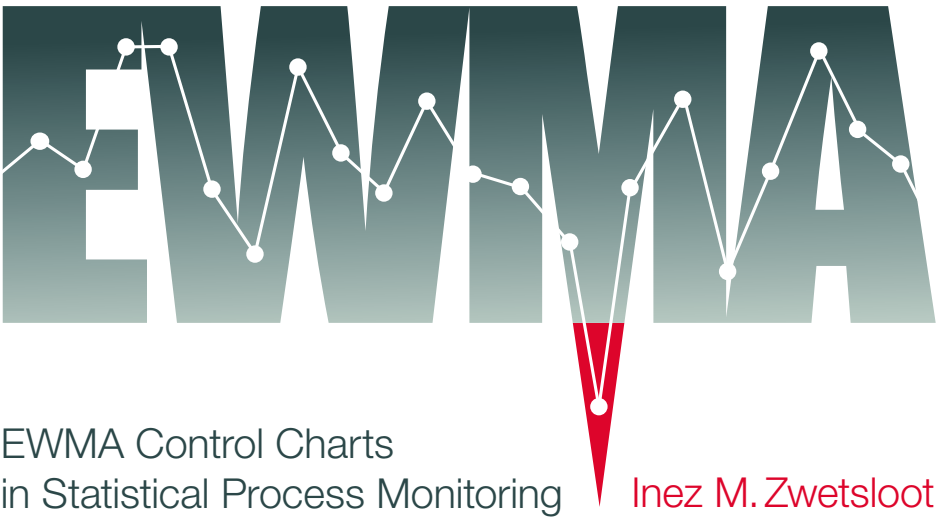
EWMA control charts in statistical process monitoring

Zwetsloot, I.M.

Publication date
2016
Document Version
Final published version



EWMA Control Charts in Statistical Process Monitoring

Inez M. Zwetsloot

EWMA Control Charts in Statistical Process Monitoring

Inez M. Zwetsloot

Publisher IBIS UvA, Amsterdam
Printed by Gildeprint - Enschede
ISBN 978-94-6233-252-2

Cover design by André Verhoek (www.averhoek.nl). Typesetting by the author. The text is set in TeX Gyre Pagella. Headings and images are set in Helvetica.

EWMA Control Charts in Statistical Process Monitoring

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het College voor Promoties
ingestelde commissie, in het openbaar te verdedigen
in de Aula der Universiteit
op vrijdag 22 april 2016, te 13:00 uur

door

Inez Maria Zwetsloot

geboren te Leidschendam

Promotiecommissie

Promotor

Prof. dr. R. J. M. M. Does	Universiteit van Amsterdam

Copromotor

Dr. M. Schoonhoven	Universiteit van Amsterdam

Overige leden

Prof. dr. H. P. Boswijk	Universiteit van Amsterdam
Prof. dr. M. R. H. Mandjes	Universiteit van Amsterdam
Prof. dr. J. de Mast	Universiteit van Amsterdam
Prof. dr. K. C. B. Roes	Universiteit Utrecht
Prof. dr. M. Salomon	Universiteit van Amsterdam
Prof. dr. J. E. Wieringa	Rijksuniversiteit Groningen
Prof. dr. W. H. Woodall	Virginia Tech

Faculteit Economie en Bedrijfskunde

To my grandparents Rie and Piet Zwetsloot-Burmanje

A good decision is based on knowledge and not on numbers
Plato

Table of Contents

List of Figures

EWMA Control Charts in Statistical Process Monitoring

List of Tables

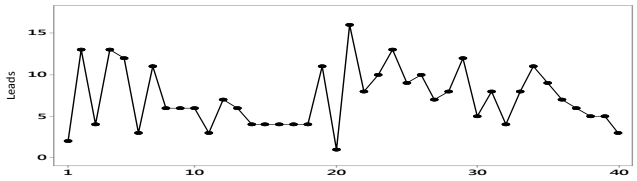
1. Introduction

In today's world, the amount of available data is steadily increasing, and it is often of interest to detect changes in the data. Statistical process monitoring (SPM) provides tools to monitor data streams and to signal changes in the data. One of these tools is the control chart. The topic of this dissertation is a special control chart: the exponentially weighted moving average (EWMA) control chart. The fact that these charts also play an important role in lean six sigma provides an additional motivation.

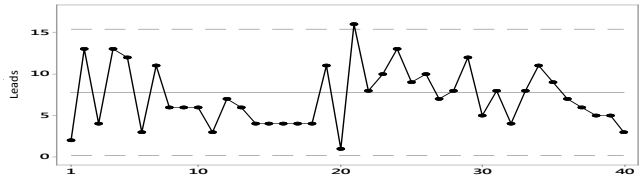
Imagine, you work at an educational institute, which offers courses and you are responsible for the recruitment process of participants in these courses. One of the important channels through which potential participants find your institute is its website. These potential participants have the option to request a brochure through an online form. From the website's data analytics system you can extract the number of requested brochures: the so-called number of 'leads'. Figure 1.1(a) shows a run chart of the total number of leads per week in the first forty weeks of 2015. What can you learn from the figure? Are you -or should you be- worried about the low number of leads in the last week?

Answers to these kinds of questions can be found using the tools and techniques studied in SPM. SPM started with the pioneering work of Walter A. Shewhart at the Bell Telephone and Western Electric companies. He introduced a chart that could be used to compare the current data with data generated by a normally operating process (Shewhart, 1926, 1931). Nowadays, we know this chart as the *Shewhart control chart*. An example is displayed in Figure 1.1(b). The added 'control limits' enable the user to distinguish 'normal' variability in the process from 'special' cause variability. Figure 1.1(b) shows one signal in week 21, indicating that in week 21 possibly something was different compared to the other weeks.

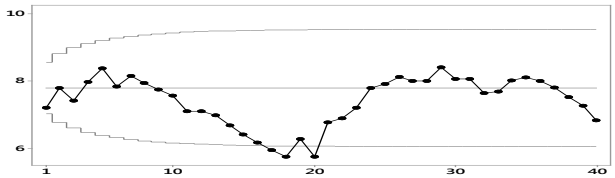
A limitation of the Shewhart control chart is its inability to detect small to moderate sized sustained changes in the process parameters. To overcome this problem, Roberts (1959) introduced the EWMA *control chart*; it is displayed in Figure 1.1(c).



(a) *Run chart*



(b) *Shewhart control chart*



(c) *EWMA control chart*

Figure 1.1: *Example charts*

Where the Shewhart control chart plots the current data from the process, the EWMA control chart plots a weighted average of the current and past data from the process. This chart therefore, has the ability to detect small sustained changes quicker. The EWMA chart in Figure 1.1(c) already signals at week 17 that a change may have occurred.

The goal of this dissertation is to investigate the properties of the EWMA control chart and to give recommendations regarding its design. More specifically, we focus on evaluating and understanding the effect of estimation error. This first chapter provides an overview of this dissertation's contributions and discusses its contributions within the literature. To this end, we first address the concepts in the dissertation's title; statistical process monitoring is discussed in Section 1.1, the control chart is discussed in Section 1.2, and the EWMA control chart is discussed in Section 1.3. Finally, in Sections 1.4 and 1.5, two problems that show up in the application of (EWMA) control charts are discussed. These lead to the motivation and outline of this dissertation in Section 1.6.

1.1 Statistical process monitoring

The term statistical process *monitoring* (SPM) might surprise the reader as statistical process *control* (SPC) is much more widely used. Before we clarify the deliberate choice for SPM, a short introduction to the field of industrial statistics and SPC is necessary.

The field of industrial statistics combines knowledge of statistics with process thinking (Vining et al., 2015). An industrial statistician uses statistical tools to generate useful information about how processes can be improved to provide better value for the customers (Bisgaard, 2012). A well-known methodology to improve processes is lean six sigma. For an introduction to this methodology the reader is referred to De Mast et al. (2012). The example of website leads, discussed above, comes from such an improvement project (see Zwetsloot and Does, 2015).

Statistical process control (SPC) is an important area of research in industrial statistics. SPC encompasses a set of problem-solving tools useful for improving and monitoring the performance of a process (Montgomery, 2013). The tools of statistical process control are widely used in the analyze and control phase of DMAIC, the framework prescribed in lean six sigma. Box et al. (2005, page 565) distinguish between two procedures in SPC: process *monitoring* and process *adjustment or control*. "By process monitoring is meant continual checking of process operation to detect unusual data....By contrast, process adjustment is used to estimate and compensate for drifts or other changes in the data." The methodologies studied in this dissertation are all related to process monitoring, hence the choice for SPM in the title of this dissertation. Hereby, we do not imply that adjustment and control are not important. However, it is beyond the scope of the work presented here.

The foundations for SPC (and hence SPM) were laid down by Shewhart (1931). For an overview of current issues and ideas in SPM see Woodall and Montgomery (2014). Shewhart introduced two types of variation. The first is *common cause variability*. This is variability due to random noise without a (clear) reason. The second type is what Shewhart called *assignable cause variability*. Control charts are designed to detect changes due to assignable causes quickly. Recall the Shewhart control chart in Figure 1.1(b). It shows one signal of a (possible) special cause of variability at week 21. An investigation of this week reveals the assignable cause for this peak in the number of leads; there was a recruitment event held in that week. The EWMA chart in Figure 1.1(c) shows signals in weeks 17, 18, and 20. An investigation of these weeks reveals the assignable cause for this downward shift in the number of leads; in week 15 the website's headings were changed resulting in a drop in the rank on online search engines. In week 19 the changes in the headings were reversed.

A process influenced only by common cause variability is called 'in-control'. To represent an in-control process, it is often modelled by a probability distribution representing the common cause variation in the process. Special cause variability is then

modelled as a change from the 'in-control' distributional parameter(s). Such a process is called 'out-of-control'. In this dissertation we use the following definition, inspired by Does et al. (1999, Chapter 1);

In-control process *A process that is stable and is influenced solemnly by common cause variability. These are causes that are inherent to the process hour by hour, day by day, and that influence everything and everybody working in the process*;

Out-of-control process *A process that is influenced by assignable or special causes. These are cause that are not continuously present in the process or which do not influence everything, but that arise from specific circumstances.*

1.2 Control charts

Control charts are designed to detect changes in a process from an in-control state to an out-of-control state. A control chart consists of plotting the information on a process characteristic against time together with so-called control limits. As soon as this plotted statistic exceeds a control limit, a signal is given. A signal indicates a possible out-of-control process.

To monitor a process, observations from the process are prospectively collected. We consider the situation were either a single observation ($n = 1$) or multiple ($n > 1$) observations are collected at each time instance. Furthermore, we consider monitoring a continuous process characteristic that can be modelled as an independent and normally distributed variable. As the normal distribution is fully determined by its mean and variance, we consider both control charts for the location (Chapters 2 and 4) as well as for the dispersion (Chapters 3 and 5).

As an alternative to the traditional Shewhart chart, Roberts (1959) introduced the EWMA chart and Page (1954) introduced the cumulative sum (CUSUM) chart. Both charts can detect small to moderate sized shifts quicker than the Shewhart chart. Numerous studies have compared the performance of the EWMA and CUSUM chart, see for example Hawkins and Wu (2014) and Zwetsloot and Woodall (2015). In this dissertation, we focus on monitoring using the EWMA chart.

There are numerous applications of control charts. Traditionally, they have been used in quality control and improvement. Examples can be found in Does et al. (1999) and Lawless et al. (2012). Many applications can also be found in other fields such as healthcare (see e.g. Woodall, 2006; Tsui et al., 2008; Spiegelhalter et al., 2012) or services (see e.g. MacCarthy and Wasusri, 2002; De Mast et al., 2012). Recently, control charts have been applied to enhance data quality (Jones-Farmer et al., 2014a) and network monitoring (Woodall et al., 2015). In other fields statistical process monitoring is often known as 'anomaly detection' or 'sequential surveillance'. These methods are closely related to the control charting techniques. For example Thottan and Ji (2003)

and Münz and Carle (2008) monitor internet traffic. And Frisén (2009) discusses applications in finance.

1.3 EWMA control charts

The exponentially weighted moving average (EWMA) control chart, as introduced by Roberts (1959), consists of plotting a weighted average of measurements, giving heaviest weights to the most recent observations. This provides the chart with the advantage of being sensitive to small- and moderate-sized sustained shifts in the process parameters. Figure 1.1(c) shows an example of an EWMA control chart. The EWMA chart has received a great deal of attention in the SPM literature. See, for example, Crowder (1987, 1989), Robinson and Ho (1978), Lucas and Saccucci (1990), Jones et al. (2001), Jones (2002), and Simões et al. (2010).

Mathematically the EWMA chart consists of plotting the EWMA statistic Z_i at time i defined as $Z_i = (1-\lambda)Z_{i-1} + \lambda M_i$, for $i = 1, 2, 3, ...$, where M_i denotes the measure of interest, based on the current information of the process characteristic. For monitoring the location (see Chapters 2 and 4), we set M_i equal to the sample mean. For monitoring the dispersion (see Chapters 3 and 5), we set M_i equal to a dispersion measure such as the sample standard deviation.

Throughout this dissertation we assume that 'the current information' consists of observations which are collected in samples of size $n \geq 1$ from the process. Furthermore, we assume that these observations are independent and normally distributed with parameters μ and σ. If the process is in control we use that $\mu = \mu_0$ and $\sigma = \sigma_0$.

The factor λ, $0 < \lambda \leq 1$, is the weighting factor and is referred to as the smoothing constant. The smaller the value of λ, the heavier the reliance on past data, and the quicker a small shift in the process parameter is detected. Under the normality assumption, Crowder (1987, 1989) and Lucas and Saccucci (1990) provided the optimal values of λ that correspond to different magnitudes of mean shifts. If $\lambda = 1$, the EWMA chart is equivalent to the Shewhart chart.

Commonly, Z_0 is set equal to a target or the expectation of M_i. Setting Z_0 larger than this expectation, is referred to as giving the chart a head start. See Lucas and Saccucci (1990) for more details on EWMA control charts with the head start feature.

The EWMA chart signals when the statistic Z_i exceeds the control limits. The upper control limit (UCL) and lower control limit (LCL) can be determined by

$$UCL_i = \mu_{Z_i} + L\sigma_{Z_i} \quad \text{and} \quad LCL_i = \mu_{Z_i} - L\sigma_{Z_i} \tag{1.1}$$

where L is a positive coefficient. Together with λ, L determines the performance of the EWMA control chart. Under the assumption of independent and normally distributed data, it follows that the expectation of Z_i, denoted by μ_{Z_i}, is equal to the expectation of M_i. Moreover, the standard deviation of Z_i, denoted by σ_{Z_i}, is equal

to

$$\sigma_{Z_i} = \sigma_M \sqrt{\frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2i}]}.$$

Where σ_M denotes the standard deviation of M_i. The standard deviation of Z_i is time dependent. This so-called time-varying standard deviation converges, as i increases, to

$$\sigma_{Z_i} = \sigma_M \sqrt{\frac{\lambda}{2-\lambda}}.$$

Steiner (1999) and Abbasi (2010) studied the difference between EWMA charts based on time-varying and asymptotic limits.

The parameters μ_M and σ_M are functions of the process parameters, which are usually unknown. Therefore, control charts are generally implemented in two phases. In the first phase, Phase I, the in-control state of the process characteristic is determined and the distributional parameters are estimated (Vining, 2009; Chakraborti et al., 2009). In the second phase, Phase II, the estimated process parameters are used to set up the EWMA control chart and the process is prospectively monitored to detect changes from the in-control state.

In this dissertation we study the EWMA chart based on estimated process parameters. The dissertation consists of two parts, referred to as 'Phase I' and 'Phase II' in analogy to the two phases of control charting. In Phase I (Chapters 2 and 3) we consider the estimation of the process parameters. In Phase II (Chapters 3 and 4) we consider the performance of the EWMA chart when parameters are estimated. The motivation for each of the two 'Phases' is discussed in the following two sections.

1.4 Phase I and contaminated data

In a survey of Phase I analysis, Jones-Farmer et al. (2014) review the major issues and developments in Phase I analysis. One of the issues is the possibility of unacceptable data in Phase I. This section discusses this motivation for Chapters 2 and 3.

A motivating example: leads

Consider again the data concerning the number of leads. To set up the control charts in Figure 1.1 estimates of the process parameters are needed. To this end, historical data were collected from 2014 (weeks 25 through 52), see Figure 1.2. As indicated in Figure 1.2 the data may contain various contaminations. It was found that in weeks 25 and 44 the webmaster requested some brochures in order to test functions on the website. Furthermore, in weeks 32 and 33 there was a problem with the server host and the brochure request form was not functioning. These contaminated observations do not belong to the in-control process, and hence need to be identified and eliminated as they will influence the estimates of the in-control process parameters.

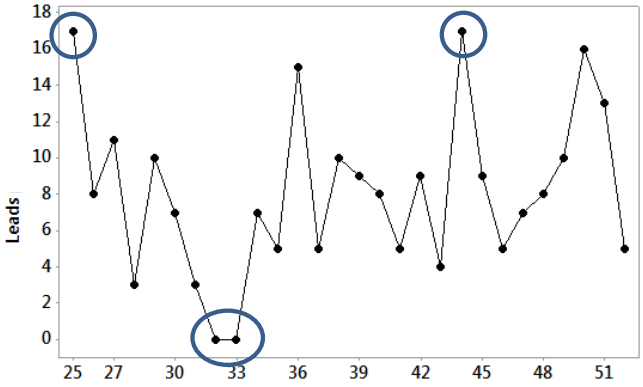


Figure 1.2: *Run chart of historical data on the number of leads*

This example does not stand alone; it is common for a (Phase I) data set to contain outliers, shifts or other forms of contaminated observations (see e.g. Vining et al., 2015). Contaminations in Phase I are problematic as they can influence the parameter estimates, resulting in Phase II control charts with less ability to detect changes in the process characteristic.

Approaches for contaminations in the Phase I data

The path that we take to deal with contaminations in Phase I is clear from the titles of Chapters 2 and 3; both start with the term *robust*. 'Robust' was introduced into industrial statistics by George Box (1953), by which he meant a (statistical) test which is still useful if the underlying assumption on the distributional properties of the data might not be true.

The use of robust estimators in Phase I is generally accepted in the SPM literature. For example, both Jensen et al. (2006) and Jones-Farmer et al. (2014b) noted that the use of robust estimators is appropriate in Phase I. It was Tukey (1960) who first considered the robustness of point estimators and detailed how to model contaminations with mixtures of normal distributions. In line with his view, we consider robust methods to have at least the following properties;

- they downgrade the influence of the contaminated observations in the data on the final estimate;
- they produce correct estimates if the measurements have not been contaminated.

The traditional robust estimators are point estimators. Some references are Rocke (1989), Tatum (1997), and Janacek and Meikle (1997). Unfortunately, as we will see in Chapters 2 and 3, robust point estimators are usually not very efficient under uncontaminated data.

Another possibility is the use of Phase I Shewhart charts which identify potentially contaminated samples, remove these from Phase I, and use the remaining samples to estimate the process parameters (e.g. Schoonhoven et al. 2011a, 2011b). An overview of Phase I charts for univariate data was given by Chakraborti et al. (2009) and Jones-Farmer et al. (2014b).

A third approach is to use change point methods. These methods are especially suited for detecting sustained changes in the process parameters. There is a long tradition of testing for sustained shifts in Phase I; for a literature overview see Amiri and Allahyari (2012).

The optimal choice of an estimation method requires knowledge of the type of contaminations. Typically, Phase I Shewhart charts are suitable when outliers can be present in the Phase I data set and change point methods are suitable if sustained shifts may occur in the Phase I data set. The aim of Chapters 2 and 3 is to introduce a new Phase I estimation methodology for the process parameters which provides reliable estimates regardless of the type of contaminations in Phase I. This is achieved by using an EWMA chart in Phase I.

In Chapter 2, we compare and evaluate various *location* estimators. In Chapter 3, we compare and evaluate various *dispersion* estimators.

1.5 Phase II and the effect of estimation

As a control chart is based on Phase I estimates, its control limits and hence performance will be conditional on the Phase I sample obtained. Each Phase I sample (from the same process) will yield different estimates and different control limits, yielding control charts which will show varying performance. In this section, we discuss the effect of estimation on the performance of the EWMA chart and give an introduction and motivation for Chapters 4 and 5.

The performance of control charts is commonly evaluated using characteristic of the run length distribution. The run length of a control chart is a random variable defined as the number of successively plotted statistics until the chart signals. One of the most common measures of control chart performance is the average run length (ARL). It is favourable to have a large ARL if the process is in control and a small ARL if the process is out of control. When parameters are estimated the control chart's performance will depend on the estimated parameters and will thus vary among practitioners. This is because practitioners use different Phase I data sets, which result in different parameter estimates, control limits, and chart performance (i.e. different ARL values). In Saleh et al. (2015a) this variation is referred to as *practitioner-to-practitioner*

variability. Equivalently, this variation can be viewed as sampling variability. Most often, charts are evaluated based on the average of the conditional ARLs (AARL), averaging across the sampling variability.

Figure 1.3 illustrates this sampling variability. It presents a boxplot of 100,000 simulated in-control ARL values for the EWMA chart for location. For each of the simulation runs, parameter estimates of the mean and standard deviation were obtained based on 50 randomly generated independent and normally distributed samples of 5 observations. Logically, the estimated values for the mean and standard deviation vary across the 100,000 drawn Phase I samples. With each pair of Phase I estimates an EWMA chart for location was set up. We use $\lambda = 0.1$ and $L = 2.454$ such that the chart has an in-control average run length of 200, if the estimates are exactly on target. The ARL was computed for each of these 100,000 estimated pairs and these conditional ARLs are displayed in Figure 1.3. From the boxplot it follows that 50% of the practitioners would have a chart with an in-control ARL value between approximately 100 and 180 and thus receiving a false alarm every 100 to 180 observations on average. Furthermore, there are also 5% of the practitioners who will receive (on average) a false alarm within the first 50 samples.

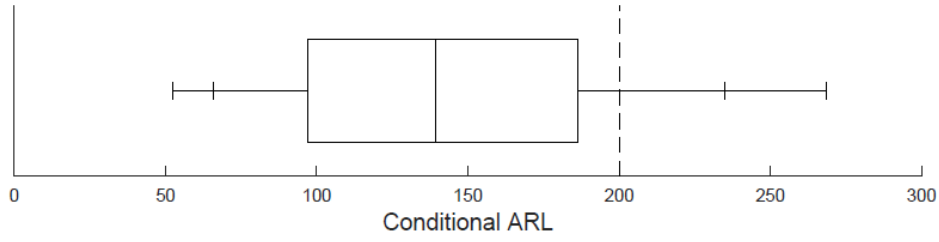


Figure 1.3: *Conditional in-control ARL. The boxplot shows the 5^{th}, 10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th}, and 95^{th} percentiles. Dotted line shows specified in-control ARL*

The effect of sampling on control charts, as illustrated in Figure 1.3, has received a great deal of attention in the SPM literature. See, for example, Quesenberry (1993), Chen (1997), Jones et al. (2001, 2002), and Saleh et al. (2015b). Jensen et al. (2006) and Psarakis et al. (2014) provided reviews of the literature on the performance of control charts with estimated parameters. The general consensus is that the use of parameter estimates results in control charts with less predictable statistical performance than those with known parameters. A specific problem is that charts give more frequent false alarms.

The results in Figure 1.3 show the necessity to take into account the effect of estimation error. In Chapter 4 we study this effect for the EWMA chart for location and suggest an alternative design procedure based on bootstrap. In Chapter 5 we study the effect of estimation on the EWMA chart for dispersion. Various designs exist of the EWMA chart for dispersion. We compare the effect of estimation across three of the

most commonly applied designs.

1.6 Contribution and outline of this dissertation

In this dissertation we contribute to the development and understanding of the EWMA control chart based on estimated parameters. We study both the EWMA chart for location as well as for dispersion.

In **Chapter 2** a new estimation method for the location based on EWMA charting is introduced. We compare this method to a wide range of existing location estimators. We consider the situation where Phase I could contain contaminated observations. This situation is relevant as in most practical applications data disturbances occur. We conclude that existing estimators are most effective given that it is known which pattern of contaminations are present in Phase I. However, if it is unknown which pattern of contaminations is present the new method gives the most precise estimate.

This chapter is based on two papers: Zwetsloot et al. (2014) and Zwetsloot et al. (2015b). The first paper entitled "A Robust Estimator of Location in Phase I Based on an EWMA Chart" has been published in the *Journal of Quality Technology*. It provides the basis for Chapter 2. The second paper, entitled "Robust Point Location Estimators for the EWMA Control Chart", has been accepted for publication in a special issue on SPM in *Quality Technology and Quantitative Management*. Both papers were combined work with dr. M. Schoonhoven and prof. dr. R.J.M.M. Does. For the paper Zwetsloot et al. (2014), I took the lead in the computation of the results presented, M. Schoonhoven took the lead in writing the paper, and R.J.M.M. Does provided supervision. For the paper Zwetsloot et al. (2015b) I took the lead and M. Schoonhoven and R.J.M.M. Does provided supervision. Both papers originated from my master's thesis "A Robust EWMA Control Chart" (Zwetsloot, 2012), written under the supervision of prof. dr. R.J.M.M. Does and prof. dr. H.P. Boswijk.

In **Chapter 3** we focus on estimating the process dispersion. We follow the same lines as in Chapter 2, and compare estimation methods for the dispersion when the data may contain various patterns of contamination. We propose a new estimation method based on EWMA charting, which is an effective estimation method, if the pattern of contamination in the data is unknown.

This chapter has been published under the title "A Robust Phase I Exponentially Weighted Moving Average Control Chart for Dispersion" in *Quality and Reliability Engineering International*. This paper was combined work with dr. M. Schoonhoven and prof. dr. R.J.M.M. Does (Zwetsloot et al., 2015a) in which I took the lead.

Chapter 4 is concerned with the effect of estimation on the monitoring (Phase II) performance of the EWMA chart for location. We show that it can be extremely difficult to lower the variation in the performance sufficiently due to practical limitation on the

amount of the Phase I data. To deal with this, we recommend an alternative design criterion and a procedure based on bootstrap.

This chapter is based on the paper "Another Look at the EWMA Control Chart with Estimated Parameters" which has been published in the *Journal of Quality Technology*. This paper was combined work with ms. N. Saleh, prof. dr. M.A. Mahmoud, prof. dr. L.A. Jones-Farmer, and prof. dr. W.H. Woodall (Saleh et al., 2015a). Most of my efforts have gone into the theory and numerical results for the bootstrap method presented in Section 5 and Appendix B of this paper (respectively Sections 4.5 and 4.8 of this dissertation).

As Box et al. (1978, Chapter 5) noted "most often we are interested in possible differences in the mean level .... Sometimes, however, it is the degree of variation of the data that is of interest." In **Chapter 5** we consider the Phase II monitoring of the dispersion. Various designs of the EWMA chart for dispersion are available. These designs vary in the choice of the dispersion measure. The most popular choice, in the literature, is the logarithm of the sample variance. Other designs are based on the sample variance or the sample standard deviation. In **Chapter 5** we compare these three EWMA dispersion charts based on estimated parameters. We argue that the chart which is less influenced by estimation error (i.e. the chart based on the sample variance) should be used in practice.

This chapter is based on the single authored paper "A Comparison of EWMA Control Charts for Dispersion based on Estimated Parameters" that has been submitted for publication (Zwetsloot, 2015).

Finally, **Chapter 6** provides a summary of this dissertation.

Phase I

Estimation from Contaminated Data

"In the majority of practical instances, the most difficult job of all is to choose the sample that is to be used as the basis for establishing the tolerance range [control limits]."

W. A. Shewhart (1939, page 76)

"Some robust estimator ... should be used on any data set...; there is little to lose, much to gain, and sufficient evidence that the result will be positive."

Rocke, Downs, and Rocke (1982, page 100)

2. Robust Estimators for Location

One of the issues in Phase I analysis is the possibility of contaminated observations in the data. To deal with the effect of contaminations, robust estimation methods are generally recommended. In this chapter, we propose a robust estimation method for the location parameter. Furthermore, we compare this new method to various existing estimation methods. This chapter is based on Zwetsloot et al. (2014) and Zwetsloot et al. (2015b).

2.1 Introduction

It is generally accepted in the literature that a control chart is implemented in two phases: Phase I, to define the stable state of the process characteristic and to estimate its distributional parameters; and Phase II to monitor the process. One of the issues in Phase I analysis is the possibility of unacceptable -contaminated- data in Phase I (Jones-Farmer et al., 2014b).

The approach we take to deal with contaminations in Phase I, is the use of robust estimation methods (see Section 1.4). Numerous robust (point) estimators have been discussed in the literature. Some robust point estimators can be found in Rocke (1989), Janacek and Meikle (1997), and Langenberg and Iglewicz (1986). Another possibility, we consider, is the use of Shewhart control charts in Phase I to screen the data. A review of these methods can be found in Chakraborti et al. (2009) and Jones-Farmer et al. (2014b). Furthermore, we consider a method particularly suited to detect sustained shifts: a change point detection method. Amiri and Allahyari (2012) provided an overview on this topic.

The optimal choice of estimation methods requires the knowledge of the type of contaminations. Typically, Phase I Shewhart charts are suitable when outliers are present in Phase I (see, e.g. Schoonhoven et al., 2011b). Change point methods are

suitable if sustained shifts occur (see, e.g. Sullivan and Woodall, 1996). In practice it is often unknown *if* contaminations are present, let alone *what patterns* of contaminations are present. Therefore, in this chapter, we introduce a new Phase I estimation methodology for the location which provides reliable estimates regardless of the pattern of contaminations. This is achieved by using an estimation method based on an EWMA chart in Phase I.

This idea follows an insight from Hunter (1986). He provided a useful analysis of the degree in which the Shewhart, CUSUM and EWMA control charts use the history of the data to detect a change in the process mean. He pointed out that the Shewhart control chart only uses the current observation and therefore has no memory while, in contrast, the CUSUM control chart uses all of the history paying equal attention to all past observations and the current observation. This is an oversimplification because, in fact, the CUSUM chart uses a decision rule whereby some of the past observations can become irrelevant. The EWMA control chart, gives less weight to data as they get older. The weight given to the current observation relative to earlier observations can be chosen by selecting a smoothing parameter between 0 and 1: a value of 1 means that all of the weight is assigned to the current observation and no weight to previous observations, equivalent to the Shewhart control chart, while a value of almost 0 results in a control chart with a long memory. We shall use these specific properties of the EWMA control chart in the proposed estimation method.

This chapter is organized as follows. In Section 2.2 we present background information on the EWMA chart for location. In Sections 2.3 and 2.4 various estimation methods are presented. In Section 2.5 we compare the effectiveness of these estimation methods. In Section 2.6 we describe the results for the Phase II context and in Section 2.7 we summarize our conclusions.

2.2 The EWMA control chart for location

In this chapter we consider estimating and monitoring the location. To this end samples of size $n \geq 1$ are collected prospectively from the process. We assume that the process characteristic is independent and normally distributed with $\mu = \mu_0$ and $\sigma = \sigma_0$ if the process is in control.

To set up the EWMA chart, we use the mean of each sample to compute the EWMA statistic

$$Z_i = (1 - \lambda)Z_{i-1} + \lambda\bar{Y}_i, \tag{2.1}$$

where $\bar{Y}_i$ is the mean of sample i. The chart signals a (possible) out-of-control condition when Z_i falls beyond the control limits. The upper control limit (UCL) and

lower control limit (LCL) are, at time i, equal to:

$$UCL_i/LCL_i = \mu_0 \pm L\frac{\sigma_0}{\sqrt{n}}\sqrt{\frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2i}]}, \tag{2.2}$$

where L is a positive coefficient which together with λ determines the performance of the EWMA control chart when the process is in control. Furthermore, we set $Z_0 = \mu_0$.

In practice, μ_0 and σ_0 are unknown and estimates need to be obtained within Phase I, the exploratory data analysis phase. To obtain these, m samples of size n are collected from the process. We denote these observations by X_{ij}, with $i = 1, 2, ..., m$ and $j = 1, 2, .., n$, and the Phase I estimates of μ_0 and σ_0 by $\hat{\mu}_0$ and $\hat{\sigma}_0$, respectively. In this chapter we consider a limited number of available Phase I data and set $m = 50$ and $n = 5$. In Zwetsloot et al. (2014) we also studied Phase I samples of size $n = 10$.

2.3 Location estimation methods

In this section, we describe various estimation methods, for the location parameter $\hat{\mu}_0$, that can be used within Phase I. We consider point estimators and - especially useful for the detection of sustained shifts - a change point method. We also present estimation methods based on control charting in Phase I.

Point estimators

Many point estimators for location have been proposed in the literature, see for example Langenberg and Iglewicz (1986), Rocke (1989), and Wang et al. (2007). In Zwetsloot et al. (2015b) we compared six of these estimators for $\hat{\mu}_0$. One of these estimators was the overall sample mean, which is known as the most efficient estimator for the location under uncontaminated normal distributed data. We discovered that, if data anomalies can be present, the use of the median of the sample means is a good alternative for the traditional overall sample mean. For conciseness we only consider these two (out of six) point estimators here. The overall sample mean $\bar{\bar{X}}$ is defined by

$$\bar{\bar{X}} = \frac{1}{m}\sum_{i=1}^{m}\bar{X}_i = \frac{1}{m}\sum_{i=1}^{m}\left(\frac{1}{n}\sum_{j=1}^{n}X_{ij}\right),$$

and the median of the sample averages $M(\bar{X})$ is defined by

$$M(\bar{X}) = \text{median}(\bar{X}_1, ..., \bar{X}_m).$$

Change point method

Detecting structural changes in a process characteristic may be done using a change point method (Amiri and Allahyari, 2012). Change point methods compare the log

likelihood of all observations, under the assumption that all observations are in control, with the log likelihood of the observations if a step change has occurred. Sullivan and Woodall (1996) proposed a change point method for exploratory data analysis and showed that this method outperforms the Shewhart chart in detecting sustained shifts. We include their change point method in our analysis.

Denote the maximum likelihood estimator of the variance of the observations in samples l through k by

$$\hat{\sigma}^2_{l:k} = \frac{1}{n(k-l+1)} \sum_{i=l}^{k} \sum_{j=1}^{n} (X_{ij} - \bar{\bar{X}}_{l:k})^2,$$

where $\bar{\bar{X}}_{l:k}$ is the overall sample average of all observations in samples l through k. To test for the existence of a sustained shift after sample τ, Sullivan and Woodall (1996) computed the likelihood ratio statistic as

$$LRT(\tau) = nm \ln[\hat{\sigma}^2_{1:m}] - n\tau \ln[\hat{\sigma}^2_{1:\tau}] - n(m-\tau) \ln[\hat{\sigma}^2_{\tau+1:m}]. \tag{2.3}$$

Because the expected value of $LRT(\tau)$ varies with τ, they first divide $LRT(\tau)$ by its expected value under in-control data. Table 2.1 presents these expected values for $m = 50$ and $n = 5$[1]. A chart can be constructed by plotting the standardized $LRT(\tau)$ versus τ and a sustained shift in Phase I is signalled if $LRT(\tau)$ exceeds the upper control limit UCL_{CP}. Every out-of-control signal indicates a possible sustained shift in the process, i.e. the process parameters in samples $1, .., \tau$ are different from the process parameters in samples $\tau+1, .., m$. When multiple signals are given, we set the estimated change point ($\hat{\tau}$) equal to the τ for which standardized $LRT(\tau)$ is largest. If there is no out-of-control signal we set $\hat{\tau} = m$.

τ	E[LRT]	τ	E[LRT]	τ	E[LRT]	τ	E[LRT]	τ	E[LRT]
1		11	2.04	21	2.02	31	2.02	41	2.04
2	2.21	12	2.03	22	2.02	32	2.02	42	2.05
3	2.14	13	2.03	23	2.02	33	2.03	43	2.06
4	2.10	14	2.03	24	2.02	34	2.03	44	2.07
5	2.08	15	2.03	25	2.02	35	2.03	45	2.08
6	2.07	16	2.03	26	2.02	36	2.03	46	2.10
7	2.06	17	2.03	27	2.02	37	2.03	47	2.13
8	2.05	18	2.02	28	2.02	38	2.03	48	2.21
9	2.04	19	2.02	29	2.02	39	2.04	49	
10	2.04	20	2.02	30	2.02	40	2.04	50	

Table 2.1: *E[LRT], the expectation of the likelihood ration statistic $LRT(\tau)$*

[1]The values in Table 2.1 differ slightly from the values for E[$LRT(\tau)$] in Table 5 in Zwetsloot et al. (2014). The values in Zwetsloot et al. (2014) are based on simulation. The values in Table 2.1 are based on an analytical expression of the expectation of a lognormally distributed variable.

To determine UCL_{CP}, we set the overall in-control false alarm rate equal to 1 percent. Using 100,000 simulations, we find that $UCL_{CP} = 5.75$ for $m = 50$ and $n = 5$.

When the change point $\hat{\tau}$ is estimated, we can determine which samples are out of control. In practice, knowledge of the process would be used to determine whether the data before or after the estimated $\hat{\tau}$ are in control. In order to prevent the deletion of a large proportion of clean observations from the Phase I data set (this problem could occur in our simulation if there is a false alarm at the beginning of the Phase I data set), we use the following decision rule: 'the majority of the samples represents the in-control process'. This implies that, if $\hat{\tau} \leq m/2$, we delete samples 1 up to $\hat{\tau}$ from Phase I. If $\hat{\tau} > m/2$, we delete samples $\hat{\tau} + 1$ up to m from Phase I. This rule ensures a more suitable comparison with the other Phase I methods considered in this chapter. The remaining samples are used to compute the overall sample mean, yielding an estimate of μ_0 based on change point analysis, which we denote by CP. We believe this is an appropriate decision rule as practitioners can investigate which sequence before or after the shift is acceptable.

The considered change point estimation method, CP, is designed to detect a single change point $\hat{\tau}$ and is at a disadvantage if multiple step changes occur in Phase I. Alternative change point methods can be designed based on recursive testing for step changes (see, e.g. Capizzi and Masarotto, 2013).

The proposed estimation method

In this section, we propose an estimation method for the location based on the use of an EWMA chart in Phase I. This new estimation method provides a robust estimate of the location when it is unknown what type of contaminations may be present in Phase I.

The proposed Phase I EWMA screening estimators consist of the following steps:

1. *Compute an initial (robust) estimate of the location and dispersion* based on all observations in Phase I, denote these by $\hat{\mu}_I$ and $\hat{\sigma}_I$, respectively. The subscript 'I' denotes that the parameter is associated with Phase I charting.

2. *Set up a Phase I* EWMA *chart* according to Equations (2.1) and (2.2) using $\mu_0 = \hat{\mu}_I$, $\sigma_0 = \hat{\sigma}_I$, $Z_0 = \hat{\mu}_I$, $L = L_I$, and $\lambda = \lambda_I$.

3. *Delete from Phase I all samples for which the corresponding* EWMA *statistic gives an out-of-control signal.*

4. *Use an efficient estimator of μ_0 based on the remaining samples.* This yields an estimate which is efficient as well as robust to various patterns of contaminations. We use the overall sample mean based on the remaining samples.

The resulting estimator is denoted by $s\widehat{\mu}_{I,\lambda_I}$, where '$s$' indicates that we use a screening Phase I chart, $\widehat{\mu}_I$ stands for the initial location estimator chosen in step 1,

and the subscript λ_I denotes the value selected for the smoothing constant in step 2. To operationalize this screening estimator, we need to select an estimator for $\widehat{\mu}_I$ and $\widehat{\sigma}_I$ and values for λ_I and L_I.

The choice of $\hat{\mu}_I$ is an important one: an efficient estimator could improve the performance of the Phase I chart under stability but inflate the Phase I control limits when disturbances are present. On the other hand, a robust estimator could result in non-optimal performance under stability, but in robust Phase I control limits when disturbances are present. Our study, in this chapter, evaluates the impact of the most efficient estimator for the location, the overall sample mean ($\bar{\bar{X}}$), and a robust estimator based on the median of the sample means ($M(\bar{X})$).

Throughout this chapter, we use a single method to estimate $\widehat{\sigma}_I$. To ensure that differences in performance are due to the difference in the estimation of the location. The estimator for $\hat{\sigma}_0$ is discussed in the next section.

As far as the choice of λ_I is concerned in step 2, we take a range of values in order to study the impact of this parameter. Small values for λ_I enable detection of sustained shifts while larger values of λ_I enable detection of scattered outliers. To assess this trade-off, we set λ_I equal to 0.2, 0.6 and 1. When $\lambda_I = 1$, we obtain the Shewhart chart. Estimators based on Shewhart charts in Phase I were also studied by Schoonhoven et al. (2011a). To obtain values for L_I, we set the false alarm rate at 1 percent, thereby following Chakraborti et al. (2009).

Table 2.2 gives an overview of the Phase I estimators considered and the corresponding values of L_I (obtained through 100,000 Monte Carlo simulations).

$\widehat{\mu}_0$	Description	L_I
$\bar{\bar{X}}$	Overall sample mean	n.a.
$M(\bar{X})$	Median of the sample averages	n.a.
CP	Change point estimator	n.a.
$s\bar{\bar{X}}_{0.6}$	Screening estimator with $\widehat{\mu}_I = \bar{\bar{X}}$ and $\lambda_I = 0.6$	2.540
$sM(\bar{X})_{0.2}$	Screening estimator with $\widehat{\mu}_I = M(\bar{X})$ and $\lambda_I = 0.2$	2.540
$sM(\bar{X})_{0.6}$	Screening estimator with $\widehat{\mu}_I = M(\bar{X})$ and $\lambda_I = 0.6$	2.610
$sM(\bar{X})_1$	Screening estimator with $\widehat{\mu}_I = M(\bar{X})$ and $\lambda_I = 1$	2.617

Table 2.2: *Phase I location estimators*

2.4 Tatum's dispersion estimator

We use an estimator for the dispersion which is known for its robustness and was proposed by Tatum (1997). This estimators was recommended in Schoonhoven and Does (2012) and Schoonhoven et al. (2011b) if contaminations may be present.

To compute Tatum's estimator, one first needs to centre each observation on its sample median creating residuals e_{ij}. If n is odd, each sample contains one residual equal to zero, which is dropped. The resulting $n'm$ residuals, with $n' = n - 1$ if n

is odd and $n' = n$ if n is even, are weighted. Large residuals are given less weight than smaller residuals, which ensures that outliers have less impact on the estimate of σ_0. This gives $u_{ij} = \frac{h_i e_{ij}}{cA^*}$, where A^* is the median of the absolute values of the $n'm$ residuals,

$$h_i = \begin{cases} 1 & E_i \leq 4.5, \\ E_i - 3.5 & 4.5 < E_i \leq 7.5, \\ c & E_i > 7.5, \end{cases}$$

$E_i = IQR_i/A^*$, where IQR_i is the interquartile range of sample i, and c is a tuning constant. Where IQR_i is defined as

$$IQR_i = X_{i(b)} - X_{i(a)}, \tag{2.4}$$

where $X_{i(o)}$ denotes the o-ordered value in sample i, $a = \lceil n/4 \rceil$, and $b = n - a + 1$. The ceiling function $\lceil z \rceil$ denotes the smallest integer not less than z. To estimate σ_0, only the residuals that are small, i.e. for which $|u_{ij}| \leq 1$, are used

$$S^* = \frac{n'm}{\sqrt{n'm - 1}} \frac{\sqrt{\sum_{i=1}^{m} \sum_{j:|u_{ij}|<1} e_{ij}^2 (1 - u_{ij}^2)^4}}{|\sum_{i=1}^{m} \sum_{j:|u_{ij}|<1} (1 - u_{ij}^2)(1 - 5u_{ij}^2)|}.$$

Tatum (1997) showed that for $c = 7$ the estimator is robust against various contaminations. An unbiased estimator of σ_0 is given by $S^*/d(n, m, c)$, where $d(5, 50, 7) = 1.068$ (obtained through $100,000$ Monte Carlo simulations). Note that we follow the implementation as set out in Schoonhoven et al. (2011b).

2.5 Phase I comparison

In this section, we evaluate the effectiveness of Phase I studies that use the methods presented in Table 2.2. One of the requirements of Phase I is to deliver accurate parameter estimates. We assess the estimation precision of the proposed methods in terms of the Mean Squared Error (MSE). In addition, the Phase I analysis is used as a tool for exploratory data analysis, allowing us to examine the data and learn from out-of-control observations. We assess this by measuring the percentage of identified out-of-control observations. First we describe the data scenarios considered in Phase I.

Contamination scenarios

Recall that the stable, in-control, Phase I data are assumed to be $N(\mu_0, \sigma_0^2)$ distributed. If Phase I contains contaminated observations, we assume that these come from a shifted normal distribution $N(\mu_0 + \delta_I \sigma_0, \sigma_0^2)$, with δ_I a constant.

Many different patterns of contaminations can occur in practice and have been studied in the literature. A distinction can be made between *scattered* and *sustained*

special causes of variation. Scattered disturbances are transient in that they affect single samples. Whereas sustained disturbances last for at least a few consecutive samples beyond their first appearance. In this dissertation, we evaluate two scattered scenarios - localized and diffuse - based on Tatum (1997) and Schoonhoven and Does (2012) and two sustained shift scenarios - single and multiple step shifts - based on Chen and Elsayed (2002) and Amiri and Allahyari (2012). These four scenarios are described below, were we set the parameters for the in-control process at $\mu_0 = 0$ and $\sigma_0 = 1$, without loss of generality. Furthermore, recall that the Phase I data consists of $m = 50$ samples of size $n = 5$.

1. A model for *localized location disturbances* in which all observations in a sample have a 90% probability of being drawn from the $N(0, 1)$ distribution and a 10% probability of being drawn from the $N(\delta_I, 1)$ distribution.

2. A model for *diffuse location disturbances* in which each observation, irrespective of the sample to which it belongs, has a 90% probability of being drawn from the $N(0, 1)$ distribution and a 10% probability of being drawn from the $N(\delta_I, 1)$ distribution.

3. A model for a *single step shift* in the location. In which the first 45 samples are drawn from the $N(0, 1)$ distribution and the last 5 samples are drawn from the $N(\delta_I, 1)$ distribution.

4. A model for *multiple step shifts* in the location. In which, at each time point, the sample has a probability p of being the first of five consecutive samples drawn from the $N(\delta_I, 1)$ distribution. After any such step shift, each sample again has a probability p of being the start of another step shift. Phase I consists of 50 samples. If sample 48 shifts, then only 3 samples (48, 49 and 50) are drawn from the $N(\delta_I, 1)$ distribution, instead of 5. To maintain the 10% (expected) contamination rate of models 1-3, we set $p = 0.023$.

The performance of the proposed estimators is evaluated for scenarios where $\delta_I = 0, 0.2, 0.4, ..., 2$. Note that for $\delta_I = 0$ the data come from the in-control distribution and hence no contaminations are present. An overview of the contamination scenarios is provided in Table 2.3.

Performance measures and simulation procedure

In order to determine the accuracy of the proposed location estimators, we determine the Mean Squared Error (MSE) for each estimation method under the contamination scenarios proposed above. The MSE is calculated as

$$MSE = \frac{1}{R}\sum_{r=1}^{R}\left(\frac{\hat{\mu}_0^r - \mu_0}{\sigma_0}\right)^2 = \frac{1}{R}\sum_{r=1}^{R}(\hat{\mu}_0^r)^2,$$

Contamination scenarios	Description
In-control	All observations from $N(0,1)$
Localized disturbances	$90\% - 10\%$ random mixture of samples from $N(0,1)$ and $N(\delta_I^2, 1)$
Diffuse disturbances	$90\% - 10\%$ random mixture of observations from $N(0,1)$ and $N(\delta_I, 1)$
Single step shift	Samples $1-45$ are $N(0,1)$ and samples $46-50$ are $N(\delta_I, 1)$
Multiple step shifts	Shifts of length 5 from $N(\delta_I, 1)$ occurring with probability 0.023

Table 2.3: *Phase I contamination scenarios affecting the location parameter*

where $\hat{\mu}_0^r$ is the value of one of the proposed estimators in the rth simulation run, and R is the total number of simulations in the Monte Carlo study.

The proposed estimators are also evaluated with two additional performance metrics: the true alarm percentage (TAP) and the false alarm percentage (FAP). These additional performance measures reflect the ability of the screening estimators to detect unacceptable observations without triggering false alarms for acceptable observations. Related measures were presented by Fraker et al. (2008), Chakraborti et al. (2009) and Frisén (2009). The TAP and FAP are calculated as

$$TAP = \frac{1}{R}\sum_{r=1}^{R} \frac{(\#\text{correct signals})^r}{(\#\text{unacceptable observations})^r} \times 100\% \tag{2.5}$$

and

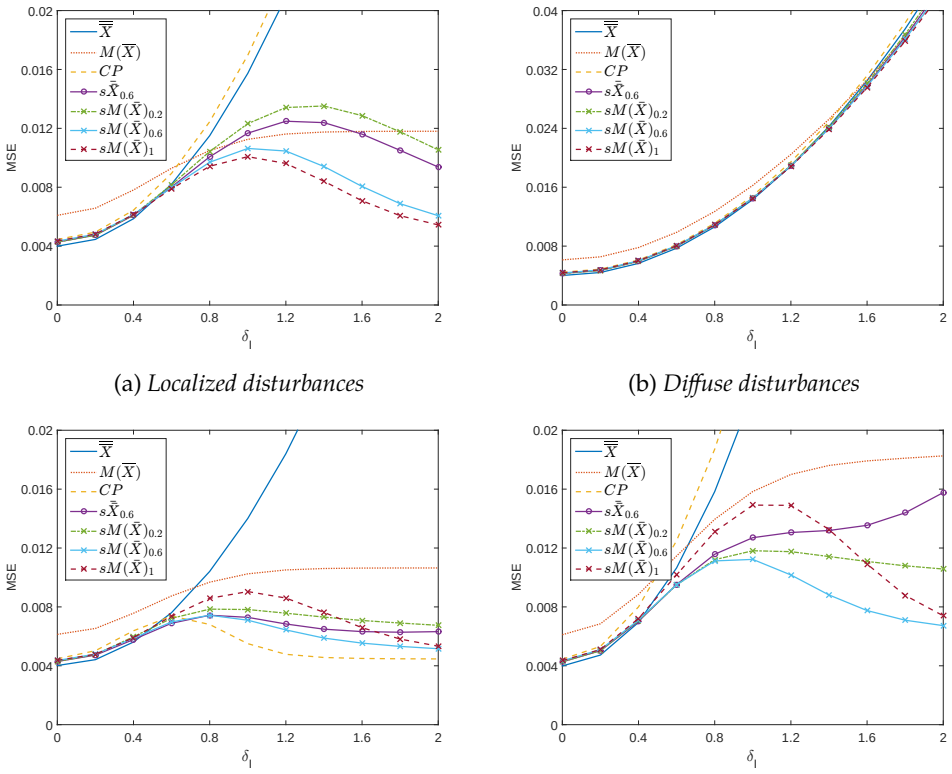
$$FAP = \frac{1}{R}\sum_{r=1}^{R} \frac{(\#\text{false alarms})^r}{(\#\text{acceptable observations})^r} \times 100\%, \tag{2.6}$$

where r denotes the r-th simulation run.

A Monte Carlo simulation study was conducted. We drew R Phase I data sets consisting of $m = 50$ samples of size $n = 5$, for each of the four contamination scenarios, and for each value of δ_I. The proposed seven location estimators, presented in Table 2.2, were calculated for each simulation run, and the three performance measures were computed based on the R runs. We set $R = 200,000$ which gives us a relative simulation error - the standard error of the estimated MSEs expressed as a percentage of the MSE - which never exceeds 0.5%.

Phase I results: estimation accuracy

First consider for each of the proposed estimation methods, the estimation accuracy as evaluated by the MSE. The MSE results are presented in Figure 2.1. The y-



(a) *Localized disturbances*

(b) *Diffuse disturbances*

(c) *Single step shift*

(d) *Multiple step shifts*

Figure 2.1: MSE *of location estimators when various types of contaminations are present in Phase I*

intercept of each subplot shows the MSE of the estimation methods when the data are in control ($\delta_I = 0$). The estimator $\bar{\bar{X}}$ shows the lowest MSE level, as expected. The other estimators show only slightly larger MSE levels, except for $M(\bar{X})$, which is least efficient under in-control data.

Next, we study the situation when contaminations are present in Phase I ($\delta_I > 0$). We see that the traditional point estimator $\bar{\bar{X}}$ is most sensitive to all data scenarios considered. The estimator $M(\bar{X})$ as well as the screening estimators are rather robust in the scenarios where the mean of an entire sample has shifted, namely localized, single, and multiple step shifts (see Figures 2.1a, 2.1c, and 2.1d), but not when diffuse disturbances are present (see Figure 2.1b). The reason is that these estimators trim the samples with a large mean rather than extreme observations within a sample. The estimator CP has the lowest MSE level when there is a single step shift (Figure 2.1c) but its performance in other situations is far worse than that of the other estimators.

Regarding the choice of λ_I for the screening methods, we had expected that the

Phase I Shewhart chart, $sM(\bar{X})_1$, would perform best for localized disturbances as the Shewhart chart is well known for its detection of single (extreme) disturbances. However, from Figure 2.1a we can see that the Phase I Shewhart chart is only slightly superior to the Phase I EWMA chart with $\lambda_I = 0.6$: $sM(\bar{X})_{0.6}$. Moreover, $sM(\bar{X})_{0.6}$ performs better when there are single or multiple step shifts. Note that an EWMA chart with a lower λ_I, for example $sM(\bar{X})_{0.2}$, does not perform as well for localized and multiple step shifts: a lower value of λ_I is more suitable for smaller single step shifts. If the disturbances in applications can be scattered as well as sustained, we recommend in Phase I the use of an EWMA chart with $\lambda_I = 0.6$ or a similar intermediate value, rather than a Shewhart chart.

As for the choice of $\hat{\mu}_I$, i.e. whether we use $sM(\bar{X})_{0.6}$ or $s\bar{\bar{X}}_{0.6}$, it is worth noting that the method based on the robust estimator $M(\bar{X})$ for the Phase I chart is as efficient under stable data ($\delta_I = 0$) as the chart based on $\bar{\bar{X}}$. This becomes clear when we realize that both charts use the efficient estimator $\bar{\bar{X}}$ to determine the mean after screening (step 4 of the algorithm). We can conclude that it does not matter for efficiency of the final estimate whether a less efficient estimator is used to construct the Phase I chart. The use of a robust estimator like $M(\bar{X})$ for the Phase I chart does pay off however: when there are large multiple step shifts (Figure 2.1d for $\delta_I > 0.8$), we see that the performance of the Phase I chart based on $\bar{\bar{X}}$ is not good. The higher is the value of δ_I, the higher will be the MSE level. When a non-robust estimator is used for the Phase I chart, disturbances might affect the Phase I limits so that the wrong observations are filtered out of the data. As the type of disturbance in Phase I is often unknown, we recommend the use of a Phase I chart based on a robust estimator such as $sM(\bar{X})_{0.6}$ rather than a Phase I chart based on an efficient estimator such as $s\bar{\bar{X}}_{0.6}$.

Finally, note that none of the proposed estimation methods perform well when there are diffuse disturbances, i.e. contaminated observations scattered over the entire Phase I data set (Figure 2.1b). Since the estimators screen whole samples, they do not identify these individual scattered outliers and therefore use all observations to estimate the location. We think that the proposed estimators can be augmented with a method that screens for individual outliers and see this as an issue for future research.

Throughout we have assumed that $m = 50$ samples of size $n = 5$ are available in Phase I. For the case of $m = 50$ samples of size $n = 10$ the conclusions are comparable and can be found in Zwetsloot et al. (2014).

To compare the performance of the proposed estimators across the various contamination schemes, we computed the Relative Mean Squared Error ($RMSE$) of the estimators. The $RMSE$ of an estimator, for a specific type of data contamination and severity δ_I, is defined as the percentage increase in the MSE level of the estimator relative to the MSE level of the estimator with the lowest MSE level. For each data scenario and each estimator, we obtain the RSME level of the estimator for all consid-

ered levels of δ_I. We present the maximum RSME over δ_I for each estimator in Table 2.4. In the presence of localized disturbances, the estimator $sM(\bar{X})_1$ has the lowest maximum $RMSE$ level (i.e. has a MSE which overall is closest to the optimal MSE for all shift sizes). When there is a single step shift, the change point estimator has the lowest $RMSE$ level: its MSE is at most 14 percent larger than the optimal estimator for all considered values of δ_I. If we consider all disturbance scenarios together (last row in Table 2.4), we find that the estimator $sM(\bar{X})_{0.6}$ is always within 36 percent of the optimal estimators' MSE level, irrespective of the pattern of contaminations.

	Phase I estimators $\hat{\mu}_0$						
Scenario	$\bar{\bar{X}}$	$M(\bar{X})$	CP	$s\bar{\bar{X}}_{0.6}$	$sM(\bar{X})_{0.2}$	$sM(\bar{X})_{0.6}$	$sM(\bar{X})_1$
Localized	836	117	936	73	94	14	**9**
Diffuse	**6**	53	11	7	7	8	9
Single	887	140	**14**	44	61	36	82
Multiple	1071	174	1247	137	59	**8**	51
All	1071	174	1247	137	94	**36**	82

Table 2.4: *Maximum Relative Mean Squared Error ($RMSE$) and, in bold, the location estimator with the lowest maximum $RMSE$ for the respective scenarios.*

Phase I results: detection probability

Apart from the precision of an estimate, it is also important that screening methods have the ability to detect unacceptable observations without triggering false alarms. Therefore, we evaluate the TAP and FAP as defined in Equations (2.5) and (2.6) for those estimation methods that have a screening procedure. The results are presented in Table 2.5.

Some interesting findings on the TAP and FAP are the following:

- When there are localized disturbances, $sM(\bar{X})_1$ shows the best performance since it has the highest TAP and lowest FAP values. Note that $sM(\bar{X})_{0.6}$, which is based on a robust initial estimator, detects more unacceptable observations than $s\bar{\bar{X}}_{0.6}$. This is because the robust Phase I control limits are not biased by any contaminations.
- All proposed estimation methods detect very few diffuse disturbances. This is not surprising given that they lack an effective way of identifying outliers within a sample.
- When there is a single step shift, the CP method performs best, followed by the methods based on a Phase I EWMA chart. The Shewhart chart performs poorly in this situation, which is to be expected, as this chart is especially designed to detect individual, scattered disturbances.

	TAP				FAP				
	δ_I				δ_I				
$\hat{\mu}_0$	0.4	1	1.6	2	0	0.4	1	1.6	2
Localized									
CP	1.4	4.1	11.2	17.0	1.1	1.3	2.6	6.7	11.1
$s\bar{\bar{X}}_{0.6}$	3.3	23.0	61.1	82.2	1.0	1.1	1.5	2.3	3.2
$sM(\bar{X})_{0.2}$	2.3	15.7	46.7	67.4	1.0	1.2	2.3	5.9	9.7
$sM(\bar{X})_{0.6}$	3.4	26.2	71.1	90.9	1.0	1.1	1.3	1.8	2.4
$sM(\bar{X})_1$	3.8	29.8	77.7	94.8	1.0	1.0	1.1	1.1	1.1
Diffuse									
CP	1.1	1.2	1.4	1.9	1.1	1.1	1.2	1.4	1.9
$s\bar{\bar{X}}_{0.6}$	1.1	1.4	2.2	2.9	1.0	1.0	1.0	1.0	1.0
$sM(\bar{X})_{0.2}$	1.0	1.2	1.7	2.2	1.0	1.0	1.0	1.0	1.1
$sM(\bar{X})_{0.6}$	1.1	1.4	2.2	3.1	1.0	1.0	1.0	1.0	1.0
$sM(\bar{X})_1$	1.1	1.5	2.5	3.4	1.0	1.0	1.0	0.9	1.0
Single step									
CP	15.1	90.7	99.3	99.8	1.0	1.7	1.3	0.2	0.1
$s\bar{\bar{X}}_{0.6}$	6.9	53.7	89.4	96.1	1.0	1.1	1.5	2.2	3.0
$sM(\bar{X})_{0.2}$	7.9	54.7	80.0	87.1	1.0	1.2	1.5	1.6	1.6
$sM(\bar{X})_{0.6}$	6.8	55.6	91.9	97.8	1.0	1.1	1.2	1.2	1.2
$sM(\bar{X})_1$	3.9	30.7	78.8	95.2	1.0	1.0	1.1	1.1	1.1
Multiple steps									
CP	4.1	28.6	48.3	53.1	1.2	1.9	7.0	13.9	16.2
$s\bar{\bar{X}}_{0.6}$	5.6	43.9	81.9	91.4	1.0	1.1	1.8	3.3	5.0
$sM(\bar{X})_{0.2}$	6.8	50.1	78.7	86.6	1.0	1.4	4.3	7.9	9.7
$sM(\bar{X})_{0.6}$	5.6	47.9	88.7	96.5	1.0	1.1	1.4	1.8	2.1
$sM(\bar{X})_1$	3.4	26.1	73.3	92.9	1.0	1.0	1.2	1.2	1.2

Table 2.5: *True Alarm Percentage (TAP) and False Alarm Percentage (FAP)*

- When there are multiple step shifts, the CP method, which we use, runs into trouble as it is designed to detect a single shift. A solution for this could be the use of a CP method which recursively identifies multiply change points. The Phase I chart based on a non-robust estimator, $s\bar{\bar{X}}_{0.6}$, deletes too many in-control observations. The EWMA chart with $\lambda_I = 0.6$ performs best.

2.6 Phase II performance

Phase I estimators are used to design Phase II control charts. In this section, we evaluate EWMA control charts in Phase II which are based on estimated parameters when

the Phase I data may or may not be contaminated.

Denote the Phase II observations by Y_{ij} with $i = 1, 2, ...$ and $j = 1, 2, ..., n$. They are assumed to be independent and $N(\mu_0, \sigma_0^2)$ distributed if the process is in control. We model out-of-control Phase II data as $N(\mu_0 + \delta_{II}\sigma_0, \sigma_0^2)$, where δ_{II} is the standardized shift size and the index 'II' indicates Phase II data.

Design of the Phase II EWMA control chart

The Phase II EWMA control chart consists of the EWMA statistic

$$Z_i = (1 - \lambda_{II})Z_{i-1} + \lambda_{II}\bar{Y}_i,$$

with $Z_0 = \hat{\mu}_0$, and control limits

$$\widehat{UCL}_i/\widehat{LCL}_i = \hat{\mu}_0 \pm L_{II}\frac{\hat{\sigma}_0}{\sqrt{n}}\sqrt{\frac{\lambda_{II}}{2 - \lambda_{II}}[1 - (1 - \lambda_{II})^{2i}]},$$

where $\hat{\mu}_0$ and $\hat{\sigma}_0$ are the Phase I estimates of μ_0 and σ_0. We consider seven EWMA Phase II control charts, each based on a different estimator $\hat{\mu}_0$, as presented in Table 2.2. For each of the seven charts we set σ_0 equal to Tatum's estimator, as presented in Section 2.4.

Furthermore, we set λ_{II} equal to 0.13, which Crowder (1989) recommended as an optimal smoothing constant to detect a shift size of $\delta_{II} = 1$. Note that λ_{II} differs from the λ_I used in the Phase I estimation methods, as Phase I is used for exploratory data analysis purposes. We take the values for L_{II} from Jones (2002): $L_{II} = 2.89$ for $n = 5$, such that the EWMA control chart has an average run length of approximately 370 when the process is in control.

Performance measures and simulation procedure

The performance of a control chart is commonly evaluated using characteristics of the run length distribution. The run length is a stochastic variable indicating the number of samples before Z_i falls outside the control limits, i.e. a signal that the process may be out of control. A common measure of control chart performance is the average run length (ARL). It is desirable to have a high ARL when the process is in control and a low ARL when the process mean has shifted. When parameters are estimated a distinction needs to be made between the conditional and unconditional average run length. The conditional ARL is the ARL given the Phase I parameter estimates $\hat{\mu}_0$ and $\hat{\sigma}_0$. In order to evaluate the overall behaviour of the EWMA control charts, we consider the average of the conditional ARLs ($AARL$), this unconditional average run length, averages over the variability of the parameter estimates $\hat{\mu}_0$ and $\hat{\sigma}_0$.

In order to obtain the $AARL$, we use the following simulation procedure: first $m = 50$ samples of size $n = 5$ are drawn from the same Phase I disturbance scenarios used to assess the MSE, with $\delta_I = 1$. Then $\hat{\mu}_0$ and $\hat{\sigma}_0$ are calculated from the data

and the control limits are computed. Observations from $N(\delta_{II}, 1)$ are drawn until the associated Z_i falls outside the control limits. The corresponding run length equals i. The calculations are made for $\delta_{II} = 0, 0.1, 0.2, 0.3, 0.4$. The entire procedure is repeated for $R = 200,000$ simulation runs. The $AARL$ is computed by averaging over all $200,000$ run lengths and the results are presented in Table 2.6.

Phase II results

First, consider the situation where the Phase I data are in control (first part of Table 2.6). We see that, for in-control data, Phase II control chart performance is similar across all estimators (i.e. they have similar in-control and out-of-control $AARL$s), except for the control chart based on the robust point estimator $M(\bar{X})$, which falls short. This confirms that it is unadvisable to use a robust point estimator: if there are no contaminations in Phase I, useful information is lost, resulting in a less powerful control chart. It is better to use a Phase I procedure that only trims Phase I observations that are considered unacceptable. When no data anomalies are found, the number of falsely deleted samples is limited.

When there are localized disturbances, the best Phase II performance is achieved by the screening estimators ($s\bar{\bar{X}}_{0.6}$, $sM(\bar{X})_{0.2}$, $sM(\bar{X})_{0.6}$, and $sM(\bar{X})_1$) and the control chart based on the robust point estimator $M(\bar{X})$. The EWMA control chart based on $\bar{\bar{X}}$ is less effective as this traditional estimator is influenced by outliers, which then also affect the resulting Phase II limits. The EWMA control chart based on CP does not work very well: this estimator deletes too many samples on the assumption that the disturbances will endure. Unfortunately, for $n = 5$, all EWMA control charts are $AARL$-biased; this means that the $AARL$ is larger for out-of-control data than for in-control data. This bias disappears partly for $n = 10$, see Zwetsloot et al. (2014). Furthermore, none of the Phase II control charts performs well when there are diffuse disturbances because these charts are typically designed to detect sample shifts instead of individual outliers.

In the presence of a single step shift, CP performs best followed by the Phase I EWMA charts, ($s\bar{\bar{X}}_{0.6}$,$sM(\bar{X})_{0.2}$, and $sM(\bar{X})_{0.6}$). Note that the Phase I Shewhart chart ($sM(\bar{X})_1$) is outperformed by the EWMA and CP methods because the Shewhart chart has no memory and does not make use of the time sequence of the observations.

Finally, in the case of multiple step shifts, the CP method does not work well at all: the resulting control chart has a relatively low $AARL$ for the acceptable situation and high $AARL$ when the process is out of control, which is not desirable. This is as expected since the CP method is designed for single step shifts. In the multiple step shift scenario, the Phase I EWMA charts ($s\bar{\bar{X}}_{0.6}$, $sM(\bar{X})_{0.2}$, and $sM(\bar{X})_{0.6}$) perform best.

To summarize, the type of disturbance and estimation method used in Phase I strongly determine the performance of the Phase II EWMA control chart. We recommend the Shewhart chart in Phase I for scattered observations, the change point

Phase I		$AARL$				
		In-control	Out-of-control			
			δ_{II}			
Scenario	$\hat{\mu}_0$		0.1	0.2	0.3	0.4
In-control	$\bar{\bar{X}}$	374	210	61	22	12
	$M(\bar{X})$	333	214	73	25	12
	CP	370	210	63	23	12
	$s\bar{\bar{X}}_{0.6}$	370	212	62	23	12
	$sM(\bar{X})_{0.2}$	369	210	62	23	12
	$sM(\bar{X})_{0.6}$	367	211	63	23	12
	$sM(\bar{X})_1$	365	212	63	23	12
Localized	$\bar{\bar{X}}$	215	338	213	72	25
($\delta_I = 1$)	$M(\bar{X})$	273	294	153	52	20
	CP	217	325	207	75	27
	$s\bar{\bar{X}}_{0.6}$	260	319	169	55	21
	$sM(\bar{X})_{0.2}$	252	324	177	58	21
	$sM(\bar{X})_{0.6}$	271	311	158	51	19
	$sM(\bar{X})_1$	279	308	152	48	19
Diffuse	$\bar{\bar{X}}$	287	517	296	82	26
($\delta_I = 1$)	$M(\bar{X})$	297	458	296	97	30
	CP	288	510	297	84	28
	$s\bar{\bar{X}}_{0.6}$	291	506	296	83	27
	$sM(\bar{X})_{0.2}$	289	509	295	83	27
	$sM(\bar{X})_{0.6}$	292	507	294	84	27
	$sM(\bar{X})_1$	293	505	294	82	27
Single step	$\bar{\bar{X}}$	210	371	210	61	22
($\delta_I = 1$)	$M(\bar{X})$	278	302	148	48	18
	CP	349	209	69	25	13
	$s\bar{\bar{X}}_{0.6}$	312	298	119	37	16
	$sM(\bar{X})_{0.2}$	306	302	125	39	16
	$sM(\bar{X})_{0.6}$	317	284	112	35	15
	$sM(\bar{X})_1$	282	321	145	44	18
Multiple steps	$\bar{\bar{X}}$	220	266	193	97	39
($\delta_I = 1$)	$M(\bar{X})$	257	264	156	67	28
	CP	247	230	143	77	43
	$s\bar{\bar{X}}_{0.6}$	287	264	137	57	24
	$sM(\bar{X})_{0.2}$	295	262	129	52	22
	$sM(\bar{X})_{0.6}$	297	257	127	51	22
	$sM(\bar{X})_1$	268	269	154	66	27

Table 2.6: *$AARL$ values for the EWMA control chart for location based on various location estimators*

method whenever single step changes are likely, and the EWMA chart with λ_I set around 0.6 when there are multiple step changes, or when there is uncertainty about the type, the length, or the magnitude of the disturbance.

2.7 Conclusion

In this chapter, we have considered several Phase I estimation methods for situations where scattered as well as sustained shifts might be present. We have studied the effectiveness of a Phase I exploratory data analysis in terms of the accuracy of the resulting estimates and the proportion of successfully identified unacceptable samples. Moreover, we have investigated the impact of data contaminations and the estimators used in Phase I on the performance of the Phase II EWMA control chart.

We have shown that data anomalies can have a huge impact on the quality of the Phase I analysis as well as on the power of the resulting Phase II EWMA control chart. There is considerable difference in the performance of the Phase I estimation methods. We have the following recommendations:

First, we recommend the use of a change point or SPC-based method instead of a single point estimator to arrive at a parameter estimate for the Phase II limits. Such methods make it possible to perform a Phase I data analysis and 'learn from the data' before any monitoring takes place. A learning stage will improve the performance of the Phase II monitoring.

Second, we have seen that most methods work well in one specific situation: when there might be scattered contaminations, the Shewhart chart works best but, when there are sustained shifts, the use of an EWMA chart or change point method is more appropriate. These methods take into account the time sequence in the samples. When multiple step changes are likely or the type of disturbance is unknown, we recommend the Phase I EWMA chart. The smoothing constant of the EWMA chart, λ_I, should ideally be set at around 0.6. A low value overemphasizes the detection of small shifts while a high value approaches the performance of the Phase I Shewhart chart.

The methods based on EWMA charts not only work well in the multiple step scenario but also in the other contamination scenarios. These estimators provide near-best estimates of the location in the presence of any pattern of Phase I contaminations.

Finally, we recommend the use of a two-step procedure, namely a robust estimator to estimate the location and construct the Phase I control chart, and an efficient estimator for post-screening estimation. The use of a robust estimator for the Phase I chart ensures that Phase I limits are not too sensitive to any disturbances, limiting the incorrect deletion of clean data, while the use of an efficient estimator subsequent to screening ensures that final estimates are efficient under stability as well.

3. Robust Estimators for Dispersion

A Phase I estimator should be efficient under in-control data and robust against contaminations in the data. In this chapter, we focus on estimating the process dispersion and consider the situation when the data may contain contaminated samples. To deal with contaminations, we propose a robust estimation method and compare this new method to various existing methods. This chapter is based on Zwetsloot et al. (2015a).

3.1 Introduction

In Chapter 2 we showed that robust estimation methods for the location can be very useful if data contaminations may be present. Apart from an estimate for the location parameter μ_0, we also need an estimate for the dispersion parameter σ_0. Within control charting an estimate of σ_0 is needed to compute the control limits for the EWMA chart for location. Next to monitoring the location, sometimes it is of interest to detect changes in the degree of variation in the data. For these so-called control chart for dispersion an estimate of σ_0 is also needed.

The motivation for this chapter, like Chapter 2, is the possibility of contaminations in the initial (Phase I) data. We focus on estimation of the dispersion. As discussed in Section 1.4 we use robust estimators to deal with contamination. We show that most existing (robust) estimation methods are either efficient or robust against either scattered disturbances or sustained shifts. Furthermore, we develop an estimation method based on EWMA charting, which is efficient under in-control data and robust to both scattered and sustained contaminations in Phase I.

This chapter is organized as follows. The EWMA control chart for dispersion is described in Section 3.2. The new estimation technique and some competing methods are described in Section 3.3. Section 3.4 compares their performance in terms of efficiency for uncontaminated and various contaminated data sets. In Section 3.5, we

study the effect of using robust estimation methods on the performance of the EWMA chart for dispersion used in Phase II. Section 3.6 offers some concluding remarks.

3.2 The EWMA control chart for dispersion

Considerable research has been published on the EWMA chart for monitoring the process dispersion. Recall that for monitoring the location we used the sample average to plot the EWMA statistic (see Equation 2.1). For monitoring the dispersion a choice for the dispersion measure should be made. Various dispersion metrics have been proposed in the literature. For example, Knoth (2015) studied the EWMA chart based on the sample variance (S^2). Another possibility is the EWMA chart based on the sample standard deviation (S) (e.g., see Ng and Case, 1989). The EWMA chart based on a logarithmic transformation of S^2 or S, was developed following an insight from Box et al. (1978). This chart is the most prevailing chart in the literature, see Crowder and Hamilton (1992), Shu and Jiang (2008), and Maravelakis and Castagliola (2009), amongst others.

Knoth (2010) studied these competing statistics and compared EWMA dispersion charts based on R (the range), S^2, S, and $\log(S^2)$ and concluded that "the best performance in terms of the average run length profile is given by the S^2 and S EWMA control charts". In this chapter, we therefore consider an EWMA chart for dispersion based on S. We use S rather than S^2 because in practice the process dispersion is most often evaluated in terms of S. In Chapter 5 we will extend the comparison performed by Knoth (2010).

To monitor the dispersion, the EWMA statistic is defined as

$$Z_i = (1-\lambda)Z_{i-1} + \lambda S_i,$$

with S_i the sample standard deviation of sample i. Again λ is a smoothing constant satisfying $0 < \lambda \leq 1$. We set $W_0 = E[S_i] = c_4(n)\sigma_0$, were $c_4(n)$ is the bias correcting coefficient

$$c_4(n) = \sqrt{\frac{2}{n-1}}\frac{\Gamma(n/2)}{\Gamma((n-1)/2)}. \tag{3.1}$$

Under the assumption of independent and normally distributed observations, the mean and variance of Z_i are equal to $E[Z_i] = c_4(n)\sigma_0$ and $V[Z_i] = \sigma_0^2(1-c_4^2(n))\frac{\lambda}{2-\lambda}(1-(1-\lambda)^{2i})$, respectively.

When monitoring dispersion, an increase in the dispersion indicates some special cause of variation that should be detected and removed, while a decrease in dispersion indicates a process improvement. As we are most interested in detecting increases, we use an one-sided EWMA control chart. Therefore, we reset the EWMA

statistic to its expected value whenever it drops below its expected value, i.e.

$$Z_i = \max[\ (1-\lambda)Z_{i-1} + \lambda S_i\ ,\ c_4(n)\sigma_0\ ]. \tag{3.2}$$

Resetting Z_i to the target instead of using the natural boundary of zero decreases the inertial problems. This occurs when the EWMA statistics is close to zero while an upward shift will happen. See Woodall and Mahmoud (2005) for a discussion on inertial properties of EWMA charts.

The EWMA chart gives an out-of-control signal whenever Z_i exceeds the upper control limit UCL_i, where

$$UCL_i = c_4(n)\sigma_0 + L\sigma_0\sqrt{(1-c_4(n)^2)}\sqrt{\frac{\lambda}{2-\lambda}}\sqrt{1-(1-\lambda)^{2i}}. \tag{3.3}$$

Here, L is a positive coefficient which, together with λ, determines the in-control performance of the chart. We use the so called time-varying control limits to enhance the charts sensitivity to shifts in the first samples (Abbasi, 2010). As σ_0 is often unknown it has to be replaced by an estimate from Phase I.

3.3 Dispersion estimation methods

In this section, we describe various estimation methods that can be used when contaminations may be present in the data. We consider efficient and robust point estimators, a change point method, and we present the new estimation method based on EWMA control charts in Phase I.

Let X_{ij}, with $j = 1, 2, ..., n$ and $i = 1, 2, ..., m$, be the Phase I observations of the process characteristic. Assume X_{ij} to be independent $N(\mu, \sigma^2)$ distributed with mean $\mu = \mu_0$ and standard deviation $\sigma = \sigma_0$ if the process is in control. Throughout this chapter we set $m = 50$ and $n = 5$.

Point estimators

Traditionally, σ_0 is estimated with the pooled sample standard deviation:

$$S_p = \sqrt{\frac{1}{m}\sum_{i=1}^{m} S_i^2}. \tag{3.4}$$

We use the unbiased estimator $\tilde{S}_p = S_p/c_4(v+1)$, where $v = m(n-1)$ and c_4 as defined in Equation (3.1). Pooling that sample standard deviations provides lower values of MSE (mean squared error) than averaging them (Mahmoud et al., 2010; Vardeman, 1999). The pooled sample standard deviation provides a basis for comparison as it is the most efficient unbiased estimator under uncontaminated, independent, and normally distributed data.

Next, we include two robust point estimators in our comparison. The first is based on the sample interquartile ranges as defined in Equation (2.4). We consider a trimmed version of the mean of the sample IQRs, as proposed by Rocke (1989);

$$\overline{IQR}_{\alpha} = \frac{1}{m - 2\lceil m\alpha \rceil} \sum_{o=\lceil m\alpha \rceil + 1}^{m-\lceil m\alpha \rceil} IQR_{(o)},$$

where $IQR_{(o)}$ denotes the o-th ordered value of the sample IQRs. We take $\alpha = 20\%$. An unbiased estimate of σ_0 is given by dividing $\overline{IQR}_{20}$ by 0.9261 (obtained through 100,000 Monte Carlo simulations).

The second robust point estimator we consider is the estimator proposed by Tatum (1997), see Section 2.4. As in Tatum (1997), this estimator is denoted by $D7$ throughout this chapter.

Change point method

Change point methods are designed to detect structural changes in the process parameters. They make use of the log likelihood of the observations. Sullivan and Woodall (1996) showed that change point methods outperform the Shewhart chart in detecting sustained shifts in Phase I. We apply a modified version of the estimator they proposed. This method was already discussed in Section 2.3. In Equation (2.3) we defined the likelihood ratio statistic, which can be used to test for a change in both the location and the dispersion parameter. Hence, the change point method, described in Section 2.3, is also relevant as estimation method in this chapter. We apply the same method and use $UCL_{CP} = 5.92$ to ensure a false alarm rate of 1 percent. Furthermore, we use the remaining samples after screening to compute the pooled standard deviation (instead of the overall samples mean). The estimator is denoted by CP.

The proposed estimation method

In this section, we propose an estimation method for the dispersion based on EWMA charting in Phase I. The new method provides a robust estimate of the dispersion when it is unknown what type of contaminations are present in Phase I. The EWMA chart can be viewed as a compromise between the Shewhart chart and methods with a memory like the CUSUM chart and the change point method. Recall that this insight comes from Hunter (1986), as discussed in Section 2.1. The proposed estimation method consists of the following steps:

1. *Use all observations in Phase I and compute an initial (robust) estimate of the dispersion*. This estimate is denoted by $\widehat{\sigma}_I$.

2. *Set up a* EWMA *chart for the Phase I data* according to Equations (3.2) and (3.3) using $\sigma_0 = \widehat{\sigma}_I$, $Z_0 = c_4(n)\widehat{\sigma}_I$, $L = L_I$, and $\lambda = \lambda_I$.

3. *Delete from Phase I all samples for which the corresponding* EWMA *statistic gives an out-of-control signal.*

4. *Compute an efficient unbiased estimator of the standard deviation, $\tilde{S}_p$, based on the remaining samples.*

The resulting estimator is denoted by $s\widehat{\sigma}_{I,\lambda_I}$, where '$s$' indicates that we use a screening Phase I chart, $\widehat{\sigma}_I$ stands for the initial dispersion estimator chosen in step 1, and the subscript λ_I denotes the value selected for the smoothing constant in step 2. To operationalize this screening estimator, we need to select an estimator for $\widehat{\sigma}_I$ and values for λ_I and L_I.

In step 1, we select the efficient estimator $\tilde{S}_p$ and the robust estimator $\overline{IQR}_{20}$ as initial estimator. This provides a comparison with an efficient estimator when no contaminations are present and with a robust estimator when contaminations occur.

In step 2, small values of λ_I enable quick detection of sustained shifts, while larger values of λ_I enable quick detection of outliers more effectively. To assess the trade-off between high and low values for λ_I, we set λ_I equal to $0.3,\ 0.5$, and 1. For $\lambda_I = 1$, the chart is equivalent to the Phase I Shewhart chart. Estimation methods based on Shewhart chart in Phase I were also studied by Schoonhoven et al. (2011b) and Schoonhoven and Does (2012).

We obtained values for L_I by setting the false alarm rate in Phase I at 1 percent, thereby following Chakraborti et al. (2009). Table 3.1 gives an overview of the screening estimators considered and the corresponding values of L_I (obtained through $100,000$ Monte Carlo simulations).

Estimator	Description	L_I
$\tilde{S}_p$	Pooled standard deviation	n.a.
$\overline{IQR}_{20}$	Trimmed interquartile range	n.a.
$D7$	Tatum	n.a.
CP	Change point estimator	n.a.
$sS_{0.5}$	Screening estimator with $\widehat{\sigma}_I = S_p$ and $\lambda_I = 0.5$	2.553
$sIQR_{0.3}$	Screening estimator with $\widehat{\sigma}_I = \overline{IQR}_{20}$ and $\lambda_I = 0.3$	2.970
$sIQR_{0.5}$	Screening estimator with $\widehat{\sigma}_I = \overline{IQR}_{20}$ and $\lambda_I = 0.5$	2.900
$sIQR_1$	Screening estimator with $\widehat{\sigma}_I = \overline{IQR}_{20}$ and $\lambda_I = 1$	2.755

Table 3.1: *Phase I dispersion estimators*

3.4 Phase I comparison

Next, we evaluate the performance of the proposed dispersion estimation methods when the Phase I data are in control as well as when the Phase I data contain contaminations. Recall that the Phase I data are $N(\mu_0, \sigma_0^2)$ distributed if the process is in control. Without loss of generality, we set $\mu_0 = 0$ and $\sigma_0 = 1$.

Contamination scenarios

Many different contamination scenarios are studied in the literature. In this chapter, like in Chapter 2, we distinguish between *scattered* and *sustained* special causes of variation. We evaluate two scattered scenarios - localized and diffuse - and two sustained shift scenarios - single and multiple step shifts-. In this chapter, we use the same four scenarios also considered in Chapter 2 (in Section 2.5). Here we apply them to the dispersion parameter:

1. A model for *localized variance disturbances* in which all observations in a sample have a 90% probability of being drawn from the $N(0,1)$ distribution and a 10% probability of being drawn from the $N(0,\gamma_I^2)$ distribution, with $\gamma_I = 1, 1.5, ..., 3.5, 4$.

2. A model for *diffuse variance disturbances* in which each observation is drawn from the $N(0,1)$ distribution and has a 10% probability of having a multiple of a χ_1^2 variable added to it, with multiplier κ_I, with $\kappa_I = 0, 0.5, 1, ..., 2.5, 3$.

3. A model for a *single step shift* in the variance. All observations in the last 5 Phase I samples are drawn from the $N(0,\gamma_I^2)$ distribution, with $\gamma_I = 1, 1.5, ..., 3.5, 4$.

4. A model for *multiple step shifts* in the variance. At each time point, the sample has a probability op p of being the first of five consecutive samples drawn from the $N(0,\gamma_I^2)$ distribution, with $\gamma_I = 1, 1.5, ..., 3.5, 4$. After any such step shift, each sample again has a probability of p of being the start of another step shift. Phase I consists of 50 samples. If say sample 48 shifts, then only 3 samples (48, 49, and 50) are drawn from the $N(0,\gamma_I^2)$ distribution, instead of 5. To maintain the 10% (expected) contamination rate of models 1-3, we set $p = 0.023$.

Note that for $\gamma_I = 1$ or $\kappa_I = 0$ all Phase I data come from the in-control distribution and hence no contaminations are present. In Zwetsloot et al. (2015a) we also considered these scenarios with an (expected) contamination rate equal to 5%. Table 3.2 presents an overview of the contamination scenarios.

Performance measures and simulation procedure

One of the requirements of Phase I is to deliver an accurate parameter estimate of σ_0, even if Phase I contains contaminated observations. In order to evaluate the accuracy of the dispersion estimators, we determine their mean squared error (MSE), which is computed as

$$MSE = \frac{1}{R}\sum_{r=1}^{R}\left(\frac{\widehat{\sigma}_0^r - \sigma_0}{\sigma_0}\right)^2 = \frac{1}{R}\sum_{r=1}^{R}(\widehat{\sigma}_0^r - 1)^2.$$

Here $\widehat{\sigma}_0^r$ denotes one of the proposed estimators presented in Table 3.1, calculated in the rth simulation run, and R is the total number of simulation runs.

Contamination scenarios	Description
In-control	All observations from $N(0,1)$
Localized disturbances	$90\% - 10\%$ random mixture of samples from $N(0,1)$ and $N(0,\gamma_I^2)$
Diffuse disturbances	$90\% - 10\%$ random mixture of observations from $N(0,1)$ and $N(0,1) + \kappa_I \chi_1^2$
Single step shift	Samples $1-45$ are $N(0,1)$ and samples $46-50$ are $N(0,\gamma_I^2)$
Multiple step shifts	Shifts of length 5 from $N(0,\gamma_I^2)$ occurring with probability 0.023

Table 3.2: *Phase I contamination scenarios affecting the dispersion parameter*

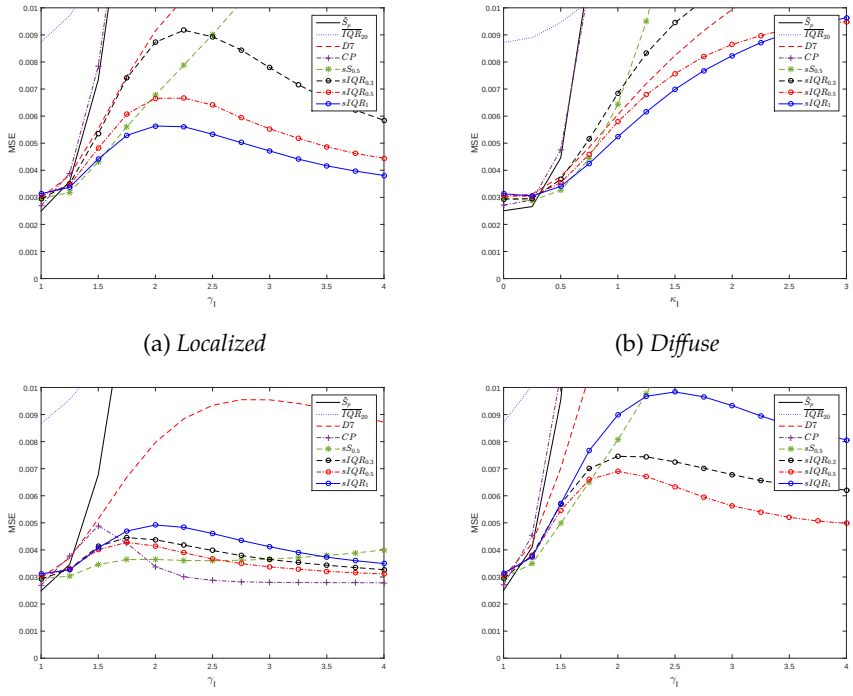
The proposed estimators are also evaluated with two additional quality characteristics: the true alarm percentage (TAP) and the false alarm percentage (FAP). These additional performance measures reflect the ability of the screening estimators to detect unacceptable observations without triggering false alarms for acceptable observations. The TAP and FAP are calculated as in Equations (2.5) and (2.6).

A Monte Carlo simulation study was conducted. We drew R Phase I data consisting of $m = 50$ samples of size $n = 5$, for each of the four contamination scenarios, and each value of γ_I or κ_I. We consider $\gamma_I = 1.25, 1.5, ..., 4$ and $\kappa_I = 0.25, .0.5, ..., 3$. The proposed eight dispersion estimators, presented in Table 3.1, were calculated for each simulation run, and the three performance measures were computed based on the R runs. The relative simulation error is defined as the standard deviation of the MSE expressed as a percentage of the MSE. We set the value of R equal to $100,000$, so that this error was never larger than 0.5%. The MSE results are presented in Figure 3.1 and the TAP and FAP metrics are shown in Table 3.3.

Phase I results

First consider the situation when the Phase I data are in control. In Table 3.3, the column corresponding to $\gamma_I = 1$ or $\kappa_I = 0$ shows the false alarm percentage if the Phase I data are in control. This percentage was set equal to 1% for all estimators. In Figure 3.1, the y-intercept in each subplot represents the MSE level of the estimators based on uncontaminated Phase I data. As expected, the pooled sample standard deviation $\tilde{S}_p$ shows the smallest MSE level followed by the CP estimator, the screening estimators, and $D7$. The robust point estimator $\overline{IQR}_{20}$ has a very large MSE level for uncontaminated data.

Furthermore, for uncontaminated Phase I data, the four screening estimators show an approximately equal MSE level. This implies that, from the perspective of efficiency under uncontaminated Phase I data, it does not matter whether we use an efficient ($sS_{0.5}$) or robust ($sIQR_{0.5}$) initial estimator for $\hat{\sigma}_I$. Furthermore, it also does not matter, from the perspective of efficiency under uncontaminated Phase I data,



(a) *Localized* (b) *Diffuse*

(c) *Single step* (d) *Multiple step*

Figure 3.1: *MSE of dispersion estimators when various types of contaminations are present in Phase I*

which λ_I we use.

Next, we consider the situation when contaminations are present in the Phase I data ($\gamma_I > 1$ or $\kappa_I > 0$). Overall, the EWMA-based methods perform reasonably well for any type of contamination, while the point estimator $\tilde{S}_p$ is most sensitive to contaminations.

When localized or diffuse shifts are present in Phase I, the screening estimators based on a robust initial estimator show the lowest MSE over all shift sizes (see Figures 4.1a and 4.1b). The screening estimators show diverging MSE levels, with $sIQR_1$ and $sIQR_{0.5}$ having the lowest MSE levels. The FAP of the estimator $sIQR_{0.3}$ is the highest among the screening estimators. This is due to the low value of λ_I, suggesting that it is undesirable to set λ_I too low.

Note that the estimation methods for the dispersion (considered in this chapter) show better ability to deal with diffuse contaminations than the estimation methods for location (considered in Chapter 2). If we compare Figure 2.1b to Figure 4.1b we see that all the considered estimators for the location are unable to deal with diffuse contaminations. However, the considered estimators for dispersion based on screen-

		TAP			FAP			
		γ_I			γ_I			
Scenario	$\widehat{\sigma}_0$	2	3	4	1	2	3	4
Localized	CP	10.7	26.4	30.0	1.0	6.4	19.0	24.5
	$sS_{0.5}$	33.0	57.7	68.4	1.0	0.6	0.8	1.1
	$sIQR_{0.3}$	30.1	68.0	86.0	1.0	2.3	7.1	12.7
	$sIQR_{0.5}$	36.4	73.0	88.4	1.0	1.3	3.7	6.8
	$sIQR_1$	39.9	74.8	89.1	1.0	0.5	0.4	0.4
		κ_I			κ_I			
		1	2	3	0	1	2	3
Diffuse	CP	7.9	21.6	25.9	1.0	6.9	20.1	24.9
	$sS_{0.5}$	7.4	14.6	17.9	1.0	2.6	5.0	6.1
	$sIQR_{0.3}$	9.4	28.8	44.7	1.0	4.8	15.7	26.8
	$sIQR_{0.5}$	9.6	26.2	39.5	1.0	4.0	11.4	18.7
	$sIQR_1$	9.2	22.7	32.7	1.0	3.2	7.1	10.1
		γ_I			γ_I			
		2	3	4	1	2	3	4
Single step	CP	90.0	98.6	99.5	0.9	1.1	0.2	0.1
	$sS_{0.5}$	58.6	85.0	91.4	1.0	0.1	0.0	0.0
	$sIQR_{0.3}$	62.1	89.6	95.9	1.0	0.4	0.2	0.2
	$sIQR_{0.5}$	59.6	90.6	96.9	1.0	0.4	0.3	0.3
	$sIQR_1$	41.1	76.2	89.9	1.0	0.5	0.4	0.4
Multiple steps	CP	39.5	53.3	53.7	1.0	11.5	17.0	17.4
	$sS_{0.5}$	45.3	69.1	76.5	1.0	0.6	0.8	0.9
	$sIQR_{0.3}$	54.8	85.9	94.0	1.0	2.1	4.6	6.4
	$sIQR_{0.5}$	51.7	85.9	94.7	1.0	1.0	1.9	2.8
	$sIQR_1$	35.4	69.6	85.2	1.0	0.6	0.5	0.5

Table 3.3: *True Alarm Percentage (TAP) and False Alarm Percentage (FAP)*

ing can (partly) deal with diffuse contaminations. Recall that diffuse contaminations effect single observations within a sample. This will increase the value of the sample standard deviation estimator, hence the estimation methods that screen for large dispersion in samples, will occasionally pick up these diffuse contaminations. This will result in deleting these samples, including the observations which are actually in control. Table 3.3 confirms this explanation and shows that for the diffuse scenario the screening estimators delete many in-control observations: the FAP is rather high.

When a single step shift is present in Phase I (Figure 4.1c), the estimator CP shows a low MSE level, as was to be expected, as CP is specifically designed for a single step shift. However, quite surprisingly, for small shift sizes ($1 < \gamma_I < 2$), its MSE level is slightly higher than the MSE level of the screening estimators. The TAP and

FAP values show the same pattern: CP has the highest TAP and it has the lowest FAP for single step shifts in Phase I.

When multiple step shifts are present in Phase I (Figure 4.1d), the screening estimators based on $\overline{IQR}_{20}$ have the lowest MSE. The estimator based on $\lambda_I = 0.5$, shows the best MSE performance. It is able to detect most of the contaminated samples with TAP values up to 94.7%, and incorrectly deletes no more than 3% of the observations.

To compare the performance of the proposed estimators across the various contamination schemes, we compute the Relative Mean Squared Error ($RMSE$) of the estimators. Recall from Section 2.5, that the $RMSE$ of an estimator, for a specific type of data contamination and severity γ_I or κ_I, is defined as the percentage increase in the MSE of the estimator relative to the MSE of the estimator with the lowest MSE. For each data scenario and each estimator, we obtain the RSME of the estimator for all considered levels of γ_I or κ_I. We present the maximum RSME for each estimator in Table 3.4.

Surprisingly, the estimator $sIQR_{0.5}$ has for *all* contamination scenarios, the lowest maximum $RMSE$ (i.e. has an MSE which overall is closest to the optimal MSE for all shift sizes). Even the CP estimator has a larger $RMSE$ for the single step shift scenario.

Overall, the best estimator is the screening estimator based on $\overline{IQR}_{20}$ and $\lambda_I = 0.5$. Irrespective of the contamination scenario in Phase I, it always has a MSE level that is within 30% of the optimal estimator for that contamination scenario and shift size.

Throughout this chapter, we have assumed that 10% of the observations in Phase I are contaminated. In Zwetsloot et al. (2015a) we also studied the situation where 5% of the observations in Phase I are contaminated. The conclusions for this scenario are comparable and can be found in the paper.

	Phase I estimators $\widehat{\sigma}_0$							
Scenario	$\tilde{S}_p$	$\overline{IQR}_{20}$	$D7$	CP	$sS_{0.5}$	$sIQR_{0.3}$	$sIQR_{0.5}$	$sIQR_1$
Localized	9955	474	181	9966	397	68	**23**	26
Diffuse	9559	248	28	7772	814	35	**23**	25
Single	12707	575	240	41	44	39	**30**	61
Multiple	9553	690	343	8659	564	25	**23**	66
All	12707	690	343	9966	814	68	**30**	66

Table 3.4: *Maximum Relative Mean Squared Error ($RMSE$) and, in bold, the estimator with the lowest maximum $RMSE$ for the respective contamination scenario.*

3.5 Phase II performance

Phase I estimation methods are used to design the Phase II control chart. In this section, we consider the effect of estimating σ_0 on the EWMA control chart for dispersion in Phase II, when the Phase I data may or may not be contaminated.

Let Y_{ij}, with $j = 1, .., n$ and $i = 1, 2, ..,$ be the Phase II data which are independent and $N(\mu_0, \gamma_{II}^2 \sigma_0^2)$ distributed, where γ_{II} is a constant. If $\gamma_{II} = 1$, the data are in control and if $\gamma_{II} \neq 1$ the data are out of control. We only consider $\gamma_{II} \geq 1$ as we study the one-sided EWMA dispersion control chart. Throughout this chapter, we set $\mu_0 = 0$ and $\sigma_0 = 1$ without loss of generality.

Design of the Phase II EWMA control chart

The Phase II EWMA control chart consists of the EWMA statistic as defined in Equation (3.2) together with the upper control limit as defined in Equation (3.3). For σ_0, we consider the eight estimation methods presented in Table 3.1, which results in eight Phase II EWMA control charts. Furthermore, λ is set equal to $\lambda_{II} = 0.3$. This value differs from the value of λ_I (0.5) chosen in Phase I, as Phase I is used for exploratory purposes. The values of L is denoted by L_{II} and presented in Table 3.5. These values were determined such that all EWMA control charts have an approximate in-control $AARL$ performance of 200, whereby we followed the design procedure originally proposed by Jones (2002).

	$\widehat{\sigma}_0$							
	$\tilde{S}_p$	$\overline{IQR}_{20}$	$D7$	CP	$sS_{0.5}$	$sIQR_{0.3}$	$sIQR_{0.5}$	$sIQR_1$
L_{II}	2.607	2.210	2.570	2.580	2.680	2.650	2.660	2.677

Table 3.5: *L_{II} for the Phase II EWMA control chart based on $\widehat{\sigma}_0$*

Performance measures and simulation procedure

Comparable to Chapter 2, we evaluated the performance of the charts in Phase II by studying the average of the conditional ARLs: the $AARL$. The $AARL$ reflects the performance of the charts averaged over all values of the Phase I parameter estimates. Solely, studying the $AARL$ does not reflect sampling variation in the performance of the chart. Therefore, we report additionally the unconditional 10th, 50th, and 90th percentiles of the run length distribution, to give a better idea of how the EWMA chart performance varies according to the different values of the parameter estimates.

A Monte Carlo simulation study was conducted to evaluate the $AARL$ and the percentiles of the unconditional run length distribution of the Phase II EWMA dispersion control chart based on the eight Phase I dispersion estimators presented in Table 3.1. We used the following simulation procedure: first $m = 50$ samples of size $n = 5$

are drawn from $N(0,1)$. We calculated $\widehat{\sigma}_0$ using the eight proposed estimators. These estimates were used to set up eight Phase II EWMA control charts according to Equations (3.2) and (3.3). Next, observations from $N(0,\gamma_{II}^2)$ are drawn until the associated Z_i falls above the control limit. The corresponding run length equals $i-1$. The calculations are made for $\gamma_{II} = 1, 1.1, 1.2, 1.4$. The entire procedure was repeated for R up to $100,000$ simulation runs, such that the relative simulation error never exceeded 0.1%. The $AARL$ was computed by averaging over all obtained run lengths and the percentiles are taken by sorting the $100,000$ run lengths and selecting the $10,000th$, the average of the $50,000th$ and $50,001st$, and the $90,000th$ run lengths.

This whole procedure was repeated for each of the four Phase I contamination scenarios, as presented in Table 3.2, where we set $\gamma_I = 2.5$ or $\kappa_I = 1.5$. The results are presented in Table 3.6. The first part of the table shows the performance of the EWMA control charts when the Phase I data are in control, followed by the results when Phase I contains out-of-control observations as defined in the four Phase I contamination scenarios.

For computational convenience and speed, we truncated the simulation at a run length equal to $30,000$ and set a run length greater than $30,000$ equal to this value. These values are therefore an underestimate of the actual $AARL$ and percentiles of the run length distribution.

Phase II results

First, consider the situation where the Phase I data are uncontaminated (first part of Table 3.6). A general observation is that under in-control data the EWMA control charts show similar performance across all estimators (i.e. they have in-control ARLs of around 200 and they show similar out-of-control ARLs and percentiles). Which is logical as the charts are designed for an $AARL$ equal to 200. One exception is the EWMA control chart based on $\overline{IQR}_{20}$, which shows smaller percentiles and thus gives more frequent false alarms.

Next, consider the situation where the Phase I data are contaminated. A general effect of contaminations in Phase I is an increased in-control ARL as well as an increased out-of-control ARL. This is of course undesirable. The control chart based on the traditional estimator $\tilde{S}_p$ is less effective if any type of contamination is present in Phase I. The control chart based on CP shows good performance for the single step shift scenario but does not work very well for the other contamination scenarios. The control chart based on $sIQR_1$ shows the best performance for localized and diffuse shifts. The estimator $sIQR_{0.5}$ has the second-best performance for these scenarios. If there are multiple step shifts present, the best performance is given by the control chart based on $sIQR_{0.5}$.

To summarize, the type of disturbance and estimation methods used in Phase I strongly determines the performance of the Phase II EWMA control chart. When it is unknown which type of contaminations are present in Phase I, we recommend the

Phase I	$AARL$ and percentiles of the unconditional run length distribution															
	In-control data				Out-of-control Phase II data											
	$\gamma_{II}=1$				$\gamma_{II}=1.1$				$\gamma_{II}=1.2$				$\gamma_{II}=1.4$			
$\widehat{\sigma}_0$	10th	50th	90th	$AARL$	10th	50th	90th	$AARL$	10th	50th	90th	$AARL$	10th	50th	90th	$AARL$
In-control																
$\tilde{S}_p$	10	86	467	201	3	23	100	42	1	9	36	15	0	3	11	5
$D7$	9	79	458	200	3	21	98	42	1	9	35	15	0	3	11	5
$\overline{IQR}_{20}$	3	36	348	207	1	12	79	39	0	6	29	13	0	2	9	4
CP	10	84	461	204	3	22	99	42	1	9	36	15	0	3	11	5
$sS_{0.5}$	10	82	469	204	3	22	99	42	1	9	35	15	0	3	11	5
$sIQR_{0.3}$	10	82	468	202	3	22	99	42	1	9	36	15	0	3	11	5
$sIQR_{0.5}$	10	82	471	204	3	22	99	42	1	9	35	15	0	3	11	5
$sIQR_1$	9	79	460	200	3	21	97	41	1	9	35	15	0	3	11	5
Localized ($\gamma_I=2.5$)																
$\tilde{S}_p$	203	7606	30000	13701	31	654	30000	5305	9	116	2650	1470	2	15	119	78
$\overline{IQR}_{20}$	8	106	1714	988	2	27	255	159	1	11	69	36	0	3	15	6
$D7$	25	277	3090	1454	6	54	378	202	2	18	94	44	0	5	18	8
CP	56	2356	30000	10674	12	268	30000	5158	4	60	4228	2225	1	10	148	263
$sS_{0.5}$	23	248	3206	1606	6	49	380	248	2	17	92	52	0	5	18	8
$sIQR_{0.3}$	16	190	2636	1389	5	40	325	213	2	14	82	46	0	4	17	7
$sIQR_{0.5}$	14	148	1541	852	4	34	224	123	2	13	63	31	0	4	15	6
$sIQR_1$	15	148	1332	708	4	34	202	103	2	13	59	27	0	4	14	6
Diffuse ($\kappa_I=1.5$)																
$\tilde{S}_p$	270	13775	30000	15775	40	1102	30000	8217	11	175	17513	3822	2	19	389	714
$\overline{IQR}_{20}$	6	78	1042	611	2	22	175	98	1	9	53	25	0	3	13	5
$D7$	23	235	1955	885	6	48	279	124	2	16	75	33	0	5	17	7
CP	74	2300	30000	10398	15	270	30000	5342	5	61	7561	2882	1	11	216	840
$sS_{0.5}$	39	480	6720	2684	9	79	690	417	3	24	142	81	1	6	24	11
$sIQR_{0.3}$	21	256	3048	1432	6	50	375	191	2	17	93	42	0	5	18	8
$sIQR_{0.5}$	19	216	2225	1057	5	45	299	144	2	16	78	35	0	4	17	7
$sIQR_1$	21	225	2155	1008	6	46	295	140	2	16	77	34	0	5	17	7
Single step ($\gamma_I=2.5$)																
$\tilde{S}_p$	462	9132	30000	13966	57	758	11502	3780	13	131	1254	685	2	16	83	39
$\overline{IQR}_{20}$	8	105	1468	836	3	27	228	127	1	11	65	30	0	3	14	6
$D7$	27	273	2283	1034	6	53	315	138	2	18	82	35	0	5	17	8
CP	10	80	450	200	3	21	97	42	1	9	35	15	0	3	11	5
$sS_{0.5}$	15	135	904	407	4	31	160	68	2	12	50	21	0	4	13	6
$sIQR_{0.3}$	13	122	865	405	4	29	153	67	2	11	48	21	0	4	13	6
$sIQR_{0.5}$	12	110	750	351	4	27	139	60	1	11	45	19	0	4	12	5
$sIQR_1$	15	145	1139	547	4	33	185	83	2	12	55	24	0	4	14	6
Multiple steps ($\gamma_I=2.5$)																
$\tilde{S}_p$	38	5501	30000	13464	10	490	30000	7684	4	90	16995	3718	1	12	369	525
$\overline{IQR}_{20}$	7	107	3282	1824	2	27	411	457	1	11	97	113	0	3	18	11
$D7$	21	265	8171	3028	5	53	800	787	2	18	156	202	0	5	24	19
CP	13	163	30000	7059	4	39	30000	5418	2	15	30000	4007	0	5	910	1593
$sS_{0.5}$	15	161	3049	1977	4	36	366	621	2	14	87	207	0	4	18	22
$sIQR_{0.3}$	12	124	1332	1019	4	30	202	232	2	12	59	71	0	4	14	11
$sIQR_{0.5}$	12	114	1079	841	3	28	174	196	1	11	53	58	0	4	13	9
$sIQR_1$	15	156	2482	1624	4	35	313	406	2	13	79	119	0	4	17	16

Table 3.6: *$AARL$ values and percentiles of the unconditional run length distribution of the EWMA control chart for dispersion, based on various dispersion estimation methods, when the data in Phase I are contaminated*

use of the EWMA chart in Phase I based on a robust initial estimator and $\lambda_I = 0.5$ ($sIQR_{0.5}$).

3.6 Conclusion

In this chapter, we have proposed a new Phase I estimation method for the dispersion of a process. This method is based on Phase I EWMA charting and provides an efficient estimator of the dispersion for in-control Phase I data and a robust estimate of the dispersion if contaminations are present in Phase I. We have compared this new method with several estimation methods from the literature, in terms of their accuracy (MSE) and the percentage of successfully identified samples (TAP and FAP). Moreover, we have investigated the impact of data contaminations in Phase I on the performance of the Phase II EWMA control chart based on the various dispersion estimators.

In this chapter we show that the existing Phase I estimation methods provide robust estimates for *specific* patterns of disturbances in Phase I. In particular, estimators based on Phase I Shewhart charts are robust primarily to outliers in Phase I and change point methods are robust to sustained shifts in Phase I. The new method, based on Phase I EWMA charting, shows MSE levels which are comparable to the MSE level of the Phase I Shewhart chart estimator if outliers are present and which are also comparable to the MSE level of the change point methods if sustained shifts are present. Thus, the proposed method provides a robust estimate for *any* of the considered patterns of contaminations in Phase I.

The choice of the smoothing constant for the Phase I EWMA chart is important as it influences robustness against the various patterns of contaminations. By studying the TAP and FAP, we have discovered that for small values of λ (0.3) the estimator deletes too many in-control samples from Phase I and for larger values of λ (1) the estimator does not identify sustained shifts in Phase I optimally. Therefore, we recommend estimating the process dispersion by means of a Phase I EWMA chart with a smoothing constant of around 0.5.

Furthermore, we recommend the use of a two-step procedure, namely a robust estimator to estimate the dispersion and construct the Phase I chart, and an efficient estimator for post-screening estimation.

Phase II

Effect of Estimation on Performance

"In these matters the only certainty is that nothing is certain."

Pliny the Elder, Roman scholar & scientist

"Control chart performance is very sensitive to errors in estimating θ_0 [the process parameters]...methods for compensating for these effect remain to be developed."

Stoumbos et al. (2000, page 996)

"To pull oneself up by one's bootstraps."

The Surprising Adventures of Baron Munchausen

4. Designing EWMA Charts for Location

When in-control process parameters are estimated, the control chart performance in Phase II will depend on the Phase I data set used. In this chapter, we focus on monitoring the location. We show that it can be extremely difficult to lower the variation in the in-control performance sufficiently due to practical limitations on the amount of the Phase I data. We recommend an alternative design criterion and a procedure based on the bootstrap approach. This chapter is based on Saleh, Mahmoud, Jones-Farmer, Zwetsloot, and Woodall (2015a).

4.1 Introduction

Phase II control charts are designed for monitoring processes and detecting deviations from the in-control values of the process parameter(s). Because the true values of the in-control parameters are rarely known in practice, practitioners typically begin by collecting baseline information on the process. The performance of the EWMA control chart for location with estimated parameters was first investigated by Jones et al. (2001), who derived the run length distribution of the chart. Jones et al. (2001) studied the run length distribution conditioned on specific values of the parameter estimates and also studied the unconditional run length distribution averaged over all possible values of the parameter estimates. They assumed uncontaminated Phase I data. They showed that the EWMA chart performance deteriorates substantially when parameters are estimated, particularly with small amounts of Phase I data. Similar to Quesenberry (1993), Jones et al. (2001) made sample-size recommendations based on the increase in the rate of early false alarms of a chart with estimated parameters over one with known parameters. This approach resulted in recommendations that more Phase I data are required for EWMA charts with small smoothing constants. Smaller values of the smoothing constant are typically recommended for detecting sustained

shifts of smaller magnitude (Crowder, 1987; Lucas and Saccucci, 1990).

In Sections 2.6 and 3.5 we evaluated the performance of the EWMA chart using characteristics of the run length distribution averaged over all values of the parameter estimates. We used the $AARL$ metric. In Section 3.5 we additionally reported the unconditional 10th, 50th, and 90th percentiles of the run length distribution, which also gave an idea of how the EWMA chart performance varies across the different values of the parameter estimates. It is difficult, however, to use multiple percentiles to make recommendations on the amount of Phase I data for control charts with estimated parameters.

Our approach, in this chapter, is to use the standard deviation of the ARL, the $SDARL$, as a measure of the amount of sampling variation in control chart performance. This variation is often referred to as practitioner-to-practitioner variability. Recently, several authors have used the $SDARL$ as a metric for determining the necessary amount of Phase I data for control charts with estimated parameters (see, e.g. Jones and Steiner, 2012; Zhang et al., 2013, 2014; Lee et al., 2013; Faraz et al., 2015). These studies showed that impractically large amounts of Phase I data are needed for a practitioner to have confidence that his/her in-control ARL is near the desired value. The extent of this phenomenon was first recognized by Albers and Kallenberg (2004, 2005).

The findings of the studies accounting for sampling (or between-practitioner) variability imply the necessity of having an alternative technique for controlling the chart's performance. Recently, Jones and Steiner (2012) and Gandy and Kvaløy (2013) proposed a design procedure, based on the bootstrap, which guarantees, with a specified probability, a certain conditional performance for control charts. Their approach is to adjust the control limits such that p% of the in-control ARL values are at least a specified value; for example, at least 90% of the charts with a particular design would have in-control ARL values of 200 or more. The main objective of this approach is to limit the proportion of low in-control ARL values resulting from the use of insufficient amounts of Phase I data. Gandy and Kvaløy (2013) showed that even with the use of relatively small amounts of Phase I data, the out-of-control ARLs using this approach increase only slightly compared to the case when the standard design method is used.

In this chapter, we extend the work of Jones et al. (2001) by evaluating the performance of the EWMA chart with estimated parameters while considering the sampling variability using the standard deviation of the average run length ($SDARL$). We also study the effect of the smoothing constant on the sampling variability. Additionally, we design the EWMA chart for location using this bootstrap approach and investigate the effect of adjusting the control limits on the out-of-control performance of the chart.

In Section 4.2, we give an overview of the EWMA chart for location with estimated parameters and present the estimators used for the in-control process parameters. In Section 4.3, we highlight the importance of incorporating the effect of sampling vari-

ability -the practitioner-to-practitioner variability- when assessing the EWMA chart. In Section 4.4, we evaluate the EWMA chart in terms of the $AARL$, $SDARL$, and some percentiles of the ARL distribution. In Section 4.5, we investigate the in-control and out-of-control performance of the EWMA chart when the control limits are determined using the bootstrap approach. Finally, we give concluding remarks in Section 4.6.

4.2 The EWMA control chart for location

We observe $Y_{i1}, Y_{i2}, ..., Y_{in}$, $i = 1, 2, 3, ...$, independent random samples of size n at regular time intervals. For each observation it is assumed that it is an independent and normally distributed random variable with mean μ and standard deviation σ. The objective is to detect any change in μ from its in-control value μ_0. We further assume that the in-control process standard deviation value is σ_0.

The EWMA chart statistic at time i is defined as

$$Z_i = (1 - \lambda)Z_{i-1} + \lambda\bar{Y}_i, \tag{4.1}$$

where $\bar{Y}_i$ is the ith sample mean and $\lambda, 0 < \lambda \leq 1$, is a smoothing constant. The initial value Z_0 is set to be equal the estimate of the mean from the Phase I data. The EWMA chart signals when the statistic Z_i exceeds the control limits given by

$$\mu_0 \pm L\sqrt{\frac{\lambda}{n(2-\lambda)}[1 - (1 - \lambda)^{2i}]}\sigma_0 \tag{4.2}$$

where L is chosen to satisfy a specific in-control performance. The time-varying control limits in Equation (4.2) are the 'exact' limits for the EWMA chart. As i increases, the term $(1 - \lambda)^{2i}$ approaches zero and the limits in Equation (4.2) converge to the asymptotic limits given by

$$\mu_0 \pm L\sqrt{\frac{\lambda}{n(2-\lambda)}}\sigma_0. \tag{4.3}$$

Throughout this chapter, we consider the EWMA chart designed using the asymptotic limits defined in Equation (4.3).

Following a similar procedure to that of Jones et al. (2001), the chart statistic in Equation (4.1) can be rewritten as

$$Z_i^* = (1 - \lambda)Z_{i-1}^* + \lambda W_i,$$

where $Z_0^* = 0$ and W_i is the standardized sample mean defined as

$$W_i = \frac{\bar{Y}_i - \mu_0}{\sigma_0/\sqrt{n}}, \quad i = 1, 2, 3, ...,$$

for any target mean value μ_0 and standard deviation σ_0. If μ_0 and σ_0 are unknown, they are typically replaced with their corresponding estimators to give

$$\hat{W}_i = \frac{\bar{Y}_i - \hat{\mu}_0}{\hat{\sigma}_0/\sqrt{n}},$$

or equivalently

$$\hat{W}_i = \frac{1}{Q}\left(\nu_i + \tilde{\delta} - \frac{Z}{\sqrt{m}}\right), \tag{4.4}$$

where $Q = \hat{\sigma}_0/\sigma_0$ is the ratio of the estimated in-control standard deviation to the actual in-control standard deviation, $\nu_i = \sqrt{n}(\bar{Y}_i - (\mu_0 + \delta\sigma_0)/\sigma_0$ is the standardized Phase II sample mean with $\delta\sigma_0$ representing the mean shift, $\tilde{\delta} = \delta\sigma_0/(\sigma_0/\sqrt{n}) = \delta\sqrt{n}$ is the standardized mean shift, and $Z = \sqrt{mn}(\hat{\mu}_0 - \mu_0)/\sigma_0$ is the standardized difference between the actual in-control mean and the estimated in-control mean. If the process is in control, then $\tilde{\delta} = \delta = 0$. We assume, without loss of generality, that $\mu_0 = 0$ and $\sigma_0/\sqrt{n} = 1$ and, because of standardization, the control limits in Equation (4.3) become

$$\pm L\sqrt{\frac{\lambda}{2-\lambda}} \tag{4.5}$$

Throughout this chapter, we assume that the Phase I data is in control and uncontaminated, therefore we consider estimating the in-control process mean μ_0 by the overall sample mean defined as

$$\hat{\mu}_0 = \frac{\sum_{i=1}^{m} \bar{Y}_i}{m}, \tag{4.6}$$

where $\bar{Y}_i$ is the ith Phase I sample mean. The process standard deviation, $\hat{\sigma}_0$, is estimated by the pooled standard deviation defined as in Equation (3.4):

$$\hat{\sigma}_0 = S_p/c_4(v+1). \tag{4.7}$$

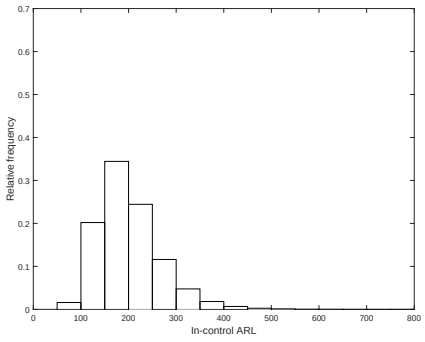
In Saleh et al. (2015a) we also consider other estimators for $\hat{\sigma}_0$ such as the sample range, the average over the sample standard deviation and S_p multiplied by $c_4(v+1)$ which is a biased estimator of σ_0. The reader is referred to the paper for the results related to these estimation methods.

4.3 The effect of sampling error

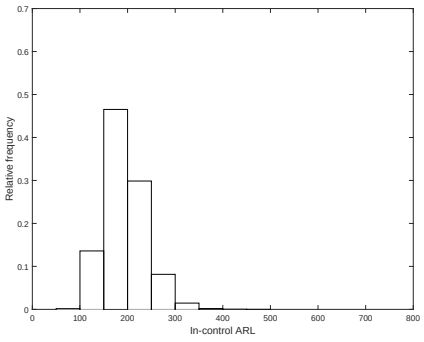
When the parameters are known, a control chart's ARL is a constant value; however, when the parameters are estimated, the ARL becomes a random variable due to the Phase I sampling. Control charts with estimated parameters have most often

been evaluated in terms of the average ARL ($AARL$). This measure was also used in Chapters 2 and 3. The use of the $AARL$, however, does not reflect other important properties of the ARL. Because the ARL distribution can be skewed, the mean of the distribution (AARL) may not give an accurate measure of the location. More importantly, the $AARL$ does not account for the variability in the ARL values. It is possible to have an $AARL$ value close to the desired value, ARL_0, but with the individual ARL values widely dispersed. The larger the variability in the in-control ARL values among practitioners, the less confident one would be in a particular chart's performance. Basically, sampling error affects each practitioner. As was illustrated also in Section 1.5.

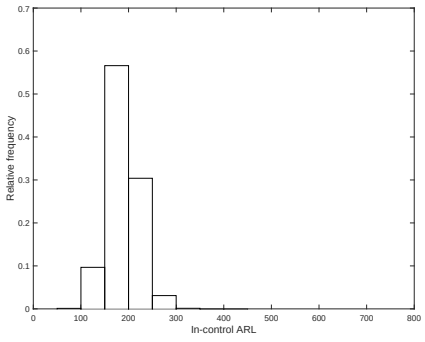
Figure 4.1 presents relative frequency histograms of 100,000 simulated in-control ARL values for four EWMA smoothing constants based on Jones et al.'s (2001) sample-size recommendations. The smoothing constant, λ, and the control chart constant, L, are those producing a chart with known parameters with a specified in-control ARL



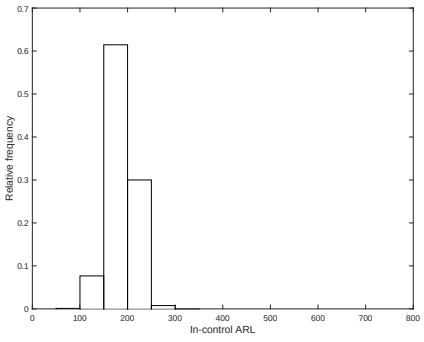
(a) $\lambda = 1.0$, $L = 2.807$, and $m = 100$



(b) $\lambda = 0.5$, $L = 2.777$, and $m = 200$



(c) $\lambda = 0.2$, $L = 2.636$, and $m = 300$



(d) $\lambda = 0.1$, $L = 2.454$, and $m = 400$

Figure 4.1: *Relative frequency histograms of in-control ARL values based on $n = 5$, various values of λ, and sample-size recommendations in Jones et al. (2001).*

value of $ARL_0 = 200$. Jones et al. (2001) determined, when $n = 5$, that m should be at least 400 if $\lambda = 0.1$, 300 if $\lambda = 0.2$, 200 if $\lambda = 0.5$, and 100 if $\lambda = 1.0$. Figure 4.1 shows that the in-control ARL values of EWMA charts designed using this amount of data are quite variable. For example in case (d), where $\lambda = 0.1$ and 400 samples of size 5 are used to estimate the parameters, the chart can have an in-control performance varying from 100 up to 300. A chart with an ARL of 100, for example, would give false signals more frequently than desired. Conversely, a chart with an ARL of 300 would give less frequent false signals than the values specified, but will be somewhat less sensitive to process changes.

4.4 Phase II performance assessment

In Section 2.6, we considered the performance of the EWMA chart in Phase II for the situation that Phase I data were in control or contained contaminations. Furthermore, we assumed that Phase I consisted of $m = 50$ samples of size $n = 5$. In the current section we extend this analysis of the Phase II performance and assess the performance of the EWMA charts based on various Phase I sample sizes m. We assume that the Phase I data are in control and hence no contaminations are present.

Performance measures and simulation procedure

The results in Figure 4.1 show the necessity of an alternative metric to measure the performance of control charts with estimated parameters. A straightforward measure of the practitioner-to-practitioner variability in control charts with estimated parameters is the standard deviation of the ARL (the $SDARL$). The $SDARL$ metric was proposed by Jones and Steiner (2012), who used it to determine the effect of the amount of Phase I data on the risk-adjusted cumulative sum (CUSUM) control chart. Saleh et al. (2015b) evaluated the Shewhart charts in terms of the $SDARL$ metric. They concluded that accounting for the sampling variability requires a far larger amount of Phase I data than that recommended by Quesenberry (1993) in order to reduce the variability among practitioners to an acceptable level. Also, Zhang et al. (2013, 2014) and Lee et al. (2013) used the $SDARL$ metric in evaluating the performance of the exponential CUSUM chart, the geometric chart, and the Bernoulli CUSUM chart, respectively. The use of the $SDARL$ metric shows that the required amount of data to adequately reduce the variation in the in-control ARL to a reasonable level is often prohibitively large.

In order to compute the $AARL$ and $SDARL$ metrics, first the conditional ARL for the EWMA chart with estimated parameters needs to be evaluated. In Chapters 2 and 3 this was done with Monte Carlo simulations, in this chapter we use the Markov chain approach described in the appendix in Section 4.7. The number of states used was 201. This number was found to balance a high level of accuracy and the acceptable time of computation. The results were validated using a Monte Carlo simula-

		m												
	λ	30	50	100	200	300	400	500	600	700	800	900	1,000	∞
$AARL$	0.1	134	147	163	177	183	**186**	189	190	191	192	193	194	200
	0.2	152	162	175	185	**189**	191	193	194	195	196	196	196	200
	0.5	184	186	191	**195**	196	197	198	198	198	198	199	199	200
	1.0	212	206	**202**	201	201	201	200	200	200	200	200	200	200
$SDARL$	0.1	81	68	51	37	30	**26**	23	20	19	17	16	15	n.a.
	0.2	97	76	55	39	**31**	27	24	22	20	19	17	16	n.a.
	0.5	124	90	61	**43**	34	30	27	24	22	21	20	19	n.a.
	1.0	143	100	**66**	45	36	31	28	25	24	22	21	20	n.a.

Table 4.1: *In-control $AARL$ and $SDARL$ values when m Phase I samples, each of size $n = 5$, are used to estimate the in-control values of the process parameters. Bolded values correspond to the sample size recommendations of Jones et al. (2001)*

tion. The process mean was estimated using the overall samples average defined in Equation (4.6) and the process standard deviation was estimated using the pooled standard deviation defined in Equation (4.7). Different values of m, ranging from 30 to $1,000$, with sample size $n = 5$ were considered. We used the same four combinations of control chart design parameters (λ, L) as considered by Jones et al. (2001): (0.1, 2.454), (0.2, 2.636), (0.5, 2.777), and (1.0, 2.807). Under the known in-control parameters assumption, these design parameters produce $ARL_0 = 200$.

Results

Table 4.1 displays the in-control $AARL$ and $SDARL$ values for various number of Phase I samples m of size $n = 5$. The last column, $m = \infty$, refers to the case when the in-control process parameters are known. The bolded values correspond to the sample size recommendations of Jones et al. (2001) on the amount of Phase I data to use.

Although Jones et al.'s (2001) recommendations regarding the amount of Phase I data were based on reducing the occurrence of early false alarms, they also provided practitioners with $AARL$ values close to ARL_0 as shown in Table 4.1. However, the results in the second part of Table 4.1 show that these values of m are associated with large values of the $SDARL$. Accounting for sampling variability in the ARL reveals that the recommended amount of data is not nearly large enough to ensure that individual practitioners will obtain an in-control ARL close to the specified value. Additionally, our results suggest that the larger the smoothing constant, the larger the $SDARL$ will be for a given amount of data. For example, given $m = 30$, $\lambda = 0.1$, the in-control $AARL = 134$ with $SDARL = 81$. If λ increases to 0.5 and 1.0, the in-control $AARL$ increases to 184 and 212 and the corresponding $SDARL$ increases as well to 124 and 143, respectively. Recall that, when $\lambda = 1$, the EWMA chart is equivalent to the Shewhart chart. Thus, we can conclude that Shewhart charts have higher variability in the chart performance than the EWMA charts.

In order to achieve stable in-control ARL performance when process parameters are estimated, the required amount of Phase I data should yield an in-control $AARL$ value close to ARL_0 and a $SDARL$ value that is sufficiently small. Zhang et al. (2014) suggested that a $SDARL$ within 10% of the ARL_0 may be reasonable, although still reflecting a significant amount of variation. Consequently, based on our results, a practitioner would need about 600 samples of size $n = 5$ if $\lambda = 0.1$, 700 if $\lambda = 0.2$, 900 if $\lambda = 0.5$, and $1,000$ if $\lambda = 1.0$ to obtain $SDARL$ values of no more than 20 (10% of 200). In most applications, it will not be realistic to obtain this amount of stable Phase I data from the process.

Furthermore, we studied the required amount of Phase I data of various values of the intended in-control ARL (ARL_0). Table 4.2 displays the in-control $AARL$ and $SDARL$ values for different values of ARL_0 when $\lambda = 0.1$. An equivalent table for $\lambda = 0.5$ can be found in Saleh et al. (2015a, Table 8). The last row of Table 4.2, entitled $m = \infty$, refers to the case when the in-control process parameters are known. The bolded $SDARL$ values in Table 4.2 are those that have a $SDARL$ value within 10% of ARL_0. As shown, the required number of samples m increases with an increase in the ARL_0 value. For example, an EWMA chart with $\lambda = 0.1$ requires about 400 in-control samples of size $n = 5$ when $ARL_0 = 100$, but this increases to $1,000$ samples of size $n = 5$ when $ARL_0 = 500$. This phenomenon occurs because the larger the in-control ARL, the wider the control limits and the further the estimated control limits are in the tails of the distribution of the control chart statistic. It is well-known that estimating more extreme quantiles of a distribution requires larger samples to achieve the same precision as when estimating more central quantiles.

To study the effect of the sample size n, we considered the in-control $AARL$ and

	$ARL_0 = 100$ ($L = 2.148$)		$ARL_0 = 200$ ($L = 2.454$)		$ARL_0 = 370$ ($L = 2.702$)		$ARL_0 = 500$ ($L = 2.815$)	
m	$AARL$	$SDARL$	$AARL$	$SDARL$	$AARL$	$SDARL$	$AARL$	$SDARL$
50	79	28	147	68	258	145	341	209
100	86	21	163	51	290	111	385	160
300	94	12	183	30	331	67	442	98
400	95	**10**	186	26	339	58	453	84
500	96	9	189	23	344	51	461	75
600	97	8	190	**20**	348	46	466	68
700	97	7	191	19	351	42	470	62
800	97	7	192	17	353	39	474	58
900	98	6	193	16	355	**36**	476	54
$1,000$	98	6	194	15	356	34	478	**51**
$1,100$	98	6	194	14	357	33	480	48
∞	100	n.a.	200	n.a.	370	n.a.	500	n.a.

Table 4.2: *In-control $AARL$ and $SDARL$ values for the EWMA chart designed for different ARL_0 values with $\lambda = 0.1$ when m Phase I samples, each of size $n = 5$, are used to estimate the in-control values of the process parameters*

$SDARL$ values for $n = 1$ and $n = 10$. In the case of $n = 10$, the control limits for $10,000$ charts were estimated. Markov chains were used to approximate the conditional ARL; the results are given in Table 4.3. This table shows that smaller number of samples m are needed than for the case of $n = 5$. For example, for the in-control $SDARL$ values to be relatively small, say within 10% of the desired ARL_0 value of 200, around 500 samples of size $n = 10$ are needed irrespective of the value of λ. Compared to 600 tot $1,000$ samples of size $n = 5$ as the results in Table 4.1 show. However, the data requirements, in terms of the total amount observations, i.e. mn, are similar for the case of $n = 5$ and $n = 10$.

		m							
	λ	30	50	100	200	500	$1,000$	$5,000$	∞
$AARL$	0.1	130	144	162	176	188	193	199	200
	0.2	144	156	172	183	192	196	199	200
	0.5	168	177	186	192	197	198	200	200
	1.0	193	194	197	198	199	200	200	200
$SDARL$	0.1	66	57	44	32	**19**	12	4	n.a.
	0.2	71	58	43	30	**18**	12	5	n.a.
	0.5	77	59	43	29	**18**	12	6	n.a.
	1.0	82	61	42	30	**19**	13	6	n.a.

Table 4.3: *In-control $AARL$ and $SDARL$ for the EWMA chart with $n = 10$ and $ARL_0 = 200$*

In the case of $n = 1$ the control limits for $10,000$ charts were estimated and the in-control conditional ARL for each chart was approximated using a Monte Carlo simulation study. For $n = 1$ the process standard deviation is estimated by the moving-range estimator, defined as

$$MR = \frac{\overline{MR}}{1.128} = \frac{1}{2/\sqrt{\pi}} \frac{1}{m-1} \sum_{i=2}^{m} |Y_i - Y_{i-1}|.$$

Table 4.4 contains the in-control $AARL$ and $SDARL$ values for the EWMA chart based on $n = 1$. Again, as with $n = 5$ and $n = 10$, several thousand observations are needed for the in-control $SDARL$ values to be relatively small, say within 10% of the desired in-control ARL value of 200. One has very little to no control over the in-control ARL value if one follows the common recommendation of $25 - 50$ individual observations in Phase I. This result was also demonstrated by Saleh et al. (2015b) for $\lambda = 1$.

We also investigated the required number of Phase I individual observations when changing the value of the desired in-control ARL for the EWMA chart. These results can be found in Saleh et al. (2015a, Tables 13 and 14). Higher numbers of observations are required for the larger value of λ. In addition, the required number of observations increases as the desired value of the ARL_0 increases.

		m							
	λ	30	50	100	200	500	1,000	5,000	∞
$AARL$	0.1	249	193	184	184	194	197	199	200
	0.2	353	248	212	202	200	199	200	200
	0.5	721	365	258	224	209	205	201	200
	1.0	976	459	275	234	210	206	200	200
$SDARL$	0.1	2,061	235	123	81	51	33	15	n.a.
	0.2	1,518	440	182	100	60	41	18	n.a.
	0.5	3,833	1,228	308	145	75	48	22	n.a.
	1.0	5,622	1,302	300	159	79	55	23	n.a.

Table 4.4: *In-control $AARL$ and $SDARL$ for the EWMA chart with $n = 1$ and $ARL_0 = 200$*

4.5 Adjusting the control limits

In order to overcome the problem of the often low in-control ARL values when using estimated parameters, Jones and Steiner (2012) and Gandy and Kvaløy (2013) argued that determining the control limits should be based on the conditional in-control ARL instead of the unconditional one. Their proposal was to adjust the control limits in a way that guarantees, with a suitable high specified probability, that the conditional in-control ARL meets or exceeds the desired level.

Gandy and Kvaløy (2013) approach is based on bootstrapping the Phase I data to construct an approximate confidence interval for the control limits. The general bootstrap procedure, introduced by Efron (1979), is a resampling technique used to estimate the sampling distribution of any sample statistic. In quality-control applications, control charts designed based on bootstrap methods have been suggested as alternatives for the standard design methods. See, for example, Bajgier (1992), Seppala et al. (1995), Liu and Tang (1996), and Jones and Woodall (1998). Recently, Chatterjee and Qiu (2009) proposed estimating the control limits of the CUSUM chart using the bootstrap. Prior work on the bootstrap methods used in quality control focused on determining estimated control limits, not on controlling the conditional ARL performance of control charts.

In order to best describe Gandy and Kvaløy's (2013) approach, let us first define P as the true in-control distribution, $\hat{P}$ as the estimated in-control distribution, $\theta = (\mu_0, \sigma_0)$ as the vector of process parameters, $\hat{\theta} = (\hat{\mu}_0, \hat{\sigma}_0)$ as the vector of estimated process parameters, and q as the control chart limit such that the chart has a specific in-control ARL. The quantities $\hat{P}$ and $\hat{\theta}$ are obtained from m in-control Phase I samples each of size n. The quantity q is a function of P and θ or their estimates. For example, $q(P; \hat{\theta})$ represents the value of control limits conditioned on $\hat{\theta}$, such that under the true in-control distribution P the performance of the chart equals the specified in-control ARL. For simplicity, in this chapter, we evaluate a limit q for

the absolute value of the EWMA chart statistic, defined in Equation (4.1), divided by the quantity $\sqrt{\lambda/(2-\lambda)}$. Therefore, the control limit q that produces the desired in-control ARL is equal to the value of L defined in Equation (4.5).

When parameters are unknown, the observed control chart performance depends on $q(P;\hat{\theta})$, which is unknown because P is unobservable. Gandy and Kvaløy (2013) proposed using the estimator $q(\hat{P};\hat{\theta})$ to build a lower one-sided confidence interval for $q(P;\hat{\theta})$ using the bootstrap technique. Let $(1-\alpha^*)\%$ be the percent of the in-control ARL values equal to or higher than the ARL_0, then we can write

$$P(q(\hat{P};\hat{\theta}) - q(P;\hat{\theta}) > p_{\alpha^*}) = P(q(P;\hat{\theta}) < q(\hat{P};\hat{\theta}) - p_{\alpha^*}) = 1-\alpha^*$$

where p_{α^*} is a constant. The quantity p_{α^*} is unknown because it represents the (α^*) quantile of the unobserved sampling distribution of $q(\hat{P};\hat{\theta}) - q(P;\hat{\theta})$. Note that Gandy and Kvaløy (2013) incorrectly referred to p_{α^*} as the $(1-\alpha^*)$ quantile. This was a typographical error because it should be the α^* quantile. Gandy and Kvaløy (2013) proposed using the bootstrap technique to estimate the distribution of $q(\hat{P};\hat{\theta}) - q(P;\hat{\theta})$ with the distribution of $q(\hat{P}^*;\hat{\theta}^*) - q(\hat{P};\hat{\theta}^*)$ where $\hat{P}^*$ and $\hat{\theta}^* = (\hat{\mu}^*, \hat{\sigma}^*)$ are the estimated in-control distribution and process parameters from the bootstrap samples, respectively. If B is the number of bootstrap samples, then p_{α^*} is approximated with $p^*_{\alpha^*}$, which represents the (α^*) quantile of $[q(\hat{P}^*_b;\hat{\theta}^*_b) - q(\hat{P};\hat{\theta}^*_b)]$, $b = 1, 2, 3, ..., B$. The upper bound $q(\hat{P};\hat{\theta}) - p^*_{\alpha^*}$ is then taken as the adjusted control limit.

The simulation steps followed in this chapter are the same as those listed in Gandy and Kvaløy (2013, page 651). In our simulation procedure, we used $m = 50$ samples of size $n = 5$, $\alpha^* = 0.1$, $\lambda = 0.1$, and $B = 1,000$ bootstrap samples. We assumed, without loss of generality, that the unknown true parameters of the in-control distribution are $\mu_0 = 0$ and $\sigma_0 = \sqrt{n}$. We assumed that the desired in-control ARL_0 is 200. Because we found that the Shewhart chart has higher levels of between-practitioner variability than the EWMA chart, we additionally designed it using this bootstrap approach. The same simulation settings were used for the Shewhart chart. The procedure followed in calculating the control limits, $q(\hat{P};\hat{\theta})$, $q(\hat{P}^*_b;\hat{\theta}^*_b)$, and $q(\hat{P};\hat{\theta}^*_b)$, for each of the Shewhart and EWMA control charts is discussed in detail in the appendix in Section 4.8. Once the limit $(q(\hat{P};\hat{\theta}) - p^*_{\alpha^*})$ was determined, the corresponding in-control and out-of-control ARLs were calculated. For the EWMA chart, the Markov chain approach described in the appendix in Section 4.7 was used in calculating the ARL.

Figures 4.2 and 4.3 contain the boxplots of the in-control and out-of-control ARL distributions, respectively, for the EWMA and Shewhart control charts. For both the EWMA and Shewhart charts, the limits computed with the bootstrap adjustment are indicated as "Adjusted Limits". For reference, charts with "Unadjusted Limits" were computed with $m = 50$ samples of size $n = 5$ using $(\lambda, L) = (0.1, 2.454)$ for EWMA

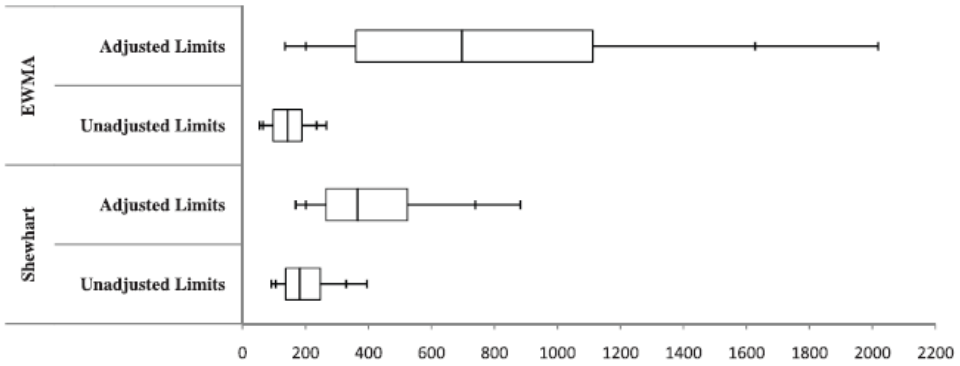


Figure 4.2: *In-control distribution of the conditional ARL when $m = 50$ and $n = 5$. The boxplots show the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles of the conditional in-control ARL distribution*

charts and $L = 2.807$ for Shewhart charts. The out-of-control ARL values were computed with a mean shift of $\tilde{\delta} = 1$, and the boxplots were constructed from $2,000$ ARL values. In Figure 4.2, one can see, as expected, that the adjusted limits resulted in about 90% of the in-control ARL values for both the EWMA and Shewhart charts of at least 200 when computed using the bootstrap approach. Interestingly, more than 75% of the EWMA charts and 50% of the Shewhart charts with unadjusted limits had an in-control ARL below 200, indicating a higher incidence of false alarms.

An interesting feature of Figure 4.2 is that the EWMA charts based on the bootstrap design have a much more variable in-control ARL distribution than the charts based on unadjusted limits. Although the in-control ARL distribution of the EWMA chart is extremely skewed to the right and more variable than that of the unadjusted limits, the out-of-control ARL distribution of the EWMA chart with the adjusted limits is very tight, as shown in Figure 4.3. The EWMA design based on the bootstrap approach has a slightly more variable out-of-control ARL distribution than the standard design. The median out-of-control ARL is around 12 for the adjusted limits and around 9 for the unadjusted limits. This small loss in out-of-control performance comes with "guaranteed" in-control performance with 90% of the bootstrap adjusted charts having in-control ARL values above 200 as compared with only 25% of the charts with unadjusted limits. Although the increased variability in the in-control ARL distribution of the EWMA charts based on the adjusted limits was initially surprising to us, we quickly realized that we are not too concerned about charts with large in-control ARL values as long as they can quickly detect an out-of-control event.

Another interesting feature of Figures 4.2 and 4.3 is that the out-of-control ARL values of the Shewhart chart with the adjusted limits are considerably higher than those of the EWMA chart with either adjusted or unadjusted limits. Hence, if the goal is to avoid frequent false alarms and to detect this sustained shift quickly, then the

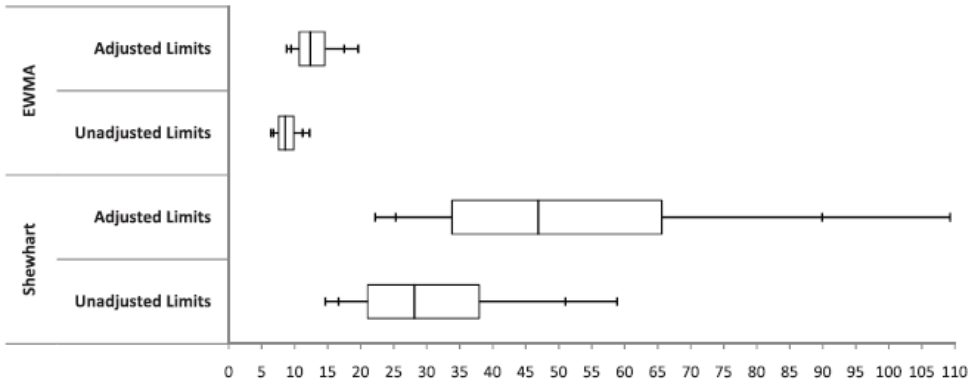


Figure 4.3: *Out-of-control distribution of the conditional ARL when $m = 50$ and $n = 5$. The boxplots show the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles of the conditional in-control ARL distribution*

EWMA chart remains much preferred to the Shewhart chart.

Figure 4.4 shows the relationship between the in-control ARL values and the out-of-control ARL values of the EWMA chart. The scatterplot presents the out-of-control ARL versus the in-control ARL, categorized by the standardized mean being underestimated (< 0) or overestimated (> 0). The lower smooth part of the scatterplot represents the case when the process mean is underestimated. Unexpectedly, the high out-of-control ARL values are associated with the lowest in-control ARL values. It can be concluded from this figure that the increase in the out-of-control ARL is due to overestimating the process mean rather than having a higher in-control ARL. Another point to note from Figure 4.4 is that a positive sustained shift along with an underestimated in-control mean increases the effective shift size and, as a consequence, results in a low out-of-control ARL value. Overestimating the in-control mean, on the other hand, leads to a decrease in the effective shift size and, thus, a significant increase in the out-of-control ARL.

4.6 Conclusion

In this chapter, we have extended the work of Jones et al. (2001) by using the $SDARL$ metric in evaluating the in-control performance of the EWMA control chart when the parameters are estimated. Accounting for the practitioner-to-practitioner variability led to some quite different conclusions regarding the chart performance. First, the EWMA chart requires more Phase I data than previously recommended in order to have consistent chart performance among practitioners. Additionally, we found that charts designed with large values of the EWMA smoothing constant have more variability in the ARL distribution; thus, we recommend more Phase I data be used with larger

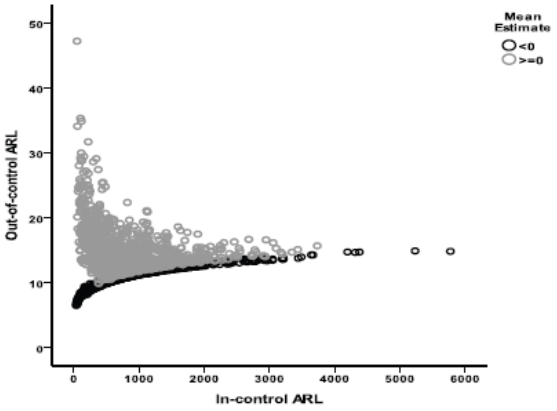


Figure 4.4: *Scatterplot of the in-control vs. the out-of-control conditional ARL values of the EWMA control chart categorized by the mean estimates overestimating or underestimating the process mean*

smoothing constants. Because EWMA charts are typically used when quickly detecting small sustained shifts is of interest, the charts are most often designed with small values of the smoothing constant ($\lambda < 0.25$).

With our recommendations regarding the required amount of Phase I data, we can easily see the difficulty in controlling the in-control ARL value of an EWMA chart. We support the use of the bootstrap-based design approach of Jones and Steiner (2012) and Gandy and Kvaløy (2013), which was recently proposed for controlling the probability of the in-control ARL being at least a specified value. Our results show that adjusting the EWMA control limits accordingly can result in a highly skewed in-control ARL distribution. However, such increases in the in-control ARL did not have much of an effect on the out-of-control performance of the chart. In our opinion, this design approach is very promising and should be considered while evaluating and comparing control charts. Controlling a percentile of the in-control ARL distribution can provide satisfactory chart performance among a wide range of practitioners.

We found that, if one considers the necessary amount of data for stable performance as determined by the $SDARL$, fewer observations are required for designing an EWMA chart than a Shewhart chart. Thus, the EWMA chart, with a small smoothing constant, has an advantage over the Shewhart chart, which would require more Phase I data to achieve similar stability in terms of ARL performance across samples. Additionally, based on the bootstrap design, we found that the EWMA chart is preferred to the Shewhart chart because, with the former, one can simultaneously avoid too frequent false alarms and detect out-of-control sustained shifts more quickly.

4.7 Appendix: calculating the $AARL$ and $SDARL$

In this chapter, the EWMA chart is evaluated using the in-control $AARL$ and $SDARL$ metrics. These metrics were calculated using the following Markov chain approach. Let $\pm h$ be the control limits given in Equation (4.5), t be the number of the subintervals between the upper and lower control limits (namely, the number of transient states), and w be the width of each subinterval defined as $w = 2h/t$. Saleh et al. (2013, Appendix B) derived the transition probabilities p_{lj}, $l = 1, 2, ...t$ and $j = 1, 2, ..t$, for the EWMA chart when process parameters are estimated. The probability p_{lj} refers to the probability of moving from the transient state l to the transient state j. They calculated p_{lj} using

$$\begin{aligned} p_{lj} =& \Phi\left(Q\left[\frac{S_j + w/2 - (1-\lambda)S_l}{\lambda}\right] - \tilde{\delta} + \frac{Z}{\sqrt{m}}|\hat{\mu}_0, \hat{\sigma}_0\right) - \\ & \Phi\left(Q\left[\frac{S_j - w/2 - (1-\lambda)S_l}{\lambda}\right] - \tilde{\delta} + \frac{Z}{\sqrt{m}}|\hat{\mu}_0, \hat{\sigma}_0\right) \end{aligned}$$

where $\Phi(.)$ is the cumulative standard normal distribution function, the quantities Z and Q are defined in Equation (4.4), and $S_{(.)}$ represents the (.)th interval midpoint.

We define $\mathbf{R}$ to be a $t \times t$ matrix consisting of the probabilities of moving from one transient state to another such that $\mathbf{R} = [p_{lj}]$, and $\mathbf{u}$ to be a $t \times 1$ vector of ones. According to Markov chain approach, the ARL vector is computed as

$$ARL = (I - \mathbf{R})^{-1}\mathbf{u} \tag{4.8}$$

where I is the identity matrix of dimension $t \times t$. Here, ARL is a $(t \times 1)$ vector containing the ARLs corresponding to all the possible initial states. We have $Z_0^* = 0$. Hence, for an odd value of t, the $(t+1)/2$-th element (middle element) corresponds to the ARL satisfying this assumption. The ARL defined in Equation (4.8) is a function of the random variables $\hat{\mu}_0$ and $\hat{\sigma}_0$, or more generally the random variables Z and Q. Hence, we can write the $AARL$ as

$$AARL = E(ARL) = \int_0^\infty \int_{-\infty}^\infty ARL g_Z(z) f_Q(q) dz dq \tag{4.9}$$

and the $SDARL$ as

$$SDARL = [E(ARL^2) - [E(ARL)]^2]^{1/2}$$

where

$$E(ARL^2) = \int_0^\infty \int_{-\infty}^\infty ARL^2 g_Z(z) f_Q(q) dz dq \tag{4.10}$$

Here, ARL is the element of the vector ARL corresponding to the initial state, and the

$g_Z(z)$ and $f_Q(q)$ are the probability density functions of the random variables Z and Q, respectively. Because the samples are assumed to be independent and normally distributed, the random variables Z and Q are independent. The variable Z follows the standard normal distribution, while Q follows a scaled chi (χ) distribution. Saleh et al. (2013) provided the functional form of the probability density function of Q. The integrations in Equations (4.9) and (4.10) were approximated using the Gaussian quadrature method. The numerical results were validated using Monte Carlo simulation.

4.8 Appendix: the bootstrap approach

In this chapter, we design the Shewhart and EWMA control charts using the bootstrap approach. We start here with providing the steps of applying Gandy and Kvaløy (2013) algorithm, then we list the calculation steps for the Shewhart and EWMA chart.

Gandy and Kvaløy's Algorithm

The steps of Gandy and Kvaløy (2013) algorithm for obtaining bootstrap-based control limits can be summarized as follows:

1. We let the true unknown in-control distribution P be $N(\mu_0, \sigma_0^2)$, with $\mu_0 = 0$ and $\sigma_0^2 = n$. We generate a Phase I data set of m samples each of size n from $N(0, n)$ and compute $\hat{\mu}_0$ and $\hat{\sigma}_0$. We then compute the quantity $q(\hat{P}; \hat{\theta})$, where $\hat{P} = N(\hat{\mu}_0, \hat{\sigma}_0^2)$ is the estimated in-control distribution and $\hat{\theta} = (\hat{\mu}_0, \hat{\sigma}_0)$ is the estimate of the parameters that are used to run the chart.

2. We then generate $B = 1,000$ bootstrap samples from $\hat{P}$ and compute $\hat{\theta}_b^* = (\hat{\mu}_0^*, \hat{\sigma}_0^*)$ for each of the $b = 1, 2, ..., B$ samples. It is important to note that $\hat{\mu}_0^*$ and $\hat{\sigma}_0^*$ are calculated the same way as $\hat{\mu}_0$ and $\hat{\sigma}_0$ were calculated.

3. We finally compute the quantities $q(\hat{P}_b^*; \hat{\theta}_b^*)$ and $q(\hat{P}; \hat{\theta}_b^*)$ for $b = 1, 2..., B$. We obtain the value of p_α^* as the α percentile of the bootstrap distribution of $q(\hat{P}_b^*; \hat{\theta}_b^*) - q(\hat{P}; \hat{\theta}_b^*)$. The final (adjusted) control limit for the chart is then taken as $q(\hat{P}; \hat{\theta}) - p_\alpha^*$.

We generated the ARL distribution by repeating the steps 1-3 for a number of times (we used $2,000$ times). Next, we explain in detail how to compute the different values of q for the Shewhart and EWMA charts.

Gandy and Kvaløy's (2013) approach is based on three limits; $q(\hat{P}; \hat{\theta})$, $q(\hat{P}_b^*; \hat{\theta}_b^*)$, and $q(\hat{P}; \hat{\theta}_b^*)$. The three limits are defined as follows:

1. The quantity $q(\hat{P}; \hat{\theta})$ represents the value of L that produces the desired in-control ARL when the Phase II data are generated from $P = N(\hat{\mu}_0, \hat{\sigma}_0^2)$ and the limits are constructed using $\theta = (\hat{\mu}_0, \hat{\sigma}_0)$.

2. The quantity $q(\hat{P}_b^*; \hat{\theta}_b^*)$, $b = 1, 2, 3, ..., B$, represents the value of L that produces the desired in-control ARL when the Phase II data are generated from $\hat{P}_b^* = N(\hat{\mu}_b^*, \hat{\sigma}_b^{*2})$ and the limits are constructed using $\theta_b^* = (\hat{\mu}_b^*, \hat{\sigma}_b^*)$.

3. The quantity $q(\hat{P}; \hat{\theta}_b^*)$, $b = 1, 2, 3, ..., B$, represents the value of L that produces the desired in-control ARL when the Phase II data are generated from $\hat{P} = N(\hat{\mu}_0, \hat{\sigma}_0^2)$ and the limits are constructed using $\theta_b^* = (\hat{\mu}_b^*, \hat{\sigma}_b^*)$.

Calculation of q for the Shewhart Control Chart

Consider the Shewhart control chart where the chart statistic is the sample mean ($\bar{Y}_i$) and the control limits are $\mu_0 \pm L\sigma_0/\sqrt{n}$, with L being the control limit constant chosen to satisfy a specific in-control performance. Finding the quantity $q(\hat{P}; \hat{\theta})$ implies finding L such that

$$P\left(\hat{\mu}_0 - L\hat{\sigma}_0/\sqrt{n} < \bar{Y} < \hat{\mu}_0 + L\hat{\sigma}_0/\sqrt{n}\right) = 1 - \alpha$$

where α is the false-alarm probability and $\bar{Y} \sim N(\hat{\mu}_0, \hat{\sigma}_0^2/n)$. This can be simplified to

$$P\left(\left|\frac{\bar{Y} - \hat{\mu}_0}{\hat{\sigma}_0/\sqrt{n}}\right| > L\right) = P(|Z| > L) = \alpha.$$

It follows that $L = Z_{1-\alpha/2}$ or equivalently $q(\hat{P}; \hat{\theta}) = Z_{1-\alpha/2}$. Similarly, $q(\hat{P}_b^*; \hat{\theta}_b^*) = L = Z_{1-\alpha/2}$, for $b = 1, 2, ...B$. For Shewhart charts, α is the reciprocal of the in-control ARL. In our study, $\alpha = 0.005$ and thus $Z_{1-\alpha/2} = 2.807$.

For the quantity $q(\hat{P}; \hat{\theta}_b^*)$, we need to find L such that

$$P\left(\hat{\mu}_b^* - L\hat{\sigma}_b^*/\sqrt{n} < \bar{Y} < \hat{\mu}_b^* + L\hat{\sigma}_b^*/\sqrt{n}\right) = 1 - \alpha$$

where $\bar{Y} \sim N(\hat{\mu}_0, \hat{\sigma}_0^2/n)$ or equivalently

$$P\left(Z < \frac{\hat{\mu}_b^* + L\hat{\sigma}_b^*/\sqrt{n} - \hat{\mu}_0}{\hat{\sigma}_0^2/\sqrt{n}}\right) - P\left(Z < \frac{\hat{\mu}_b^* - L\hat{\sigma}_b^*/\sqrt{n} - \hat{\mu}_0}{\hat{\sigma}_0^2/\sqrt{n}}\right) = 1 - \alpha. \quad (4.11)$$

In such a case, a search algorithm is required for finding the value of L that satisfies Equation (4.11). A search algorithm could be of a binary search type or any other trial and error type.

Our studies showed that L is always bounded between two values. Refer to the quantity in Equation (4.11) as $P(Z < b_2) - P(Z < b_1) = 1 - \alpha$. It can be deduced that, if $\hat{\mu}_0^* \geq \hat{\mu}_0$, then necessarily $b_1 \geq Z_{\alpha/2}$, which implies that

$$L \leq \frac{\hat{\sigma}_0}{\hat{\sigma}_0^*} Z_{1-\alpha/2} + \frac{\hat{\mu}_0^* - \hat{\mu}_0}{\hat{\sigma}_0^*/\sqrt{n}}.$$

Also, if $\hat{\mu}_0^* \leq \hat{\mu}_0$, then necessarily $b_2 \leq Z_{1-\alpha/2}$, which implies that

$$L \leq \frac{\hat{\sigma}_0}{\hat{\sigma}_0^*} Z_{1-\alpha/2} + \frac{\hat{\mu}_0 - \hat{\mu}_0^*}{\hat{\sigma}_0^*/\sqrt{n}}.$$

Hence, an upper bound for L can be defined as

$$\frac{\hat{\sigma}_0}{\hat{\sigma}_0^*} Z_{1-\alpha/2} + \frac{|\hat{\mu}_0 - \hat{\mu}_0^*|}{\hat{\sigma}_0^*/\sqrt{n}}.$$

Additionally, $b_2 - b_1 \geq 2Z_{1-\alpha/2}$ because the symmetric interval is the shortest interval for a standard normal density containing a given probability. Thus, substituting with the expressions for b_1 and b_2 and solving for L provides that $L \geq (\hat{\sigma}_0/\hat{\sigma}_0^*)Z_{1-\alpha/2}$. Consequently, we can say that $L \in (L_1, L_2)$, where

$$L_1 = \frac{\hat{\sigma}_0}{\hat{\sigma}^*} Z_{1-\alpha/2} \text{ and } L_2 = L_1 + \Delta,$$

where $\Delta = \sqrt{n}(|\hat{\mu}_0 - \hat{\mu}_0^*|/\hat{\sigma}_0^*)$. These bounds can help to speed up the search algorithm to solve Equation (4.11) for L.

Alternatively, to solve Equation (4.11) for L one can use a very strong relationship between L/L_2 and Δ/L_2 discussed in Saleh et al. (2015a, Appendix B). Recently, Goedhart et al. (2015) provided an analytical expression to approximate L, this makes the bootstrap approach for the Shewhart chart redundant.

Calculation of q for the EWMA Control Chart

For the EWMA chart, finding the quantity $q(\hat{P}; \hat{\theta})$ is similar to the case of finding $q(P; \theta)$; i.e., the value of L that produces the desired in-control ARL when the in-control process parameters are known. This is because the in-control distribution is defined with the same estimated parameters ($\hat{\theta}$) used in building the control chart limits. Hence, it follows that $q(\hat{P}; \hat{\theta})$ is equal to 2.454 for $\lambda = 0.1$. Similarly, $q(\hat{P}_b^*; \hat{\theta}_b^*) = 2.454$ for $b = 1, 2, 3, ..., B$.

For finding $q(\hat{P}; \hat{\theta}_b^*)$, were the estimates of the in-control distribution differ from that corresponding to the control limits, a search algorithm is required for finding the value of L satisfying an in-control ARL of 200 under data from $\hat{P}$. The search algorithm was of a trial-and-error type and was validation using the Markov chain technique described in Section 4.7 for each and every iteration. For the Markov chain approach the standardized sample mean $\hat{W}$, as defined in Equation (4.4), should be based on the quantities $Q = \hat{\sigma}_b^*/\hat{\sigma}_0$, $\nu_i = \sqrt{n}(\bar{Y}_i - (\hat{\mu}_0 + \delta\hat{\sigma}_0))/\hat{\sigma}_0$, $\tilde{\delta} = \sqrt{n}\delta$, and $Z = \sqrt{mn}(\hat{\mu}_b^* - \hat{\mu}_0)/\hat{\sigma}_0$.

5. Comparing EWMA Charts for Dispersion

Various designs of the EWMA chart for dispersion exist. As shown in Section 1.4, using so-called Phase I estimates effects the performance of control charts. In this chapter, we compare the various designs of the EWMA chart for dispersion based on estimated parameters and argue that the chart for which the performance is least influenced by sampling error (the chart based on the sample variance) should be used in practice. This chapter is based on Zwetsloot (2015).

5.1 Introduction

Various EWMA charts for dispersion have been proposed in the literature; they differ in the dispersion measure used. For example, Knoth (2015) studied the EWMA chart based on the sample variance (S^2). Another possibility is the EWMA chart based on the sample standard deviation (S), which was the studied in Chapter 3. The EWMA chart based on a logarithmic transformation of S^2 or S, is the most prevailing chart. For example, see Crowder and Hamilton (1992), Shu and Jiang (2008), and Maravelakis and Castagliola (2009), amongst others.

The process parameters need to be estimated in most practical applications, and, as discussed in Chapter 1, using estimated parameters severely affects the Phase II performance of control charts. In Chapter 4 we showed that sampling error severely affects the performance of the EWMA chart for location. Less research has been conducted on the effect of estimation on the EWMA chart for dispersion, even though Jensen et al. (2006) noted that "it seems reasonable to expect that the effect of parameter estimation would be more severe on charts that are used to monitor the variance than for charts to monitor the mean".

An exception is Maravelakis and Castagliola (2009) who studied the effect of estimation on the EWMA chart based on a variant of the $\ln(S^2)$ statistic. They compared

the effect of various Phase I sample sizes on the performance of the chart. They concluded that 80 samples of size 5 are not yet enough data to have comparable performance to the known parameter case. Our findings in Chapter 3 are in line with these results. There we showed that the performance of the EWMA chart for dispersion based on the sample standard deviation S is strongly influenced by sampling variability. The effect of sampling variability has also been studied for other types of dispersion charts. For example, Epprecht et al. (2015) and Faraz et al. (2015) both studied the Shewhart chart for dispersion based on estimated parameters. All these studies focus on a single chart and quantified the effect of sampling error by comparing the performance of that chart based on estimated parameters to the performance of the same chart based on known parameters.

In this chapter, we look at the effect of estimation on the chart's performance from a different perspective. Rather than quantifying the effect of Phase I estimation for a *single* chart, we use the effect of estimation on the performance as a comparison metric for *multiple* charts. We argue that the EWMA chart for dispersion which is least influenced by sampling variability should be used in practical applications.

Therefore, the objective of this chapter is to give an advice on which EWMA chart for dispersion to use when process parameters are estimated. That is, which EWMA chart should be used in practice? We conduct an extensive simulation study to evaluate and compare the effect of Phase I estimation on the various EWMA charts for dispersion. We evaluate the performance in terms of both the conditional performance as well as the unconditional (marginal) performance.

This chapter is organized as follows. In Section 5.2, we discussed how to set up the EWMA charts for dispersion. Next, in Section 5.3 we use an example to illustrate the effect of estimation on the three considered charts. In Section 5.4, we discuss the simulation set up. Afterwards, in Section 5.5, a performance comparison of the charts is given. Furthermore, in Section 5.6, we discuss the out-of-control performance. In Section 5.7 we extend the comparison to various designs. Finally, in Section 5.8, we provide some concluding remarks.

5.2 Three EWMA control charts for dispersion

Throughout this chapter, we assume that we observe process observations in samples of size $n > 1$ at regular time intervals $i = 1, 2, ...$ Furthermore, we assume that the process observations are independent and normally distributed with mean μ and standard deviation σ. Data are assumed in control if $\mu = \mu_0$ and $\sigma = \sigma_0$.

When monitoring the dispersion, it is often of interest to detect increases in the dispersion as these indicate some special cause of variation or performance deterioration. Contrary, a decrease in dispersion indicates a process improvement. With this in mind, most research on charts for dispersion has focussed on the one-sided charts, where the statistic is reset to a so-called reflection boundary, if it drops below this boundary. As in Chapter 3, we consider one-sided charts here.

Recall, from Equation (3.2), that for one-sided monitoring the EWMA statistic at time i is defined as

$$Z_i = \max\left[(1-\lambda)Z_{i-1} + \lambda M_i, E[M_i]\right], \text{ for } i = 1, 2, ..., \tag{5.1}$$

where M_i is a measure for the dispersion in the ith sample (in Chapter 3 we used the sample standard deviation for M_i). Again λ, $0 < \lambda \leq 1$, is a smoothing constant. The initial value Z_0 is set equal to the expectation of M_i.

The EWMA chart signals when the statistic Z_i exceeds the upper control limit given by

$$\mu_M + L\sqrt{\frac{\lambda}{2-\lambda}}\sigma_M. \tag{5.2}$$

where L is chosen to satisfy a specific in-control performance, and μ_M and σ_M are the mean and standard deviation of M_i, respectively. For ease of computation we consider the EWMA chart with this asymptotic limit, for the difference between asymptotic and exact limits see Section 1.3 and Abbasi (2010).

Several dispersion measures (M_i) have been studied in the literature. For example, the sample variance S^2 was used by Knoth (2005, 2015). Another possibility is the sample standard deviation S, which was studied in Chapter 3 (or equivalently Zwetsloot et al., 2015a). Following an idea of Box et al. (1978, Chapter 5.4) it became popular to use a dispersion measure based on the logarithm of the sample variance. The rationale behind this is that tests for comparing variances become less sensitive to non normality of the original data, as the statistic is approximately more normally distributed than S^2 itself. Also taking a logarithm of S^2 or S changes the model from a variance shift to a location shift model. Thus an increase in the underlying process dispersion will show as an increase in the mean level of the plotted EWMA statistic, whereas the variance of the plotted EWMA statistics remains equal (Crowder and Hamilton, 1992). This idea has been used and studied quite extensively, see for example Crowder and Hamilton (1992), Castagliola (2005), and Shu and Jiang (2008).

In this chapter, we consider three measures of dispersion (M_i): the sample variance S^2, the sample standard deviation S, and the logarithm of the sample variance $\ln(S^2)$. For conciseness, we only consider these three charts, as the other proposed statistics in, for example, Castagliola (2005) or Abbas et al. (2013), are all extensions of the S, S^2, and $\ln S^2$ based charts. Therefore, we expect that the results for these charts will be comparable to conclusions drawn here.

The EWMA chart statistic in Equation (5.1) can be rewritten as

$$W_i = \max\left[(1-\lambda)W_{i-1} + \lambda D_i, E[D_i]\right], \text{ for } i = 1, 2, ..., \tag{5.3}$$

where D_i is the standardized dispersion measure defined as

$$D_i = \frac{S_i^2}{\sigma_0^2}, \quad D_i = \frac{S_i}{\sigma_0}, \text{ or } D_i = \ln\left(\frac{S_i^2}{\sigma_0^2}\right),$$

depending on the considered chart for dispersion. Here σ_0 is the in-control process standard deviation. In practice, σ_0 is unknown, and is typically replaced with an estimator $\hat{\sigma}_0$ to give

$$\hat{D}_i = \frac{S_i^2}{\hat{\sigma}_0^2}, \quad \hat{D}_i = \frac{S_i}{\hat{\sigma}_0}, \text{ or } \hat{D}_i = \ln\left(\frac{S_i^2}{\hat{\sigma}_0^2}\right). \tag{5.4}$$

Following a similar procedure as in Epprecht et al. (2015) we rewrite $\hat{D}_i$ as

$$\hat{D}_i = \chi_i^2 \frac{1}{n-1}\frac{\gamma^2}{Q^2}, \quad \hat{D}_i = \chi_i \frac{1}{\sqrt{n-1}}\frac{\gamma}{Q}, \text{ or } \hat{D}_i = \ln\left(\chi_i^2 \frac{1}{n-1}\frac{\gamma^2}{Q^2}\right). \tag{5.5}$$

Here the information of the sample at time i is represented by the random variable $\chi_i^2 = (n-1)S_i^2/\sigma^2$. Which follows a chi-squared distribution with $n-1$ degrees of freedom, as we assumed that the process observations are independent and normally distributed. The constant $\gamma = \sigma/\sigma_0$ is the ratio of the actual standard deviation at time i to the in-control standard deviation. If the process is in control $\gamma = 1$. Finally, the variable $Q = \hat{\sigma}_0/\sigma_0$ represents the precision of the Phase I estimate. Note that Q is a random variable, each realization comes from a Phase I sample.

As estimator for the in-control process standard deviation we use the pooled sample standard deviation, as defined in Equation 3.4. Mahmoud et al. (2010) recommended to use the pooled sample standard deviation in quality monitoring applications. Note that S_p is a biased estimator of σ_0, whereas S_p^2 is an unbiased estimator of σ_0^2.

When we use the standardized D_i for the EWMA statistic, the control limit as given in Equation (5.2) becomes

$$\mu_D + L\sqrt{\frac{\lambda}{2-\lambda}}\sigma_D. \tag{5.6}$$

As D_i is standardized, the values μ_D and σ_D are independent of σ_0, see the appendix in Section 5.9 for details. Therefore, the control limit equals a constant depending on the chosen statistic D, the set value for λ, the sample size n, and the chosen ARL_0 which determines L. We denote the control limit as

$$UCL(D, \lambda, n, ARL_0). \tag{5.7}$$

In the appendix (Section 5.9) the values of $UCL(D, \lambda, n, ARL_0)$ are provided.

Throughout this chapter, we refer to the EWMA chart based on the sample variance

as s^2-EWMA, to the EWMA chart based on the sample standard deviation as S-EWMA, and to the EWMA chart based on the logarithm of the sample variance as $\ln s^2$-EWMA. Table 5.1 gives an overview of the three proposed EWMA charts for dispersion.

Chart	Statistic	Description
s^2-EWMA	$\frac{S^2}{\sigma_0^2}$	EWMA chart based on the sample variance
S-EWMA	$\frac{S}{\sigma_0}$	EWMA chart based on the sample standard deviation
$\ln s^2$-EWMA	$\ln \frac{S^2}{\sigma_0^2}$	EWMA chart based on the log of the sample variance

Table 5.1: *Overview of the EWMA charts for dispersion*

5.3 The effect of sampling error

For each chart, its performance will depend on the Phase I estimate. We are interested in comparing the effect of estimating the process dispersion across the three EWMA charts. An overestimate of the in-control process dispersion leads to wider control limits, or equivalently, to an over-correction of the standardized statistic defined in Equation (5.5). Hence increases in the process dispersion will be more difficult to detect. Vice versa, if a practitioner obtains an underestimate of the process dispersion, the control chart will give more frequent false alarms than expected. This variability in performance is often referred to as practitioner-to-practitioner variability and can be alternatively seen as sampling variation for a single user.

To illustrate the sampling variability, consider a practitioner who collected a reference sample of $m = 50$ samples of size $n = 5$ and estimated the process standard deviation using the pooled sample standard deviation (S_p). Table 5.2 provides the percentiles of the cumulative distribution function of S_p or equivalently Q for $\sigma_0 = 1$. The distribution function of $Q = \hat{\sigma}_0/\sigma_0$ can be found in Chen (1997). From Table 5.2 we see that with a probability of 0.95 the practitioner will obtain an estimate of σ_0 between 0.903 and 1.099 when $\sigma_0 = 1$. Note that, in contrast with Chapter 3, we only consider in-control Phase I data.

To illustrate the effect of estimation on the performance, we design three EWMA

	Percentiles								
	2.5	5	10	25	50	75	90	95	97.5
Q	0.903	0.918	0.936	0.966	1.000	1.034	1.064	1.083	1.099

Table 5.2: *Percentiles of the distribution of Q based on $m = 50$ samples of size $n = 5$*

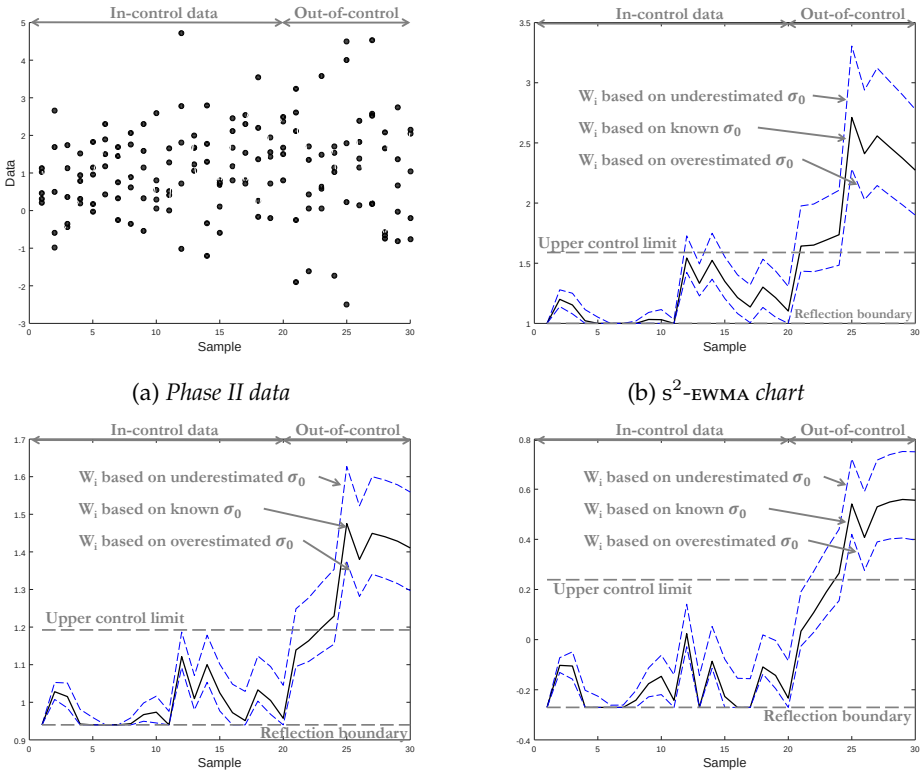
charts for dispersion based on known ($\hat{\sigma}_0 = \sigma_0$), underestimated ($\hat{\sigma}_0 = 0.903\sigma_0$), and overestimated ($\hat{\sigma}_0 = 1.099\sigma_0$) process dispersion. The charts are set up with $\lambda = 0.15$ and UCL such that under known parameters the in-control ARL equals 200. The derivations of the values for UCL can be found in the appendix in Section 5.9 and the UCL equals 1.5894, 1.1924, and, 0.2389 for the s^2-, s-, and $\ln s^2$-EWMA chart, respectively.

Figure 5.1(a) shows the Phase II data, which were generated for this example. We simulated the first twenty Phase II samples as in-control observations and the next ten samples as out-of-control observations with $\gamma = 1.5$. Figures 5.1(b)-(d) show the three EWMA charts used to monitor these data. The dotted horizontal lines equal the control limit and reflection boundary, respectively. The solid line shows the EWMA statistic based on known process dispersion. Around this line two dotted lines show the EWMA statistic based on under- and overestimated process dispersion.

First consider Figure 5.1(b); it shows the s^2-EWMA chart with three plotted EWMA statistics. The solid line, based on known process dispersion, signals an out-of-control condition at sample 21, just after the actual shift in the process. Whereas the top dotted line, based on underestimated process dispersion, signals (a false alarm) at sample 12. The lower dotted lined, based on overestimated process dispersion, signals at sample 25. Thus, depending on the obtained Phase I estimate a practitioner can get, for this situation with a probability of 0.95, a signal between sample 12 and 25.

Next, consider Figures 5.1(c) and (d) which show the s-EWMA and $\ln s^2$-EWMA charts, respectively. The charts based on known process dispersion both signal at sample 24, a bit later than the s^2-EWMA chart. This is in line with the results in Knoth (2010), who showed that, for the case of known parameters, the s^2-EWMA chart signals shifts of size $\gamma = 1.5$ the quickest. For the underestimated case the S and $\ln S^2$ charts do not signal a false alarm around sample 12, contrary to the s^2-EWMA chart, and signal the shift in the process at sample 21, respectively, sample 22. Hence there are real differences in performance between the three EWMA charts for dispersion. Figures 5.1(b)-(d) show the necessity to take into account the effect of sampling variability.

Note that for $\lambda = 1$ the EWMA chart is equivalent to the Shewhart chart. Epprecht et al. (2015) study the Shewhart chart for dispersion based on S^2 and illustrate the sampling variability. Contrary to the EWMA chart, it makes no differences for the Shewhart chart which statistic (S^2, S, or $\ln S^2$) is used for monitoring. Figures 5.1(b), (c) and (d) would be identical for the Shewhart chart. This difference is explained by the time dependence in the plotted EWMA statistic. The plotted statistic for the Shewhart control chart is time independent, hence the probability that a statistic, for example S, exceeds the control limits is equal to the probability that a monotone transformation of that statistic, for example S^2, exceeds the control limits (which has also been transformed according to the same monotone transformation). However, for the EWMA chart the plotted statistic W_i is a weighted average of S_i or S_i^2 and these weighted averages are no long one-on-one transformations. Therefore, the EWMA charts signal at



(a) *Phase II data*

(b) S^2-EWMA *chart*

(c) S-EWMA *chart*

(d) $\ln S^2$-EWMA *chart*

Figure 5.1: *Example Phase II data and three EWMA charts for dispersion applied to these data*

different times if they are based on different statistics. Consequently, the EWMA charts based on S^2 and S, or any other dispersion metric, will show different performance as illustrated in Figures 5.1(b), (c), and (d). We will show that as λ increases towards 1 the difference in performance becomes smaller.

5.4 Performance measures and simulation procedure

Above it was illustrated that the performance of the EWMA control chart can vary depending on the chosen dispersion statistic and the obtained Phase I data set. For a comprehensive comparison of these charts we evaluate the performance using the conditional ARL, as well as the chart's overall performance evaluated in terms of the $AARL$ and $SDARL$ metrics, as introduced in Sections 1.5 and 4.4. We will show that for the EWMA charts for dispersion the $AARL$ values can be similar but the charts show very different patterns in the variability of the conditional ARLs. The larger

this variability the larger the uncertainty about the performance of a single chart.

We compute the conditional ARL with the statistical software package R using the functions `ewma.arl` and `lns2ewma` in the package `spc` (Knoth, 2014). These functions are designed to compute ARL values for charts based on known parameters. They can be used to compute conditional ARLs for charts based on estimated parameters by using Equation (5.5). Note that the plotted statistic D_i is proportional to γ/Q. This equivalence is used to compute the conditional ARLs. For more details see Section 5.6. We compute the conditional ARL for $50,000$ random generated realizations of Q. The mean and standard deviation of these $50,000$ conditional ARLs give the $AARL$ and $SDARL$ values. Furthermore, these results were validated using a Markov chain approach and integral equations comparable to the Markov chain approach and integral equations discussed in Section 4.7.

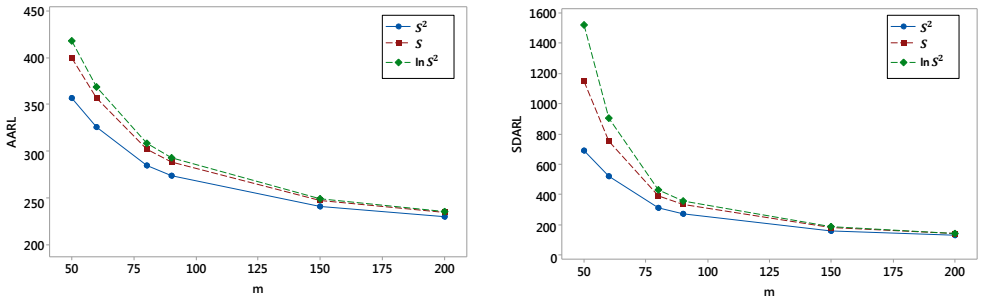
Implementation of the EWMA charts involves several design choices - the chart itself (S^2, S, or $\ln S^2$), the size of the shift the chart is tuned to (λ), the desired in-control performance (ARL_0) which also determines the value of L, the size n of the collected samples, and the number of samples m collected in Phase I. We explore the impact of these factors on the performance of the EWMA charts with simulation. We assume that the in-control process follows a normal distribution where we set, without loss of generality, $\mu_0 = 0$ and $\sigma_0 = 1$. As soon as the process goes out of control, the process dispersion equals γ, which will have a value > 1.

5.5 Comparison of the in-control performance

In this section, we first consider a single configuration of the EWMA charts for dispersion, and use this chart as an illustrative comparison of the three dispersion charts. We use the same setting as in Section 5.3; $\lambda = 0.15$, $n = 5$, and $UCL = 1.5894, 1.1924$ and 0.2389, such that the charts have an $ARL_0 = 200$ under known parameters. In Section 5.7, we will discuss the effect of these selected design settings and generalize our comparison to a range of design settings.

Figure 5.2 displays the $AARL$ and $SDARL$ values for each of the three dispersion charts versus the number of Phase I samples $m = 50, 100, ..., 200$ used to estimate $\hat{\sigma}_0$.

Figure 5.2(a) shows, as expected, that all three charts suffer from estimation error. The smaller the value of m, the larger the difference between the $AARL$ and the designed value of $ARL_0 = 200$. The figure also shows that the three charts react differently to Phase I estimation; the s^2-EWMA chart has an $AARL$ which is the closest to the designed value of $ARL_0 = 200$. Whereas, the S-EWMA has a higher $AARL$ followed by $\ln S^2$ with the highest $AARL$ for each considered values of m. However, recall that for an in-control process a high ARL is desirable. From this perspective all three charts show acceptable performance. Note that by adjusting the UCL one can influence the $AARL$ value. We have designed charts such that the ARL_0 equals 200 under known parameters. Of course the chart can also be designed to have $AARL = 200$



(a) $AARL$ (b) $SDARL$

Figure 5.2: *In-control $AARL$ and $SDARL$ of each of the three EWMA charts for dispersion versus m, the number of Phase I samples*

instead of $ARL_0 = 200$.

The $AARL$ is an average of conditional ARLs. To study the variability in these values we studied the $SDARL$ displayed in Figure 5.2(b). Note that the y-axis has a different scale than Figure 5.2(a). The results may be surprising - it shows that the $SDARL$ differs strongly across the three dispersion charts. The $\ln s^2$-EWMA chart has a $SDARL$ of over 1500, when the chart is designed with a Phase I estimate based on $m = 50$ samples of size $n = 5$. Whereas the $SDARL$ of the s^2-charts is half of this values and lies around 700. Concluding, the $AARL$ of the charts are relatively close together, however the $SDARL$ values differ strongly across the three charts especially for m in the range of $50-100$. Note that the $SDARL$ value is dependent on the $AARL$ value.

Another aspect seen from Figure 5.2(b) is the effect of increasing the sample size in Phase I. The $SDARL$ goes down substantially for m up to 150, afterwards it only declines very slowly as m increases to 200. In order to achieve a predictable ARL performance when process parameters are estimated, the $SDARL$ should be small. Zhang et al. (2014) suggested that a $SDARL$ within 10% of the ARL_0 may be reasonable, this implies a $SDARL$ value < 20. For $m = 5000$ (not shown in the figure) the $SDARL$ of the three charts equals around 20. Hence only for 5000 samples (or more) of size $n = 5$ the charts will have predictable conditional ARL performance.

As a comparison consider the EWMA chart for location. In Chapter 4 we showed that for this EWMA chart for location, designed with $\lambda = 0.1$, around $m = 600$ to 800 samples of size $n = 5$ are needed to have a $SDARL$ within 10% of the $AARL$. For the EWMA chart for dispersion designed with $\lambda = 0.1$, around 5000 samples are needed. This difference in required number of samples confirms the prediction in Jensen et al. (2006) that charts designed for monitoring the dispersion are more affected by sampling error than charts designed to detect location shifts. It seems that impractical large sample sizes are needed to decrease the variability in the performance of the

EWMA chart for dispersion.

To better understand these very large $SDARL$ values, we study the conditional ARL of the three dispersion charts. Figure 5.3 shows, in the lower panel, the curves of the conditional ARL for the three EWMA charts for dispersion versus the standardized Phase I estimate Q. Note that the conditional ARLs are on a logarithmic scale. The horizontal dotted line represents the $ARL_0 = 200$ value. For $Q = 1$ the charts obtain this value, as designed. The top panel in Figure 5.3 shows the probability density function of Q based on $m = 50$ and $m = 500$ Phase I samples of size $n = 5$. The functional formula of the density of Q can be found in Chen (1997).

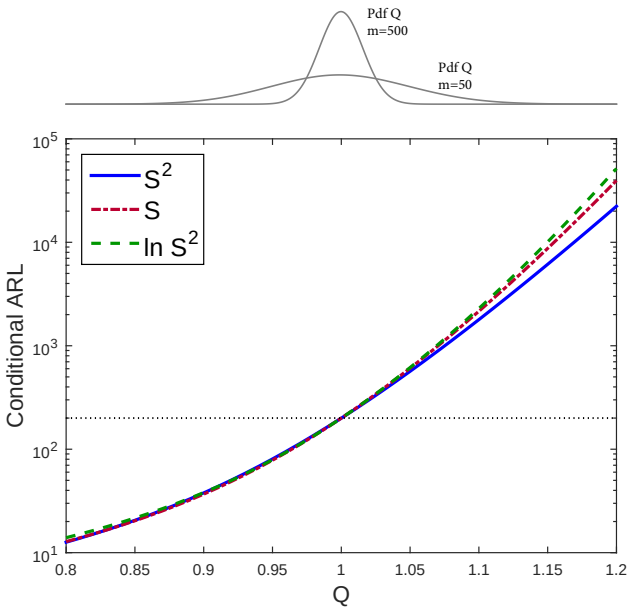


Figure 5.3: *Conditional ARL versus Q, the standardized Phase I estimate. Dotted horizontal line represents the pre-specified ARL_0. Top panel shows probability density plots of Q*

Figure 5.3 shows that the conditional ARL increases as Q increases, i.e. if the process dispersion is overestimated in Phase I. What is - perhaps - surprising is that the conditional ARL profile shows values up to (and over) 10,000 runs. This phenomenon occurs because the larger the estimated in-control process dispersion, the smaller the values of D_i (as it is divided by this estimate) and thus the EWMA statistic stays small and has a (very) small chance to exceed the upper control limit. The density plots shows that conditional ARLs of 10,000 have very small probability of occurring if the sample size is large ($m = 500$) and a larger, but still small, probability of occurring if the sample size is small ($m = 50$).

Another feature that can be seen from Figure 5.3 is that the choice for a dispersion

statistic (i.e. D_i) has a large effect on the conditional ARL. Note that, especially for overestimated values of Q the s- and $\ln s^2$-EWMA charts have conditional ARLs which are much larger than the conditional ARLs of the s^2-EWMA chart. For overestimating the standard deviation by 10 percent ($Q = 1.1$), the conditional ARL equals 1789 for the s^2-EWMA chart, 2156 for the s-EWMA chart, and 2291 for the $\ln s^2$-EWMA chart. The probability of obtaining $Q = 1.1$ depends on the number of Phase I samples as can be seen in the top panel of Figure 5.3. This difference in conditional ARL coincides with the different values for the $SDARL$ displayed in Figure 5.2(b) for $m = 50$ and with the near identical $SDARL$ values for $m > 150$.

With respect to in-control data, the difference of overestimated standard deviation ($Q > 1$) is not really a concern as the ARL is very high; one expects it to be 200 and if it is a lot higher than this should not be a problem. Of course, the out-of-control performance should not be much higher than expected. This is discussed in Section 5.6. A more problematic feature in Figure 5.3 is when $Q < 1$, i.e. when the process dispersion is underestimated. All three charts have very low conditional ARLs. Thus all three charts give more frequent false alarms than expected.

Concluding, the in-control performance of the three EWMA charts for dispersion based on estimated parameters is strongly influenced by sampling error. Especially, if the process dispersion is overestimated by at least 10%, the conditional ARL increases heavily. Furthermore, when the process dispersion is underestimated, the charts yield more frequent false alarms. We showed that increasing the number of Phase I samples decreases the sampling variability. However at least $m = 5000$ samples are needed to have a reasonable low variability. All three EWMA charts suffer from this, although they are affected differently. The s^2-EWMA chart has the lowest conditional ARL values. This translates into the smallest $SDARL$. Hence the s^2-EWMA chart has the most predictable in-control performance, when the process dispersion is estimated based on limited Phase I data ($m < 150$).

5.6 A note on the out-of-control performance

Apart from in-control performance, the out-of-control performance of a chart is very important. It reflects the ability to detect process changes, and therefore reflects the main goals of process monitoring. Usually, the performance of charts is compared for a range of possible out-of-control scenarios, whereby for the EWMA chart one focusses on sustained shifts in the process parameter. As we focus on one-sided charts we study increases in the process dispersion from the in-control level σ_0 to $\sigma > \sigma_0$, or equivalently $\gamma > 1$.

The out-of-control performance of the three charts based on known parameters is compared by Knoth (2010). In his analysis he showed that differences in out-of-control performance are small, especially for increases in the process dispersion. He showed that the s-EWMA chart detects shifts of size $1 < \gamma < 1.3$ slightly better than the other two charts and that the s^2-EWMA chart detects shifts of size $\gamma > 1.3$ the quickest.

Knoth (2010) considered two-sided charts.

In this section, we are interested in the out-of-control performance of the three charts when process parameters are estimated. More specifically, we are interested in the *difference* between the out-of-control performance for the three considered EWMA charts. To evaluate the out-of-control performance it is clarifying to consider the standardized version of D_i from Equation (5.5)

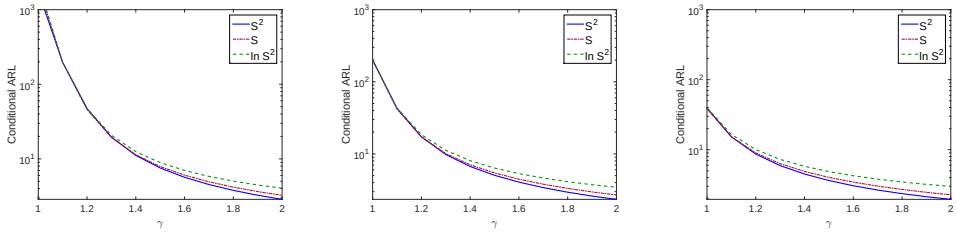
$$D_i \sim \chi_i^2 \frac{1}{n-1} \frac{\gamma^2}{Q^2}, \tag{5.8}$$

where "$\sim$" implies that D_i is proportional to the right-hand side of this equation.

The out-of-control performance is modelled as $\gamma > 1$. As long as the ratio of γ/Q remains constant, the value of D_i is constant, and thus the performance of the EWMA chart remains equivalent. Thus the conditional ARL for out-of-control data with $\gamma = 1.25$ where the chart is set up with an on target Phase I estimate ($Q = 1$), is by definition equal to the in-control conditional ARL ($\gamma = 1$) where the chart was set up with an Phase I estimate $Q = 0.8$ as $1.25/1 = 1/0.8$. Therefore, studying out-of-control situations for $\gamma > 1$, is equivalent to studying the in-control conditional performance for $Q < 1$. Recall that in Figure 5.3 we studied the conditional in-control ARL for a range of values of Q. For the low values of Q the three considered EWMA charts for dispersion showed similar performance. Hence, the out-of-control performance of the chart based on estimated parameters will be similar across the three charts. Therefore, the out-of-control performance is not very relevant for our comparison as the three charts have similar out-of-control performance if the parameters are estimated.

To illustrate this point, Figure 5.4 shows the conditional out-of-control ARLs for each chart versus the out-of-control shift size γ. In each subfigure the charts are based on a different Phase I estimate; in Figure 5.4(a) the charts are designed based on an overestimated Phase I estimate of the dispersion ($Q = 1.099$), in Figure 5.4(b) the charts are designed based on known dispersion ($Q = 1$), and in Figure 5.4(c) the charts are designed based on an underestimated Phase I estimate of the dispersion ($Q = 0.903$).

Figure 5.4 shows that the out-of-control performance depends on the Phase I estimate; for overestimated dispersion the out-of-control ARL is larger than for underestimated dispersion. Given a Phase I estimate, the three charts show similar out-of-control ARL values. The conditional ARL values vary slightly for large shift sizes; the out-of-control conditional ARL differ at most 2 across the charts. In conclusion: the out-of-control performance is nearly equivalent across the three considered EWMA charts for dispersion, irrespective of the Phase I estimate (Q) used to set up the chart.



(a) *Overestimated* Q (b) *Known* Q (c) *Underestimated* Q

Figure 5.4: *Out-of-control conditional* ARL *versus* γ*, the ratio of the out-of-control to the in-control process dispersion. Panels show different Phase I estimate*

5.7 Extending the comparison to various designs

So far we have studied a single set up of the EWMA charts for dispersion, where we set $\lambda = 0.15$, the in-control $ARL_0 = 200$, and we used a sample of size $n = 5$. In this section, we consider the effect of changing these setting on the relative performance of the three dispersion charts.

To explore the effect of the desired in-control performance, we set the ARL_0 equal to $100, 370$, and 500 which covers the commonly used choices. The corresponding values of L can be found in the appendix in Section 5.9. Figures 5.5(a)-(c) show the curves of the conditional ARL for the three EWMA charts versus the standardized Phase I estimate Q, for each value of ARL_0, respectively. Unsurprisingly, as the ARL_0 increases the conditional ARL also increases for all three charts. The difference between the three charts is similar for the ARL_0 equal to $100, 370$, and 500. As well as similar to the case considered as illustrative example (Figure 5.3) were $ARL_0 = 200$. Hence, the choice of ARL_0 does influence the level of the conditional ARL, however it does not influence our comparison. Our conclusion that the s^2-EWMA chart has the most predictable conditional ARL performance, holds for a range of ARL_0 values.

Next, we study the effect of λ on the EWMA chart's performance. The values of λ should be chosen based on the out-of-control shift that the charts should detect quickly. EWMA charts are typically used when quick detection of small sustained shifts is of interest, and the charts are most often designed with small values of the smoothing constant ($\lambda < 0.3$). Crowder and Hamilton (1992) advise $\lambda \approx 0.15$ for a shift of $25 - 30$ percent of the target value. Recall that for $\lambda = 1$ the charts are equivalent to the Shewhart charts and that there is no difference in performance between Shewhart charts based on S^2, S, or $\ln S^2$. Hence, as λ tends to 1 we expect the difference in conditional ARL between the charts to become smaller.

This expectation is confirmed by Figures 5.5(d)-(f), which display the curves of the conditional ARL versus the standardized Phase I estimate Q for three values of $\lambda = 0.1, 0.3$, and 1, respectively. For $\lambda = 0.3$ there is still substantial difference between the conditional ARLs. As expected for $\lambda = 1$, there is no difference in performance

across the three chart. As EWMA charts are typically designed with $\lambda < 0.3$, one can expect that the EWMA charts show different performance in practice, in line with the conclusions drawn in the previous sections for $\lambda = 0.15$.

Finally, we consider the choice of n. In this chapter we have considered monitoring samples of size $n > 1$. Thereby we compare within sample variation over time. The size of the sample is set at $n = 3, 7$, and 9. Figures 5.5(g)-(i) show the curves of the conditional ARL for each EWMA chart versus the standardized Phase I estimate Q for $n = 3, 7$, and 9, respectively. For a given overestimated value $Q > 1$, the conditional ARL is larger for larger n. Note that n influences how many Phase I observations there are, hence the distribution of Q depends on n. Therefore, the probability of a given Q to occur differs across the values of n. We see that in comparing the three EWMA chart for dispersion, the conditional ARL patterns prevail; the differences between the charts that we saw in Figures 5.3 prevail for the various values of n.

Overall we see that the conclusions based on the illustrative setting used in the previous sections hold when these settings are changed.

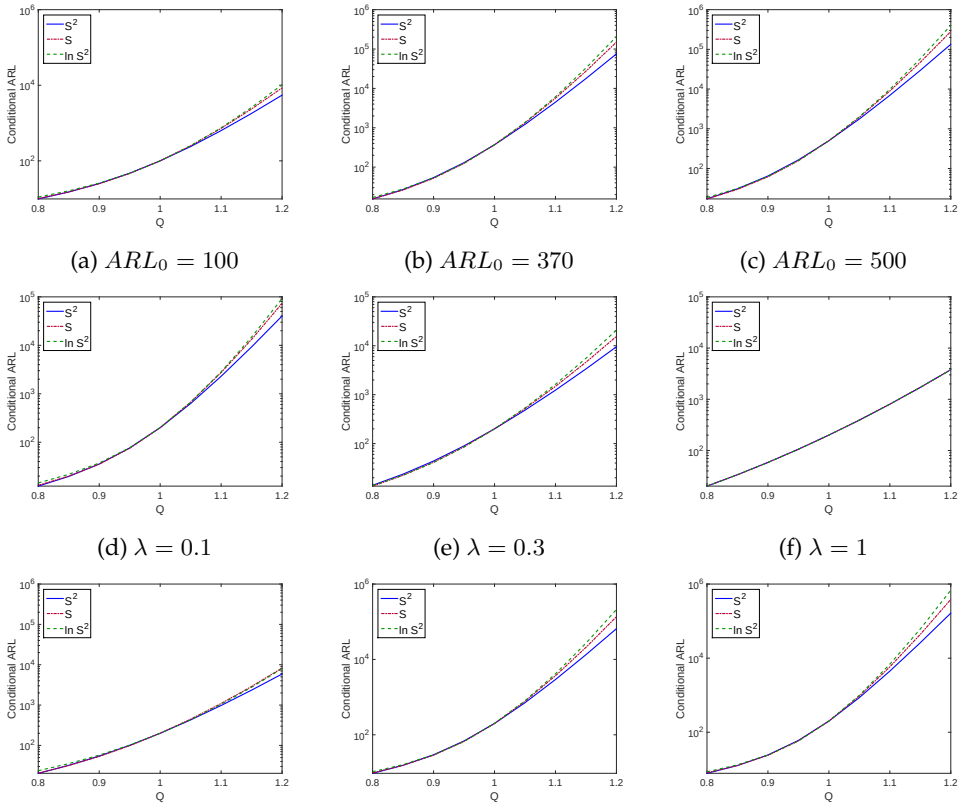
5.8 Conclusion

In this chapter, we have extended the work of Knoth (2010) by comparing EWMA charts for dispersion based on estimated parameters. We have used the conditional ARL to compare performance across charts. First, the s^2-EWMA chart has the most predictable performance; the conditional ARL lies the closest to the designed value. Additionally, we found that sampling error more strongly affects the charts for dispersion than the chart for location. Thus, we recommend more Phase I data to be used than typically advised for the EWMA chart for location.

By considering the conditional performance of the charts, we discovered that the difference in performance arises from large overestimation of the process dispersion. Furthermore, the conditional ARL values can be surprisingly high. The charts based on $\ln S^2$ and S show the largest conditional ARL values compared to the chart based on S^2.

Furthermore, the three charts have near equivalent out-of-control conditional ARL performance. The s^2-EWMA chart shows the largest out-of-control ARL values. However, this difference is so small that we consider it to be of no practical value to emphasize it. Hence, irrespective of the Phase I estimate that they were designed with, all three charts can detect sustained shifts equally quick.

The common view in the literature is that the EWMA chart based on a logarithmic transformation of the sample variance should be used to monitor the dispersion. In this chapter we found that this chart, if based on estimated parameters, has the largest variability in the conditional ARLs, measured by the $SDARL$. Therefore, we recommend to use the EWMA chart based on the sample variance in practice, as it has the most predictable performance. Initially Box et al. (1978, page 122) suggested using $\ln(S^2)$ as it "is more robust to non normality in the original data". Hence the



(a) $ARL_0 = 100$ (b) $ARL_0 = 370$ (c) $ARL_0 = 500$

(d) $\lambda = 0.1$ (e) $\lambda = 0.3$ (f) $\lambda = 1$

(g) $n = 3$ (h) $n = 7$ (i) $n = 9$

Figure 5.5: *In-control conditional ARL values versus Q, the standardized Phase I estimates. Panels show charts designed with various values of ARL_0, λ, and n*

comparison of the dispersion charts probably alters for different distributions of the process characteristic.

We found that the necessary number of data in Phase I, in order to obtain predictable performance is impractically large ($m = 5000$). In Chapter 4 we applied a different design approach for the EWMA chart for location: rather than designing the chart to have a specified ARL_0 the charts were designed to have at most say 10 percent of the charts with an conditional ARL smaller than the specified ARL_0. This so-called guaranteed conditional performance design has been studied for the Shewhart S^2-chart by Faraz et al. (2015). This approach can also be carried out for the EWMA charts for dispersion. We have no reason to doubt that our conclusion that the s^2-EWMA chart has the most predictable performance would carry over to this guaranteed performance design scheme.

In this chapter we have considered a single Phase I estimator; the pooled sample

standard deviation. The choice for this estimator influences the charts performance (see, e.g. Chapter 3). However, since we use the same estimator for all three charts, we expect that our conclusions considering the comparison across the charts, will carry over to other choices of Phase I estimators.

We have only considered one-sided EWMA charts for dispersion. Our comparison is probably influenced by this choice. Recall that the difference between the charts mostly prevails for overestimated Q. In the case of a two-sided chart many more false alarms will probably be observed, hence the conditional ARL will not be as high as shown here for the one-sided charts. This is left for future research.

5.9 Appendix: computation of the upper control limit

The EWMA chart for dispersion signals when the EWMA statistic (as defined in Equation 5.3), exceeds the upper control limit as given in Equation (5.7):

$$UCL(D, \lambda, n, ARL_0).$$

In this appendix we provide the values of the UCL. These values depend of the choice of the dispersion statistics used to monitor D, the selected sample size n, the choice of λ, and the choice of ARL_0. For a specific value of λ the UCL can be determined by referring to Figure 5.6. Each figure shows the UCL versus λ for a specified ARL_0 for known parameters. A search algorithm was used to obtain the values for UCL.

Next we derive the functional form of $UCL(D, \lambda, n, ARL_0)$. Recall that $UCL = \mu_D + L\sqrt{\lambda/(2-\lambda)}\sigma_D$ where μ_D and σ_D are the mean and standard deviation of $\hat{D}_i$ which is defined as in Equation (5.4):

$$\hat{D}_i = \frac{S_i^2}{\hat{\sigma}_0^2}, \quad \hat{D}_i = \frac{S_i}{\hat{\sigma}_0}, \text{ or } \hat{D}_i = \ln\left(\frac{S_i^2}{\hat{\sigma}_0^2}\right).$$

Because we have assumed that the process characteristic is normally distribution, it follows that S^2 is gamma distributed. The values of μ_D and σ_D follow from the gamma and log-gamma distribution. They are given in Table 5.3 and derivations can be found in Mahmoud et al. (2010) and Crowder and Hamilton (1992).

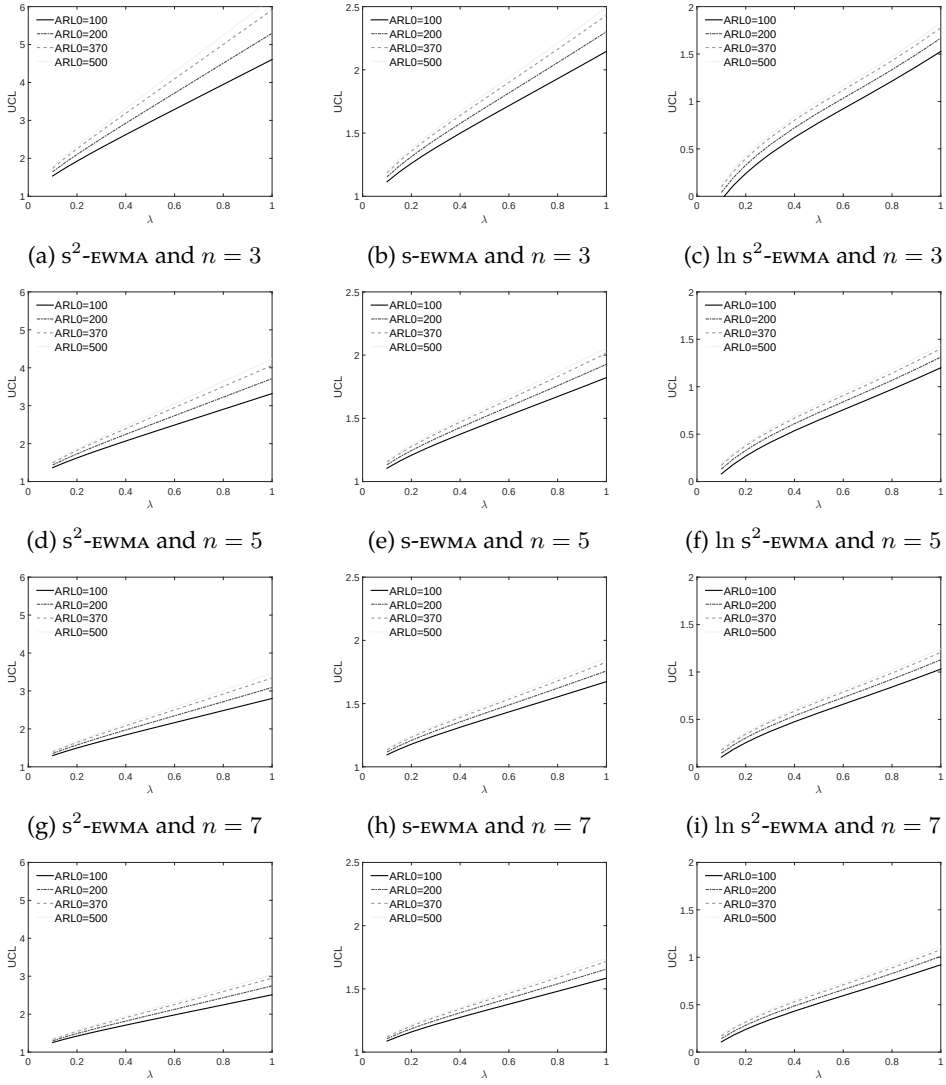
After some simplification it follows that $UCL(D, \lambda, n, ARL_0) =$

$$\begin{cases} 1 + L_{S^2}\sqrt{\frac{\lambda}{2-\lambda}}\sqrt{\frac{2}{(n-1)}} & \text{if } D = \frac{S^2}{\sigma_0^2}; \\ c_4(n) + L_S\sqrt{\frac{\lambda}{2-\lambda}}\sqrt{1-c_4^2(n)} & \text{if } D = \frac{S}{\sigma_0}; \\ \ln\left(\frac{2}{n-1}\right) + \psi\left(\frac{n-1}{2}\right) + L_{lS^2}\sqrt{\frac{\lambda}{2-\lambda}}\sqrt{\psi'\left(\frac{n-1}{2}\right)} & \text{if } D = \ln\left(\frac{S^2}{\sigma_0^2}\right). \end{cases}$$

Where L_{S^2}, L_S, and L_{lS^2} are the corresponding control chart constants, $\psi()$ is the digamma function, and $\psi'()$ is the trigamma function. Furthermore, the control charting constant $c_4(.)$ is defined in Equation 3.1.

D	μ_D	σ_D
$\frac{S^2}{\sigma_0^2}$	1	$\sqrt{\frac{2}{n-1}}$
$\frac{S}{\sigma_0}$	$c_4(n)$	$\sqrt{1-c_4^2(n)}$
$\ln\frac{S^2}{\sigma_0^2}$	$\ln\left(\frac{2}{n-1}\right) + \psi\left(\frac{n-1}{2}\right)$	$\sqrt{\psi'\left(\frac{n-1}{n}\right)}$

Table 5.3: *Mean and standard deviation of dispersion measures D*



(a) s^2-EWMA and $n = 3$ (b) S-EWMA and $n = 3$ (c) $\ln s^2$-EWMA and $n = 3$

(d) s^2-EWMA and $n = 5$ (e) S-EWMA and $n = 5$ (f) $\ln s^2$-EWMA and $n = 5$

(g) s^2-EWMA and $n = 7$ (h) S-EWMA and $n = 7$ (i) $\ln s^2$-EWMA and $n = 7$

(j) s^2-EWMA and $n = 9$ (k) S-EWMA and $n = 9$ (l) $\ln s^2$-EWMA and $n = 9$

Figure 5.6: *UCL versus λ for various* EWMA *charts for dispersion and various sample sizes n.*

6. *Summary*

EWMA Control Charts in Statistical Process Monitoring

In this dissertation we contribute to the development and understanding of the EWMA control chart based on estimated parameters. The presented research is based on Zwetsloot, Schoonhoven, and Does (2014, 2015a, 2015b), Saleh, Mahmoud, Jones-Farmer, Zwetsloot, and Woodall (2015a), and Zwetsloot (2015). In this concluding chapter, we summarize our findings.

6.1 EWMA control charts

All processes show variation. Part of this variation is *inherent* to the process when it functions in a predictable manner. However, sometimes variation may occur because of *special causes* of variation. Deciding which observation occurs because of special causes is not always easy. As Box, Hunter and Hunter (2005, page 566) pointed out "nothing looks so unrandom as a series of random numbers". Control charts are used to operationalize 'common' and 'special' causes of variability.

The original control chart was developed by W. A. Shewhart in a memorandum issued on May 16, 1924. The exponentially weighted moving average (EWMA) control chart was introduced by S. W. Roberts in 1959. The EWMA statistic is a weighted average of measurements, giving heaviest weights to the most recent observations. This provides the chart the advantage of being sensitive to small- and moderate-sized sustained shifts in the process parameters.

It is generally accepted that a control chart is implemented in two phases: Phase I, to define the stable state of the process characteristic and to estimate the process parameters; and Phase II to monitor the process. For an overview of Phase I methods, see Chakraborti et al. (2009) and Jones-Farmer et al. (2014b).

In this dissertation, we study the EWMA chart based on estimated parameters. We consider monitoring both the location as well as the dispersion. We assume that the process characteristic can be modelled as an independent and normally distributed random variable.

6.2 Motivation

We consider the problem of *estimating process parameters from data which can contain contaminated observations*. In practice, data sets often contain outliers, step shifts, recording errors, and other data quality issues (Jones-Farmer et al., 2014b; Vining et al., 2015). Al these types of 'contaminations' are problematic as they can influence the parameter estimates, resulting in control charts with less predictable statistical performance. The possible presence of contaminations in Phase I, provides the motivation for Chapters 2 and 3.

Moreover, we consider *the effect of sampling variability on the monitoring performance* of the EWMA chart. It is well known that the performance of control charts is influenced by Phase I estimation (Jones et al., 2001; Jensen et al., 2006; Psarakis et al., 2014). Generally, control charts based on estimated parameters are evaluated by studying the performance averaged over all possible Phase I estimates. However, the performance of a control chart will depend on the 'actual' estimate. This conditional performance can be very different from the 'average' performance. It is this variability in conditional performance that motivates Chapters 4 and 5.

6.3 Methods

We study *robust estimation methods*, to ensure accurate estimates of the in-control process parameters if contaminations are present. Using a simulation study, we compare the performance of the considered estimation methods for in-control data and for various contamination scenarios. The effectiveness of the estimation methods is evaluated in terms of the accuracy of the resulting estimates and the proportion of successfully identified contaminated observations. Furthermore, we propose new estimation methods for the location and the dispersion, based on EWMA charts.

To evaluate the effect of sampling variability on the EWMA chart's performance, we *study the variability of the conditional average run lengths*. Furthermore, in Chapter 4, we consider an alternative procedure for controlling the EWMA chart's performance proposed by Jones and Steiner (2012) and Gandy and Kvaløy (2013). This procedure is based on bootstrapping to guarantee, with a specified probability, a certain conditional performance for each chart. The main objective of this approach is to limit the proportion of EWMA charts with a low actual in-control ARL.

Furthermore, in Chapter 5, we compare three designs for the EWMA chart for dispersion. *We use the effect of sampling variation on the performance as a comparison metric*.

The main objective of this comparison is to give a recommendation on which EWMA chart for dispersion to use if parameters need to be estimated.

6.4 Results

The results show that data contaminations can have a huge impact on the accuracy of the estimates obtained in Phase I. Existing estimation methods provide accurate estimates for *specific* patterns of contaminations. The new method, based on EWMA charting, provides accurate estimates for *any* of the considered contamination scenarios.

Furthermore, we show that to set up EWMA charts, more Phase I data are required than previously recommended, in order to decrease the sampling variability to a reasonable level. Moreover, we found that sampling variation has a larger effect on the EWMA chart for dispersion than on the EWMA chart for location.

Also, our results show that using a bootstrap-based design approach, results in highly skewed in-control ARL distributions. However, such increases of the in-control ARL did not have much of an effect on the out-of-control performance.

Finally, we found that the EWMA chart based on the sample variance (S^2) has the most predictable performance, compared to the charts based on the sample standard deviation (S), or on the logarithm of the sample variance ($\ln S^2$).

6.5 Discussion and recommendations

We recommend the **use of estimation methods based on screening** if contaminations may be present in Phase I. These methods are efficient under stability and robust to contaminations. We recommend to implement the robust estimation as follows: (1) first screen the initial data to delete the contaminated observations and 'learn from the data'; (2) perform the screening with an EWMA chart designed with a smoothing constant around 0.5. This balances the ability to detect outliers and sustained shifts; (3) use a two-step procedure, namely a robust estimator to construct the initial chart, and an efficient estimator for post-screening estimation.

Automatic implementation of these methods would be unwise. As noted by Jones-Farmer et al. (2014b) careful consideration prior to eliminating process observations in a Phase I analysis is important. The method can be used to signalling those samples which need to be investigated. Furthermore, it ensures that - even if we oversee a contaminated observation - the resulting estimates are near to the target.

A limitation of the proposed method for location is its inaccuracy if single scattered outliers (diffuse disturbances) are present. Furthermore, we have only compared the estimation methods for independent and normally distributed in-control data.

We recommend to **take the effect of sampling variability into account** when choosing and designing EWMA control charts, as parameter estimation has a large impact on performance. Furthermore, we **support the use of the bootstrap-based design** approach of Jones and Steiner (2012) and Gandy and Kvaløy (2013), which was recently proposed for controlling the probability of the in-control ARL being at least a specified value. Charts based on bootstrap show less frequent false alarms and are still able to detect sustained shifts quickly.

A limitation of the bootstrap-based design is its relative complexity; it requires advanced statistical knowledge to implement the design. Another limitation is that we have only considered the parametric bootstrap method for independent and normally distributed data. Gandy and Kvaløy (2013) also proposed a non-parametric version of the bootstrap algorithm. The implementation of this method for the EWMA chart is left for future study.

We recommend the **use of the EWMA chart based on the sample variance** (S^2). Because, it has the lowest variability in conditional ARL performance compared to the other considered EWMA charts for dispersion.

Throughout this dissertation, we primarily considered data collected in samples of size $n = 5$. The methods to estimate and monitor the dispersion, as presented in Chapter 3 and 5, are not easily implemented for $n = 1$. As the standard deviation estimators are based on the assumption that $n > 1$. However, we feel confident that the results will hold for other sample sizes $n > 1$ and can be modified to $n = 1$.

References

Abbas, N., Riaz, M., and Does, R. J. M. M. (2013). CS-EWMA Chart for Monitoring Process Dispersion. *Quality and Reliability Engineering International*, 29(5):653–663.

Abbasi, S. A. (2010). On Sensitivity of EWMA Control Chart for Monitoring Process Dispersion. In *Proceedings of the World Congress on Engineering*, 3:2027–2032.

Albers, W. and Kallenberg, W. C. M. (2004). Are Estimated Control Charts in Control? *Statistics*, 38(1):67–79.

Albers, W. and Kallenberg, W. C. M. (2005). New Corrections for Old Control Charts. *Quality Engineering*, 17(3):467–473.

Amiri, A. and Allahyari, A. (2012). Change Point Eestimation Methods for Control Chart Postsignal Diagnostics: A Literature Overview. *Quality and Reliability Engineering International*, 28(7):673–685.

Bajgier, S. M. (1992). The Use of Bootstrapping to Construct Limits on Control Charts. In *Proceedings of the Decision Science Institute*, 1:1611–1613.

Bisgaard, S. (2012). The Future of Quality Technology: From a Manufacturing to a Knowledge Economy & From Defects to Innovations. *Quality Engineering*, 24(1):30–36.

Box, G. E. P. (1953). Non-Normality and Tests on Variances. *Biometrika*, 40(3/4):318–335.

Box, G. E. P., Hunter, W., and Hunter, J. (1978). *Statistics for Experimenters*. John Wiley and Sons, New York, NY, 1st edition.

Box, G. E. P., Hunter, W., and Hunter, J. (2005). *Statistics for Experimenters*. John Wiley and Sons, New York, NY, 2nd edition.

Capizzi, G. and Masarotto, G. (2013). Phase I Distribution-Free Analysis of Univariate Data. *Journal of Quality Technology*, 45(3):273–284.

Castagliola, P. (2005). A New S^2-EWMA Control Chart for Monitoring the Process Variance. *Quality and Reliability Engineering International*, 21(8):781–794.

Chakraborti, S., Human, S. W., and Graham, M. A. (2009). Phase I Statistical Process Control Charts: An Overview and Some Results. *Quality Engineering*, 21(1):52–62.

Chatterjee, S. and Qiu, P. (2009). Distribution-Free Cumulative Sum Control Charts Using Bootstrap-Based Control Limits. *The Annals of Applied Statistics*, 3(1):349–369.

Chen, A. and Elsayed, E. A. (2002). Design and Performance Analysis of the Exponentially Weighted Moving Average Mean Estimate for Processes Subject to Random Step Changes. *Technometrics*, 44(4):379–389.

Chen, G. (1997). The Mean and Standard Deviation of the Run Length Distribution of $\bar{X}$ Charts when Control Limits are Estimated. *Statistica Sinica*, 7:789–798.

Crowder, S. V. (1987). A Simple Method for Studying Run-Length Distributions of Exponentially Weighted Moving Average Charts. *Technometrics*, 29(4):401–407.

Crowder, S. V. (1989). Design of Exponentially Weighted Moving Average Schemes. *Journal of Quality Technology*, 21(3):155–155.

Crowder, S. V. and Hamilton, M. D. (1992). An EWMA for Monitoring a Process Standard Deviation. *Journal of Quality Technology*, 24(1):12–21.

De Mast, J., Does, R. J. M. M., De Koning, H., and Lokkerbol, J. (2012). *Lean Six Sigma for Services and Healthcare*. Beaumont Quality Publications, Alphen aan de Rijn, the Netherlands.

Does, R. J. M. M., Roes, K. C. B., and Trip, A. (1999). *Statistical Process Control in Industry; Implementation and Assurance of SPC*. Kluwer Academic, Dordrecht.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1):1–26.

Epprecht, K. E., Loureiro, L. D., and Chakraborti, S. (2015). Effect of the Amount of Phase I Data on the Phase II Performance of S^2 and S Control Charts. *Journal of Quality Technology*, 47(2):135–155.

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

Fraker, S. E., Woodall, W. H., and Mousavi, S. (2008). Performance Metrics for Surveillance Schemes. *Quality Engineering*, 20(4):451–464.

Frisén, M. (2009). Optimal Sequential Surveillance for Finance, Public Health, and Other Areas (with discussion). *Sequential Analysis: Design Methods and Applications*, 28(3):310–337.

Gandy, A. and Kvaløy, J. T. (2013). Guaranteed Conditional Performance of Control Charts via Bootstrap Methods. *Scandinavian Journal of Statistics*, 40(4):647–668.

Goedhart, R., Schoonhoven, M., and Does, R. J. M. M. (2015). Guaranteed In-Control Performance of the Shewhart X and $\bar{X}$ Control Charts. *Submitted for publication*.

Hawkins, D. M. and Wu, Q. (2014). The CUSUM and the EWMA Head-to-Head. *Quality Engineering*, 26(2):215–222.

Hunter, J. S. (1986). The Exponentially Weighted Moving Average. *Journal of Quality Technology*, 18(4):203–210.

Janacek, G. J. and Meikle, S. E. (1997). Control Charts Based on Medians. *The Statistician*, 46(1):19–31.

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. (2002). The Statistical Design of EWMA Control Charts with Estimated Parameters. *Journal of Quality Technology*, 34(3):277–288.

Jones, L. A., Champ, C. W., and Rigdon, S. E. (2001). The Performance of Exponentially Weighted Moving Average Charts with Estimated Parameters. *Technometrics*, 43(2):156–167.

Jones, L. A. and Woodall, W. H. (1998). The Performance of Bootstrap Control Charts. *Journal of Quality Technology*, 30(4):362–375.

Jones-Farmer, L. A., Ezell, J. D., and Hazen, B. T. (2014a). Applying Control Chart Methods to Enhance Data Quality. *Technometrics*, 56(1):29–41.

Jones-Farmer, L. A., Woodall, W. H., Steiner, S. H., and Champ, C. W. (2014b). An Overview of Phase I Analysis for Process Improvement and Monitoring. *Journal of Quality Technology*, 46(3):265–280.

Jones, M. A. and Steiner, S. H. (2012). Assessing the Effect of Estimation Error on Risk-Adjusted CUSUM Chart Performance. *International Journal for Quality in Health Care*, 24(2):176–181.

Knoth, S. (2005). Accurate ARL Computation for EWMA-S^2 Control Charts. *Statistics and Computing*, 15(4):341–352.

Knoth, S. (2010). Control Charting Normal Variance-Reflections, Curiosities, and Recommendations. In *Frontiers in Statistical Quality Control*, 9:3–18.

Knoth, S. (2014). SPC: Statistical Process Control - Collection of Some Usefull Functions. *R package version 3.2.2*.

Knoth, S. (2015). Run Length Quantiles of EWMA Control Charts Monitoring Normal Mean or/and Variance. *International Journal of Production Research*, 53(15):4629–4647.

Langenberg, P. and Iglewicz, B. (1986). Trimmed Mean $\bar{X}$ and R Charts. *Journal of Quality Technology*, 18(2):152–161.

Lawless, J. F., Crowder, M. J., and Lee, K. A. (2012). Monitoring Warranty Claims With Cusums. *Technometrics*, 54(3):269–278.

Lee, J., Wang, N., Xu, L., Schuh, A., and Woodall, W. H. (2013). The Effect of Parameter Estimation on Upper-Sided Bernoulli Cumulative Sum Charts. *Quality and Reliability Engineering International*, 29(5):639–651.

Liu, R. Y. and Tang, J. (1996). Control Charts for Dependent and Independent Measurements based on Bootstrap Methods. *Journal of the American Statistical Association*, 91(436):1694–1700.

Lucas, J. M. and Saccucci, M. S. (1990). Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements. *Technometrics*, 32(1):1–12.

MacCarthy, B. L. and Wasusri, T. (2002). A Review of Non-Standard Applications of Statistical Process Control (SPC) Charts. *International Journal of Quality & Reliability Management*, 19(3):295–320.

Mahmoud, M. A., Henderson, G. R., Epprecht, E. K., and Woodall, W. H. (2010). Estimating the Standard Deviation in Quality-Control Applications. *Journal of Quality Technology*, 42(4):348–357.

Maravelakis, P. E. and Castagliola, P. (2009). An EWMA Chart for Monitoring the Process Standard Deviation when Parameters are Estimated. *Computational Statistics & Data Analysis*, 53(7):2653–2664.

Montgomery, D. C. (2013). *Introduction to Statistical Quality Control*. John Wiley & Sons, Hoboken, NJ., 7th edition.

Münz, G. and Carle, G. (2008). Application of Forecasting Techniques and Control Charts for Traffic Anomaly Detection. In *Proc. 19th ITC Specialist Seminar on Network Usage and Traffic, Berlin, Germany*.

Ng, C. H. and Case, K. E. (1989). Development and Evaluation of Control Charts using Exponentially Weighted Moving Averages. *Journal of Quality Technology*, 21(4):242–250.

Page, E. S. (1954). Continuous Inspection Schemes. *Biometrika*, 41(1/2):100–115.

Psarakis, S., Vyniou, A. K., and Castagliola, P. (2014). Some Recent Developments on the Effects of Parameter Estimation on Control Charts. *Quality and Reliability Engineering International*, 30(8):1113–1129.

Quesenberry, C. P. (1993). The Effect of Sample Size on Estimated Limits for $\bar{X}$ and X Control Charts. *Journal of Quality Technology*, 25(4):237–247.

Roberts, S. W. (1959). Control Chart Tests Based on Geometric Moving Averages. *Technometrics*, 1(3):239–250.

Robinson, P. B. and Ho, T. Y. (1978). Average Run Lengths of Geometric Moving Average Charts by Numerical Methods. *Technometrics*, 20(1):85–93.

Rocke, D. M. (1989). Robust Control Charts. *Technometrics*, 31(2):173–184.

Rocke, D. M., Downs, G. W., and Rocke, A. J. (1982). Are Robust Estimators Really Necessary? *Technometrics*, 24(2):95–101.

Saleh, N. A., Mahmoud, M. A., and Abdel-Salem, G. A. S. (2013). The Performance of the Adaptive Exponentially Weighted Moving Average Control Chart with Estimated Parameters. *Quality and Reliability Engineering International*, 29(4):595–606.

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

Saleh, N. A., Mahmoud, M. A., Keefe, M. J., and Woodall, W. H. (2015b). The Difficulty in Designing Shewhart $\bar{X}$ and X Control Charts with Estimated Parameters. *Journal of Quality Technology*, 47(2):127–138.

Schoonhoven, M. and Does, R. J. M. M. (2012). A Robust Standard Deviation Control Chart. *Technometrics*, 54(1):73–82.

Schoonhoven, M., Nazir, H. Z., Riaz, M., and Does, R. J. M. M. (2011a). Robust Location Estimators for the $\bar{X}$ Control Chart. *Journal of Quality Technology*, 43(4):363–379.

Schoonhoven, M., Riaz, M., and Does, R. J. M. M. (2011b). Design and Analysis of Control Charts for Standard Deviation with Estimated Parameters. *Journal of Quality Technology*, 43(4):307–333.

Seppala, T., Moskowitz, H., Plante, R., and Tang, J. (1995). Statistical Process Control via the Subgroup Bootstrap. *Journal of Quality Technology*, 27(2):139–153.

Shewhart, W. A. (1926). Quality Control Charts. *Bell System Technical Journal*, 5(4):593–603.

Shewhart, W. A. (1931). *Economic Control of Quality of Manufactured Product*. Van Nostrand Company, New York.

Shewhart, W. A. (1939). *Statistical Method from the Viewpoint of Quality Control*. Graduate School of the Department of Agricultures, Washington, D.C. [Reprinted by Dover Publications, New York, 1986].

Shu, L. and Jiang, W. (2008). A New EWMA Chart for Monitoring Process Dispersion. *Journal of Quality Technology*, 40(3):319–331.

Simões, B. F. T., Epprecht, E. K., and Costa, A. F. B. (2010). Performance Comparisons of EWMA Control Chart Schemes. *Quality Technology and Quantitative Management*, 7(3):249–261.

Spiegelhalter, D., Sherlaw-Johnson, C., Bardsley, M., Blunt, I., Wood, C., and Grigg, O. (2012). Statistical Methods for Healthcare Regulation: Rating, Screening and Surveillance. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 175(1):1–47.

Steiner, S. H. (1999). Exponentially Weighted Moving Average Control Charts with Time-Varying Control Limits and Fast Initial Response. *Journal of Quality Technology*, 31(1):75–86.

Stoumbos, Z. G., Reynolds Jr, M. R., Ryan, T. P., and Woodall, W. H. (2000). The State of Statistical Process Control as we Proceed into the 21st Century. *Journal of the American Statistical Association*, 95(451):992–998.

Sullivan, J. H. and Woodall, W. H. (1996). A Control Chart for Preliminary Analysis of Individual Observations. *Journal of Quality Technology*, 28(3):265–278.

Tatum, L. G. (1997). Robust Estimation of the Process Standard Deviation for Control Charts. *Technometrics*, 39(2):127–141.

Thottan, M. and Ji, C. (2003). Anomaly Detection in IP Networks. *IEEE Transactions on Signal Processing*, 51(8):2191–2204.

Tsui, K.-L., Chiu, W., Gierlich, P., Goldsman, D., Liu, X., and Maschek, T. (2008). A Review of Healthcare, Public Health, and Syndromic Surveillance. *Quality Engineering*, 20(4):435–450.

Tukey, J. W. (1960). *A Survey of Sampling from Contaminated Distributions in Contributions to Probability and Statistics.* Chapter in 'Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling' . Stanford University Press.

Vardeman, S. B. (1999). A Brief Tutorial on the Estimation of the Process Standard Deviation. *IIE transactions*, 31(6):503–507.

Vining, G. (2009). Technical Advice: Phase I and Phase II Control Charts. *Quality Engineering*, 21(4):478–479.

Vining, G., Kulahci, M., and Pedersen, S. (2015). Recent Advances and Future Directions for Quality Engineering. *Quality and Reliability Engineering International*, DOI: 10.1002/qre.1797.

Wang, T., Li, Y., and Cui, H. (2007). On Weighted Randomly Trimmed Means. *Journal of Systems Science and Complexity*, 20(1):47–65.

Woodall, W. H. (2006). The Use of Control Charts in Health-Care and Public-Health Surveillance (with discussion). *Journal of Quality Technology*, 38(2):89–104.

Woodall, W. H. and Mahmoud, M. A. (2005). The Inertial Properties of Quality Control Charts. *Technometrics*, 47(4):425–436.

Woodall, W. H. and Montgomery, D. C. (2014). Some Current Directions in the Theory and Application of Statistical Process Monitoring. *Journal of Quality Technology*, 46(1):78–94.

Woodall, W. H., Zhao, M., Paynabar, K., Sparks, R., and Wilson, J. D. (2015). An Overview and Perspective on Social Network Monitoring. *Manuscript in preparation*.

Zhang, M., Megahed, F. M., and Woodall, W. H. (2014). Exponential CUSUM Charts with Estimated Control Limits. *Quality and Reliability Engineering International*, 30(2):275–286.

Zhang, M., Peng, Y., Schuh, A., Megahed, F. M., and Woodall, W. H. (2013). Geometric Charts with Estimated Control Limits. *Quality and Reliability Engineering International*, 29(2):209–223.

Zwetsloot, I. M. (2012). A Robust EWMA Control Chart. *MPhil Theses Series 158, Tinbergen Institute, Amsterdam*.

Zwetsloot, I. M. (2015). A Comparison of EWMA Control Charts for Dispersion Based on Estimated Parameters. *Manuscript in preparation*.

Zwetsloot, I. M. and Does, R. J. M. M. (2015). Quality Quandaries: Improving Revenue by Attracting More Clients Online. *Quality Engineering*, 27(1):130–139.

Zwetsloot, I. M., Schoonhoven, M., and Does, R. J. M. M. (2014). A Robust Estimator of Location in Phase I Based on an EWMA Chart. *Journal of Quality Technology*, 46(4):302–316.

Zwetsloot, I. M., Schoonhoven, M., and Does, R. J. M. M. (2015a). A Robust Phase I Exponentially Weighted Moving Average Control Chart for Dispersion. *Quality and Reliability Engineering International*, 31(6):989–999.

Zwetsloot, I. M., Schoonhoven, M., and Does, R. J. M. M. (2015b). Robust Point Location Estimators for the EWMA Control Chart. *Quality Technology and Quantitative Management*, to appear.

Zwetsloot, I. M. and Woodall, W. H. (2015). A Head-to-Head Comparative Study of the Conditional Performance of Control Charts based on Estimated Parameters. *Manuscript in preparation*.

EWMA Control Charts in
Statistical Process Monitoring

Samenvatting

EWMA Regelkaarten in Statistische Procesmonitoring

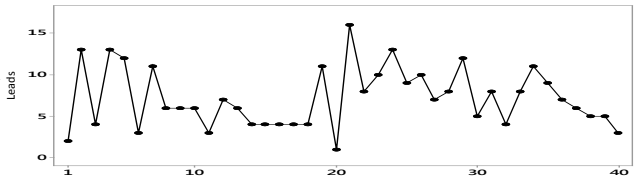
De hoeveelheid beschikbare data neemt alsmaar toe. Het detecteren van veranderingen in datastromen is vaak van belang. Statistische procesmonitoring biedt een aantal instrumenten om verandering in data te detecteren. Eén van deze instrumenten is de regelkaart. Het onderwerp van dit proefschrift is een speciale regelkaart, de EWMA regelkaart, die gebaseerd is op het exponentieel voortschrijdende gemiddelde van de data.

In deze samenvatting beschrijven we eerst de concepten van een regelkaart en vervolgens van de EWMA regelkaart. Daarna vatten we de bevindingen van dit proefschrift samen in twee gedeelten - corresponderend met de opsplitsing van dit proefschrift. We sluiten af met een conclusie.

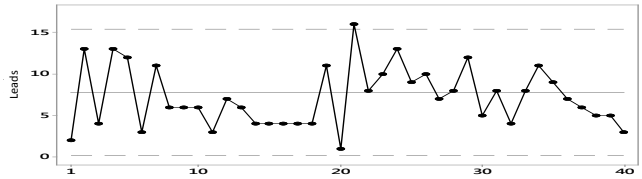
De regelkaart: een voorbeeld

Stelt u zich voor: u werkt bij een onderwijsinstituut en bent verantwoordelijk voor het rekruteren van deelnemers aan een programma. Veel potentiële deelnemers vinden uw instituut online en bezoeken vervolgens de website. Hier kan deze potentiële deelnemer een brochure aanvragen via een online formulier. Vanuit de database van de website houdt u het aantal brochure-aanvragen per week bij: de zogeheten 'leads'. Figuur NL.1(a) laat een tijdgrafiek zien van het aantal leads in de eerste veertig weken van 2015. Wat kunnen we leren van deze grafiek? Bent u ongerust - of zou u het moeten zijn - over het lage aantal leads in de laatste week?

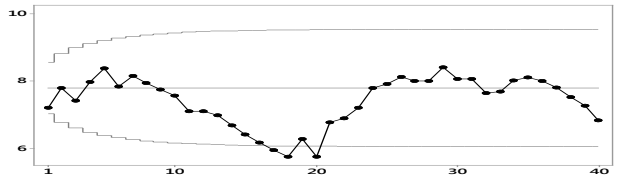
De methoden en technieken van statistische procesmonitoring (ook statistische procesbeheersing genoemd) geven antwoord op dit soort vragen. Dit vakgebied begon met de introductie van de regelkaart door Walter A. Shewhart (op 16 mei 1924). De regelkaart is een grafische weergave van de proceskarakteristiek (in het voorbeeld het aantal leads) tezamen met twee 'regelgrenzen'. Deze grenzen laten de gebruiker onderscheid maken tussen 'normale' procesvariatie en variatie veroorzaakt door



(a) *Tijdgrafiek*



(b) *Shewhart regelkaart*



(c) *EWMA regelkaart*

Figuur NL.1: *Voorbeelden van regelkaarten*

'speciale' oorzaken. Figuur NL.1(b) laat een voorbeeld van een *Shewhart regelkaart* zien. Het signaal in week 21 duidt erop dat er in week 21 misschien iets anders gebeurde dan in de andere weken.

Een regelkaart functioneert goed, wanneer de kaart 'speciale' oorzaken van variatie snel detecteert en dan een zogeheten 'out-of-control' signaal geeft. Hierbij is snelheid van detectie van belang; dit wordt gekwantificeerd met behulp van de 'Average Run Length (ARL)', oftewel de gemiddelde tijd om tot een signaal te komen. Daarnaast stelt men altijd een 'in-control' ARL in; dit is de gemiddelde tijd om tot een signaal te komen als het proces normaal functioneert, dus in feite de tijd om tot een vals alarm te komen. Een regelkaart presteert goed als de 'in-control ARL' lang is en de 'out-of-control ARL' kort.

De EWMA regelkaart

Een beperking van de Shewhart regelkaart is dat deze kleine veranderingen in het proces moeilijk detecteert en dus een (te) hoge out-of-control ARL heeft bij kleine verschuivingen in het proces. Om dit probleem op te lossen ontwikkelde Roberts in 1959 de EWMA regelkaart. Deze kaart detecteert kleine verschuiving sneller, doordat deze regelkaart gebruik maakt van zowel de huidige meting als van de voorafgaande

metingen. Figuur NL.1(c) laat een EWMA regelkaart zien met een mogelijke verschuiving in week 17 en week 20.

De EWMA regelkaart geeft gewicht aan alle data; hoe verder de data in het verleden liggen hoe kleiner het gewicht. Wiskundig bestaat de EWMA regelkaart uit de 'EWMA grootheid' Z_i, gedefinieerd als $Z_i = (1-\lambda)Z_{i-1} + \lambda M_i$, op tijdstip i. Hier is M_i de maatstaaf van het proces waarin men geïnteresseerd is, zoals bijvoorbeeld het gemiddelde of de spreiding van de proces karakteristiek. The factor λ is de wegingsconstante (smoothing constant). Deze constante λ wordt gekozen tussen 0 en 1: een waarde van 1 betekent dat alleen gewicht wordt toegekend aan de huidige data. Hoe kleiner λ wordt gekozen, hoe meer gewicht data in het verleden krijgt.

De EWMA kaart signaleert een verschuiving in het proces zodra Z_i één van de twee regelgrenzen overschrijdt. Om de regelgrenzen te kunnen bepalen, dienen eerst het gemiddelde en de standaarddeviatie (spreiding) van de proceskarakteristiek te worden vastgesteld. In de praktijk zijn deze parameters vaak niet bekend en dienen ze te worden geschat. Het is daarom gebruik dat een regelkaart in twee fasen wordt geïmplementeerd. In *Fase I* worden de proces parameters geschat op basis van historische data. Waarna in *Fase II* deze schattingen worden gebruikt om de regelgrenzen op te stellen zodat het proces gemonitord kan worden.

In dit proefschrift bestuderen we de EWMA regelkaart gebaseerd op schattingen van de proces parameters. Dit proefschrift bestaat uit twee delen: in Fase I (Hoofdstukken 2 en 3) bestuderen we het schatten van de parameters. In Fase II (Hoofdstukken 4 en 5) bestuderen we de prestatie van de EWMA regelkaart gebaseerd op deze schatters.

Fase I: schatten op basis van vervuilde data

Data vervuiling zorgt voor slechte prestaties van de regelkaart. In de eerste fase van dit proefschrift (Hoofdstukken 2 en 3) wordt gekeken naar *het schatten van parameters gebaseerd op data welke vervuilde waarnemingen kan bevatten*. In de praktijk is een dataset vaak niet perfect representatief voor het 'normale' proces. Deze Fase I dataset kan bijvoorbeeld verstoringen, meetfouten, uitschieters of andere data vervuilingen bevatten. Schattingen van parameters op basis van vervuilde Fase I data zijn vaak onzuiver en dit kan resulteren in een regelkaart met een statistische prestatie die minder voorspelbaar is.

Robuuste schatters als alternatief. Als oplossing voor vervuilde data in Fase I bestuderen we *robuuste schatters*. De term robuust betekent dat een schatter het effect van vervuilde data op de uiteindelijke schatting verlaagt. We bestuderen en vergelijken verschillende schattingsmethoden, zowel bestaande als een nieuwe methode. We vergelijken hoe robuust deze methoden zijn bij verschillende scenario's met vervuiling in Fase I. Daarnaast kijken we ook of onder zuivere Fase I data de schatters correcte en precieze ('efficiënte') schattingen opleveren.

EWMA regelkaart als robuuste schatter geeft consistente schattingen. De resultaten laten zien dat bestaande robuuste schattingsmethoden accurate schattingen opleveren voor *specifieke* scenario's met vervuiling. De nieuw geïntroduceerde schat-

tingsmethode, gebaseerd op een EWMA regelkaart, geeft accurate schattingen voor *alle* bestudeerde scenario's, maar is niet optimaal bij een specifiek scenario.

Fase II: effect van schatting op prestatie

Prestatie regelkaart gebaseerd op geschatte parameters slechter. In het tweede gedeelte van dit proefschrift (Hoofdstukken 4 en 5) kijken we naar het *effect van schatten op de prestatie van de regelkaart in Fase II*. Vanuit de literatuur is bekend dat dit effect groot kan zijn. Doorgaans wordt de prestatie van een regelkaart, gebaseerd op Fase I schattingen, geëvalueerd op basis van de gemiddelde prestatie. Daarbij wordt gemiddeld over alle mogelijke Fase I schattingen. Echter de prestatie van de regelkaart hangt af van de 'eigenlijk' geschatte waarde. Deze conditionele prestatie kan heel erg afwijken van de 'gemiddelde' prestatie. Het is de variabiliteit in de conditionele prestatie die Hoofdstukken 4 en 5 motiveert.

Gebruik de conditionele prestatie. Om het effect van schatten op de prestatie van de regelkaart in Fase II te bestuderen, kijken we naar de variabiliteit in de conditionele prestatie van de ARL. Daarnaast bekijken we in Hoofdstuk 4 een alternatieve procedure om de regelkaart op te zetten, gebaseerd op de conditionele prestatie. Deze methode, wordt gebaseerd op het 'bootstrappen' van de Fase I data.

In Hoofdstuk 5, vergelijken we drie verschillende EWMA regelkaart voor het monitoren van de spreiding. *We vergelijken de regelkaarten op basis van de conditionele prestatie*, met als doel een advies te geven over welke regelkaart te gebruiken indien parameters geschat zijn in Fase I.

Bootstrapmethode is een goed alternatief. De resultaten van Hoofdstuk 4 laten zien dat het effect van schattingen op de EWMA regelkaart groter is dan voorheen aangetoond. We raden aan om een regelkaart op te zetten met de bootstrapmethode waarmee het aantal regelkaarten dat vroegtijdig een vals alarm geeft beperkt kan worden.

Gebruik EWMA regelkaart voor de spreiding op basis van steekproef variantie In Hoofdstuk 5 bestuderen we de EWMA regelkaart voor de spreiding. We laten zien dat het effect van schattingen op de prestatie van deze kaart groot is. Het is zelfs groter dan voor het gemiddelde (zoals bestudeerd in Hoofdstuk 4). We raden aan om de EWMA kaart op basis van de steekproef variantie (S^2) te gebruiken, omdat deze de meest voorspelbare prestatie laat zien.

Aanbevelingen, discussie en conclusie

We hebben het effect van data vervuiling in Fase I bestudeerd. We raden aan om een robuuste schatter gebaseerd op het filteren van de data te gebruiken indien vervuiling aanwezig kan zijn. Deze methoden zijn efficiënt onder zuivere data en robuust voor verstoringen. We raden de volgende procedure aan: (1) filter de data en verwijder de vervuilde observaties; (2) gebruik een EWMA regelkaart met een wegingsconstante van

ongeveer 0.5 voor de filtering; (3) gebruik een twee-staps procedure, waarbij eerst een robuuste schatter wordt gebruikt voor de initiële regelkaart en daarna een efficiënte schatter voor de data na filtering.

Een beperking is dat deze methode niet goed werkt bij het scenario met diffuus verspreide verstoringen. Verder is een beperkende factor dat we alleen data beschouwen welke onafhankelijk en normaal verdeeld zijn.

Verder hebben we laten zien dat het effect van een schatter groot is op de prestatie van de EWMA regelkaart. Daarom is het gebruik van de conditionele ARL aan te raden. Hierbij wordt het effect van Fase I schattingen op de Fase II prestatie namelijk heel inzichtelijk. We ondersteunen het gebruik van de bootstrapmethode voor het opzetten van de EWMA regelkaart. Op deze manier kan het aantal valse alarmen dat de regelkaart geeft beperkt worden.

Een beperking van de bootstrapmethode is dat deze relatief complex is. Verder hebben we alleen gekeken naar de parametrische bootstrap voor onafhankelijke en normaal verdeelde data.

We hebben een vergelijking gemaakt tussen drie EWMA regelkaarten voor de spreiding. Wanneer deze regelkaarten gebaseerd zijn op schattingen, is de prestatie van deze kaarten verschillend. We raden aan de EWMA kaart op basis van de steekproef variantie te gebruiken. Deze kaart geeft namelijk, in vergelijking met de andere twee, de kleinste variabiliteit in de (conditionele) prestatie.

Acknowledgements

It is unbelievable; I have written a book. Two years of research master and three years of PhD wrapped up together in a hundred pages. It makes me smile from ear to ear. It would have been impossible without the help and support of many.

Marit, your persistence and patience have learned me how to write and your always constructive questions have encouraged me to keep (re)evaluating my work. *Ronald*, you have always baffled me by being able to pinpoint the main question, problem, or contribution after a single read. Your cheerful confidence together with your ability to push me just outside of my comfort zone, made me work very effective with a lot of join and laughter. I would like to thank you both for all your time, effort, inspiration, and laughs.

I would like to thank all the other people who co-authored papers with me: *prof.dr. Bill Woodall, prof.dr. Allison Jones-Farmer, prof.dr. Mahmoud Mahmoud, Nesma Saleh, Marly Buitenhuis, Bart Lameijer, and Yara Basta.* Working with such a variety of enthusiastic and motivated people was very inspiring, thank you.

Working at IBIS UvA has been a great experience. It is a very stimulating and fun environment to do research, teaching, and consulting. I would like to thank all my (former) colleagues - *Alex, Atie, Benjamin, Jeroen, Marit, Rob, Ronald, Tashi, and Thomas*- for this wonderful time.

Laura, Eszter, Karo, Chris, Patrick, Thomas, and Alex, the years at Tinbergen and IBIS UvA have been a great experience due to your friendship. Thank you for the hours of study, work, teaching, lunches, inspirational Monday diners, sate's at Koosjes, and our Friday afternoon drinks (even if it wasn't Friday).

This book would not have had such a fitting cover without the networking skill of *Kasper* and the enthusiastic help of *André Verhoek*. Thank you both.

My PhD at IBIS UvA started with two coincidences. First, I stumbled into science and data collection under supervision of prof. dr. *Maarten Pieter Schinkel*; he encouraged me to go into academia. Next, I bumped into *Benjamin Kemper* at Café de Roeter; he introduced me to IBIS UvA. Both have, probably accidentally, unknowingly, and coincidentally, brought me to start my PhD. Thank you.

Finally, three people deserve a special thanks. These go to *Doreth* for your never

ending trust in my abilities. To *Sabine (Nala)* for making home a cozy place and feeding me throughout the last few months of endless work. And to, of course, *Wim Jan*. Even though you sometimes got frustrated by not grasping all the statistics in this book, your never ending interest in my work is an enormous gift, thank you.

It is not just this book that makes me happy. Life outside of work and research is enormously good. This is essentially due to all my loving family and friends: *mama, papa, Willem, Janneke, Oma and Opa, Wim Jan, Doreth, Hanna (Adje), Annabelle (Tommie), Sabine (Nala)*, and so many more, thank you!

Though the expression says that 'one should take life into ones own hands', I can't stop marveling at life's lucky coincidences. I would like to finish with a quote (of which I have always hoped that it is true), from Simon Van Booy's collection of stories 'Love Begins in Winter': *"Coincidences mean you're on the right path."*

Inez Zwetsloot
Amsterdam, February 2016

EWMA Control Charts in
Statistical Process Monitoring

Curriculum Vitae

Inez Maria Zwetsloot (1988) holds a MSc degree in Economics with a specialization in Econometrics (2012) from the mPhil research programme at the Tinbergen Institute, the Netherlands. Inez also obtained a BSc in Econometrics (2010) from the University of Amsterdam.

Inez' research is on industrial statistics with a focus on statistical process monitoring and process improvement. This dissertation was written from 2012 to 2015 at the Operations Management department of the Amsterdam Business School of the University of Amsterdam.

During the same period, Inez was also statistical consultant at the Institute for Business and Industrial Statistics, a consultancy firm which is part of the Amsterdam Business School. She trained project leaders (Lean Six Sigma Green and Black Belts) from various organizations in the Netherlands. Also, she was supervisor of (Lean Six Sigma) improvement projects.

At the Amsterdam Business School, Inez taught (parts of) various courses in applied statistics and process improvement. She was lecturer and coordinator for the course Quantitative Methods for the regular MBA at the Amsterdam Business School.

In addition to her interest in the application of statistics, Inez is an active player of field hockey and badminton. Furthermore, she has a taste for traveling across the world.

The **Institute for Business and Industrial Statistics** (IBIS UvA) is a part of the Amsterdam Business School of the University of Amsterdam. It is a center of expertise in Lean Six Sigma and industrial statistics. It supports quality and efficiency improvement initiatives in Dutch and international organizations. The interaction between scientific research on the one hand, and the application of technology in business and industry lies at the core of the institute.

IBIS UvA