.

Woodall, W.H. (1997). Control charts based on attribute data: Bibliography and review. *Journal of Quality Technology*, 29(2), 172-183.

Wu, Z., Yang, W., Jiang, W. and Khoo, M.B.C. (2008). Optimization designs of the combined Shewhart-CUSUM control charts. *Computational Statistics & Data Analysis*, 53(2), 496-506.

Yang, K. and Hancock, W.M. (1990). Statistical quality control for correlated samples. *International Journal of Production Research*, 28(3), 595-608.

Yang, S.-F. and Wu, S.-H. (2017a). A double sampling scheme for process mean monitoring. *IEEE Access*, 5, 6668-6677.

Yang, S.-F. and Wu, S.-H. (2017b). A double sampling scheme for process variability monitoring. *Quality and Reliability Engineering International*, 33(8), 2193-2204.

You, H.W. (2017). Run length distribution of synthetic double sampling chart. *International Journal of Applied Engineering Research*, 12(24), 14268-14272.

You, H.W. (2018). Performance of synthetic double sampling chart with estimated parameters based on expected average run length. *Journal of Probability and Statistics*, Article ID: 7583610, 6pages.

You, H.W., Khoo, M.B.C., Lee, M.H. and Castagliola, P. (2015). Synthetic double sampling $\bar{X}$ chart with estimated process parameters. *Quality Technology and Quantitative Management*, 12(4), 579-604.

Zaman, B., Lee, M.H., Riaz, M. and Abujiya, M.R. (2017). An adaptive EWMA scheme-based CUSUM accumulation error for efficient monitoring of process location. *Quality and Reliability Engineering International*, 33(8), 2463-2482.

Zhang, Y., Castagliola, P., Wu, Z. and Khoo, M.B.C. (2011). The synthetic $\bar{X}$ chart with estimated parameters. *IIE Transactions*, 43(9), 676-687.