# STATISTICAL PROCESS CONTROL USING MODIFIED ROBUST HOTELLING'S $T^2$ CONTROL CHARTS

**FIRAS SALEEM FARES HADDAD**

**DOCTOR OF PHILOSOPHY**
**UNIVERSITI UTARA MALAYSIA**
**2013**

## Permission to Use

In presenting this thesis in fulfilment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the Universiti Library may make it freely available for inspection. I further agree that permission for the copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor(s) or, in their absence, by the Dean of Awang Had Salleh Graduate School of Arts and Sciences. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part, should be addressed to :

Dean of Awang Had Salleh Graduate School of Arts and Sciences
UUM College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok

# Abstrak

Carta Hotelling's $T^2$ adalah alat yang popular bagi memantau kawalan proses berstatistik. Walau bagaimanapun, carta ini sensitif pada titik terpencil. Bagi mengatasi masalah ini, tiga pendekatan terhadap carta Hotelling's $T^2$ teguh telah dicadangkan iaitu pendekatan pemangkasan, peWinsoran dan berasaskan median. Kesemua pendekatan ini menggunakan penganggar lokasi teguh dan penganggar skala teguh yang masing-masing menggantikan min biasa dan matriks kovarians. Bagi setiap pendekatan, tiga penganggar skala teguh: $MAD_n$, $S_n$ dan $T_n$ diperkenalkan, dan penganggar ini berfungsi sewajarnya mengikut pendekatan. Pendekatan pertama, ditandai sebagai $T_t^2$, menggunakan konsep pemangkasan melalui jarak Mahalanobis. Penganggar skala teguh digunakan untuk mengganti matriks kovarians dalam jarak Mahalanobis. Min terpangkas dan matriks kovarians terpangkas merupakan penganggar lokasi dan skala bagi carta $T_t^2$. Pendekatan kedua, $T_w^2$, menggunakan setiap penganggar skala sebagai kriteria Winsor. Pendekatan ini mengaplikasikan penganggar M-satu langkah terubahsuai terWinsor dan kovarians terWinsor yang sepadan, masing-masing sebagai penganggar lokasi dan matrik skala bagi carta $T_w^2$. Manakala dalam pendekatan ketiga, $T_H^2$, penganggar skala teguh berperanan sebagai matriks skala dengan Hodges-Lehman sebagai penganggar lokasi. Pendekatan ini menggunakan data asal tanpa sebarang pemangkasan atau peWinsoran. Secara keseluruhannya, sembilan carta kawalan teguh telah dicadangkan. Prestasi setiap carta kawalan teguh dinilai berdasarkan kadar penggera palsu dan kebarangkalian mengesan. Bagi mengkaji kekuatan dan kelemahan carta yang dicadangkan, pelbagai keadaan diwujudkan dengan memanipulasi empat pembolehubah iaitu bilangan ciri-ciri kualiti, kadar data terpencil, tahap anjakan min dan sifat ciri-ciri kualiti (bebas dan bersandar). Secara umumnya, carta yang dicadangkan menunjukkan prestasi yang baik dari segi kadar penggera palsu. Dari sudut kebarangkalian mengesan, prestasi kesemua carta yang dicadangkan mengatasi carta Hotelling's $T^2$ tradisional. Keseluruhannya, kajian mendapati carta Hotelling's $T^2$ teguh yang dicadangkan boleh dijadikan alternatif yang baik kepada carta tradisional yang dipertikaikan.

Katakunci: Hotelling's $T^2$, Carta kawalan, Penganggar teguh

## Abstract

Hotelling's $T^2$ chart is a popular tool for monitoring statistical process control. However, this chart is sensitive to outliers. To alleviate the problem, three approaches to the robust Hotelling's $T^2$ chart namely trimming, Winsorizing and median based were proposed. These approaches used robust location and scale estimators to substitute for the usual mean and covariance matrix, respectively. For each approach, three robust scale estimators: $MAD_n$, $S_n$ and $T_n$ were introduced, and these estimators functioned accordingly to the approach. The first approach, denoted as $T_t^2$, applied the concept of trimming via Mahalanobis distance. The robust scale estimator was used to replace the covariance matrix in Mahalanobis distance. The trimmed mean and trimmed covariance matrix were the location and scale estimators for the $T_t^2$ chart. The second approach, $T_w^2$, employed each scale estimator as the Winsorized criterion. This approach applied Winsorized modified one step M-estimator and its corresponding Winsorized covariance as the location and the scale matrix for $T_w^2$ chart, respectively. Meanwhile, in the third approach, $T_H^2$, the robust scale estimator took the role of the scale matrix with Hodges-Lehman as the location estimator. This approach worked with the original data without any trimming or Winsorizing. Altogether, nine robust control charts were proposed. The performance of each robust control chart was assessed based on false alarm rates and probability of detection. To investigate on the strengths and weaknesses of the proposed charts, various conditions were created by manipulating four variables, namely number of quality characteristics, proportion of outliers, degree of mean shifts, and nature of quality characteristics (independent and dependent). In general, the proposed charts performed well in terms of false alarm rates. With respect to probability of detection, all the proposed charts outperformed the traditional Hotelling's $T^2$ charts. The overall findings showed that, the proposed robust Hotelling's $T^2$ control charts are viable alternatives to the disputed traditional charts.

Keywords: Hotelling $T^2$, Control chart, Robust estimator

## Acknowledgement

I am grateful to the Almighty Allah for giving me the opportunity to complete my PhD thesis.

In completing this thesis, I owe a debt of gratitude and thanks to many persons and institutions that have supported me throughout this difficult yet challenging journey. While being thankful to all of them, I must register my gratitude to some in particular. Primarily, I would like to express my deepest appreciation to my supervisor Associate Professor Dr. Sharipah Soaad Syed Yahaya who has been very patient in guiding me and supporting from the first day arrival here in Malaysia and throughout this thesis. She assisted me immensely in focusing my thinking and ideas towards the right direction and gave me her valuable ideas, insights, comments and suggestions towards understanding the empirical predicaments I have encountered. I would like to also thank my co-supervisor Dr. Nor Idayu who supported me and help me in all stages of the writing of the thesis. To all academic and administrative staff in College of Art and Science, my sincere gratitude goes to you.

I would like to express my never ending appreciation and gratitude to Prof. Jose Luis Alfaro and the people in Jordan. First and foremost, I would like to thank my father who had been a great and wise teacher in my life, my lovely mother for her infinite patience especially during my absence, and her sincere flow of love has accompanied me all the way in my long struggle and has pushed me to pursue my dreams. My lovely wife Maiss for her love and supported me, infinite patience especially during my absence and for her pushed me to pursue my dream. Big thanks and appreciate for my two sons Yazeed and Saleem and my brothers and sisters for their patients until finish this journey of the studying.

I would like to thank all my friends, Dr. Mustafa Abu- Shaweish, Dr. Ossama Badawi, Dr. Moatasem Smadi, Dr. Raed Khasawnweh, Hussam Haddad, Dr. Malek Kasasbeh, Mr. Mohammed Kasasbeh, Dr. Hamzeh Smadi, Dr. Belal Al-Wadi, Dr. Abed Alftah Al-Azam, Dr. Haidar Al-Dreybe, Dr. Mahmoud Al-Eqab, Dr. Tareg

Abusaa, Mr. Ramzi Al-Tarazi, Dr. Aymen Abu Alhija, Dr. Salem Al-Shra'a, Dr. Abdallah Alshamari, Dr. Ali Naimat, Dr. Abed Al-hameed Al-Eneze, Dr. Obaideh Alhazimeh, Mr Amer Alhazimeh Dr. Alla'a Alsiad, Dr. Eid Hassan, Dr. Adnan Almulhem, Dr. Saleh Al-Rasheed, Mr. Basem Ayoub, Dr. Younis Megdad, Dr. Amer Abu- Rashed, Dr. Mahmoud Megdadi, Dr. Hatim Megdadi. To all of you, I have this to say: I love you, respect you, pray for you, and may Allah bless you.

# Table of Contents

# List of Tables

# List of Figures

# List of Appendices

# List of Abbreviations

| | |
|---|---|
| ***MOM*** | Modified One-step *M*-estimator |
| ***HL*** | Hodges and Lehmann estimator |
| ***Med*** | Median |
| $\boldsymbol{MAD_n}$ | Median absolute deviation |
| $\boldsymbol{S_n}$ | A scale estimator |
| $\boldsymbol{T_n}$ | A scale estimator |
| **FA** | False Alarms |
| **POD** | Probability of Detection |
| **ARE** | *A*symptotic Relative Efficiency |
| ***MD*** | Mahalonobis Distance |

# CHAPTER ONE
# MULTIVARIATE QUALITY CONTROL CHARTS

## 1.1 Introduction

The invention of Statistical Process Control (SPC) chart was pioneered by Dr. Walter Shewhart while he was working for Bell Labs in 1920. He aimed to monitor the quality of a process mathematically. Since then, this tool has received tremendous attention and interest from many researchers and practitioners from various fields including statistics, engineering and education to name just a few. There are some definitions of SPC charts tool. We refer to Montgomery (2005), who defined the SPC charts as tool for optimizing the amount of information needed for decision-making purposes. In addition, Nedumaran and Pignatiello (2000) defined the charts as tools to monitor performance or state of the process.

In general, SPC charts are graphical presentations that display the stability of a process. Unlike other common charts, such as bar chart, line chart or pie charts, SPC charts have some main features such as the following:

(i) The upper limit and lower limit's lines that create a range to where a process output is considered "in control"

(ii) A center line which located in the middle of the lower and upper limits that reflects the average state of the process.

The contents of the thesis is for internal user only

# References


Abu-Shawiesh, M. O. & Abdullah, M. (2001). A new Robust Bivariate Control Chart For Location. *Communication in statistics, 30*, 513 – 529.

Alfaro, J. L. & Ortega, J. F. (2008). A Robust Alternative to Hotelling's $T^2$ Control Chart Using Trimmed Estimators. *Qual. Reliab. Eng. Int., 24,* 601-611.

Alfaro, J. L.& Ortega, J. F. (2009). A comparison of robust alternative to Hotelling $T^2$ control chart. *Applied Statistics, 36*(12), 1385-1396.

Alloway, J. A. J. & Raghavachari, M. (1990). Multivariate Control Charts Based on Trimmed Mean. *ASQC Quality Congress Transactions – San Francisco,* 449-453.

Alloway, J. A. & Raghavachari, M. (1991). An introduction to Multivariate Control Charts. *ASQC Quality Congress Transactions – Milwaukee,* 773-781.

Alt, F. B. (1985). Multivariate quality control. *New York, John Wiley & Sons*.

Ansell, J. G., & Margaret. (2009). Robustness of Location Estimators to Asymmetry. *Journal of the Royal Statistical Society. Series C ( Applied statistics),* 249-254.

Ben-Gal, I. (2005). Outliers detection. In: Maimon O. and Rockach L. (Eds.) Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers, *Kluwer Academic Publishers, ISBN* 0-387-24435-2.

Bersimis, S., Psarakis, S. & Panaretos, J. (2006). Multivariate Statistical Process Control Charts: An Overview. *Qual. Reliab. Engng. Int., 23, 517–543.*

Besterfield, D. H. (Ed.) (2004). *Quality Control* (7 Ed.): Pearson Education, Inc., Upper Saddle River, New Jersey 07458.

Bickel, P. J. (1965). On some robust estimates of location. *The Annals of Mathematical Statistics, 36,* 847-858.

Bluman, & Allan, G. (2009). *Elementary Statistics: A step by step approach* (7 ed.).

Bradley, J. V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology. 31*, 144-152.

Brooks, A. (1985). Heat treating shows why SPC is no curve-all for manufacturers production Engineering. *32*(9), 66-77.

Brown, B. M. & Kildea, D. G. (1978). Reduced u-statistics and the Hodges-Lehmann estimator. *The Annals of Statistics, 6, (*4), 828-835.

Cheng, L. M., Away, Y. & Hasan, M. K. (2006). The Algorithm and Design of Real-time Multivariate Statistical Process Control System. *Multimedia Cyberspace*, *4*(2), 18-23.

Chenouri, S., Variyath, A., M. & Steiner, S. H. (2009). A Multivariate Robust Control Chart for Individual Observations. *Quality Technology*. *41*(3), 259-271.

Chernobai, A. & Rachev, S. T. (2006). Applying robust methods to operational risk modeling *Operational Risk, 1*(1), 27-41.

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

Everitt, B. S. (1979). A Monte Carlo investigation of robustness of Hotelling' one and two sample T- square test. *American statistical Association*, *74*(365), 48-51.

Geyer, C., J. (2006). Breakdown Point Theory Notes. Retrieved from http://www.stat.umn.edu/geyer/s06/5601/theo/break.pdf

Hampel, F. R. (1974). The influence curve and its role in robust estimation. *J. Amer. Statist. Assoc., 69*, 383-393.

Hampel, F. R. (2000). Robust inference. *Seminar for Statistic. Research Report No. 93.* Switzerland.

Hellerstein, J. M. (2008). Quantitative Data of Cleaning for Large Data basis. *EECS Computer Science Division. UC Berkeley*

Hodges, J. L.& Lehmann, E. L. (1963). Estimates of location based on rank tests. *Annals of Mathematical Statistics. 34*, 598-611.

Hodges, R. V. (1967). Some Observations on Robust Estimation. *American Statistical Association, 62(320)*, 1179-1186.

Hogg, R. V. (1974). Adaptive robust procedures: a partial review and some suggestions for future applications and theory. *American Statistical Association. 69(348),* 909-923.

Holmes, D. S. & Mergen, A. E. (1996). Identifying the sources for out of control signals when the $T^2$ control chart is used. *Quality Engineering, 8*(1), 137-143.

Hotelling, H. & McGraw-Hill (1947). Multivariate Quality Control, Illustrated by the Air Testing of Sampling Bombsights. *Techniques of Statistical Analysis, Eisenhart, C., Hastay, M., Wallis, W. A. (eds.),* 111-184.

Jackson, J. E. (1985). Multivariate Quality Control. *Communication in statistics: Theory and methods*, *14*(110), 2657-2688.

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

Jensen, W. A., Birch, J. B. & Woodall, W. H. (2007). High Breakdown Estimation Methods for Phase I Multivariate Control Charts. *Quality Reliability Engineering International*, 23, 615-629.

Johnson, M. E. (Ed.) (1987). *Multivariate statistical simulation*: New York, John Wiley and sons.

Johnson, N. (2007). *A comparative simulation study of robust estimators of standard error.*, Brigham Young University

Linden, A., Adams, J. L. & Robert, N. (2004). Using an Empirical Method for Establishing Clinical Outcome Targents in Disease Management Programs. *Disease Management. 7(2),93-101.*

Majid, A., Haron, K. & Midi, H. (2010). Hodges Lehmann Estimator for Robust Design Using Dual Response Surface Approach. *Australian Journal of Basic and Applied Science. 4(8), 3418 – 3423, ISSN 1991 – 8178. INSI net publication*

Mahir A., R., & Al-Khazaleh, A. M. (2009). New Method to Estimate Missing Data by Using the Asymmetrical Winsorized Mean in a Time Series. *Applied Mathematical Sciences, 3*(35), 1715 - 1726.

Mason, R. & Young, J. (1999). Improving the sensitivity of the $T^2$ statistic in multivariate process control. *Quality Technology*, *31*, 155-165.

Mason, R. L., Chou, W., Y. & Young, J. C. (2001). Applying Hotelling's $T^2$ statistic to Batch Processes. *Quality Technology, 33(4), 466.*

Mason, R. L., Chou, Y., Sullivan, J. H., Stoumbos, Z. G. & Young, J. C. (2003). Systematic patters in $T^2$ charts. *Quality Technology, 35*, 47-58.

Mason, R. L. & Young, J. C. (2002). Multivariate Statistical Process Control with Industrial Applications. *Vol. 9* of *American Statistical Association-Society for Industrial and Applied Mathematics (ASA-SIAM).*

Math works Matlab Version 7.8 (R2009a).

Mehrotra, K., Jackson, P. & Schick, A. (1991). On choosing an optimally trimmed mean. *Communication in Statistics-Simulation and Computation, 20*(1), 73-80.

Midi H., Shabbak, A., Talib, B. A. & Hassan, M. N. (2009). *Multivariate control charts based on robust Mahalonobis distance.* Proceedings of the 5th Asian Mathematical Conference, Malaysia

Mingxin W. (2006). *Abstract trimmed and winsorized estimator.* PhD. Michigan State University.

Mohammadi M., Midi H., Arasan J. & Al-Talib B. (2011). High Breakdown Estimators to Robustify Phase II Control Charts. *Applied Sciences, 11 (3),* 503-511

Montgomery, D. C. (Ed.) (2005). *Introduction to Statistical Quality Control* (5 ed.): John Wiley and Sons, New York, NY, USA.

Nedumaran, G.& Pignatiello, J. J. (2000). On constructing $T^2$ control charts for retrospective examination. *Communication in Statistics-Simulation and Computation*, *29*, 621-632.

Oktay A. & Aricigil A. (2001). Multivariate quality control: A historical perspective. *Statistics statistical techniques*, 54-65.

Othman, A. R., Keselman, H. J., Padmanabhan, A. R., Wilcox, R. R. & Fradette, K. (2004). Comparing measures of the 'typical' score across treatment groups. *British Journal of Mathematical and Statistical Psychology*. *57,* 215 - 234.

Pei-Chen, W. (2002). *The central limit theorem and comparing means, trimmed means, one step M-estimators and modified one step M-estimators under non normality.* Southern of California, Los Angeles, California.

Peña& Prieto, F. J. (2001). Multivariate Outlier Detection and Robust Covariance Matrix Estimation. *Technometrics August, 43*(3), 286-301.

Pei-Chen. W. (2007). Modern One-Way ANOVA F Methods: Trimmed Means, One Step M-estimators. *Quantitative Research, 12(1),* 151-169.

Randles, R. H., Wolf, D. A. (Ed.) (1979). *Introduction to the theory of nonparametric statistics*. John Wiley & Sons. ISBN: 0471042455.

Rock, D. M., Downs, G. W. & Rock, A. J. (1982). Are robust estimators necessary? *Technimetrics, 24*, 95 - 101.

Rocke, D. M. & Woodruff, D. L. (1998). Some Statistical Tools for Data Mining Applications. *Center for Image Processing and Integrated Computing.*

Rosenberger, J. L.& Gasko & (1983). Comparing location estimators: Trimmed means, medians and trimmed mean. In Hoaglin, D.C., Mosteller, F. & Tukey, J.W. (Eds). *Understanding robust and exploratory data analysis, New York: John Wiley & Sons*, 297-336.

Rousseeuw, P. J.& Croux, C. (1992). Time-efficient algorithms for two highly robust estimators of scale. *Computational Statistics, 1*, 411-428.

Rousseeuw, P. J.& Croux, C. (1993). Alternatives to the Median Absolute Deviation. [Theory and Methods]. *American Statistical Association. 88(424).*

Rousseeuw, P. J. & Van Zomeren, R. C. (1990). Unmasking multivariate outliers and leverage points. *American statistical Association, 85(633-651).*

Sepulveda, A.& Nicholas, J. A. (1997). A Simulation Approach to Multivariate Quality Control. *Computers in. Engng, 33 (*No's I-2), 113-I 116.

Siegel, A. F. (Ed.). (1988). *Statistics and data analysis: an introduction*: New York: John Wiley & Sons.

Steiner, S. H. (1994). *Quality Control and Improvement Based on Grouped Data*. (Doctoral dissertation). Retrieved from http//digitalcommons.mcmaster.ca/open dissertations. Paper 3883.

Sullivan, J. H.& Woodal, W. H. (1996). A comparison of Multivariate Control Charts for Individual Observations. *Quality Technology, 28*, 398 - 408.

Tiku, M. L.& Singh, M. (1982). Robust statistics for testing mean vectors of multivariate distribution. *Communication in statistics - theory and methods, 11*(9), 985-1001.

Thompson, J.R. and Koronacki, J. (1993). *Statistical Process Control for Quality Improvement*. Chapman & Hall, 391 pgs.

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

Vargas, J. A. (2003). Robust Estimation in Multivariate Control Charts for Individual Observation. *Quality Technology, 35*, 367 - 376.

Wei, G. C. (2007). *Comparison Between a Robust Control Chart Based on Hodges Lehmann Estimator and Classical Hotelling T square Chart*. Thesis of Master of Applied Statistics. Department of Mathematics, Institute of Mathematical Research, University Putra Malaysia.

Wilcox, R. (2003). Multiple Comparisons Among Dependent Groups Based on a Modified One -step M-estimator. *Biometrical, 44*, 466-477.

Wilcox, R., & Keselman, H. J. (2003). Repeated Measures ANOVA Based on a Modified One-Step M-Estimator. *British Mathematical and Statistical Psychology, 56(1),* 15 - 25.

Wilcox, R.R. & Keselman, H.J. (2003). Modern robust data analysis methods: Measures of central tendency. *Psychological Methods 8 (3)*: 254–274.

Wilcox, R. R. (1995). ANOVA: A paradigm for low power and misleading measures of effect size? *Review of Educational Research, 65(1),* 51-77.

Wilcox, R. R. (1997). Introduction to Robust Estimation and Hypothesis Testing. *Academic Press*.

Williams, J. D., Sullivan, J. H. & Birch, J. B. (2006). Maximum Value of Hotelling's $T^2$ Statistics Based on the Successive Differences Covariance Matrix Estimator. *Communication in statistics-theory and Method. 38(4),* 471 - 483.

Williams, J. D. Woodall, W. H., Birch, J. B. & Sullivan, J. H. (2006). Distribution of Hotelling’s $T^2$ Statistics Based on the Successive Differences Estimator. *American Statistical Association, 38(3), 217 - 229*

Woodall, W. H., & Montgomery, D. C. (1999). Research Issues and Ideas in Statistical Process Control. *Quality Technology. 31*, 376-386.

Syed Yahaya, S. S., Ali, H. & Omar, Z. (2011). Alternative Hotelling $T^2$ Control Charts Based on Minimum Vector Variance (MVV). *Modern Applied Science*. 5(4), ISSN 1913-1844 (print) ISSN 1913-1852 (online).

Syed Yahaya, S. S., Othman, A. R., & Keselman, H. J. (2006). Comparing the Typical Score Across Independent Groups Based on Different Criteria for Trimming. *Metodoloski zvezki, 3*(1), 49-62.

Syed Yahaya, S. S., Othman, A. R., &Keselman, H. J. (2004). Testing the equality of location parameters for skewed distributions using $S_1$ with high breakdown robust scale estimators. In M. Hubert, G. Pison, A. Struyf and S. Van Aelst (Eds.), *Theory and Applications of Recent Robust Methods, Series: Statistics for Industry and Technology* (319-328). Basel, Switzerland: Birkhauser.