DETERMINING OUT-OF-CONTROL VARIABLE(S) IN A MULTIVARIATE QUALITY CONTROL CHART FOR INDIVIDUAL OBSERVATIONS

BY

AGOG, NATHAN SAMUEL

B.Sc STATISTICS (ABU, 2010)

M.Sc/SCI/0207/2011-2012

A THESIS SUBMITTED TO THE SCHOOL OF POSTGRADUATE STUDIES, AHMADU BELLO UNIVERSITY, ZARIA NIGERIA

IN PARTIAL FULFILMENT FOR THE AWARD OF MASTER OF SCIENCE (M.Sc) DEGREE IN STATISTICS,

DEPARTMENT OF MATHEMATICS

AHMADU BELLO UNIVERSITY, ZARIA

NIGERIA.

AUGUST, 2014

DECLARATION

I hereby declare that the work in this thesis entitled "Determining Out-of-Control variables in a Multivariate Quality Control Chart for Individual Observations" is a record of my own research work under the supervision of Dr. H.G Dikko and Prof. O.E Asiribo. All literatures cited in this work have been duly acknowledged in the text and a list of references provided. No part of this work has been submitted for the award of degree in any other University.

_____ _____

AGOG, Nathan Samuel **Date**

CERTIFICATION

This thesis entitled "DETERMINING OUT-OF-CONTROL VARIABLE(S) IN A MULTIVARIATE CONTROL CHART FOR INDIVIDUAL OBSERVATIONS" by AGOG, Nathan Samuel (M.Sc/Sci/207/2011-2012) meets the regulations governing the award of the degree of Master of Science of the Ahmadu Bello University, Zaria and is approved for its contribution to knowledge and literary presentation.

Dr. H.G. Dikko _____ _____

Chairman, Supervisory Committee Signature Date

Prof. O.E. Asiribo _____ _____

Member, Supervisory Committee Signature Date

Dr. B. Sani _____ _____

Head of Department Signature Date

Prof. Adebayo A. Joshua _____ _____

Dean, School of Postgraduate Studies Signature Date

DEDICATION

This research work is dedicated to God Almighty for His faithfulness endures forever and also to my late father Mr. Samuel Agog who has laid this foundation.

ACKNOWLEDGEMENT

I am grateful to God Almighty, who has been faithful to me in all areas of my endeavor. Thank you Lord for the wisdom and knowledge you have blessed me with.

I am very grateful to my supervisors, Dr. H.G Dikko and Prof. O.E. Asiribo for their relentless support and time to ensure that this work became a successful one.

My deep appreciation goes to all academic and non-academic staff members of the Department of Mathematics, Ahmadu Bello University Zaria, particularly to the Head of Department, Dr. Babangida Sani, the Postgraduate Coordinator in person of Dr. Abubakar Yahaya, the seminar Coordinator in person of Dr. A. Mohammed for their immense contributions. I also want to say a big thank you to Mr. N.N. Chibuike for his time and contribution to the success of this work.

A special thanks to my lovely mother, Mrs. Lami Samuel Agog for her relentless support, fervent prayers and encouragement throughout this programme. Mum, may the Almighty God continue to bless you. My gratitude goes to Mr. and Mrs. Joel Agog, Mr. Japheth Agog, Miss Joan Agog, Miss Jael Agog and also my cousins Ajibola Mustapha, Abraham Philemon and Zumunta Zubairu. Thank you for your patience and understanding.

I am very grateful to the family of Bulus Basahuwa Agog for their support throughout my stay in Zaria. To my friends; Miss Afiniki B. Kantiok Mr. Alhamdu Woji, Mr. Hosea Matoh, Mr. Patrick Neyu Achi, Mr. Isaac Bobai, Mr. Ariko Joseph David, Miss. Rebecca Wada Bitiyong, Miss Kinan Kabila, Mr. Amos Matoh, Mr. Aaron Adams, Mr. David John and all my M.Sc classmates, thank you for your wonderful support during this research work. And for those who have contributed in one way or the other, space and time will not permit me to mention your names, thank you very much for your assistance.

ABSTRACT

The decomposition of the Hotelling's T^2 statistic into orthogonal components is considered to be one of the most effective methods for detecting variable(s) responsible for an out-of-control signal. In this work, an extension of the T^2 decomposition from three variables (p=3) to four (p=4) variables, where the number of decompositions increased from 3! = 6 to 4! = 24 and the decomposition terms also increased from 18 to 96 terms having 32 distinct terms were provided. These distinct terms are the ones that were examined for possible contribution to the T^2 signal. A dataset obtained from an Indomie company in Northern Nigeria was used to assess the validity of constructed model by demonstrating the invariance property of the Hotelling's T^2 statistic. The model was also used to identify the variable(s) that significantly contribute to an out-of-control signal. By comparing the critical values with the corresponding T^2 values, we were able to detect variation between the four (4) variables in their mean and also their variance-covariance structure.

TABLE OF CONTENTS

Declaration i

Certification ii

Dedication iii

Acknowledgement iv

Abstract v

Table of Contents vi

List of Tables viii

List of Figures ix

List of Appendices x

List of Abbreviations xi

CHAPTER ONE: GENERAL INTRODUCTION 1

1.0 Introduction 1

1.1 Motivation of the Study 2

1.2 Assumptions of Statistical Process Control (SPC) 3

1.3 Aim and Objectives of the Study 5

1.4 Significance of the Study 5

1.5 Traditional Statistical Process 5

1.5.1 Univariate Control Charts 6

1.6 Multivariate Statistical Process Control (MSPC) 8

1.6.1 Advantages of MSPC 9

1.6.2 Disadvantages of MSPC 10

1.7 Application of Multivariate Quality Control 10

CHAPTER TWO: LITERATURE REVIEW 12

2.0 Introduction 12

2.1 Multivariate Chart 12

2.1.1 Hotelling's T^2 Control Chart 12

2.1.2 Multivariate Exponentially–Weighted Moving Average Control Chart 16

2.1.3 Multivariate Cumulative Sum Control Chart 17

2.2 Identifying Out- of- Control Variable 17

2.2.1 Using Bonferroni Control Limits. 18

2.2.2 Application of Principal Components 20

2.2.3 Application of T^2 Decomposition 21

2.2.4 Application of Neural Networks 23
2.2.5 Using Graphical Techniques 24
2.2.6 Cause-Selecting Control Chart and Regression Adjusted Variables. 25
CHAPTER THREE: METHODOLOGY 27
3.0 Introduction 27
3.1 Interpretation of Out-of-Control 27
3.2 The Decomposition of Hotelling's T^2 Statistic 30
3.3 Model for the T^2 Decomposition Using Four Variables 31
3.4 Computing the MYT Decomposition Terms 35
3.6 Method of Data Collection and Data Analysis 39
CHAPTER FOUR: RESULTS AND DISCUSSIONS 41
4.0 Introduction 41
4.1 Normality Test 41
4.2 Hotelling's Control Chart (Phase I) 42
4.3 Hotelling's Control Chart (Phase II) 42
4.4 Computation of the T^2 Decomposition Terms 43
4.5 Hotelling's T^2 Control Chart after Taking out Abnormal Observations 49
4.6 The Invariance Property of the Hotelling's T^2 Statistic 50
CHAPTER FIVE: SUMMARY, CONCLUSION AND RECOMMENDATION 52
5.0 Introduction 52
5.1 Summary 52
5.2 Conclusion 54
5.3 Recommendation 54
5.4 Contribution To Knowledge 55
5.5 Further Research 55
REFERENCES 56
APPENDICES 64

LIST OF TABLES

Table 3.1:Unique Decomposition Terms (cited from Mason *et al.,* 1997) 34
Table 4.1:MYT decomposition terms of three signaling points. 444

LIST OF FIGURES

Figure 4. 1:The Q-Q plot of individual variables in the four machines 41

Figure 4. 2:Hotelling's T^2 control chart for historical dataset 42

Figure 4. 3:Hotelling's T^2 Control Chart for Phase II 42

Figure 4. 4:Hotelling's T^2 control chart for 177 observations 49

Figure 4. 5:Phase II Hotelling's T^2 control chart for 177 observations 50

LIST OF APPENDICES

APPENDIX I **65**
APPENDIX II **70**
APPENDIX III **75**
APPENDIX IV **77**
APPENDIX V **78**
APPENDIX VI **79**

LIST OF ABBREVIATIONS

ARL: AVERAGE RUN LENGTH

CUSUM: CUMULATIVE SUM

EWMA: EXPONENTIALLY WEIGHTED MEAN AVERAGE

FITs: FINITE INTERSECTION SETS

MCUSUM: MULTIVARIATE CUMULATIVE SUM

MEWMA: MULTIVARIATE EXPONENTIALLY WEIGHTED MEAN AVERAGE

MSPC: MULTIVARIATE STATISTICAL PROCESS CONTROL

MSQC: MULTIVARIATE STATISTICAL QUALITY CONTROL

MP: MULTIVARIATE PROFILE

NNI: NEURAL NETWORK IDENTIFIER

PCA: PRINCIPAL COMPONENT ANALYSIS

SPC: STATISTICAL PROCESS CONTROL

Q-Q plot: QUANTILE-QUANTILE PLOT

CHAPTER ONE

GENERAL INTRODUCTION

1.0 INTRODUCTION

Statistical Process Control (SPC) has played a significant role in controlling the product quality for decades since Shewhart (1931) illustrated the technique of the control charts by applying statistical concepts in the manufacturing process.

According to MacCarthy and Wasusri (2002), statistical process control is a powerful tool for monitoring and control processes and has been widely used in manufacturing and non-manufacturing processes.

With the advancement in technology, there has been an increase in customer expectations and the need to monitor correlated variables simultaneously. Process monitoring in which several variables are of interest is called Multivariate Statistical Process control (MSPC).

Multivariate control charts is widely used in practice to monitor the simultaneous performance of several related quality characteristics. The origin of multivariate control chart can be attributed to Hotelling (1947). A multivariate control scheme has a better sensitivity than one based on the univariate control charts in monitoring multivariate quality process. (Lu *et al.,* 1998)

Woodall and Montgomery (1999) stated that multivariate process control is one of the most rapidly developing sections of statistical process control. The demand to implement MSPC in a production process for quality improvements increases daily. Statistical methods play a very important role in quality improvement in manufacturing industries (Woodall, 2000).

The quality of any product is usually determined by several correlated quality variables. One of the popular multivariate control charts is based on Hotelling's T^2 statistic which is used

to simultaneously monitor those quality variables and taking their correlations into consideration. There are a lot of literatures focusing on multivariate control charting methods based on Hotelling's T^2 statistic in detecting mean shift such as Sullivan and Woodall (1996), Mason and Young (1999), Tong *et al.,* (2005), and many others.

1.1 MOTIVATION OF THE STUDY

When Hotelling T^2 detects a change in the mean vector, corrective action is required. A T^2 value, however, does not provide direct information about which variable is responsible for the overall out-of-control condition. This information is of practical importance because quality engineers/analysts need to know which variable requires adjustments after the process is declared out-of-control. The most challenging issue about multivariate quality control chart is the ability to identify the variable which is responsible for an out-of-control condition.

Many literatures have discussed and presented methods which can be used to identify out-of-control variable or variables and much credit has been to the method proposed by Mason *et al*.,(1995), for more information see the works of Bersimis *et al.*,(2007). This method involves the decomposition of the T^2 statistic into orthogonal components which reflects the contribution of each variable in an observation vector. Much application of the decomposition technique has been on two and three variables as seen in the case of Yarmohammadi and Ebrahimi (2010), Ulen and Demir (2013), Sani and Abubakar (2013) and so many others.

Holmes and Mergen (1993), Sullivan and Woodall (1996), and Vargas (2003) have noted that the Phase I Hotelling's T^2 control chart for individual observations is less sensitive in detecting trend or process mean shifts. In this work we would apply both Phase I and Phase II of the Hotelling's T^2 control chart and also provide the T^2 decomposition model for p=4, using

Mason, Young, Tracy (MYT) decomposition technique which would be used for identifying out-of-control variable .

1.2 ASSUMPTIONS OF STATISTICAL PROCESS CONTROL (SPC)

The standard assumptions in SPC are that the observed process values are normally, independently and identically distributed (iid) with fixed mean (μ) and standard deviation σ when the process is in control. Before MSPC can be implemented, the p variables must be related to each other.

Correlation analysis is a technique used to show the strength of the relationship between pairs of variables. Del Castillo (2002) defined correlation as the departure of two or more variables from independence. Montgomery (2001), defined correlation as a degree to which two or more quantities are associated.

When two or more random variables are defined on a probability space, it is useful to measure the relationship between the variables. A common measure of the relationship between two variables is called covariance. The covariance between random variables X and Y, denoted as COV(X, Y) or σ_{XY} is

$$\sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)]$$

Covariance provides an idea of the strength of correlation. In the case of two variables X and Y, the correlation is considered to be very strong if X is far from its mean and Y is also far from its mean. Hence, the covariance between the two variables X and Y describes the variation between the two variables.

In the multivariate case, the population covariance is represented in a matrix denoted as **Σ**. The covariance matrix is also called the variance-covariance matrix. The variance-covariance matrix is a symmetrical matrix that contains the covariance among a set of random variables. The

main diagonal elements of the matrix are the variances of the random variables, and the off-diagonal elements are the covariance between the *p* variables (Neter *et al*.,1996).

The $p \times p$ variance-covariance matrix, **S** is as follows;

$$S=\begin{pmatrix} S_1^2 & S_{12} & \cdots & S_{1P} \\ S_{21} & S_2^2 & \cdots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_p^2 \end{pmatrix}$$

In a two-dimensional plot, the degree of correlation between the values on the axes is quantified by the so-called correlation coefficient. The most common correlation coefficient is the Moment Correlation, which is found by dividing the covariance of the two variables by the product of their standard deviation. This correlation coefficient (r) is a measure of the degree of linear relationship between two variables X and Y. The square of (r) is called the coefficient of determination and denotes the portion of total variance explained by the regression model (Walpole and Myers, 1993). The sample correlation coefficient is calculated by

$$r_{xy} = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{(n-1)s_x s_y}$$

where $\overline{x}$ and $\overline{y}$ are the sample means, S_x and S_y are the sample standard deviations of x_i and y_i respectively.

The correlation coefficient

$$\rho_{xy} = \frac{\operatorname{cov}(x,y)}{\sigma_x \sigma_y} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

The correlation coefficient may take value between -1.0 and +1.0.

1.3 AIM AND OBJECTIVES OF THE STUDY

Much research has been done on multivariate process control for variable data in various situations. Moreover, interpretation of out-of-control signals and how to identify the quality characteristics contributing to out-of-control signals have also been discussed. However, identifying influential variable(s) that contribute to out-of-control signal is still a difficult task especially when the quality characteristics are beyond three.

The aim of this research work is to:

Determine variable(s) that significantly contributes to an out-of-control signal in a multivariate quality control chart. This is achieved through the following objectives:

1. Application of Hotelling's T^2 components using MYT decomposition technique.
2. Identifying out-of-control condition and out-of-control variable(s)
3. Illustration of invariance property of the Hotelling's T^2 statistic from the derived components.

1.4 SIGNIFICANCE OF THE STUDY

The significance of this study is geared toward detecting out-of-control variable(s) in a multivariate quality control chart. In achieving this, we explored and compared various methods and techniques use in checking out-of-control condition in multivariate control chart. This work is also aimed at helping practitioners in the field of quality control to be able to determine variables that causes out-of-control signal in a monitoring process.

1.5 TRADITIONAL STATISTICAL PROCESS

Control charts were developed in 1931 by Shewhart to be used for process monitoring. Control charts are widely used for detecting assignable and chance causes of variation. Some definitions of control charts are presented as follows.

According to Shewhart (1931), "the control chart may serve these purposes, first, it is used to define the goal or standard for a process that management strives to attain and secondly, it may be used as an instrument for attaining that goal and thirdly, it may serve as means to judging whether the goal has been reached." Control chart may also be viewed as a statistical tool as defined by Duncan in 1956.

Feigenbaum (1983) defined control charts as "…a graphical comparison of the actual product characteristics with limits reflecting the ability to produce as shown by past experience on the product characteristics."

Therefore, control chart is a graphical display used to monitor a process. It usually consists of a horizontal centerline corresponding to the in-control value of the parameter that is being monitored and the upper and lower control limits. Control limits are not determined arbitrarily, nor are they related to specification limits but rather by statistical criteria. Sample points that fall within the control limits are said to be in-control while those points that fall beyond the control limits are said to be an out-of-control process.

The traditional statistical process control is generally referred to as the univariate control charts. This is due to the fact that it considers only a variable for monitoring quality characteristics.

1.5.1 Univariate Control Charts

One major setback of the Shewhart chart is that it regards only the last data point and does not carry a memory of the previous data. As a result, small changes in the mean of a random variable are less likely to be detected rapidly. Exponentially weighted moving average (EWMA) chart improves upon the detection of small process shifts. Rapid detection of small

changes in the quality characteristic of interest and ease of computations through recursive equations are some of the many good properties of EWMA chart that make it attractive.

EWMA chart was first introduced by Roberts (1959) to achieve faster detection of small changes in the mean. The EWMA is a statistic for monitoring the process that averages the data in a way that gives less weight to data as they are further removed in time. EWMA is defined as:

$$Z_i = \lambda \bar{X}_i + (1-\lambda) Z_{i-1} \qquad \text{with } 0 \le \lambda < 1, Z_0 = \mu_0$$

It is used as the basis of a control chart. The procedure consists of plotting the EWMA statistic Z_i versus the sample number on a control chart with center line $CL = \mu_0$.

The upper control limit (UCL) is

$$UCL = \mu_0 + K\sigma_{\bar{x}} \sqrt{\frac{\lambda}{2-\lambda}} [1-(1-\lambda)^{2i}]$$

and lower control limits (LCL) is

$$LCL = \mu_0 - K\sigma_{\bar{x}} \sqrt{\frac{\lambda}{2-\lambda}} [1-(1-\lambda)^{2i}]$$

where

μ_0 =mean

K=constant

$\sigma_{\bar{X}}$ =standard deviation

λ=smoothing parameter

The term $[1-(1-\lambda)^{2i}]$ approaches unity as *i* get larger, so after several sampling intervals, the control limits will approach the steady state values

$$UCL = \mu_0 + K\sigma_{\bar{x}} \sqrt{\frac{\lambda}{2-\lambda}}$$

$$LCL = \mu_0 - K\sigma_{\bar{x}}\sqrt{\frac{\lambda}{2-\lambda}}$$

The CUSUM (cumulative sum) chart is an effective way of monitoring small deviations in the process mean when small deviations are of interest. The CUSUM chart, originally developed by Page (1954), incorporates all information in the sequence of sample values and plots the cumulative sums of the deviations from a target value using samples from prior observations.

1.6 MULTIVARIATE STATISTICAL PROCESS CONTROL (MSPC)

According to Montgomery and Klatt (1972), a lot of attention has been given to the design of control charts where only one quality characteristic is of interest. However, based on the two authors industrial products and processes are characterized by more than one measurable quality characteristic and their joint effect describes product quality.

Process monitoring in which several variables are of interest is called MSPC. Multivariate charts are better than the simultaneous operation of several univariate control charts. Process monitoring using control charts can be seen in two- stage process, Phase I and Phase II (Woodall, 2000). Each phase has a role in monitoring the quality of a product. In Phase I, charts are used for retrospectively testing whether the process was in control using historical dataset. This Phase aids the practitioners in bringing a process to an in-control state. In Phase II, the main concern is to further monitor the historical data set when subsequent samples are drawn.

The parameter of the run-length distribution is often used for measuring the performance of the control chart methods, where the run length is the number of samples taken before an out-of-control condition occurs.

Hotelling (1947) was the pioneer to develop a quality control chart for several related variables and the control chart is well known as the Hotelling T^2 control chart. The Hotelling T^2 control chart is rated as the most widely used multivariate control chart that deals with changes in the mean vector of *p* correlated quality characteristics (Aparisi and Haro, 2001). Hotelling's T^2 control chart is a direct analogue of the Shewhart $\bar{X}$ control chart.

The main tool used for monitoring MSPC is through the use of the quality control chart. MSPC procedure involves fulfilling four conditions:

1. One should be able to state if the process is in control or not.
2. Should be able to know if there was/is a false signal.
3. Should be able to know the relationship among variables, attributes taken into consideration.
4. If the process is out of control, one should know the reasons why it is out of control (Bersimis *et al.,* 2007).

1.6.1 Advantages of MSPC

MSPC has several advantages as compared to its univariate equivalent. As considered by Hotelling (1947), Alt (1985), and Lowry and Montgomery (1995).

According to Hotelling (1947), MSPC has the ability to combine measures in several dimensions into a single measure of performance. In addition, MSPC offers an easier graphical tool to the practitioner. The practitioner can only use one chart instead of multiple univariate charts to evaluate the product or system quality as a whole rather than the sum of many individual parts (Hotelling, 1947 and Montgomery, 2001).

Montgomery (2001) gave a demonstration that multivariate control charts will produce an acceptable type I error or in-control run length while maintaining the original data of means and

variances and correlations. Multivariate statistics also consider the relationship between the variables since the variance-covariance matrix is part of the computations (Hotelling, 1947). And hence, multivariate charts can detect changes in the relationships among variables being monitored, which would not be noticeable from separate univariate chart (Lowry and Montgomery, 1995).

Also, MSPC provides the appropriate control region for the application. If the assumption of independence does not hold, then the assumed performance of the traditional Shewhart approaches can be misleading. The MSPC can guarantee error protection from a variety of different types of shifts in the process. Another advantage of the MSPC is that it moves away from the application of run rules (Sullivan and Woodall, 1996).

1.6.2 Disadvantages of MSPC

Much satisfying evidence has been presented concerning the benefits of applying the MSPC, the following limitations were noticed.

According to Mason *et al.,* (1997a), Ryan (2000) and Montgomery (2001), multivariate control charting procedures are computationally intensive. And hence, it works well when the number of variables is not too large, that is when $p > 10$. As the number of variables grows, multivariate control chart lose its efficiency with regard to process shift detection. Also, a multivariate control chart procedure does not directly provide the information an operator need when the control chart signals an out-of-control condition. It doesn't give information on which variable or set of variables is out-of-control (Hawkins, 1991).

1.7 APPLICATION OF MULTIVARIATE QUALITY CONTROL

Control charts are originally developed for individual processes and have been applied within a number of areas, including:

1. Hospital infection control (Sellick,1993)
2. Prediction of business failures (Theodossiou, 1993)
3. Monitoring the impact of human disturbance of ecological systems (Anderson and Thompson, 2004)
4. Quality Management of higher education (Mergen *et al.,*2000)
5. Corroborating bribery (Charnes and Gitlow, 1995)
6. Improving athletic performance (Clark and Clark, 1997)
7. Improving the quality of Pharmaceutical products (Ulen and Demir, 2013)

CHAPTER TWO

LITERATURE REVIEW

2.0 INTRODUCTION

This chapter seeks to review related literatures on the works of different scholars regarding multivariate statistical quality control chart. Many authors have proposed quality control procedures for several related variables that led to the development of multivariate quality control.

2.1 MULTIVARIATE CHART

Many control charts have been proposed for multivariate data, with the most popular being Hotelling's T^2 or χ^2 chart, the Multivariate Exponentially-Weighted Moving Average (MEWMA) chart and the Multivariate Cumulative Sum (MCUSUM) chart. Mason *et al*., (1997b) presented an assessment of many multivariate techniques. In their paper, they recommended when multivariate charts should be used and also discussed some problems associated with the selection of multivariate charts. The problems involve the violation of the assumption of multivariate normality that is required for many charts, the estimation of the covariance matrix, missing data, and the effect of autocorrelation. Bersimis *et al.,*(2007) provided a review on some multivariate charts.

2.1.1 Hotelling's T^2 Control Chart

Hotelling's T^2 control chart is the most common multivariate control chart which was proposed by Hotelling (1947). Hotelling's T^2 control chart has attracted much attention in the area of multivariate quality control by researchers. Hotelling's T^2 control chart is purposely used for monitoring the mean vector of quality characteristics of a process. There are two types of T^2 chart, one for subgroup data and the other for individual observations.

Assuming there are p variables to be monitored. Let the p variables be independent and identically distributed random variables of size (n) is taken at each time period and the vector of sample means $\bar{X} = (\bar{X}_1, \bar{X}_2, ..., \bar{X}_p)'$ is calculated. If the population has a multivariate normal distribution with a mean $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$ when the mean vector and the covariance matrix are known, the multivariate analogue of the t statistic, namely the T^2 statistic is given as;

$$T^2 = n(\bar{X} - \bar{\bar{X}})S^{-1}(\bar{X} - \bar{\bar{X}})'$$

where

$\bar{X} = (\bar{X}_1, \bar{X}_2, ..., \bar{X}_p)$ is the sample means for each variable,

$\bar{\bar{X}} = (\bar{\bar{X}}_1, \bar{\bar{X}}_2, ..., \bar{\bar{X}}_p)$ is the overall mean vector computed from the in-control samples where

$\bar{\bar{X}}_j = \frac{1}{m}\sum_{k=1}^{m} \bar{X}_{jk}$, j=1,…,$p$ and S is the sample covariance matrix defined as

$$S = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(X_i - \bar{X})'$$

The upper control limit of the T^2 control chart for a phase I (retrospective) analysis is given as

$$UCL = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{\alpha,p,mn-m-p+1}$$

where $UCL = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{\alpha,p,mn-m-p+1}$ represents the $100(1-\alpha)^{th}$ percentile of the F distribution with parameters p and $mn-m+p+1$.

For a phase II (monitoring phase) analysis, the upper control limit for the T^2 chart is given by

$$UCL = \frac{p(m+1)(n-1)}{mn-m-p+1} F_{\alpha,p,mn-m-p+1}$$

When monitoring individual vectors and not subgroups (n=1), the T^2 charting statistic reduces to

$$T^2 = (X - \bar{X})' S^{-1} (X - \bar{X})$$

and the upper control limit is

$$UCL = \frac{p(m+1)(m-1)}{m^2 - mp} F_{\alpha, p, m-p}$$

Some researchers use the upper control limit in phase I presented by Tracy *et al.,* (1992) given by

$$UCL = \frac{(m-1)^2}{m} F_{\alpha, p/2, (m-p-1)/2},$$

where $F_{\alpha, p/2, (m-p-1)/2}$ represents the $100(1-\alpha)^{th}$ percentile of the F distribution with parameters $p/2$ and $(m-p-1)/2$.

The lower control limit in phase I and phase II for monitoring individual and subgroup observations is given by

LCL= 0.

Since the test statistic is a generalized measurement of distance, the lower control limit is always zero. The reason for this is that any shift in the mean will always lead to an increase in the T^2 statistic, and hence the lower control limit may be neglected. If the computed T^2 statistic exceeds the upper control limit, then the process mean is out-of-control.

Mason *et al.,* (2001) studied the effectiveness of using the T^2 control chart for batch (subgroup) processes. They recommended that when batch data are collected from the same multivariate normal distribution, T^2 statistic should be applied for detecting out-of-control signal. Chou and Mason (2001) studied the effect of multivariate non-normality on the T^2 chart, showing

that the Phase II UCL for the chart based on the *F* distribution may be very inaccurate. They used a kernel smoothing technique to estimate the distribution of the T^2 statistic so that a more accurate UCL could be found. They also developed a sample size requirement for data taken from a multivariate exponential distribution. Sullivan and Woodall (1996) recommended the use of sample mean vector and covariance matrix. If any value of T^2 statistic exceeds an upper control limit an out-of-control signal is generated.

Aparisi *et al.,* (2004) applied the T^2 chart to runs rules to increase its power to detect small to moderate shifts in the process. Champ *et al.,* (2005) studied the effect of parameter estimation on the T^2 chart. They found that the traditional control limits based on the *F* distribution creates a chart that is not quick in detecting a change when the parameters are estimated. They also developed corrected control limits for the chart such that estimated parameters and parameters assumed to be known will achieve a nearly identical ARL.

Mason *et al.,*(1997b) constructed multivariate profile chart by superimposing an $\bar{x}$ chart of univariate statistics on the T^2 chart. By performing discrimination analysis, it enables distinguishing in-control conditions from out-of-control conditions. It provides an idea on where the assignable causes are occurring. Lowry and Montgomery (1995) proposed poly plots and multivariate control webs to superimpose univariate statistics on multivariate statistics in order to test trends in individual statistics and realize how they affect other variables.

Jarrett and Pan (2007) used the T^2 chart to monitor the residuals of a vector autoregressive model. Champ and Aparisi (2007) presented two double sampling T^2 charts.

2.1.2 Multivariate Exponentially–Weighted Moving Average Control Chart

The multivariate exponentially-weighted moving average (MEWMA) chart is the multivariate extension of the popular univariate exponentially-weighted moving average (EWMA) chart. The MEWMA chart was initially proposed by Lowry *et al*., 1992. Unlike Hotelling's chart, which is based solely on the most recent observation, the MEWMA chart uses information from the recent history of multiple observations up until the current time point. This enables the chart to detect smaller shifts in the process mean. The MEWMA chart is generally used in phase II with individual data and uses the statistic

$$T_i^2 = Z_i' \Sigma_{Z_i}^{-1} Z_i$$

where

$$Z = \lambda x_i + (1-\lambda) Z_{i-1}$$

and the covariance matrix is given by

$$\Sigma_{Z_i} = \frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2i}]\Sigma$$

with the scalar constant λ, $0 < \lambda \leq 1$ (which may be adjusted to change the weighting of the past observations) x_i, is the vector of observations at point *i* and $Z_0 = 0$. Tables are provided in multiple papers for the upper control limit (Prabhu and Runger, 1997).

Sullivan and Stoumbos (2001) studied the process of achieving robust performance with the chart. Stoumbos and Sullivan (2002) further studied the robustness of the MEWMA chart to multivariate non-normality where they showed that the chart often performs poorly when the normality assumption is violated with common values of the charting constant (greater than 0.05).

The charting constant may be chosen in a way so that similar average run lengths are achieved under a wide range of distributions. The values they recommend for the charting constant are often very small, which means putting the majority of the weight on the past observations (instead of the most current). Champ and Jones-Farmer (2007) studied the properties of the MEWMA control chart when parameters are estimated. Lee and Khoo (2006), studied the optimal design of MEWMA charts using the average run length and the median run length. Joner, Jr. *et al.,*(2008) developed a one-sided MEWMA control chart for health surveillance.

2.1.3 Multivariate Cumulative Sum Control Chart

The Multivariate Cumulative Sum (MCUSUM) chart is the multivariate extension to the univariate Cumulative Sum (CUSUM) chart. Similar to the MEWMA, the MCUSUM charting statistic uses information from multiple past observations as opposed to only the previous observation, giving it the ability to detect smaller shifts. Many MCUSUM control charts have been proposed, including those by Woodall and Ncube (1985), Healy (1987), Crosier (1988), and Pignatiello and Runger (1990). These charts have also been a focus of a substantial amount of research, but the consensus seems to be that the EWMA charts are easier to apply and have similar efficiency as the CUSUM charts. Champ and Jones-Farmer (2007) studied the effects of parameter estimation on these charts, along with the effect on the MEWMA chart.

2.2 IDENTIFYING OUT- OF- CONTROL VARIABLE

If a univariate control chart gives an out-of-control signal, one can easily detect the problem and give a solution to the problem. This is different in the case of a multivariate control chart, since the number of variables is more than one. Identifying an out-of-control variable or variables after a multivariate control chart signals has been an interesting topic for many

researchers over the last few years. In this section, methods for detecting which of the *p* different variables are out-of-control are presented.

2.2.1 Using Bonferroni Control Limits.

The first approach to this problem was proposed by Alt (1985), who recommended using a set of Bonferroni limits on each *p* individual variable as the method of choice for interpreting an out-of-control signal on a multivariate control chart. Hence, the *p* individual control charts would be constructed, each with a probability that the test statistic plots beyond the control limits under an in-control state equal to α/p and not α. Doganaksoy *et al.,* (1991) later made an improvement on this technique by combining several procedures that resulted in a priority ranking of the variables to be investigated along with the use of Bonferroni type limits. The result is a procedure that is more likely to indicate that a particular variable is out of control than the traditional use of Bonferroni intervals.

Hayter and Tsui (1994) buttressed the idea of Bonferroni-type control limits by providing a procedure for exact simultaneous control confidence intervals for each of the variable means, using simulation. Hence, for a known variance–covariance matrix **Σ** and a chosen probability that the test statistic plots beyond the control limits under an in-control state, equal to α, the experimenter first evaluates the critical point $C_{R,\alpha}$. The choice of $C_{R,\alpha}$ depends on the correlation matrix R. The authors give guidance and various tables for choosing the critical point $C_{R,\alpha}$, following any observation vector X^t, the experimenter constructs intervals for the statistic $(X-\sigma_i C_{R,\alpha}, X+\sigma_i C_{R,\alpha})$, where σ_i is the standard deviation of the i^{th} variable, for each of the *p* variables. This procedure triggers an alarm when $M=\max[\mid X_i-\mu_{0i}\mid/\sigma_i]>C_{R,\alpha}$, i=1,2,…,p. A graphical control display can be created by charting the M statistic for each

multivariate observation. This method offers easy identification of errant variables as well as easy quantification of any changes in the variable means.

Alt (1985) and Jackson (1991) discussed the use of an elliptical control region. However, this process has the disadvantage that it can be applied only in the special case of two quality characteristics. In this case, an elliptical control region is constructed where the elliptical region is centered at $\mu_0' = (\mu_1, \mu_2)$ and can be used in place of the Phase II χ^2 - chart. All points lying on ellipse have the same value of χ^2. The χ^2 - chart gives a signal every time the process is out of control, while the elliptical region is useful for indicating which of the variables led to an out-of-control signal.

Chua and Montgomery (1992), used the elliptical control region to provide a solution to the interpretation problem. They use a MEWMA control chart for identifying out-of-control observations and the hyperplane method for identifying the variable or variables that caused the problem. Mader *et al*.,(1996) presented the use of the elliptical control region for power supply calibration as a process-monitoring technique.

Sepulveda and Nachlas (1997), introduced a new control chart called the simulated minimax control chart. This chart provides evidence about which variable caused the out-of-control signal. The chart monitors the maximum and minimum standardized sample mean of samples taken from the multivariate process. It is assumed that the data are normally distributed and that the variance–covariance matrix is known and constant over time. Hence, by monitoring the maximum and minimum standardized sample mean, an out-of-control signal is directly connected with the corresponding out-of-control variable. Sepulveda and Nachlas (1997) also discussed the statistical properties and the Average Run Length (ARL) performance of the minimax control chart.

2.2.2 Application of Principal Components

Principal components Analysis (PCA) can be used to check which of the p variables are responsible for an out-of-control signal. Various methods have been proposed which use principal components for interpreting an out-of control signal. The most common practice is to use the first k most significant principal components, in the case that a T^2 control charts gives an out-of-control signal.

Jackson (1985) and Pignatiello and Runger (1990) recommended the use of principal components to help in the interpretation of an out-of-control signal. They pointed out that by using both the individual variables and the principal components with the univariate charts, the information about the correlation effect of the variables is not lost. Tracy *et al.,*(1992) expanded the previous work and provided an interesting bivariate setting in which the principal components have meaningful interpretations.

Kourti and MacGregor (1996) provided another approach based on principal component analysis and partial least squares. The T^2 is expressed in terms of normalized principal components scores of the multinormal variables. When an out-of-control signal is received, the normalized score with high values are detected, and contribution plots are used to find the variables responsible for the signal. A contribution plot indicates how each variable involved in the calculation of that score contributes to it. Computing variable contributions eliminates much of the criticism that principal components posed on lack of physical interpretation. This approach is particularly applicable to large ill conditioned data sets due to the use of principal components.

Wasterhuis *et al.,* (2000) proposed the contribution plots which can be constructed for the normalized principal component scores with high values, to find the variables responsible for the out-of-control signal.

Maravelakis *et al.,* (2002) proposed a new method based on PCA to identify the variable or variables responsible for an out-of-control signal in the χ^2 control chart. Theoretical control limits were derived and a detailed investigation of the properties and limitations of the new method was given. Also, a graphical technique which is applicable to these limiting situations was provided. Choi *et al*.,(2005) proposed a fault detection technique based on a maximum likelihood-PCA mixture. The basic problem of this method is that the principal components do not always have a physical interpretation.

2.2.3 Application of T^2 Decomposition

Many authors have suggested using decomposition techniques for identifying variables that cause an out-of-control signal. The T^2 decomposition technique is considered to be the most promising method for detecting variable(s) that significantly contributes to a signal.

Mason *et al.,* (1995) proposed the use of T^2 decomposition as a technique for identifying out-of-control signal. The T^2 decomposition is considered to be the most valuable technique in detecting influential variable(s) that is responsible for an out-of control signal. The main idea behind this method is to decompose the T^2 statistic into independent parts, each of which reflects the contribution of an individual variable.

Mason *et al.,* (1997a), provides an appropriate computing scheme that can greatly reduce the computational effort. This method was designed to deal with individual observations, but it can easily be generalized to handle rational subgroups.

Mason *et al.,* (1996) presented a control technique for monitoring a step process, which is based on a double decomposition of Hotelling's T^2 statistic.

Mason *et al.,*(1997b), presented an improved method for interpreting the unique components resulting from applying the MYT decomposition (Mason *et al.,* 1995) to a signaling

T^2 statistic. In their work, they recommended a shortened sequential computational scheme for isolating the variables contributing to the signal using a bivariate decomposition.

Mason and Young (1999) showed using T^2 decomposition, that by improving the model specification at the time that the historical data set is constructed, it may be possible to increase the sensitivity of the T^2 statistic to signal detection. This is accomplished by using plots that are similar to cause-selecting control charts. Yarmohammadi and Ebrahimi (2010), used the beta distribution to improve the detection of out-of-control observations especially when the number of observations is small and later applied the MYT decomposition using three process variables.

Murphy (1987) proposed a method that came from the idea of discriminant analysis and uses the overall T^2 value, comparing it with a T_{p1}^2 value based on a subset of p_1 variables, which are suspected to be responsible for the out-of-control signal. Then, T_p^2 is the full square distance and T_{p1}^2 is the reduced distance corresponding to the subset of the p_1 variables which are suspected of being associated with the out-of-control signal. The difference is calculated finally as, $D = T_p^2 - T_{p1}^2$, having a χ^2 distribution with p_1 degrees of freedom, under the null hypothesis that the subvector $\bar{X}_{p1}$ follows p_1 dimensional distribution with mean μ_{01} and variance-covariance Σ_{01}.

Doganaksoy *et al.,*(1991) introduced the univariate t statistic ranking procedure using the test statistic

$$t = (\bar{\bar{X}}_f - \bar{\bar{\bar{X}}})[S(n_f^{-1} + n^{-1})]^{-\frac{1}{2}}$$

where *f* stands for the observation that gave the out-of-control signal, *n* is the sample size, $\overline{\overline{X}}$ is the sample mean and S is the standard deviation of the reference data. This method is based on the use of *p* unconditional T^2 terms.

Wierda (1994) recommended a step-down procedure, assuming that there is an *a priori* ordering (which variable is the most sensitive to shifts) among the means of the *p* variables and sequentially tested subsets using this ordering to determine the sequence. The test statistic has the form

$$F_j = (T_j^2 - T_{j-1}^2)\{1 + [T_{j-1}^2 / (n-1)]\}^{-1}$$

where T_j^2 represents the unconditional T^2 for the first *j* variables in the chosen group, F_j is the charting statistic under the null hypothesis and follows an $(n_f - 1)j(n_f - j)^{-1}F_{pj,n_f-j}$ distribution. This procedure can be considered as an alternative to using the regular T^2-chart and not only as a supplement.

Timm (1996) proposed the use of finite intersection tests (FITs). He assumes that there is *a priori* ordering among the means of the *p* variables. Although T^2 is optimal for finding a general shift in the mean vector, it is not optimal for shifts that occur for certain subsets of variables. Runger *et al*.,(1996) simplified previous recommendations given by Wierda (1994) and Timm (1996), considering all subsets of variables.

2.2.4 Application of Neural Networks

Hwarng (2008) and Hwarng and Wang (2008) proposed a neural network based control chart called the Neural Network Identifier (NNI), for multivariate autocorrelated data. Their chart indicates the individual out-of-control variables that are responsible for the signal. The NNI is an interesting approach to forecasting and control.

Alfaro *et al*., (2009) proposed an alternative method for interpreting out of control variables called "Classification problem". It is solved through the application of booting with classification Trees. The classifier is then used to determine which variable or variables caused the change in the process.

Verron *et al.,* (2010) developed a new method of detection and isolation using Bayesian network. They combined two approaches which are the causal decomposition of the T^2 statistic and the detection of fault with Bayesian network. According to them, the method permits the isolation of variables implicated in the fault.

Aparisi and Sanz (2010) developed a software interpretation based on the use of neural networks for MEWMA chart. They described that user of the MEWMA control chart has an easy tool that helps to take a decision when the MEWMA control chart detects an out-of-control state.

The potential drawback to this approach is that a very large training (in-control) data set is required to create the chart. The sample size required also increases as the number of variables being monitored increases.

2.2.5 Using Graphical Techniques

Blazek *et al.,* (1987) developed a graphical approach based on a polyplot that is very helpful in determining which variables is out-of-control. The length of the range on the polyplot provides information about the value of each original variable relative to its target value. It is customary to use a T^2 statistic on the polyplot. This can be replaced with a more effective MEWMA statistic.

Fuchs and Benjamini (1994) presented a method for simultaneously controlling a process and interpreting out-of-control signals. The Multivariate Profile (MP) chart is a symbolic scatter plot in which a symbol is constructed for each group of observations. The symbol used represent

the adoption of a profile plot that encodes visually the size and sign of each variable from its reference value. Summaries of data for individual variables are displayed via a symbol, and global information about the group is displayed by means of the location of the symbol on the scatter plot. The graphical display used stressed the need for fast interpretation of an out-of-control signal.

Sparks *et al.,* (1997), presented a method for monitoring multivariate process data based on the Gabriel biplot. This approach allows them to detect changes in location, variation and correlation structure accurately but still display concisely a large amount of information. They illustrated the use of the biplot on an example of industrial data. Nottingham *et al.,* (2001), developed radial plots as SAS based data visualization tool that can improve process control practitioner's ability to monitor, analyze and control a process.

The main drawback to these graphical methods is that their operation is tedious and cumbersome, because of their own graphical nature. Nevertheless, the main problem is that a graphical approach requires the user to interpret the results. This means that it is a complex task to measure their effectiveness in an objective way.

2.2.6 Cause-Selecting Control Chart and Regression Adjusted Variables.

Hawkins (1991) developed a procedure called the regression-adjustment. The scheme involves plotting univariate control charts of the residuals from each variable obtained when that variable is regressed on all the others. Hawkins showed that the performance of ARL for this scheme competes with other methods depending on the type of control charts applied to the residuals.

Wade and Woodall (1993) also use the same method of regression adjusted where they considered a two-step process in which the steps are not independent. In particular, when the

incoming variable X_1 (the first step of the process) is charted, the outgoing quality X_2 (the second step of the process) is monitored after adjusting for the incoming quality X_1. A chart for X_1 and $Z = X_2 - \hat{X}_2$ respectively, where $\hat{X}_2$ is the predicted value of X_2 based on the regression line connecting X_1 and X_2.

Kalagonda and Kulkarni (2003) proposed a diagnostic procedure called 'D-technique,' using dummy variables in a multiple-regression method. Their method enables the identification of causative factors, such as the mean shift or relationship shift which responsible for out-of-control signal. The method also indicates the direction of the mean shift (whether the mean is increased or decreased).

CHAPTER THREE

METHODOLOGY

3.0 INTRODUCTION

Hotelling's T^2 control chart is the most commonly used multivariate control charts, when all the quality characteristics are normally distributed.

Let $X_1, X_2, ..., X_p$ be the p quality characteristics we are interested to monitor and assuming that they are normally distributed with multivariate mean, μ and covariance matrix Σ .

Then the multivariate analogue of t statistic is

$$t^2 = \frac{(\bar{X} - \mu_0)^2}{s^2 / n}$$

$$t^2 = n(\bar{X} - \mu_0)' s^{-2} (\bar{X} - \mu_0)$$

When t^2 is generalized to p variables, it becomes

$$T^2 = n(\bar{X} - \mu_0)' \Sigma_0^{-1} (\bar{X} - \mu_0)$$

where

μ_0 is a $(p \times 1)$ vector of population mean

$\bar{X}$ is a $(p \times 1)$ vector of sample mean and

Σ_0 is a $(p \times p)$ variance-covariance matrix

3.1 INTERPRETATION OF OUT-OF-CONTROL

It is not easy to determine which of the p variables is responsible for an out-of-control signal. The standard practice is to plot individual $\bar{X}$ charts on individual variables, which may not be successful due to correlation among variables. However, much research has been carried

out in order to address the issue of interpreting an out-of-control signal of multivariate control chart. Among all the researches carried out, the decomposition of Hotelling's T^2 approach (Mason, 1995), has been widely known and used due to its ability in providing clear interpretation of the situation with respect to the original variable.

This approach is to decompose the T^2 statistic into independent components, each of which reflects the contribution of an individual variable. The general decomposition procedure is outlined below. The T^2 statistic for a p dimensional observation vector $X' = (x_1, x_2, ..., x_p)$ can be represented as

$$T^2 = (X - \bar{X})' S^{-1} (X - \bar{X}) \tag{3.1}$$

Suppose the vector $(X - \bar{X})$ is partition as

$$(X - \bar{X}) = [(X^{(p-1)} - \bar{X}^{(p-1)}), (x_p - \bar{x}_p)] \tag{3.2}$$

where

$X^{(p-1)} = (x_1, x_2, ..., x_{p-1})$ represents the $(p-1)$ dimensional variable vector excluding the p^{th} variable x_p.

$\bar{x}_p$ represents the mean of the p^{th} variable .

and $\bar{X}^{(p-1)}$ represents the corresponding $(p-1)$ elements of the mean vector.

Suppose the matrix S is partition so that

$$S = \begin{pmatrix} S_{XX} & S_{xX} \\ S'_{xX} & S_p^2 \end{pmatrix}, \tag{3.3}$$

where

S_{XX} = the $(p-1)\times(p-1)$ covariance matrix for the first $(p-1)$ variables.

S_p^2 = the variance of x_p and

$S_{xX} = (p-1)$ dimensional vector containing the covariances between x_p and the remaining $(p-1)$ variables.

The T^2 statistic in (3.1) can be partitioned into two independent parts (Rencher, 1993). These components are given by

$$T^2 = T^2_{p-1} + T^2_{p.1,2,...,p-1} \tag{3.4}$$

The first term on the RHS (3.4) can be expressed as

$$T^2_{p-1} = (X^{(p-1)} - \bar{X}^{(p-1)})' S_{XX}{}^{-1} (X^{(p-1)} - \bar{X}^{(p-1)}), \tag{3.5}$$

using the first (p-1) variables and is itself a T^2 statistic.

According to (Mason, *et al.,* 1995), the last term in (3.4) can be shown to be the square of the component of the vector *X* adjusted by the estimates of the mean and standard deviation of the conditional distribution of x_p given $(x_1, x_2, ..., x_{p-1})$. It is given as

$$T^2_{p.1,2,...,p-1} = (x_p - \bar{x}_{p.1,2,...,p-1})^2 \Big/ s^2_{p.1,2,...,p-1}$$

where

$$\bar{x}_{p.1,2,...,p-1} = \bar{x}_p + B_p^{'}(X^{(p-1)} - \bar{X}^{(p-1)}), \tag{3.6}$$

and

$$B_p^{'} = S_{XX}^{-1} s_{xX}$$

is the $(p-1)$ dimensional vector estimate of the coefficients from the regression of x_p on the $(p-1)$ variables $x_1, x_2, ..., x_{p-1}$. It can be shown that the estimate of the conditional variance is given as

$$s^2_{p.1,2,...,p-1} = s^2_p - s^{'}_{xX} S^{-1}_{XX} s_{xX}$$

3.2 THE DECOMPOSITION OF HOTELLING'S T^2 STATISTIC

The decomposition of the Hotelling's T^2 statistic is presented as follows:

From (3.4), recall that,

$$T^2 = T^2_{p-1} + T^2_{p.1,2,...,p-1}$$

Since the first term of the above equation is a T^2 statistic, it can also be separated into two orthogonal parts.

$$T^2_{p-1} = T^2_{p-2} + T^2_{p-1.1,2,...,p-2}.$$

The first term, T^2_{p-2}, is a T^2 statistic on the first (p-2) components of the X vector, and the second term, $T^2_{p-1.1,2,...,p-2}$, is the square of x_{p-1} adjusted by the estimates of the mean and standard deviation of the conditional distribution of x_{p-1} given $(x_1, x_2, ..., x_{p-2})$.

Continuing to iterate and partition in this fashion yields one of the many possible MYT decompositions of a T^2 statistic. It is given by

$$T^2 = T^2_1 + T^2_{2.1} + T^2_{3.1,2} + ... + T^2_{p.1,2,3,...,p-1}. \tag{3.7}$$

The T^2_1 term above is an unconditional Hotelling's T^2 statistic for the first variable of the observation vector X, it is given as,

$$T^2_1 = \frac{(x_1 - \bar{x}_1)^2}{s^2_1} \tag{3.8}$$

The first $(p-1)$ terms of (3.4) correspond to the T^2 value of the subvector $X'_{p-1} = (x_1, x_2, ..., x_{p-1})$, i.e.,

$$T^2_{(x_1, x_2, ..., x_{p-1})} = T^2_1 + T^2_{2.1} + T^2_{3.1,2} + ... + T^2_{p-1.1,2,...,p-2}. \tag{3.9}$$

Similarly, the first $(p-2)$ terms of (3.9) correspond to the subvector $X'_{p-2} = (x_1, x_2, ..., x_{p-2})$,

i.e,

$$T^2_{(x_1,x_2,\ldots,x_{p-2})} = T^2_1 + T^2_{2.1} + T^2_{3.1,2} + \ldots + T^2_{p-2.1,2,\ldots p-3}. \quad (3.10)$$

Using this approach the T^2 values for all subvectors of the original vector X are computed. The decomposition of two process variables as shown by Mason *et al.,* (1997a) as

$$T^2 = \begin{cases} T_1^2 + T_{2.1}^2 \\ T_2^2 + T_{1.2}^2 \end{cases}$$

Sani and Abubakar (2013), presented the decomposition of three variables by (Mason andYoung, 2002) to demonstrate the invariance property of the Hotelling's T^2 statistic,

$$T^2 = \begin{cases} T_1^2 + T_{2.1}^2 + T_{3.1,2}^2 \\ T_1^2 + T_{3.1}^2 + T_{2.1,3}^2 \\ T_2^2 + T_{3.2}^2 + T_{1.2,3}^2 \\ T_2^2 + T_{1.2}^2 + T_{3.1,2}^2 \\ T_3^2 + T_{2.3}^2 + T_{1.2,3}^2 \\ T_3^2 + T_{1.3}^2 + T_{2.1,3}^2 \end{cases}$$

Here in this work, we extended the decomposition of Hotelling's T^2 statistic from three variable (p=3) to four variables (p=4). The number of decompositions increased from 6 to 24 possible ways (i.e. $p!$ = 3! = (6 decompositions) to $p!$ = 4! = (24 decompositions)). These decompositions give the same overall T^2 statistic as stated by (Mason, *et al.,* 1995).

3.3 MODEL FOR THE T^2 DECOMPOSITION USING FOUR VARIABLES

Starting by selecting any one of the p variables, followed by any of the (p-1) remaining variables to condition on the first selected variable, the next step is to select any of the remaining (p-2) variables to condition on the first two selected variables. Finally, any of the (p-3) variables to condition on the first three selected variables are selected. Iterating the same procedure will generate all the decomposition equations which give the same overall T^2 statistic.

For instance, taking the first p variables, assuming we select p=1 is selected (remaining 2nd, 3rd and 4th variables). The decomposition results are presented as follows;

$$T^2 = T_1^2 + T_{2.1}^2 + T_{3.1,2}^2 + T_{4.1,2,3}^2$$

$$T^2 = T_1^2 + T_{3.1}^2 + T_{4.1,3}^2 + T_{2.1,3,4}^2$$

$$T^2 = T_1^2 + T_{4.1}^2 + T_{2.1,4}^2 + T_{3.1,2,4}^2$$

The next step is to keep the selected p variable and the (p-1) remaining variables conditioned on the p selected variable constant (that is, the first and the second decomposition terms in the above decompositions). Then, any of the remaining (p-2) variables not used in the first decomposition to be condition on the first two kept variables is selected. Finally, any of the remaining (p-3) variables not used during the first decomposition to condition on the first three selected variables are selected. These yield the following results:

$$T^2 = T_1^2 + T_{2.1}^2 + T_{4.1,2}^2 + T_{3.1,2,4}^2$$

$$T^2 = T_1^2 + T_{3.1}^2 + T_{2.1,3}^2 + T_{4.1,2,3}^2$$

$$T^2 = T_1^2 + T_{4.1}^2 + T_{3.1,4}^2 + T_{2.1,3,4}^2$$

Computing in the same procedure will generate all the decomposition equations which produce the same overall T^2 statistic. Hence, the complete T^2 decomposition results using four variables ($p = 4$) are presented as follows;

$$T_1^2 + T_{2.1}^2 + T_{3.1,2}^2 + T_{4.1,2,3}^2$$
$$T_1^2 + T_{3.1}^2 + T_{4.1,3}^2 + T_{2.1,3,4}^2$$
$$T_1^2 + T_{4.1}^2 + T_{2.1,4}^2 + T_{3.1,2,4}^2$$
$$T_1^2 + T_{2.1}^2 + T_{4.1,2}^2 + T_{3.1,2,4}^2$$

$$
\begin{aligned}
&T_1^2 + T_{3.1}^2 + T_{2.1,3}^2 + T_{4.1,2,3}^2 \\
&T_1^2 + T_{4.1}^2 + T_{3.1,4}^2 + T_{2.1,3,4}^2 \\
&T_2^2 + T_{1.2}^2 + T_{3.1,2}^2 + T_{4.1,2,3}^2 \\
&T_2^2 + T_{3.2}^2 + T_{4.2,3}^2 + T_{1.2,3,4}^2 \\
&T_2^2 + T_{4.2}^2 + T_{1.2,4}^2 + T_{3.1,2,4}^2 \\
&T_2^2 + T_{1.2}^2 + T_{4.1,2}^2 + T_{3.1,2,4}^2 \\
&T_2^2 + T_{3.2}^2 + T_{1.2,3}^2 + T_{4.1,2,3}^2 \\
&T_2^2 + T_{4.2}^2 + T_{3.2,4}^2 + T_{1.2,3,4}^2 \\
&T_3^2 + T_{1.3}^2 + T_{2.1,3}^2 + T_{4.1,2,3}^2 \\
&T_3^2 + T_{2.3}^2 + T_{4.2,3}^2 + T_{1.2,3,4}^2 \\
&T_3^2 + T_{4.3}^2 + T_{1.3,4}^2 + T_{2.1,3,4}^2 \\
&T_3^2 + T_{1.3}^2 + T_{4.1,3}^2 + T_{2.1,3,4}^2 \\
&T_3^2 + T_{2.3}^2 + T_{1.2,3}^2 + T_{4.1,2,3}^2 \\
&T_3^2 + T_{4.3}^2 + T_{2.3,4}^2 + T_{1.2,3,4}^2 \\
&T_4^2 + T_{1.4}^2 + T_{2.1,4}^2 + T_{3.1,2,4}^2 \\
&T_4^2 + T_{2.4}^2 + T_{3.2,4}^2 + T_{1.2,3,4}^2 \\
&T_4^2 + T_{3.4}^2 + T_{1.3,4}^2 + T_{2.1,3,4}^2 \\
&T_4^2 + T_{1.4}^2 + T_{3.1,4}^2 + T_{2.1,3,4}^2 \\
&T_4^2 + T_{2.4}^2 + T_{1.2,4}^2 + T_{3.1,2,4}^2 \\
&T_4^2 + T_{3.4}^2 + T_{2.3,4}^2 + T_{1.2,3,4}^2
\end{aligned}
\tag{3.11}
$$

The first terms are unconditional while others are conditional. The importance of the above result is that it allows one to examine the T^2 statistic from many different perspectives. From the decomposition, it shows that an increase in the number of variables will lead to an increase in the number of terms. These make computation become too challenging.

Detailed information is given in Table 3.1.

Table 3.1: Unique Decomposition Terms (cited from Mason *et al.*, 1997a)

Number of variables (p)	Number of unique terms $p \times 2^{(p-1)}$
2	4
3	12
4	32
5	80
10	5120
15	245760
20	10485760

The following is a sequential computational scheme that has the potential of further reducing the computations to a reasonable number when the overall T^2 signals, as proposed by Mason *et al.,* 1997a.

Step 0. Conduct a T^2 test with a specified nominal confidence level. If an out-of-control condition is signaled then continue with step 1.

Step 1. Compute the individual T^2 statistic for every component of the X vector. Remove variables whose observations produce a significant T_i^2. With significant variables removed a reduced set of variables is achieved. Check the subvector of the remaining k variables if there is any signal.

Step 2. Optional but useful for very large p: examine the correlation structure of the reduced set of variables. Remove any variable having a very weak correlation (0.3 or less) with all the other variables. The contribution of a variable that falls in this category is measured by T_i^2 component.

Step 3. If a signal remains in the subvector of k variables not deleted, compute all $T^2_{i.j}$ terms. Remove from the study all pairs of variables, (x_i, x_j) that have a significant $T^2_{i.j}$ term. This indicates that something is wrong with the bivariate relationship. When this occurs it will further reduce the set of variables under consideration. Examine all removed variables for cause of the signal. Compute the T^2 terms for the remaining subvector. If no signal is present, the source of the problem is with the bivariate relationships and those variables that were out of individual control.

Step 4. If the subvector of the remaining variables still contains a signal, compute all $T^2_{i.j.k}$ terms. Remove any triple (x_i, x_j, x_k) of variables that show significant results and check the remaining subvector for a signal.

Step 5. Continue computing the higher order terms in this fashion until there are no variables left in the reduced set. The worst case situation is that all unique terms will have to be computed.

3.4 COMPUTING THE MYT DECOMPOSITION TERMS

The values: $T^2_{(x_1,x_2,\ldots,x_p)}, T^2_{(x_1,x_2,\ldots,x_{p-1})}, \ldots, T^2_{(x_1)}$ are computed using the general formula

$$T^2_{(x_1,x_2,\ldots,x_j)} = (X^{(j)} - \bar{X}^{(j)})' S^{-1}_{jj} (X^{(j)} - \bar{X}^{(j)}) \tag{3.12}$$

where

j=1,2,…p

$X^{(j)}$ = represents the appropriate subvector

$\bar{X}^{(j)}$ = is the corresponding subvector mean and

S_{jj} = denotes the corresponding covariance sub matrix obtained from the overall S matrix given above by deleting all unused rows and columns.

One of the MYT decomposition for the observation vector (x_1, x_2, x_3, x_4) is given as

$$T^2_{(x_1,x_2,x_3,x_4)} = T^2_1 + T^2_{3.1} + T^2_{2.1,3} + T^2_{4.1,2,3} \tag{3.13}$$

The computation begins by first determining the value of the conditional term $T^2_{4.1,2,3}$. From (3.13) above,

$$T^2_{4.1,2,3} = T^2_{(x_1,x_2,x_3,x_4)} - T^2_{(x_1,x_2,x_3)} \tag{3.14}$$

From the above expression, the variance-covariance matrix and the mean vector of the observation vector is given as;

$$S_{44} = \begin{pmatrix} S_1^2 & S_{12} & S_{13} & S_{14} \\ S_{21} & S_2^2 & S_{23} & S_{24} \\ S_{31} & S_{32} & S_3^2 & S_{34} \\ S_{41} & S_{42} & S_{43} & S_4^2 \end{pmatrix} \text{ and } \bar{X}^{(4)} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \bar{x}_4 \end{pmatrix} \tag{3.15}$$

Thus the computation of $T^2_{(x_1,x_2,x_3,x_4)}$ is as follows;

$$T^2_{(x_1,x_2,x_3,x_4)} = (X^{(4)} - \bar{X}^{(4)})' S^{-1}_{44} (X^{(4)} - \bar{X}^{(4)}) \tag{3.16}$$

To obtain $T^2_{(x_1,x_2,x_3)}$, the original estimates of the mean vector and covariance structure is partitioned to obtain the mean vector and covariance matrix of the sub vector $X^{(3)} = (x_1, x_2, x_3)$. The corresponding partition is given as

$$S_{33} = \begin{pmatrix} S_1^2 & S_{12} & S_{13} \\ S_{21} & S_2^2 & S_{23} \\ S_{31} & S_{32} & S_3^2 \end{pmatrix} \text{ and } \bar{X}^{(3)} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{pmatrix} \tag{3.17}$$

Thus the computation $T^2_{(x_1,x_2,x_3)}$ is as follows;

$$T^2_{(x_1,x_2,x_3)} = (X^{(3)} - \bar{X}^{(3)})' S^{-1}_{33} (X^{(3)} - \bar{X}^{(3)}) \tag{3.18}$$

Also the decomposition of $T^2_{(x_1,x_2,x_3)}$ is given by

$$T^2_{(x_1,x_2,x_3)} = T^2_1 + T^2_{3.1} + T^2_{2.1,3} \tag{3.19}$$

Equation (3.19) is obtained by first computing the conditional term $T^2_{2.1,3}$ as follows

$$T^2_{2.1,3} = T^2_{(x_1,x_2,x_3)} - T^2_{(x_1,x_3)}$$

(3.20)

To obtain the term $T^2_{(x_1,x_3)}$, the original estimates of the mean vector and covariance structure is partitioned to obtain the mean vector and covariance matrix of the sub vector $X^{(2)} = (x_1, x_3)$. The corresponding partition is given as

$$S_{22} = \begin{pmatrix} S_1^2 & S_{13} \\ S_{31} & S_3^2 \end{pmatrix} \text{and } \bar{X}^{(2)} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_3 \end{pmatrix} \tag{3.21}$$

Hence, the computation of the term $T^2_{(x_1,x_3)}$ is as follows;

$$T^2_{(x_1,x_3)} = (X^{(2)} - \bar{X}^{(2)})' S_{22}^{-1} (X^{(2)} - \bar{X}^{(2)}) \tag{3.22}$$

Also, the decomposition for $T^2_{(x_1,x_3)}$ is given by

$$T^2_{(x_1,x_3)} = T_1^2 + T^2_{3.1} \tag{3.23}$$

Thus, $T^2_{3.1}$ is computed as

$$T^2_{3.1} = T^2_{(x_1,x_3)} - T^2_{(x_1)} \tag{3.24}$$

But $T^2_{(x_1)} = T_1^2$

$$T^2_{3.1} = T^2_{(x_1,x_3)} - T_1^2 \tag{3.25}$$

Hence, the unconditional term T_1^2 is computed by

$$T_1^2 = \frac{(x_1 - \bar{x}_1)^2}{s_1^2} \tag{3.26}$$

$$T_j^2 = \frac{(x_j - \bar{x}_j)^2}{s_j^2} \tag{3.27}$$

j=1,2,…,*p,* is the square of a univariate *t* statistic for the observed value of the j[th] variable of the vector X.

3.5 IDENTIFICATION OF SIGNALING VARIABLES

In identifying variables that significantly contribute to an out-of-control condition, the T^2 signal is interpreted in relation to the MYT decomposition. Assuming a signaling observation vector $X^{'} = (x_1, x_2, ..., x_p)$ such that

$$T^2_{(x_1, x_2, ..., x_p)} > UCL \tag{3.28}$$

Recall that T^2 statistic can be constructed on any subset of the variables $(x_1, x_2, ..., x_p)$. Construct the T^2 statistic for each individual variable X_j , j=1,2,…,*p*. so that

$$T_j^2 = \frac{(x_j - \bar{x}_j)^2}{s_j^2}$$

where $\bar{X}_j$ and S_j^2 are the corresponding mean and variance estimates as determined from the historical dataset. Compare these individual T_j^2 values to the UCL which given as

$$UCL = \left(\frac{p(n+1)(n-1)}{n(n-p)} \right) F_{\alpha, p, n-p}$$

$$= \left(\frac{n+1}{n} \right) F_{\alpha, 1, n-1} \tag{3.29}$$

is computed for an appropriate α level and for a value of p=1. Exclude from the original set of variables all X_j for which

$$T_j^2 > UCL_{(x_j)}$$

Compute the T^2 statistic for all possible pairs of variables. For all (x_i, x_j) where $i \neq j$ compute $T^2_{(x_i, x_j)}$ and compare these values to the upper control limit,

$$UCL_{(x_i,x_j)} = \left(\frac{2(n+1)(n-1)}{n(n-2)} \right) F_{\alpha,2,n-2} \tag{3.30}$$

is computed for an appropriate α level and for a value of $p=2$. Exclude from this group all pairs of variables for which

$$T^2_{(x_i,x_j)} > UCL_{(x_i,x_j)}$$

For all (x_i, x_j, x_k) with $i \neq j \neq k$, compute $T^2_{(x_i,x_j,x_k)}$ and compare these values to the upper control limit, where

$$UCL_{(x_i,x_j,x_k)} = \left(\frac{3(n+1)(n-1)}{n(n-3)} \right) F_{\alpha,3,n-3} \tag{3.31}$$

is computed for an appropriate α level and for a value of $p=3$. Exclude from this group all pairs of variables for which

$$T^2_{(x_i,x_j,x_k)} > UCL_{(x_i,x_j,x_k)}$$

The excluded pair of variables, in addition to the excluded single variables comprises the group of variables contributing to the overall signal. Continue to iterate in this fashion as to exclude from the remaining group all variables of signaling groups of four, five etc. the procedure produce a set of variables that contribute to the signal.

3.6 METHOD OF DATA COLLECTION AND DATA ANALYSIS

To illustrate the application of the derived Hotelling's T^2 decomposition, dataset for the analysis was obtained from an Indomie Company in Northern Nigeria for percentage free fatty acid (% FFA) recorded from four (4) different machines in the production process. In each of the machines, there were one hundred and eighty samples (180) recorded at the same time interval for thirty days. Let x_1, x_2, x_3 and x_4 denote % FFA readings from the four machine readings. The

R software is used for the data analysis with the aid of Multivariate Statistical Quality control (MSQC) package which is downloaded through the Comprehensive R Archive Network (CRAN).

CHAPTER FOUR

RESULTS AND DISCUSSIONS

4.0 INTRODUCTION

This chapter seeks to apply the derived Hotelling's T^2 decomposition for using four variables and to identify the variable(s) that contribute to an out-of-control condition. The Q-Q plot is first constructed to assess the normality assumption on the four variables.

4.1 NORMALITY TEST

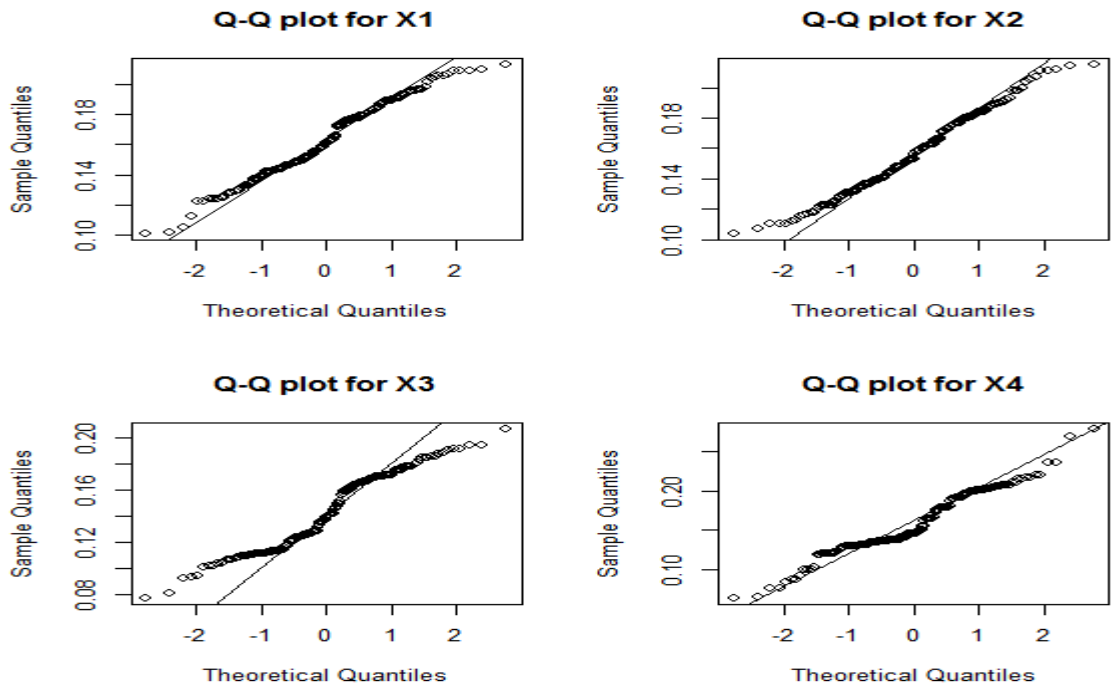


Figure 4. 1:The Q-Q plot of individual variables in the four machines

The plots in figure 4.1 are visual inspection on the four variables. It indicated that the four variables are normally distributed since the points are not too far away from the line.

4.2 HOTELLING'S CONTROL CHART (PHASE I)

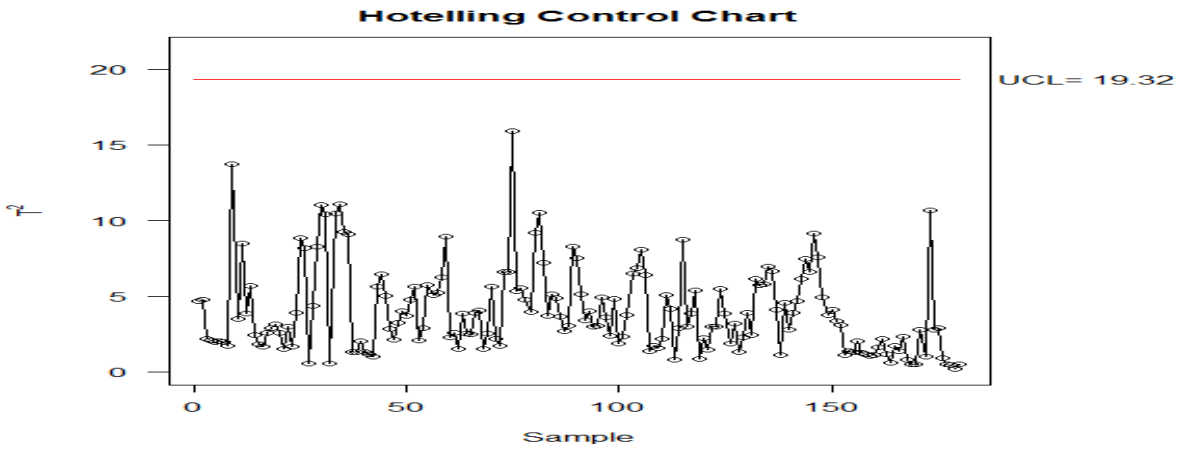


Figure 4. 2:Hotelling's T^2 control chart for historical dataset

The T^2 chart for the historical dataset shows that the monitoring process on the four machines under study is in-control, since none of the T^2 value is above the upper limit of the control chart. The next step is to construct the Phase II control chart to see whether the process is in-control.

4.3 HOTELLING'S CONTROL CHART (PHASE II)

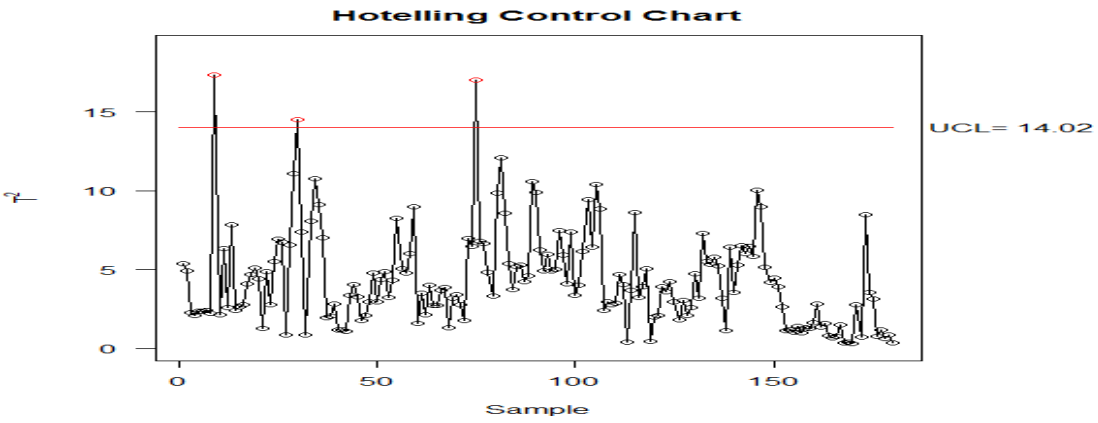


Figure 4. 3:Hotelling's T^2 Control Chart for Phase II

The construction of the Phase II control chart reveals the existence of three points outside the upper control limit (14.02). Thus, we source for the variable(s) which are responsible for the

abnormal observations in the sub vectors. The samples that fall outside the control limits are, sample 9 (T^2=17.34), sample 30 (T^2=14.51) and sample 75 (T^2=17.02). (see Appendix II,III,IV,V,VI).

4.4 COMPUTATION OF THE T^2 DECOMPOSITION TERMS

In Appendix IV, the T^2 decomposition of the observation vector in sample 75 is given. The unconditional T^2 statistic, T_1^2, T_2^2, T_3^2 and T_4^2 is obtained directly from the table. The unconditional T^2 terms are;

T_1^2=0.8067, T_2^2 =0.0138, T_3^2 =0.3122 and T_4^2=10.1150. The conditional Hotelling's T^2 terms $T_{1.4}^2, T_{2.1}^2, T_{1.3,4}^2, T_{2.1,4}^2, T_{1.2,3,4}^2$ and $T_{4.1,2,3}^2$ are obtained as follows;

$T_{1.4}^2 = T_{(x_1,x_4)}^2 - T_4^2$ = 11.0161-10.1150=0.9011

$T_{2.1}^2 = T_{(x_1,x_2)}^2 - T_1^2$=3.7598-0.8067=2.9531

$T_{1.3,4}^2 = T_{(x_1,x_3,x_4)}^2 - T_{(x_3,x_4)}^2$=14.0132-13.9664=0.0468

$T_{2.1,4}^2 = T_{(x_1,x_2,x_4)}^2 - T_{(x_1,x_4)}^2$=14.8638-11.0161=3.8477

$T_{1.2,3,4}^2 = T_{(x_1,x_2,x_3,x_4)}^2 - T_{(x_2,x_3,x_4)}^2$=17.0154-14.6149=2.4005

$T_{4.1,2,3}^2 = T_{(x_1,x_2,x_3,x_4)}^2 - T_{(x_1,x_2,x_3)}^2$=17.0154-5.4042=11.6112

Other conditional terms of the MYT decomposition can be obtained using the same technique as shown above. The most important thing is to correctly decompose the T^2 statistic into orthogonal components.

Table 4.1:MYT decomposition terms of three signaling points.

Components of decomposition	T^2 values of Sample 9	T^2 values of sample 30	T^2 values of sample 75	Critical Value
T_1^2	0.3267	0.2817	0.8067	6.8163
T_2^2	4.8246	3.8462	0.0138	6.8163
T_3^2	0.2390	0.0000	0.3122	6.8163
T_4^2	0.2857	0.6429	**10.1150***	6.8163
$T_{1.2}^2$	**12.3066***	9.5009	3.6218	9.5587
$T_{1.3}^2$	0.1105	0.4942	2.8059	9.5587
$T_{1.4}^2$	0.1144	0.0145	0.9011	9.5587
$T_{2.1}^2$	**16.8045***	**13.0647***	2.9531	9.5587
$T_{2.3}^2$	6.0473	6.5684	0.3892	9.5587
$T_{2.4}^2$	5.0178	3.2607	3.1219	9.5587
$T_{3.1}^2$	0.0228	0.2125	2.3114	9.5587
$T_{3.2}^2$	1.4617	2.7222	0.6876	9.5587
$T_{3.4}^2$	0.0922	0.1167	3.8514	9.5587
$T_{4.1}^2$	0.0734	0.3757	**10.2094***	9.5587
$T_{4.2}^2$	0.4789	0.0574	**13.0989***	9.5587
$T_{4.3}^2$	0.1389	0.7596	**13.6542***	9.5587

$T^2_{1.2,3}$	11.0515	7.7811	4.7028	11.8802
$T^2_{1.2,4}$	11.8288	9.5619	1.6269	11.8802
$T^2_{1.3,4}$	0.0401	0.1497	0.0468	11.8802
$T^2_{2.1,3}$	**16.9883***	**13.8553***	2.2861	11.8802
$T^2_{2.1,4}$	**16.7322***	**12.8081***	3.5988	11.8802
$T^2_{2.3,4}$	6.2467	5.8162	0.6485	11.8802
$T^2_{3.1,2}$	0.2066	1.0031	1.6444	11.8802
$T^2_{3.1,4}$	0.0179	0.2519	2.9971	11.8802
$T^2_{3.2,4}$	1.3211	2.6722	1.3780	11.8802
$T^2_{4.1,2}$	0.0011	0.1191	11.1040	11.8802
$T^2_{4.1,3}$	0.0685	0.4151	10.8951	11.8802
$T^2_{4.2,3}$	0.3383	0.0074	**13.9135***	11.8802
$T^2_{1.2,3,4}$	10.7133	7.9341	2.4005	14.0213
$T^2_{2.1,3,4}$	**16.9199***	13.6006	3.0022	14.0213
$T^2_{3.1,2,4}$	0.2056	1.0444	2.1516	14.0213
$T^2_{4.1,2,3}$	0.0001	0.1604	11.6112	14.0213

* denotes significance at 0.001 level.

Considering sample 9 from Table (4.1) above, the values of the unconditional T^2 component did not indicate any significant contribution. But there is significant contribution

coming from the variance-covariance structure. Since the terms $T^2_{1.2}, T^2_{2.1}, T^2_{1.2,3}, T^2_{1.2,4}, T^2_{2.1,3}, T^2_{2.1,4}$, $T^2_{1.2,3,4}$ and $T^2_{2.1,3,4}$ are relatively high when compared to other components. The values of $T^2_{1.2}$ =12.3066*, $T^2_{2.1}$=16.8045*, $T^2_{2.1,3}$=16.9883*, $T^2_{2.1,4}$=16.7322* and $T^2_{2.1,3,4}$ = 16.9199* have contributed significantly to the out-of-control condition in sample 9. Hence, the quality engineer should look at the correlation structure between variable 1 and variable 2 with special attention on variable 2. Addressing the abnormality between variables 1 and 2 would make the observation vector in sample 9 to be in control.

$T^2 - T^2_{1.2}$= 17.3379-12.3066 =5.0313<14.0213(UCL)

$T^2 - T^2_{2.1}$= 17.3379-16.8045=0.5334<14.0213(UCL)

The above result shows that the sub vector is not significant anymore after removing quality characteristic 1 and 2.

In the case of sample 30, the T^2 value of variable 1 conditioned on variable 2 ($T^2_{1.2}$ =9.5009) is relatively high when compared to other variables conditioned on variable 2, even though it is not significant but it is an indication that the variance-covariance structure deviates from the historical data. The components $T^2_{2.1}$=13.0647*, $T^2_{2.1,3}$=13.8553*and $T^2_{2.1,4}$=12.8081* have significantly contributed to the out-of-control condition in sample 30.

Thus, an isolation of variable 1 and 2 from the process would make the observation vector in sample 30 to be in control.

$T^2 - T^2_{1.2}$= 14.5099-9.5009 =5.009<14.0213(UCL)

$T^2 - T^2_{2.1}$= 14.5099-13.0647=1.4452<14.0213(UCL)

For sample 75, the value of the unconditional T^2 component associated with quality characteristic 4 (T_4^2=10.1150) is the highest. Thus, the operating personnel should focus on quality characteristic 4 for this out-of-control point. Also, it can be seen that the relationship between quality characteristic 4 and other quality characteristics causes a change in variance-covariance structure since the components $T_{4.1}^2, T_{4.2}^2, T_{4.3}^2, T_{4.1,2}^2, T_{4.1,3}^2, T_{4.2,3}^2$ and $T_{4.1,2,3}^2$ are relatively high when compared to other components from Table (4.1) above. Also, $T_{4.1}^2$=10.2094*, $T_{4.2}^2$ =13.0989*, $T_{4.3}^2$ =13.6542 and $T_{4.1,2,3}^2$ =13.9135* are significantly different when compared to their corresponding critical values. Thus, removing variable 4 from the observation vector in sample 75 would make the sub vector to be in-control.

$T^2 - T_4^2$=17.0154-10.1150=6.9004<14.0213(UCL)

Correlation matrix of the reference data, R

$$R = \begin{pmatrix} 1 & 0.9181 & 0.6510 & 0.5228 \\ 0.9181 & 1 & 0.6405 & 0.5104 \\ 0.6510 & 0.6405 & 1 & 0.3926 \\ 0.5228 & 0.5104 & 0.3926 & 1 \end{pmatrix}$$

Variance-covariance matrix, S is

$$S = \begin{pmatrix} 0.00060 & 0.00057 & 0.00046 & 0.00049 \\ 0.00057 & 0.00065 & 0.00047 & 0.00049 \\ 0.00046 & 0.00047 & 0.00082 & 0.00042 \\ 0.00049 & 0.00049 & 0.00042 & 0.00140 \end{pmatrix}$$

and the process mean vector is

$$\bar{X} = \begin{pmatrix} 0.1629 \\ 0.1565 \\ 0.1416 \\ 0.1578 \end{pmatrix}$$

For sample 9, the observation point is

$$X_9 = \begin{pmatrix} 0.146 \\ 0.104 \\ 0.126 \\ 0.140 \end{pmatrix}$$

where $T^2 = (X - \bar{X})' S^{-1} (X - \bar{X}) = 17.3379$ and is computed as ,

$$T_1^2 + T_{3.1}^2 + T_{2.1,3}^2 + T_{4.1,2,3}^2 = 0.3267 + 0.0228 + 16.9883 + 0.0001 = 17.3379$$

For sample 30, the observation point is

$$X_{30} = \begin{pmatrix} 0.147 \\ 0.110 \\ 0.140 \\ 0.130 \end{pmatrix}$$

where $T^2 = (X - \bar{X})' S^{-1} (X - \bar{X}) = 14.5099$ and is computed as,

$$T_1^2 + T_{3.1}^2 + T_{2.1,3}^2 + T_{4.1,2,3}^2 = 0.2817 + 0.2125 + 13.8553 + 0.1604 = 14.5099$$

For sample 75, the observation point is

$$X_{75} = \begin{pmatrix} 0.182 \\ 0.163 \\ 0.124 \\ 0.279 \end{pmatrix}$$

where $T^2 = (X - \bar{X})' S^{-1} (X - \bar{X}) = 17.0154$ and is computed as,

$$T_1^2 + T_{3.1}^2 + T_{2.1,3}^2 + T_{4.1,2,3}^2 = 0.8067 + 2.3114 + 2.2861 + 11.6112 = 17.0154$$

UCL=14.0213

4.5 HOTELLING'S T^2 CONTROL CHART AFTER TAKING OUT ABNORMAL OBSERVATIONS

PHASE I

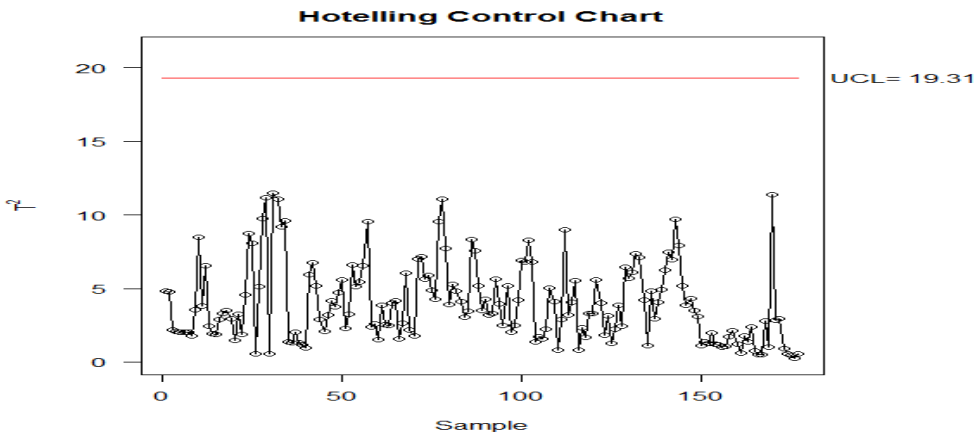


Figure 4. 4:Hotelling's T^2 control chart for 177 observations

The above control chart indicated that Phase I chart is still in-control after removing the three signaling points from sample 9, sample 30 and sample 75.

PHASE II

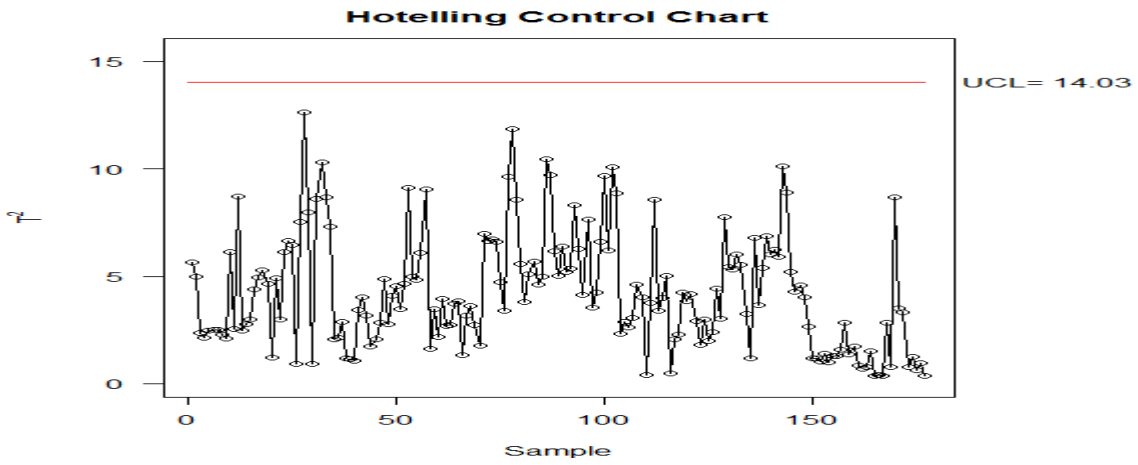


Figure 4. 5:Phase II Hotelling's T^2 control chart for 177 observations

In figure 4.4 it also shows that Phase II chart is in-control when the three signaling points are removed from sample 9, 30 and 75. Thus, the 177 samples are said to be the in-control dataset.

4.6 THE INVARIANCE PROPERTY OF THE HOTELLING'S T^2 STATISTIC.

The invariance property of the Hotelling's T^2 statistic is a property that emphasized that any ordering of an observation vector will produce the same overall T^2 value. Consider the observation vector in sample 75 where $T^2 = (X - \bar{X})' S^{-1} (X - \bar{X})$=17.0154.

The demonstration of the invariance property of the Hotelling's T^2 statistic using equation (3.11) is as follows;

$T_1^2 + T_{2.1}^2 + T_{3.1,2}^2 + T_{4.1,2,3}^2$ =0.8067+2.9531+1.6444+11.6112=17.0154
$T_1^2 + T_{3.1}^2 + T_{4.1,3}^2 + T_{2.1,3,4}^2$ =0.8067+2.3114+10.8951+3.0022=17.0154
$T_1^2 + T_{4.1}^2 + T_{2.1,4}^2 + T_{3.1,2,4}^2$ =0.8067+10.2094+3.8477+2.1516=17.0154
$T_1^2 + T_{2.1}^2 + T_{4.1,2}^2 + T_{3.1,2,4}^2$ =0.8067+2.9531+11.1040+2.1516=17.0154
$T_1^2 + T_{3.1}^2 + T_{2.1,3}^2 + T_{4.1,2,3}^2$ =0.8067+2.3114+2.2861+11.6112=17.0154
$T_1^2 + T_{4.1}^2 + T_{3.1,4}^2 + T_{2.1,3,4}^2$ =0.8067+10.2094+2.9971+3.0022=17.0154
$T_2^2 + T_{1.2}^2 + T_{3.1,2}^2 + T_{4.1,2,3}^2$ =0.0138+3.6218+1.6444+11.6112=17.0154

$T_2^2 + T_{3.2}^2 + T_{4.2,3}^2 + T_{1.2,3,4}^2$ =0.0138+0.6876+13.9135+2.4005=17.0154

$T_2^2 + T_{4.2}^2 + T_{1.2,4}^2 + T_{3.1,2,4}^2$ =0.0138+13.0989+1.6269+2.1516=17.0154

$T_2^2 + T_{1.2}^2 + T_{4.1,2}^2 + T_{3.1,2,4}^2$ =0.0138+3.6218+11.1040+2.1516=17.0154

$T_2^2 + T_{3.2}^2 + T_{1.2,3}^2 + T_{4.1,2,3}^2$ =0.0138+0.6876+4.7028+11.6112=17.0154

$T_2^2 + T_{4.2}^2 + T_{3.2,4}^2 + T_{1.2,3,4}^2$ =0.0138+13.0989+1.3780+2.4005=17.0154

$T_3^2 + T_{1.3}^2 + T_{2.1,3}^2 + T_{4.1,2,3}^2$ =0.3122+2.8059+2.2861+11.6112=17.0154

$T_3^2 + T_{2.3}^2 + T_{4.2,3}^2 + T_{1.2,3,4}^2$ =0.3122+0.3892+13.9135+2.4005=17.0154

$T_3^2 + T_{4.3}^2 + T_{1.3,4}^2 + T_{2.1,3,4}^2$ =0.3122+13.6542+0.0468+2.1516=17.0154

$T_3^2 + T_{1.3}^2 + T_{4.1,3}^2 + T_{2.1,3,4}^2$ =0.3122+2.8059+10.8951+2.1516=17.0154

$T_3^2 + T_{2.3}^2 + T_{1.2,3}^2 + T_{4.1,2,3}^2$ =0.3122+0.3892+4.7028+11.6112=17.0154

$T_3^2 + T_{4.3}^2 + T_{2.3,4}^2 + T_{1.2,3,4}^2$ =0.3122+13.6542+0.6485+2.4005=17.0154

$T_4^2 + T_{1.4}^2 + T_{2.1,4}^2 + T_{3.1,2,4}^2$ =10.1150+0.9011+3.8477+2.1516=17.0154

$T_4^2 + T_{2.4}^2 + T_{3.2,4}^2 + T_{1.2,3,4}^2$ =10.1150+3.1219+1.3780+2.4005=17.0154

$T_4^2 + T_{3.4}^2 + T_{1.3,4}^2 + T_{2.1,3,4}^2$ =10.1150+3.8514+0.0468+2.1516=17.0154

$T_4^2 + T_{1.4}^2 + T_{3.1,4}^2 + T_{2.1,3,4}^2$ =10.1150+0.9011+2.9971+3.0022=17.0154

$T_4^2 + T_{2.4}^2 + T_{1.2,4}^2 + T_{3.1,2,4}^2$ =10.1150+3.1219+1.6269+2.1516=17.0154

$T_4^2 + T_{3.4}^2 + T_{2.3,4}^2 + T_{1.2,3,4}^2$ =10.1150+3.8514+0.6485+2.4005=17.0154

Therefore any of the above decompositions when considered would give an idea concerning the variable that contributes significantly to an out-of-control condition in any four process variable(s).

CHAPTER FIVE

SUMMARY, CONCLUSION AND RECOMMENDATION

5.0 INTRODUCTION

In this chapter, we present the summary, conclusion and recommendation based on the results obtained from this research work.

5.1 SUMMARY

This research work was aimed at deriving the T^2 decomposition model which can be used to identify variable(s) that contribute significantly to out-of-control signal. The model was derived and the invariance property of the Hotelling's T^2 statistic was demonstrated to show that all the 24 decompositions have the same overall T^2 value. This technique is extremely useful in production supervision by providing information as to when and where to control %FFA (percentage Free Fatty Acid) for quality improvement.

Phase I Hotelling's T^2 control chart indicated that the process is in control using the historical dataset while on drawing further samples after the 180^{th} sample, Phase II Hotelling's T^2 control chart shows that three points falls above the control limit and these points are located in sample 9, 30 and 75.

The Hotelling's T^2 control chart is able to account simultaneously for variation in several variables since it requires both an overall measure of deviation from the targets as well as an assessment of the data correlation structure. This was noticed in the relationship between variable 1 and 2 conditioned on other variables as seen in sample 9 and 30. Also variable 4 (X_4) conditioned on the other quality characteristics indicates deviation in the variance-covariance structure with reference to the historical dataset.

It is an established fact that the most difficult part of applying multivariate control chart

is the interpretation of out-of-control variable. In this research work, the T^2 decomposition approach is employed. This is because it provides a comprehensive interpretation of the situation with respect to the original variable, though it is known to be computationally intensive.

In this study, the direct comparison between the values of each unconditional component is used. The quality characteristic with the highest value of unconditional component is the characteristic that significantly contributes to an out-of-control signal. Also, the quality characteristic with the highest value of conditional component is considered to have significantly contributed to an out-of-control signal.

Each of the T^2 components was compared to its critical value to determine if it is significant. The differences between the overall T^2 value for the given observation and the values of each component, was computed. This was to see if the differences are significant or not and it was discovered that the difference between the overall T^2 and any of the unconditional component, T_1^2 or T_2^2 in sample 9 would still leave the sub vector significant. When the conditional component between variables 1 and 2 is removed, it leaves the sub vector not significant. In sample 30, removing the unconditional component T_2^2=3.8462 from the observation vector makes the sub vector not significant. In sample 75, removing variable 1, 2 or 3 leaves the sub vector significant. While removing variable 4 from the observation vector makes the sub vector not significant.

Phase I control chart is constructed after removing the potential causes in the abnormal observation. It shows that the process is in-control. Also, Phase II control chart indicated that the process is in-control when further samples were drawn. And hence the process was considered to be in-control and the dataset was considered to be the in-control historical dataset.

5.2 CONCLUSION

This research is an extension of the T^2 decomposition of the Hotelling's T^2 statistic into orthogonal components from three variables to four variables with the aid of the MYT decomposition approach. The T^2 decomposition is aimed at identifying the variable(s) that significantly contribute to an out-of-control condition. The T^2 decomposition is chosen in this research work due its ability to comprehensively display the potential cause of out-of-control condition.

The use of multivariate control chart is very effective than the univariate control chart, although the interpretation of out-of-control signal in a multivariate control chart is more complicated than the univariate control chart. Thus, the T^2 decomposition is being derived and applied in this research. This is to encourage quality control engineers on the significance of the T^2 decomposition in detecting variables that significantly contribute to out-of-control signal.

The value of each unconditional component is directly compared with its corresponding upper control limit to assess the quality characteristic that significantly contribute to out-of-control condition. Also, quality characteristic that has the highest value of conditional components is considered to have significantly contributed to out-of-control signal. Hence, such quality characteristic is considered to be the potential cause of the abnormal observation.

The invariance property of the Hotelling's T^2 statistic was demonstrated using the derived components where the decompositions had the same overall result irrespective of the ordering of the observation vector.

5.3 RECOMMENDATION

Whenever a multivariate statistical process is to be monitored, it is important to construct the Phase I and Phase II control chart so as not to rush into wrong conclusion. The Phase II

control chart is very sensitive in detecting shifts unlike the Phase I. In this research, we noticed that our Phase I control chart using the historical dataset was unable to detect an out-of-control condition, but the Phase II chart did.

5.4 CONTRIBUTION TO KNOWLEDGE

In this work, the following contributions are made which are vital for academic and industrial consumption.

1. This research work has succeeded in deriving the T^2 decomposition when four variables are considered for monitoring any process control.
2. The decomposition model will help quality control engineers as well as researchers in the field of quality control for monitoring process control.
3. This work has also provided a guide on how to derive models for the T^2 decomposition for monitoring process control beyond three variables.

5.5 FURTHER RESEARCH

1. An extension on this work can be made regarding the decomposition of the Hotelling's T^2 statistics.
2. A program/algorithm can be design for easy computing of the Hotelling's T^2 decomposition.

REFERENCES

Alt F.B., (1985) Multivariate Quality Control. The Encyclopedia of Statistical Sciences, Kotz S., Johnson, NL, Read CR(eds), New York: John Wiley, pp. 110-122.

Alfaro, E., Alfaro, J.L., Gamez, M. and Garcia, N. (2009). A Booting Approach for understanding Out-of-control Signals in Multivariate Control Charts. *International Journal of Production Research* 47(24): 6821-6834

Anderson, M. and Thompson, A. (2004). Multivariate control charts for ecological and environmental monitoring. *Ecological Applications*, 14(6): 1921-1935

Aparisi, F. and Haro, C.L. (2001). Hotelling's T^2 control chart with variable sampling intervals. *International Journal of Production Research*, 39: 3127-3140

Aparisi, F., Champ, C. W., and Garcia-Diaz, J. C. (2004). A performance analysis of Hotelling's T^2 control chart with supplementary runs rules. *Quality Engineering*, 16(3), 359-368.

Aparisi F. and Sanz J., (2010). Interpreting the Out-of-Control Signals of Multivariate Control charts Employing Neural Networks. *World Academy of Science, Engineering and Technology* 37: 226-230.

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

Blazek, L.W., Novic B and Scott M.D.(1987). Displaying Multivariate Data Using Polyplots. *Journal of Quality Technology*, 19(2): 69-74.

Champ, C.W., Jones-Farmer, L.A. and Rigdon, S.E. (2005). Properties of the T^2 control chart when parameters are estimated. *Technometrics*, 47(4): 437-445.

Champ, C. W., and Jones-Farmer, L. A. (2007). Properties of multivariate control charts with estimated parameters. *Sequential Analysis*, 26(2): 153-169.

Champ, C.W. and Aparisi, F. (2007). Double sampling Hotelling's T^2 charts. *Quality and Reliability Engineering International,* 24(2): 153-166

Charnes, J.M. and Gitlow, H. (1995). Using control charts to corrobate bribery in jai alai. *The American Statistician*, 49(4): 386-389.

Choi, S.W, Martin, E.B and Morris, A.J. (2005). Fault detection based on a maximum-likelihood principal component analysis mixture. *Industrial and Engineering Chemistry Research*, 44:2316–2327.

Chou, Y., Mason, R. L. and Young, J. C. (2001). The control chart for individual observations from a multivariate non-normal distribution. *Communications in Statistics - Theory and Methods*, 30(8): 1937-1949.

Chua M.K. and Montgomery D.C., (1992). Investigation and characterization of a control scheme for multivariate quality control. *Quality and Reliability Engineering International*, 8:37-44.

Clark, T. and Clark, A. (1997). Continuous improvement on the free throw line. *Quality Progress*, 30(10): 78-80.

Crosier, R.B. (1988). Multivariate generalizations of cumulative sum quality control schemes. *Technometrics*, 30(3): 291–303.

Del Castillo, E. (2002). Statistical Process Adjustment for Quality control, John wiley and Sons, Inc., New York.

Doganaksoy N., Faltin F.W and Tucker W.T.(1991).Identification of out-of-control multivariate characteristic in a multivariable manufacturing environment. *Communications in Statistics - Theory and Methods*, 20(9): 2775-2790.

Duncan, A.J. (1956). The Economic Design of $\bar{x}$ chart used to maintain current control of a Process. *Journal of the American Statistical Association*, 51(274): 228-242.

Feigenbaum, A.V. (1983). Total quality control, New York, NY: MacGraw-Hill

Fuchs C. and Benjamini Y.(1994) Multivariate profile charts for statistical process control. *Technometrics*, 36(2): 182-195.

Hawkins, D.M. (1991). Multivariate Quality Control based on regression-adjusted variables. *Technometrics*, 33(1): 61-75.

Hayter A.J. and Tsui K.L.,(1994) Identification and quantification in multivariate quality control problems. *Journal of Quality Technology*, 26(3): 197-208.

Healy, J. D. (1987). A note on multivariate CUSUM procedures. *Technometrics*, 29(4): 409-412.

Holmes, D. S. and Mergen, A.E. (1993). Improving the performance of the T^2 chart. *Quality Engineering*, 5(4): 619–625.

Hotelling, H. (1947). Multivariate quality control. In C. Eisenhart, M. W. Hastay, and W. A.Wallis, eds. *Techniques of Statistical Analysis*. New York: McGraw-Hill, pp 111-184.

Jackson, J.E., (1985). Multivarite quality control. *Communication in Statistics.-Theory and Methods* 14(11): 2657-2688.

Jackson, J.E., (1991). Multivariate quality control: 40 years later. Statistical Process control in Manufacturing, Keats J.B, Montgomery, D.C (eds). Marcel Dekker, New York. PP. 123-138.

Jarrett, J. E. and Pan, X. (2007). The quality control chart for monitoring multivariate autocorrelated processes, *Computational Statistics and Data Analysis*, 51(8): 3862-3870.

Joner Jr., M. D., Woodall, W. H., Reynolds Jr., M. R. and Fricker Jr., R. D. (2008). A one-sided MEWMA chart for health surveillance. *Quality and Reliability Engineering International*, 24(5): 503-518.

Kalagonda, A. A., and Kulkarni, S. R. (2003). Diagnosis of multivariate control chart signal based on dummy variable regression technique. *Communications in Statistics – Theory and Methods,* 32(8): 1665-1684.

Kourti, T., and MacGregor, J. F. (1996). Multivariate SPC methods for process and product monitoring. *Journal of Quality Technology,* 28(4): 409-428.

Lee, M. H. and Khoo, M. (2006). Optimal statistical design of a multivariate EWMA chart based on ARL and MRL. *Communications in Statistics: Simulation and Computation,* 35(3): 831-847.

Lowry, C. A. and Montgomery, D.C. (1995). A review of multivariate control charts. *IIE Transactions*, 2(6): 800–810.

Lowry, C.A, Woodall, W.H., Champ C.W. and Ridgon, S.E. (1992). A multivariate exponentially weighted moving average control chart. *Technometrics*, Vol. 34(1): 46-53.

Lu, X.S., M. Xie, T.N. Goh and C.D. Lai. (1998). Control chart for multivariate attribute

processes. *International Journal of Production Research*. 36(12): 3477-3489.

MacCarthy, B.L., and Wasusri, T. (2002). A review of non-statistical applications of statistical process control (spc) charts. *International Journal of Quality and Reliability management* 19(3): 295-320.

Mader,D.P, Glycenfer,J.J. and Prins, J.(1996). An application in multivariate statistical process control for power supply calibration. *Quality Engineering*.9:99–106.

Maravelakis, P. E., Bersimis, S., Panaretos, J., and Psarakis, S. (2002). Identify the out of control variable in a multivariate control chart. *Communications in Statistics – Theory and Methods* 31(2): 2391-2408.

Mason R.L., Tracy N.D. and Young J.C., (1995). Decomposition of T^2 for multivariate control chart interpretation. *Journal of Quality Technology* 27(2): 99-108.

Mason R.L., Tracy N.D. and Young J.C.,(1996). Monitoring a multivariate step process. *Journal of Quality Technology,* 28:39-50.

Mason, R.L. and Young, J.C. (1999). "Improving the sensitivity of the T^2 statistic in multivariate process control", *Journal of Quality Technology*, 31(2): 155-165.

Mason, R.L., Tracy, N.D., and Young, J.C. (1997a). A practical Application for Interpreting Multivariate T^2 Control Chart Signals, *Journal of Quality Technology*, 29(4): 396-501.

Mason, R. L.,Tracy, N.D., and Young, J.C. (1997b). Assessment of multivariate process control techniques, *Journal of Quality Technology*, 29(2): 140-143.

Mason, R. L., Chou, Y., and Young, J. C. (2001). Applying Hotelling's T^2 statistic to batch processes, *Journal of Quality Technology*, 33(4): 466-479.

Mason, R.L. and Young, J.C. (2002). Multivariate Statistical Process Control with Industrial Applications, Philadelphia: ASA-SIAM.

Mergen, E., Grant, O. and Widrick, M. (2000). Quality management applied to higher education. *Total Quality Management*, 11:345-352.

Montgomery, D.C. and Klatt, P.J (1972). Economic design of T^2 control charts to maintain current control of a process. *Management Sciences*, 19:76-89.

Montgomery, D.C. (2001). Introduction to Statistical Quality Control (4^{th} ed.). Canada. John Wiley and Sons, Inc.

Murphy B.J., (1987). Selecting out-of-control variables with T^2 multivariate quality procedures. *The Statistician*, 36: 571-583.

Neter, J., Kutner, M.H., Nachtscheim, C.J., and Wasserman, W. (1996). Applied Linear statistical models, Fourth edition, WCB/MacGraw-Hill companies, Inc., New York.

Nottingham Q.J., Cook D.F. and Zobel C.W.,(2001). Visualization of multivariate data with radial plots using SAS. *Computers and Industrial Engineering*, 41(1): 17-35.

Page, E.S. (1954). Continuous inspection schemes. *Biometrika* 41: 100–114

Pignatiello, J.J. and Runger, G.C (1990). Comparison of multivariate CUSUM charts. *Journal of Quality Technology*, 22:173–186.

Prabhu, S.S. and Runger, G.C. (1997). Designing a Multivariate EWMA Control Chart. *Journal of Quality Technology*, 29:8-15.

Rencher, A.C. (1993). The contribution of individual variables to Hotelling's T^2, Wilks' □ and R^2, *Biometrics*, 49:479-489.

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

Runger G.C., Alt F.B. and Montgomery D.C., (1996). Contributors to a multivariate SPC chart signal. *Communications in Statistics - Theory and Methods*, 25(2): 2203-2213.

Ryan, T.P. (2000). Statistical Methods for Quality Improvement (2nd ed.), John Wiley. New York.

Sani, S.A., and Abubakar, Y., (2013). Demonstrating the Invariance Property of the Hotelling's T^2 statistic. *International Journal of Innovative Research in Science, Engineering and Technology,* 2(7):2515-2519.

Sellick, J.J. (1993). The using of statistical process control charts in hospital epidemiology. *Infection control and Hospital Epidemiology*, 14(11): 649-656.

Sepulveda, A. and Nachlas, J.A.,(1997). A simulation approach to multivariate quality control. *Computers and Industrial Engineering*, 33:113–116.

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

Sparks R.S., Adolphson A. and Phatak A.(1997). Multivariate process monitoring using the dynamic biplot. *International Statistical Review*, 65(3): 325-349.

Sullivan, J.H and Woodall, W.H. (1996). A comparison of multivariate control charts for individual observations. *Journal of Quality Technology*, vol. 28(4): 398-408.

Sullivan, J. H. and Stoumbos, Z. G. (2001). Achieving robust performance with the MEWMA control chart. *ASA Proceedings of the Joint Statistical Meetings, American Statistical Association (Alexandria, VA).*

Stoumbos, Z. G. and Sullivan, J. H. (2002). Robustness to non-normality of the multivariate EWMA control chart. *Journal of Quality Technology*, 34(3): 60-276.

Theodossiou, P. (1993). Predicting the shifts in the mean of a multivariate time series process: An application to predicting business failures. *Journal of the American Statistical Association*, 88: 441-449.

Timm N.H., (1996).Multivariate quality control using finite intersection tests. *Journal of Quality Technology*, 28(2): 233-243.

Tracy N.D., Young J.C. and Mason R.L., (1992). Multivariate control charts for individual observations. *Journal of Quality Technology*, 24(2): 88-95.

Tong, L.I. Wang, C.H. and Huang, C.L. (2005). "Monitoring defects in IC fabrication using a Hotelling T^2control chart", *IEEE Trans. On Semiconductor Manufacturing*, 18(1): 140-148.

Ulen, M. and Demir, I. (2013). Application of multivariate statistical quality control in pharmaceutical industry. *Balkan Journal of Mathematics*, 1(8):93-105.

Vargas, N. J. A. (2003). Robust estimation in multivariate control charts for individual observations. *Journal of Quality Technology*, 35(4): 367–376.

Verron, S, Jing, L. and Teodor, T. (2010). Fault detection and isolation of faults in a multivariate process with Bayesian networks. *Journal of Process control* 20(8): 902-911.

Wade, M.R and Woodall, W.H. (1993). A Review and Analysis of Cause-selecting control charts. *Journal of Quality Technology*, 25(3): 61-170.

Wasterhuis, J. A., Gurden, S. P. and Smilde, A. K. (2000). Generalized contribution plots in multivariate statistical process monitoring. *Chemometrics and Intelligent Laboratory Systems* 51: 95-114.

Walpole, R.E., and Myers, R.H., (1993). Probability and statistics for Engineers and scientists, Fifth edition, Macmillan Publishing company, New York.

Wierda, S.J.(1994). Multivariate statistical process control—recent results and directions for future research. *Statistica Neerlandica*, 48:147–168.

Woodall, W. H. and Ncube, M. M. (1985). Multivariate CUSUM quality control procedures, *Technometrics*, 27(3): 285-292.

Woodall W.H, and Montgomery, D.C., (1999). Research issues and ideas in statistical process control. *Journal of Quality Technology*. 31(4): 376-386.

Woodall W.H (2000). Controversies and contradictions in statistical process control. *Journal of Quality Technology*, 32(4): 341-350.

Yarmohammadi, M. and Ebrahimi, J. (2010). Identification of out-of-control variables in multivariate statistical process control. *Journal of Statistical Modeling and Analytics*, 1(1): 95-112.

APPENDICES

APPENDIX I

Percentage Free Fatty Acid (%FFA) reading for daily dough moulding from four machines

Sample	Time	X_1	X_2	X_3	X_4
1	8:00am	0.125	0.116	0.141	0.144
2	10:00am	0.124	0.120	0.144	0.138
3	12:00am	0.128	0.123	0.109	0.120
4	2:00pm	0.129	0.125	0.109	0.122
5	4:00pm	0.130	0.123	0.111	0.120
6	6:00pm	0.130	0.123	0.110	0.120
7	8:00am	0.133	0.125	0.106	0.122
8	10:00am	0.135	0.127	0.108	0.124
9	12:00am	0.146	0.104	0.126	0.140
10	2:00pm	0.131	0.139	0.128	0.125
11	4:00pm	0.105	0.118	0.112	0.119
12	6:00pm	0.124	0.129	0.103	0.140
13	8:00am	0.132	0.107	0.104	0.128
14	10:00am	0.127	0.122	0.106	0.140
15	12:00am	0.144	0.132	0.106	0.138
16	2:00pm	0.146	0.132	0.111	0.136
17	4:00pm	0.146	0.128	0.109	0.134
18	6:00pm	0.152	0.132	0.111	0.136
19	8:00am	0.161	0.140	0.116	0.137
20	10:00am	0.158	0.138	0.118	0.137
21	12:00am	0.161	0.160	0.121	0.138
22	2:00pm	0.073	0.153	0.123	0.151
23	4:00pm	0.146	0.132	0.111	0.136
24	6:00pm	0.144	0.123	0.112	0.164
25	8:00am	0.101	0.112	0.109	0.13
26	10:00am	0.102	0.110	0.117	0.120
27	12:00am	0.150	0.145	0.142	0.140
28	2:00pm	0.148	0.124	0.142	0.146
29	4:00pm	0.149	0.117	0.140	0.164
30	6:00pm	0.147	0.110	0.140	0.130
31	8:00am	0.155	0.174	0.102	0.132
32	10:00am	0.150	0.145	0.142	0.140
33	12:00am	0.166	0.184	0.147	0.099
34	2:00pm	0.179	0.180	0.093	0.130
35	4:00pm	0.174	0.173	0.092	0.131
36	6:00pm	0.165	0.178	0.102	0.136
37	8:00am	0.165	0.159	0.160	0.134
38	10:00am	0.173	0.166	0.167	0.146

39	12:00am	0.181	0.174	0.178	0.155
40	2:00pm	0.180	0.179	0.171	0.179
41	4:00pm	0.185	0.183	0.166	0.176
42	6:00pm	0.183	0.180	0.166	0.179
43	8:00am	0.165	0.180	0.171	0.179
44	10:00am	0.161	0.177	0.168	0.190
45	12:00am	0.164	0.175	0.169	0.199
46	2:00pm	0.178	0.184	0.172	0.192
47	4:00pm	0.175	0.173	0.177	0.190
48	6:00pm	0.176	0.176	0.181	0.203
49	8:00am	0.163	0.153	0.170	0.205
50	10:00am	0.160	0.163	0.163	0.205
51	12:00am	0.159	0.160	0.175	0.202
52	2:00pm	0.156	0.160	0.175	0.202
53	4:00pm	0.177	0.163	0.163	0.199
54	6:00pm	0.175	0.159	0.168	0.199
55	8:00am	0.180	0.153	0.169	0.191
56	10:00am	0.213	0.212	0.171	0.191
57	12:00am	0.210	0.211	0.170	0.218
58	2:00pm	0.209	0.207	0.162	0.237
59	4:00pm	0.193	0.188	0.165	0.270
60	6:00pm	0.185	0.188	0.170	0.196
61	8:00am	0.202	0.192	0.170	0.195
62	10:00am	0.190	0.182	0.171	0.191
63	12:00am	0.187	0.183	0.185	0.22
64	2:00pm	0.189	0.185	0.180	0.207
65	4:00pm	0.194	0.189	0.178	0.207
66	6:00pm	0.205	0.204	0.175	0.215
67	8:00am	0.206	0.205	0.164	0.203
68	10:00am	0.179	0.177	0.160	0.201
69	12:00am	0.196	0.187	0.166	0.211
70	2:00pm	0.177	0.191	0.165	0.206
71	4:00pm	0.197	0.189	0.160	0.196
72	6:00pm	0.188	0.184	0.159	0.202
73	8:00am	0.190	0.181	0.121	0.218
74	10:00am	0.172	0.166	0.126	0.237
75	12:00am	0.182	0.163	0.124	0.279
76	2:00pm	0.193	0.175	0.128	0.196
77	4:00pm	0.181	0.165	0.113	0.195
78	6:00pm	0.189	0.184	0.123	0.191
79	8:00am	0.177	0.177	0.137	0.22
80	10:00am	0.196	0.185	0.111	0.207
81	12:00am	0.197	0.177	0.11	0.207
82	2:00pm	0.195	0.177	0.125	0.215
83	4:00pm	0.198	0.181	0.182	0.203
84	6:00pm	0.191	0.200	0.188	0.201

85	8:00am	0.176	0.171	0.190	0.211
86	10:00am	0.178	0.159	0.164	0.206
87	12:00am	0.176	0.158	0.161	0.196
88	2:00pm	0.178	0.160	0.163	0.202
89	4:00pm	0.209	0.193	0.194	0.136
90	6:00pm	0.207	0.190	0.184	0.134
91	8:00am	0.197	0.190	0.185	0.136
92	10:00am	0.183	0.171	0.178	0.137
93	12:00am	0.176	0.159	0.177	0.137
94	2:00pm	0.178	0.161	0.168	0.138
95	4:00pm	0.179	0.161	0.173	0.151
96	6:00pm	0.161	0.137	0.161	0.135
97	8:00am	0.188	0.166	0.170	0.164
98	10:00am	0.178	0.163	0.151	0.130
99	12:00am	0.179	0.158	0.161	0.120
100	2:00pm	0.175	0.162	0.165	0.140
101	4:00pm	0.178	0.163	0.170	0.146
102	6:00pm	0.184	0.161	0.169	0.164
103	8:00am	0.191	0.166	0.166	0.130
104	10:00am	0.152	0.151	0.081	0.132
105	12:00am	0.153	0.134	0.101	0.076
106	2:00pm	0.154	0.131	0.103	0.099
107	4:00pm	0.164	0.153	0.126	0.13
108	6:00pm	0.163	0.150	0.121	0.131
109	8:00am	0.176	0.165	0.139	0.136
110	10:00am	0.155	0.148	0.163	0.134
111	12:00am	0.147	0.149	0.175	0.146
112	2:00pm	0.162	0.162	0.184	0.155
113	4:00pm	0.173	0.173	0.156	0.179
114	6:00pm	0.173	0.165	0.184	0.176
115	8:00am	0.156	0.149	0.077	0.179
116	10:00am	0.137	0.129	0.121	0.179
117	12:00am	0.137	0.131	0.126	0.19
118	2:00pm	0.134	0.131	0.124	0.199
119	4:00pm	0.145	0.144	0.128	0.155
120	6:00pm	0.148	0.144	0.113	0.179
121	8:00am	0.152	0.141	0.123	0.176
122	10:00am	0.143	0.130	0.137	0.179
123	12:00am	0.144	0.132	0.111	0.179
124	2:00pm	0.133	0.137	0.110	0.190
125	4:00pm	0.146	0.148	0.125	0.199
126	6:00pm	0.142	0.139	0.145	0.162
127	8:00am	0.144	0.142	0.157	0.173
128	10:00am	0.189	0.181	0.159	0.165
129	12:00am	0.183	0.179	0.137	0.149
130	2:00pm	0.188	0.182	0.144	0.129

131	4:00pm	0.182	0.176	0.149	0.131
132	6:00pm	0.125	0.113	0.149	0.131
133	8:00am	0.112	0.110	0.123	0.144
134	10:00am	0.193	0.198	0.165	0.133
135	12:00am	0.206	0.215	0.178	0.167
136	2:00pm	0.204	0.214	0.170	0.175
137	4:00pm	0.192	0.198	0.187	0.181
138	6:00pm	0.142	0.137	0.113	0.147
139	8:00am	0.189	0.171	0.191	0.187
140	10:00am	0.152	0.146	0.146	0.102
141	12:00am	0.177	0.167	0.171	0.118
142	2:00pm	0.189	0.171	0.191	0.192
143	4:00pm	0.209	0.211	0.194	0.166
144	6:00pm	0.184	0.191	0.206	0.206
145	8:00am	0.191	0.198	0.16	0.128
146	10:00am	0.159	0.156	0.156	0.064
147	12:00am	0.154	0.146	0.136	0.063
148	2:00pm	0.149	0.148	0.141	0.083
149	4:00pm	0.128	0.123	0.114	0.087
150	6:00pm	0.137	0.134	0.135	0.087
151	8:00am	0.142	0.137	0.135	0.092
152	10:00am	0.137	0.135	0.094	0.145
153	12:00am	0.142	0.137	0.113	0.147
154	2:00pm	0.140	0.137	0.111	0.145
155	4:00pm	0.142	0.140	0.112	0.138
156	6:00pm	0.138	0.138	0.107	0.146
157	8:00am	0.140	0.138	0.114	0.136
158	10:00am	0.146	0.141	0.112	0.134
159	12:00am	0.145	0.139	0.113	0.136
160	2:00pm	0.157	0.149	0.117	0.137
161	4:00pm	0.159	0.145	0.117	0.137
162	6:00pm	0.137	0.139	0.108	0.138
163	8:00am	0.144	0.135	0.114	0.151
164	10:00am	0.149	0.144	0.122	0.135
165	12:00am	0.149	0.153	0.123	0.164
166	2:00pm	0.149	0.151	0.125	0.130
167	4:00pm	0.148	0.152	0.121	0.120
168	6:00pm	0.150	0.150	0.126	0.140
169	8:00am	0.151	0.148	0.124	0.146
170	10:00am	0.155	0.152	0.128	0.164
171	12:00am	0.122	0.118	0.113	0.130
172	2:00pm	0.142	0.141	0.123	0.132
173	4:00pm	0.124	0.140	0.137	0.076
174	6:00pm	0.143	0.135	0.111	0.099
175	8:00am	0.122	0.115	0.110	0.130
176	10:00am	0.143	0.141	0.125	0.131

177	12:00am	0.151	0.142	0.127	0.136
178	2:00pm	0.152	0.148	0.131	0.134
179	4:00pm	0.157	0.148	0.133	0.146
180	6:00pm	0.155	0.152	0.128	0.165

APPENDIX II

Computation of Hotelling's T^2 statistic for Phase I

Sample	Time	X_1	X_2	X_3	X_4	T^2
1	8:00am	0.125	0.116	0.141	0.144	4.68
2	10:00am	0.124	0.12	0.144	0.138	4.76
3	12:00am	0.128	0.123	0.109	0.12	2.18
4	2:00pm	0.129	0.125	0.109	0.122	2.08
5	4:00pm	0.130	0.123	0.111	0.120	2.00
6	6:00pm	0.130	0.123	0.110	0.120	2.02
7	8:00am	0.133	0.125	0.106	0.122	1.98
8	10:00am	0.135	0.127	0.108	0.124	1.75
9	12:00am	0.146	0.104	0.126	0.140	13.72
10	2:00pm	0.131	0.139	0.128	0.125	3.52
11	4:00pm	0.105	0.118	0.113	0.119	8.47
12	6:00pm	0.124	0.129	0.103	0.140	3.84
13	8:00am	0.132	0.107	0.104	0.128	5.71
14	10:00am	0.127	0.122	0.106	0.140	2.43
15	12:00am	0.144	0.132	0.106	0.138	1.82
16	2:00pm	0.146	0.132	0.111	0.136	1.69
17	4:00pm	0.146	0.128	0.109	0.134	2.58
18	6:00pm	0.152	0.132	0.111	0.136	2.92
19	8:00am	0.161	0.140	0.116	0.137	3.14
20	10:00am	0.158	0.138	0.118	0.137	2.58
21	12:00am	0.161	0.160	0.121	0.138	1.52
22	2:00pm	0.073	0.153	0.123	0.151	3.01
23	4:00pm	0.146	0.132	0.111	0.136	1.69
24	6:00pm	0.144	0.123	0.112	0.164	3.92
25	8:00am	0.101	0.112	0.109	0.130	8.86
26	10:00am	0.102	0.110	0.117	0.120	8.16
27	12:00am	0.150	0.145	0.142	0.140	0.57
28	2:00pm	0.148	0.124	0.142	0.146	4.35
29	4:00pm	0.149	0.117	0.140	0.164	8.27
30	6:00pm	0.147	0.110	0.140	0.130	11.05
31	8:00am	0.155	0.174	0.102	0.132	10.51
32	10:00am	0.150	0.145	0.142	0.140	0.57
33	12:00am	0.166	0.184	0.147	0.099	10.45
34	2:00pm	0.179	0.180	0.093	0.130	11.07
35	4:00pm	0.174	0.173	0.092	0.131	9.26
36	6:00pm	0.165	0.178	0.102	0.136	9.12
37	8:00am	0.165	0.159	0.160	0.134	1.34
38	10:00am	0.173	0.166	0.167	0.146	1.32
39	12:00am	0.181	0.174	0.178	0.155	2.04
40	2:00pm	0.180	0.179	0.171	0.179	1.34
41	4:00pm	0.185	0.183	0.166	0.176	1.18

42	6:00pm	0.183	0.180	0.166	0.179	1.00
43	8:00am	0.165	0.180	0.171	0.179	5.64
44	10:00am	0.161	0.177	0.168	0.190	6.46
45	12:00am	0.164	0.175	0.169	0.199	5.03
46	2:00pm	0.178	0.184	0.172	0.192	2.82
47	4:00pm	0.175	0.173	0.177	0.190	2.15
48	6:00pm	0.176	0.176	0.181	0.203	3.26
49	8:00am	0.163	0.153	0.17	0.205	3.99
50	10:00am	0.160	0.163	0.163	0.205	3.73
51	12:00am	0.159	0.160	0.175	0.202	4.76
52	2:00pm	0.156	0.160	0.175	0.202	5.63
53	4:00pm	0.177	0.163	0.163	0.199	2.07
54	6:00pm	0.175	0.159	0.168	0.199	2.91
55	8:00am	0.180	0.153	0.169	0.191	5.77
56	10:00am	0.213	0.212	0.171	0.191	5.08
57	12:00am	0.210	0.211	0.170	0.218	5.22
58	2:00pm	0.209	0.207	0.162	0.237	6.24
59	4:00pm	0.193	0.188	0.165	0.270	8.96
60	6:00pm	0.185	0.188	0.170	0.196	2.30
61	8:00am	0.202	0.192	0.170	0.195	2.60
62	10:00am	0.190	0.182	0.171	0.191	1.50
63	12:00am	0.187	0.183	0.185	0.220	3.85
64	2:00pm	0.189	0.185	0.180	0.207	2.55
65	4:00pm	0.194	0.189	0.178	0.207	2.47
66	6:00pm	0.205	0.204	0.175	0.215	3.91
67	8:00am	0.206	0.205	0.164	0.203	4.05
68	10:00am	0.179	0.177	0.160	0.201	1.51
69	12:00am	0.196	0.187	0.166	0.211	2.53
70	2:00pm	0.177	0.191	0.165	0.206	5.65
71	4:00pm	0.197	0.189	0.160	0.196	2.17
72	6:00pm	0.188	0.184	0.159	0.202	1.71
73	8:00am	0.190	0.181	0.121	0.218	6.62
74	10:00am	0.172	0.166	0.126	0.237	6.58
75	12:00am	0.182	0.163	0.124	0.279	15.93
76	2:00pm	0.193	0.175	0.128	0.196	5.32
77	4:00pm	0.181	0.165	0.113	0.195	5.53
78	6:00pm	0.189	0.184	0.123	0.191	4.75
79	8:00am	0.177	0.177	0.137	0.22	3.96
80	10:00am	0.196	0.185	0.111	0.207	9.22
81	12:00am	0.197	0.177	0.110	0.207	10.52
82	2:00pm	0.195	0.177	0.125	0.215	7.22
83	4:00pm	0.198	0.181	0.182	0.203	3.73
84	6:00pm	0.191	0.200	0.188	0.201	5.11
85	8:00am	0.176	0.171	0.190	0.211	4.86
86	10:00am	0.178	0.159	0.164	0.206	3.63
87	12:00am	0.176	0.158	0.161	0.196	2.68

88	2:00pm	0.178	0.160	0.163	0.202	3.04
89	4:00pm	0.209	0.193	0.194	0.136	8.28
90	6:00pm	0.207	0.190	0.184	0.134	7.51
91	8:00am	0.197	0.190	0.185	0.136	5.16
92	10:00am	0.183	0.171	0.178	0.137	3.40
93	12:00am	0.176	0.159	0.177	0.137	4.00
94	2:00pm	0.178	0.161	0.168	0.138	2.98
95	4:00pm	0.179	0.161	0.173	0.151	3.04
96	6:00pm	0.161	0.137	0.161	0.135	4.94
97	8:00am	0.188	0.166	0.170	0.164	3.63
98	10:00am	0.178	0.163	0.151	0.130	2.38
99	12:00am	0.179	0.158	0.161	0.120	4.84
100	2:00pm	0.175	0.162	0.165	0.140	1.90
101	4:00pm	0.178	0.163	0.170	0.146	2.36
102	6:00pm	0.184	0.161	0.169	0.164	3.77
103	8:00am	0.191	0.166	0.166	0.13	6.48
104	10:00am	0.152	0.151	0.081	0.132	6.87
105	12:00am	0.153	0.134	0.101	0.076	8.06
106	2:00pm	0.154	0.131	0.103	0.099	6.41
107	4:00pm	0.164	0.153	0.126	0.130	1.35
108	6:00pm	0.163	0.150	0.121	0.131	1.70
109	8:00am	0.176	0.165	0.139	0.136	1.60
110	10:00am	0.155	0.148	0.163	0.134	2.17
111	12:00am	0.147	0.149	0.175	0.146	5.06
112	2:00pm	0.162	0.162	0.184	0.155	4.14
113	4:00pm	0.173	0.173	0.156	0.179	0.79
114	6:00pm	0.173	0.165	0.184	0.176	2.91
115	8:00am	0.156	0.149	0.077	0.179	8.72
116	10:00am	0.137	0.129	0.121	0.179	2.98
117	12:00am	0.137	0.131	0.126	0.190	3.88
118	2:00pm	0.134	0.131	0.124	0.199	5.37
119	4:00pm	0.145	0.144	0.128	0.155	0.84
120	6:00pm	0.148	0.144	0.113	0.179	2.24
121	8:00am	0.152	0.141	0.123	0.176	1.46
122	10:00am	0.143	0.130	0.137	0.179	2.99
123	12:00am	0.144	0.132	0.111	0.179	2.99
124	2:00pm	0.133	0.137	0.110	0.190	5.49
125	4:00pm	0.146	0.148	0.125	0.199	3.87
126	6:00pm	0.142	0.139	0.145	0.162	1.85
127	8:00am	0.144	0.142	0.157	0.173	3.18
128	10:00am	0.189	0.181	0.159	0.165	1.32
129	12:00am	0.183	0.179	0.137	0.149	2.27
130	2:00pm	0.188	0.182	0.144	0.129	3.92
131	4:00pm	0.182	0.176	0.149	0.131	2.44
132	6:00pm	0.125	0.113	0.149	0.131	6.14
133	8:00am	0.112	0.110	0.123	0.144	5.73

134	10:00am	0.193	0.198	0.165	0.133	5.81
135	12:00am	0.206	0.215	0.178	0.167	6.98
136	2:00pm	0.204	0.214	0.170	0.175	6.66
137	4:00pm	0.192	0.198	0.187	0.181	4.11
138	6:00pm	0.142	0.137	0.113	0.147	1.13
139	8:00am	0.189	0.171	0.191	0.187	4.59
140	10:00am	0.152	0.146	0.146	0.102	2.82
141	12:00am	0.177	0.167	0.171	0.118	3.92
142	2:00pm	0.189	0.171	0.191	0.192	4.69
143	4:00pm	0.209	0.211	0.194	0.166	6.16
144	6:00pm	0.184	0.191	0.206	0.206	7.48
145	8:00am	0.191	0.198	0.160	0.128	6.60
146	10:00am	0.159	0.156	0.156	0.064	9.17
147	12:00am	0.154	0.146	0.136	0.063	7.58
148	2:00pm	0.149	0.148	0.141	0.083	4.91
149	4:00pm	0.128	0.123	0.114	0.087	3.78
150	6:00pm	0.137	0.134	0.135	0.087	4.10
151	8:00am	0.142	0.137	0.135	0.092	3.34
152	10:00am	0.137	0.135	0.094	0.145	3.12
153	12:00am	0.142	0.137	0.113	0.147	1.13
154	2:00pm	0.140	0.137	0.111	0.145	1.38
155	4:00pm	0.142	0.140	0.112	0.138	1.31
156	6:00pm	0.138	0.138	0.107	0.146	2.01
157	8:00am	0.140	0.138	0.114	0.136	1.26
158	10:00am	0.146	0.141	0.112	0.134	1.17
159	12:00am	0.145	0.139	0.113	0.136	1.04
160	2:00pm	0.157	0.149	0.117	0.137	1.10
161	4:00pm	0.159	0.145	0.117	0.137	1.61
162	6:00pm	0.137	0.139	0.108	0.138	2.17
163	8:00am	0.144	0.135	0.114	0.151	1.15
164	10:00am	0.149	0.144	0.122	0.135	0.63
165	12:00am	0.149	0.153	0.123	0.164	1.74
166	2:00pm	0.149	0.151	0.125	0.130	1.37
167	4:00pm	0.148	0.152	0.121	0.120	2.33
168	6:00pm	0.150	0.150	0.126	0.140	0.79
169	8:00am	0.151	0.148	0.124	0.146	0.51
170	10:00am	0.155	0.152	0.128	0.164	0.50
171	12:00am	0.122	0.118	0.113	0.130	2.81
172	2:00pm	0.142	0.141	0.123	0.132	1.04
173	4:00pm	0.124	0.140	0.137	0.076	10.68
174	6:00pm	0.143	0.135	0.111	0.099	2.79
175	8:00am	0.122	0.115	0.110	0.130	2.87
176	10:00am	0.143	0.141	0.125	0.131	0.91
177	12:00am	0.151	0.142	0.127	0.136	0.50
178	2:00pm	0.152	0.148	0.131	0.134	0.47
179	4:00pm	0.157	0.148	0.133	0.146	0.20

180	6:00pm	0.155	0.152	0.128	0.165	0.53

APPENDIX III

Hotelling's T^2 results for Phase II Control Chart

Sample No	T^2 value	Sample No	T^2 value	Sample No	T^2 value	Sample No	T^2 value
1	5.38	40	1.21	79	3.32	118	5.08
2	4.94	41	1.18	80	9.85	119	0.46
3	2.30	42	1.09	81	12.06	120	2.00
4	2.07	43	3.39	82	8.55	121	2.09
5	2.37	44	4.04	83	5.37	122	3.92
6	2.38	45	3.29	84	3.73	123	3.60
7	2.43	46	1.76	85	5.21	124	4.26
8	2.21	47	2.11	86	5.27	125	2.96
9	**17.34**	48	2.97	87	4.25	126	1.81
10	2.13	49	4.77	88	4.60	127	3.08
11	6.33	50	2.98	89	10.57	128	2.10
12	2.58	51	4.36	90	9.88	129	2.58
13	7.83	52	4.87	91	6.25	130	4.75
14	2.40	53	3.24	92	4.91	131	3.19
15	2.65	54	4.31	93	5.99	132	7.29
16	2.80	55	8.27	94	4.93	133	5.51
17	4.10	56	5.07	95	5.01	134	5.32
18	4.68	57	4.77	96	7.46	135	5.81
19	5.10	58	6.01	97	5.94	135	5.26
20	4.43	59	8.99	98	4.13	137	3.20
21	1.29	60	1.60	99	7.38	138	1.16
22	4.89	61	3.62	100	3.37	139	6.44
23	2.80	62	2.16	101	4.03	140	3.56
24	5.51	63	4.00	102	6.14	141	5.30
25	6.94	64	2.72	103	9.46	142	6.51
26	6.74	65	2.75	104	6.43	143	6.01
27	0.86	66	3.71	105	10.40	144	6.47
28	6.58	67	3.88	106	8.87	145	5.84
29	11.06	68	1.33	107	2.41	146	10.05
30	**14.51**	69	3.18	108	2.94	147	8.97
31	7.39	70	3.44	109	2.77	148	5.17
32	0.86	71	2.78	110	2.89	149	4.19
33	8.08	72	1.77	111	4.69	150	4.45
34	10.78	73	6.96	112	4.05	151	3.92
35	9.13	74	6.49	113	0.42	152	2.66
36	7.02	75	**17.02**	114	3.67	153	1.16
37	1.97	76	6.80	115	8.62	154	1.16
38	2.06	77	6.64	116	3.25	155	1.04
39	2.81	78	4.83	117	3.94	156	1.27

157	1.00	163	1.62	169	0.41	175	3.15
158	1.30	164	0.82	170	0.34	176	0.77
159	1.27	165	0.68	171	2.77	177	1.19
160	1.65	166	0.76	172	0.75	178	0.63
161	2.84	167	1.52	173	8.48	179	0.88
162	1.36	168	0.38	174	3.55	180	0.36

APPENDIX IV

MYT Decomposition Components for the Observation Vector in Sample 9

T^2 decomposition	UCL	P-value	1	2	3	4
0.3267	6.8163	0.5683	1	0	0	0
4.8246	6.8163	0.0293	2	0	0	0
0.2390	6.8163	0.6255	3	0	0	0
0.2857	6.8163	0.5936	4	0	0	0
17.1312	9.5587	0.0000	1	2	0	0
0.3495	9.5587	0.7055	1	3	0	0
0.4001	9.5587	0.6709	1	4	0	0
6.2863	9.5587	0.0023	2	3	0	0
5.3035	9.5587	0.0058	2	4	0	0
0.3779	9.5587	0.6858	3	4	0	0
17.3378	11.8802	0.0000	1	2	3	0
17.1323	11.8802	0.0000	1	2	4	0
0.4180	11.8802	0.7403	1	3	4	0
6.6246	11.8802	0.0003	2	3	4	0
17.3379	14.0213	0.0000	1	2	3	4

APPENDIX V

MYT Decomposition Components for the Observation Vector in Sample 30

T^2 decomposition	UCL	P-value	1	2	3	4
0.2817	6.8163	0.5963	1	0	0	0
3.8462	6.8163	0.0514	2	0	0	0
0.0000	6.8163	1.0000	3	0	0	0
0.6429	6.8163	0.4237	4	0	0	0
13.3464	9.5587	0.0000	1	2	0	0
0.4942	9.5587	0.6109	1	3	0	0
0.6574	9.5587	0.5194	1	4	0	0
6.5674	9.5587	0.0018	2	3	0	0
3.9036	9.5587	0.0219	2	4	0	0
0.7596	9.5587	0.4694	3	4	0	0
14.3495	11.8802	0.0000	1	2	3	0
13.4655	11.8802	0.0000	1	2	4	0
0.9093	11.8802	0.4377	1	3	4	0
6.5758	11.8802	0.0003	2	3	4	0
14.5099	14.0213	0.0000	1	2	3	4

APPENDIX VI

MYT Decomposition Components for Observation Vector in Sample 75

T^2 decomposition	UCL	P-value	1	2	3	4
0.8067	6.8163	0.3703	1	0	0	0
0.0138	6.8163	0.9065	2	0	0	0
0.3122	6.8163	0.5770	3	0	0	0
10.1150	6.8163	0.0017	4	0	0	0
3.7598	9.5587	0.0251	1	2	0	0
3.1181	9.5587	0.0467	1	3	0	0
11.0161	9.5587	0.0000	1	4	0	0
0.7014	9.5587	0.4973	2	3	0	0
13.2369	9.5587	0.0000	2	4	0	0
13.9664	9.5587	0.0000	3	4	0	0
5.4042	11.8802	0.0014	1	2	3	0
14.8638	11.8802	0.0000	1	2	4	0
14.0132	11.8802	0.0000	1	3	4	0
14.6149	11.8802	0.0000	2	3	4	0
17.0154	14.0213	0.0000	1	2	3	4