NEW VARIABLE PARAMETERS CHART BASED ON AUXILIARY INFORMATION AND MULTIVARIATE CHARTS FOR SHORT PRODUCTION RUNS

CHONG NGER LING

UNIVERSITI SAINS MALAYSIA

2019

NEW VARIABLE PARAMETERS CHART BASED ON AUXILIARY INFORMATION AND MULTIVARIATE CHARTS FOR SHORT PRODUCTION RUNS

by

CHONG NGER LING

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

November 2019

ACKNOWLEDGEMENT

First and foremost, I would like to extend my sincere gratitude to my supervisor Prof. Michael Khoo Boon Chong for his continuous guidance, encouragement and insightful suggestions throughout my Ph.D. journey and the completion of my thesis. I am grateful to Prof. Michael Khoo for sharing his immense knowledge on quality control with me and for always providing me valuable feedbacks in my journey towards a Ph.D. Without his guidance and unwavering support, the successful completion of my Ph.D. would not be possible. My sincere thanks goes to Prof. Michael Khoo for giving me the opportunity to work as a Graduate Research Assistant (GRA) which has aided me tremendously especially in the funding of my Ph.D. study.

A very special gratitude goes to my parents, Chong Siew Leng and Chiam Hong Choo, for their constant encouragement and support. Thanks for always being by my side through thick and thin, for constantly believing in me and for supporting my every endeavour. I am eternally grateful to my parents who have always given me the strength to persevere.

Many thanks to the School of Mathematical Sciences (PPSM) for the sponsorship to conference and various workshops which has given me the opportunity to acquire valuable knowledge. I am also grateful for the USM Graduate Assistant (GA) scheme that has provided me the opportunity to work as a GA during the first year of my Ph.D. study. I would like to sincerely acknowledge the Dean (Prof. Hailiza Kamarulhaili), lecturers and staff of PPSM for their kind assistance and support throughout my Ph.D. journey. I also owe my gratitude to my friends for their support, assistance, words of encouragement and for always being there for me. Lastly, I would also like to thank all of those who may have contributed indirectly to the successful completion of my thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Auxiliary information
ARL	Average run length
ARL_0	In-control ARL
ARL_1	Out-of-control ARL
ASI_0	In-control average sampling interval
ASS	Average sample size
ASS_0	In-control ASS
ASS_1	Out-of-control ASS
ATS	Average time to signal
ATS_0	In-control ATS
ATS_1	Out-of-control ATS
CL	Center line
CRL	Conforming run length
cdf	Cumulative distribution function
CUSUM	Cumulative sum
EATS	Expected ATS
$EATS_0$	In-control EATS
$EATS_1$	Out-of-control EATS
EASS	Expected ASS
$EASS_0$	In-control EASS
$EASS_1$	Out-of-control EASS
E(P(I))	Expected P(I)
ETARL	Expected TARL
$ETARL_1$	Out-of-control ETARL
ETSDRL	Expected TSDRL
$ETSDRL_0$	In-control ETSDRL
$ETSDRL_1$	Out-of-control ETSDRL
EWMA	Exponentially weighted moving average

EWMA-AI	EWMA with auxiliary information
FSS	Fixed sample size
LCL	Lower control limit
MCUSUM	Multivariate CUSUM
MEWMA	Multivariate EWMA
pdf	Probability density function
P(*I*)	Probability of getting a signal within the number of scheduled inspections
pmf	Probability mass function
RL	Run length
RS-AI	Run sum with auxiliary information
SAS	Statistical Analysis System
SH-AI	Shewhart with auxiliary information
SDRL	Standard deviation of the run length
$SDRL_0$	In-control SDRL
$SDRL_1$	Out-of-control SDRL
SPC	Statistical Process Control
SYN-AI	Synthetic with auxiliary information
TARL	Truncated average run length
$TARL_0$	In-control TARL
$TARL_1$	Out-of-control TARL
TFT-LCD	Thin-film transistor-liquid crystal display
tpm	Transition probability matrix
TRL	Truncated run length
TSDRL	Truncated standard deviation of the run length
$TSDRL_0$	In-control TSDRL
$TSDRL_1$	Out-of-control TSDRL
UCL	Upper control limit
VSS	Variable sample size
VSI	Variable sampling interval
VSSI	Variable sample size and sampling interval

VSSI-AI	VSSI with auxiliary information
VP	Variable parameters
VP-AI	VP with auxiliary information

LIST OF NOTATIONS

δ	Shift size
$\delta_{\min}$	Minimum shift
$\delta_{\max}$	Maximum shift
$(\delta_{\min}, \delta_{\max})$	Shift interval
n	Sample size
n_L	Large n
n_S	Small n
t	Sampling interval
t_L	Long t
t_S	Short t
ρ	Correlation coefficient
C	Auxiliary variable
S	Study variable
μ_S	Population mean of S
μ_{S0}	In-control population mean
$\bar{S}_v$	Sample mean of S
σ_S^2	Population variance of S
μ_C	Population mean of C
$\bar{C}_v$	Sample mean of C
σ_C^2	Population variance of C
v	Sample number
$X_{S_v}^*$	Regression estimator of μ_S
(S, C)	Bivariate random sample of the study and auxiliary variables
$\Phi(\, . \,)$	Cdf of the standard normal distribution
$\mathbf{s}$	Steady-state probability vector

$\mathbf{I}$	Identity matrix
I	Number of scheduled inspections
L	LCL of the CRL sub-chart
$\mathbf{1}$	Vector with all elements equal to unity
γ	Smoothing constant
Z_v	EWMA statistic of the vth sample
E_h	Midpoint of the hth subinterval
$+U_f$	Positive score for the fth region
$-U_f$	Negative score for the fth region
Q_{+f}	Region f above CL
Q_{-f}	Region f beneath CL
I_1	Central region
I_2	Warning region
I_3	Out-of-control region
Y_v	Charting statistic of the VSSI-AI and VP-AI charts
$\mathbf{t}$	Vector of sampling intervals
$\mathbf{d}$	Steady-state probability vector of the VSSI-AI and VP-AI charts
b	Number of quality characteristics
$\boldsymbol{\mu}_0$	In-control mean vector
$\boldsymbol{\mu}_1$	Out-of-control mean vector
$\boldsymbol{\Sigma}_0$	Covariance matrix
λ	Non-centrality parameter
M	Finite production horizon

CARTA PARAMETER BOLEH BERUBAH BERDASARKAN MAKLUMAT TAMBAHAN DAN CARTA-CARTA MULTIVARIAT UNTUK LARIAN PENGELUARAN PENDEK BAHARU

ABSTRAK

Kebelakangan ini, perusahaan-perusahaan berusaha untuk terus menambahbaik kualiti yang merupakan asas kepada kepuasan pelanggan. Banyak penambahbaikan dalam skema carta kawalan telah dilakukan untuk menambahbaik pemantauan proses. Dalam tesis ini, carta parameter boleh berubah dengan maklumat tambahan (disingkatkan sebagai VP-AI) telah dicadangkan. Carta VP-AI direka bentuk dengan penganggar regresi yang mempunyai ketepatan yang lebih baik disebabkan penggunaan pembolehubah tambahan untuk menganggar min populasi. Dengan menggunakan kaedah rantai Markov, formula masa untuk berisyarat purata (ATS) dan jangkaan ATS (EATS) diterbitkan untuk saiz anjakan yang diketahui dan tidak diketahui. Penemuan menunjukkan bahawa carta VP-AI mengatasi carta VP asas dan merupakan justifikasi integrasi maklumat tambahan untuk menambahbaik kepekaan carta VP. Perbandingan carta VP-AI dengan carta-carta persaingannya menunjukkan bahawa, untuk semua anjakan, prestasi carta VP-AI mengatasi carta Shewhart AI (SH-AI), sintetik AI (SYN-AI), dan saiz sampel dan selang pensampelan boleh berubah AI (VSSI-AI) dengan ketara. Tambahan pula, untuk kebanyakan anjakan, carta VP-AI mempunyai prestasi yang lebih baik berbanding dengan carta-carta purata bergerak berpemberat eksponen AI (EWMA-AI) dan hasil tambah larian AI (RS-AI). Aplikasi carta VP-AI ditunjukkan dengan contoh ilustrasi berdasarkan dataset sebenar. Dalam banyak situasi, proses adalah bersifat multivariat, yang mana lebih daripada satu ciri kualiti perlu dipantau serentak. Selain itu, banyak syarikat telah menggunakan teknik larian pengeluaran pendek untuk menjadi lebih fleksibel dan

khusus. Oleh itu, dalam tesis ini, carta larian pendek saiz sampel tetap (FSS) T^2 telah dibangunkan. Memandangkan keberkesanan pendekatan adaptif, carta larian pendek saiz sampel boleh berubah (VSS) T^2 juga dicadangkan dalam tesis ini dengan mengubah saiz sampel yang diambil berdasarkan maklumat proses lalu. Berdasarkan penemuan yang diperoleh, carta larian pendek VSS T^2 adalah lebih cepat berbanding dengan carta FSS dalam pengesanan keadaan luar kawalan bagi kebanyakan anjakan dan kebaikan prestasi carta terdahulu berbanding dengan carta yang kemudian meningkat dengan bilangan pemeriksaan (I) dan saiz sampel purata dalam kawalan (ASS_0). Oleh sebab serakan taburan panjang larian singkat carta larian pendek VSS T^2 adalah lebih kecil daripada serakan yang sama untuk carta FSS, carta terdahulu mengatasi carta yang kemudian. Tambahan pula, carta larian pendek VSS T^2 mempunyai keberangkalian yang lebih tinggi untuk memberi isyarat dalam I pemeriksaan; oleh itu, ia adalah lebih peka daripada carta FSS. Suatu contoh berdasarkan dataset sebenar yang menunjukkan aplikasi carta larian pendek VSS T^2 telah ditunjukkan dalam tesis ini.

NEW VARABLE PARAMETERS CHART BASED ON AUXILIARY INFORMATION AND MULTIVARIATE CHARTS FOR SHORT PRODUCTION RUNS

ABSTRACT

Contemporarily, enterprises strive to continuously enhance quality which is a basis of customer satisfaction. Numerous advancements to the control charting scheme have been made to enhance process monitoring. In this thesis, the variable parameters chart with auxiliary information (abbreviated as VP-AI) is proposed. The VP-AI chart is designed with a regression estimator that has an improved precision due to the use of auxiliary variable to estimate the population mean. By adopting the Markov chain method, the average time to signal (ATS) and expected ATS (EATS) formulae are derived for known and unknown shift sizes. The findings show that the VP-AI chart prevails over the basic VP chart and justifies the integration of auxiliary information to improve the sensitivity of the VP chart. A comparison of the VP-AI chart with its competing charts shows that, for all shifts, the performance of the VP-AI chart surpasses the Shewhart AI (SH-AI), synthetic AI (SYN-AI) and variable sample size and sampling interval AI (VSSI-AI) charts considerably. Additionally, for most shifts, the VP-AI chart has a superior performance in comparison with the exponentially weighted moving average AI (EWMA-AI) and run sum AI (RS-AI) charts. The application of the VP-AI chart is shown using an illustrative example based on a real dataset. In many situations, the process is multivariate in nature, where more than one quality characteristic has to be monitored simultaneously. Furthermore, many companies have adopted the short production runs technique to be more flexible and specialized. Hence, in this thesis, the fixed sample size (FSS) T^2 short-run chart is developed. In view of the effectiveness of the adaptive approach, the variable sample size (VSS) T^2 short-run chart is further proposed in this thesis by varying the sample size taken according to past process information. Based on the findings, the VSS T^2 short-run chart is quicker than its FSS counterpart in detecting an out-of-control condition for most shifts and the outperformance of the former to the latter increases with the number of inspections (I) and in-control average sample size (ASS_0). As the spread of the truncated run length distribution of the VSS T^2 short-run chart is smaller than its FSS counterpart, the former surpasses the latter. Moreover, the VSS T^2 short-run chart has a higher probability of signaling an alarm within the I inspections; thus, it is more sensitive than its FSS counterpart. An example based on a real dataset that demonstrates the application of the VSS T^2 short-run chart is presented in this thesis.

CHAPTER 1
INTRODUCTION

1.1 Statistical Process Control

Over the years, quality has been regarded by customers as a paramount decision factor when selecting a product or service and there is an ever-increasing demand for continuous improvement of quality. Quality can be defined as products with features that adhere to the requirements of customers and freedom from errors that lead to customer dissatisfaction (Juran & Godfrey, 1999). Consequently, many companies have adopted methods to continuously improve quality which is a key factor in a company's growth and competitiveness. According to Montgomery (2013), the quality of a product can be enhanced by reducing the amount of variation present in the quality characteristic of a product. In fact, companies strive to lower variation of the product's quality characteristic from the target value in order to manufacture products that meet the expectations of customers.

To illustrate the concept of variation, suppose that the quality characteristic studied is the thickness of a blade. As every product manufactured has a certain degree of variation, there are no identical blades. If there is a slight variation in the blade thickness, the variation may not be noticed and has no impact. On the other hand, if there is a considerable variation in the blade thickness, customers will notice the deviation from their expectations and may consider the blade to be unacceptable. The causes of variation can be divided into two categories: (i) common or chance and (ii) assignable or special.

The common causes of variation (i.e. machine deterioration, low-grade raw materials, temperature, lighting) are an inherent part of a process and is always present (Mason & Antony, 2000). In contrast, assignable causes of variation (i.e. measurement error, surges in power, equipment malfunction) have a larger magnitude and are not an inherent part of the process (Mason & Antony, 2000). Additionally, assignable causes of variation arise occasionally and cause the process to operate at a substandard level.

Statistical Process Control (SPC) is an extensively adopted approach where statistical tools are employed to monitor process behaviour and reduce assignable causes of variation. Assignable causes of variation lead to the production of wastes and an increase in cost associated with repairs. SPC ensures that the operation of a process is at its fullest potential and can be implemented on various processes. Mason and Antony (2000) mentioned that the benefits obtained from the implementation of SPC include improved consistency of process output, a more predictable process, enhanced company reputation and lowered costs.

By adopting SPC, more products that conform to specifications are manufactured; thus, minimizing scrap and rework (doing work all over again) which results in a reduction of costs. Specifications refer to the requirements for the quality characteristics of a product or service. More precisely, for a product, specifications are the desired measurements of the quality characteristics. As an illustration, a tyre must meet the required measurements specified in its design so that it can be adequately aligned to its assembly. In terms of service, specifications refer to the maximum time interval required for a service to be provided. For example, the specification for a food delivery is 20 minutes; thus, the food ordered has to be delivered within 20 minutes to ensure customer satisfaction.

Control chart, histogram, check sheet, cause-and-effect diagram, scatter diagram, Pareto chart and defect concentration diagram are the seven major tools of SPC (Montgomery, 2013). Among the seven SPC tools, control chart is the most sophisticated and serves as a visual diagram that can be used to monitor process behaviour. Hence, control chart has been extensively adopted in manufacturing and non-manufacturing sectors as it provides useful insight to practitioners regarding the type of variation (assignable or common) present in a process. By employing control charts, companies can reduce process variation and improvement in the process leads to an enhanced financial and competitive position.

1.2 Background and Applications of Control Charts

In the 1920s, Walter A. Shewhart from Bell Telephone Laboratories developed the Shewhart control chart (Montgomery, 2013). Shewhart discovered that it is paramount to comprehend the causes of variation before any action can be taken to improve a process. The main purpose of using a control chart is to differentiate common and assignable causes of variation which will guide quality practitioners prior to any quality improvement actions. It is essential for a process to be in a state of statistical control such that only common causes of variation are present in a process (Best & Neuhauser, 2006).

The control chart plots the measured quality characteristic obtained from a sample against time or sample number. Following the order in which the sample is obtained, the quality characteristic is plotted on the control chart. Figure 1.1 shows a graphical representation of a typical control chart. The center line is a representation of the average of the quality characteristic used; thus, indicating the center of the process. Meanwhile, a control chart typically has two control limits, which are the upper and lower control limits that help practitioners in making a decision. A process is in statistical control when the plotted points are within the control limits and is out-of-control when a point is plotted beyond the control limits.

The presence of a point beyond the control limits indicates the presence of assignable cause(s); thus, swift corrective actions have to be taken by practitioners to return the process to its in-control condition. Note that a control chart that can detect an out-of-control condition as soon as possible is more effective and allows practitioners to take quicker corrective actions. It is important for practitioners to take corrective actions swiftly so that less wastes will be produced and the incurrence of additional costs will be reduced. A control chart should also have as few false alarms as possible. A false alarm or Type-1 error occurs when a control chart indicates that the process is in an out-of-control condition even though the process is actually in-control.

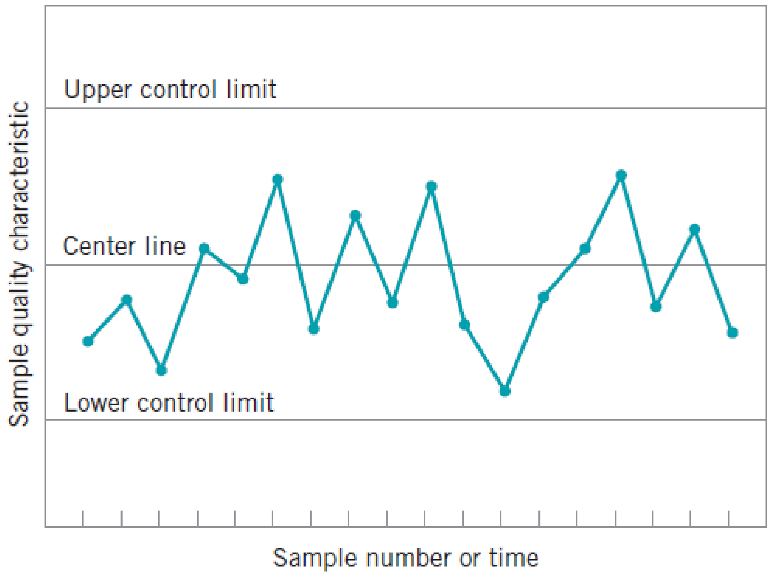

Figure 1.1 A schematic representation of a typical control chart
Source: Montgomery (2013)

Control charts can be divided into two categories (univariate or multivariate) depending on the number of quality characteristics. A control chart is univariate when only one quality characteristic is studied to evaluate product quality. In contrast, a control chart is multivariate in nature when more than one quality characteristic is studied. In practice, more than one quality characteristic is usually studied to assess the quality of a product manufactured. When there is more than one quality characteristic to be monitored, it is more convenient and economical to employ a single multivariate control chart instead of multiple univariate control charts. This is because the number of univariate control charts required will increase with the number of quality characteristics. Furthermore, it is common for product quality to depend on multiple quality characteristics that are correlated. Although the univariate chart is effective in process monitoring and is widely adopted, it does not take the correlation between the quality characteristics into consideration. As the multivariate control chart considers the relationship between variables, it is more suitable in monitoring the correlated quality characteristics of a multivariate process (Topalidou & Psarakis, 2009). As an illustration, a multivariate control chart can be adopted to monitor a chemical process with three correlated variables, i.e. temperature, pressure and flow rates simultaneously.

To date, control charts have been extensively adopted in a wide range of application domains. Researchers have employed control charts in various processes in order to distinguish assignable and common causes of variation, leading to the improvement of numerous processes. Control charts have been primarily applied to the manufacturing process. Other than the manufacturing domain, control charts have also been applied to the healthcare, environmental and service industry domains.

In the manufacturing domain, control charts are used to assist practitioners in decision-making while lowering the amount of scrap, rework and costs. With the utilization of control charts, manufacturing organizations can have a better financial performance and gain a competitive advantage. Maul et al. (1996) used several $\bar{X}$ and R control charts to monitor the variables current, voltage, power and resistance of a gas metal arc welding process. By adopting several control charts to monitor the variables, corrective actions can be taken if the control limits are exceeded by any of the variables; thus, ensuring the quality of weld. In a sputtering process, the uniformity of the sputtering coating thickness is important to ensure the quality of the thin-film transistor-liquid crystal display (TFT-LCD) manufactured. Hence, Yang and Cheng (2008) adopted a multivariate control chart to monitor the mean, range and standard deviation of the sputtering coating thickness in a TFT-LCD manufacturing process.

Control charts have also been employed to improve the quality of healthcare. Marshall and Mohammed (2003) used the Shewhart control chart to study variation in the prescription of antibiotics among general practitioners and provided recommendations for improvement. Moreover, Rogers et al. (2004) discussed the cumulative sum (CUSUM) approach that can be used to study the quality of cardiac surgical performance. For example, a graphical display of the cumulative number of surgical failures against the operation number can be plotted. The cumulative number remains unchanged if the surgery is successful or increases by one if there is a failure. Mohammed et al. (2008) adopted the p chart to monitor the proportion of deaths among patients who had a fractured neck femur.

Control charts also have applications in the environmental domain where they are adopted to study environmental variables. Hunt et al. (1978) used control chart to monitor the measurements of air pollutants. By using control chart, the shifts in the mean and range of the air pollutants can be simultaneously studied and displayed graphically; hence, aiding air pollution control agencies. In addition, Zimmerman et al. (1996) monitored the water quality of an estuary in terms of temperature, salinity and oxygen saturation level. By using control chart to highlight the changes present in the quality characteristic of water, a better decision can be made to improve water quality. The quality of air temperature, vapour pressure and wind speed obtained from a weather station were assessed with a control chart by Eching and Snyder (2003).

In the service industry, control charts are used to improve customer satisfaction level as a result of the service provided. Gardiner and Mitra (1994) showed the use of the $\bar{X}$ and S control charts to study the customer waiting time in the banking industry. The customer waiting time is the quality characteristic measured. Hence, the schedules for the front-counter staff in a bank are arranged such that the customer waiting time must not exceed the maximum value of three minutes set by the bank management. Meanwhile, Zolkepley et al. (2018) adopted the generalized variance control chart to study the variability present in the process of teaching and learning. The midterm exam results of a secondary school were collected and plotted on the chart to monitor the performance of the students.

1.3 Problem Statements

The Shewhart control chart is extensively adopted due its simple implementation which enables quality practitioners to easily employ it to monitor the behavior of a process (Montgomery, 2013). The Shewhart chart is known to be sensitive in detecting large process shifts. However, the Shewhart chart has a limitation of being insensitive to small and moderate shifts as it does not take past process information into consideration. As a measure to overcome this drawback, researchers have developed adaptive control charting schemes that have an enhanced sensitivity to small and moderate shifts.

Traditionally, control charting schemes adopt fixed process parameters (i.e. sample size, sampling interval, control and warning limits). On the other hand, the process parameters adopted for the adaptive control charting scheme can vary according to the position of the previous sample statistic on the chart. Adaptive control charts can be divided into four categories which are the (i) variable sample size (VSS) (ii) variable sampling interval (VSI) (iii) variable sample size and sampling interval (VSSI) and (iv) variable parameters (VP) charts. For the VSS chart, only the sample size varies while the VSI chart varies only the sampling interval. As for the VSSI chart, two parameters (sample size and sampling interval) are allowed to vary. Meanwhile, the VP chart allows all parameters of the chart to alternate between two values depending on the position of the prior sample statistic. In comparison to the Shewhart chart, many research works have discovered that adaptive charts are more sensitive to process shifts owing to the consideration of past process information. Hence, the effectiveness of process monitoring can be considerably improved by employing adaptive charts (Tagaras, 1998).

Another approach to improve process monitoring is by incorporating auxiliary information (AI) into control charts. In a real-life setting, measuring the quality characteristic of interest or study variable may incur high costs and is time-consuming. Thus, it may be difficult to acquire a precise estimator of the study variable's population mean. To improve the accuracy of the estimator, practitioners can employ information from both study and auxiliary variables. Therefore, the concept of integrating an auxiliary or supplementary variable which is correlated to the study variable has been widely employed to enhance the precision of estimators. In fact, the incorporation of auxiliary variable has been adopted in various domains. To illustrate, the auxiliary variable concept can be adopted in the platinum refinery where the amount of platinum metal is the study variable while the amount of another metal is taken as the auxiliary variable (Ahmad et al., 2014). As it may be costly or time-consuming to measure the amount of platinum metal, practitioners can measure and incorporate information from the amount of another metal which is correlated with the former to obtain a more precise estimator. By improving the precision of the estimator, a control chart becomes more sensitive to process shifts; thus, resulting in a more powerful control chart.

Motivated by the effectiveness of adaptive charts in improving process monitoring and the improved sensitivity of control charts through the incorporation of auxiliary information, a new VP chart with auxiliary information (abbreviated as VP-AI) is proposed in this thesis. To obtain an in-depth understanding of the effect of incorporating auxiliary information on the VP chart, the performance of the VP-AI chart is compared with the standard VP chart. At the same time, the performance of the VP-AI chart is compared with the existing Shewhart AI (SH-AI), synthetic AI (SYN-AI), exponentially weighted moving average AI (EWMA-AI), run sum AI (RS-AI) and variable sample size and sampling interval AI (VSSI-AI) charts. The comparisons are made in terms of the steady-state average time to signal (ATS) and expected ATS (EATS) when the exact shift is known and unknown in advance, respectively.

In a traditional setting, a considerable volume of products is manufactured and the run of a process is long (infinite production horizon). However, in many industries, there is an increased interest in short production runs or a finite production horizon where the run of a process only lasts for a few days or hours. The products manufactured in a short production runs setting are typically manufactured in low volume with a high variety. For example, in the clothes manufacturing domain, the short production runs approach is adopted when customers request for a diverse selection of clothes at a low quantity (Celano et al., 2011). Additionally, companies have adopted the short production runs approach as a measure to increase flexibility in manufacturing and to be more specialized. The use of short production runs can be seen in the Just-in-Time technique where products are only manufactured to meet demand; hence, reducing the amount of wastes and excess inventory.

The simultaneous monitoring of several correlated quality characteristics using a multivariate control charting scheme has seen an increased interest in recent years. This is due to the fact that usually more than one quality characteristic of interest is monitored in practice. The Hotelling's T^2 chart is an extensively adopted multivariate control charting scheme as it can be implemented easily. In view of the widespread application of the Hotelling's T^2 chart and the importance of considering short production runs, the fixed sample size (FSS) T^2 short-run chart is developed in this thesis. Considering the effectiveness of the adaptive approach, the VSS T^2 short-run chart is further proposed in this thesis. The VSS approach is an adaptive procedure where the sample size can vary and is dependent on the prior sample statistic. Subsequently, the run length properties of both FSS and VSS T^2 short-run charts are compared for cases when the exact shift size can be specified in advance and when the exact shift size is unknown beforehand.

To summarize, a new VP-AI chart is developed in this thesis by integrating two powerful control charting approaches which are the adaptive procedure and auxiliary information technique. In control charting literature, it is a common practice to integrate effective control charting procedures to construct a new and superior control chart to improve process monitoring. Additionally, most of the prior short-run charts present in the literature are univariate in nature, where only one quality characteristic is monitored. To fill the gap of limited multivariate short-run control charting schemes, a new FSS T^2 short-run chart is proposed. As the adaptive approach is effective in enhancing process monitoring, it is adopted in the multivariate short production runs setting by proposing a new VSS T^2 short-run chart in this thesis.

1.4 Objectives of the Thesis

The objectives of this thesis are outlined below:

(i) To propose a new control chart that incorporates auxiliary information to monitor the process mean, namely the VP-AI chart.

(ii) To propose the new multivariate FSS and VSS T^2 charts for monitoring short production runs.

1.5 Organization of the Thesis

In this section, the organization of this thesis is outlined. This thesis begins with Chapter 1 and concludes with Chapter 5. In Chapter 1, a summary of SPC is first provided. Subsequently, the background of control charts is discussed and the applications of control charts in various domains are enumerated. This is followed by the research motivation and objectives of the thesis. The organization of the thesis is then given.

Chapter 2 presents a discussion of the performance measures that are employed in this thesis. The performance measures, such as truncated average run length (TARL), truncated standard deviation of the run length (TSDRL), average sample size (ASS), probability of getting a signal within the number of scheduled inspections (P(I)) and ATS are used when the exact shift size can be specified in advance. Meanwhile, the expected TARL (ETARL), expected TSDRL (ETSDRL), expected ASS (EASS), expected P(I) (E(P(I)) and EATS are adopted when the exact shift size cannot be specified in advance. Additionally, a review of the existing AI charts, i.e. SH-AI, SYN-AI, EWMA-AI, RS-AI and VSSI-AI charts are provided in this chapter.

Chapter 3 presents a detailed discussion on the methodology and performance measures of the new VP-AI chart for monitoring the process mean. Furthermore, the optimization algorithms used to design the VP-AI chart by minimizing the out-of-control ATS (ATS_1) and EATS ($EATS_1$) are explained in this chapter. In terms of the ATS and EATS criteria, the performance of the VP-AI chart is studied and compared with the existing AI charts. To demonstrate the implementation of the VP-AI chart, an illustrative example based on a real dataset is given.

In Chapter 4, a review on the T^2 statistic and short production runs approach is provided. The methodologies of the FSS and VSS T^2 short-run charts are discussed. In addition, the optimization designs of the VSS T^2 short-run chart that minimize the out-of-control TARL ($TARL_1$) and ETARL ($ETARL_1$) are explained in this chapter. The performance of the VSS T^2 short-run chart is compared with its FSS counterpart. Additionally, the application of the VSS T^2 short-run chart is illustrated with an example based on a real dataset.

Chapter 5 completes the thesis with conclusions. The main contributions and findings of the thesis are enumerated in this chapter. This chapter also identifies topics and suggestions that can be explored for further research.

At the end of this thesis, the references and appendices are provided. Appendix A consists of the optimization programs for the proposed charts written in the ScicosLab software. To verify the results of the proposed charts, simulation programs are written in the Statistical Analysis System (SAS) software and are provided in Appendix B. Moreover, the MATLAB programs that are used to optimally design the existing AI charts can be obtained from Appendix C. Lastly, Appendix D presents the optimal parameters of the existing AI charts.

CHAPTER 2
A REVIEW ON PERFORMANCE MEASURES AND RELATED CONTROL CHARTS

2.1 Introduction

The statistical efficiency of a control chart is determined by the speed a process shift is detected by the chart (Costa, 1997). In control charting literature, performance measures have been adopted by researchers to study the performance of control charts in detecting shifts. The performance measures average run length (ARL) and standard deviation of the run length (SDRL) have been extensively adopted in the literature to evaluate the performance of charts. However, there are various situations where ARL and SDRL are not suitable for adoption, i.e. when the sampling interval is not fixed or the production run is short. Hence, other performance measures have to be employed and are discussed in Section 2.2. The performance measures discussed are the ATS, TARL, TSDRL, ASS and P(I) when the shift size δ is known and can be specified in advance. Alternatively, the performance measures EATS, ETARL, ETSDRL, EASS and E(P(I)) are discussed for cases when δ is unknown and cannot be specified in advance.

The Shewhart chart has prevailed in detecting large shifts but it is not as effective in detecting small and moderate shifts. In view of this, various control charts that incorporate auxiliary information have been developed in the literature to improve process monitoring. In fact, proposing a new VP-AI chart that integrates auxiliary information in monitoring the process mean is an objective of this thesis. In Section 2.3, detailed explanations on the construction and performance measures of prior research works on AI charts that are related to the proposed VP-AI chart are provided.

The AI charts reviewed are the SH-AI, SYN-AI, EWMA-AI, RS-AI and VSSI-AI charts.

2.2 Performance Measures of Control Charts

In the control charting literature, performance measures are employed by researchers to evaluate the sensitivity of a control chart. In other words, it is essential to adopt performance measures to assess the ability of a chart in detecting process shifts. A chart with a superior ability in detecting shifts is more effective in process monitoring. Hence, performance measures serve as an aid to practitioners in selecting the best or most suitable chart.

Typically, a characteristic of the run length (RL) distribution is used to assess the performance a chart such that RL refers to the number of points plotted on a chart until the occurrence of a signal that implies an out-of-control condition. The most adopted performance measure to study the performance of a chart is the ARL. The employment of ARL in control charting schemes can be seen in Lu (2015), Guo and Wang (2015), Yang and Arnold (2016), Tran and Knoth (2018) and Maravelakis et al. (2019), to name a few. ARL is defined as the expected number of samples plotted on a chart until a signal is triggered. Note that ARL_0 and ARL_1 denote the ARL when the process is in-control and out-of-control, respectively.

During process monitoring, false alarm or Type-1 error where the process is in-control even though an out-of-control signal is triggered by the chart may occur (Testik, 2007). A false alarm occurs when a point is plotted beyond the control limits when the process is in-control. To illustrate, if $ARL_0 = 370$, this implies that, on average, a point falling beyond the control limits will occur (or a signal will be triggered) by the chart in every 370 samples even though the process is in an in-control condition (Montgomery, 2013). When the process is in-control or on target, having a value of the performance measure which is as large as possible is preferable to lower the false alarm rate.

In contrast, when the process is out-of-control or off-target, it is desirable to have performance measures with lower values so that the chart signals an out-of-control condition earlier. Thus, quality practitioners can take swift corrective actions to eliminate the assignable cause(s) earlier. The performance measure when the process is out-of-control can be used to compare the effectiveness of several charts where the chart with a lower performance measure is more effective in detecting shifts. However, the false alarm rate of the charts has to be the same to provide a common ground for comparison. For example, with the same value of ARL_0, a chart with a lower ARL_1 value is more sensitive to process shifts compared to its counterpart with a higher ARL_1 value.

Another performance measure that is typically used in tandem with the ARL to evaluate the performance of a chart is the SDRL. The SDRL is employed to study the variability in the run length distribution. The in-control and out-of-control SDRLs are denoted as $SDRL_0$ and $SDRL_1$, respectively. When comparing several control charts with the same ARL_0, the one with the lowest $SDRL_1$ value has the smallest spread in the run length distribution; thus, surpassing the competing charts.

2.2.1 Average Time to Signal (ATS) and Expected ATS (EATS)

Even though ARL is a commonly used performance measure, it is only suitable as a performance measure when the sampling interval is fixed or constant. When the sampling interval of a process can vary according to the value of the prior sample statistic, ATS should be used as a performance measure instead of ARL. In other words, ATS is a more suitable performance measure when the sampling interval is variable. This is because the time to signal is no longer a constant multiple of the ARL when the sampling interval can be varied (Li et al., 2014). Numerous research works with variable sampling interval, such as Liu et al. (2015), Chew et al. (2016), Khaw et al. (2017), Yeong et al. (2017) and Zhou (2017) have adopted ATS to evaluate the performance of the charts.

For a chart with variable sampling interval, its sensitivity in monitoring process shifts is measured by the length of time until the chart signals. When the zero-state performance of a chart is assessed, ATS is defined as the average time from the beginning of a process until the time in which the occurrence of an out-of-control condition is signalled by the chart. The zero-state case assumes that the process begins off-target (Prabhu et al., 1997). On the other hand, when the steady-state performance of a chart is assessed, ATS can be defined as the average time from the time an assignable cause occurs to the time the occurrence of an out-of-control condition is signalled by the chart. For the steady-state case, it is assumed that the process begins on-target. However, a shift occurs in the future, causing the process to be off-target at a random time (Prabhu et al., 1997).

The ATS can be divided into two categories which are the in-control ATS (ATS_0) and ATS_1. Note that ATS_0 has to be fixed for a fair comparison to be made when the performance of various charts is compared. When comparing several charts such that ATS_0 is fixed, the chart with the lowest ATS_1 is the most effective as the least time is required for the chart to signal an out-of-control condition. With a shorter time to detect an out-of-control condition, practitioners are able to take swift corrective actions to remove the assignable causes(s) present in a process and return the process to its in-control state.

In the preceding paragraphs, the ATS is adopted with the assumption that shift δ is known beforehand. However, in practice, the exact δ is usually unknown and cannot be specified by quality practitioners prior to the occurrence of a shift. Hence, the EATS computed for an interval of shifts $(\delta_{\min}, \delta_{\max})$, where $\delta_{\min}$ and $\delta_{\max}$ denote the minimum and maximum shifts, respectively, should be employed as a performance measure when the exact δ cannot be specified in advance. Similar to ATS, EATS is a performance indicator used by practitioners to select the best control charting scheme. Note that EATS can be divided into the in-control EATS (EATS_0) and EATS_1. When $\text{EATS}_0 = \text{ATS}_0$ is fixed, the chart with a lower EATS_1 value for the interval $(\delta_{\min}, \delta_{\max})$ has a superior performance.

Additionally, a larger δ has to be swiftly detected by control charts as it leads to a more considerable loss in quality compared to a smaller δ. Hence, the ATS_1 values are expected to decrease as δ increases. This indicates that the time to signal is shorter for a larger δ; thus, implying that the chart is more sensitive as δ increases. Furthermore, EATS_1 is expected to exhibit a similar trend, where EATS_1 decreases when $(\delta_{\min}, \delta_{\max})$ extend over larger δ values. As a result, the chart has a better ability in detecting $(\delta_{\min}, \delta_{\max})$ for larger δ values.

2.2.2 Truncated Average Run Length (TARL) and Expected TARL (ETARL)

The ARL is a suitable performance measure when the production run is long. However, when the production run is short, the process may end without the issuance of any out-of-control signal by the chart. In other words, there is a possibility that an out-of-control condition does not occur during the short production runs. Taking this phenomenon into account, the TARL which was introduced by Nenes and Tagaras (2010) is a more suitable performance measure for short production runs. In fact, TARL has been adopted in numerous research works involving short production runs, i.e. Celano et al. (2011, 2013), Castagliola et al. (2013, 2015), Amdouni et al. (2015)

and Chong et al. (2019). The definition of TARL is the average number of samples taken until a signal that indicates an out-of-control condition is triggered or until the completion of a production process, whichever occurs first.

It is worth noting that only the zero-state TARL can be considered where it is assumed that short production runs begin off-target. The steady-state case assumes that an out-of-control condition happens randomly in the future. However, only a small number of scheduled inspections I is taken in short production runs. Hence, the steady-state TARL cannot be considered due to the limited number of inspections. Similar to ARL, TARL consists of the in-control TARL ($TARL_0$) and $TARL_1$. In short production runs, practitioners can use $TARL_1$ to compare the performance of several charts such that the value of $\text{TARL}_0 = I$ is fixed for all the charts considered. The chart which is the most sensitive in detecting a process shift has the lowest $TARL_1$ value. Note that $TARL_1$ is adopted with the assumption that the magnitude of δ is known in advance.

As the value of δ is typically unknown beforehand in practice, ETARL is considered as an alternative to TARL and it is used to evaluate the performance of a chart. For shift interval $(\delta_{\min}, \delta_{\max})$, the chart with the lowest $ETARL_1$ has the best performance when $\text{TARL}_0 = I$ is fixed for all charts under comparison. Similar to control charts with long production runs, short-run control charts are quicker in detecting larger δ that results in a higher loss of quality in comparison to smaller δ. Hence, as δ increases, the $TARL_1$ values are expected to decrease. Meanwhile, $ETARL_1$ is expected to exhibit a similar trend where $ETARL_1$ decreases when $(\delta_{\min}, \delta_{\max})$ extend over larger δ values. Thus, when δ is unknown, the chart is more sensitive in detecting shifts in the interval $(\delta_{\min}, \delta_{\max})$ for larger δ values.

2.2.3 Truncated Standard Deviation of the Run Length (TSDRL) and Expected TSDRL (ETSDRL)

SDRL is a suitable performance measure to measure variability in the RL distribution when the production horizon is infinite. However, when a finite production horizon is considered, the TSDRL that measures the spread of the truncated run length (TRL) should be used as an alternative to SDRL. The in-control and out-of-control TSDRL are denoted as TSDRL_0 and TSDRL_1, respectively.

With a fixed TARL_0 value to ensure a common ground for comparison, practitioners can compare the TSDRL_1 values of various charts. A chart with a lower TSDRL_1 value has a smaller spread of the TRL distribution which indicates that the chart is superior to the other charts. As the use of TSDRL assumes that the exact δ is known in advance, it may not be a suitable performance measure in practice.

Thus, the ETSDRL which is a counterpart to the TSDRL when the exact δ cannot be specified in advance can be adopted for the shift interval $(\delta_{\min}, \delta_{\max})$. The in-control ETSDRL is denoted as ETSDRL_0 while the out-of-control ETSDRL is denoted as ETSDRL_1. Given that $\text{TARL}_0 = I$ is fixed, practitioners can compare the spread of various charts using ETSDRL_1. The lower the value of ETSDRL_1, the smaller is the spread of the TRL of the chart. Hence, the chart with lower ETSDRL_1 values outperforms the other competing charts.

2.2.4 Average Sample Size (ASS) and Expected ASS (EASS)

For the VSS chart with sample size that varies according to past process information, the ASS is computed to ensure a fair comparison with other charts. The in-control ASS is denoted as ASS_0 while its out-of-control counterpart is denoted as ASS_1. For a common ground of comparison, the ASS_0 of the VSS chart is set to be equal to the fixed sample size (n) of the FSS chart.

From one sample to another, the sample size of the VSS chart varies. Thus, ASS_1 can be used as a performance measure of the VSS chart. If the value of ASS_1 is more than n, the cost required to implement the VSS chart may be higher due to the more intensive sampling. Thus, the VSS chart may be deemed as less attractive due to the increase in cost. On the other hand, if the value of ASS_1 is less than n, this implies that the implementation of the VSS chart will be less costly. Hence, practitioners may adopt the VSS chart in this case due to its cost saving feature.

When the exact δ is unknown, the EASS that considers $(\delta_{\min}, \delta_{\max})$ can be adopted as an alternative to ASS. The in-control and out-of-control EASS are represented by $EASS_0$ and $EASS_1$, respectively. If the $EASS_1$ value of the VSS chart is lower than n, it has a superior performance due to the reduction in the cost of implementing the chart.

2.2.5 Probability of Getting a Signal within the Number of Scheduled Inspections (P(*I*)) and Expected P(*I*) (E(P(*I*)))

The probability of getting a signal within I inspections (P(I)) can be used to study the sensitivity of a short-run chart and has been adopted in research works involving short production runs, like those by Celano et al. (2011, 2013), Celano and Castagliola (2018), and Chong et al. (2019), to name a few. The short-run chart with a higher P(I) has a better ability in detecting a process shift. Hence, a short-run chart that has a greater P(I) outperforms other short-run charts.

For the case where the exact shift size is unknown and cannot be specified in advance, E(P(I)) can be adopted to evaluate the sensitivity of the short-run chart in detecting $(\delta_{\min}, \delta_{\max})$. The short-run chart that has the highest E(P(I)) is the most sensitive and outperforms the other short-run charts.

2.3 Auxiliary Information (AI) Control Charts

In order to acquire estimates of the population parameters with an improved precision, the auxiliary information approach is commonly employed. As information from related auxiliary variables can be typically obtained alongside the quality characteristic of interest or study variable, it can be thus utilized to improve the precision of an estimator. The use of auxiliary information can be seen in coal energy generation, where the total energy generated from coal is the study variable (Ahmad et al., 2014). Meanwhile, the auxiliary variable which is correlated to the study variable can be the air temperature or the flue gas quantity (Ahmad et al., 2014).

Additionally, the auxiliary information approach can be adopted in the manufacturing of textile fibres where the study variable is the break factor of a single textile strand while the correlated auxiliary variable is the textile fibre's weight (Haq & Khoo, 2016). It may be expensive or time consuming to obtain a large sample of the study variable to improve the accuracy of the estimator. Thus, researchers have incorporated the auxiliary information approach into control charts as a method to enhance process monitoring. Specifically, the information from the auxiliary or supplementary variable is incorporated into the regression estimator; thus, improving the accuracy of the estimator and enhancing the sensitivity of the control chart in monitoring shifts.

To improve the performance of the Shewhart chart in monitoring the process mean, Riaz (2008) incorporated auxiliary information into the Shewhart chart and developed the SH-AI chart. By utilizing information from an auxiliary variable and the correlation coefficient ρ between the auxiliary and study variables, the SH-AI chart has an enhanced performance in process monitoring. It was discovered that the performance of the SH-AI chart improves as ρ increases.

Additionally, the EWMA-AI chart that monitors process mean was proposed by Abbas et al. (2014). The EWMA-AI chart was found to be effective in detecting small and moderate shifts and outperforms other competing charts. Furthermore, Haq and Khoo (2016) proposed the SYN-AI chart which integrates auxiliary information with the synthetic chart. It was discovered that the SYN-AI chart is superior to the basic synthetic chart and surpasses the EWMA chart when $\rho \geq 0.75$.

The RS-AI chart was developed by Ng et al. (2018) and the former was compared with the SH-AI, SYN-AI and EWMA-AI charts. The RS-AI chart was designed using 4 and 7 scoring regions and Ng et al. (2018) discovered that the performance of the RS-AI chart improves by adding more scoring regions. When ρ is large, it was shown that the RS-AI chart surpasses the competing control charting schemes for all δ values. Meanwhile, Saha et al. (2018) proposed the VSSI-AI chart which is a VSSI chart combined with auxiliary information and found that the former outperforms the EWMA-AI and SYN-AI charts. More existing AI charts present in the literature can be found in Riaz (2011), Riaz et al. (2013), Ahmad et al. (2014), Lee at al. (2015) and Abbasi and Riaz (2016), to name a few.

In this thesis, the performance of the VP-AI chart is compared with the VSSI-AI chart as both charts are adaptive charts involving parameters that vary according to the prior sample statistic's position on the chart. All the parameters (sampling interval, sample size, control and warning limits) of the VP-AI chart can be varied while only two parameters (sampling interval and sample size) of the VSSI-AI chart can be varied. When both VP-AI and VSSI-AI charts are compared, the improvement in process monitoring by varying the control and warning limits can be studied.

Another competing chart that will be compared with the VP-AI chart is the RS-AI chart. Note that only the RS-AI chart with 7 regions is considered in this thesis even though Ng et al. (2018) developed RS-AI charts with 4 and 7 regions, as the latter was found to be superior to the former in detecting process mean shifts. As the RS-AI chart was compared with the SH-AI, SYN-AI and EWMA-AI charts, the VP-AI chart is similarly compared with the three competing AI charts in this thesis. Note that Abbas et al. (2014) computed the run length properties of the EWMA-AI chart using simulation. However, the Markov chain approach is adopted to derive the run length properties of the EWMA-AI chart in this thesis.

2.3.1 Shewhart AI (SH-AI) Chart

The SH-AI chart was developed by Riaz (2008) to enhance the performance of the Shewhart chart through the incorporation of auxiliary information. Riaz (2008) showed that the SH-AI chart which has a more precise estimator by considering auxiliary information is a more powerful control charting scheme in comparison to the standard Shewhart chart in monitoring process mean. Suppose that there are two correlated variables in a process, i.e. auxiliary variable *C* and study variable *S*. A regression estimator that incorporates information from *C* and *S* is employed in the design of an SH-AI chart. However, the SH-AI chart only monitors the process mean shifts of *S*. It is worth noting that bivariate normality is assumed for (S, C), i.e. $(S, C) \sim N_2\left(\mu_S, \mu_C, \sigma_S^2, \sigma_C^2, \rho\right)$.

Given that $\delta = \dfrac{|\mu_S - \mu_{S0}|}{\sigma_S}$, then $\mu_S\left(= \mu_{S0} + \delta\sigma_S\right)$ and σ_S^2 represent the population mean and variance of *S*, respectively, such that μ_{S0} is the in-control population mean. At the same time, μ_C and σ_C^2 represent the population mean and variance of *C*, respectively, while ρ denotes the correlation coefficient between the variables *S* and *C*.

Based on a bivariate normal distribution, suppose that (S_j, C_j), such that $j = 1, 2, ..., n$ is a bivariate random sample and has a fixed sample size denoted by n. Hence, the regression estimator of μ_S is given by (Riaz, 2008)

$$X^*_{S_v} = \bar{S}_v + \beta(\mu_C - \bar{C}_v), \quad (2.1)$$

where $\beta = \rho\left(\frac{\sigma_S}{\sigma_C}\right)$. Note that S has a sample mean, $\bar{S}_v = \frac{1}{n}\sum_{j=1}^{n} S_j$ and C has a sample mean $\bar{C}_v = \frac{1}{n}\sum_{j=1}^{n} C_j$ for sample number v. Meanwhile, the mean and variance of $X^*_{S_v}$ is as follows:

$$E\left(X^*_{S_v}\right) = \mu_S \quad (2.2)$$

and

$$Var\left(X^*_{S_v}\right) = \frac{1}{n}\sigma_S^2\left(1-\rho^2\right), \quad (2.3)$$

respectively. Thus, $X^*_{S_v} \sim N\left(\mu_S, (1/n)\sigma_S^2\left(1-\rho^2\right)\right)$, where $X^*_{S_v}$ follows the normal distribution.

The upper control limit (UCL), center line (CL) and lower control limit (LCL) of the SH-AI chart are (Riaz, 2008)

$$\text{UCL} = \mu_{S0} + 3\frac{\sigma_s}{\sqrt{n}}\sqrt{1-\rho^2}, \quad (2.4a)$$

$$\text{CL} = \mu_{S0} \quad (2.4b)$$

and

$$\text{LCL} = \mu_{S0} - 3\frac{\sigma_s}{\sqrt{n}}\sqrt{1-\rho^2}, \quad (2.4c)$$