UNIVERSITI PUTRA MALAYSIA

ROBUST CONTROL CHARTS FOR CHANGE POINTS DETECTION IN PRESENCE OF OUTLIERS

NG KOOI HUAT

IPM 2012 2

ROBUST CONTROL CHARTS FOR CHANGE POINTS DETECTION IN PRESENCE OF OUTLIERS

By

NG KOOI HUAT

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

February 2012

© COPYRIGHT UPM

DEDICATIONS

- To my family for having unconditional love for me.
- To my beloved supervisors, lecturers, teachers and friends who uplifted my life.

© COPYRIGHT UPM

Abstract of Thesis Presented to the Senate of Universiti Putra Malaysia in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

ROBUST CONTROL CHARTS FOR CHANGE POINTS DETECTION IN PRESENCE OF OUTLIERS

By

NG KOOI HUAT

February 2012

Chairman: Habshah Midi, PhD

Faculty: Institute for Mathematical Research

Control charts are used to detect whether or not a process has changed. When a control chart signals indicating that a process has changed, practitioners must initiate a search for the special cause. However, given a signal from a control chart, practitioners generally do not know what caused the process situation to change or when the process has changed. Identifying the time of the process change would simplify the seeking of the special cause. It is now evident that outliers have great impact on the parameters estimation in the setting of a control chart. The violation of assumption from normality for change point hypothesis testing method can also gravely mislead the inferential statistics. Hence, the main focus of this research is to take remedial measures for these issues on the occasion that there is a violation of normality assumption and in the presence of contamination. We have presented

© COPYRIGHT UPM

a robust individuals control chart in the context of exploratory analysis for the purpose of locating the step change position. This type of chart offers some significant advantages over the existing individuals control chart. It is about adopting the M-Scale estimator into the proposed modified procedure in the estimation of process standard deviation. The results signify that the proposed method offers substantial improvements over the existing method. On the same ground, to further enhance hypothesis testing approach in the presence of outlier for the change point statistics, the Huber Maximum-Type testing method is incorporated into the proposed modified framework. The findings indicate that the proposed approach is more efficient in detecting the correct step change position, both in normal shift and the shift in the existence of disturbances.

We also proposed a robust MM control chart for monitoring the change in process mean when there is a contamination in data collection. The newly proposed control chart is formulated through the use of S-scale estimate, which in turn yields the MM-location estimate, possessing 50% high breakdown point and 95% efficiency when the errors are under normality (Salibian-Barrera, 2004). From the results, it appears to suggest that the proposed robust MM control chart is more reliable and performs superbly in the presence of outliers.

Finally, the new robust subsample-based Modified Biweight A Scale (MBAS) chart which is resistant to outliers is proposed. A novel scale measure, namely the Modified Biweight A (MBAS) scale estimator is incorporated which provides a choice for practitioners who are interested in the detection of permanent shifts in

© COPYRIGHT UPM

process variance. It is evident that the proposed chart outperforms the conventional charts when contaminated data are present. In summary, the proposed robust control-charting methodologies appear to efficiently monitor contaminated data situations and process shift, while the classical charts are not a preference for process monitoring where contamination may exist. In this thesis, all the proposed procedures were examined by real data sets and Monte Carlo simulation studies. Comparative studies among the classical and the proposed robust methods reveal that the proposed robust methods are able to rectify the issues in relation to the presence of outliers. On the contrary, the classical approaches seem to perform poorly in these circumstances.

© COPYRIGHT UPM

Abstrak Tesis yang Dikemukakan kepada Senat Universiti Putra Malaysia
sebagai Memenuhi Keperluan untuk Ijazah Doktor Falsafah

CARTA KAWALAN TEGUH DALAM PENGENALPASTIAN PERUBAHAN KEDUDUKAN TITIK DENGAN KEHADIRAN TITIK TERPENCIL

Oleh

NG KOOI HUAT

Februari 2012

Pengerusi: Habshah Midi, PhD

Fakulti: Institut Penyelidikan Matematik

Carta kawalan digunakan untuk pengenalpastian samaada suatu proses telah berubah. Apabila carta kawalan menunjukkan isyarat di mana suatu proses telah berubah, pengamal perlu memulakan pemeriksaaan untuk mencari sebab khas. Walau bagaimanapun, apabila carta kawalan memberikan isyarat, kebiasaannya pengamal tidak mengerti sebab-sebab yang mencetuskan situasi tersebut dan masa perubahan proses. Penentuan masa perubahan proses akan memudahkan pencarian sebab khas. Kini, jelasnya titik terpencil memberikan kesan buruk ke atas penganggaran parameter dalam pembentukkan carta kawalan. Penyimpangan daripada andaian kenormalan juga mempunyai kesan serius terhadap statistik pentakbiran. Oleh yang demikian, tumpuan utama kajian ini adalah untuk mengambil langkah pemulihan terhadap isu-isu yang melibatkan penyimpangan

© COPYRIGHT UPM

daripada andaian kenormalan dan dalam situasi kewujudan titik terpencil. Kami telah memperkenalkan carta kawalan teguh individu dalam kontek analisis penjelajahan dengan tujuan untuk menjejaki kedudukan perubahan titik. Carta kawalan yang dikemukakan mempunyai nilai yang tinggi berbanding dengan carta kawalan individu yang sedia ada. Dalam cadangan pindaan tatacara ini, penganggar M-Skala digabungkan dalam operasi penganggaran sisihan piawai proses. Keputusan menunjukkan bahawa pindaan tatacara yang dicadangkan menawarkan kemajuan besar berbanding dengan kaedah yang sedia ada. Dengan asas yang sama, demi mempertingkatkan tatacara ujian hipotesis terhadap perubahan statistik titik dengan kehadiaran titik terpencil, ujian hipotesis Jenis-Maksimum Huber digabungkan dalam kerangka cadangan terubahsuai. Hasil penyiasatan menunjukkan bahawa pendekatan yang dicadangkan lebih cekap dalam usaha penentuan titik perubahan dengan tepat, samada dalam perubahan titik yang biasa ataupun perubahan titik dengan kehadiran gangguan.

Kami juga mengemukakan carta kawalan MM yang baharu bagi memantau penukaran dalam proses purata dengan kewujudan pencemaran data. Cadangan carta kawalan yang baharu ini telah menggunakan penganggar jenis S-skala di mana penganggar titik MM yang dikemukakan mempunyai titik musnah yang tinggi sehingga mencapai 50 peratus dan 95 peratus kecekapan apabila ralatnya berada dalam situasi normal. Keputusan menunjukkan bahawa carta kawalan MM lebih diyakini dan berprestassi tinggi dalam keadaan kehadiran titik terpencil.

© COPYRIGHT UPM

Akhir sekali, cadangan carta kawalan (MBAS) baharu yang berasaskan data subsampel diperkenalkan. Carta kawalan ini teguh terhadap titik terpencil. Suatu penganggar skala yang baru iaitu penganggar MBAS digabungkan bertujuan untuk menawarkan pilihan lain kepada pengamal proses yang berminat dalam perubahan tetap dalam proses varian. Nyatanya, carta kawalan cadangan baru ini lebih efisien berbanding carta kawalan piawai terutamanya dalam kehadiran titik terpencil. Sebagai kesimpulan, pembentukkan carta kawalan yang teguh sememangnya lebih cekap dalam pemantauan proses yang tercemar dan dalam perubahan proses. Pada masa yang sama, carta kawalan tradisional bukan pilihan tepat untuk pemerhatian proses yang mungkin tercemar. Dalam tesis ini, setiap tatacara telah diperiksa dengan data set yang sebenar dan juga kajian simulasi Monte Carlo. Perbandingan antara kaedah klasik dan kaedah teguh telah mendedahkan bahawa semua kaedah teguh yang dicadangkan berjaya memperbaiki masalah dalam kehadiran titik terpencil. Sebaliknya, kaedah klasik mempunyai prestasi yang rendah dalam situasi-situasi tersebut.

© COPYRIGHT UPM

ACKNOWLEDGEMENTS

I do not have much word to express how I am grateful to my supervisor, Prof. Dr. Habshah Midi, who has the merit of teaching me how to do research and who has transmitted to me her enthusiasm through her activities. Her willingness to share her ideas in research problems, and the energy and time she put in advising my thesis work is highly obliged. I am benefited enormously from her continuous support and confidence throughout my research. Without her help and support, this dissertation would have been impossible. Moreover, her advices and her remarks have proven to be very useful and simulating. I also greatly value her friendship, kindness, and elegant personality. I feel truly privileged to have been her student.

I acknowledge my internal co-supervisors Dr. Jayanthi Arasan and Dr. Bashar Abdul Aziz Majeed Al-Talib, senior lecturers of my institute, for their help. I am indebted to my external co-supervisor Dr. A.H.M. Rahmatullah Imon, Professor of Statistics, Department of Mathematical Sciences, Ball State University, U.S.A., who has helped me a lot by responding my constant volley of electronic messages regarding my research problem. I am truly grateful that I have such a great mentor.

Special thanks to Dr. Sarker, S.K., fellow researcher of my institute, for his important suggestions and cooperation in my research work. His valuable words always inspired me so much. I am very grateful to Dr. Mohd Bakri Adam, Dr. Kuddus and Dr. Hossein for helping me in R and S-Plus programming.

© COPYRIGHT UPM

I gratefully acknowledge moral supports of my friends and their continuous encouragement. I remember my friends Arezoo, Saniza, Hossein, Kourash, Askan, Vello, Sohel, NKH, YHK, WTZ, LFP, Lily Wong, KLF, YLK and SHS. Thank you to all my friends for all of your generosity and kindness.

Finally, I would like to thank Universiti Putra Malaysia for the financial support. They have created an excellent environment for my research here. My sincere thanks are extended to all the staff of the Institute for Mathematical Research (INSPEM), UPM, for their cordial assistance during this research work.

© COPYRIGHT UPM

I certify that a thesis Examination Committee has met on 29 February 2012 to conduct the final examination of Ng Kooi Huat on his thesis entitled "Robust Control Charts for Change Points Detection in Presence of Outliers" in accordance with Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Examination Committee are as follows:

Mohd Rizam Bin Abu Bakar, PhD
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Chairman)

Mohd Bakri Bin Adam, PhD
Senior Lecturer
Faculty of science
Universiti Putra Malaysia
(Internal Examiner)

Abdul Ghapor Bin Hussin, PhD
Associate Professor
Centre for Foundation Studies
Universiti Malaya
(External Examiner)

Muhammad Hanif Mian, PhD
Professor
Lahore University of Management Sciences
Pakistan
(External Examiner)

SEOW HENG FONG, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 23 April 2012

© COPYRIGHT UPM

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfillment of the requirements for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Habshah Binti Midi, PhD
Professor
Faculty of Science
Universiti Putra Malaysia
(Chairman)

Jayanthi Arasan, PhD
Senior Lecturer
Faculty of science
Universiti Putra Malaysia
(Member)

A. H. M. Rahmatullah Imon, PhD
Professor
Ball State University
Muncie, IN 47306, U.S.A.
(Member)

Bashar Abdul Aziz Majeed Al-Talib, PhD
Senior Lecturer
Faculty of science
Universiti Putra Malaysia
(Member)

BUJANG BIN KIM HUAT, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

© COPYRIGHT UPM

DECLARATION

I declare that the thesis is my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously, and is not concurrently, submitted for any other degree at Universiti Putra Malaysia or at any other institutions.

NG KOOI HUAT

Date: 29 February 2012

© COPYRIGHT UPM

TABLE OF CONTENTS

		Page
DEDICATIONS		ii
ABSTRACT		iii
ABSTRAK		vi
ACKNOWLEDGEMENTS		ix
APPROVAL		xi
DECLARATION		xiii
LIST OF TABLES		xviii
LIST OF FIGURES		xxiii
LIST OF APPENDICES		xxvi
LIST OF ABBREVIATIONS		xxvii
CHAPTER		
1	**INTRODUCTION**	
	1.1 Background of the Study	1
	1.2 Importance and Motivation of the Study	3
	1.3 Objectives of Study	10
	1.4 Plan of the Study	11
	1.5 Benefits and Contributions	15
2	**LITERATURE REVIEW**	
	2.1 Introduction	17
	2.2 Change Point Analysis	17
	2.2.1 Test Statistics	22
	2.2.2 Critical Values	25
	2.3 Statistical Process Control	27
	2.3.1 Control Chart	28
	2.3.2 Control Limits	30
	2.3.3 Individuals Chart and Moving Range Chart	31
	2.3.4 Assessing the Performance of Control Charts (Average Run Length)	34
	2.4 Robust Statistics	36
	2.4.1 Outliers	40
	2.4.2 Resistance and Robustness of Efficiency	42
	2.4.3 M-Measures of Location	42
	2.4.4 Breakdown Bound	45
	2.4.5 Sample Median	45
	2.4.6 Measures of Scale	46
	2.4.7 Median Absolute Deviation	48
	2.4.8 q - Quantile Range	50

© COPYRIGHT UPM

2.4.9 Numerical Computation of M-Estimates 51
2.4.10 Location with Previously Computed Dispersion Estimation 52
2.4.11 Scales Estimates 53
2.4.12 Simultaneous Estimation of Location and Dispersion 54

3 **PERMANENT SHIFT IN PROCESS MEAN DETECTION WITH ROBUST INDIVIDUALS CONTROL CHART IN EXPLORATORY ANALYSIS**
3.1 Introduction 56
3.2 M-Scale Estimate 58
3.3 De Mast and Roes Control-Charting Procedure 61
3.4 The Proposed Control-Charting Procedure 68
3.5 Illustrative Example: Application to Wood Moisture Content Data 70
3.6 Simulation Study, Numerical Results and Discussion 75
3.7 Conclusion 80

4 **ROBUST INDIVIDUALS CONTROL CHART BASED ON HUBER MAXIMUM-TYPE TESTING FORMULATION FOR CONTAMINATED DATA**
4.1 Introduction 82
4.2 Maximum-Type Test Statistics Estimator 84
4.2.1 Maximum-Type Critical Values 86
4.3 Huber Maximum-Type M-Test 87
4.4 The Proposed Control-Charting Procedure 90
4.5 Numerical Example: Modified Hybrid Microcircuits Data 92
4.6 Simulation Study, Numerical Results and Discussion 96
4.7 Conclusion 112

5 **CHANGE POINT DETECTION WITH ROBUST MM CONTROL CHART IN THE PRESENCE OF OUTLIERS**
5.1 Introduction 114
5.2 MM-Location Estimates 117
5.3 Control Limits 119
5.4 Process Step Change Model 122

© COPYRIGHT UPM

5.5 Control-Charting Procedure 123
5.6 Illustrative Example: Vane-Opening Measurements Data 124
5.7 Simulation Study, Numerical Results and Discussion 129
5.8 Conclusion 149

6 **ROBUST MODIFIED BIWEIGHT A CONTROL CHART FOR CHANGE IN PROCESS VARIANCE IN THE PRESENCE OF DISTURBANCES**
6.1 Introduction 151
6.2 Modified Biweight A Pooled-Based Scale Estimator 153
6.3 Modified Biweight A Subsample-Based Scale Estimator 157
6.4 Control Limits 158
6.5 Change Point Estimator 160
6.6 Process Step Change Model 161
6.7 Numerical Example: Modified Server-to-Car Times Data (In Minutes) 162
6.8 Simulation Study 168
6.9 Numerical Results and Discussion 170
6.9.1 Average Change Point and Expected Length 170
6.9.2 Root Mean Squared Deviation 178
6.10 Conclusion 184

7 **SUMMARY, CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER STUDIES**
7.1 Introduction 187
7.2 Summary 187
7.2.1 Permanent Shift in the Process Mean Detection with Robust Individuals Control Chart in Exploratory Analysis 188
7.2.2 Robust Individuals Control Chart Based on Huber Maximum-Type Formulation for Contaminated Data 189
7.2.3 Change Point Detection with Robust MM Control Chart in the Presence of Outliers 190

© COPYRIGHT UPM

7.2.4 Robust Modified Biweight A Control Chart for Change in Process Variance in the Presence of Disturbances 191
7.3 Conclusion 192
7.4 Areas of Further Research 193

REFERENCES 196
APPENDICES 204
BIODATA OF STUDENT 232
LIST OF PUBLICATIONS 233
LIST OF PRESENTATIONS 234
AWARDS 235

© COPYRIGHT UPM