ROBUST DIAGNOSTICS AND ESTIMATION FOR DUAL RESPONSE SURFACE FUNCTION WITH HETEROSCEDASTIC ERRORS USING NEW OPTIMIZATION TECHNIQUE IN THE PRESENCE OF OUTLIERS

By

NASUHAR BINTI AB. AZIZ

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

December 2022

FS 2022 64

DEDICATION

To my parents;

My mother Rohani Che Mat &

My Father Late Ab Aziz Yusoff (May his soul rest in perfect peace, Ameen)

and

To my husband (Mohd Bahri Ab Rahman) and my son (Abderrahman Mohd Bahri)

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

ROBUST DIAGNOSTICS AND ESTIMATION FOR DUAL RESPONSE SURFACE FUNCTION WITH HETEROSCEDASTIC ERRORS USING NEW OPTIMIZATION TECHNIQUE IN THE PRESENCE OF OUTLIERS

By

NASUHAR BINTI AB. AZIZ

December 2022

Chairman : Professor Habshah binti Midi, PhD
Faculty : Science

Robust design is a process and quality improvement method that focuses on determining the optimal factor settings to minimize variation in the quality characteristics while keeping a process mean at the customer-identified target value. The dual response surface optimization has become increasingly popular in robust design in order to achieve such aims whereby two empirical models, namely the process mean and process standard deviation of the response are established at the outset. Various dual response optimizations have been proposed, such as Vining and Myers (VM), Lin and Tu (LT), weighted mean square errors (WMSE) and Penalty function (PM) methods. However, the existing methods are not able to obtain an estimated mean response closer to the target value with small variation. In addressing the problem, a new optimization method called penalty function based on decision maker's (PFDMM) preference method is developed. The PFDMM is developed in two stages whereby the penalty constant, ξ, is determined in the first stage and subsequently the PFDMM functions are optimized to obtain the optimal factor settings to estimate the optimal mean response. The performance of the PFDMM is compared to VM, LT, WMSE and PM methods. Empirical evidences via simulation experiments and numerical data have shown that the PFDMM method is more efficient than the existing methods as it has the least bias and RMSE.

In the dual response surface methodology, sample mean and sample standard deviation, are the most commonly used measures for estimating the mean and standard deviation of the response variables. Moreover, the ordinary least squares (OLS) method is usually employed to estimate the parameters of the process mean and process standard deviation response functions. However, the OLS method is tremendously affected by the presence of outliers. Hence, we propose a robust penalty function optimization scheme based on MM-estimator, denoted as $PFDMM_R$. Simulation studies and numerical examples have

proven that the newly proposed $\mathrm{PFDMM_R}$ outperforms the existing methods with the least bias and RMSE values for data with and without outliers.

This research also addresses the combined problem of outliers and heteroscedastic errors for the dual response surface model. A robust reweighted least square (RRWLS) method is proposed and successfully tackles both problems. The newly proposed method consists of two steps to simultaneously solve the problem of outliers and heteroscedastic errors by considering robust weight. The first weight function is used to reduce heteroscedasticity effect, and the second weight function is formulated to reduce the effects of outliers. The reweighted least squares (RLS), two stage robust method based on MM-estimator (TSR-MM) and robust reweighted least squares (RRWLS) are integrated in the algorithm of VM, LT, WMSE, PM and PFDMM optimization methods. The result of simulation study and real datasets show that the PFDMM-RRWLS based method is superior compared to the existing methods discussed in this thesis.

An augmented desirability function (AADF) approach was proposed to optimize the process mean and process standard deviation functions for multiple responses in which each response is replicated *t* times. The AADF employs the OLS method to estimate the parameters of the process mean and process standard deviation functions of the responses. The AADF is formulated by augmenting the overall desirability function of the predicted process mean and the overall desirability function of the predicted process standard deviation into a single overall desirability function, DS_λ. It is observed that the DS_λ can be expressed as geometric mean. However, it is now evident that the OLS and the geometric mean are easily affected by outliers. Thus, a new robust augmented approach desirability function (RAADF) is developed by integrating the MM-estimator, geometric median and geometric MM in its establishment. The results indicate that the performance of the RAADF-geometric MM surpasses the RAADF-geometric median and AADF-geometric mean methods.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

DIAGNOSTIK DAN PENGANGGARAN TEGUH BAGI DUA PERMUKAAN SAMBUTAN DENGAN HETEROSEDASTIK MENGGUNAKAN TEKNIK PENGOPTIMUMAN YANG BARU DENGAN KEHADIRAN TITIK TERPENCIL

Oleh

NASUHAR AB. AZIZ

Disember 2022

Pengerusi : Profesor Habshah binti Midi, PhD
Fakulti : Sains

Reka bentuk teguh ialah proses dan kaedah penambahbaikan kualiti yang menumpukan pada penentuan tetapan faktor optimum untuk meminimumkan variasi dalam ciri kualiti sambil mengekalkan min proses pada nilai sasaran yang dikenal pasti oleh pelanggan. Pengoptimuman dua permukaan sambutan telah menjadi semakin popular dalam reka bentuk teguh untuk mencapai matlamat sedemikian di mana dua model empirikal, iaitu min proses dan sisihan piawai proses bagi sambutan diwujudkan pada awalnya. Pelbagai pengoptimuman dua permukaan sambutan telah dicadangkan, seperti kaedah Vining dan Myers (VM), Lin dan Tu (LT), ralat min kuasa dua berpemberat (WMSE) dan kaedah fungsi Penalti (PM). Walau bagaimanapun, kaedah sedia ada tidak dapat memperoleh anggaran sambutan min yang lebih hampir kepada nilai sasaran dengan variasi yang kecil. Dalam menangani masalah tersebut, kaedah pengoptimuman baharu yang dipanggil fungsi penalti berdasarkan keutamaan pembuat keputusan (PFDMM) dibangunkan. Prestasi PFDMM dibandingkan dengan kaedah VM, LT, WMSE dan PM. Bukti empirikal melalui kajian simulasi dan data berangka telah menunjukkan bahawa kaedah PFDMM adalah lebih cekap daripada kaedah sedia ada kerana ia mempunyai paling kurang bias dan RMSE.

Dalam metodologi dua permukaan sambutan, min sampel dan sisihan piawai sampel adalah ukuran yang biasa digunakan untuk menganggar min dan sisihan piawai bagi pembolehubah sambutan. Selain itu, kaedah biasa kuasa dua terkecil (OLS) biasanya digunakan untuk menganggarkan parameter bagi fungsi sambutan min proses dan sisihan piawai proses. Walau bagaimanapun, kaedah OLS sangat dipengaruhi oleh kehadiran titik terpencil. Oleh itu, kami telah mencadangkan skema pengoptimuman fungsi penalti teguh berdasarkan penganggar MM, yang dilambangkan sebagai $PFDMM_R$. Kajian simulasi dan contoh berangka telah membuktikan bahawa $PFDMM_R$

dicadangkan mengatasi kaedah sedia ada dalam kajian ini dengan paling sedikit pincang dan nilai RMSE untuk data yang mempunyai titik terpencil dan tiada titik terpencil.

Penyelidikan ini juga menangani masalah titik terpencil dan ralat berheterosedastik secara serentak untuk model dua permukaan sambutan. Kaedah kuasadua terkecil berpemberat teguh (RRWLS) dicadangkan dan berjaya menangani masalah kedua-duanya. Kaedah yang baru dicadangkan terdiri daripada dua langkah untuk menyelesaikan masalah titik terpencil dan heterosedastik secara serentak dengan mengambil kira pemberat teguh. Fungsi pemberat pertama digunakan untuk mengurangkan kesan heterosedastik, dan fungsi pemberat kedua dirumuskan untuk mengurangkan kesan titik terpencil. Kaedah kuasadua terkecil berpemberat (RLS), kaedah dua peringkat teguh berdasarkan penganggar MM (TSR-MM) dan kaedah kuasadua terkecil berpemberat teguh (RRWLS) disepadukan dalam algoritma kaedah pengoptimuman VM, LT, WMSE, PM dan PFDMM. Hasil kajian simulasi dan set data sebenar menunjukkan kaedah berasaskan PFDMM-RRWLS adalah lebih unggul berbanding kaedah sedia ada yang dibincangkan dalam tesis ini.

Pendekatan yang diperkukuhkan dengan fungsi *desirability* (AADF) telah dicadangkan untuk mengoptimumkan fungsi proses min dan proses sisihan piawai untuk sambutan berganda di mana setiap sambutan direplikasi sebanyak t kali. AADF menggunakan kaedah OLS untuk menganggarkan parameter fungsi sambutan bagi proses min dan proses sisihan piawai. AADF diformulasi dengan menggabungkan fungsi *desirability* keseluruhan bagi proses min yang diramalkan dan fungsi *desirability* keseluruhan bagi proses sisihan piawai yang diramalkan kepada satu fungsi *desirability* keseluruhan, DS_{λ}. Adalah diperhatikan bahawa, DS_{λ} yang boleh dinyatakan sebagai min geometri. Walaubagaimanapun, ianya terbukti bahawa OLS dan min geometri mudah dipengaruhi oleh titik terpencil. Oleh itu, pendekatan yang diperkukuhkan untuk fungsi *desirability* keseluruhan teguh (RAADF) dibangunkan dengan mengintegrasikan penganggar MM, median geometri dan MM geometri. Keputusan menunjukkan bahawa prestasi RAADF-MM geometrik mengatasi kaedah RAADF-median geometrik dan kaedah AADF-min geometric.

ACKNOWLEDGEMENTS

All praises and glorification are due to Allah (SWT). I thank Him for the ability, wisdom and strength bestowed on me to conduct this study. I really thank Him.

I would like to express my sincere gratitude to my supervisor, Prof. Dr. Habshah Midi, who has taught and helped me over the years. I have benefited enormously from her continuous support and confidence throughout my research. Without her help and support, this dissertation would have been impossible, I feel truly privileged to have been her student. I am also very grateful to my committee members Associate Prof. Dr. Leong Wah June, Dr. Mohd Shafie Bin Mustafa for their guidance, encouragement and support.

I would also like to express my gratitude to my mother Madam Rohani for her prayers, support and encouragement, to also my brothers, sisters and relative for their moral support and prayers. My special thanks go to my beloved husband Bahri for his prayers, endurance, support and encouragement. Similarly, I would like to appreciate my friends, colleagues in research, housemate and well-wishers for their prayers.

Lastly, I gratefully acknowledge the financial support from the Universiti Teknologi MARA (UiTM) as my main sponsor during my studies, without their support this study wouldn’t have been possible.

This thesis was submitted to the Senate of the Universiti Putra Malaysia and has been accepted as fulfillment of the requirement 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)

Leong Wah June, PhD
Professor
Faculty of Science
Universiti Putra Malaysia
(Member)

Mohd Shafie bin Mustafa, PhD
Senior Lecturer
Faculty of Science
Universiti Putra Malaysia
(Member)

ZALILAH MOHD SHARIFF, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 13 April 2023

Declaration by the Graduate Student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software

Signature: ________________________ Date: _______________

Name and Matric No: Nasuhar Ab. Aziz

Declaration by Members of the Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) were adhered to.

Signature:	
Name of Chairman of Supervisory Committee:	Professor Dr. Habshah binti Midi
Signature:	
Name of Member of Supervisory Committee:	Professor Dr. Leong Wah June
Signature:	
Name of Member of Supervisory Committee:	Dr. Mohd Shafie bin Mustafa

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

AADF	Augmented Approach to the Desirability Function
AWLSRE	Adaptive Weighted Least Square Ratio Estimator
BLUE	Best Linear Unbiased Estimator
BP	Breakdown Point
CCC	Central Composite Circumscribed
CCD	Central Composite Design
CCF	Central Composite Face Centered
CCI	Central Composite Inscribe
CN	Copeland and Nelson
DPR	desirability function of the predicted responses, $\hat{y}_i$
DSR	desirability function for the standard deviation of the predicted responses, $\mathrm{sd}(\hat{y}_i)$
DR	Dual Response
DRSM	Dual Response Surface Methodology
DRSO	Dual Response Surface Optimization
IQR	Inter-Quartile Range
IRWLS	Iterative Reweighted Least Squares
GRD	Generalized Reduced Gradient
G-mean	Geometric mean
G-median	Geometric median
G-MM	Geometric MM
$\mathrm{G}t_i$	Generalized Studentized Residuals
LMS	Least Median of Squares
LT	Lin and Tu

LTB	larger-the-better
LTS	Least Trimmed Square
MGQ	Modified Goldfeld-Quandt Test
MAD	Mean Absolute Deviation
MSE	Mean Square Error
MVUE	Minimum Variance Unbiased Estimator
NTB	normal-the-better
OLS	Ordinary Least Squares
PF	Penalty Function Method
PFDMM	Penalty Function Decision Maker's Preference
$PFDMM_R$	Robust Penalty Function Decision Maker's Preference
RGQ	Robust Goldfeld-Quandt Test
RRWLS	Robust Reweighted Least Squares
RPD	Robust Parameter Design
RSM	Response Surface Methodology
RLS	Reweighted Least Squares
RAADF	Robust Augmented Approach to the Desirability Function
STB	smaller-the-better
TSR-MM	Two Stage Robust MM-estimator
TSRWLS	Two-Step Robust Weighted Least Squares
VM	Vining and Myers
WMSE	Weighted Mean Square Error
WLS	Weighted Least Squares

CHAPTER 1

INTRODUCTION

1.1 Background and Purpose

Response surface methodology (RSM) was first developed by Box and Draper (1987). It is an important tool to find the optimal factor settings of the design point, which can either maximize or minimize the given response function. RSM is widely used in many disciplines, such as in manufacturing industries, engineering, and agricultural sciences. The food industry, in particular, has been a prime user of RSM since the early 1970s (Myers et al., 1989). For example, Dey and Dora (2011), studied the effects of temperature, pH, enzyme concentration / substrate concentration (E/S) ratio on the response, i.e., degree of hydrolysis (DH) for marine shrimp. They noted that RSM was successfully applied to determine the optimum factor settings on the control variables for maximum DH value. The traditional RSM emphasizes on locating the optimal process parameters to achieve the target mean value where it assumes a homogeneous variance. However, in a real situation, this assumption may not be achieved. In this situation, both the mean and the standard deviation of the responses should be considered when determining the optimum factor settings for the control variables which lead to the concept of dual response surface methodology (DRSM). The DRSM is the most efficient method for simultaneously optimizing the estimated mean and the estimated standard deviation functions of the responses to achieve the desired target while keeping the standard deviation small (Myers et al., 1989). Several optimization methods such as VM, MSE, WMSE and Penalty function methods have been proposed to simultaneously optimizing both functions of the estimated mean and the estimated standard deviation of the responses to obtain the optimal factor settings for the control variables. Various optimization models have been developed, and the widespread application of these techniques has resulted in significant improvement in product quality (Manojkumar, 2022).

1.2 Response Surface Designs

In this section, the central composite design (CCD) is often employed in fitting the second-order polynomial model since it is the most popular class of designs in RSM. The three types of CCD designs and its characteristics are discussed in the following subsections.

1.2.1 The Structure of Central Composite Design

The central composite design is the most popular class of region of second-order designs. In fact, the basic of CCD is derived from a sequential experiment and turns out to be an effective tool for non-sequential response surface experiment. CCD is an efficient design that is ideal for sequential experimentation and allows a reasonable amount of information for testing lack of fit while not involving a large number of designs points. The CCD consists of three significant components:

(a) 2^k corner points which is the base of CCD formed by a two-level full factorial design. This component provides the information in estimation of linear term and two level interactions effects. The corner points take the coded coordinates of the form $(\pm 1, \pm 1, \ldots, \pm 1)$ (Montgomery et al., 2021). For example, for $k = 2$;

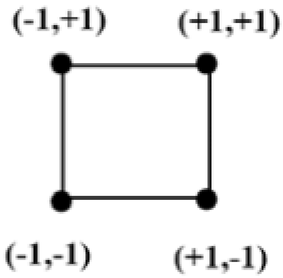


Figure 1.1 : The CCD with 2^k corner points

(b) 2^k star points (axial points). These points permit the estimation of all quadratic terms. When $\alpha > 1.0$, significant test for higher-order curvature effects can be conducted. However, the axial points take the coordinates $(\pm\alpha, 0, \ldots, 0), (0, \pm\alpha, \ldots, 0)$, etc. For example, for $k = 2$;

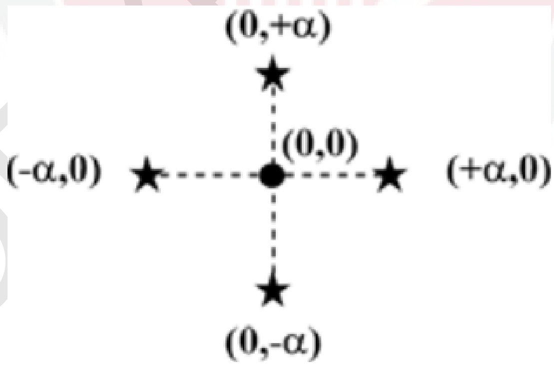


Figure 1.2 : The CCD with axial points $\alpha > 1.0$

(c) n_0 center points provide the information about the existence of curvature in the system. Hence, it provides an internal estimation of error (pure error) and contributes toward the estimation of quadratic terms. The coded coordinates of the center point with replication are $(0,0, \ldots, 0)$. For example, for $k = 2$;

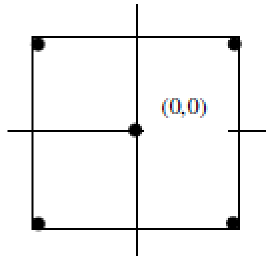


Figure 1.3 : The CCD with corner, axial and center points

The Table 1.1 shows the settings of the axial points is as follows:

Table 1.1 : The settings of the axial point

x_1	x_2	…	x_k
$-\alpha$	0	…	0
α	0	…	0
0	$-\alpha$	…	0
0	α	…	0
.	.		.
.	.		.
.	.		.
0	0	…	$-\alpha$
0	0	…	α

Hence, the setting of CCD for $k = 2$ variables with a center point is shown in Table 1.2.

Table 1.2 : The full settings of CCD for $k = 2$

x_1	x_2
-1	-1
1	-1
-1	1
1	1
$-\sqrt{2}$	0
$\sqrt{2}$	0
0	$-\sqrt{2}$
0	$\sqrt{2}$
0	0
0	0

1.2.2 Types of Central Composite Design

In general, there are three main types of CCD: circumscribed, face-centered, and inscribed. Each design has its own characteristics where the experimenter often chooses the design based on the region of interest and number of levels for each factor. Many experimenters are often doubtful in choosing the best type of CCD to use in a given study. In order to make the right choice, the experimenter must first understand the differences between these types in terms of the experimental region of interest and region of operability (Montgomery et al., 2021).

1.2.2.1 Central Composite Circumscribed (CCC)

Central Composite Circumscribed (CCC) provides a spherical region and it requires five levels for each factors. CCC is obtained by augmenting the factorial. The distance between axial point and center point, α is greater than 1. Essentially, a CCC design provides high quality predictions over the entire prediction space. Figure 1.4 presents a CCC for time and temperature using the setting in Table 1.1. As the figure shows, the rotatable CCD uses an α value 1.4 to describe a circular design geometry.

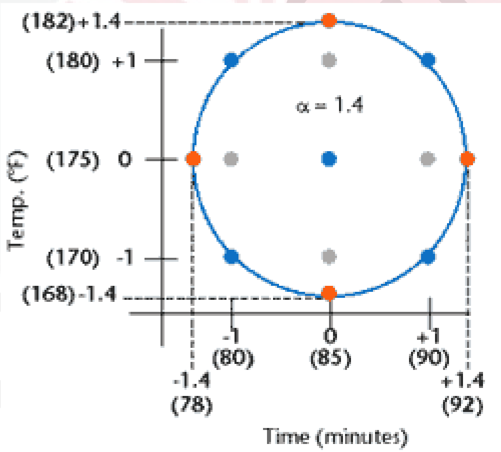


Figure 1.4 : Two variables CCC design

1.2.2.2 Central Composite Face Centered (CCF)

Central Composite Face Centered (CCF) provides a cubical region and it requires three levels for each factors. The distance between the axial point and center point, α is equal to 1. For this reason, CCF is not a rotatable design. Figure 1.5 presents a CCF design for two study variables: time and temperature. Both code and natural variable level settings for time and temperature are shown in the figure. The design consists of a center point, four factorial points (corner points) and four axial points (points parallel to each variable axis on a circle of radius equal to 1.0 and origin at the center). The points in Figure 1.5

identify the variable level setting combinations that constitute the nine design points (experiments runs). According to Montgomery et al. (2021), as the value of α increases, the axial points extend beyond the faces of the square and the design region becomes more spherical.

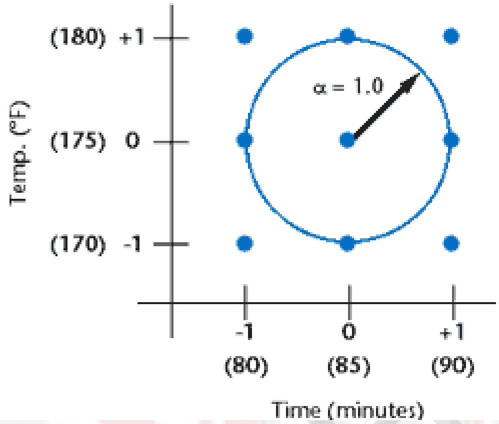


Figure 1.5 : Two variables CCF design

1.2.2.3 Central Composite Inscribe (CCI)

Central Composite Inscribe (CCI) is applicable in the situation where the range of a factors setting is limited. CCI uses the factors setting (± 1) as the axial points and creates a factorial or fractional factorial design with this range. In other words, this design is a scaled down CCC design with each factor level of CCC design divided by α to generate the CCI design. Similarly, CCI requires five levels of each factor. Since it uses only the points within the factor ranges that originally specified, it does not provide the same high quality prediction over the entire space when compared to the CCC.

Figure 1.6 presents a CCI design for time and temperature, again using the previously defined lower and upper variable bounds. The CCI also uses an α value of 1.4 to describe a circular geometric region. However, inscribing restricts the actual design region to the defined variables ranges by locating the axial points at the lower and upper bounds of the variables ranges. The factorial points are brought into the interior of the design space (inscribed) and set at a distance from the center point that preserves the proportional distance of the factorial points to the axial points.

The inscribed option is a convenient way to generate a rotatable CCD that enables the experimenter to study the full ranges of the experiment variables while excluding non-allowable operating conditions at one or more of the extremes of the design region. The excluded portion of the original CCF region is shown in grey in Figure 1.6 as are the excluded CCF points (Montgomery et al., 2021).

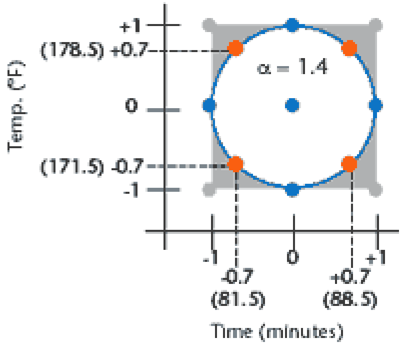


Figure 1.6 : Two variables CCI design

Hence, the setting for each design is also distinct. For example, the setting for $k = 2$ CCC, CCF and CCI are given in Table 1.3.

Table 1.3 : CCC, CCF and CCI settings for $k = 2$, $n_0 = 1$

CCC		CCF		CCI	
x_1	x_2	x_1	x_2	x_1	x_2
-1	-1	-1	-1	-0.707	-0.707
1	-1	1	-1	0.707	-0.707
-1	1	-1	1	-0.707	0.707
1	1	1	1	0.707	0.707
-1.414	0	-1	0	-1	0
1.414	0	1	0	1	0
0	-1.414	0	-1	0	-1
0	1.414	0	1	0	1
0	0	0	0	0	0

1.3 Basic Concepts of Robust Estimators

There are three main concepts related to robust estimators. The three key concepts are discussed in the following subsections.

1.3.1 Efficiency

Efficiency refers to the extent to which the estimator executes the least squares under the conditions where the normal distributed errors occur. It is generally represented as a percentage and can be calculated with the help of the mean squared error or variance of the least squares fit divided by the mean squared error or variance of the robust fit. An estimator can be referred to as efficient if the minimum variance unbiased estimator (MVUE) reaches the minimum variance for all the parameter estimates.

According to Simpson (1995), the desirable range for efficiencies is 90%-95%. For normal distribution error, the efficiency can be represented as

$$\text{Efficiency} = \frac{\text{MSE}\ (\text{OLS}_C)}{\text{MSE (robust)}}\ x\ 100$$

where OLS_C refers to OLS for clean dataset.

1.3.2 Breakdown Point

The idea of a breakdown is more understood with respect to the concept of the robust regression estimating methodologies. The estimator is noted by the several outliers considered for rendering it as ineffective. The breakdown point is expressed in the form of a percent value or a fraction and it is described as the quantity of contamination which is allowed in a data set before all estimates present incorrect information regarding the parameters. There are a few regression estimators which have extremely small possible breakdown points of $1/n$, or 0%. In these cases, a single outlier would make the regression equation useless. Some other robust regression estimators have maximal breakdown points of $n/2$ or 50% such as least median of square, least trimmed square, S and MM-estimator. The highest possible breakdown point (BP) is 50%, because the estimate keeps bounded when fewer than 50% of the data are replaced by outlying observations (Maronna et al., 2019). The breakdown point of an estimator T_x given the data matrix X_y, is defined as:

$$\text{BP}\left(T_x/X_y\right) = \min\left\{\frac{m}{n} : \text{SUP}_{X*y}\left\|T_x\left(X_y\right) - T_x\left(X_y^*\right)\right\|\right\} = \infty$$

where the supremum represents the overall possible data matrix X_y^*, which includes m contaminated points and *n-m* observation – refer to Donoho and Huber, (1983), Leroy Rousseeuw, (1987), Maronna et al., (2019). In layman's terms, breakdown point is defined as the smallest fraction of contamination that can cause an estimator to take an arbitrary large value. The least squares estimator possesses a breakdown point that goes as low as $1/n$, which means that even a single outlying observation can render an OLS estimator useless. Alternatively, some robust regression estimators such as least median of squares, S and MM-estimator, and least trimmed squares have high BP values of about 50%. The highest possible value for BP is equal to 50% due to the fact that the estimate stays bounded when the outlying observation replaces less than 50% of the data (Rousseeuw & Croux, 1993).

1.3.3 Bounded Influence

Another vital parameter in terms of the robust technique is associated with the bounded influence. This pertains to the capacity of the method to manage the effect on the estimation of the model and which is presented by the outlier points in the *X*-space (also known as the high leverage point) (Simpson, 1995). Certain outliers, especially the ones found on the X-axis, can have a significant impact on the estimation of the coefficient for the regression equations. A bounded influence should be possessed by the estimator, as it limits the effect that the outlier data points have (Birkes & Dodge, 1993). The least squares have high susceptibility to the high leverage points. However, an unbounded influence is possessed by some of the robust methodologies.

1.4 Motivation of the Study

As will be discussed in Chapter 2 (Literature Review), many optimization methods have been proposed to simultaneously optimize the fitted functions of the mean and standard deviation of the response surface models (Vining & Myers, 1990; Lin & Tu, 1995; Copeland & Nelson, 1996; Ding et al., 2014). Among all of the methods that will be mentioned, most of them failed to obtain estimated mean response closer to the target value with small variation. Therefore, Baba et al. (2015) proposed a penalty function-based approach (PM) as another alternative optimization scheme. Nonetheless, this method concentrates more on the bias and forces the estimated mean response to be close to the target value, which is not adequately efficient. The PM method also has another shortcoming whereby no specific rule is imposed for determining the penalty constant, ξ. It is difficult to be used because the value of penalty constant, ξ is determined on trial and error basis. Hence, their work has motivated us to propose a new procedure which is called penalty method based on decision maker's (PFDMM) preference to obtain the penalty constant, ξ. The PFDMM is expected to be more efficient optimization scheme that is capable of obtaining estimated mean response closer to the target value while keeping small variations in the response variables.

In the classical dual response approach problem, the sample mean and the sample standard deviation are usually used to estimate the process mean and process standard deviation of the responses. Moreover, the OLS methods are popularly used to estimate the parameters of the process mean and process standard deviation functions of the responses. It is now evident that the sample mean and sample standard deviations and the OLS methods are easily affected by outliers. In the presence of outliers, Park and Cho (2003) introduced replacement of sample mean with sample median, and used mean absolute deviation (MAD) and inter-quartile range (IQR) as an alternative to sample standard deviation. These estimators have been successfully used in different areas and well-known as outlier-resistant estimators. Although these estimators are resistant to outliers, they are not efficient under normal errors (Bakar & Midi, 2015). To overcome this problem, Park and Leeds (2016) used Hodges-Lehman estimator and Shamos estimator as an alternative to the location and scale estimates, respectively. According to Hettmansperger and McKean (2010), the breakdown points of Hodges-Lehman estimator and Shamos estimator are 29.3%. Yohai (1987) proposed MM-estimator which is not only highly efficient and robust, but it also has high breakdown points up to 50%. Hence, to get more efficient estimates of the mean and standard deviation of the response when outliers are present in a data, we integrate the MM-estimator in the establishment of the robust PFDMM optimization scheme. The MM-estimates of regression and MM-estimator of mean and standard deviation are used to down weight the effect of the outliers in the responses.

This thesis also addresses the issue of parameter estimation of the response functions for the process mean and process standard deviation in the presence of heteroscedasticity and outliers. The reweighted least square (RLS) proposed by Goethals and Cho (2011) is used to estimate the model parameters when the assumptions of constant error variances are violated. Their work did not investigate the effect of outliers on the parameters estimates and consequently the estimated mean optimal response will be affected in their presence. Several methods are used to remedy the combined problem of outliers and heteroscedasticity in dual response surface models that include the two-stage robust (TSR-MM) MM-estimator of Mustafa and Midi (2019). The TSR-MM estimator are formulated using two stages whereby in the first stage, the RLS was employed and subsequently the MM procedure is employed in stage 2. Unfortunately, the TSR-MM is not effective enough since this procedure does not take into consideration in identifying the exact number of vertical outliers which are present in the dataset so that their effect can be minimized. A good robust estimator is one that can identify the exact number of outliers in a data set and minimize their effect (Dhhan et al., 2017; Rashid et al., 2021). The weakness of this estimator has inspired us to develop a new robust reweighted least squares method (RRWLS) that can simultaneously rectify the problems of heteroscedasticity and outliers. Subsequently, the RRWLS is integrated in the algorithm of PFDMM.

The desirability function approach which was introduced by Harrington (1965) has been widely used for simultaneously optimizing multiple responses problem. However, the main problem in classical desirability function is that, this approach is only interested to optimize optimal factor settings based on predicted responses, $\hat{y}_l$ assuming variability between responses is constant. In practice, this assumption is not always satisfied. As a results, Chen et al. (2012) proposed an augmented approach to the desirability function (AADF) method, whereby besides having introduced the transformation d_l for the predicted responses, $\hat{y}_l$, they also introduced the transformation d_l for the standard deviation of the predicted responses, $\text{sd}(\hat{y}_l)$. The overall desirability functions for the predicted responses, $\hat{y}_l$ and the standard deviation of the predicted responses, $\text{sd}(\hat{y}_l)$ were then constructed. In order to obtain the optimal factor settings, they combined the overall desirability function of the predicted responses, $\hat{y}_l$ (DPR) and the overall desirability function for the standard deviation of the predicted responses, $\text{sd}(\hat{y}_l)$ (DSR) into a single overall desirability function, denoted as DS desirability function by assigning weight to DPR and DSR overall desirability functions. Unfortunately, the DS desirability function is not effective enough, since the transformation d_l for the standard deviation of the predicted responses, $\text{sd}(\hat{y}_l)$ is not desirable to minimize the variation of each response and also their approach is not applicable for the experimental data with replicates for each factor setting. Therefore, Chen et al. (2013), attempted to solve simultaneously optimization of the process mean and process standard deviation for multiple responses problem with the main aim to minimize the process standard deviation of the response. Therefore, Chen et al. (2013) used the smaller-the-better (STB) type of the transformation, d_l introduced by Derringer and Suich (1980) to the transformation, d_l for the process standard deviation, meanwhile for the process mean, larger-the-better (LTB), normal-the-better (NTB) and also smaller-the-better (STB) type is used as the transformation d_l depending on whether the optimal process mean of the response has to maximized, attained a target value or minimized. The DS desirability function approach proposed by Chen et al. (2012) has been used to simultaneously optimize the process mean and process standard deviation of the responses. The augmented desirability function (AADF) approach of Chen et al. (2013) is applicable for experimental data with replicates in each of the response variable for each factor setting.

We observed that the DS used by Chen et al. (2012) and Chen et al. (2013) can be expressed as geometric mean. Moreover, their approaches are based on the OLS estimator and geometric mean which is very sensitive to outliers. Midi et al. (2013) proposed AADF-MM estimator for multiple responses. MM- estimator has been used to estimate the parameters of the second-order polynomial models. Then, they improvised the AADF method of Chen et al. (2012). In their method, they proposed geometric median for the overall desirability function for the standard deviation of the predicted responses, $\mathrm{sd}(\hat{y}_l)$. However, it is not clearly written, how the overall desirability function of the AADF-MM is formulated. Hence, practitioners find this method difficult to use. Moreover, the AADF-MM can only be applied to the experimental data with un replicate response variable for each factor settings. Thus, this motivates us to develop robust augmented approach to the desirability function (RAADF) which is based on the MM estimator, geometric median and geometric MM which are outlier resistant estimator.

1.5 Research Objectives

The foremost objective of our research can be outlined systematically as follows:

1. To establish a new penalty constant function optimization (PFDMM) based on the newly developed optimal penalty constant, ξ, in the dual response methodology in accordance with a decision maker's preferences.
2. To improve the performance of the proposed PFDMM method when outliers are present in the data by employing the MM-estimator in the algorithm of PFDMM method.
3. To establish a new PFDMM technique based on the newly developed robust estimation technique (RRWLS) for dual response function models in the presence of heteroscedasticity and outliers.
4. To develop a robust augmented desirability function technique for multiple responses based on MM-estimator, geometric median, and geometric MM in the presence of outliers.

1.6 Scope and Limitation of the Study

Dual response surface methodology is the widely used method in many fields of study such as engineering, clinical trials and high-tech industries. Most of the datasets collected using an experimental procedure. The experimenters such as scientists and engineers have to follow research ethics to publish their datasets. As such, not much referred real datasets in the literature.

The response surface design has different types of experimental design, such as central composite design (CCD), factorial design and Box Behnken design. All types of design have their limitation, the most crucial part is to consider total design points that will be used. For CCD design, the coded coordinates for the n_o center point, let $x_{ij} = (x_{i1}, x_{i2}, \ldots, x_{ip})$ is written as $x_i = (0,0, \ldots 0)$. The more design points used; the more n_o center points should be introduced and this caused problems in the computation of the fitted response functions for the process mean and process standard deviation. Therefore, throughout the thesis we only conduct the simulation study with 27 and 45 design points with 3^2 designs.

Since robust statistic is relatively new technique in dual response surface model, there are not so many algorithms and statistical software related to dual response surface model are available. Writing our coded algorithms by R-programming is the most challenging job.

1.7 Outline of the Thesis

In accordance with the objectives and the scope of the study, the contents of this thesis are organized in seven chapters. The thesis chapters are structured so that the study objectives are apparent and are conducted in the sequence outlined.

Chapter Two: This chapter briefly presents the literature reviews of the dual response surface optimization. This chapter also reviews some robust estimators and robust regression methods for parameter estimation in the presence of outliers. Different types of outliers in regression and experimental data and the existing techniques to detect them are also presented. This chapter also reviews on heteroscedasticity with example and its consequences and, heteroscedasticity with its usual detection and estimation technique are also reviewed by highlighting the strengths and weaknesses of existing methods. Finally, the desirability function methods are discussed.

Chapter Three: This chapter discusses the development of new penalty function optimization based on decision maker's preference (PFDMM). The Monte Carlo simulation study and two real datasets are discussed to evaluate the performance of the proposed method.

Chapter Four: This chapter deals with development of robust design optimization-based MM-estimates of process mean and process standard deviation and MM- estimates of regression with the new proposed penalty function optimization based on decision maker's preference (PFDMM). The derivation of MM-estimator for the process mean and process standard deviation functions of dual response surface models are presented. The new proposed robust optimization approach is described. Numerical examples and simulation study are presented.

Chapter Five: This chapter deals with the new proposed weighting method and estimation technique for dual response surface model in the presence of outliers and heteroscedastic errors. The classical reweighted least squares estimators are presented. The new proposed method which is called robust reweighted least squares (RRWLS) algorithm is presented. Simulation study and numerical examples are presented to evaluate the performances of the proposed RRWLS estimator and the proposed PFDMM optimization based on the RRWLS estimator.

Chapter Six: This chapter deals with the development of the augmented approach to the desirability function (AADF) to optimize dual response surface model for multiple responses with replicates for each factor settings in experimental data in the presence of outliers. The classical AADF method which can be expressed as geometric mean, used OLS method to estimate the parameters of the response functions models that can be adversely affected by outliers. The new proposed method which is called the robust augmented to the desirability function (RAADF) algorithm is presented to resolve the issues. The simulation study data and real data examples are presented.

Chapter Seven: This chapter provides summary and detailed discussions of the thesis conclusions. Areas for future research are also recommended.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

The response surface methodology (RSM) has played a significant role in designing, developing, formulating, and analyzing of new products and scientific studies. RSM typically involves experimental design, regression model and optimization for more than one response. The primary objective of RSM is to find the optimal settings conditions on a set of exploratory variables that minimizes or maximizes the fitted second-order polynomial model. The widespread application of RSM in solving real life industrial problem make it necessary to find a more sophisticated and efficient approach that can achieve the needed results for a decision maker. Most of the early research in RSM was based on a single response case. This approach was introduced and developed in Box and Wilson (1951) and further study was presented by (Draper, 1963; Khuri & Conlon, 1981). Continuous research in this area lead to the development of dual response approach, where Myers and Carter (1973) suggested the used of mean and standard deviation functions as separate function in their effort to tackle the unequal variance problem. Dual response surface model is an approach that uses RSM for modelling second-order polynomial process mean and process standard deviation which subsequently optimize both of their fitted functions. This approach is tremendously affected by the presence of outliers. The presence of outliers leads to getting incorrect optimal factor settings and this resulting to less efficient predicted optimal mean and optimal standard deviation of the response variable. Other than that, the effect also in terms of biasness and wrong inferences due to the violation of the normally distributed errors assumption of the model. Biased estimates also produced when the assumptions of homoscedasticity errors are violated (Ismaeel et al., 2021).

This chapter focuses on the literature reviews of the most important issues discussed in this thesis which are related to optimization in dual response surface model, robust estimation methods, identification of outliers, heteroscedasticity diagnostic methods, remedial techniques for heteroscedasticity and the desirability function methods.

2.2 Dual Response Optimization

A Japanese engineer known as Genich Taguchi in 1980’s first introduced robust design optimization. Taguchi's (1985) philosophy and quality-focused methods, which are the foundation of robust parameter design (RPD) or, briefly, robust design optimization, evolved to provide a cost-effective approach to improving product and process quality. The customer's perspective, according to Taguchi's concept, requires manufacturers to evaluate both the mean (and goal specification) and the variability for a given quality measure in order to improve the quality of the supplied product as well as the quality of the manufacturing process. Despite the fact that some scholars such as Vining and Myers (1990), Myers et al. (1992) and Tan et al. (2019) agree with Taguchi's philosophy, his technical approaches to RPD have been heavily criticised due to a variety of perceived mathematical flaws. Kackar (1985), Vining and Myers (1990), Myers et al. (1992), have addressed the controversy about his assumptions, approaches to experimental design, and analytical procedures.

Taguchi's methodology, according to a number of scholars such as Robinson et al. (2004) and Sabarish et al. (2020), just assisted process improvement rather than optimization. As a result, various RPD models have been proposed that study a number of alternative optimization strategies that are more solidly founded in well-established experimental design and response surface methodology (RSM). As a result, many research efforts have been proposed to rectify these weaknesses. Myers and Carter (1973) first introduced dual response (DR) model based on RSM, and then were extended by Vining and Myers (1990). Vining and Myers (1990), suggested that two separate models are established for the optimization response process, the process mean and the process standard deviation. Vining and Myers (1990) formulated the dual response problem as a primary response subject to an equality constraint on the secondary response. The primary and secondary responses can be expressed as the second-order polynomial fitted model response functions.

This approach includes two significant contributions: (1) in terms of estimations, the mean and variance output responses are considered and estimated as two separate functions of input factors based on the OLS; and (2) in terms of optimization, a robust design model is used to investigate the trade-off between the estimated process mean and process standard deviation functions. Based on the DR model approach, the second development stage of RPD has three sequential stages; the design of experiments, estimations and optimization.

Various optimization methods of dual response surface optimization (DRSO), which attempt to simultaneously optimize both the fitted functions of the process mean and process standard deviation of the responses, were proposed in the literature (Tang & Xu, 2002; Kovach et al., 2008; Baba et al., 2015; Bertsimas et al., 2022). Vining and Myers (1990), were the first to introduce dual response surface optimization (DRSO) using response surface methodology (RSM). We denote this method as the VM method. However, their method received criticism from Del Castillo and Montgomery (1993) due to using Lagrange multipliers to simultaneously optimize the process mean and process standard deviation. They noted that this approach may be impractical and imprecise for a global best solution due to the restriction of the process mean which was made to be equal to a specific target value. Thus, they proposed the generalized reduced gradient (GRD) technique with inequality constraints to rectify the drawback of the VM optimization method.

Similarly, Lin and Tu (1995) also pointed out the weakness of VM method in which it did not guarantee a global optimal due to its constraint to a specific value. Lin and Tu (1995) proposed an optimization scheme based on squared-loss model, which is known as mean square error (MSE). This model deals with square bias and minimizes the variance component of the factor settings. The Lin and Tu (LT) method is to minimize the MSE function. This optimization scheme is widely used in order to find the optimum settings in robust design (Park & Cho, 2003; Goethals & Cho, 2011; Boylan & Cho, 2013; Park & Leeds, 2016). Park and Leeds (2016) used the LT optimization scheme with the robust sample mean and sample standard deviation estimators of the response variables. Chelladurai et al. (2021) suggested using RSM and the analysis of variance technique to optimize the process parameters in various manufacturing processes. However, as already mentioned, the RSM has its weakness whereby it only focused on optimizing the mean response and assuming that the variance of the response is constant which is not valid in practice.

Kim and Lin (1998) also pointed out that the objective function of LT does not consider the measure of the violation of constraint, and it does not clearly specify how far the estimated mean response $\hat{\mu}(x)$ is from the desired target value τ. In an attempt to reduce the effect of bias and variance in obtaining the optimal setting conditions for the input variables, Ding et al. (2004) modified the LT method by imposing relative weights, w on the bias and the variance terms of the MSE optimization scheme. Their method is called weighted MSE (WMSE), where W = relative weight factor $(0 \leq w \leq 1)$. According to Jeong et al. (2005) the data-driven approach is used to determine the value of W. However, they pointed out that it has major shortcoming whereby it utilizes no input from the decision maker (DM). Jeong et al. (2005) and Lee and Kim (2012) then introduced the proper procedure to determine the value of W by taking into consideration the decision maker's preference in consideration. Nevertheless, most of the existing methods do not show much improvement in terms of acquiring an estimated mean response that is closer to the target value with small variation.

Although the LT optimization work well for experimental data, it can be a burden to process engineers when dealing with a large number of experiments. Lee et al. (2018) noted that the LT optimization is not practical for large volume of operational data such as from manufacturing lines. Then, Lee et al. (2018) proposed an optimization scheme based on data mining approach for operational data. Nonetheless, the LT optimization scheme is suspected not to be that efficient since it does not achieve the target with minimal bias in the estimated process mean response due to some little bias introduced in its formulation (Baba et al., 2015). The weakness of LT methods has motivated Baba et al. (2015) to propose optimization scheme based on penalty function and call it penalty function optimization (PM). This PM approach has been used to optimize the factor settings for unbalanced experimental data combination with weighted least square (WLS) estimator (Baba et al., 2022). The result shows that, this PM method is more efficient when compared with LT method for unbalanced data. However, one of major issues of PM method is that it has neither a clear rule nor systematic approach to determine the value of the penalty constant, ξ. Hence, a proper procedure to determine the penalty constant, ξ, is needed prior to developing new optimization method so that more accurate predicted optimal mean response is obtained with smaller bias and smaller variance.

2.3 Robust Estimation Methods

It is important to highlight that in most optimization methods of dual response surface model, the ordinary least squares (OLS) is the commonly used method for estimating the models' parameters. The classical sample mean and classical sample standard deviation are also popular used measures of location/central tendency and measures of spread, respectively. Unfortunately, many statistics practitioners are not aware that the classical mean, classical standard deviation and OLS method are not resistant to outliers (Dhhan et al., 2017). Outlier(s) is a single or group of observations which are markedly different among the bulk of data or pattern set by most observations. It is now evident that outliers may have an unduly effect on the sample mean, sample standard deviation, and OLS estimates (Rana et al., 2012; Alguraibawi et al., 2015; Maronna et al., 2019; Uraibi & Midi, 2020; Rashid et al., 2021).

Since the resulting optimum responses are inefficiently determined by the sample mean and sample standard deviation, Park and Cho (2003) and Lee et al. (2007) suggested replacement of sample mean with sample median, and, mean absolute deviation (MAD) and inter quartile range (IQR) were used as an alternative to sample standard deviation. These estimators have been successfully applied in various areas and are well-known as outlier-resistant estimators. Median is chosen since the measure can easily be derived and provide robustness up to 50% with respect to outliers (Maronna et al., 2019). However, at normal errors, median is known to be less efficient than the mean which is measured at only 64% (Lee et al., 2007; Bakar & Midi, 2015; Maronna et al., 2019).

To overcome this problem, Park and Leeds (2016) used the Hodges-Lehman estimator and Shamos estimators as an alternative to the sample mean and sample standard deviation estimate. The MAD and Shamos estimator are Fisher-consistent (Fisher, 1922) for the standard deviation under the normal distribution, that is, as the sample size goes to infinity, it converges to the standard deviation under the normal distribution. However, due to the cost of sampling and measurements in practical applications, the experimenter often needs to find parameter estimates from data with small sample sizes, whereas these estimates have serious biases, especially when the sample size is not large enough (Williams, 2011). Therefore, Park et al. (2022) investigated the performance for the MAD and Shamos estimators for small data. The analysis is based on unbiasing factors and relative efficiency. The results show, unbiasing factors significantly improves the small sample sizes (also known as finite-sample) performance for both estimators. According to Hettmansperger and McKean (2010), the breakdown points for both Hodges-Lehman estimator and Shamos estimators is 29.3%. Hence a more robust method which is more efficient and have high breakdown points are needed to estimate the sample mean and sample standard deviation of the responses of the RSM model.

Since the beginning of 1960's, many robust estimators have been introduced in the ordinary, non-response surface model to provide resilient estimates. Robust estimators are known to be resistant to outliers. Many robust estimation methods, such as M-estimator, Generalized M-estimator (GM), least median square (LMS), least trimmed square (LTS), S-estimator, and MM-estimator, have been proposed (Huber & Ronchetti, 1981; Leroy & Rousseeuw, 1987; Yohai, 1987; Huber, 2005; Wilcox, 2005).

Huber and Ronchetti (1981) proposed M-estimator which is a very popular robust estimator, which has its basis on the concept of the replacement of the squared residuals, e_i^2 with a different residual function. The M-estimators are seen to be as effective as the OLS methodology and they also resist the outlier values which are present in the *Y*-direction. The M-estimator is determined by minimizing the sum of a less rapidly increasing function of residuals, as follows:

$$\min_{\beta} \sum_{i=1}^{n} \rho(r_i) = \min_{\beta} \sum_{i=1}^{n} \rho\left(y_i - \sum_{i=1}^{n} x_{ij}\hat{\beta}_j\right)$$

where $\rho(.)$ is an objective function. Differencing $\rho(.)$ function give rise to ψ function.

The ψ-function of M-estimator controls the weight assigned to each residual and hence, dampening the influence of outlying cases. There are different existing ρ- function, ψ-function and ω-weight functions in the literature (Andrews, 1974; Hampel et al., 1986; Huber, 2005). Among the functions is monotonic Huber function (Bickel, 1975; Huber, 2005) and redescending function Tukey's biweight or bisquare function by Beaton and Tukey (1974). Figure 2.1 presents $\rho(.), \psi(.)$ and $\omega(.)$ for least square (LS), Huber and Tukey bisquare estimates, respectively. We can see clearly that the $\rho(.)$ for both the LS and Huber increase without bound as the error, r deviates from 0, but the $\rho(.)$ for LS increases more rapidly. In contrast, the $\rho(.)$ for bisquare is bounded. The Huber $\psi(.)$is a monotonic while bisquare has a redescending $\psi(.)$.

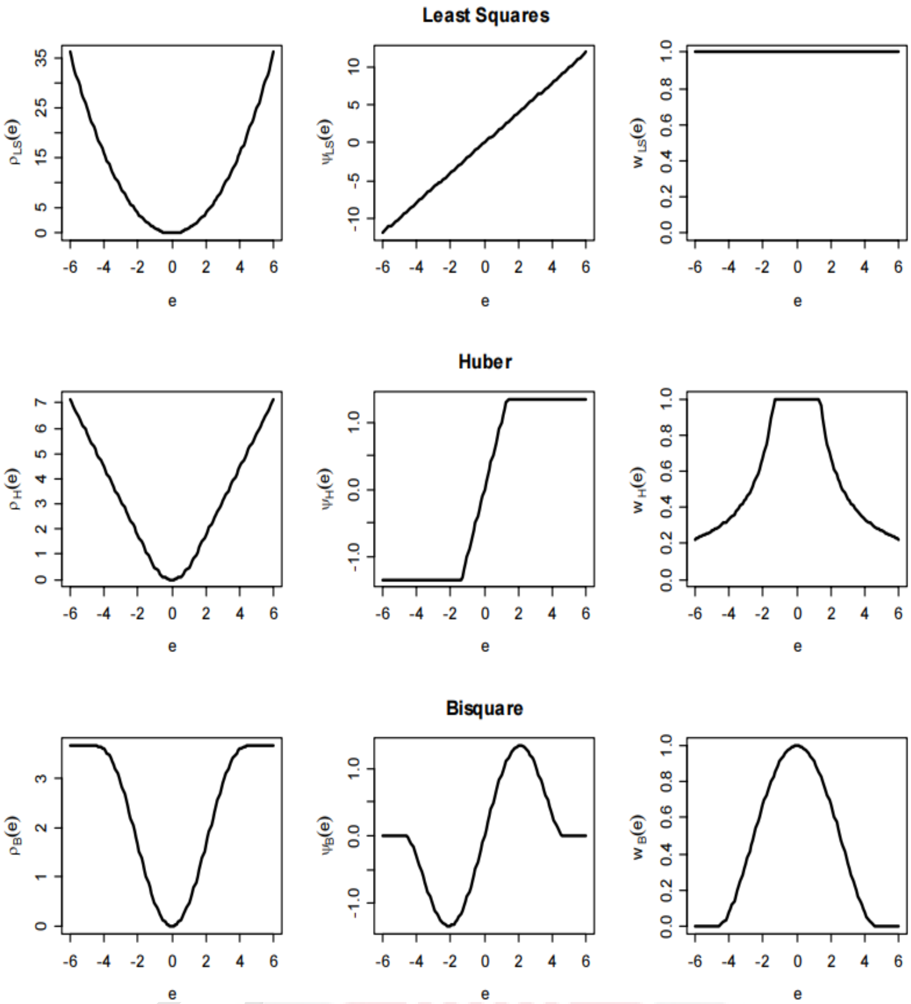


Figure 2.1 : $\rho - function$, $\psi - function$ and $\omega - function$ for LS, Huber (with $k = 1.345$) and bisquare (with $k = 1.345$) estimates

Yohai and Zamar (1988) pointed out that one of the aims of robust regression is to achieve high efficiency, high breakdown point (close to 50%), and bounded influence properties. The breakdown of M-estimator is very low which is equals to ($1/n$) and is non bounded influence. It can handle vertical outliers but not successful in handling high leverage points (also known as outliers in the X-direction).

To reduce this shortcoming, the bounded influence Generalized M-estimators (GM-estimators) had been proposed by Schweppe (see Hill, 1977; Anderson, 2008). The main purpose of GM-estimator is to produce weights that consider both Y and X-directions. Many types of GM-estimators were proposed in literature (Coakley & Hettmansperger, 1993; Wilcox, 2005; Anderson, 2008; Midi et al., 2021). However, in this thesis, our study is limited to experimental data with fixed independent variables, X; hence only consider outlier in the Y-direction not in X-direction. Therefore, GM-estimator is not suitable estimator to estimate the parameter for response surface model.

Rousseeuw and Yohai (1984) suggested a robust method named S-estimator which is high breakdown point. These estimators are the solution that finds the smallest possible dispersion or scatter of the errors (Campbell et al., 1998)

$$\min \hat{\sigma} = \left(r_1(\hat{\beta}), r_2(\hat{\beta}), \ldots, r_n(\hat{\beta})\right)$$

The objective function for S-estimator is given by

$$\min \frac{1}{n}\sum_{i=1}^{n} \rho\left(\frac{r_i}{\hat{\sigma}_r}\right) = b$$

Despite the S-estimators have high breakdown points, but it has low relative efficiency of 30% (Croux et al., 1994).

Rousseeuw (1984) proposed a least median of squares (LMS) as an alternative to robust technique. The LMS estimator is obtained by minimizing the median (Med) of squared errors rather than the sum of squares error in least squared method, as

$$\text{minMed}\ (r_i^2) = \text{minMed}\left(y_i - \sum x_{ij}\,\hat{\beta}\right)^2$$

Despite, the LMS estimator has high breakdown points of 50%, but it has low relative efficiency of 37%. Rousseeuw (1985) also proposed another robust estimator which is called the least trimmed squares (LTS) estimator. It is minimizing the sum of trimmed squared error, as $\left(\min \sum_{i=1}^{h} r_{(i)}^2\right)$. Although the LTS achieves a high breakdown point of 50% but its relative efficiency is so low, it's about 8% respectively (Rousseeuw & Croux, 1993; Stromberg et al., 2000).

Lee et al. (2007), introduced M-estimator Huber function to estimate the process mean and process standard deviation, then optimize the optimum settings using LT method. Meanwhile, Bashiri and Moslemi (2013), introduced M-estimator Huber function for multi-response surfaces. However, their approach did not produce good results due to using M-estimator Huber function which is known to have low breakdown point. Yohai, (1987) developed a MM-estimation method that has a very high efficiency and a higher breakdown point which is quite resistant to the outlier points. As the name suggested, M-estimation procedures are used twice to calculate the final estimates. The first part comes from computing its initial estimates by S-estimator. Since S-estimator minimizes a robust M-estimates of residual scale, this is where the first 'M' comes from. The second 'M' comes from the second part whereby M-estimates of the regression coefficient are computed based on the initial estimates. The MM-estimators have high breakdown point property at 50% and asymptotically efficient (Simpson, 1995; Uraibi et al., 2009). This MM-estimator is widely used in order to estimate the parameter for regression model (Smucler et al., 2017; Van den Boogaart et al., 2021; Rabbi et al., 2022; Khedmati & Niaki, 2022). Midi et al. (2012) also introduced MM-estimator to estimate the parameters of second-order polynomial RSM model when outliers exist in experimental data. They used experimental data with one response variable and compared the performance of MM-estimator with the OLS estimator. They concluded that when the MM-estimator is incorporated in the LT optimization method, it gives better result when compared to LT method using OLS estimator. Since MM-estimator when combined with the existing optimization methods produce better optimum process mean response, hence, it is a need to integrate the MM-estimator with several developed methods in this thesis, to improve the efficiency of the estimated process mean response.

2.4 Identifications of Outliers

It is important to point out that in real life situation, data sets may contain some contamination and heavy-tailed distribution (Hampel et al., 1986). In regression, observations that correspond to residuals that exhibit an unusual pattern are typically noted as outliers. A few useful procedures can be employed to detect outliers. Statistical literature has suggested various identification processes for outliers. To identify outliers, residual plots that have their basis on various residual types have been suggested (Ryan, 1997; Chatterjee & Hadi, 2006; Montgomery et al., 2021). Chatterjee and Hadi (2006) stated that residual patterns are frequently more informative compared to their magnitudes. Graphical displays of residuals are hence much more beneficial compared to formal test procedures. However, an extensive search is still taking place (Barnett & Lewis, 1994, Rashid et al., 2022; Zahariah & Midi, 2022; Baba et al., 2022).

A good amount of analytical detection methods for outliers are available in the literature, such as standardised residuals which is represented by $d_i = e_i/\hat{\sigma}$, $i = 1,2,\ldots,n$, where e_i refers to the OLS residual for the ith case as well as the $\hat{\sigma}^2 = \frac{1}{n-p}\sum_{i=1}^{n} e_i^2$. Rocke and Woodruff (1996) noted that any observations in which their standardized residuals fall outside (-3,3) is a potential outlier. Srikantan (1961) developed studentized residual technique for the identification of outliers represented by $d_i = e_i/\hat{\sigma}\sqrt{1-h_{ii}}$, $i = 1,2,\ldots,n$, where e_i refers to the OLS residual for the ith case as well as the $\hat{\sigma}^2 = \frac{1}{n-p}\sum_{i=1}^{n} e_i^2$ and h_{ii} represents the diagonal elements for the hat matrix H. Ellenberg (1976) proposed another detection method called deletion studentized (also known as R-Student) to identify outliers. The ith deletion studentized residual is defined as,

$$t_i = \frac{y_i - x_i'\hat{\beta}_{(i)}}{\hat{\sigma}_{(i)}\sqrt{1-\mathrm{h_{ii}}}}, \qquad i = 1,2,\ldots,n$$

where $\hat{\beta}_{(i)}$ and $\hat{\sigma}_{(i)}$ are the respective OLS parameter estimates and MSE when the ith observation is deleted, respectively. Unfortunately, all of these methods are not successful in detecting the genuine outliers since they are based on the OLS which is easily affected by outliers and the identification approach may be misleading.

To overcome this problem, robust methods were introduced to detect outlier. Huber (2005) utilised the standardised LTS residuals in determining outliers, i.e., $r_i = \frac{e_i}{\hat{\sigma}}, i = 1,2,\ldots,n$, where e_i represents the LTS residuals that correspond to the ith observation. Mustafa (2015), proposed an improvised method of identifying outliers by employing the highly efficient and high breakdown point robust MM-estimator in the formulation of the robust standardized residuals, which is referred as Standardized residuals based on MM denoted as d_{MM}. The standardized residuals based on MM are defined as follows:

$$d_{MM} = \frac{y_i - x_i'\hat{\beta}_{MM}}{\hat{\sigma}_{MM}} = \frac{e_i}{\hat{\sigma}_{MM}} \quad i = 1,2,\ldots,n$$

where, e_i is the residual from the MM for the ith case and $\hat{\sigma}_{MM}$ is the estimated value of σ calculated using MM-estimator.

With similar approach as d_{MM} standardized residual, Mustafa (2015) defined studentized residuals and deletion studentized residuals by employing the MM estimates, respectively as follows:

$$r_{MM\,i} = \frac{e_i}{\hat{\sigma}_{MM}\sqrt{1 - h_{ii}}} \quad i = 1,2,\ldots,n$$

and

$$t_{MM\,i} = \frac{e_i}{\hat{\sigma}_{MM(i)}\sqrt{1 - h_{ii}}} \quad i = 1,2,\ldots,n$$

where $\hat{\sigma}_{MM(i)}$ is the estimated of σ based on MM-estimator without the ith observation. As a rule of thumb, the ith observation is considered as an outlier if $|r_{MM\,i}|$ or $|t_{MM\,i}|$>2.5 or 3.

However, the d_{MM}, r_{MM}, and t_{MM} are not be able to correctly identify multiple outliers. Imon (2005) proposed generalized studentized residuals ($\mathrm{G}t_i$), which is based on group deletion procedure to identify multiple outlying observations in multiple linear regression.

In response surface model, unfortunately, little work has been explored in the identification of outlier. Goupy (1996) proposed to use the normal probability plot to detect a single outlier of the designed analysis of experiment data. Nevertheless, he pointed out that using graphical method is not very successful in detecting small outliers. Myers et al. (2009) employed studentized residual, R-students, and Cook's distance based on OLS estimator to indicate any problem with outliers in response surface model. Hund et al. (2002), suggested the half-normal probability plot and normal probability plot based on OLS residuals in order to examine multiple outliers in the experimental design data. However, such plots also not able to determine genuine outliers. Fang and He (2010) showed the effects of outlying observations on the estimates of regression coefficients and its standard error for different types of central composite design (CCD). Bashiri and Moslemi (2013) used studentized residual to detect outliers in Box-Behnken design. Rheem and Oh (2019) pointed out that, in RSM, central composite design (CCD) with multiple runs at the center points is frequently used. Usually, outliers at the center point ruin statistical analysis. Therefore, they proposed to conduct data screening at the center point using 3D scatterplots. If the outliers are observed at the center point, examine them. If they are due to errors in measuring or typing, correct them, otherwise, delete them. Unfortunately, all of these methods are not very successful in detecting multiple outliers in response surface model (Rheem & Oh., 2019).

The $\mathrm{G}t_i$ is very successful in detecting multiple outliers in linear regression model. Following Imon, Mustafa (2015) adapted the $\mathrm{G}t_i$ in the formulation of a new measures for identifying multiple outliers in the response surface model. They formulated $\mathrm{MGRS}t_i$ by utilizing the MM-estimator in the first step of the modified generalized studentized residuals. It has been shown that the $\mathrm{MGRS}t_i$ is very successful in detecting multiple outliers in the response surface model (Mustafa, 2015). Details explanation is given in Appendix A5.

2.5 Heteroscedasticity

One of the assumptions of linear regression model is that the variance of the error terms is constant. If this assumption failed, this implies that the error terms do not come from population with constant variance. The model in this case is said to be heteroscedasticity regression model. Many papers and books discussed the problem of heteroscedasticity (Kutner, et al., 2004; Rana et al., 2008; Imon, 2009; Alih & Choon, 2015; Montgomery et al., 2021). The heteroscedasticity is a common feature of data. Having violated this assumption, the OLS estimator is no longer the best linear unbiased estimator (BLUE). The OLS remains unbiased in the presence of heteroscedasticity. However, it lost efficiency (it does not have the minimum variance in the class of unbiased estimators). Heteroscedasticity also can give severe problems such as, significance tests will run either too high or too low and the standard errors of the estimates tend to be biased (Montgomery et al., 2021).

Some few sources of heteroscedasticity as stated by Greene (2008) are:

1. Incorrect data transformation: Heteroscedasticity usually occurs when wrong transformation technique is applied to the data set.
2. Skewness in the distribution: When one or more regressors are included in the model, heteroscedasticity is likely to occur.
3. Violation of the assumption of classical linear regression model i.e., when the regression model is not correctly specified.

2.5.1 Heteroscedasticity Diagnostic Methods

It is important to first identify whether or not the errors of the model is heteroscedastic, before taking any remedial measures. There are a number of specific tests for heteroscedasticity and they are classified into two ways: informal tests and formal tests. A useful first procedure is the informal way of examining the pattern of the residuals to see whether estimated variances differ from observation to observation. To do this, few plots are very useful such as Residuals-Predictors (R-P) plot, Residuals-Fits (R-F) plot, Squared Residuals-Predictor (SR-P) plot. The discussion of these plots can be seen in many standard books of regression and econometrics (Judge et al., 1985; Pindyck & Rubinfeld, 1997; Gujarati & Porter, 2002; Kutner et al., 2004; Chatterjee & Hadi, 2006; Greene, 2008, Montgomery et al., 2021).

The are many formal methods of detecting the presence of heteroscedasticity in the literature, such as Park test (Park, 1966), Glejser test (Glejser, 1969), Spearman's Rank Correlation test (Yule & Kendall, 1953; Gujarati, 2003), Goldfeld-Quant test (Goldfeld & Quandt, 1965), Breusch-Pagan (BP) test (Breusch & Pagan, 1979), Whites' General Heteroscedasticity test (White, 1980). These tests are established based on the different assumptions of heteroscedasticity.

White test and Breusch-Pagan test depend on the significance of an auxiliary regression in which the dependent variable and regressors with higher orders and cross-product are defined as explanatory variables in the auxiliary regression model. In these tests, some (or all) of the regressors are taken as independent variable and the OLS residual are taken as dependent variable. The Ramsey test also takes the least squares residuals as dependent variable, but, the fitted values and their higher order as regressors. The Park test regresses the natural log of the squared residuals against natural log of independent variable. Then, Glejser test, regressing the absolute values of the squared residuals with close related to the independent variable which is suspected is causing the heteroscedasticity behavior. The Spearman rank correlation test is formed based on the rank correlation between regressors and residuals. If replicative data is available, there are other useful tests like Bartlett test (Glaser, 1976), and Hartley test (Kutner, et al., 2004). However, if the heteroscedasticity pattern does not depend on the regressors, these tests are not able to detect the existence of heteroscedasticity. Similarly, in the presence of outlying observations these classical diagnostic measures fail to detect the existence of heteroscedasticity.

In the literature not much robust diagnostic technique for heteroscedasticity are available. Rana et al., (2008) suggested robust test procedure to detect heteroscedasticity which is unaffected in the presence of outliers. They modified one of the most popular and commonly used tests, the Goldfeld-Quandt (GQ) test, by replacing its nonrobust components by robust alternatives, and the methods are called Modified Goldfeld-Quandt test (MGQ). Nonetheless, the MGQ is not efficient since this method only usefull for one explanatory variable. Then, Alih and Choon (2015) proposed robust Goldfeld-Quandt (RGQ) test to remedy this problem.

2.5.2 Remedial Techniques of Heteroscedasticity in Dual Response Surface Model

In the presence of heteroscedasticity, the beta estimates are considered to be non-bias and consistent but the covariance matrix estimator which is based on the ordinary least squares (OLS) estimates tend to become bias and consistent (White, 1980; MacKinnon & Webb, 2015; Stock & Watson, 2006). In the literature, a few robust estimators have been proposed to deal with the presence of outliers and heteroscedasticity for multiple linear regression model such as; robust weighted least squares (RWLS) proposed by Habshah et al. (2009). In their technique they employed robust method twice, first in estimating the group variances and second in determining the weight. Rana et al. (2012) suggested using robust Wild Bootstrap for stabilizing variance to estimate the parameters of regression model. Rasheed et al. (2014) proposed a method based on M-estimation using iterative reweighted least squares (IRWLS) with Huber and Turkey bisquare weights functions to estimate model parameters. Habshah et al. (2014) proposed a new estimation technique named two-step robust weighted least squares (TSRWLS), in the presence of outliers and heteroscedasticity by applying robust procedures for regression model. Zafar and Aslam (2021) proposed adaptive weighted least square ratio estimator (AWLSRE) to cope with the issues of heteroscedasticity when dependent variable contains outlier. They extended the method of least squares ratio (LSR) given by Akbilgic and Akinci (2009). They noted that performance of AWLSRE is found to be more efficient as compared to the traditional methods available in the literature.

A number of researchers have proposed new methods as an alternative method to OLS in order to solve heteroscedasticity problem in dual response surface model. For example, Cho and Park (2005) used a weighted least squares (WLS) approach in dual response model application involving unbalanced data sets in which the number replicates between design points vary, which in turn can induce non-constant variance in the residuals. They suggested the use of a weighting scheme based on the number of replicates at a particular design point to overcome this shortfall, which proved to outperform the OLS approach typically applied in such situation. Cho and Park (2005) also used LT optimization in order to optimize optimal settings. Nonetheless, the LT optimization scheme is suspected not to be that efficient since it does not achieve the target with minimal bias in the estimated mean response due to some little bias introduced in its formulation. To overcome this shortcoming, Baba et al. (2022), proposed a new method, where they combined the Penalty Function (PM) with the WLS method. The WLS was employed to rectify the problems of heteroscedasticity for unbalanced data and the PM was used to offer substantial improvements over the LT method. However, the weakness of this method is that, it uses the PM optimization technique which is already noted in Section 2.2 to have several shortcomings.

Similarly, Goethals and Cho (2011) proposed the reweighted least squares (RLS) to handle the problems of heteroscedasticity in dual response surface model. Their work did not take into consideration when outliers are present in the data. There is strong evidence that the RLS is not reliable in this situation. Therefore, Boylan and Cho (2013) proposed using coefficient of variation to identify influential sources of variability between design points, then removed these sources and then applied optimal design theory to rebalance the experimental framework. The problem of this method is that when abnormal data or outliers are contained within heteroscedasticity condition its performance becomes less efficient.

Mustafa and Midi (2019) developed robust estimator in dual response models in order to provide more consistent and efficient estimator in order to solve the combined problems of outliers and heteroscedastic errors. The Two Stage Robust (TSR-MM) estimator of Mustafa and Midi (2019) incorporated robust MM-estimator in the algorithm of TSR-MM to rectify the problem of vertical outliers and heteroscedastic errors simultaneously. The first stage is to find weight for correcting heteroscedasticity errors and then using MM-estimator to estimate the parameters of the process mean and process standard deviation model. Once the fitted response functions for the process mean and process standard deviation have been constructed, the LT optimization method is employed to estimate the optimum factor settings of control variables. Although the TSR-MM can improve the parameter estimates of the model, it is not effective enough because no proper method is used to identify the exact number of vertical outliers in the datasets. Therefore, their total effect cannot be minimized. Moreover, the LT optimization scheme which has been mentioned to have several shortcomings (Section 2.2) were used. Thus, more efficient robust method needs to be proposed to overcome not only the problem of heteroscedasticity but also to minimize the total effects of outliers in the experimental data.

2.6 Desirability Function

The desirability function which was introduced by Harrington (1965) has been widely used methods for simultaneously optimize multiple responses (Subrahmanyam et al., 2018; da Silva et al., 2019; Bikbulatov & Stepanova., 2021; Jawad et al., 2022; Guedri & Msahli, 2022). The objective of desirability function approach is to assign an individual desirability function of a set of responses and at the same time chooses optimal factor settings that maximize the overall desirability function, which represents the most desirability response values. The desirability function approach consists of three stages that are, model building, transformation into individual desirability function and combination into an overall desirability function.

Suppose that each of the m responses $y_1, y_2, \dots, y_m$ can be modeled as a function of a vector of p given factor settings, $x = (x_1, x_2, \dots, x_p)$. Using second-order polynomial model, the OLS method is used to estimate the parameters of the model, yielding the predicted values of each response at the point $x = (x_1, x_2, \dots, x_p)$. The desirability function is used to transform each of the predicted responses $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_m$ to an individual desirability function, d_l', $0 \leq d_l \leq 1$, $l = 1,2, \dots, m$, where 0 representing an undesirable value of $\hat{y}_l$ and 1 representing a desirable value of $\hat{y}_l$. Then, the m individual desirability values, $d = (d_1, d_2, \dots, d_m)$ are combined into an overall desirability function of the predicted responses $\hat{y}_l$,denoted as D, where $0 \leq D \leq 1$.

Harrington (1965) proposed to transform each predicted responses, $\hat{y}_l$ to an individual desirability function, d_l using exponential transformation that can be shown as follow:

For a one-sided transformation,

$$d_l = \exp\big(-\exp(-\hat{y}_l)\big)$$

For a two-sided transformation,

$$d_l = \exp(-|\hat{y}_l|^r)$$

where r is a user-selected shape parameter. Extending the work of Harrington (1965), Gatza and McMillan (1972) modified the Harrington's transformation as follow:

$$d_l = \frac{\{\exp[\exp(-\hat{y}_l)] - \exp(-1)\}}{[1 - \exp(-1)]}$$

Derringer and Suich (1980) pointed out that, the transformations d_l of Harrington (1965) and Gatza and McMillan (1972) restricted to particular members of the exponential family. Several transformations method can also be found in the literatures (Fuller & Schere, 1998; Wu & Hamada, 2000; Pote et al., 2022). However, all of the transformation methods are quite complex and not appealing to practitioners because they are difficult to apply. Therefore, Derringer and Suich (1980) introduced a generalized transformation scheme for the predicted responses, $\hat{y}_l$ to desirability values. Their individual desirability transformation function varies according to the type of response, i.e., nomimal-the-best (NTB), smaller-the-best (STB), and larger-the-better (LTB) type, depending on whether the response has to be maximized, minimized, or attained a target value, respectively. For the LTB type, the individual desirability function is defined as

$$d_l = \begin{cases} 0 & \text{for} \quad \hat{y}_l \le y_l^{\min} \\ \left(\dfrac{\hat{y}_l - y_l^{\min}}{y_l^{\max} - y_l^{\min}}\right)^r & \text{for} \quad y_l^{\min} \le \hat{y}_l \le y_l^{\max} \\ 1 & \text{for} \quad \hat{y}_l \ge y_l^{\max} \end{cases} \tag{2.1}$$

where $y_l^{\min}$ and $y_l^{\max}$ are respectively, the lower and upper bounds of y_l, r is the shape parameter $(r > 0)$, and all user-selected. Similarly, the STB type, the individual desirability function is

$$d_l = \begin{cases} 1 & \text{for} \quad \hat{y}_l \le y_l^{\min} \\ \left(\dfrac{y_l^{\max} - \hat{y}_l}{y_l^{\max} - y_l^{\min}}\right)^r & \text{for} \quad y_l^{\min} \le \hat{y}_l \le y_l^{\max} \\ 0 & \text{for} \quad \hat{y}_l \ge y_l^{\max} \end{cases} \tag{2.2}$$

If the lth response is of the NTB type, the individual desirability function with the target T_l is given by

$$d_l = \begin{cases} \left(\dfrac{\hat{y}_l - y_l^{\min}}{T_l - y_l^{\min}}\right)^{r_1} & \text{for} \quad y_l^{\min} \le \hat{y}_l \le T_l \\ \left(\dfrac{y_l^{\max} - \hat{y}_l}{y_l^{\max} - T_l}\right)^{r_2} & \text{for} \quad T_l \le \hat{y}_l \le y_l^{\max} \\ 0 & \text{for} \quad y_l^{\min} < \hat{y}_l \text{ or } \hat{y}_l > y_l^{\max} \end{cases} \tag{2.3}$$

where $r_1 > 0$ and $r_2 > 0$ are the two shape parameters. After that, the individual desirability function, d_l is selected for each predicted response. Harrington (1965) proposed that all individual desirability function, d_l be converted to overall desirability function denoted as D which lies between 0 and 1 using geometric mean as follows:

$$D = (d_1 d_2 \dots d_m)^{1/m} \tag{2.4}$$

where D lies between 0 and 1. Subsequently, D is optimized to obtain the optimal factor settings. If the m responses are not equally important, Deringger (1994) proposed individual desirability function, d_l having weights for the m characteristics. The overall desirability function proposed by Derringer (1994) is as follows:

$$D_w = \left(d_1^{w_1} d_2^{w_2} \dots d_m^{w_m}\right) \tag{2.5}$$

where the w_l are user-selected weight based on the importance of each response variables and satisfy the interval [0,1] and $\sum w_l = 1$. The higher the value of D, the more desirable is the overall product. The value of D also has the property that if any $d_l = 0$ (that is, if one of the response variables is unacceptable) then $D = 0$ (that is, the overall product is unacceptable). This is generally not desirable. As such, when $D = 0$, it may cause inaccurate results in obtaining the optimal factor settings. Their work have motivated Khuri and Conlon (1981), Tabucanon (1988), Kim and Lin (2000), and Kim and Lin (2006) to propose an overall desirability function, D by employing arithmetic mean or weighted averages of the d_j's or adding some penalty functions to the d_l's.

The weakness of the overall desirability functions in Equation (2.4) and (2.5) is that it is based on geometric mean. Another weakness of the overall desirability function, D in Equation (2.4) is that the variability of each predicted responses, $\hat{y}_l$ is not taken into consideration by assuming that all observations have equal variation. Evidences from real examples, however, indicates that the equal variation assumption may not be valid in practice. In such a case, the existing multiple responses optimization approaches can be misleading. Various researchers attempted to solve simultaneous optimization of the predicted responses, $\hat{y}_l$ and the standard deviation (sd) of the predicted responses, $\text{sd}(\hat{y}_l)$ for multiple responses problem (Chiao & Hamada, 2001; Kim & Lin, 2006; Costa et al.,2011).

Chen et al. (2012) proposed an augmented approach to the desirability function (AADF), to simultaneously optimize the overall desirability function of the predicted responses, $\hat{y}_l$ and the overall desirability function of the standard deviation of the predicted responses, $\mathrm{sd}(\hat{y}_l)$ for multiple responses. Chen et al. (2012) suggested to use the transformation functions of Deringger and Suich (1980) which is based on the LTB, STB and NTB in order to obtain individual desirability function of the predicted responses, $\hat{y}_l$. Then, the overall desirability function of the predicted response, $\hat{y}_l$ is formulated as in Equation (2.4). However, in order to get the individual desirability function of the standard deviation of the predicted responses, $\mathrm{sd}(\hat{y}_l)$, Chen et al. (2012) proposed individual d_{s_l} defined as follows:

$$d_{s_l} = \begin{cases} \left(\dfrac{c_l - \sqrt{v_l(x)}}{c_j}\right)^r & \text{for } 0 < v_l(x) < c_l^2 \\ 0 & \text{for } v_l(x) \geq c_l^2 \end{cases} \tag{2.6}$$

where r is the shape parameter $(r > 0)$ (see Chen at al., 2012).

Chen et al. (2012) then defined the overall desirability function of the standard deviation of the predicted responses, $\mathrm{sd}(\hat{y}_l)$ denoted as S, is defined as follows:

$$S = \left(d_{s_1} d_{s_2} \ldots d_{s_m}\right)^{1/m} \tag{2.7}$$

where S lies between 0 and 1. Chen et al. (2012), pointed out that, the performance of the overall desirability function of the predicted responses, $\hat{y}_l$ and the overall desirability function of the standard deviation of the predicted responses, $\mathrm{sd}(\hat{y}_l)$ are not equally important. Therefore, Chen et al. (2012) suggested to combine the overall desirability function, D in Equation (2.4) and the overall desirability function of the standard deviation of the predicted responses, S in Equation (2.7) by assigning each function with different weights and called this method as augmented approach to the desirability function (AADF). They proposed to simultaneously optimize multiple responses after assigning weights to reflect the relative importance of the predicted mean response and the predicted standard deviation response, effects. The overall AADF, denoted as DS_λ desirability function of Chen et al. (2012) is then defined as follows:

$$DS_\lambda = (d_1 d_2 \ldots d_m)^{\lambda/m}\left(d_{s_1} d_{s_2} \ldots d_{s_m}\right)^{(1-\lambda)/m} \tag{2.8}$$

where λ are user-selected and satisfy interval [0,1]. The weakness of this augmented desirability function, DS_λ is that the transformation method of d_{s_l} in Equation (2.6) does not consider the STB type of Deringger and Suich (1980) which is already proven to be very desirable in terms of minimizing the standard deviation of the predicted responses. Moreover, the DS_λ desirability function approach of Chen et al. (2012) is not applicable for experimental data with replicates for each factor settings. For such experimental data, Chen et al. (2013) suggested using the AADF, DS_λ of Equation (2.8) with slight modification where STB type in Equation (2.3) is used to obtain individual desirability function of the standard deviation, d_l.

Marinkovic (2020) pointed out that, most desirability functions known in the literature are piecewise, non-differentiable function. Therefore, Marinkovic (2020) proposed a novel desirability function which is continuous and differentiable in its domain. This method is more suitable for applying some of the efficient gradient-based optimization methods. However, the disadvantage of this methods is that it does not take into account the variability of each predicted response, $\hat{y}_l$ and hence only focused on optimizing the overall desirability function for the predicted responses, $\hat{y}_l$ by assuming that the variability of the response is constant, which is not valid in practice (Chen et al., 2013).

Most of the recent work in the desirability function considers response surface methodology whereby the ordinary least squares (OLS) method is used to fit the second-order polynomial model. Moreover, the existing overall AADF method used geometric mean (Chen et al. 2012, 2013). Nevertheless, Chen et al. (2013) did not conduct a simulation study to assess their method's performance. Two real data sets were used to compare their method with the Harrington (1965) method and concluded that their method performed better. Unfortunately, when abnormal data or outliers contained in the data series, it may have significantly affected the OLS estimates. It is now evident that geometric mean is easily affected by outliers. Midi et al. (2013) proposed an augmented approach to the desirability function by integrating MM- estimator and geometric median in the formulation of the overall desirability function. They called this as method an augmented approach to the desirability function based on MM-estimator (AADF-MM). In order to rectify the problem of outliers in the data, Midi et al. (2013) extended the AADF of Chen et al. (2012). Instead of using geometric mean, they formulated geometric median for the overall desirability. However, it is not clearly written, how the overall desirability function of the AADF-MM is formulated. Hence, the AADF-MM is not appealing to practitioners since no proper steps shown in its establishment. Moreover, the AADF-MM can only be applied to experimental data with unreplicated responses for each factor setting. The AADF of Chen et al. (2013) is the widely used method to simultaneously optimize multiple responses in which each response variable has replicates for each factor setting. To the best of our knowledge, no work has been done to investigate the effect of outliers on the AADF of Chen et al. (2013). Thus, a new robust AADF for multiple responses with replicates for each factor setting is needed to rectify the problem of outliers in the data.

CHAPTER 3

PENALTY FUNCTION OPTIMIZATION IN DUAL RESPONSE BASED ON DECISION MAKER'S PREFERENCE

3.1 Introduction

The dual response surface methodology is a widely used technique in industrial engineering for simultaneously optimizing both the process mean and process standard deviation functions of the response variables. Many optimization techniques have been proposed to optimize the two fitted response surface functions that include the Lin and Tu (1995) optimization method which is based on mean squared error (MSE) objective function. However, it is not able to obtain an estimated mean response closer to the target value with small variation. As a solution to this problem, a penalty function (PM) method proposed by Baba et al. (2015) was put forward. The PM method has been shown to be more efficient than some existing methods. However, the drawback of the PM method is that it does not have a specific rule for determining the penalty constant, ξ. Thus, in practice, practitioners will find this method difficult since it depends on subjective judgments.

In this chapter, a new method called the penalty function based on decision maker's preferences structure in obtaining the penalty constant, ξ denoted as PFDMM is developed. The PFDMM is the extension work of PM optimization method, which is based on our newly developed method of determining penalty constant, ξ. At the outset, the value ξ is determined. Afterwards, it is incorporated in the algorithm of the proposed PFDMM optimization scheme. The performance of our proposed method is investigated by a Monte Carlo simulation study and real examples.

3.2 The Dual Response Surface Optimization

The dual response surface method aims to find the optimum factor settings condition of the controllable factors to diminish the variability and deviation from the desired target of the decision maker. Three basic strategies are considered to achieve such aim: experimental design, regression fitting and optimization aspect. Let's say an experiment was conducted to analyze the effect of some control variables for a given experimental design. Suppose the characteristics of interest, i.e., the response variable depends on a set of jth control variables $(j = 1,2, \dots, p)$, (coded) $x_{ij} = (x_{i1}, x_{i2}, \dots, x_{ip})$. Let y_{it} represents the response variable at the ith design point $(i = 1,2, \dots, n)$ and the tth replicate $(t = 1,2, \dots, r)$. Table 3.1 presents the summary of the observed data where $x, \bar{y}_i, s_i^2$ represent a set of controlled variables or factors, the sample mean and sample standard deviation of the response, respectively.

Table 3.1 : The presentation of the observed data for a given experimental design

Run	x				Replication				$\bar{y}_i$	s_i
1	x_{11}	x_{12}	$\cdots$	x_{1p}	y_{11}	y_{12}	$\cdots$	y_{1r}	$\bar{y}_1$	s_1
2	x_{21}	x_{22}	$\cdots$	x_{2p}	y_{21}	y_{22}	$\cdots$	y_{2r}	$\bar{y}_2$	s_2
.	.	.		.	.	.		.	.	.
.	.	.		.	.	.		.	.	.
.	.	.		.	.	.		.	.	.
i	x_{i1}	x_{i2}	$\cdots$	x_{ip}	y_{i1}	y_{i2}	$\cdots$	y_{ir}	$\bar{y}_i$	s_i
.	.	.		.	.	.		.	.	.
.	.	.		.	.	.		.	.	.
.	.	.		.	.	.		.	.	.
n	x_{n1}	x_{n2}	$\cdots$	x_{np}	y_{n1}	y_{n2}	$\cdots$	y_{nr}	$\bar{y}_n$	s_n

At each of the design point, the sample mean and sample standard deviation of the responses are usually calculated as in Equation (3.1).

$$\bar{y}_i = \frac{1}{r}\sum_{t=1}^{r} y_{it} \text{ and } s_i = \sqrt{\frac{1}{r-1}\left(\sum_{t=1}^{r} y_{it} - \bar{y}_i\right)^2} \tag{3.1}$$

Following Myers et al. (2009), the control variables are firstly translated from the natural units (original) into coded units using Equation (3.2).

$$x_{ij} = \frac{\delta_{ij} - \left[\max(\delta_{ij}) + \min(\delta_{ij})\right]/2}{\left[\max(\delta_{ij}) - \min(\delta_{ij})\right]/2} \tag{3.2}$$

$$i = 1,2,\ldots,n \,;\, j = 1,2,\ldots,p$$

where δ is the original value of x. Max (.) and min (.) refer to the maximum and minimum values of the original values of x, respectively.

Subsequently, the second-order polynomial model for the process mean and process standard deviation functions of the response variables are formulated (Montgomery et al., 2021). The ordinary least squares (OLS) method is usually used to estimate the parameters of the models. The fitted second-order polynomial model for the process mean and process standard deviation functions of the response variables at each design point, denoted as $\hat{\mu}(x)$ and $\hat{\sigma}(x)$ respectively, are written as follows:

$$\hat{\mu}(x) = \hat{\beta}_0 + \sum_{j=1}^{p} \hat{\beta}_j x_{ij} + \sum_{j=1}^{p} \hat{\beta}_{jj} x_{ij}^2 + \sum\sum_{j<k=2} \hat{\beta}_{jk} x_{ij} x_{ik} \tag{3.3}$$

and

$$\hat{\sigma}(x) = \hat{\gamma}_0 + \sum_{j=1}^{p} \hat{\gamma}_j x_{ij} + \sum_{j=1}^{p} \hat{\gamma}_{jj} x_{ij}^2 + \sum \sum_{j<k=2} \hat{\gamma}_{jk} x_{ij} x_{ik} \tag{3.4}$$

where $\hat{\beta}$ and $\hat{\gamma}$ are the estimates of the model parameters.

The control variables, x_{ij} in Equation (3.3) and Equation (3.4) are firstly translated from the natural units into coded units. Then an optimization method is used to simultaneously optimize both the process mean and process standard deviation functions of the Equation (3.3) and Equation (3.4) to obtain the estimated optimal factor settings. Then, the estimated mean and estimated standard deviation of the responses can be determined.

In this section, we briefly summarized some of the commonly used optimization methods. Prior to formulating the optimization scheme, the process mean and process standard deviation functions need to be established at the outset. The VM method simultaneously optimizes the objective function of the process standard deviation in which it restricts the process mean towards the desired target. Using the second-order polynomial model in Equation (3.3) and Equation (3.4), the VM method of optimization scheme is to

$$\text{minimize } \hat{\sigma}(x)$$

$$\text{subject to } \hat{\mu}(x) = \tau$$

where τ is the target value.

Lin and Tu (1995) proposed an optimization scheme based on squared-loss model which is known as mean square error (MSE). This model deals with square bias and minimizes the variance component of the factor settings. The Lin and Tu (LT) method is to

$$\text{minimize MSE} = (\hat{\mu}(x) - \tau)^2 + \left(\hat{\sigma}(x)\right)^2$$

Ding et al. (2004) proposed a natural extension of LT method by imposing relative weights on the bias and the variance terms. Their method which is called weighted MSE (WMSE) is to

$$\text{minimize WMSE} = w(\hat{\mu}(x) - \tau)^2 + (1 - w)\hat{\sigma}(x)$$

where w = relative weight factor, $(0 \leq w \leq 1)$. The data-driven approach is used to determine the value of w.

Baba et al. (2015,2022) proposed a penalty function-based approach as another alternative optimization scheme. The Penalty function (PM) method is to

$$\text{minimize}= \left(\frac{\xi}{2}\right) (\hat{\mu}(x) - \tau)^2 + \left(\hat{\sigma}(x)\right)^2 \tag{3.5}$$

where ξ is the penalty constant, $0 \leq \xi \leq \infty$. Nonetheless, this method concentrates more on the bias and forces the estimated process mean response to be close to the target value, which is not adequately efficient. Moreover, no specific penalty constant value, ξ was given to Equation (3.5). Hence, practitioners will find these approaches difficult since they must do trial and error subjective judgment.

3.3 The Proposed Penalty Function Optimization in Dual Response based on Decision Maker's Preference Method

In Chapter 2, we have reviewed some of the optimization schemes of the dual response surface methodology, and their weaknesses are noted. The penalty function-based approach (PM) seems to be better than the other methods but still has shortcoming whereby the determination of the penalty constant, ξ, is not clearly specified. Hence, we propose a new method which is called the penalty method based on decision maker's (PFDMM) preference method, where the penalty constant, ξ, is determined by considering decision maker's judgments. The proposed method is less complicated and can be implemented by practitioners.

Prior to the development of the method, the sample mean and sample standard deviation of Equation (3.1) are computed. Then, the fitted response models of Equation (3.3) and Equation (3.4) are estimated using the OLS estimator.

The new optimization approach (PFDMM) consists of two stages whereby in the first stage, penalty constant, ξ, is determined. Subsequently, in the second stage, Equation (3.7) is optimized to obtain the estimated optimal factor settings.

Stage 1: **Determining the penalty constant value,** ξ

Determining the value ξ is critical and challenging because of its large interval (Lee & Kim, 2012). The optimal penalty constant denoted as ξ^* is defined as the value that can provide the estimated mean with zero or reasonably closed to zero bias and at the same time minimizes the estimated standard deviation of the responses. To fix the idea, we illustrate how the value of penalty value, ξ is determined, by using, the roman catapult problem data. This data set was also used by Kim and Lin (1998) and Baba et al. (2015) to assess their proposed method. The aim of the experiment is to identify a combination of factor settings for which a variety of projectile types can be delivered to a target with minimal variability. The response variable is the distance of the projectile from the base of the catapult. The distance was measured in inches. The input variables are x_1 is the arm length (0.32-3.68 inches), x_2 is stop angle (30-90^0) and x_3 is the pivot height (2.5-5.5 inches). The target of distance of projectile from the base is set to $\tau = 80$. Figure 3.1

depicts the corresponding estimated mean and estimated variance of responses for Catapult study data, for some values of penalty constant, ξ, in the interval of 1-200, i.e., [1,200]. From the figure, it can be seen that the estimated process variance of the response increases as the ξ value increases. On the contrary, the estimated mean decreases when the ξ value increases. However, when ξ value reaches a certain point, the performance of both estimated responses becomes stable. In such circumstances, it is essential to diminish the unnecessary value of ξ.

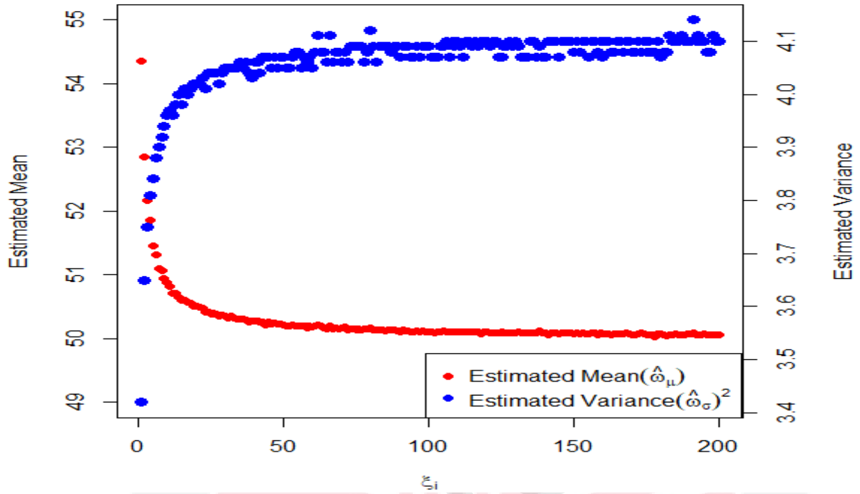


Figure 3.1 : The estimated mean and estimated variance versus penalty constant, (ξ)

Jeong et al. (2005) stated that in such case, it is necessary to narrow down the interval to make a substantial choice of ξ. Following the idea of Jeong et al. (2005), we firstly defined a vector $z^i = \left(z_1^i, z_2^i\right)^T$ where $z_1 = (\bar{y}_i - \tau)^2$ and $z_2 = s_i^2$. The penalty function (PF) denoted as PF_ξ for each ξ, $0 \leq \xi \leq \infty$ can be expressed as follows:

$$\text{PF}_{\xi_i} = \left(\frac{\xi}{2}\right) z_1^i + \left(z_2^i\right) \tag{3.6}$$

The idea behind this basic premise is that the ξ value should be congruent with the decision maker's (DM's) preference structure whereby the ranking of the z^igiven by the DM should be consistent with the ordering of the corresponding PF_ξ values. For example, if the DM prefers z^i to z^q (denoted as $z^i \succ z^q$), $\text{PF}_\xi^{z^i}$ should be less than $\text{PF}_\xi^{z^q}$. We employed the pairwise ranking scheme that compares only one pair of vectors at each time. The maximum number of pairwise rank is C_2^a. If the expression $\text{PF}_\xi^{z^i} < \text{PF}_\xi^{z^q}$ is true, then the possible interval of ξ is obtained, otherwise, it will not be considered as the interval of ξ. This procedure is repeated for all pairwise vectors. Next, the set of all interval $\xi_{i,q}^n$ that satisfies inequalities $\text{PF}_\xi^{z^i} < \text{PF}_\xi^{z^q}$ is obtained by intersecting the $\xi_{i,q}^n$ for all possible pairs of vectors.

Owing to this, we propose a systematic algorithm to find the optimal penalty constant ξ^*. Hence, the proposed algorithm of determining the penalty constant, ξ method can be summarized as follows:

Step 1: Firstly compute $\bar{y}_i$ and s_i as in Equation (3.1).

Step 2: Compute a vector $z^i = \left(z_1^i, z_2^i\right)^T$ where $z_1 = (\bar{y}_i - \tau)^2$ and $z_2 = s_i^2$.

Step 3: Calculate criteria value, i.e., $\mathrm{CV}_i = |\bar{y}_i - \tau| + s_i$. We suggest a simple algorithm of reducing the number of pairwise comparisons of z^i by computing $\triangle$, where $\triangle$ is an outlier-resistant estimator, i.e., median of CV_i. If CV_i value exceeds the $\triangle$ value, then vectors of $z^i = \left(z_1^i, z_2^i\right)^T$ will not be considered in the ranking of z^i.

To clarify the idea, the process mean, $\bar{y}_i$, process standard deviation, s_i, criteria values, CV_i along with the value z^i, are presented in Table 3.2 for Catapult data. In this example, the design points, $i = 1,3,6,8,9,10,11,12,13,14$ were removed from the analysis since their corresponding CV_i exceed the median $\mathrm{CV}_i = 27.72$, Consequently, only design points, $i = 2,4,5,7,15,16,17,18,19,20$ are remained. After that, the remaining design points were rank in ascending order, i.e., $z^1, \dots, z^{10}$ based on the value of CV_i. The number of pairwise comparisons to be made is reduced to $C_2^a = 45$, i.e., $\{(z^1, z^2), (z^1, z^3), (z^1, z^4), \dots, (z^9, z^{10})\}$, where

Table 3.2 : Summary of ranking process

i	$\bar{y}_i$	s_i	CV_i	$z^i = \left(z_1^i, z_2^i\right)^T$	Code
1	41.0	1.7	40.7	(1521,1.7)	
2	80.7	10.0	10.7	(0.49,10.0)	z^6
3	47.0	4.4	37.4	(1089,4.4)	
4	75.0	19.5	24.5	(25,19.5)	z^9
5	60.3	7.5	27.2	(388.09,7.5)	z^{10}
6	114.7	11.6	46.3	(1204.09,11.6)	
7	69.0	7.8	18.8	(121,7.8)	z^8
8	94.7	30.7	45.4	(216.09,30.7)	
9	56.7	4.9	28.2	(542.89,4.9)	
10	111.3	8.1	39.4	(979.69,8.1)	
11	50.0	7.0	37.0	(900,7.0)	
12	60.0	24.4	44.4	(400,24.4)	
13	54.7	5.0	30.3	(640.09,5.0)	
14	116.7	6.8	43.5	(1346.89,6.8)	
15	84.7	5.9	10.6	(22.09,5.9)	z^5
16	83.3	3.8	7.1	(10.89,3.8)	z^1
17	85.3	3.8	9.1	(28.09,3.8)	z^3
18	86.0	3.7	9.7	(36,3.7)	z^4
19	84.3	4.7	9.0	(18.49,4.7)	z^2
20	85.7	5.9	11.6	(32.49,5.9)	z^7
	Median CV_i		27.72		

Step 4: Conduct C_2^a pairwise comparisons. For each pair of (z^i, z^q), where $z^i \succ z^q$ obtain the values for $\mathrm{PF}_\xi^{z^i}$ and $\mathrm{PF}_\xi^{z^q}$ for some values of ξ in an interval, i.e., $0 \leq \xi \leq \infty$ that satisfy the condition $\mathrm{PF}_\xi^{z^i} < \mathrm{PF}_\xi^{z^q}$. Denote the interval of the corresponding ξ for each pair in term of sets, i.e., $S_1 \cap S_2 \cap \ldots \cap S_{C_2^a}$. Then the intersection of these sets, i.e., $S_1 \cap S_2 \cap \ldots \cap S_{C_2^a}$ will be considered as the final set of the optimal penalty constant, ξ.

Stage 2: Optimization stage

In this stage, optimization models are formulated and solved. This yield estimated optimum values of process mean and process standard deviation of the responses. As mentioned earlier, the new procedure introduced in this thesis is called a penalty method based on the decision maker's (PFDMM) preference which simultaneously optimizes the process mean and process standard deviation functions of the response variables. The PFDMM is defined to

$$\text{minimize} = \left(\frac{\xi^*}{2}\right)\,(\hat{\mu}(x) - \tau)^2 + \left(\hat{\sigma}(x)\right)^2 \tag{3.7}$$

where ξ^* is the optimal penalty constant determined in Step 4 of Stage 1. The optimal setting (x^*) with the corresponding ξ^* value within the interval that gives the smallest RMSE and bias will be considered as the final optimal settings. The $\hat{\mu}(x)$ and $\hat{\sigma}(x)$ from Equation (3.3) and Equation (3.4) are the fitted second-order polynomial model for the process mean and process standard deviation functions of the response variables at each design point based on OLS regression estimator.

The PFDMM in Equation (3.7) is optimized by using any unconstrained optimization method such as Newton's method, BFGS method, Conjugate gradient method, and Steepest ascent (descent) method. In addition, nonlinear optimization software can be used to find the optimal factor settings for dual response surface problem. In this study, the Rsolnp package in R software is utilized to analyze this problem. The resultant estimated optimal factor settings are then substituted in Equation (3.3) and Equation (3.4) to obtain the estimated optimum process mean and process standard deviation of the responses. We expect that the PFDMM could give an optimum response with the least bias and least variability, compared to other methods.

3.4 Monte Carlo Simulation Study

Monte Carlo simulation study is designed to assess the performance of our new proposed PFDMM method and to compare its performance with other methods namely the VM, LT, WMSE and PM. Following Baba et al. (2015), using 3^3 factorial design and at each control factor settings $x_{ij} = (x_{i1}, x_{i2}, x_{i3})$, $i = 1,2,\ldots,27$, five responses $(y_{i1}, y_{i2}, \ldots, y_{i5})$ are randomly generated from normal distribution with mean $\mu(x)$ and standard deviation $\sigma(x)$, as in Equation (3.8).

$$\mu(x) = 500 + (x_1 + x_2 + x_3)^2 + x_1 + x_2 + x_3 \qquad (3.8)$$
$$\sigma(x) = \sqrt{100 + (x_1 + x_2 + x_3)^2 + x_1 + x_2 + x_3}$$

where the desired target is assumed to be 500, i.e.,$\tau = 500$. The values 500 and 100 in Equation (3.8) were chosen so that the mean will be closer to the target value with certain variation. The five methods are then applied to each set of the generated data. In each simulation run, *REP*=1000 iterations is considered. The performance of the proposed method is evaluated based on bias, standard error (SE) and root mean squares errors (RMSE) of the estimated optimal mean response, $\hat{\mu}(x)$. The estimated mean and variance of the estimated optimal process mean response, computed over *REP* iterations are given by $\bar{\hat{\mu}} = \sum_{i=1}^{REP} (\hat{\mu}(x)/REP)$ and $\mathrm{var}(\hat{\mu}(x)) = \sum_{i=1}^{REP} (\hat{\mu}(x) - \bar{\hat{\mu}})/REP$, respectively. As per Midi (1999) and Riazoshams and Midi (2015) the bias is computed as bias $= \bar{\hat{\mu}} - 500$ and the root mean square error (RMSE) $= \sqrt{(\mathrm{bias})^2 + \mathrm{var}(\hat{\mu}(x))}$. The MSE consists of two components; one measures the variability (precision) and the other measures its bias (accuracy). A good method is one that has the least values of bias, standard error (SE) and RMSE.

Other factorial design settings such as 4^2 are also considered. To save space, we only present the results of 3^3 factorial design. However, the results of 4^2 are consistent. The results of Table 3.3 shows the estimated bias, SE and RMSE for the five optimization methods were calculated based on the OLS methods. It can be observed that our proposed PFDMM-optimization method outperforms the other four methods evident by having the smallest bias and RMSE.

Table 3.3 : Comparing estimated bias, SE and RMSE for five optimization methods

Methods		**1000 iterations**	
	bias	**SE**	**RMSE**
VM	1.302	4.412	4.600
LT	1.304	3.747	3.967
WMSE	1.472	3.783	4.060
PM	0.071	0.222	0.233
PFDMM	0.016	0.188	0.189

3.5 Numerical Examples

In this section, real-world experimental data with three responses is used to demonstrate the performance of the proposed technique compared to other multiple optimization methods.

3.5.1 The Catapult Study Data

The roman style catapult problem data is revisited (already described in Stage 1 of the PFDMM method) to show the merit of our proposed PFDMM method when compared to the other four methods. The central composite design (CCD) with three replicates was evaluated at each design point. Following Kim and Lin (1998), the target of interest for distance point is assumed to be 80, i.e., $\tau = 80$. The data is shown in the Appendix A.1.

Let us first illustrate how the penalty constant value, ξ is determined.

If all of the 20 alternatives vectors were to be compared, the total number of pairwise comparisons would be 190. In addition, the number of pairwise comparisons can be reduced by screening out the vectors in which their CVs exceed the median of CV. From Table 3.4, it can be observed that only 10 remaining vectors for data were considered out of 20 vectors. Pairwise comparisons among the vectors are conducted and it consists of C_2^{10} pairwise comparisons.

Figure 3.4 shows that the proposed method has established a set value of ξ as a result of intersection of the $I(i,q)$. The resultant intersection plot is determined by $I(3,5)$ and $I(6,7)$ which corresponds to the lower and upper limits of ξ, i.e., $\xi = 5$ and $\xi = 6$. Then the final set of optimal penalty constant ξ^* to be used is $\xi^* \epsilon [5,6]$. Subsequently, the optimal factor settings (x^*) for each function is obtained by minimizing the PFDMM in Equation (3.7) using the resultant ξ^* values. The optimal factor settings (x^*) with the corresponding ξ value within the interval that gives the smallest RMSE and bias will be considered as the final optimal settings.

Table 3.4 : Process mean and process standard deviation with the criteria value (CV) and ranking code for Catapult study data set

i	$\bar{y}_i$	s_i	CV_i	Code
1	41.0	1.7	40.7	
2	80.7	10.0	10.7	z^6
3	47.0	4.4	37.4	
4	75.0	19.5	24.5	z^9
5	60.3	7.5	27.2	z^{10}
6	114.7	11.6	46.3	
7	69.0	7.8	18.8	z^8
8	94.7	30.7	45.4	
9	56.7	4.9	28.2	
10	111.3	8.1	39.4	
11	50.0	7.0	37.0	
12	60.0	24.4	44.4	
13	54.7	5.0	30.3	
14	116.7	6.8	43.5	
15	84.7	5.9	10.6	z^5
16	83.3	3.8	7.1	z^1
17	85.3	3.8	9.1	z^3
18	86.0	3.7	9.7	z^4
19	84.3	4.7	9.0	z^2
20	85.7	5.9	11.6	z^7
	Median CV_i		**27.72**	

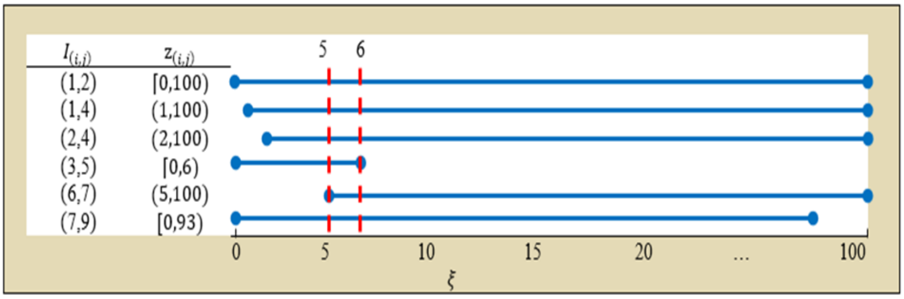


Figure 3.2 : The individual set of ξ for each model for Catapult study data set

The VM, LT, WMSE and PM methods are then applied to the data and their performances are compared with the PFDMM. The estimated optimal factor settings, estimated optimal mean, estimated standard deviation and RMSE are exhibited in Table 3.5. It is interesting to observe the results of our proposed PFDMM in Table 3.5. The PFDMM method shows a slightly better result when compared to other methods. It can be observed from Table 3.5 that the estimated mean optimal response of our proposed PFDMM is the closest to the target value i.e., 80 and with the least standard deviation and RMSE compared to other methods which is very desirable. The results of this data set agree reasonably well with the results of the simulation study. Here, we present the interpretation of the results, in particular for PFDMM method, after the coded units of the estimated factor settings are translated back to natural units using Equation (3.2). The results indicate that the optimum combination of factors where arm length at 2.0279 inches, stop angle at 48.771° and pivot height at 3.733 inches would give a maximum value of 79.8639 inches of the response, i.e., distance point where a projectile lands from the base of the catapult.

Table 3.5 : Estimated optimal factor settings, estimated mean, estimated standard deviation and RMSE for Catapult study data set

Method	Optimal settings $(\boldsymbol{x}^*)$	$\hat{\boldsymbol{\mu}}(\boldsymbol{x}^*)$	$\hat{\boldsymbol{\sigma}}(\boldsymbol{x}^*)$	RMSE
VM	(0.0444,-0.3093,-0.2394)	79.5818	3.0092	3.0381
LT	(0.0449,-0.3100,-0.2374)	79.6228	3.0143	3.0378
WMSE	(0.0440,-0.3087,-0.2414)	79.5409	3.0041	3.0389
PM	(0.0471,-0.3138,-0.2262)	79.8474	3.0023	3.0062
PFDMM	(0.0166,-0.3743,-0.1778)	79.8639	2.9985	3.0016

3.5.2 The Printing Process Study Data

The second real example in this thesis is the printing process study data. The experiment introduced by Box and Draper (1987) was conducted to determine the effect of the three control variables: x_1 (speed), x_2 (pressure), and x_3 (distance) on the characteristics of the printing process, which include the machine's index to apply colored inks to package labels. The experiment is a factorial design with three replicates at each of the 27 design points, assuming that the target of interest for distance point is $\tau = 500$. The experimental data is given in Appendix A.2.

Firstly, let us discuss how the penalty constant value, ξ is determined.

If all 27 alternative vectors were to be compared, the total number of pairwise comparisons would be 351. In addition, the number of pairwise comparisons can be reduced by screening out the vectors in which their CVs exceed the median of CV. From Table 3.6, it can be observed that there are only 13 remaining vectors of data. Thus, C_2^{13} pairwise comparisons among the vectors are considered. Figure 3.5 shows that the proposed method has established a set value of ξ as a result of intersection of the $I(i, q)$. The resultant intersection plot is determined by $I(3,4)$ which corresponds to the lower and upper limits of ξ, i.e., $\xi = 3$ and $\xi = 4$. Then the final set of optimal penalty constant ξ^* to be used is $\xi^* \epsilon [3,4]$.

Table 3.6 : Process mean and process standard deviation with the criteria value (CV) and ranking code for Printing study data set

$\boldsymbol{i}$	$\bar{\boldsymbol{y}}_{\boldsymbol{i}}$	$\boldsymbol{s}_{\boldsymbol{i}}$	**CV$_i$**	**Code**
1	24	12.49	488.5	
2	120.3	8.39	388.1	
3	213.7	42.8	329.1	
4	86	3.46	417.5	
5	136.7	80.41	443.7	
6	340.7	16.17	175.5	z^7
7	112.3	27.57	415.3	
8	256.3	4.62	248.3	z^9
9	271.7	23.63	251.9	z^{11}
10	81	0	419.0	
11	101.7	17.67	416.0	
12	357	32.91	175.9	z^8
13	171.3	15.01	343.7	
14	372	0	128.0	z^4
15	501.7	92.5	94.2	z^2
16	264	63.5	299.5	z^{13}
17	427	88.61	161.6	z^6
18	730.7	21.08	251.8	z^{10}
19	220.7	133.8	413.1	
20	239.7	23.46	283.8	z^{12}
21	422	18.52	96.5	z^3
22	199	29.45	330.5	
23	485.3	44.64	59.3	z^1
24	673.7	158.2	331.9	
25	176.7	55.51	378.8	
26	501	138.9	139.9	z^5
27	1010	142.5	652.5	
	Median CV$_i$		**329.1**	

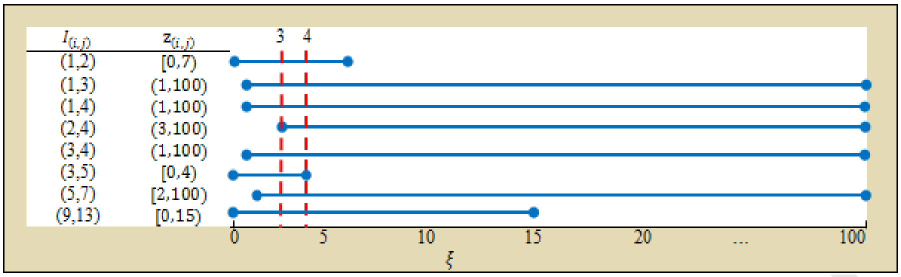


Figure 3.3 : The individual set of ξ for each model for Printing study data set

All five methods were used to determine the optimal factor setting, optimal mean, optimal standard deviation and RMSE, which are exhibited in Table 3.7 respectively. Similar to the results of Catapult data (Table 3.5), it can be observed from Table 3.7 that the PFDMM shows a slightly better result when compared to other methods for clean data. The results of the real examples were consistent with the results of the simulation study. It is very interesting to note that the results show the estimated mean optimal response of the PFDMM is very close to the target value, i.e., 500 with the least standard deviation and least RMSE which is very desirable.

Table 3.7 : Estimated optimal factor settings, estimated mean, estimated standard deviation and RMSE for Printing study data set

Method	Optimal settings (x^*)	$\hat{\mu}(x^*)$	$\hat{\sigma}(x^*)$	RMSE
VM	(1.00,0.06,-0.24)	494.65	43.40	43.73
LT	(1.00,0.07,-0.25)	494.69	43.46	43.78
WMSE	(1.00,0.08,-0.25)	496.44	43.52	43.67
PM	(1.00,0.12,-0.26)	500.00	42.75	42.75
PFDMM	(1.00,0.10,-0.25)	500.00	42.00	42.00

3.6 Conclusion

The dual response surface optimization is the prominent method in quality engineering whereby it simultaneously optimizes the fitted functions of the process mean and process standard deviation. To date, various optimization techniques have been thoroughly explored. Recently optimization scheme based on penalty function received considerable attention in literature. However, there is no proper procedure to determine the penalty constant, ξ as yet. Hence, we propose a systematic method to determine the penalty constant based on decision maker's (DM's) preference structure. The performances of the proposed method are beneficial and favourable for implementation by the practitioners. Moreover, the practical procedures were discussed to substantiate the effectiveness of the proposed method by comparing the four different optimization scheme approaches. Owing to the tabulated and discussed results, it is evident that the proposed method is more efficient and effective compared to the current existing methods as it has the least bias and RMSE.

CHAPTER 4

HIGH BREAKDOWN ESTIMATOR FOR DUAL RESPONSE OPTIMIZATION IN THE PRESENCE OF OUTLIERS

4.1 Introduction

In most dual response optimization methods, the sample mean and sample standard deviation of the responses often use non-outlier-resistant estimators. The ordinary least squares (OLS) method is also usually used to estimate the parameters of the process mean and process standard deviation functions. Nevertheless, many statisticians are not aware that the OLS procedure and the classical sample mean and sample standard deviation are easily influenced by outliers (Maronna et al., 2019; Rashid et al., 2021). The PFDMM method which is proposed in Chapter 3 is shown to be more efficient than the commonly used optimization method to simultaneously optimize the process mean and process standard deviation functions of the response variables. Nonetheless, the PFDMM is based on the OLS estimator. Hence, in the presence of outliers, the PFDMM will be much affected.

In this chapter, we improvised the proposed PFDMM method as discussed in Chapter 3. Alternatively, instead of using those classical methods, we propose using a high breakdown and highly efficient robust MM-mean, robust MM-standard deviation, and robust MM regression estimators to overcome these shortcomings. To reduce the effect of outliers, robust MM-estimators are incorporated in the algorithm of PFDMM to estimate the mean and the standard deviation of the responses as well as to obtain the fitted functions of the process mean and process standard deviation of the response. The proposed method is called the robust penalty function based on decision maker's preference using MM estimator, denoted as $\mathbf{PFDMM_R}$. The performance of the proposed method is compared to VM, LT, WMSE, PM and PFDMM methods. Simulation study and two real datasets, namely the Catapult study data and Printing process study data are used to assess the performance of the proposed $\mathbf{PFDMM_R}$ method.

4.2 MM-estimator

Prior to the development of the $\mathrm{PFDMM_R}$ method, the MM-estimator is used to estimate the sample mean and sample standard deviation of the responses and to obtain the fitted response models. The resultant MM-estimates are highly robust and less affected by outliers.

4.2.1 The Derivation of the MM-estimator for mean and standard deviation of the responses

Mean and standard deviation are popularly used measures of location/central tendency and measure spread of data, respectively. However, there is evidence that they may perform poorly when outliers occur in a data (Maronna et al., 2019). In other words, replacing only one of n observations with large value can negatively affect the value of the sample mean. Thus, robust MM-estimator which is highly robust and highly efficient (95%) with high breakdown point (50%) is employed to provide greater efficiency in the dual response surface estimation.

This section discusses the algorithm of the MM robust estimator as an alternative to the classical methods presented in Equation (3.1) for estimating the mean and standard deviation of the response values at each design point. Consider the following location-scale model

$$y_i = \mu + \sigma\varepsilon_i \tag{4.1}$$

where $y_1, y_2, \ldots, y_n$ be n responses and ε_i, $i = 1,2, \ldots, n$, are independent and identically distributed (i.i.d) observation with variance equals to 1. The following algorithm summarized the MM-estimator algorithm for estimating μ, and the scale σ.

Step 1: The initial consistent estimator of the location μ_0 and scale σ_0 were computed using high breakdown point S-estimator (Yohai, 1987).

Step 2: From Step (1), compute the residuals e_i and subsequently compute the M-estimate of standard deviation parameter (scale), $\hat{\sigma}$ where $\hat{\sigma}$ is the solution to

$$\frac{1}{n}\sum \rho_0\left(\frac{e_i}{\hat{\sigma}}\right) = 0.5$$

where ρ_o is an objective function that guarantee the M estimator to have high breakdown point, and $\psi = \rho'$ has to be redescending ρ function such as Hampel, Tukey's Biweight and Tanh functions. Tukey's Biweight function is employed in this thesis, and it is defined in Section 4.2.2.

Step 3: Compute the M estimate of $\hat{\mu}$ using ρ_1, where ρ_1 is an objective function that is chosen so that the M estimator achieves high efficiency. It has to be redescending ρ function. Yohai (1987) noted that for Tukey's Tukey's Biweight function, employing, $c_1 = 4.68$, results in high efficiency. $\hat{\mu}$ is a solution to

$$\sum \psi\left(\frac{e_i}{\hat{\sigma}}\right) x_i = 0 \tag{4.2}$$

where $\psi = \rho'(t/c_1)$, and $t = e_i/\hat{\sigma}$. Upon convergence, $\hat{\mu}$ and $\hat{\sigma}$ are the estimates for the mean (MM-mean) and standard deviation (MM-standard deviation) based on MM estimator. In the subsequent sections, they are referred as $\hat{\mu}_{\text{MM}}$ and $\hat{\sigma}_{\text{MM}}$ respectively.

4.2.2 The Derivation of MM-Estimator for the Process Mean and Process Standard Deviation Functions of Dual Response Surface Models

As already discussed in Chapter 3, the statistical models for the mean and standard deviation of the responses are established by fitting second order polynomial model, where the fitted functions are represented as in Equation (3.3) and Equation (3.4). The second-order polynomial model of Equation (3.3) can be written as follows:

$$\mu(x_i) = \beta_0 + \sum_{j=1}^{p} \beta_j\, x_{ij} + \sum_{j=1}^{p} \beta_{jj}\, x_{ij}^2 + \sum \sum_{j<k=2} \beta_{jk}\, x_{ij} x_{ik} + \varepsilon_i \qquad (4.3)$$

Let's $\mu(x_i)$ is written as y_i.

The true error ε_i, can be computed as, $\varepsilon_i = \left(y_i - E(y_i)\right)$. The true value of $\hat{\beta}$ can be estimated by using the OLS by minimizing the error sum of squares,

$$L_{\text{OLS}} = \sum_{i=1}^{n} \left(y_i - E(y_i)\right)^2 = \sum_{i=1}^{n} e_i^2 = e^t e = (y - X\beta)^T\,(y - X\beta) \qquad (4.4)$$

$$\frac{\partial L_{\text{OLS}}}{\partial \beta} = 0$$

The above partial derivatives were solved and yield the OLS estimates

$$\hat{\beta} = (X^T X)^{-1} X^T y \qquad (4.5)$$

$\hat{\beta}$ is known to be very sensitive to outliers. As an alternative, robust methods are put forward that can achieve a maximum of 50% breakdown point. We demonstrate the derivation of the MM-estimates for the second order polynomial model as follows:

Step 1: Compute an initial consistent estimator of β using high breakdown estimator preferably using high breakdown point S-estimator (Rousseeuw and Yohai, 1984). These estimators are the solution that finds the smallest possible dispersion or scatter of the errors.

$$\hat{\sigma} = \underset{\sigma}{\text{argmin}}\ \hat{\sigma}\left(r_1(\hat{\beta}), r_2(\hat{\beta}), \ldots, r_n(\hat{\beta})\right) \qquad (4.6)$$

where $r_1(\hat{\beta}), r_2(\hat{\beta}), \ldots, r_n(\hat{\beta})$ is the residual of the first, second and …*n*th observation, respectively and $\hat{\sigma}$ is the solution to

$$\frac{1}{n}\sum_{i=1}^{n}\rho\left(\frac{r_i}{\hat{\sigma}_n}\right)=b \qquad (4.7)$$

where b is a constant, defined as $b = E_{\Phi}[\rho(r)]$ and Φ denotes for the standard normal distribution.

The influence function ρ must satisfy the following condition:

i) ρ is symmetric and continuously differentiable and $\rho(0) = 0$.

ii) there exist $c > 0 \in \rho$ is strictly increasing on $[0, c]$ and constant on $[c, \infty]$.

From this condition, $\psi(t) = \rho'(t)$ where $t = r_i/\hat{\sigma}$ always be zero from certain value of t. So, ψ redescending Tukey's Biweight function is a good choice. A common choice of influence function, $\rho(t)$ is the Tukey's Biweight function (also called Bisquare function) proposed by Tukey (1967), and is given by

$$\rho(t)=\begin{cases}1-[1-(t/c)^2]^3 & if\ \ |t|\le c\\ 1 & if\ \ |t|>c\end{cases} \qquad (4.8)$$

Differentiating $\rho(t)$ gives $\psi(t)$ where $\psi(t)$ is given by

$$\psi(t)=\begin{cases}t\left[1-\left(\frac{t}{c}\right)^2\right]^2 & if\ \ |t|\le c\\ 0 & if\ \ |t|>c\end{cases} \qquad (4.9)$$

with derivative $\rho'(t) = 6\,\psi/c^2$ and the weight function use for estimating σ is denoted as $\omega(t)$ is defined as follows (Maronna et al., 2019).

$$\omega(t)=\begin{cases}\frac{\rho(t)}{t^2} & ,\quad t\ne 0\\ \rho''(0) & ,\quad t=0\end{cases} \qquad (4.10)$$

From Equation (4.9) $\rho''(t) = \psi'(t)$; by using product rule of derivative,

let $u = t$, and $u' = 1, v = \left[1-\left(\frac{t}{c}\right)^2\right]^2, v' = 2\left[1-\left(\frac{t}{c}\right)^2\right]\left[\frac{-2t}{c}\right]\left[\frac{1}{c}\right]$

$\psi'(t) = uv' + vu'$

$$= (t)\left[\frac{-4t}{c^2}\right]\left[1-\left(\frac{t}{c}\right)^2\right] + \left(1-\left(\frac{t}{c}\right)^2\right)^2 (1)$$

$$= \left[1 - \frac{t^2}{c^2}\right]\left[\frac{-5t^2 + c^2}{c^2}\right]$$

$$= \frac{1}{c^4}\left[-5c^2t^2 + c + 5t^4 - t^2c^2\right]$$

$$= \frac{1}{c^4}\left[-6c^2t^2 + c^4 + 5t^4\right]$$

$\psi'(0)$ or $\rho''(0) = \frac{c^4}{c^4} = 1$

Therefore, the weight function uses for estimating σ is denoted as $\omega(t)$ is defined as follows

$$\omega(t) = \begin{cases} \dfrac{\rho(t)}{t^2} & , \quad t \neq 0 \\ 1 & , \quad t = 0 \end{cases}$$

From Equation (4.7), $\hat{\sigma}$ is the solution to the following equation

$$\frac{1}{n}\sum_{i=1}^{n} \frac{\rho\left(\frac{r_i}{\hat{\sigma}}\right)}{\left(\frac{r_i}{\hat{\sigma}}\right)^2}\left(\frac{r_i}{\hat{\sigma}}\right)^2 = b \qquad (4.11)$$

$$\frac{1}{n}\sum_{i=1}^{n} \omega_i\, r_i^2 = b\hat{\sigma}^2$$

where $\rho\left(\frac{r_i}{\hat{\sigma}}\right)/\left(\frac{r_i}{\hat{\sigma}}\right)^2$ as defined in (4.11) is the weight for $t \neq 0$

Hence, can express $\hat{\sigma}_{k+1}^2$ as

$$\hat{\sigma}_{k+1}^2 = \frac{\sum_{i=1}^{n} \omega_i r_i^2}{nb} \qquad (4.12)$$

Solve equation 4.12 iteratively until convergence, i.e., if $\left|\frac{\hat{\sigma}_{k+1}^2}{\hat{\sigma}_k} - 1\right| < \varepsilon$, where ε is very small value, say 10^{-5}. At convergence, $\hat{\sigma}$, $\hat{\beta}$ are the S estimator for σ and β.

Step 2: Compute $r_i = y_i - \hat{y}_i$ based on initial S estimator to obtain the M-estimate of the standard deviation or scale. $\hat{\sigma}$ is the solution to $\frac{1}{n}\sum_{i=1}^{n} \rho_0\left(\frac{r_i}{\hat{\sigma}}\right) = b$, where $\rho\prime = \psi$ function has to be properly redescending, i.e. Hampel, Biweight and Tanh function. Here, again we use Biweight function.

In order to achieve MM estimator that guarantee high breakdown point equals 0.5

i) c_0 should be selected using $\rho_0(t/c_0)$.

ii) in the above step, redefined $b = E_\Phi\big(\rho(t/c_0)\big)$. Find c_0 iteratively, by numerical integration

$$E\big(\rho(t/c_0)\big) = 0.5$$
$$\int \rho(t/c_0)\frac{1}{\sqrt{2\pi}}e^{-(t/c_0)^2/2}\,dt = 0.5$$

For Biweight function in Equation (4.8)

$$E\big(\rho(t/c_0)\big) = \int_{-c_0}^{c_0} 1 - [1-(t/c)^2]^3 \frac{1}{\sqrt{2\pi}}e^{-(t/c_0)^2/2}dt + \int_{c_0}^{\infty}(1)\frac{1}{\sqrt{2\pi}}e^{-(t/c_0)^2/2}dt + \int_{-\infty}^{-c_0}\frac{1}{\sqrt{2\pi}}e^{-(t/c_0)^2/2}dt$$

At convergence $c_0 = 1.56$.

Step 3: Compute the M-estimate of $\hat{\beta}$ using ρ_1. To achieve MM with high efficiency, c_1 is chosen such as $\rho_1(t/c_1)$. In order $e_1 \leq e_0$, c_1 must be larger than c_0, $c_1 \geq c_0$. The larger the value of c_1, the higher the efficiency at normal distribution.

$\hat{\beta}_1$ is the solution to $\sum_{i=1}^{n} \psi\left(\frac{r_i}{\hat{\sigma}}\right) r_i$=0, which satisfy $s_1(\hat{\beta}_1) \leq s(\hat{\beta}_0)$, where $s(\hat{\beta}) = \sum_{i=1}^{n} \rho\,(r_i/\hat{\sigma})$

To find c_1 with high efficiency, say 95%, need to solve the following expression iteratively,

$$\text{Eff} = \left\{E_\Phi\left(\psi'\left(\frac{t}{c_1}\right)\right)\right\}^2 \Big/ \left\{E_\Phi\left(\psi'\left(\frac{t}{c_1}\right)\right)\right\}$$

$$0.95 = \frac{\left\{\frac{1}{\sqrt{2\pi}}\int \psi'\left(\frac{t}{c_1}\right) e^{-(t/c_1)^2/2}\,dt\right\}^2}{\left\{\frac{1}{\sqrt{2\pi}}\int \psi'\left(\frac{t}{c_1}\right) e^{-(t/c_1)^2/2}\,dt\right\}}$$

At convergence $c_1 = 4.68$.

To illustrate the M-estimator of step 3

Let x_i be p dimensional column vectors with co-ordinate $(x_{i1}, x_{i2}, \dots, x_{ip})$, $y_i = x_i'\beta$, $y = X\beta + \varepsilon$, y & ε are vectors with element y_i & ε_i.

$\hat{\beta}$ is an M-estimate of regression parameter model that minimizes the objective function ρ,

$$\hat{\beta}_M = \underset{\beta}{\operatorname{argmin}} \left[\sum \rho \left(\frac{y_i - x_i'\beta}{\sigma} \right) \right]$$

To minimize, take first partial derivative of ρ with respect to β for $j = 0,1,2, \dots p$ and equate to zero. This gives system of $p + 1$ equations.

$$\sum_{i=1}^{n} x_{ij}\, \psi \left(\frac{y_i - x_i'\hat{\beta}}{\hat{\sigma}} \right) = 0, \qquad j = 0,1,2, \dots, p \tag{4.13}$$

where $\psi = \rho'$, $x_{i0} = 1$ and x_{ij} is the ith observation on the jth predictor variable.

It is noted that.

$$\omega_i = \begin{cases} \dfrac{\psi(t)}{t} & , \quad t \neq 0 \\ \psi'(t) & , \quad t = 0 \end{cases}$$

For Biweight $\psi(.)$ Function in Equation (4.9), at 95% efficiency, weight, ω_i can be written as

$$\omega_i = \begin{cases} \left[1 - \left(\dfrac{t_i}{4.68} \right)^2 \right]^2 & if \;\; |t| \leq 4.68 \\ 0 & if \;\; |t| > 4.68 \end{cases}$$

By formulating weight, Equation (4.13) becomes

$$\frac{\sum x_{ij}\,\psi\left(\frac{y_i - x_i'\hat{\beta}^{(0)}}{\hat{\sigma}}\right)}{\left(\frac{y_i - x_i'\hat{\beta}^{(0)}}{\hat{\sigma}}\right)}\left(\frac{y_i - x_i'\hat{\beta}^{(0)}}{\hat{\sigma}}\right) = 0$$

This becomes

$$\sum_{i=1}^{n} x_{ij}\,\omega_i\left(y_i - \sum_{j=1}^{p} x_{ij}\hat{\beta}\right) = 0, \qquad j = 0,1,2,\dots,p \tag{4.14}$$

Solve Equation (4.14) iteratively by using iterative reweighted least squares (IRLS) with initial $\hat{\beta}^{(0)}$ and $\hat{\sigma}$.

In matrix notation Equation (4.14) can be written as

$$X'WX\hat{\beta} = X'Wy$$

$$\hat{\beta} = (X'WX)^{-1}X'Wy$$

Although MM-estimators have high breakdown point property as high as 50% and are asymptotically efficient, they may not perform especially well in the presence of high leverage points i.e., outlying observations in the X-coordinate (Simpson, 1995; Habshah et al, 2009; Uraibi, 2009).

4.3 Development of Robust Optimization Approach

The robust PFDMM technique is similar to the PFDMM in Chapter 3 with slight modification whereby the OLS in the algorithm of PFDMM is replaced with MM-estimator, we denote it as $\mathrm{PFDMM_R}$. It consists of two stages whereby in the first stage, penalty constant, ξ, is determined based on MM-estimator. Subsequently, in the second stage, Equation (4.17) is optimized to obtain the estimated optimal factor settings.

Stage 1: Determining the penalty constant value, ξ

Determining the value ξ is critical and challenging because of its large interval (Lee & Kim, 2012). The optimal penalty constant, denoted as ξ^* is defined as the value that can provide an estimated mean response with zero or reasonably close to zero bias and at the same time minimizes the estimated standard deviation of the responses.

Jeong et al. (2005) stated that in such case, it is necessary to narrow down the interval to make a substantial choice of ξ. Following the idea of Jeong et al. (2005), we firstly defined a vector $z^i = \left(z_1^i, z_2^i\right)^T$ where $z_1 = (\hat{\mu}_{MM} - \tau)^2$ and $z_2 = (\hat{\sigma}_{MM})^2$. The penalty function (PF) denoted as $\mathrm{PF}_{R\xi}$ for each ξ, $0 \le \xi \le \infty$ can be expressed as follows:

$$PF_{R\xi_i} = \left(\frac{\xi}{2}\right) z_1^i + \left(z_2^i\right)$$

The idea behind this basic premise is that the ξ value should be congruent with the DM's preference structure whereby the ranking of the z^i given by the DM should be consistent with the ordering of the corresponding $\mathrm{PF}_{R\xi}$ values. For example, if the DM prefers z^i to z^q (denoted as $z^i \succ z^q$), $\mathrm{PF}_{R\xi}^{z^i}$ should be less than $\mathrm{PF}_{R\xi}^{z^q}$. We employed the pairwise ranking scheme that compares only one pair of vectors at each time. The maximum number of pairwise rank is C_2^a. If the expression $\mathrm{PF}_{R\xi}^{z^i} < \mathrm{PF}_{R\xi}^{z^q}$ is true, then the possible interval of ξ is obtained, otherwise, it will not be considered as the interval of ξ. This procedure is repeated for all pairwise vectors. Next, the set of all interval $\xi_{i,q}^n$ that satisfies inequalities $\mathrm{PF}_{R\xi}^{z^i} < \mathrm{PF}_{R\xi}^{z^q}$ is obtained by intersecting the $\xi_{i,q}^n$ for all possible pairs of vectors.

With similar approach as the PFDMM method described in Chapter 3, we propose a systematic algorithm to find the optimal penalty constant, ξ^* for $\mathrm{PFDMM_R}$. Hence, the proposed algorithm of determining the penalty constant, ξ method can be summarized in two stages as follows:

Stage 1: Determining the penalty constant value, ξ based on MM-estimator.

Step 1: Firstly compute $\hat{\mu}_{MM}$ and $\hat{\sigma}_{MM}$ for each response variable.

Step 2: Compute a vector $z^i = \left(z_1^i, z_2^i\right)^T$ where $z_1 = (\hat{\mu}_{MM} - \tau)^2$ and $z_2 = (\hat{\sigma}_{MM})^2$.

Step 3: Calculate criteria value, i.e., $\mathrm{CV}_i = |\hat{\mu}_{MM} - \tau| + \hat{\sigma}_{MM}$. We suggest a simple algorithm of reducing the number of pairwise comparisons of z_i by computing Δ where Δ is an outlier-resistant estimator, i.e., median of CV_i. If CV_i value exceeds the Δ value, then vectors of $z^i = \left(z_1^i, z_2^i\right)^T$ will not be considered in the ranking of z^i.

Step 4: Conduct C_2^a pairwise comparisons. For each pair of (z^i, z^q), where $z^i \succ z^q$), obtain the values for $\mathrm{PF}_{R\xi}^{z^i}$ and $\mathrm{PF}_{R\xi}^{z^q}$ for some values of ξ in an interval, $0 \le \xi \le \infty$ that satisfy the condition $\mathrm{PF}_{R\xi}^{z^i} < \mathrm{PF}_{R\xi}^{z^q}$. Denote the interval of the corresponding ξ for each pair in term of sets, i.e $S_1 \cap S_2 \cap \ldots \cap S_{C_2^a}$. Then, the intersection of these sets, i.e. $S_1 \cap S_2 \cap \ldots \cap S_{C_2^a}$ will be considered as the final set of the optimal penalty constant, ξ.

Stage 2: **Optimization stage**

In this stage, optimization models are formulated and solved. These yields estimated optimum values of process mean and process standard deviation of the responses. As mentioned earlier, the new procedure introduced in this chapter is called the robust penalty method based on the decision maker's ($PFDMM_R$) preference which simultaneously optimizes the predicted responses of the process mean and process standard deviation functions.

The $\hat{\mu}_{MM}(x)$ and $\hat{\sigma}_{MM}(x)$ in Equation (4.15) and Equation (4.16) are the fitted values based on the MM-regression estimator.

$$\hat{\mu}_{MM}(x) = \hat{\beta}_0 + \sum_{j=1}^{p} \hat{\beta}_j x_{ij} + \sum_{j=1}^{p} \hat{\beta}_{jj} x_{ij}^2 + \sum \sum_{j<k=2} \hat{\beta}_{jk} x_{ij} x_{ik} \tag{4.15}$$

and

$$\hat{\sigma}_{MM}(x) = \hat{\gamma}_0 + \sum_{j=1}^{p} \hat{\gamma}_j x_{ij} + \sum_{j=1}^{p} \hat{\gamma}_{jj} x_{ij}^2 + \sum \sum_{j<k=2} \hat{\gamma}_{jk} x_{ij} x_{ik} \tag{4.16}$$

The $PFDMM_R$ is defined to

$$\text{minimize} = \left(\frac{\xi^*}{2}\right) (\hat{\mu}_{MM}(x) - \tau)^2 + \left(\hat{\sigma}_{MM}(x)\right)^2 \tag{4.17}$$

where $\xi^*$is the optimal penalty constant based on MM-estimator which is determined in Step 4 of Stage 1. It is also noted that the MM-estimator denoted as $\hat{\mu}_{MM}$ and $\hat{\sigma}_{MM}$ is used to estimate the mean and the standard deviation of the response variables.

4.4 Monte Carlo Simulation Results

In this study, the performances of five optimization methods, namely the classical VM, LT, WMSE, PM, PFDMM are compared to the $PFDMM_R$ using Monte Carlo simulation. Codes for the simulations are written using R programming. The robust package in R, i.e. lmrob, rrcov and Rsolnp are used to obtain the robust MM estimates of model parameters, process mean and process standard deviation, respectively. The same simulation design as in Chapter 3 is used. As per Park and Cho (2003), at each control factor settings $x_{ij} = (x_{i1}, x_{i2}, x_{i3})$, $i = 1,2,\ldots,27$, five responses $(y_{i1}, y_{i2}, \ldots, y_{i5})$ are randomly generated from normal distribution with mean $\mu(x)$ and standard deviation $\sigma(x)$, as follows:

$$\mu(x) = 500 + (x_1 + x_2 + x_3)^2 + x_1 + x_2 + x_3$$
$$\sigma(x) = \sqrt{100 + (x_1 + x_2 + x_3)^2 + x_1 + x_2 + x_3}$$

According to Park and Cho (2003), the values in the above equations were chosen so that the mean will be closer to the target value ($\tau = 500$) with certain variation. The performance of the proposed method when compared to other methods is evaluated under two different scenario, i.e., clean and contaminated data (data with outlier/s).

In order to see the effect of outliers on the different methods, three observations of the original y responses are randomly replaced with contaminated y values which are drawn from normal distribution, i.e., $N(150, 10^2)$. The five methods are then applied to each set of the generated data. All simulations experiments are conducted by performing 1000 Monte Carlo replications. The bias and standard error (SE) for each method is recorded together with its root mean square error (RMSE). A good method is one that has the lowest bias, SE, and RMSE. The bias, standard error (SE) and RMSE of the estimated optimal mean response for clean and contaminated data are exhibited in Table 4.1 and Table 4.2, respectively. Several interesting points emerge from these results.

First, the focus is on the results of Table 4.1 for the clean data. It can be observed from Table 4.1 that the biases and RMSE of VM, LT and WMSE are reasonably closed to each other. However, optimization based on PM, PFDMM and PFDMM_R methods outperformed the other three existing methods as evident by having smaller standard errors of estimates. This is due to the use of the classical methods to compute the estimates of mean and standard deviation and also the use of the OLS method for VM, LT and WMSE in obtaining the estimates of the predicted models in Equation 3.3 and Equation 3.4. However, our proposed PFDMM_R method is still the most efficient method among the six optimization schemes because it has the smallest value of RMSE, followed by the PFDMM, PM, LT, WMSE and VM methods.

Table 4.1 : Estimated bias, SE and RMSE of the optimal mean response for clean data

Methods	1000 iterations		
	bias	SE	RMSE
VM	1.302	4.412	4.600
LT	1.304	3.747	3.967
WMSE	1.472	3.783	4.060
PM	0.071	0.222	0.233
PFDMM	0.016	0.188	0.189
PFDMM_R	0.011	0.027	0.029

The bias, SE and RMSE of the estimated optimal mean responses for contaminated data are presented in Table 4.2. It is interesting to observe that the bias, SE and RMSE for all estimates of the existing methods have markedly increased. The results of our proposed PFDMM_R are very appealing. These attractive results are due to the usage of the robust MM estimator, which is resistant to outliers. Similar to clean data, the proposed PFDMM_R emerges to be conspicuously the most efficient method. It has the smallest bias, SE and RMSE among the six methods. Again, the performance of PFDMM comes second and is followed by the PM, LT, WMSE and VM methods. Our proposed PFDMM_R optimization method based on robust MM fitted response surface function has substantially improved the estimated optimal mean responses in terms of having the smallest bias, SE and RMSE. Smaller RMSE indicates that the combination of optimum factor settings result in the estimated mean response being closer to the target value with less bias and less variability, which is desirable.

Table 4.2 : Estimated bias, SE and RMSE of the optimal mean response for with outliers

Methods	1000 iterations		
	bias	SE	RMSE
VM	3.576	4.609	5.834
LT	2.193	1.760	2.811
WMSE	2.626	2.477	3.609
PM	0.930	0.031	0.931
PFDMM	0.564	0.128	0.579
$PFDMM_R$	0.005	0.017	0.018

4.5 Numerical Example

In this section, real world data is considered to evaluate the performance of the proposed PFDMM_R method. Their performances are also compared to the VM, LT, WMSE, PM and PFDMM method. The numerical examples are also taken from Kim and Lin (1998) and Baba et al. (2015) which is a Catapult study data and Box and Draper (1987) the Printing process study data. The original data together with the contaminated or modified ones are used to study the performances of the newly proposed methods.

Catapult study data

The Catapult study data set is revisited. This data set has been used in Chapter 3 to assess the performance of the PFDMM method. This data set was also used by Kim and Lin (1998) and Baba et al. (2015) by assuming that the target of interest for distance point is $\tau = 80$. In order to see the effect of an outlier on the estimated optimal mean response, we deliberately changed one data point, that is the $y_{3,20}$ observation corresponding to y value equals to 79 (in bold) with higher values of 790. The data set is shown in Appendix A1 (the control variables, x_i are coded) while in Table 4.3, the mean, standard deviation, MM-mean, MM-standard deviation and criteria value (CV) are presented. Figure 4.1(a) and Figure 4.1(b) show the boxplot of the original and modified data sets with one outlier. It can be observed from the box-plot that the distribution of the response variable of the original data set is fairly closed to normal distribution. However, the response variables for the modified data is skewed to the right with one outlier detected.

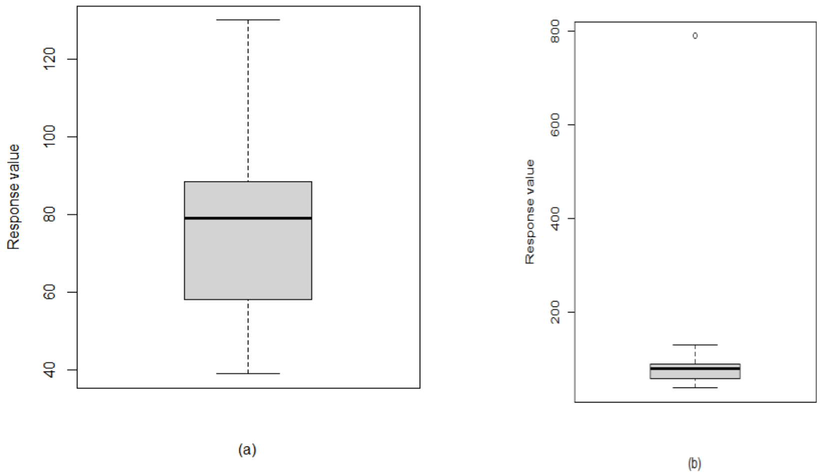


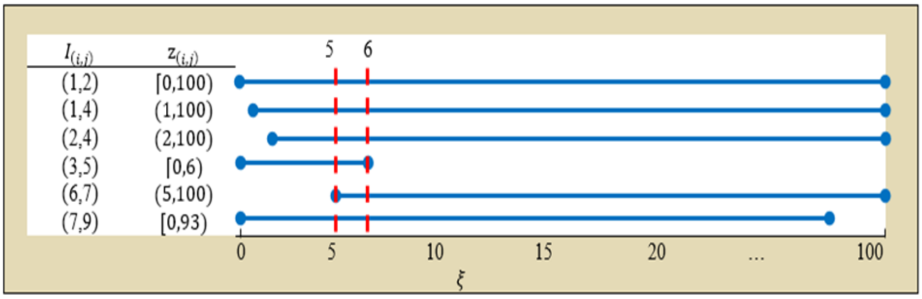
Figure 4.1 : Boxplot for (a) original data (b) modified (with outlier) for Catapult data

What is immediately clear from Table 4.3 is that the process mean and process standard deviation based on the sample mean and sample standard deviation are very sensitive to the contaminated data (shown in bold) when compared with the proposed estimation approach using MM-mean and MM-standard deviation. From Table 4.3, it can be observed that only 9 remaining vectors for data when using PFDMM optimization with OLS estimation method and 10 vectors for PFDMM_R optimization scheme using MM estimation method were considered out of 20 vectors. Pairwise comparisons among the vectors are conducted. For OLS method, it consists of C_2^9 while for MM method, C_2^{10} pairwise comparisons.

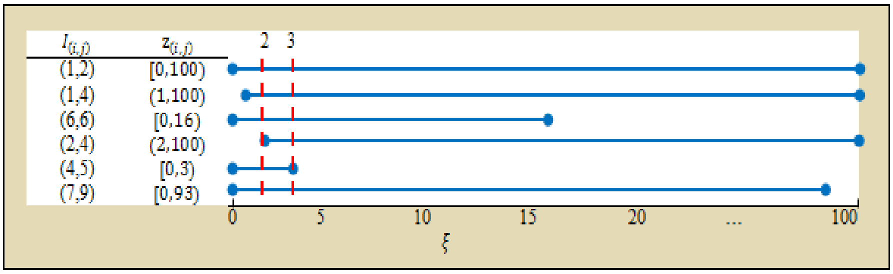
Table 4.3 : Process mean and process standard deviation with the criteria value (CV) and ranking code for Catapult study data set

i	$\bar{y}_i$	s_i	CV$_i$	Code PFDMM	MM$_l$	MM$_s$	CV$_i$	Code PFDMM$_R$
1	41.0	1.7	40.7		41.1	1.1	40.0	
2	80.7	10.0	10.7	z^6	80.5	8.0	8.5	z^5
3	47.0	4.4	37.4		44.5	1.1	36.6	
4	75.0	19.5	24.5	z^8	64.0	5.9	21.9	z^{10}
5	60.3	7.5	27.2	z^9	60.3	7.5	27.2	
6	114.7	11.6	46.3		114.4	6.3	40.7	
7	69.0	7.8	18.8	z^7	64.5	1.1	16.6	z^9
8	94.7	30.7	45.3		77.0	4.2	7.2	z^2
9	56.7	4.9	28.3		59.5	1.1	21.6	
10	111.3	8.1	39.5		116.0	2.1	38.1	
11	50.0	7.0	37.0		50.0	4.5	34.5	
12	60.0	24.4	44.4		46.0	5.1	39.1	
13	54.7	5.0	30.4		54.6	4.3	29.6	
14	116.7	6.8	43.5		117.8	3.2	41.0	
15	84.7	5.9	10.5	z^5	87.3	2.1	9.4	z^7
16	83.3	3.8	7.1	z^1	85.5	1.1	6.6	z^1
17	85.3	3.8	9.1	z^3	87.5	1.1	8.6	z^6
18	86.0	3.7	9.6	z^4	86.2	2.1	8.3	z^4
19	84.3	4.7	9.1	z^2	85.2	2.1	7.3	z^3
20	85.7 **(322.7)**	5.9 **(404.7)**	11.5 **(647)**		84.5 **(84.5)**	6.5 **(6.5)**	11.5 **(11.5)**	z^8
	Median CV$_{i,A}$		27.72 **(29.3)**		Median CV$_{i,B}$		21.6 **(21.6)**	

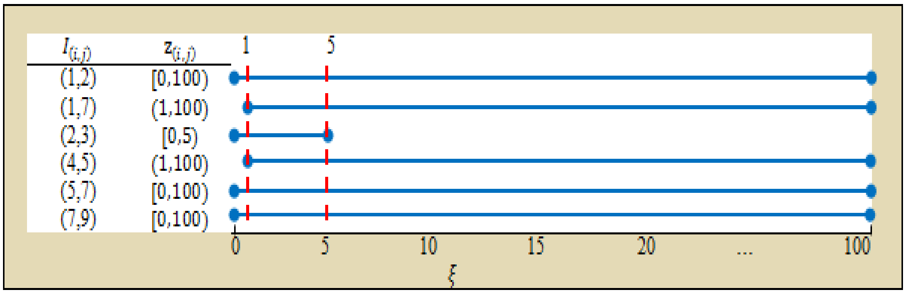
Figure 4.2 shows that the proposed method has established a set value of ξ as a result of intersection of the $I(i,q)$. Figure 4.2(a) shows that the resultant intersection PFDMM optimization with clean data has a set value of ξ as a result of intersection of the $I(i,q)$. The resultant intersection plot is determined by $I(3{,}5)$ and $I(6{,}7)$ which corresponds to the lower and upper limits of ξ, i.e., $\xi = 5$ and $\xi = 6$. Then the final set of optimal penalty constant ξ^* to be used is $\xi^*_{PFDMM} \epsilon [5{,}6]$. Second Figure 4.2(b), the resultant intersection when using the PFDMM optimization with contaminated data, the lower and upper limits are $\xi = 2$ and $\xi = 3$ which are determined by $I(2{,}4)$ and $I(4{,}5)$. Hence, the optimal penalty constant is $\xi^*_{PFDMM} \epsilon [2{,}3]$. Finally, Figure 4.2(c) shows resultant intersection the PFDMM$_R$ optimization method for clean and contaminated data, i.e., $\xi^*_{PFDMM_R} \epsilon [1{,}5]$. Subsequently, the optimal setting (x^*) when using MM estimation method is obtained by minimizing the PFDMM$_R$ in Equation (4.15) using the resultant ξ^* values. The optimal setting (x^*) with the corresponding ξ value within the interval that gives the smallest RMSE and bias will be considered as the final optimal settings.



(a) Individual set of ξ^*_{PFDMM} for clean data



(b) Individual set of ξ^*_{PFDMM} for contaminated data



(c) Individual set of $\xi^*_{PFDMM_R}$ for clean & contaminated data

Figure 4.2 : The individual set of ξ for Catapult study data set

The VM, LT, WMSE, PM, and PFDMM methods are then applied to the data and their performances are compared with the $\mathbf{PFDMM_R}$. The estimated optimal factor setting, estimated optimal mean response, estimated standard deviation and RMSE based on the OLS and MM estimators are exhibited in Table 4.4 and Table 4.5, respectively. It is interesting to observe the results of our proposed $\mathbf{PFDMM_R}$ method has the smallest standard deviations and RMSE values among the six methods. As displayed in Table 4.5 for contaminated data using MM estimation methods to obtain various estimates, our proposed $\mathbf{PFDMM_R}$ method has also outperformed the other methods evident by having the smallest value of standard deviation and RMSE. Moreover, the estimated mean optimal response of the $\mathbf{PFDMM_R}$ is very close to the target value, i.e., 80 with the least standard deviation, which is very desirable. The results of this data set agree reasonably well with the results of the simulation study. Here, we present the interpretation of the results, in particular for $\mathbf{PFDMM_R}$ method, as in Table 4.5 after the coded units of the estimated factor settings are translated back to natural units using Equation (3.2). The results indicate that the optimum combination of factors where arm length at 4.7182 inches, stop angle at 44.54 and pivot height at 3.023 inches would give a maximum value of 80.0588 inches of the response, i.e., distance point where a projectile lands from the base of the catapult.

Table 4.4 : Estimated optimal settings, estimated mean response, estimated standard deviation of response and RMSE for Catapult study data set (original clean data)

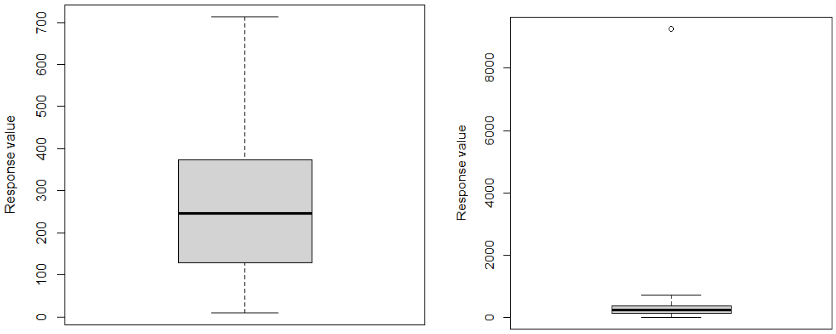
Method	Optimal settings (x^*)	$\hat{\omega}_\mu(x^*)$	$\hat{\omega}_\sigma(x^*)$	RMSE
VM	(0.0444,-0.3093,-0.2394)	79.5818	3.0092	3.0381
LT	(0.0449,-0.3100,-0.2374)	79.6228	3.0143	3.0378
WMSE	(0.0440,-0.3087,-0.2414)	79.5409	3.0041	3.0389
PM	(0.0471,-0.3138,-0.2262)	79.8474	3.0023	3.0062
PFDMM	(0.0166,-0.3743,-0.1778)	79.8639	2.9985	3.0016
$\mathbf{PFDMM_R}$	(1.6180,1.4180,0.6820)	80.0588	0.1072	0.1223

Table 4.5 : Estimated optimal settings, estimated mean response, estimated standard deviation of response and RMSE for Catapult study data set (with outlier)

Method	Optimal settings (x^*)	$\hat{\omega}_\mu(x^*)$	$\hat{\omega}_\sigma(x^*)$	RMSE
VM	(0.7948,0.0526,-1.4118)	78.8875	5.0207	5.1425
LT	(0.7959,0.0519,-1.4105)	78.9945	5.0420	5.1413
WMSE	(0.7932,0.0533,-1.4135)	78.7549	4.9942	5.1471
PM	(0.7959,0.0519,-1.4105)	78.9945	5.0420	5.1413
PFDMM	(0.5231,0.1060,-1.5008)	78.9826	4.8517	4.9572
$\mathbf{PFDMM_R}$	(1.6180,1.4180,0.6820)	80.0588	0.1072	0.1223

Printing process study data

The printing process data set is revisited. This data set has been used in Chapter 3 to assess the performance of the PFDMM method. This data set also used by Lin and Tu (1995) and Baba et al. (2015) by assuming that the target of distance point is $\tau = 500$. In order to see the effect of an outlier on the estimated mean optimal response, we deliberately changed one data point, that was the $y_{1,8}$ observation corresponding to y value equals to 259 (in bold) with higher values of 9259. Figure 4.3(a) and Figure 4.3(b) display the boxplot of the original and the modified data sets with an outlier. From the box-plot, it suggests that the response variable of the original data is reasonably close to normal distribution. On the other hand, the modified data of the response variables are skewed to the right with one outlier detected.



(a) (b)

Figure 4.3 : Boxplot for (a) original (no outlier) data set and (b) modified data set with outlier for Printing process data

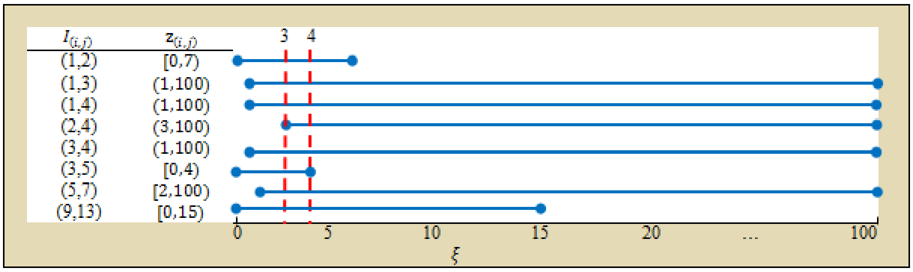
The data set is shown in Appendix A2 (the control variable, x_i are coded) while in Table 4.6, the mean, standard deviation, MM-mean, MM-standard deviation, and criteria value (CV) are presented.

If all 27 alternative vectors were to be compared, the total number of pairwise comparisons would be 351. In addition, the number of pairwise comparisons can be reduced by screening out the vectors in which their CVs exceed the median of CV. Similar to the simulation experiment and Catapult data example, we consider two scenarios, i.e., using clean and contaminated data. From Table 4.6, it can be observed that there are only 13 remaining vectors of data that include all two estimators. Thus, C_2^{13} pairwise comparisons among the vectors are considered.

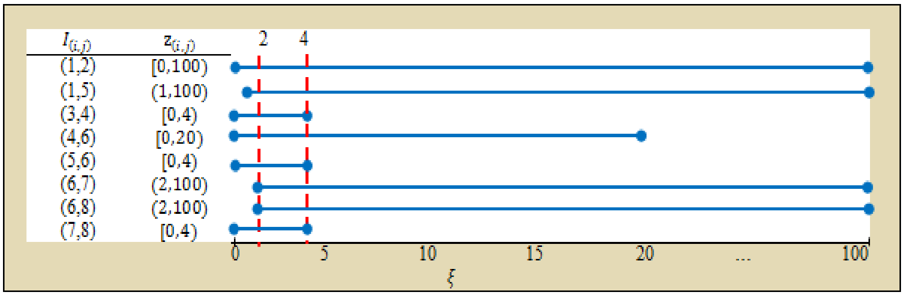
Table 4.6 : Process mean and process standard deviation with the criteria value (CV) and ranking code for Printing process data set

i	$\bar{y}_{i,A}$	$s_{i,A}$	$CV_{i,A}$	Code PFDMM	MM_l	MM_s	$CV_{i,B}$	Code $PFDMM_R$
1	24	12.49	488.5		26.05	5.80	479.6	
2	120.3	8.39	388.1		115.5	1.07	385.6	
3	213.7	42.8	329.1	z^{13}	189	5.13	316.1	
4	86	3.46	417.5		86.59	1.73	415.1	
5	136.7	80.41	443.7		183	6.62	323.6	
6	340.7	16.17	175.5	z^7	341.5	12.31	170.8	z^7
7	112.3	27.57	415.3		111.8	17.56	405.76	
8	256.3 **(3256.33)**	4.62 **(5198.5)**	248.3 **(7954.8)**		255 **(255)**	8.55 **(8.55)**	253.6 **(253.6)**	z^{12}
9	271.7	23.63	251.9	z^{10}	285	6.62	221.6	z^{10}
10	81	0	419.0		81.86	0	418.1	
11	101.7	17.67	416.0		91.5	3.20	411.7	
12	357	32.91	175.9	z^8	358.1	30.69	172.6	z^8
13	171.3	15.01	343.7		171.8	14	342.2	
14	372	0	128.0	z^4	371.7	0	128.3	z^5
15	501.7	92.5	94.2	z^2	554.5	10.89	65.4	z^2
16	264	63.5	299.5	z^{12}	300	10.26	210.3	z^9
17	427	88.61	161.6	z^6	427.4	85.88	158.5	z^6
18	730.7	21.08	251.8	z^9	729.4	12.82	242.2	z^{11}
19	220.7	133.8	413.1		149	106.84	457.8	
20	239.7	23.46	283.8	z^{11}	232.9	9.44	276.5	z^{13}
21	422	18.52	96.5	z^3	413	6.55	93.6	z^3
22	199	29.45	330.5		197.4	22.42	325.0	
23	485.3	44.64	59.3	z^1	511	5.92	16.9	z^1
24	673.7	158.2	331.9		668.3	112.18	280.5	
25	176.7	55.51	378.8		175.6	44.88	369.3	
26	501	138.9	139.9	z^5	421.5	23.94	102.4	z^4
27	1010	142.5	652.5		934.5	22.27	456.8	
	Median $CV_{i,A}$		329.1 **(330.5)**		Median $CV_{i,B}$		280.5 **(280.5)**	

Figure 4.4 shows that the proposed method has established a set value of ξ as a result of intersection of the $I(i,q)$. First, in Figure 4.4(a), the resultant intersection plot for PFDMM optimization with clean and contaminated data is determined by $I(3,4)$ which corresponds to the lower and upper limits of ξ, i.e., $\xi = 3$ and $\xi = 4$. Then the final set of optimal penalty constant ξ^* to be used for OLS estimator is $\xi^*_{PFDMM} \epsilon [3,4]$. Finally, Figure 4.4(b) shows the final penalty constant for $PFDMM_R$ optimization with clean and contaminated data estimator, i.e., $\xi^*_{PFDMM} \epsilon [2,4]$.



(a) Individual set of ξ^*_{PFDMM} for PFDMM clean and contaminated data



(b) Individual set of $\xi^*_{PFDMM_R}$ for $PFDMM_R$ clean and contaminated data

Figure 4.4 : The individual set of ξ for Printing process study data set.

All six methods were used to determine the optimal factor setting, optimal mean response, standard deviation and RMSE, which are exhibited in Table 4.7 and Table 4.8 for clean and contaminated datasets, respectively. Similar to the Catapult data, it can be observed from Table 4.7 and Table 4.8, that PFDMM$_R$ consistently provides the smallest standard deviation and RMSE in the absence and presence of outliers in the dataset. The results of the real examples were consistent with the results of the simulation study. It is very interesting to note that the results that show the estimated mean optimal response of the PFDMM$_R$ is very close to the target value, i.e., 500 with the least standard deviation, which is very desirable.

Table 4.7 : Estimated optimal settings, estimated mean response, estimated standard deviation of response and RMSE for Printing process study data set, (original clean data)

Method	Optimal settings (x^*)	$\hat{\omega}_\mu(x^*)$	$\hat{\omega}_\sigma(x^*)$	RMSE
VM	(1.00,0.06,-0.24)	494.65	43.40	43.73
LT	(1.00,0.07,-0.25)	494.69	43.46	43.78
WMSE	(1.00,0.08,-0.25)	496.44	43.52	43.67
PM	(1.00,0.12,-0.26)	500.00	42.75	42.75
PFDMM	(1.00,0.10,-0.25)	500.00	42.00	42.00
$PFDMM_R$	(0.60,0.40,-0.10)	500.00	38.81	38.81

Table 4.8 : Estimated optimal settings, estimated mean response, estimated standard deviation of response and RMSE for Printing process study data set data with outlier

Method	Optimal settings (x^*)	$\hat{\omega}_\mu(x^*)$	$\hat{\omega}_\sigma(x^*)$	RMSE
VM	(0.93,1.00,-1.00)	493.28	46.72	47.20
LT	(0.92,1.00,-1.00)	493.94	46.81	47.20
WMSE	(0.92,1.00,-1.00)	493.28	46.72	47.20
PM	(0.92,1.00,-1.00)	493.93	46.80	47.19
PFDMM	(0.91,1.00,-1.00)	493.96	46.55	46.94
$PFDMM_R$	(0.60,0.40,-0.10)	500.00	38.81	38.81

4.6 Conclusion

The main aim of this study is to develop a more reliable alternative scheme ($PFDMM_R$) that simultaneously optimizes both the mean and standard deviation functions of the response variables in the response surface model. Robust MM-mean, robust MM-standard deviation estimator, and MM regression estimator which are highly efficient and have high breakdown point were employed instead of using the classical non-robust methods. We also establish a systematic method to determine the penalty constant ξ which is based on the decision maker's (DM) preference structure and integrate this value in the $PFDMM_R$ optimization scheme. The existing optimization scheme is not effective enough as they fail to obtain an estimated mean response that is closer to the target value with small variation. Moreover, those methods are easily affected by outliers due to the use of non-robust methods to estimate the mean and the standard deviation of the response variables, and the non-robust least squares (OLS) method to estimate the parameters of the mean and standard deviation functions. The results obtained from the real data set and the Monte Carlo simulation study show that our proposed $PFDMM_R$ method is the most efficient method compared to VM, LT, WMSE, PM and PFDMM methods, in the presence and absence of outlier in the data set.

CHAPTER 5

NEW ROBUST PARAMETER ESTIMATION METHOD FOR DUAL RESPONSE SURFACE MODELS IN THE PRESENCE OF HETEROSCEDASTICITY AND OUTLIERS

5.1 Introduction

Much of dual response surface model over the last few decades has been devoted to determining the operating parameters that resulted in the best estimated mean response under homogeneity assumptions on the variances of the error terms. However, when the homoscedasticity assumption is violated, the OLS estimate is still unbiased but becomes inefficient due to the inconsistency of the variance-covariance matrix of the estimates. As a consequence, the inference will become unreliable. In the classical dual response surface analysis, to solve this problem, the reweighted least squares (RLS) method is employed to estimate the parameters of the dual response models. Indeed, the RLS approach is known to be very sensitive to vertical outliers which causes bias in the parameter estimates. There is strong evidence that the RLS is not reliable in this situation (Mustafa & Midi, 2019).

In this chapter, the RLS method is improved to tackle the simultaneous problem of heteroscedasticity and outliers. The improved method is called robust reweighted least squares (RRWLS) whereby it is established by integrating a weight function obtained from the proposed method of detection of outliers of Mustafa (2015). The proposed RRWLS method is integrated with five optimization algorithms, namely the VM, LT, WMSE, and PM including our proposed PFDMM method as discussed in Chapter 3. The performance of each method is evaluated by using simulation study and a real dataset namely the semiconductor manufacturing process data.

5.2 The Proposed Robust Reweighted Least Squares (RRWLS)

For linear regression model, the OLS estimates of β is $\hat{\beta} = (X'X)^{-1}X'y$ and the covariance matrix

$$\mathrm{cov}(\hat{\beta}) = (X'X)^{-1}X'\Omega\, X(X'X)^{-1} \tag{5.1}$$

where $E(\varepsilon\varepsilon') = \Omega$ is a positive definite matrix.

For homoscedastic errors, $\Omega = \sigma^2 I_n$ where I_n is identity matrix and it follows that the

$$\mathrm{cov}(\hat{\beta}) = \sigma^2 (X'X)^{-1}X'\mathrm{Z}\, X(X'X)^{-1} \tag{5.2}$$

where Z is a diagonal matrix.

It can be shown that the weighted least squares estimator is

$$\hat{\beta}_{\text{WLS}} = (X'W\,X)^{-1}X'W\,y \tag{5.3}$$

and

$$\text{cov}(\hat{\beta}_{\text{WLS}}) = \hat{\sigma}^2_{\text{WLS}}(X'WX)^{-1}$$

where W is defined as $W = Z^{-}$ where W is a diagonal matrix with diagonal elements or weights $w_1, w_2, \ldots, w_n$. The $\text{cov}(\hat{\beta}_{WLS})$ can be estimated by $\hat{\sigma}^2_{WLS}(X'WX)^{-1}$ where $\hat{\sigma}^2_{WLS} = \frac{\sum w_i \hat{\varepsilon}_i^2}{(n-p)}$.

It is important to note that when W matrix is known usually when the heteroscedastic error structures are known and under normality of errors, the WLS gives the best linear unbiased estimator. However, in real applications this situation almost never happens. As such, estimated weights are employed instead. It is generally believed that small variation in the weights due to estimation do not often affect a regression analysis or its interpretation much. However, outliers have an effect on the determination of weights.

The objective of this chapter is to develop a new parameter estimation method that can rectify the combined problem of heteroscedasticity and outliers. To achieve this goal, a reliable weight matrix, W needs to be determined when the heteroscedasticity error structure is unknown. The W matrix should perform well in the presence of heteroscedasticity and outliers. Hence two weight functions should be formulated whereby the first weight is used to reduce the effect of heteroscedasticity and the second weight function to reduce the effect of outliers.

A two stage Robust Reweighted Least Square (RRWLS) estimator is proposed in this regard. The RRWLS is an extension work of Mustafa and Midi (2019), who introduced two stage robust estimation (TSR-MM) method whereby in the first stage, they employed weight based on KNN (Kutner, Nachtsheim and Neter) estimator of Kutner et al. (2004). They employed MM estimator in the algorithm of KNN instead of the OLS method. Using this weight function, they used Equation (5.3) to estimate the parameters of the respond surface function.

The shortcoming of this method is that it does not used diagnostic method to detect outliers in order to reduce their effects. According to Dhhan et al. (2016) and Rashid et al. (2021), a weight function can be formulated based on the diagnostic method of detection of outliers. The effectiveness of the weight function depends on the efficient diagnostic method of detection of outliers. If potential outliers are not correctly identified, they will definitely affect the parameter estimation and other aspects of a weighted least squares analysis.

Therefore prior to developing a new parameter estimation method for rectifying heteroscedasticity and outliers, a good diagnostic method of identifying outliers needs to be formulated at the outset. We attempted to improve the diagnostic method of detections of outliers of Mustafa (2015) in response surface model $(\mathrm{MGRS}t_i)$ but the results are not much differed. Hence, we decided to formulate a weight function based on $\mathrm{MGRS}t_i$ to reduce the effect of outliers on the predicted responses of process mean and process standard deviation. Hence, in formulating RRWLS, two weight functions are established in which the initial weights function in the first step uses the KNN method based on MM-estimator to reduce the effect of heteroscedasticity and outliers and on the second step, a weight function is introduced to reduce the effect of remaining outliers. The final weight is obtained by multiplying both weights.

The first weight is based on the KNN estimator for correcting heteroscedasticity and outliers problem whereby the MM-estimator is used instead of the OLS estimator. The second weights were used to down weight the effects of the remaining outliers. The RRWLS is then applied to the fitted second-order polynomial model, $\hat{\mu}(x)$ and $\hat{\sigma}(x)$ for predicting the process mean and process standard deviation functions of the response variables at each design point, as already given in Equation (3.3) and Equation (3.4).

The algorithm of RRWLS consists of two stages summarized as follows:

Stage 1.

Step 1: Firstly, compute the process mean**,** $\hat{\mu}_{MM}$ and the process standard deviation, $\hat{\sigma}_{MM}$ using MM-estimators as discussed in Chapter 4.

Step 2: Fit the process mean model in Equation (3.3) and the process standard deviation model in Equation (3.4) by using MM-estimator to obtain $\hat{\mu}(x)$ and $\hat{\sigma}(x)$, Then the residual e_i is computed.

Step 3: Regress the absolute residuals, denoted as s_i where $s_i = |e_i|$ (obtained from Step 2) on $\hat{\mu}(x)$ and $\hat{\sigma}(x)$ then find the fitted values of $\hat{s}_i$ also by using MM-estimator.

Step 4: The square of the inverse fitted values would be the initial robust weights, i.e., $w_{1i} = 1/(\hat{s}_i)^2$.

Stage 2

Step 1: According to Mustafa (2015), any observation that corresponds to $\mathrm{MGRS}t_i$ which is larger than the cut-off points of $\mathrm{MGRS}t_i$ is considered as outliers. The cut-off point $\mathrm{CP}_{\mathrm{MGRS}t_i}$is given by,

$$\mathrm{CP}_{\mathrm{MGRS}ti} = \mathrm{median}(\mathrm{MGRS}t_i) + 3\mathrm{MAD}(\mathrm{MGRS}t_i)$$

where $\mathrm{MAD}(\mathrm{MGRS}t_i) = \mathrm{median}\{|\mathrm{MGRS}t_i - \mathrm{median}(\mathrm{MGRS}t_i)|\}/0.6745$

We have thoroughly studied the MGRSt_i method and have reproduced the results. The formulation of the MGRSt_i and its results are shown in Appendix A5.

Step 2: Furno (1996), Dhhan et al. (2016) and Rashid et al. (2021) formulated a weight function based on the method of identification of outliers together with their respective cut-off points. Following their ideas, the weight function based on MGRSt_i is formulated as follows:

$$w_{2i} = \min\left[1, \frac{CP_{\mathrm{MGRS}t_i}}{\mathrm{MGRS}t_i}\right]$$

where weight equal 1 is given to regular observations and weight equals to $\frac{CP_{MGRSt_i}}{MGRSt_i}$ is given to outliers. We multiply the weight w_{1i} with the weight w_{2i} to get the final weight w_i. The regression coefficients obtained from this RRWLS are the desired estimate of the heteroscedastic dual response model in the presence of outliers.

5.3 Monte Carlo Simulation Study

In this section, we evaluate and compare the performance of proposed RRWLS with the classical KNN (or simply call reweighted least square (RLS) and TSR-MM estimator of Mustafa and Midi (2019) by conducting Monte Carlo simulation study. As per Park and Cho (2003) simulation design, five responses are generated from each distribution with $\mu(x)$ and $\sigma(x)$ at each control factor settings $x_{il} = (x_{i1}, x_{i2}, x_{i3})$. Here, we consider two sets of design points, i.e., 27 and 45 design points. The total number of iterations is $REP = 1000$, each having 27 design points, and 45 design points with 135 responses and 225 responses, respectively. The $\mu(x)$ and $\sigma(x)$ are given in the following equations as per Park and Cho (2003) so that the mean will be closer to the target value ($\tau = 50$) with certain variations.

$$\mu(x) = 50 + 5(x_1^2 + x_2^2 + x_3^2)$$

$$\sigma(x) = \sqrt{100 + 5(x_1 - 0.5)^2 + x_2^2 + x_3^2}$$

Following Mustafa and Midi (2019), in order to induce heteroscedasticity of the error variances, we generated $\sigma^2(x)$ according to this relation, $\sigma^2(x) \times \exp(0.8x_1 + 0.8x_2^2 + 0.8x_3^3)$. To learn more about the impact of outliers, the data were contaminated by producing different contamination levels (2.2%, 3.1% and 5%) where 2.2%, 3.1% and 5% corresponds to 3,4, and 7 outliers for 27 design points, while 5,7, and 11 outliers for 45 design points, respectively. The contamination was created by inflating regular response Y with observation generated from $N(250, 10^2)$ distribution for all sample sizes.

The objective of the first simulation study is to show that the proposed RRWLS is the most efficient method when compared to the RLS and TRS-MM methods. Mean squared error (MSE) using RLS, TSR-MM and RRWLS methods were considered to evaluate the performance of the estimators. For each model, we simulated $REP = 1000$ data sets. Following Croux et al. (2003), we use the total mean squared error (MSE) as a metric criterion to assess the performance of the proposed method compared to the existing methods.

$$\text{MSE} = \frac{1}{REP}\sum_{i=1}^{REP}\|y_i - \hat{y}_i\|^2.$$

where $\|.\|$ indicates the Euclidean norm.

Tables 5.1-5.2 presents the result of the proposed RRWLS method and the existing methods (RLS and TSR-MM). The results show that when there is no outlier in the data, the MSE values of the RLS is the best compared to the TSR-MM and RRWLS. In this situation, the RLS has the smallest values of the MSE. These results are as expected because it is now evident that the RLS is unbiased and efficient estimator under normal errors. Nonetheless, when both outliers and heteroscedasticity come together in a data, the performance of the RLS becomes very poor. However, the RRWLS performs excellently followed by TSR-MM and RLS as evidenced by having the smallest MSE values.

Table 5.1 : The mean square error (MSE) of the estimated process mean

i	**% of Outliers**	**RLS**	**TRS-MM**	**RRWLS**
27	0	45.80	54.89	55.25
	2.2	125.66	77.33	76.56
	3.1	167.19	77.21	75.67
	5.0	305.32	77.36	73.54
45	0	50.58	67.47	62.57
	2.2	176.27	70.22	68.02
	3.1	233.39	70.52	68.62
	5.0	373.32	71.75	68.29

Table 5.2 : The mean square error (MSE) of the estimated process standard deviation

i	% of Outliers	RLS	TRS-MM	RRWLS
27	0	29.54	41.58	42.12
	2.2	274.54	65.35	62.45
	3.1	456.32	63.15	56.98
	5.0	854.91	63.38	57.33
45	0	32.23	40.27	38.53
	2.2	481.63	45.68	42.24
	3.1	752.37	46.53	44.87
	5.0	1327.38	48.72	47.62

The second objective of the simulation study is to show that the PFDMM optimization is more efficient than the VM, LT, WMSE and PM optimization methods when combined with the RLS, TSR-MM and RRWLS methods. The same simulation design as described before is used. The estimated mean of the optimal mean response, computed over m iterations are given by $\bar{\hat{\mu}} = \sum_{i=1}^{REP}(\hat{\mu}\,(x)/REP)$, Bias $=\bar{\hat{\mu}} - 50$, and $\text{var}\big(\hat{\mu}(x)\big) = \sum_{i=1}^{REP}(\,\hat{\mu}(x) - \bar{\hat{\mu}})^2/REP$. The mean squared error (MSE) is written as $\text{MSE} = (\text{Bias})^2 + \text{var}(\hat{\mu})$. As per Park and Cho (2003), the user-specified target value is assumed to be 50, i.e., $\tau = 50$.

Tables (5.3- 5.4) and Tables (5.5-5.6) present the values of bias, SE and RMSE when all the optimization methods (VM, LT WMSE, PM, PFDMM) are combined with the RLS, TSR-MM and RRWLS methods in simulated heteroscedastic errors without and in the presence of vertical outliers, respectively. To see the effect of outliers, the good data were then replaced by three outliers: the first observation of the second response variable, the 27th of observation of the third response variables, and the 14th observation of the fourth response variable.

First, let us focus on the results shown in Table 5.3 and Table 5.4 for the PFDMM and four other optimizations methods without vertical outliers. The results show that all the five optimization methods are fairly closed to each other. However, the PFDMM provides the best results because it has the smallest values of SE and RMSE. It is also observed that, the PFDMM-RRWLS based method is more efficient compared to PFDMM-RLS based and PFDMM-TSR-MM based methods which have the smallest values of bias, SE and RMSE.

Table 5.3 : Estimated bias, standard error (SE) and MSE for five optimizations method without outlier, 27 design points

	RLS			TSR-MM			RRWLS		
	Bias	SE	RMSE	Bias	SE	RMSE	Bias	SE	RMSE
VM	1.57	3.57	3.89	1.28	2.43	2.74	1.21	2.31	2.60
LT	1.49	3.26	3.58	1.15	2.38	2.64	1.10	2.35	2.59
WMSE	1.20	3.05	3.27	1.37	2.76	3.08	1.28	2.65	2.94
PM	1.17	2.63	2.87	1.05	1.72	2.02	1.14	1.64	1.99
PFDMM	0.92	2.03	2.23	1.20	1.46	1.89	0.98	1.37	1.68

Table 5.4 : Estimated bias, standard error (SE) and MSE for five optimizations method without outlier, 45 design points

	RLS			TSR-MM			RRWLS		
	Bias	**SE**	**RMSE**	**Bias**	**SE**	**RMSE**	**Bias**	**SE**	**RMSE**
VM	1.84	3.87	4.29	1.53	2.86	3.24	1.39	2.73	3.06
LT	1.57	3.54	3.87	1.37	2.59	2.93	1.28	2.31	2.64
WMSE	1.32	3.42	3.67	1.46	2.84	3.19	1.29	2.34	2.67
PM	1.28	2.86	3.13	1.34	2.46	2.80	1.20	2.23	2.53
PFDMM	1.10	2.52	2.75	1.15	2.37	2.72	0.95	2.16	2.47

It is interesting to observe the results from Table 5.5 to Table 5.6 when vertical outliers were introduced in the data. As expected, the presence of outliers changes situation drastically. The optimization method based on RLS estimates has been strongly affected by vertical outliers. We can see that all optimization methods with RLS based have higher values of bias, SE and RMSE compared with TSR-MM and RRWLS based methods. Nevertheless, it is interesting to observe that our proposed PFDMM-RRWLS based method is superior and more reliable compared to other methods in this study.

Table 5.5 : Estimated bias, standard error (SE) and MSE for five optimizations method with 2.2% of outlier, 27 design points

	RLS			TSR-MM			RRWLS		
	Bias	**SE**	**RMSE**	**Bias**	**SE**	**RMSE**	**Bias**	**SE**	**RMSE**
VM	2.29	4.84	5.35	1.58	2.67	3.10	1.41	2.42	2.80
LT	2.05	3.75	4.27	1.36	3.01	3.30	1.28	3.11	3.36
WMSE	2.58	4.07	4.82	1.84	3.29	3.76	1.88	3.16	3.68
PM	2.06	3.72	4.25	1.37	2.12	2.52	1.29	2.17	2.52
PFDMM	2.10	3.65	4.21	1.24	1.58	2.01	1.19	1.29	1.76

Table 5.6 : Estimated bias, standard error (SE) and MSE for five optimizations method with 2.2% of outlier, 45 design points

	RLS			TSR-MM			RRWLS		
	Bias	**SE**	**RMSE**	**Bias**	**SE**	**RMSE**	**Bias**	**SE**	**RMSE**
VM	3.65	5.23	6.38	1.78	2.71	3.24	1.49	2.63	3.02
LT	3.54	5.10	6.20	1.64	2.84	3.28	1.54	2.80	3.13
WMSE	3.26	6.21	7.01	1.76	2.96	3.44	1.53	2.71	3.11
PM	3.11	5.09	5.96	1.60	2.51	2.98	1.26	2.65	2.99
PFDMM	3.06	4.85	5.73	1.52	2.48	2.91	1.12	2.49	2.73

5.4 Numerical Example

In this section, a numerical example is presented to investigate the performance of the proposed method. The optimization of the semiconductor manufacturing process in 13 experiment runs was obtained from an experiment conducted by Shin and Cho (2005).

Two control factor or input variables are considered, namely the temperature of the molding stage (x_1) measured in degrees Fahrenheit, and injection flow rate (x_2) measured in pounds per second, to achieve the target response variable coating thickness of silicon wafers. The experiment is a factorial design with four replicates at each of the 13 design points, assuming that the target of interest for distance point is $\tau = 72.8$. The data for the mean, variance, and MM-estimators for the process mean and process standard deviation at each design point are shown in Appendix A3. In order to see the effect of an outlier on the estimated mean optimal response, we deliberately changed two data points, that were the $y_{13,2}$ and $y_{12,3}$ observations corresponding to y value equals to 70.6 (in bold) with higher values of 300 and 72.5 (in bold) with higher values of 200. To confirm the variability trends, the Breusch-Pagan hypothesis test is applied. In the second-order model, the hypothesis are $p = 5$ parameters. Since $11.520 > \chi^2(0.05,5) = 11.070$ we can conlude that H_1: that is, the error variance is not constant.

Table 5.7 and Table 5.8 show the estimated coefficients of the predicted process mean function, $\hat{\mu}(x)$ and the predicted process standard deviation function, $\hat{\sigma}(x)$ without and with outliers obtained using RLS, TSR-MM and RRWLS approaches. Table 5.7 displays the parameter estimates for the process mean and process standard deviation, in the situation when there is no outlier. In the absence of outlier, all estimators perform equally good as their estimates are relatively closed. As already mentioned, the RLS is more efficient than the other methods for data having heteroscedastic errors in the absence of outlier. Therefore, the RLS for clean data can be used as benchmark for the purpose of comparison. It can be observed from Table 5.8, that in the presence of one outlier, things alter substantially. The results show that outlier has an impact on the RLS method. As can be observed, not only the RLS estimated values of $\hat{\mu}(x)$ and $\hat{\sigma}(x)$ have significantly increased, but some of the estimates changed signed. For example, $\hat{\beta}_{11}$ has changed sign form 1.21 to -3.39. It is interesting to observe the results of RRWLS and TSR-MM estimates which are not much affected by the existence of outliers, since their estimates do not change much from the RLS estimates for clean data.

Table 5.7 : Estimated coefficients of the predicted process mean, $\hat{\mu}(x)$ and the predicted process standard deviation, $\hat{\sigma}(x)$ using RLS, TSR-MM and RRWLS methods without outlier

Coefficients	RLS		TSR-MM		RRWLS	
	$\hat{\mu}(x)$	$\hat{\sigma}(x)$	$\hat{\mu}(x)$	$\hat{\sigma}(x)$	$\hat{\mu}(x)$	$\hat{\sigma}(x)$
$\hat{\beta}_0$	71.86	1.57	71.87	1.58	71.77	2.07
$\hat{\beta}_1$	0.83	0.23	0.79	0.24	2.09	1.17
$\hat{\beta}_2$	-0.12	-0.24	-0.16	-0.24	-2.73	-0.14
$\hat{\beta}_{11}$	1.21	0.63	1.21	0.63	1.05	0.37
$\hat{\beta}_{22}$	1.59	1.21	1.58	1.19	0.11	1.64
$\hat{\beta}_{12}$	-1.91	-0.15	-1.91	-0.15	-0.15	0.58

Table 5.8 : Estimated coefficients of the predicted process mean, $\hat{\mu}(x)$ and the predicted process standard deviation, $\hat{\sigma}(x)$using RLS, TSR-MM and RRWLS methods with outlier

Coefficients	RLS		TSR-MM		RRWLS	
	$\hat{\mu}(x)$	$\hat{\sigma}(x)$	$\hat{\mu}(x)$	$\hat{\sigma}(x)$	$\hat{\mu}(x)$	$\hat{\sigma}(x)$
$\hat{\beta}_0$	81.17	21.09	72.31	1.52	71.79	2.10
$\hat{\beta}_1$	0.73	0.25	0.85	0.23	0.75	1.13
$\hat{\beta}_2$	-0.05	-0.28	-0.03	-0.23	-0.09	-0.12
$\hat{\beta}_{11}$	-3.39	-9.18	0.99	0.66	1.34	0.35
$\hat{\beta}_{22}$	-3.08	-8.49	1.36	1.23	1.83	1.57
$\hat{\beta}_{12}$	-1.77	-0.29	-1.92	-0.16	-2.27	0.68

Tables 5.9-5.10 present the results of the five optimization methods based on RLS, TSR-MM and RRWLS. It is observed that, the PFDMM-RRWLS based method is more efficient compared to PFDMM-RLS based and PFDMM-TSR-MM based methods evident by having the smallest values of bias, SE and RMSE. It is interesting to see the results from Table 5.10 when vertical outliers were introduced in the data. As expected, the presence of outliers changes situation drastically. The optimization method with RLS estimates have been strongly affected by vertical outliers. We can see that all optimization methods with RLS based have higher values of bias, SE and RMSE compared with TRS-MM and RRWLS based methods. Nevertheless, it is interesting to observe that our proposed PFDMM-RRWLS based method is superior and more reliable compared to the other two methods in this study.

Table 5.9 : Estimated optimum settings, Bias, SE and RMSE for the five optimization methods based on RLS, TSR-MM and RRWLS (without outlier)

Method	RLS				TSR-MM				RRWLS			
	Optimal Settings	Bias	SE	RMSE	Optimal Settings	Bias	SE	RMSE	Optimal Settings	Bias	SE	RMSE
VM	(0.086,-0.086,)	1.02	2.32	2.54	(-1.414,-0.474)	0.46	0.34	0.57	(-1.414,-0.242)	0.21	0.33	0.39
LT	(-0.214,0.186)	0.95	2.35	2.53	(-1.414,-0.224)	0.19	0.43	0.47	(-1.414,-0.5251)	0.15	0.40	0.18
WMSE	(-0.214,0.106)	1.01	2.31	2.52	(-1.414,-0.524)	0.14	0.47	0.49	(-1.414,-0.561)	0.11	0.33	0.35
PM	(-0.324,0.266)	0.83	2.21	2.36	(-1.414,-0.374)	0.12	0.33	0.35	(-1.414,-0.287)	0.15	0.28	0.32
PFDMM	(-0.634,0.346)	0.39	2.18	2.21	(-1.414,-0.414)	0.16	0.24	0.29	(-1.414,-0.1803)	0.12	0.20	0.23

Table 5.10 : Estimated optimum settings, Bias, SE and RMSE for the five optimization methods based on RLS, TSR-MM and RRWLS (with outlier)

Method	RLS				TSR-MM				RRWLS			
	Optimal Settings	Bias	SE	RMSE	Optimal Settings	Bias	SE	RMSE	Optimal Settings	Bias	SE	RMSE
VM	(-0.561,-0.205)	3.35	8.35	8.99	(-1.154,0.206)	0.78	1.04	1.30	(-1.153,0.242)	0.80	1.01	1.29
LT	(0.784,-0.221)	3.40	8.45	9.10	(-1.054,0.206)	0.94	1.08	1.43	(-1.153,0.256)	0.99	1.21	1.56
WMSE	(0.235,-0.023)	3.56	8.01	8.77	(-1.044,0.196)	0.83	1.05	1.33	(-1.152,0.188)	0.81	1.10	1.36
PM	(0.130,-0.167)	2.86	5.89	6.55	(-1.164,0.206)	0.64	1.05	1.22	(-1.164,0.215)	0.53	1.01	1.14
PFDMM	(-0.566,-0.205)	3.35	8.35	8.99	(-1.154,0.206)	0.60	1.03	1.19	(-1.154,0.301)	0.45	0.94	1.04

5.5 Conclusion

The main goal of this chapter was to establish a new optimization scheme in the dual response model for correcting heteroscedastic errors in the presence of outliers. In a heteroscedasticity data without outliers, it appears that the estimated optimum process mean response of the RLS, TSR-MM and the RRWLS based estimators work equally good. The RLS based estimator is an excellent strategy for dealing with heteroscedasticity, although it is susceptible to outliers. Hence, they are not reliable. The proposed RRWLS-based strategy in this chapter can solve both problem of heteroscedasticity and outliers at the same time. The results of the simulation experiment and real data show that the PFDMM-RRWLS based approach outperforms the other current methods in terms of handling outliers and heteroscedastic errors in the response surface models.

CHAPTER 6

IMPROVISED DESIRABILITY FUNCTION (IDF) OPTIMIZATION TECHNIQUE FOR MULTIPLE RESPONSES BASED MODIFIED GEOMETRIC MEAN (MGM) AND MM-ESTIMATOR IN THE PRESENCE OF OUTLIERS

6.1 Introduction

Harrington's desirability function has been widely used extensively to optimizing multiple responses in order to find optimal factor settings for the predicted responses, $\hat{y}_l$. This method assumes that the variability between responses is constant. However, this assumption doesn't hold in a real situation. As a solution to this problem, an augmented approach to the desirability function (AADF) proposed by Chen et al. (2012) is developed to simultaneously optimize multiple responses that consider variability between responses in its establishment. The AADF is developed by augmenting the overall desirability function of the predicted responses, $\hat{y}_l$ and the overall desirability function of the standard deviation of predicted responses, $\mathrm{sd}(\hat{y}_l)$ into a single overall desirability function. We observed that the AADF can be presented as geometric mean. Chen et al. (2013) extended the AADF method to optimize dual response surface model for multiple responses with replicates for each response variable. The AADF of Chen et al. (2013) is also observed to be based on geometric mean. Furthermore, the OLS method is used to estimate the parameters of the response functions for the process mean and process standard deviation. It is already mentioned that the OLS and the geometric mean can be adversely affected by outliers. Subsequently, the estimated process mean and estimated process standard deviation of the responses will give inaccurate results.

In this chapter, the robust AADF method is proposed, which we called a robust augmented approach to the desirability function denoted as (RAADF). The RAADF is the extension work of AADF of Chen et al. (2013), whereby the overall desirability function, DS_λ that combined the overall desirability function of the predicted responses, $\hat{y}_l$ and overall desirability function of the standard deviation of the predicted responses, $\mathrm{sd}(\hat{y}_l)$, is formulated based on geometric median and geometric MM. To the best of our knowledge, no research has been done to improve the performance of Chen et al. (2013) method when outliers are present in the experimental data with replicates for each factor setting. The performance of our proposed RAADF method is investigated by a Monte Carlo simulation study and real examples.

6.2 Classical Harrington's Desirability Function

The Harrington's desirability function was first introduced by Harrington (1965). It is one of the most widely used methods in optimizing multiple responses. Suppose that there are m responses, $y = (y_1, y_2, \dots, y_m)$ and for each response can be modelled as a function of a vector of p factor settings of input variable, $x = (x_1, x_2, \dots, x_p)$. Given n experimental runs or design points, a total of t replications at each design point are conducted for each response variable. The layout for this experiment design is shown in Table 6.1. Let y_{ilt} be the ith design point, the lth response variable and at the tth replicate at each design point, where $i = 1,2, \ldots, n$, $l = 1,2, \ldots, m$, and $t = 1,2, \ldots, r$. The sample mean and sample standard deviation of the lth response variable, is given in Equation (6.1).

$$\bar{y}_{il} = \frac{\sum_{t=1}^{r} y_{ilt}}{r} \qquad s_{il} = \sqrt{\frac{\sum_{t=1}^{r} (y_{ilt} - \bar{y}_{il})^2}{r-1}} \tag{6.1}$$

Table 6.1 : Experimental layout for multiple responses

Run	x			y_1			$\bar{y}_1$	s_1	…	y_m			$\bar{y}_m$	s_m
1	x_{11}	…	x_{1p}	y_{111}	…	y_{11r}	$\bar{y}_{11}$	s_{11}	…	y_{1m1}	…	y_{1mr}	$\bar{y}_{1m}$	s_{1m}
2	x_{21}	…	x_{2p}	y_{211}	…	y_{21r}	$\bar{y}_{21}$	s_{21}	…	y_{2m1}	…	y_{2mr}	$\bar{y}_{2m}$	s_{2m}
.	.		.	.		.	.	.	…	.		.	.	.
.	.		.	.		.	.	.	…	.		.	.	.
.	.		.	.		.	.	.	…	.		.	.	.
i	x_{i1}	…	x_{ip}	y_{i11}	…	y_{i1r}	$\bar{y}_{i1}$	s_{i1}	…	y_{im1}	…	y_{imr}	$\bar{y}_{im}$	s_{im}
.	.		.	.		.	.	.	…	.		.	.	.
.	.		.	.		.	.	.	…	.		.	.	.
.	.		.	.		.	.	.	…	.		.	.	.
n	x_{n1}	…	x_{np}	y_{n11}	…	y_{n1r}	$\bar{y}_{n1}$	s_{n1}	…	y_{nm1}	…	y_{nmr}	$\bar{y}_{nm}$	s_{nm}

The fitted second-order polynomial model for the process mean and process standard deviation functions of the response variables, denoted as $\hat{y}_{\mu_l}(x)$ and $\hat{y}_{s_l}(x)$, respectively, are given as follows:

$$\hat{y}_{\mu_l}(x) = \hat{\beta}_0 + \sum_{j=1}^{p} \hat{\beta}_j x_{ij} + \sum_{j=1}^{p} \hat{\beta}_{jj} x_{ij}^2 + \sum \sum_{j<k=2} \hat{\beta}_{jk} x_{ij} x_{ik} \tag{6.2}$$

$$\hat{y}_{s_l}(x) = \hat{\gamma}_0 + \sum_{j=1}^{p} \hat{\gamma}_j x_{ij} + \sum_{j=1}^{p} \hat{\gamma}_{jj} x_{ij}^2 + \sum \sum_{j<k=2} \hat{\gamma}_{jk} x_{ij} x_{ik} \tag{6.3}$$

The aim of desirability function is to overcome the problem of multiple responses to be a single response problem. During the desirability function approach process, each of the m predicted responses $\hat{y}_{1,}\hat{y}_{2,} \ldots, \hat{y}_m$ will be transform into individual desirability function d_l, where $0 \leq d_l \leq 1$, $l = 1,2, \ldots, m$. Then, combine all individual desirability function into overall desirability function D using geometric mean where $0 \leq D \leq 1$.

Myers and Carter (1973) was first to introduce dual response surface method to solve multiple responses where the response functions are separately model and simultaneously optimize the process mean and the process standard deviation to achieve the desired target while keeping the variance small. Numerous researchers have developed extensions of these methods, contributing to the breadth of knowledge in the field (Del Castillo & Montgomery, 1993; Lin & Tu, 1995; Copeland & Nelson, 1996; Baba et al., 2015).

Chen et al. (2013) proposed some secondary information to augment the overall desirability function and make the optimal solution more practical. The individual desirability function for the process mean, d_{μ_l} and for the process standard deviation, d_{s_l}, is formulated where $0 \leq d_{\mu_l} \leq 1$, $0 \leq d_{s_l} \leq 1$, respectively. For the process mean, the response variables can be classified into three types of desirability functions: larger-the-better (LTB), smaller-the better (STB), and nominal-the-best (NTB), depending on whether the mean response has to attain a user-specified target value, maximize or minimize, respectively. The three types of individual desirability function for LTB, STB and NTB are illustrated in Figure 6.1.

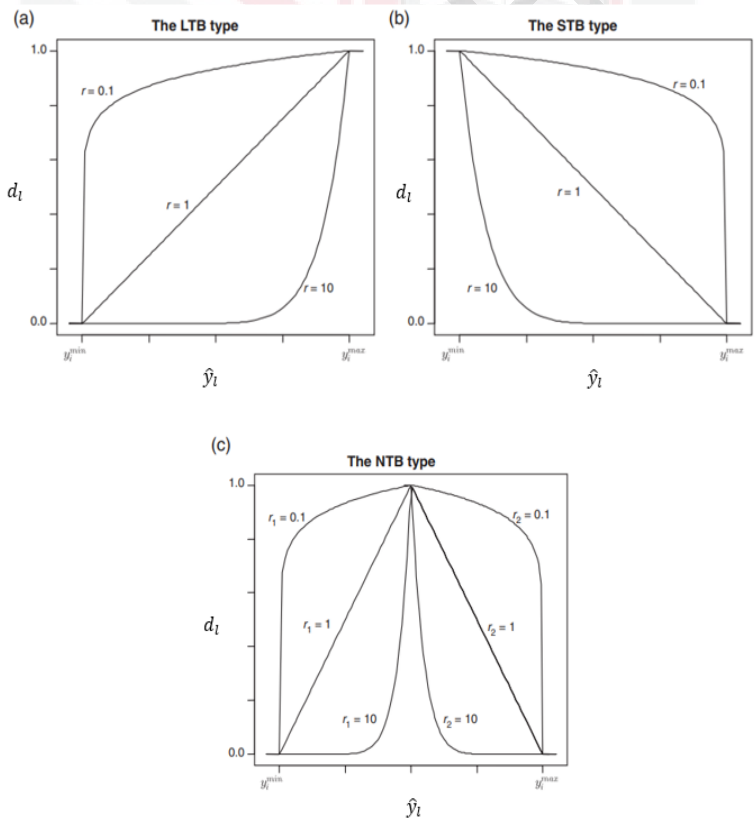


Figure 6.1 : Individual desirability functions for simultaneous optimization response variables, y_l

(a) The LTB type. (b) The STB type. (c) The NTB type

The individual desirability function for the NTB type when the response is maximized is defined as follows:

$$d_{\mu_l} = \begin{cases} \left(\dfrac{\hat{y}_{\mu_l}(x) - \bar{y}_{\min}}{\tau_{\mu_l} - \bar{y}_{\min}}\right)^{r_1} & \text{for } \bar{y}_{\min} \le \hat{y}_{\mu_l}(x) \le \tau_{\mu_l} \\ \left(\dfrac{\bar{y}_{\max} - \hat{y}_{\mu_l}(x)}{\bar{y}_{\max} - \tau_{\mu_l}}\right)^{r_2} & \text{for } \tau_{\mu_l} \le \hat{y}_{\mu_l}(x) \le \bar{y}_{\max} \end{cases} \quad (6.4)$$

For the LTB and STB type, the desirability function is defined as in Equation (6.5) and Equation (6.6) respectively.

$$d_{\mu_l} = \begin{cases} 0 & \text{for } \hat{y}_{\mu_l}(x) \le \bar{y}_{\min} \\ \left(\dfrac{\hat{y}_{\mu_l}(x) - \bar{y}_{\min}}{\bar{y}_{\max} - \bar{y}_{\min}}\right)^{r} & \text{for } \bar{y}_{\min} \le \hat{y}_{\mu_l}(x) \le \bar{y}_{\max} \\ 1 & \text{for } \hat{y}_{\mu_l}(x) \ge \bar{y}_{\max} \end{cases} \quad (6.5)$$

$$d_{\mu_l} = \begin{cases} 1 & \text{for } \hat{y}_{\mu_l}(x) \le \bar{y}_{\min} \\ \left(\dfrac{\hat{y}_{\mu_l}(x) - \bar{y}_{\max}}{\bar{y}_{\min} - \bar{y}_{\max}}\right)^{r} & \text{for } \bar{y}_{\min} \le \hat{y}_{\mu_l}(x) \le \bar{y}_{\max} \\ 0 & \text{for } \hat{y}_{\mu_l}(x) \ge \bar{y}_{\max} \end{cases} \quad (6.6)$$

where $\hat{y}_{\mu_l}(x)$ is the fitted second-order polynomial model for the process mean of the lth response variable as in Equation (6.2), $\bar{y}_{\max}$, $\bar{y}_{\min}$ and $\tau_{\mu_l}$are upper and lower limits and the target value for the response $\hat{y}_{\mu l}(x)$, respectively. The weights $r > 0$, r_1>0 and $r_2 > 0$ is a user-selected shape parameter.

After selecting an appropriate individual process mean desirability function for each response, all individual process mean desirability function d_{μ_l} are combined into an overall process mean desirability function using geometric mean as follows:

$$D = \left(d_{\mu_1} d_{\mu_2} \dots d_{\mu_m}\right)^{1/m} \quad (6.7)$$

For the process standard deviation effect, the STB type is considered as individual desirability since it is desirable to minimize the variation. The desirability function is defined as follows:

$$d_{s_l} = \begin{cases} 1 & \text{for } \hat{y}_{s_l}(x) \le s_{\min} \\ \left(\dfrac{s_{\max} - \hat{y}_{s_l}(x)}{s_{\max} - s_{\min}}\right)^r & \text{for } s_{\min} \le \hat{y}_{s_l}(x) \le s_{\max} \\ 0 & \text{for } \hat{y}_{s_l}(x) \ge s_{\max} \end{cases} \tag{6.8}$$

The $\hat{y}_{s_l}(x)$ is the fitted second-order polynomial model for the process standard deviation of the jth response variable as in Equation (6.3), $s_{\max}$ and $s_{\min}$ are the upper and the lower limits for the response $\hat{y}_{s_l}(x)$. The weight $r > 0$ is a user-selected shape parameter.

After defining all individual process standard deviation desirability function d_{s_l} are combined into overall process standard deviation desirability function using geometric mean:

$$S = \left(d_{s_1} d_{s_2} \dots d_{s_m}\right)^{1/m} \tag{6.9}$$

Several researchers have discussed techniques for establishing the shape parameter settings r, r_1 and r_2 (Derringer, 1994; Jeong & Kim, 2009). The choices for r, r_1 and r_2 are subjective, and they reflect the experts' opinion on how desirable each value of the predicted value of the outcome is. Different choices will result in different desirability shapes. Most practitioners will typically utilize a linear shape in the desirability $(r = 1)$ to assess penalties for deviating from a desired target value; the formulation does offer the flexibility to experiment with greater restriction. The larger values r_1 and $r_2 > 1$ imply the importance of being close to τ_{μ_l}. On the other hand, small values r_1 and $r_2 < 1$ suggest that $\hat{y}_{\mu_l}(x)$ does not have very close to τ_{μ_l}.

Another point that should be taken into consideration is the selection of an acceptable upper and lower limit for each response variable. While specification limits $\bar{y}_{\max}$ and $\bar{y}_{\min}$ typically exist with respect to the process mean response for a given characteristics, the identification of the limit for each variance response $s_{\max}$ and $s_{\min}$ may also subjective in nature. Frequently, however, within the design and manufacturing of a particular product where process monitoring occurs, there is prior knowledge of what may be considered unacceptable in terms of variability.

To simultaneously optimize both the process mean and process standard deviation of multiple responses, Chen et al. (2013) proposed to combine the overall process mean desirability function, D defined in Equation (6.7) and the overall process standard deviation desirability function, S defined in Equation (6.9) and assign a weight λ to reflect the relative importance of the process mean and standard deviation effects.

The overall desirability function of Chen et al. (2013) is denoted as DS_λ and defined as follows:

$$\mathrm{DS}_{\lambda} = \left(d_{\mu_1} d_{\mu_2} \dots d_{\mu_m}\right)^{\lambda/m} \left(d_{s_1} d_{s_2} \dots d_{s_m}\right)^{(1-\lambda)/m} \tag{6.10}$$

Afterwards, the DS_{λ} is optimized to obtain optimal factor settings. All analyses were computed using R software. The desirability package in R is used to obtain the overall desirability function, DS_{λ} values and optimal factor settings.

6.3 Proposed Robust Augmented Approach to the Desirability Function (RAADF)

The sample mean and sample standard deviation in Equation (6.1) are known to be easily affected by outliers. As an alternative, MM-estimator that has been discussed in Chapter 4 which has a very high breakdown point is used to estimate the process mean and the process standard deviation.

In practice, the fitted second-order polynomial models for the process mean and the process standard deviation functions of the response variables are often estimated by the OLS method. Data analysis based on the least squares estimator is less efficient and not reliable when outliers are present in the data (Riazoshams et al., 2010). To remedy this problem, robust regression technique has been considered to dampen the effects of outliers. To remedy this problem, the MM-regression estimator is used to estimate the parameters of the process mean and process standard deviation functions instead of using the OLS method.

We notice that the overall desirability function, DS_{λ} of Chen et al. (2013) in Equation (6.10) has a drawback. It can be shown that the DS_{λ} can be expressed as geometric mean. Geometric mean is known to be sensitive to outliers. Relying on the DS_{λ} by geometric mean will produce inefficient estimates when outliers are present in the dataset. Subsequently the predicted values of the optimal process mean and optimal process standard deviation of the response variables are not accurate. Hence, the AADF of Chen et al. (2013) does not provide good estimates when outliers are present in a dataset. Since geometric mean is not resistant to outliers, the formulation of DS_{λ} needs to be based on resistant estimator such as geometric median and geometric MM. Here, we show the derivation of our proposed overall desirability function based on geometric median and geometric MM.

Let us first show that Equation (6.10) can be expressed as Equation (6.13).

By taking log on both sides of Equation (6.10), gives

$$\log(\mathrm{DS}_{\lambda}) = \log\left[\left(d_{\mu_1} d_{\mu_2} \dots d_{\mu_m}\right)^{\lambda/m} \left(d_{s_1} d_{s_2} \dots d_{s_m}\right)^{(1-\lambda)/m}\right] \tag{6.11}$$

$$= \log\left[\left(d_{\mu_1} d_{\mu_2} \dots d_{\mu_m}\right)^{\lambda/m}\right] + \log\left[\left(d_{s_1} d_{s_2} \dots d_{s_m}\right)^{(1-\lambda)/m}\right]$$

$$= \frac{\lambda}{m}\left[\log d_{\mu_1} + \log d_{\mu_2} + \dots + \log d_{\mu_m}\right]$$

$$+ \frac{(1-\lambda)}{m}\left[\log d_{s_1} + \log d_{s_2} + \dots + \log d_{s_m}\right]$$

Conventionally, we take natural logarithm instead of logarithm with any arbitrary base. Thus, Equation (6.11) can be written as follows:

$$\ln(\mathrm{DS}_\lambda) = \frac{\lambda}{m}\left[\ln d_{\mu_1} + \ln d_{\mu_2} + \dots + \ln d_{\mu_m}\right] + \frac{(1-\lambda)}{m}\left[\ln d_{s_1} + \ln d_{s_2} + \dots + \ln d_{s_m}\right]$$

$$= \lambda \frac{\sum_{l=1}^{m}[\ln d_{\mu_l}]}{m} + (1-\lambda)\frac{\sum_{l=1}^{m}[\ln d_{s_l}]}{m} \qquad (6.12)$$

It can be observed that in each component of Equation (6.12) is the sample mean of $\ln d_{\mu_l}$ and $\ln d_{s_l}$.

To obtain expression for DS_λ, take exponent on both sided of Equation (6.12).

Therefore, the DS_λ can be written as

$$DS_\lambda = \exp\left[\lambda\left\{\frac{\sum_{l=1}^{m}\ln\left(d_{\mu l}\right)}{m}\right\} + (1-\lambda)\left\{\frac{\sum_{l=1}^{m}\ln(d_{sl})}{m}\right\}\right] \qquad (6.13)$$

It can be shown that the DS_λ in Equation (6.10) is equivalent to Equation (6.13).

It can be observed that in each component of Equation (6.13), $\sum_{l=1}^{m} ln\left(d_{\mu l}\right)/m$ and $\sum_{l=1}^{m} ln(d_{sl})/m$ represents the sample mean. As already mentioned, sample mean is popularly used measure of location. Nevertheless, it is now evident that sample mean performs poorly in the presence of outliers. Hence, it is needed to find an alternative robust method to replace the sample mean. Following the idea of Das and Imon (2014), each component of Equation (6.13) which is the sample mean, be replaced with median and sample mean based on MM-estimator. Median and MM estimator is known to be outlier-resistant estimator. Therefore, both estimators are not easily affected by outliers and are reliable alternatives.

Hence, the proposed overall desirability function based on median and sample mean based on MM estimator, can be defined as follows:

$$DS_{\lambda\text{median}} = \exp\left[\lambda\left\{\text{median}\left[\ln\left(d_{\mu_l}\right)\right]\right\} + (1-\lambda)\left\{\text{median}\left[\ln\left(d_{s_l}\right)\right]\right\}\right] \quad (6.14)$$

$$DS_{\lambda\text{MM}} = \exp\left[\lambda\left\{\text{MM}\left[\ln\left(d_{\mu_l}\right)\right]\right\} + (1-\lambda)\left\{\text{MM}\left[\ln\left(d_{s_l}\right)\right]\right\}\right] \quad (6.15)$$

Afterwards, Equation (6.13), Equation (6.14) and Equation (6.15) are optimized to obtain optimal factor settings. All analyses were computed using R software. The desirability package in R is used to optimize the overall desirability functions, $DS_{\lambda}, DS_{\lambda\text{median}}$ and $DS_{\lambda\text{MM}}$to obtain the estimated optimal factor settings.

We called Equation (6.13), Equation (6.14) and Equation (6.15) as AADF-G-mean, RAADF-G-median and RAADF-G-MM, respectively. For simplicity, we just refer to them as G-mean, G-median and G-MM.

The algorithm for the proposed robust augmented approach desirability function (RAADF) method is summarized as follows:

Step 1: For each level of the p factors $x = \left(x_1, x_2, \ldots, x_p\right)$ obtain the value of the process mean and process standard deviation for each response variable using the MM-estimator as discussed in Chapter 4, while for AADF the values are obtained using Equation (6.1).

Step 2: Using MM estimator, obtain the fitted second-order polynomial model for the process mean, $\hat{y}_{\mu_l}(x)$ and the fitted second-order polynomial model for the process standard deviation, $\hat{y}_{s_l}(x)$, of the response variables, while for AADF, the fitted second-order polynomial model is obtained using OLS.

Step 3: Compute the individual desirability function for the predicted responses of the process mean, d_{μ_l} according to three type of desirability functions, i.e., NTB, LTB and STB. The choice is depending whether the process mean response has to obtain a user-specified target value, maximize or minimize.

Step 4: Compute the individual desirability function for the predicted responses of the process standard deviation, d_{s_l} according to STB type of desirability function.

Step 5: Formulate the robust augmented approach to the desirability function (RAADF) based on geometric median and geometric MM as follows:

$$DS_{\lambda\text{median}} = \exp\left[\lambda\left\{\text{median}\left[\ln\left(d_{\mu_l}\right)\right]\right\} + (1-\lambda)\left\{\text{median}\left[\ln\left(d_{s_l}\right)\right]\right\}\right]$$

$$DS_{\lambda\text{MM}} = \exp\left[\lambda\left\{\text{MM}\left[\ln\left(d_{\mu_l}\right)\right]\right\} + (1-\lambda)\left\{\text{MM}\left[\ln\left(d_{s_l}\right)\right]\right\}\right]$$

Step 6: Find the optimal factor settings, $x^* = (x_1^*, x_2^*, \dots, x_p^*)$ that maximizes the overall desirability function DS_λ, $DS_{\lambda\text{median}}$ and $DS_{\lambda\text{MM}}$ over the experimental region using desirability package in R software.

Step 7: Compute the optimum predicted responses process mean, $\hat{y}_{\mu_l}(x^*)$ by substituting the values of x^* in the Equation (6.2).

Step 8: Compute the optimum predicted responses process standard deviation, $\hat{y}_{s_l}(x^*)$ by substituting the values of x^* in the Equation (6.3).

6.4 Monte Carlo Simulation Study

The performances of the newly proposed G-median and G-MM methods are evaluated via Monte Carlo simulation experiment and to compare its performance with the existing G-mean. As per Baba et al. (2015), at each control factor settings $x_{ij} = (x_{i1}, x_{i2}, x_{i3})$, $i = 1,2, \dots, 27$, a second-order polynomial model with three and six response variables, ($m = 3$ and 6) with five replicates for each factor setting are randomly generated from normal distribution with mean $\mu(x)$ and standard deviation $\sigma(x)$, as in Equation (6.16).

$$\mu(x) = 500 + (x_1 + x_2 + x_3)^2 + x_1 + x_2 + x_3 \tag{6.16}$$

$$\sigma(x) = \sqrt{100 + (x_1 + x_2 + x_3)^2 + x_1 + x_2 + x_3}$$

As per Park and Cho (2003), the values 500 and 100 were chosen so that the mean will be closer to the target value ($\tau = 500$) with certain variations. All simulation experiments are done based on $REP = 1000$ Monte Carlo simulations. In order to see the effect of outliers, this data set has been modified to contain three outliers for $m = 3$, and five outliers for $m = 6$. The outliers were generated from $N(50, 10^2)$ distribution. The performance of the proposed method is evaluated based on the bias and root mean squares errors (RMSE) of the estimated optimal process mean responses, $\hat{y}_{\mu_l}(x)$.

In this simulation study, we also be interested to see the performance of our methods on different values of λ in the range of (0,1). This is due to the fact that as the value of λ decreases, less weight is placed on optimization of the process mean. When $\lambda = 0$, DS_λ depends only on process standard deviation. We consider the value of λ, i.e.,$\lambda =$ 0.5,0.8 and 0.9 and compare our proposed method with those obtained from the existing methods.

Results in Table 6.2 to Table 6.7 present the overall desirability functions, the bias and the RMSE for the G-mean, G-median and G-MM with different value of λ. A good method is one that has higher values of desirability function. Let us first focus on the results of desirability function of G-mean, G-median and G-MM exhibited in Table 6.2 and Table 6.3. It is interesting to observe that as the value of λ increases, the overall desirability values also increased. These results show that the overall desirability function depends on the given weight, λ. It can be seen that all the three methods perform best at $\lambda = 0.9$. It can also be observed that the G-MM method provides highly appealing results with higher values of the overall desirability function, $DS_{\lambda\mathrm{MM}}$, followed by the G-median and G-mean irrespective of whether or not outliers exist in the dataset.

Table 6.2 : The overall desirability function values, $m = 3$

Cont. Level	λ	Methods		
		DS_{λ}	$DS_{\lambda\mathrm{median}}$	$DS_{\lambda\mathrm{MM}}$
0%	0.5	0.7504	0.8346	0.8418
	0.8	0.8454	0.9162	0.9540
	0.9	0.9061	0.9454	0.9540
3 Outliers	0.5	0.5456	0.8143	0.8342
	0.8	0.5487	0.8907	0.9040
	0.9	0.5621	0.9142	0.9283

Table 6.3 : The overall desirability function values, $m = 6$

Cont. Level	λ	Methods		
		DS_{λ}	$DS_{\lambda\mathrm{median}}$	$DS_{\lambda\mathrm{MM}}$
0%	0.5	0.6573	0.8114	0.9402
	0.8	0.7978	0.9101	0.9212
	0.9	0.8508	0.9253	0.9394
5 Outliers	0.5	0.6302	0.8016	0.8192
	0.8	0.7421	0.8825	0.8955
	0.9	0.7902	0.9219	0.9342

It is important to note that a good multiple response optimization method using desirability function approach must not only able to provide higher value of overall desirability function but also must be efficient in terms of producing the smallest bias and RMSE for predicted responses of the process mean and process standard deviation. Tables 6.4-6.7 present the bias and RMSE for the three methods. It can be observed that the G-MM consistently provides the smallest bias and RMSE whether or not outliers are present in the dataset. The G-mean of Chen et al. (2013) performs poorly in the presence of outliers. The G-median performs better than the G-mean but it cannot outperform the G-MM. The overall results show that the G-MM performs excellently, followed by the G-median and G-mean irrespective of the number of outliers, number of response variables and the value of λ.

Table 6.4 : Bias and MSE of $\hat{y}_{\mu_l}$ without outlier, $m = 3$

λ	Method	$\hat{y}_{\mu 1}$		$\hat{y}_{\mu 2}$		$\hat{y}_{\mu 3}$	
		Bias	**RMSE**	**Bias**	**RMSE**	**Bias**	**RMSE**
$\lambda = 0.5$	G-mean	0.3061	1.8814	0.2953	1.8369	0.1333	1.6889
	G-median	0.2385	2.4796	0.2312	2.3869	0.2389	0.5468
	G-MM	0.1318	0.6543	0.1305	0.6725	0.1549	0.4325
$\lambda = 0.8$	G-mean	0.5834	1.8467	0.4729	0.4729	0.2954	1.6624
	G-median	0.2775	1.2678	0.1861	0.3643	0.1495	1.3304
	G-MM	0.1584	0.8423	0.1624	0.0254	0.1027	0.9548
$\lambda = 0.9$	G-mean	0.4691	1.9247	0.2333	1.5849	0.2954	1.6624
	G-median	0.1437	1.4265	0.1624	1.5197	0.2004	1.3786
	G-MM	0.1187	1.1546	0.1300	0.7259	0.1628	0.358

Table 6.5 : Bias and MSE of $\hat{y}_{\mu_l}$ without outlier, $m = 6$

λ	Method	$\hat{y}_{\mu 1}$		$\hat{y}_{\mu 2}$		$\hat{y}_{\mu 3}$	
		Bias	**RMSE**	**Bias**	**RMSE**	**Bias**	**RMSE**
$\lambda = 0.5$	G-mean	0.3061	1.8814	0.2953	1.8369	0.1333	1.6889
	G-median	0.1523	1.2845	0.2146	0.5203	0.1025	2.2130
	G-MM	0.1342	0.5423	0.1145	0.1487	0.1006	0.7901
$\lambda = 0.8$	G-mean	0.5834	1.8467	0.4729	0.8437	0.2954	1.6624
	G-median	0.2187	1.6548	0.2105	0.6487	0.1217	1.6254
	G-MM	0.2014	0.8456	0.1454	0.2459	0.1010	1.1547
$\lambda = 0.9$	G-mean	0.4691	1.9247	0.2333	1.5849	0.2954	1.6624
	G-median	0.1587	1.6423	0.1746	1.4235	0.1347	1.2574
	G-MM	0.1154	1.2548	0.1405	0.7852	0.1136	1.2456

λ	Method	$\hat{y}_{\mu 4}$		$\hat{y}_{\mu 5}$		$\hat{y}_{\mu 6}$	
		Bias	**RMSE**	**Bias**	**RMSE**	**Bias**	**RMSE**
$\lambda = 0.5$	G-mean	0.0366	0.7405	0.0640	0.6820	0.2548	0.8949
	G-median	0.0143	0.4251	0.0241	0.4251	0.2345	0.7123
	G-MM	0.0105	0.3679	0.0214	0.3478	0.1467	0.5214
$\lambda = 0.8$	G-mean	0.0759	0.3910	0.1946	0.5998	0.0571	0.7413
	G-median	0.0497	0.1450	0.1279	0.4250	0.0154	0.6321
	G-MM	0.0254	0.1056	0.1146	0.3458	0.0121	0.4671
$\lambda = 0.9$	G-mean	0.0480	0.3831	0.3300	0.7348	0.0897	0.7372
	G-median	0.0156	0.1974	0.1462	0.5479	0.0425	0.4219
	G-MM	0.0745	0.1252	0.0950	0.3458	0.0143	0.2146

Table 6.6 : Bias and MSE of $\hat{y}_{\mu_l}$ with outlier, $m = 3$

λ	Method	$\hat{y}_{\mu1}$		$\hat{y}_{\mu2}$		$\hat{y}_{\mu3}$	
		Bias	**RMSE**	**Bias**	**RMSE**	**Bias**	**RMSE**
$\lambda = 0.5$	G-mean	1.4874	2.6167	1.4036	2.9432	1.1534	2.0769
	G-median	0.2125	0.4796	0.1112	1.4969	0.4389	0.5568
	G-MM	0.1516	0.3541	0.0955	0.5725	0.1329	0.4425
$\lambda = 0.8$	G-mean	1.2804	2.5679	1.4883	1.5708	1.1363	2.7084
	G-median	0.2123	1.1688	0.1661	0.3543	0.3295	1.3244
	G-MM	0.1421	0.6443	0.1434	0.1254	0.1067	0.9648
$\lambda = 0.9$	G-mean	1.5426	2.4810	1.3516	2.4873	1.1031	2.5944
	G-median	0.1358	1.3255	0.1324	1.6593	0.1304	1.2586
	G-MM	0.1100	1.1356	0.0950	0.3259	0.9328	0.3381

Table 6.7 : Bias and MSE of $\hat{y}_{\mu_l}$ with outlier, $m = 6$

λ	Method	$\hat{y}_{\mu1}$		$\hat{y}_{\mu2}$		$\hat{y}_{\mu3}$	
		Bias	**RMSE**	**Bias**	**RMSE**	**Bias**	**RMSE**
$\lambda = 0.5$	G-mean	1.4061	2.3315	1.2548	1.8949	1.0366	1.7405
	G-median	0.1345	0.4523	0.0441	0.4251	0.2146	0.5203
	G-MM	0.0467	0.2314	0.0114	0.3478	0.1145	0.1487
$\lambda = 0.8$	G-mean	1.6835	2.7457	1.0571	1.7413	1.0759	1.3910
	G-median	0.0154	0.6321	0.1279	0.4250	0.2105	0.6487
	G-MM	0.0121	0.4671	0.1146	0.3458	0.1454	0.2459
$\lambda = 0.9$	G-mean	1.6691	2.6247	1.0897	1.7372	1.0480	1.3831
	G-median	0.0425	0.4219	0.1462	0.5479	0.1746	1.4235
	G-MM	0.0143	0.2146	0.0950	0.3458	0.1405	0.7852

λ	Method	$\hat{y}_{\mu4}$		$\hat{y}_{\mu5}$		$\hat{y}_{\mu6}$	
		Bias	**RMSE**	**Bias**	**RMSE**	**Bias**	**RMSE**
$\lambda = 0.5$	G-mean	1.1333	2.6889	1.0640	1.6820	1.2953	2.8369
	G-median	0.1025	2.2130	0.2146	0.5203	0.1523	1.2845
	G-MM	0.1006	0.7901	0.1145	0.1487	0.1342	0.5423
$\lambda = 0.8$	G-mean	1.2954	2.6624	1.1946	1.5998	1.4729	1.8437
	G-median	0.1217	1.6254	0.2105	0.6487	0.2187	1.6548
	G-MM	0.1010	1.1547	0.1454	0.2459	0.2014	0.8456
$\lambda = 0.9$	G-mean	1.2954	2.6624	1.3300	1.7348	1.2333	2.5849
	G-median	0.1347	1.2574	0.1746	1.4235	0.1587	1.6423
	G-MM	0.1136	1.2456	0.1405	0.7852	0.1154	1.2548

6.5 Numerical Example

The Colloidal gas aphrons experiment data which was taken from Jauregi et al. (1997) is used to evaluate the performance of our proposed methods compared to the G-mean method. Chen et al. (2013) also used this dataset to assess the performance of their methods. This dataset is presented in Appendix A4. The purpose of this experiment was to investigate the effects of the concentration of surfactant (x_1), the concentration of salt (x_2), and the time of stirring (x_3) during the optimization process on the three properties of the colloidal gas aphrons, including stability (y_1), volumetric ratio (y_2), and temperature (y_3). Each of the response variables is replicated three times. The response y_1 was the larger-the-better (LTB) type, y_2 was the smaller-the-better (STB) type, and y_3 was the nominal-the-best (NTB) type. The experimental region was the three-dimensional cube defined by $-1 \leq x_{ij} \leq 1,\ i = 1,2,\ldots,15$ and $j = 1,2,3$. According to Hu et al. (2007), the value of the constant $r = 1$ is chosen. The associated goals, specifications and target values for the characteristics are shown in Table 6.8.

Table 6.8 : The quality characteristics – goals and specification limit for Colloidal gas aphrons experiment data

Quality characteristics	Goal	Specifications limits	Target value
Y_1 stability	LTB	$3.0 \leq \hat{y}_{\mu 1} \leq 7.0$	-
		$0 \leq \hat{y}_{s1} \leq 0.1$	
Y_2 volumetric	STB	$0.1 \leq \hat{y}_{\mu 2} \leq 0.6$	-
		$0 \leq \hat{y}_{s2} \leq 0.1$	
Y_3 temperature	NTB	$15 \leq \hat{y}_{\mu 3} \leq 45$	$\tau_{\mu_3} = 30$
		$1 \leq \hat{y}_{s3} \leq 2$	

In order to see the effect of an outlier on the estimated mean optimal response, we deliberately changed two data points, that was the y_{111} observation and y_{1533} observation. The usual analysis was done using the coded values. The G-mean, G-median and G-MM desirability functions methods were then applied to the data. The results are exhibited in Table 6.9 and 6.10.

The results are exhibited in Table 6.9 and 6.10 for the G-mean, G-median and G-MM. From the tables, as expected, the G-median and G-MM yield higher values of the overall desirability function with and without contaminated data. However, for the G-mean, the overall desirability function is shown to be badly affected by the presence of the outliers. As can be seen in Table 6.10, $\hat{y}_{\mu_2}$ is 0.01, which is less than the lower bound of 0.1, $\hat{y}_{s_2}$ is 0.75 which is greater than the upper bound of 0.1 and $\hat{y}_{s_3}$ is -0.48 which is less than the lower limit of 1. The results give an indication that the G-mean is no longer useful in the presence of the outliers. On the other hand, the G-median and G-MM give reliable estimates, where the predicted responses of the proses mean and process standard deviation lie between the specification limits for each variable.

Table 6.9 : Comparisons of optimization results for the colloidal gas aphrons experiment without outliers

Optimization results	$\hat{y}_{\mu 1}$	$\hat{y}_{\mu 2}$	$\hat{y}_{\mu 3}$	$\hat{y}_{s1}$	$\hat{y}_{s2}$	$\hat{y}_{s3}$
G-mean $x^* = (-0.7, -0.4, -1.0)$ $DS_\lambda = 0.5687$	4.45	0.27	26.23	0.05	0.05	1.71
G-median $x^* = (1.0, -0.4, -0.8)$ $DS_{\lambda\text{median}}$=0.823	6.31	0.50	27.3	0.07	0.02	1.89
G-MM $x^* = (1.0, -0.6, -0.8)$ $DS_{\lambda\text{MM}}$ = 0.8337	6.32	0.50	27.3	0.09	0.02	1.89

Table 6.10 : Comparisons of optimization results for the colloidal gas aphrons experiment with outliers

Optimization results	$\hat{y}_{\mu 1}$	$\hat{y}_{\mu 2}$	$\hat{y}_{\mu 3}$	$\hat{y}_{s1}$	$\hat{y}_{s2}$	$\hat{y}_{s3}$
G-mean $x^* = (0.3, -0.4, -1.0)$ $DS_\lambda = 0.7321$	4.54	0.01	30.59	0.75	-0.48	1.79
G-median $x^* = (1.0, 0.7, -1.0)$ $DS_{\lambda\text{median}}$=0.7331	5.92	0.48	25.73	0.08	0.05	1.63
G-MM $x^* = (0.8, -0.7, -0.9)$ $DS_{\lambda\text{MM}}$ = 0.7640	6.28	0.49	25.12	0.09	0.05	1.65

6.6 Conclusion

The main aim of this study was to develop a reliable alternative method to simultaneously optimize multiple responses with replicates for each response variable when outliers are present in the data. A robust augmented approach to the desirability function (RAADF) is developed in this regard. The RAADF is developed by formulating robust overall desirability function using geometric median and geometric MM. Prior to developing the RAADF, the fitted response functions of the sample mean and sample standard deviation of the response are obtained using MM-estimator. The classical AADF based on geometric mean performs poorly in the presence of outliers. Our proposed G-MM outperforms the G-median and G-mean evident by having the smallest value of RMSE, biases and larger value of desirability function irrespective of the number of response variables and the number of outliers. Hence, it can be a good alternative in dealing with multiple responses with replicate when outliers are present in a dataset.

CHAPTER 7

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER STUDIES

7.1 Introduction

This thesis conducts extensive research to develop optimization techniques for multiple responses experimental data. We focus on the impact of outlier and heteroscedasticity problems in the data sets. Several robust techniques have been proposed to handle the problem of outliers and heteroscedasticity in multiple responses data. The contributions of the study, conclusions, and recommendations for future studies are presented in this chapter, which is organized as follows:

7.2 Research Contributions

The overall conclusions of the current study are described in the following subsections which are based on the main contributions of each chapter, from Chapter 3 to Chapter 6.

7.2.1 Penalty Function Optimization in Dual Response Based on Decision Maker's Preference

One of the major concerns in performing the dual response surface optimization is on getting the results of optimal factor settings that can used to estimate the optimal mean response. The mean squared error (MSE) optimization scheme, also known as LT methods has been widely used to simultaneously optimizing process mean and process standard deviation. One of the shortcomings of LT methods is that it does not take into account the measure of violation of the constraint function during the conversion of the constraint optimization into unconstraint. Additionally, it does not specify how large the magnitude of the target value should deviate from the actual target of the experimenter. As a solution to this shortcoming, penalty function (PM) optimization method is put forward. However, it is still not very efficient since this method concentrates more on bias and force the estimated mean response to be close to the target value. Moreover, the determination of penalty constant is not clearly specified. Thus, a new penalty function optimization method which is based on decision maker's (PFDMM) preference is proposed in Chapter 3. The PFDMM is developed in two stages whereby the penalty constant, ξ is determined in the first stage and subsequently the PFDMM functions is optimize to obtain the optimal factor settings to estimate the optimal mean response. The performance of the PFDMM is compared with the VM, LT, WMSE and PM method. Empirical evidence via simulation experiments and numerical data have shown that the PFDMM method is more efficient compared to the existing methods in this study as it has the least bias and RMSE.

7.2.2 High Breakdown Estimator for Dual Response Optimization in the Presence of Outliers

Dual response surface approach is a tool for simultaneously optimizing the process mean and the process standard deviation functions of the response variables by assuming that experimental data are approximately normal and no outlier in the data set. In dual response approach, a second-order polynomial model for the process mean and process standard deviation of the response y are constructed. The sample mean and sample standard deviation are widely used to estimate the mean and the standard deviation of the responses. Moreover, the OLS method is usually employed to estimate the parameters of both models. In the existence of outliers, all the existing methods (VM, LT, WMSE and PM) performs poorly due to the used of the OLS estimator in their algorithm. The PFDMM proposed in Chapter 3 is shown to be more efficient to simultaneously optimize the process mean and process standard deviation functions of the response variables. Nonetheless, it is based on the OLS estimator. Hence, in the presence of outliers, the PFDMM will be much affected. As a solution to this problem in Chapter 4 robust penalty function based on decision maker's preference denoted as $\mathrm{PFDMM_R}$ is developed. Robust MM-mean, robust MM-standard deviation and robust MM regression estimators which are highly efficient and have high breakdown points are integrated in the formulation of robust $\mathrm{PFDMM_R}$ method. The $\mathrm{PFDMM_R}$ is compared with the VM, LT, WMSE, PM, and PFDMM methods. The results show that $\mathrm{PFDMM_R}$ produces more accurate and reliable results when outliers are present in the dataset.

7.2.3 New Robust Parameter Estimation method for Dual Response Surface models in the presence of Heteroscedastic Errors and Outliers

The reweighted least square (RLS) method is often used by practitioners to estimate the parameters of the process mean and process standard deviation functions for dual response surface models with heteroscedastic errors. Nonetheless, in the presence of outliers and heteroscedastic errors, the RLS performs poorly. To remedy these problems, a two stage robust estimator based on MM-estimator methods (TSR-MM) was put forward. However, this method is not effective enough because no proper method is used to identify the exact number of vertical outliers. Moreover, the TSR-MM is incorporated in the LT optimization method used to optimize the process mean and the process standard deviation response functions. It is now evident that the LT optimization scheme is less efficient since it does not achieve the target minimum bias of the estimated process mean response. As a solution to this problem, the robust reweighted least square (RRWLS) is proposed in chapter 5 to withstand the influence of outliers and simultaneously encounter the heteroscedastic condition. The establishment of the RRWLS consists of two weight functions whereby the initial weight function used the KNN method based on MM-estimator to reduce the effect of heteroscedasticity and outliers and the second weight function is based on weight function of the modified generalized residual studentized ($\mathrm{MGRS}t_i$) method used to down weight the effects of remaining outliers. Then, both weights are multiplied to get the final weight. The RRWLS is found to be more efficient than the RLS and TSR-MM. The RLS, TSR-MM and RRWLS are then integrated in the VM, LT, WMSE, PM and PFDMM optimization methods. Empirical evidence via simulation experiments and numerical data have shown that the PFDMM-RRWLS method outperforms the PFDMM-TSR-MM and PFDMM-RLS optimization method.

7.2.4 Improvised Desirability Function Optimization Technique for Multiple Responses based Modified Geometric Mean and MM-estimator in the Presence of Outliers

An augmented desirability function (AADF) method is introduced by Chen et al. (2012) to simultaneously optimize multiple responses that consider variability between responses. The AADF is developed by augmenting the overall desirability function of the predicted responses, $\hat{y}_j$ and the overall desirability function of the standard deviation of predicted responses, $\mathrm{sd}(\hat{y}_j)$ into a single overall desirability function using geometric mean. The AADF method is extended by Chen et al. (2013) to be applicable to multiple responses with replicates for each response variable. Prior to the construction of the individual desirability function, the fitted response functions of the process mean and process standard deviation are obtained using the OLS estimator. It is now evident that the OLS and the geometric mean are easily affected by outliers. To remedy these problems, robust augmented approach desirability function (RAADF) is proposed in chapter 6. The RAADF is formulated by augmenting the overall desirability function of the predicted process mean responses and the overall desirability function of the process standard deviation of predicted responses into a single overall desirability function using geometric median and geometric MM. In order to obtain the individual desirability functions of the predicted responses of the process mean and process standard deviation, the transformation functions based on LTB, STB and NTB type and based on STB type are used, respectively. Moreover, the MM-regression estimator is used to estimate the parameters of the process mean and process standard deviation functions instead of using the OLS method. The results of the simulation study and real dataset signify that the RAADF-geometric MM desirability function method consistently outperforms the RAADF-geometric median and AADF-geometric mean when outliers are present in the dataset.

7.3 Areas of Further Studies

Outliers in a set of data are defined to be an observation that appears to be inconsistent with the remainder of that set of data in different fields of studies. It is now evident that outliers have on adverse effect on the calculated values of various estimates. When outliers are present in a dataset, the entire set up of design experiment is disturbed. In this thesis, our proposed methods are applied only to balanced design data, where the number of replicates is complete for each design point. It is important to note that unbalanced design makes errors of a model become heteroscedastic which produce large standard errors of estimates. It would be a good idea to explore optimization techniques dealing with unbalanced design for future studies. Moreover, we can suggest some topics for further studies as follows:

- Further research can be done on robust heteroscedasticity diagnostic method in dual response surface model.
- In Chapter 6, the robust augmented approach desirability function (RAADF) is developed using geometric MM and geometric median. With the similar manner, RAADF can be investigated as further research for Hybrid Genetic Algorithmn Designs.

REFERENCES

Akbilgic, O., and Akinci, E. D. (2009). A novel regression approach: Least squares ratio. *Communications in Statistics-Theory and Methods*, 38(9), 1539-1545.

Alguraibawi, M., Midi, H., and Imon, A. H. M. (2015). A new robust diagnostic plot for classifying good and bad high leverage points in a multiple linear regression model. *Mathematical Problems in Engineering*, 1, 1-12.

Alih, E., and Choon, H. O. (2015). An outlier-resistant test for heteroscedasticity in linear models. *Journal of Applied Statistics*, 42(8), 1617-1634.

Anderson, R. (2008). *Modern methods for robust regression*. The United States of America: Sara Miller McCune. SAGE publications.

Andrews, D. F. (1974). A robust method for multiple linear regression. *Technometrics*, 16(4), 523-531.

Baba, A. M., Midi, H., & Abd Rahman, N. H. (2022). Spatial Outlier Accommodation Using a Spatial Variance Shift Outlier Model. *Mathematics*, 10(17), 1-19.

Baba, I., Midi, H., Ibragimov, G., and Rana, S. (2022). A New Optimization Scheme for Robust Design Modeling with Unbalanced Data. *Sains Malaysiana*, 51(5), 1577-1586.

Baba, I., Midi, H., Rana, S., and Ibragimov, G. (2015). An alternative approach of dual response surface optimization based on penalty function method. *Mathematical Problems in Engineering*, 2015.

Bakar, N. M. A., and Midi, H. (2015). Robust centering in the fixed effect panel data model. *Pakistan Journal of Statistics*, 31(1).

Baltagi, B. H., Egger, P., and Pfaffermayr, M. (2013). A generalized spatial panel data model with random effects. *Econometric Reviews*, 32(5-6), 650-685.

Barnett, V. and Lewis, T. (1994). *Outliers in statistical data*. 3rd edition. New York: Wiley.

Bashiri, M., and Moslemi, A. (2013). Simultaneous robust estimation of multi-response surfaces in the presence of outliers. *Journal of Industrial Engineering International*, 9(1), 1-12.

Bashiri, M., and Moslemi, A. (2013). The analysis of residuals variation and outliers to obtain robust response surface. *Journal of Industrial Engineering International*, 9(1), 1-10.

Beaton, A. E., and Tukey, J. W. (1974). The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. *Technometrics*, 16(2), 147-185.

Bertsimas, D., Shtern, S., & Sturt, B. (2022). Two-stage sample robust optimization. *Operations Research*, 70(1), 624-640.

Bickel, P. J. (1975). One-step Huber estimates in the linear model. *Journal of the American Statistical Association,* 70(350), 428-434.

Bikbulatov, E. S., and Stepanova, I. E. (2011). Harrington's desirability function for natural water quality assessment. *Russian Journal of General Chemistry*, 81(13), 2694-2704.

Birkes, D., and Dodge, Y. (1993). *Alternative methods of regression*. John Wiley and Sons.

Box, G. E. P., and Wilson (1951). On the Experimental Attainment of Optimum Conditions. *Journal of the Royal Statistical Society: Series, B* 13, 1-45.

Box, G. E., and Draper, N. R. (1987). *Empirical model-building and response surfaces*. John Wiley and Sons.

Boylan, G. L., and Cho, B. R. (2013). Comparative studies on the high-variability embedded robust parameter design from the perspective of estimators. *Computers and Industrial Engineering,* 64(1), 442-452.

Breusch, T. and Pagan, A. (1979). A simple test for heteroscedasticity and random coefficient variation, *Econometrica*, 47, 1287–1294.

Campbell, N. A., Lopuhaä, H. P., and Rousseeuw, P. J. (1998). On the calculation of a robust S-estimator of a covariance matrix. *Statistics in medicine*, 17(23), 2685-2695.

Chatterjee, S., and Hadi, A. S. (2006). *Regression analysis by example*. 4th edition. New York: Wiley.

Chelladurai, S. J. S., Murugan, K., Ray, A. P., Upadhyaya, M., Narasimharaj, V., and Gnanasekaran, S. (2021). Optimization of process parameters using response surface methodology: A review. *Materials Today: Proceedings,* 37, 1301-1304.

Chen, H. W., Wong, W. K., and Xu, H. (2012). An augmented approach to the desirability function. *Journal of Applied Statistics,* 39(3), 599-613.

Chen, H. W., Xu, H., and Wong, W. K. (2013). Balancing location and dispersion effects for multiple responses. *Quality and Reliability Engineering International*, 29(4), 607-615.

Chiao, C. H., and Hamada, M. (2001). Analyzing experiments with correlated multiple responses. *Journal of Quality Technology*, 33(4), 451-465.

Ch'ng, C. K., Quah, S. H., and Low, H. C. (2005). The MM-estimator in response surface methodology. *Quality Engineering*, 17(4), 561-565.

Cho, B. R., and Park, C. (2005). Robust design modeling and optimization with unbalanced data. *Computers and Industrial Engineering,* 48(2), 173-180.

Coakley, C. W., and Hettmansperger, T. P. (1993). A bounded influence, high breakdown, efficient regression estimator. *Journal of the American Statistical Association,* 88(423), 872-880.

Copeland, K. A., and Nelson, P. R. (1996). Dual response optimization via direct function minimization. *Journal of Quality Technology,* 28(3), 331-336.

Costa, N. R., Lourenço, J., and Pereira, Z. L. (2011). Desirability function approach: a review and performance evaluation in adverse conditions. *Chemometrics and Intelligent Laboratory Systems*, 107(2), 234-244.

Croux, C. and Haesbroeck G. (2003). Implementing the Bianco and Yohai estimator for logistic regression. *Computational Statistics and Data Analysis, 44*(1), 273-295.

Croux, C., Rousseeuw, P. J., and Hössjer, O. (1994). Generalized S-estimators. *Journal of the American Statistical Association*, 89(428), 1271-1281.

da Silva, B., Gonzaga, L. V., Fett, R., and Costa, A. C. O. (2019). Simplex-centroid design and Derringer's desirability function approach for simultaneous separation of phenolic compounds from Mimosa scabrella Bentham honeydew honeys by HPLC/DAD. *Journal of Chromatography, A*, 1585, 182-191.

Das, K. R., and Imon, A. R. (2014). Geometric median and its application in the identification of multiple outliers. *Journal of Applied Statistics,* 41(4), 817-831.

Del Castillo, E., and Montgomery, D. C. (1993). A nonlinear programming solution to the dual response problem. *Journal of Quality Technology,* 25(3), 199-204.

Derringer G., and Suich R. (1980). Simultaneous Optimization of Several Response Variables. *Journal of Quality Technology,* 12(4), 214-219.

Derringer, G. C. (1994). A balancing act-optimizing a products properties. *Quality Progress*, 27(6), 51-58.

Dey, S. S., and Dora, K. C. (2014). Optimization of the production of shrimp waste protein hydrolysate using microbial proteases adopting response surface methodology. *Journal of Food Science and Technology,* 51(1), 16-24.

Dhhan, W., Rana, S., and Midi, H. (2017). A high breakdown, high efficiency and bounded influence modified GM estimator based on support vector regression. *Journal of Applied Statistics,* 44(4), 700-714.

Ding, R., Lin, D. K., and Wei, D. (2004). Dual-response surface optimization: A weighted MSE approach. *Quality Engineering,* 16(3), 377-385.

Donoho, D. L. and Huber, P. J. (1983). *The Notion of Breakdown Point. A festschrift for Erich L. Lehmann*, 157-184.

Draper, N. R. (1963). Ridge analysis of response surfaces. *Technometrics*, 5(4), 469-479.

Ellenberg, J. H. (1976). Testing for a single outlier from a general linear regression. *Biometrics*, 637-645.

Fang, J., and He, Z. (2010, May). Analysis of response surface designs to outlier. In *2010 International Conference on E-Business and E-Government* (pp. 2648-2651). IEEE.

Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A: Containing Papers of a Mathematical or Physical Character* 222(594-604), 309-368.

Fuller, D., and Scherer, W. (1998, October). The desirability function: underlying assumptions and application implications. In *SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 98CH36218)* (Vol. 4, pp. 4016-4021). IEEE.

Gatza, P. E., and McMillan, R. C. (1973). *The Use of Experimental Design and Computerized Data Analysis in Elastomer Development Studies*. Army Mobility Equipment Research and Development Center Fort Belvoir Va.

Glejser, H. (1969). A new test for heteroskedasticity. *Journal of AmericanStatistical Association.* 64, 316–323.

Goethals, P. L., and Cho, B. R. (2011). Solving the optimal process target problem using response surface designs in heteroscedastic conditions. *International Journal of Production Research*, 49(12), 3455-3478.

Goldfeld, S.M. and Quandt, R.E. (1965). Some tests for homoskedasticity.J*ournal of the American Statistical Association,* 60, 539–547.

Goupy, J. (1996). Unconventional experimental designs theory and application. *Chemometrics and intelligent laboratory systems,* 33(1), 3-16.

Greene, W. (2008). *Econometric Analysis*, New York: Pearson.

Guedri, W., Jaoudi, M., and Msahli, S. (2022). Evaluating farmer's satisfaction of different agrotextile bunch covers using desirability function. *Textile Research Journal*, 92(17-18), 3337-3350.

Gujarati, D. and D. Porter (2002). *Basic econometrics,* McGraw-Hill/Irwin.

Habshah M., Rana, S., and Imon, A. H. M. (2014). Two-step robust estimator in heteroscedastic regression model in the presence of outliers. *Economic computation & economic cybernetics studies & research*, 48(3).

Habshah Midi, Md. Sohel Rana, and A. H. M. Rahmatullah Imon (2009). The performance of robust weighted least squares in the presence of outliers and heteroscedastic errors. *Wseas transactions on mathematics*, 7(8), 351-361.

Habshah, M., Norazan, M. R., and Rahmatullah Imon, A. H. M. (2009). The performance of diagnostic-robust generalized potentials for the identification of multiple high leverage points in linear regression. *Journal of Applied Statistics,* 36(5), 507-520.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986). *Robust Statistics*. New York: Wiley.

Harrington (1965). The Desirability Function. *Industrial Quality Control* 12, 494- 498.

Hettmansperger, T. P., and McKean, J. W. (2010). *Robust nonparametric statistical methods.* CRC Press.

Hill, R. W. (1977). *Robust regression when there are outliers in the carriers*. Unpublished Ph.D. thesis. Harvard University, Boston, MA.

Huber, P. and E. Ronchetti (1981). Robust Statistics, ser. *Wiley Series in Probability and Mathematical Statistics. New York, NY, USA, Wiley-IEEE* 52:54.

Huber, P. J. (2005). *Robust statistics,* John Wiley and Sons.

Hund, E., Massart, D. L., and Smeyers-Verbeke, J. (2002). Robust regression and outlier detection in the evaluation of robustness tests with different experimental designs. *Analytica Chimica Acta,* 463(1), 53-73.

Imon, A.H.M.R. (2005). Identifying multiple influential observations in linear regression. *Journal of Applied Statistics,* 32(9), 929-946.

Imon, A.H.M.R. (2009). Deletion residuals in the detection of heterogeneity ofvariances in linear regression. *Journal of Applied Statistics*, 36, 347–358.

Ismaeel, S. S., Midi, H., & Sani, M. (2021). Robust Multicollinearity Diagnostic Measure For Fixed Effect Panel Data Model. *Malaysian Journal Fundamental Applied. Science*, 17(5), 636-646.

Jauregi, P., Gilmour, S., and Varley, J. (1997). Characterisation of colloidal gas aphrons for subsequent use for protein recovery. *Chemical Engineering, Journal* 65(1): 1-11.

Jawad, A. H., Sahu, U. K., Mastuli, M. S., ALOthman, Z. A., and Wilson, L. D. (2022). Multivariable optimization with desirability function for carbon porosity and methylene blue adsorption by watermelon rind activated carbon prepared by microwave assisted H3PO4. *Biomass Conversion and Biorefinery*, 1-15.

Jeong, I. J., and Kim, K. J. (2009). An interactive desirability function method to multiresponse optimization. *European Journal of Operational Research*, 195(2), 412-426.

Jeong, I. J., Kim, K. J., and Chang, S. Y. (2005). Optimal weighting of bias and variance in dual response surface optimization. *Journal of Quality Technology*, 37(3), 236-247.

Judge, G.G., Griffiths W.E., Hill, R.C., Lutkepohl H. and Lee T.C. (1985). *The Theory and Practice of Econometrics*. Wiley, New York.

Kackar, R. N. (1985). Off-line quality control, parameter design, and the Taguchi method. *Journal of Quality Technology*, 17(4), 176-188.

Khattree, R. (1996). Robust parameter design: A response surface approach. *Journal of Quality Technology*, 28(2), 187-198.

Khedmati, M., and Niaki, S. T. A. (2022). Phase-I robust parameter estimation of simple linear profiles in multistage processes. *Communications in Statistics-Simulation and Computation*, 51(2), 460-485.

Khuri , A. I., and Conlon, M. (1981). Simultaneous optimzation of multiple response represented by polynomial regression functions. *Technometric*, 23(4), 363-375.

Kim, K. J., and Lin, D. K. (1998). Dual response surface optimization: a fuzzy modeling approach. *Journal of Quality Technology*, 30(1), 1-10.

Kim, K. J., and Lin, D. K. (2000). Simultaneous optimization of mechanical properties of steel by maximizing exponential desirability functions. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 49(3), 311-325.

Kim, K. J., and Lin, D. K. (2006). Optimization of multiple responses considering both location and dispersion effects. *European Journal of Operational Research*, 169(1), 133-145.

Kim, Y. J., and Cho, B. R. (2002). Development of priority-based robust design. *Quality Engineering*, 14(3), 355-363.

Kovach, J., Cho, B. R., and Antony, J. (2008). Development of an experiment-based robust design paradigm for multiple quality characteristics using physical programming. *The International Journal of Advanced Manufacturing Technology*, 35(11), 1100-1112.

Kutner, M. H., Nachtsheim, C. J., Neter, J. and Li, W. (2004). A*pplied linear regression models*. 5th edition. New York: MacGRAW-Hill.

Lee, D. H., and Kim, K. J. (2012). Interactive weighting of bias and variance in dual response surface optimization. *Expert Systems with Applications,* 39(5), 5900-5906.

Lee, D.H., Yang, J.K. dan Kim, K.J. (2018). Dual-response optimization using a patient rule induction method. *Quality Engineering,* 30(4), 610-620.

Lee, M. S., and Kim, K. J. (2007). Expected desirability function: consideration of both location and dispersion effects in desirability function approach. *Quality Technology and Quantitative Management,* 4(3), 365-377.

Lee, S. B., Park, C., and Cho, B. R. (2007). Development of a highly efficient and resistant robust design. *International Journal of Production Research,* 45(1), 157-167.

Leroy, A. M. and P. J. Rousseeuw (1987). *Robust regression and outlier detection*. Wiley Series in Probability and Mathematical Statistics, New York: Wiley.

Lin, D. K., and Tu, W. (1995). Dual response surface optimization. *Journal of Quality Technology,* 27(1), 34-39.

MacKinnon, J. G., and White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics,* 29(3), 305-325.

Manojkumar, N., Muthukumaran, C., & Sharmila, G. (2022). A comprehensive review on the application of response surface methodology for optimization of biodiesel production using different oil sources. *Journal of King Saud University-Engineering Sciences*, *34*(3), 198-208.

Marinković, V. (2020). A novel desirability function for multi-response optimization and its application in chemical engineering. *Chemical Industry and Chemical Engineering Quarterly,* 26(3), 309-319.

Maronna, R. A., Martin, R. D., Yohai, V. J., and Salibián-Barrera, M. (2019). *Robust statistics: theory and methods (with R)*. John Wiley and Sons.

Midi, H. (1999). Preliminary estimators for robust non-linear regression estimation. *Journal of Applied Statistics*, *26*(5), 591-600.

Midi, H., Ismaeel, S. S., Arasan, J., and A.,Mohammed. (2021). Simple and fast Generalized-M (GM) estimator and its application to real data set. *Sains Malaysiana,* 50(3), 859-867.

Midi, H., Mustafa, M. S., and Fitrianto, A. (2012). Performance of optimum response surface methodology based on MM-estimator. *International Journal of Mathematical Models and Methods in Applied Sciences* 6(6), 757-764.

Midi, H., Mustafa, M. S., and Fitrianto, A. (2013, April). Augmented approach to desirability function based on MM estimator. In *AIP Conference Proceedings* (Vol. 1522, No. 1, pp. 1240-1247). American Institute of Physics.

Montgomery, D. C., Peck, E. A., and Vining, G. G. (2021). *Introduction to linear regression analysis*. John Wiley and Sons.

Mustafa, M. S. (2015). Robust estimation and outlier detection in response surface methodology. Ph.D. thesis. Institute for Mathematical Research, Universiti Putra Malaysia.

Mustafa, M. S., and Midi, H. (2019). The TSR_MM based on robust location and scales measures in dual response optimization in the presence of outliers and heteroscedastic errors. *ASM Science Journal*, 12(1), 310-319.

Myers, R. H., and Carter, W. H. Jr. (1973). Response surface techniques for dual response systems. *Technometric*, 15(2), 301-317.

Myers, R. H., Khuri, A. I., and Carter, W. H. (1989). Response surface methodology: 1966–l988. *Technometrics,* 31(2), 137-157.

Myers, R. H., Khuri, A. I., Vining, G. G. (1992). Response surface alternative to the Taguchi robust design problem. *The American Statistician,* 46(2), 131-139.

Myers, R. H., Montgomery, D. C., Anderson-Cook, C. M. (2009). *Response surface methodology: process and product optimization using designed experiments*. 2nd Edition. Canada. John Wiley and Sons, Inc.

Park, C., and Cho, B. R. (2003). Development of robust design under contaminated and non-normal data. *Quality Engineering*, 15(3), 463-469.

Park, C., and Leeds, M. (2016). A highly efficient robust design under data contamination. *Computers and Industrial Engineering,* 93, 131-142.

Park, C., Kim, H., and Wang, M. (2022). Investigation of finite-sample properties of robust location and scale estimators. *Communications in Statistics-Simulation and Computation*, 51(5), 2619-2645.

Park, R.E. (1966). Estimation with heteroscedastic error terms. *Econometrica*, 34,888

Pindyck, S. R., and Rubinfeld, L. D. (1997). *Econometric models and econometric forecasts*, 4th Edition. New York: Irwin/McGraw-Hill.

Pote, R. N., Patil, R. K., and Badadhe, A. M. (2022). Optimisation of performance and emission parameters of diesel engine using tyre pyrolysis oil. *Australian Journal of Mechanical Engineering*, 20(4), 1172-1184.

Rabbi, F., Khalil, A., Khan, I., Almuqrin, M. A., Khalil, U., and Andualem, M. (2022). Robust model selection using the out-of-bag bootstrap in linear regression. *Scientific Reports,* 12(1), 1-10.

Rana M.S., H. Midi, and A.H.M.R. Imon, (2008). A Robust Modification of the Goldfeld-Quandt Test for the Detection of Heteroscedasticity in the Presence of Outliers, *Journal of mathematics and Statistics,* 4(4), 277-283.

Rana, Md. S., Midi, H., and Imon A.H.M.R. (2012). Robust wild bootstrap for stabilizing the variance of parameter estimates in heteroscedastic regression models in the presence of outliers. *Mathematical Problem in Engeering.*

Rasheed, B. A., Adnana, R., Saffarib, S. E., and Dano Patia, K. (2014). Robust weighted least squares estimation of regression parameter in the presence of outliers and heteroscedastic errors. *Journal Technology,*71(1), 11-18.

Rashid, A. M., Midi, H., Dhhan, W., & Arasan, J. (2022). Detection of outliers in high-dimensional data using nu-support vector regression. *Journal of Applied Statistics*, *49*(10), 2550-2569.

Rashid, A. M., Midi, H., Slwabi, W. D., and Arasan, J. (2021). An efficient estimation and classification methods for high dimensional data using robust iteratively reweighted SIMPLS algorithm based on nu-support vector regression. *IEEE Access*, *9*, 45955-45967.

Rheem, S., and Oh, S. (2019). Improving the quality of response surface analysis of an experiment for coffee-supplemented milk beverage: I. Data screening at the center point and maximum possible R-Square. *Food Science of Animal Resources*, 39(1), 114.

Riazoshams, H., Midi, H. B., and Sharipov, O. S. (2010). The performance of robust two-stage estimator in nonlinear regression with autocorrelated error. *Communications in Statistics-Simulation and Computation,* 39(6), 1251-1268.

Robinson, T. J., Borror, C. M., & Myers, R. H. (2004). Robust parameter design: a review. *Quality and reliability engineering international*, *20*(1), 81-101.

Rocke, D. M., and Woodruff, D. L. (1996). Identification of outliers in multivariate data. *Journal of the American Statistical Association,* 91(435), 1047-1061.

Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of the American Statistical Association,* 79(388), 871-880.

Rousseeuw, P. J. (1985). Multivariate estimation with high breakdown point. *Mathematical statistics and applications,* 8, 283-297

Rousseeuw, P. J., and Croux C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association,* 88(424), 1273-1283.

Rousseeuw, P. J., and Yohai, V. (1984). *Robust regression by means of S-estimators, robust and nonlinear time series analysis*. Lecture Notes in Statistics.

Ryan, T. P. (1997). *Modern Regression Methods*. NewYork: Wiley.

Sabarish, K. V., & Pratheeba, P. (2020). An experimental analysis on structural beam with Taguchi orthogonal array. *Materials Today: Proceedings*, 22, 874-878.

Shin, S., and Cho, B. R. (2005). Bias-specified robust design optimization and its analytical solutions. *Computers and Industrial Engineering,* 48(1), 129-140.

Simpson, J. R. (1995). New Methods and comparative evaluations for robust and biased-robust regression estimation. Ph.D. thesis. Arizona State University.

Smucler, E., and Yohai, V. J. (2017). Robust and sparse estimators for linear regression models. *Computational Statistics and Data Analysis,* 111, 116-130.

Srikantan, K. S. (1961). Testing for the single outlier in a regression model. *Sankhyā: The Indian Journal of Statistics, Series A*, 251-260.

Stock, J. H., and Watson, M. W. (2008). Heteroskedasticity-robust standard errors for fixed effects panel data regression. *Econometrica,* 76(1), 155-174.

Stromberg, A. J., et al. (2000). The least trimmed differences regression estimator and alternatives. *Journal of the American Statistical Association*, 95(451), 853-864.

Subrahmanyam, A. P. S. V. R., Rao, C. M., and Raju, B. N. (2018). Taguchi based desirability function analysis for the optimization of multiple performance characteristics. *IJMTER*, *5*(5), 2349-9745.

Tabucanon, M. T. (1988). Multiple criteria decision making in industry. *Elsevier Science Limited*, 8.

Taguchi G., Wu Y. (1985). Introduction to OO-Line Quality Control. Nagoya, Japan: Central Japan Quality Control Assoction, 1985.

Tang, L. C., and Xu, K. (2002). A unified approach for dual response surface optimization. *Journal of Quality Technology,* 34(4), 437-447.

Tang, L. N., Ma, Y. Z., Wang, J. J., Ouyang, L. H., & Byun, J. H. (2019). Robust parameter design of supply chain inventory policy considering the uncertainty of demand and lead time. *Scientia Iranica*, *26*(5), 2971-2987.

Tukey, J. W. (1967). A survey of sampling from contaminated distributions. *Contributions to Probability and Statistics*, 448-485.

Uraibi, H. S., Midi, H., Talib, B. A., and Yousif, J. H. (2009). Linear regression model selection based on robust bootstrapping technique. *American Journal of Applied Sciences,* 6(6), 1191.

Uraibi, H., and Midi, H. (2020). Robust variable selection method based on huberized lars-lasso regression. *Economic Computation and Economic Cybernetics Studies and Research,* 54(3).

Van den Boogaart, K. G., Filzmoser, P., Hron, K., Templ, M., and Tolosana-Delgado, R. (2021). Classical and robust regression analysis with compositional data. *Mathematical Geosciences*, 53(5), 823-858.

Vining, G. G., and Myers, R. H. (1990). Combining Taguchi and response surface philosophies: A dual response approach. *Journal of Quality Technology*, 22, 34-45.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 48, 817-838.

Wilcox, R. R. (2005). *Introduction to robust estimation and hypothesis testing*. 2nd edition. The United States of America: Elsevier academic.

Williams, D. C. (2011). Finite sample correction factors for several simple robust estimators of normal standard deviation. *Journal of Statistical Computation and Simulation*, 81(11), 1697-1702.

Wu, C. J., & Hamada, M. S. (2011). *Experiments: planning, analysis, and optimization*. John Wiley & Sons.

Yohai, V. J., and Zamar, R. H. (1988). High breakdown-point estimates of regression by means of the minimization of an efficient scale. *Journal of the American Statistical Association*, 83(402), 406-413.

Yohai, Victor J. "High breakdown-point and high efficiency robust estimates for regression." *The Annals of statistics,* (1987), 642-656.

Zafar, Z., and Aslam, M. (2021). An adaptive weighted least squares ratio approach for estimation of heteroscedastic linear regression model in the presence of outliers. *Communications in Statistics-Simulation and Computation*, 1-11.

Zahariah, S., & Midi, H. (2022). Minimum regularized covariance determinant and principal component analysis-based method for the identification of high leverage points in high dimensional sparse data. *Journal of Applied Statistics*, 1-19.

APPENDICES

Appendix A1

The Catapult Study Data Set (Kim and Lin, 1998)

i	x_1	x_2	x_3	y_{i1}	y_{i2}	y_{i3}
1	-1(0.32)	-1(30)	-1(2.5)	39	42	42
2	-1(0.32)	-1(30)	1(5.5)	80	91	71
3	-1(0.32)	1(90)	-1(2.5)	52	45	44
4	-1(0.32)	1(90)	1(5.5)	97	60	68
5	1(3.68)	-1(30)	-1(2.5)	60	68	53
6	1(3.68)	-1(30)	1(5.5)	113	127	104
7	1(3.68)	1(90)	-1(2.5)	78	65	64
8	1(3.68)	1(90)	1(5.5)	130	75	79
9	-1.682(-0.823)	0(2)	0(2)	59	60	51
10	1.682(4.826)	0(2)	0(2)	115	117	102
11	0(2)	-1.682(-48.46)	0(2)	50	57	43
12	0(2)	1.682(52.46)	0(2)	88	43	49
13	0(2)	0(2)	-1.682(-0.523)	54	60	50
14	0(2)	0(2)	1.682(4.523)	122	119	109
15	0(2)	0(2)	0(2)	87	89	78
16	0(2)	0(2)	0(2)	86	85	79
17	0(2)	0(2)	0(2)	88	87	81
18	0(2)	0(2)	0(2)	89	87	82
19	0(2)	0(2)	0(2)	86	88	79
20	0(2)	0(2)	0(2)	88	90	79**(790)**

Appendix A2

The Printing Process Study Data Set (Box and Draper, 1987)

i	x_1	x_2	x_3	y_{i1}	y_{i2}	y_{i3}
1	-1	-1	-1	34	10	28
2	0	-1	-1	115	116	130
3	1	-1	-1	192	186	263
4	-1	0	-1	82	88	88
5	0	0	-1	44	178	188
6	1	0	-1	322	350	350
7	-1	1	-1	141	110	86
8	0	1	-1	259**(9259)**	251	259
9	1	1	-1	290	280	245
10	-1	-1	0	81	81	81
11	0	-1	0	90	122	93
12	1	-1	0	319	376	376
13	-1	0	0	180	180	154
14	0	0	0	372	372	372
15	1	0	0	541	568	396
16	-1	1	0	288	192	312
17	0	1	0	432	336	513
18	1	1	0	713	725	754
19	-1	-1	1	364	99	199
20	0	-1	1	232	221	266
21	1	-1	1	408	415	443
22	-1	0	1	182	233	182
23	0	0	1	507	515	434
24	1	0	1	846	535	640
25	-1	1	1	236	126	168
26	0	1	1	660	440	403
27	1	1	1	878	991	1161

Appendix A3

The Semiconductor Manufacturing Process Data Set (Shin and Cho, 2005)

Run	Coded units		Quality characteristics of interest coating thickness (μm)			
	Temp (x_1)	Flow rate (x_2)	y (4 replication)			
1	-1	-1	76.3	73.5	74.2	68.8
2	1	-1	70.4	81.2	79.6	76.7
3	-1	1	76.6	72.0	78.5	77.7
4	1	1	72.3	67.5	72.7	75.7
5	-1.414	0	70.6	75.8	71.5	69.9
6	1.414	0	74.1	80.2	77.1	76.2
7	0	-1.414	78.5	68.7	75.3	76.2
8	0	1.414	70.2	76.3	75.9	79.2
9	0	0	74.1	71.8	71.9	72.5
10	0	0	72.1	70.4	74.2	73.3
11	0	0	74.2	69.8	72.2	71.2
12	0	0	70.1	69.3	72.5**(200)**	71.6
13	0	0	69.8	70.6**(300)**	74.1	71.6

Appendix A4

The Colloidal Gas Aphrons Data Set (Jauregi et al., 2005)

Run	(x_1)	(x_2)	(x_2)	(y_1)	(y_2)	(y_3)
1	-1	-1	-1	4.50	0.17	29.0
	-1	-1	-1	4.50	0.26	23.0
	-1	-1	-1	4.50	0.27	25.1
2	1	-1	-1	6.04	0.50	23.0
	1	-1	-1	6.39	0.53**(53.0)**	25.4
	1	-1	-1	6.31	0.55	25.3
3	-1	1	-1	3.81	0.17	22.0
	-1	1	-1	4.09	0.20	27.0
	-1	1	-1	4.08	0.22	26.3
4	1	1	-1	5.67	0.44	25.5
	1	1	-1	5.19	0.40	21.0
	1	1	-1	5.18	0.41	23.5
5	-1	-1	1	4.67	0.32	20.0
	-1	-1	1	4.22	0.32	41.0
	-1	-1	1	4.25	0.33	25.3
6	1	-1	1	6.73	0.57	35.5
	1	-1	1	6.57	0.57	18.0
	1	-1	1	6.61	0.58	23.0
7	-1	1	1	3.40	0.12	43.0
	-1	1	1	4.32	0.28	20.0
	-1	1	1	4.41	0.25	25.3
8	1	1	1	5.72	0.46	19.0
	1	1	1	5.09	0.50	34.0
	1	1	1	5.08	0.53	31.9
9	-1	0	0	4.09	0.27	36.0
	-1	0	0	4.38	0.23	24.0
	-1	0	0	4.41	0.22	26.3
10	1	0	0	5.52	0.52	30.0
	1	0	0	5.39	0.51	24.0
	1	0	0	5.42	0.53	23.0
11	0	-1	0	5.92	0.61	32.0
	0	-1	0	5.93	0.59	23.4
	0	-1	0	5.91	0.62	23.5
12	0	1	0	4.74	0.36	36.0
	0	1	0	4.50	0.30	21.0
	0	1	0	4.62	0.32	23.0
13	0	0	-1	5.01	0.36	27.0
	0	0	-1	4.70	0.25	24.0
	0	0	-1	4.32	0.22	25.3
14	0	0	1	4.94	0.53	38.0
	0	0	1	5.01	0.51	25.0
	0	0	1	5.02	0.53	26.3
15	0	0	0	4.85	0.47	34.0
	0	0	0	4.94	0.46	34.0
	0	0	0	4.96	0.45	32.0**(112)**

Appendix A5

Generalized Studentized Residual Based on MM-estimators

The modified generalized response surface studentized residuals denoted as $\mathrm{MGRS}t_i$. The $\mathrm{MGRS}t_i$ is summarized as follows:

Step 1: Identify several observations which are the suspected outliers by using the MM-estimator to separate the data into the Deletion and Remaining group

Step 2: Fit the linear model of the clean 'R' group and obtain the parameter estimate $\hat{\beta}_R$ given as,

$$\hat{\beta}_R = (X_R^T X_R)^{-1} X_R^T Y_R$$

Step 3: Compute the externally studentized residuals based on MM (t_i^*) for the remaining group, R is given as follows:

$$t_i^* = \frac{y_i - x_i^T \hat{\beta}_R}{\hat{\sigma}_{R-i}\sqrt{1 - w_{ii(R)}}} = \frac{\hat{\varepsilon}_{i(R)}}{\hat{\sigma}_{R-i}\sqrt{1 - w_{ii(R)}}} \qquad (1)$$

where $w_{ii(R)}$ is the diagonal hat matrix as follows

$$w_{ii(R)} = x_i^T (X_R^T X_R)^{-1} x_i, \;\; i = 1,2,\dots,n$$

Step 4: By using Rao (1995) results, the set R with an additional point i is defined as

$$w_{ii(R+i)} = x_i^T (X_R^T X_R + x_i x_i^T)^{-1} x_i, = \frac{w_{ii(R)}}{1 + w_{ii(R)}}$$

and

$$\begin{aligned}\hat{\beta}_{R+i} &= (X_R^T X_R + x_i x_i^T)^{-1} (X_R^T Y_R + x_i y_i) \\ &= \hat{\beta}_R + \frac{(X_R^T X_R)^{-1}}{1 + w_{ii(R)}} \hat{\varepsilon}_{i(R)}\end{aligned}$$

Step 5: Consider the formulation of the externally studentized residual which is defined in Equation (1) and compute the externally studentized residual for $i \notin$ as follows:

$$t_i^* = \frac{y_i - x_i^T \hat{\beta}_{(R+i)}}{\hat{\sigma}_R \sqrt{1 + w_{ii(R)}}} = \frac{\hat{\varepsilon}_{i(R)}}{\hat{\sigma}_R \sqrt{1 + w_{ii(R)}}} \tag{2}$$

Step 6: Compute the modified generalized response surface studentized residuals (MGRSt_i) for the whole dataset which is formulated by combining Equation (1) and Equation (2) as follows

$$\text{MGRS}t_i^* = \begin{cases} \dfrac{\hat{\varepsilon}_{i(R)}}{\hat{\sigma}_{R-i}\sqrt{1 - w_{ii(R)}}} & \text{for} \quad i \in R, \\ \dfrac{\hat{\varepsilon}_{i(R)}}{\hat{\sigma}_R \sqrt{1 + w_{ii(R)}}} & \text{for} \quad i \in D, \end{cases}$$

Monte Carlo Simulation Results

In this section, a series of Monte Carlo simulation study that is designed to evaluate the performance of our new proposed methods namely, the MGRSt$_i$, the MM-studentized residuals (r_{MM}), and the MM-deletion studentized residuals (t_{MM}) as compared to some existing diagnostics, namely the OLS-studentized residuals (r_{OLS}), and OLS-deletion studentized residuals (t_{OLS}) in the identification of outlier. The performances of these measures are evaluated based on the rate of correct detection of outliers and the rate of masking and swamping effects. A good measure is the one that has a higher percentage of correction detection of outliers and smaller rate of masking and swamping effects. In each experiment, three samples of size $n = 27{,}36{,}45$ and different number of outlier points ($\alpha = 1{,}2{,}3{,}4$) are considered. Here we considered 3 factors. The variable Y is generated according to a normal distribution with $\mu(x_i)$ and $\sigma(x_i)$ at each control factor settings $x_i = (x_{i1}, x_{i2}, x_{i3})$, $i = 1{,}2, \ldots, 3$. The total number of iterations is 1000, each having n x 3 design points and n responses. Following Park and Cho (2003), $\mu(x_i)$ and $\sigma(x_i)$ are generated as follows:

$$\mu(x_i) = 50 + (x_1^2 + x_2^2 + x_3^2)$$
$$\sigma(x_i) = 10 + \{(x_1 - 0.5)^2 + x_2^2 + x_3^2\}$$

In order to generate outlier points in a dataset, the first observations of the regular data (x_i, y_i) is replaced with a new y value which is generated as random $N(150{,}20)$. The results of the Monte Carlo simulation study are summarized in Tables 1 which represents the percentage of correct detection of outliers and the masking and swamping rates at different number of outlier and different sample sizes.

The results of the Monte Carlo simulation study are summarized in Table 1 which represents the percentage of correct detection of outliers and the masking and swamping rates at different number of outlier and different sample sizes. The results in Table 1 clearly indicate that all detection measures able to correctly detect one outlier in the data. However, for 2 or more outliers, the performance of r_{OLS} and t_{OLS} become worst especially for small n. Both measures have lower percentage of correct detection and higher masking rates compared to other measures. On the other hand, the r_{MM} performs better than the t_{MM}, but cannot outperform the $\mathrm{MGRS}t_i$. The $\mathrm{MGRS}t_i$ consistently having the highest percentage of correct detection and smallest rate of masking and swamping effect.

Table 9.1 : Percentage of Correctly Identified Outliers, Masking and Swamping for Simulation Study

No. of Outlier	n	% Correct detection					% Masking					% Swamping				
		r_{OLS}	t_{OLS}	r_{MM}	t_{MM}	$MGRSt_i$	r_{OLS}	t_{OLS}	r_{MM}	t_{MM}	$MGRSt_i$	r_{OLS}	t_{OLS}	r_{MM}	t_{MM}	$MGRSt_i$
1	27	100	100	100	100	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.6	0.1	0.2
	36	100	100	100	100	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.6	0.1	0.2
	45	100	100	100	100	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.1
2	27	45	44.8	100	99.5	100	55	55.2	0.0	0.5	0.0	0.0	0.0	0.9	0.0	0.1
	36	90.3	89.2	100	100	100	9.7	10.2	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.1
	45	100	100	100	100	99.7	0.0	0.0	0.0	0.0	0.3	0.0	0.0	1.0	0.0	0.1
3	27	0.0	0.0	97.9	94.4	100	100	100	2.1	5.6	0.0	0.0	0.0	0.4	0.0	0.0
	36	38.7	38.3	100	99.9	100	61.3	61.7	0.0	0.1	0.0	0.0	0.0	0.6	0.0	0.0
	45	86.9	86.2	100	100	100	13.1	13.8	0.0	0.0	0.0	0.0	0.0	0.8	0.0	0.0
4	27	13.7	13.5	81.2	79.5	99.8	86.3	86.5	18.8	20.2	0.2	0.0	0.0	0.1	0.0	0.0
	36	13.3	12.9	100	99.7	100	86.7	87.2	0.0	0.3	0.0	0.0	0.0	0.4	0.0	0.0
	45	52.1	51.8	100	100	100	47.9	48.2	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0

Numerical Example

An artificial data set with 20 observations, three predictors, and the output response variable data set introduced by Myers and Montgomery (2002) is used. In this data set, a central composite design (CCD) with six center points is applied to describe the model. All the controllable variables are $-1.682 \leq x_i \leq 1.682, i = 1,2,3$. In order to see the effect of outliers, this data set has been modified to contain three outliers such that the 7th, 10th, and 13th observations are replaced with a large values which are generated from N(150, 20). The experimental data are given in Table 3. The results of Appendix A6 point out that the r_{OLS} based fails to identify any point as outlier. The t_{OLS} can identify one outlier. However, r_{MM} and t_{MM} can identify three outliers, but wrongly detect the true outliers. The MGRSt_i successfully detect the tree observations as genuine outliers.

Table 9.2 : Five measures for Artificial Data

Index	$\boldsymbol{r_{OLS}}$	$\boldsymbol{t_{OLS}}$	$\boldsymbol{r_{MM}}$	$\boldsymbol{t_{MM}}$	$\boldsymbol{MGRSt_i}$
(cut of point)	**(2.5)**	**(2.5)**	**(2.5)**	**(2.5)**	**(3.58)**
1	0.62	0.59	0.54	-0.12	0.29
2	-1.45	-1.55	**14.47**	**-12.07**	0.38
3	0.13	0.12	0.23	-0.05	0.94
4	-0.43	-0.42	1.39	-0.29	1.14
5	0.43	0.41	0.87	0.19	3.05
6	-0.13	-0.12	0.23	-0.05	1.68
7	1.45	1.55	0.46	-0.10	**4.59**
8	-0.62	-0.60	0.08	0.02	1.0
9	-1.43	-1.52	**11.45**	**-3.76**	-0.03
10	1.44	1.53	0.60	0.13	**4.53**
11	0.29	0.28	0.25	-0.05	1.32
12	-0.29	-0.27	0.86	0.18	2.17
13	0.62	0.60	0.96	0.20	0.52
14	-0.62	-0.60	0.34	-0.07	-1.00
15	-0.41	-0.39	0.30	0.06	0.23
16	-0.44	-0.42	0.19	0.04	0.15
17	-0.40	-0.38	0.34	0.07	0.25
18	2.48	**3.80**	**10.96**	**3.40**	**5.57**
19	-0.76	-0.74	0.99	-0.21	-0.69
20	-0.48	-0.46	0.02	0.00	0.03

Appendix A6

Artificial Data set

Run	x_1	x_2	x_3	y
1	-1	-1	-1	49.80
2	1	-1	-1	52.82
3	-1	1	-1	52.98
4	1	1	-1	54.34
5	-1	-1	1	58.03
6	1	-1	1	56.04
7	-1	1	1	53.09
8	1	1	1	56.42
9	-1.682	0	0	50.03
10	1.682	0	0	52.56
11	0	-1.682	0	53.86
12	0	1.682	0	58.71
13	0	0	-1.682	51.10
14	0	0	1.682	49.39
15	0	0	0	52.41
16	0	0	0	51.94
17	0	0	0	52.94
18	0	0	0	50.79
19	0	0	0	46.90
20	0	0	0	51.22

Appendix B

R Programming Code

Dual Response Optimization Technique

```
rm(list=ls(all=TRUE))
set.seed(6)
sig1<-list()
mu1<-list()
y<-list()
mu2<-list()
sg2<-list()
mui3<-list()
mui4<-list()
mui5<-list()
for(j in 1:5){
x1<-c(-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-
1,0,1)
x2<-c(-1,-1,-1,0,0,0,1,1,1,-1,-1,-1,0,0,0,1,1,1,-1,-1,-
1,0,0,0,1,1,1)
x3<-c(-1,-1,-1,-1,-1,-1,-1,-1,-
1,0,0,0,0,0,0,0,0,0,1,1,1,1,1,1,1,1,1)
mu<-500+(x1+x2+x3)^2 +x1+x2+x3
sig<-sqrt(100+(x1+x2+x3)^2 +x1+x2+x3)



q<-matrix(0,nrow=27,ncol=3)
for(i in 1:27){
q[i,]<-rnorm(3,mu[i],sig[i])
}
ybar<-apply(q,1,mean)
s<-apply(q,1,var)
p<-round(q,0)
colnames(p)<-c("y1","y2","y3")
y[[j]]<-cbind(x1,x2,x3,p,ybar,s)
}


mu1[[j]]<-
c(lm(ybar~(x1+x2+x3)^2+I(x1^2)+I(x2^2)+I(x3^2),data=y)$coef)
sig1[[j]]<-c(lm(s~(x1+x2+x3)^2+I(x1^2)+I(x2^2)+I(x3^2),
data=y)$coef)



#for(j in 1:5){
mn<-function(x){
mu1[[j]][1]+mu1[[j]][2]*x[1]+mu1[[j]][3]*x[2]+mu1[[j]][4]*x[3]+m
u1[[j]][5]*x[1]^2+mu1[[j]][6]*x[2]^2+mu1[[j]][7]*x[3]^2+
mu1[[j]][8]*x[1]*x[2]+mu1[[j]][9]*x[1]*x[3]+mu1[[j]][10]*x[2]*x[
3]
}
```

```
sigma<-function(x){
sig1[[j]][1]+sig1[[j]][2]*x[1]+sig1[[j]][3]*x[2]+sig1[[j]][4]*x[
3]+sig1[[j]][5]*x[1]^2+sig1[[j]][6]*x[2]^2+sig1[[j]][7]*x[3]^2+
sig1[[j]][8]*x[1]*x[2]+sig1[[j]][9]*x[1]*x[3]+sig1[[j]][10]*x[2]
*x[3]
}


fn1<-function(x){
(sig1[[j]][1]+sig1[[j]][2]*x[1]+sig1[[j]][3]*x[2]+sig1[[j]][4]*x
[3]+sig1[[j]][5]*x[1]^2+sig1[[j]][6]*x[2]^2+sig1[[j]][7]*x[3]^2+
sig1[[j]][8]*x[1]*x[2]+sig1[[j]][9]*x[1]*x[3]+sig1[[j]][10]*x[2]
*x[3])^2+

mu1[[j]][1]+mu1[[j]][2]*x[1]+mu1[[j]][3]*x[2]+mu1[[j]][4]*x[3]+m
u1[[j]][5]*x[1]^2+mu1[[j]][6]*x[2]^2+mu1[[j]][7]*x[3]^2+
mu1[[j]][8]*x[1]*x[2]+mu1[[j]][9]*x[1]*x[3]+mu1[[j]][10]*x[2]*x[
3]-500)
}

x.lb<-c(-1,-1,-1)
x.ub<-c(1,1,1)
ineqc1<-function(x){
z1<-x[1]
z2<-x[2]
z3<-x[3]
return(c(z1,z2,z3))
}
result<-solnp(c(-1,-1,-1),fun = fn1,ineqfun =ineqc1, ineqLB =
x.lb, ineqUB=x.ub)
mu2[[j]]<-mn(c(result$pars))
sg2[[j]]<-sigma(c(result$pars))
```

Robust reweighted least square

```
library(rrcov)
library(MASS)
library(robustbase)
library(robust)
x1=c(-1,1,-1,1,-1.414,1.414,0,0,0,0,0,0,0)
x2=c(-1,-1,1,1,0,0,-1.414,1.414,0,0,0,0,0)
x3<-x1^2
x4<-x2^2
x5<-x1*x2
y1<-
c(76.3,70.4,76.6,72.3,70.6,74.1,78.5,70.2,74.1,72.1,74.2,7
0.1,69.8)
y2<-
c(73.5,81.2,72,67.5,75.8,80.2,68.7,76.3,71.8,70.4,69.8,69.
3,70.6)
y3<-
c(80,79.6,78.5,72.7,71.5,77.1,75.3,75.9,71.9,74.2,72.2,72.
5,74.1)
```

```
y4<-
c(68.8,76.7,77.7,75.7,69.9,76.2,76.2,79.2,72.5,73.3,71.2,7
1.6,71.6)
x<-cbind(x1,x2,x3)
q<-data.frame(y1,y2,y3,y4)
ybar<-apply(q,1,CovMMest)
d1=matrix(0,13,1)
d2=matrix(0,13,1)
for(i in 1:13){
s1=ybar[[i]]
loc=getCenter(s1)
scl=getCov(s1)
d1[i,1]=loc
d2[i,1]=scl

}
data2=cbind(d1,d2)
y<-d1
y2<-sqrt(d2)
n=dim(x)[1]
pp=rep(0,n)
D=0
cov2=lmrob(d1~(x1+x2)^2+I(x1^2)+I(x2^2),method          =
"MM")$residuals
cov1<-abs(cov2)
cutoff.rd=3

for(i in 1:n){
if(cov1[i] > cutoff.rd)
D=c(D,i)}

X=as.matrix(cbind(1,x))
D=c(D)
D=D[-1]
if(length(D)==0){
h=matrix(hat(X), n, 1)

for(i in 1:n){
pp[i] = h[i,1]/(1-h[i,1])}}
print(pp)
else{
for(i in 1:n){
if(cov1[i] > cutoff.rd){
pp[i]=(t(X[i,])%*%ginv(t(X[-D,])%*% X[-D,]) %*% X[i,])}
else{
pp[i]=(t(X[i,])%*%ginv(t(X[-D,])%*% X[-D,]) %*% X[i,])/(1-
(t(X[i,] )%*%
ginv(t(X[-D,])%*% X[-D,]) %*% X[i,]))}
MAD.pp=median(abs(pp-median(pp)))/0.6745
cutoff=median(pp)+3*MAD.pp
pp=matrix(pp,n,1)}}
```

```
w11 <-abs(cutoff/pp)
vecone<- c(rep(1, length(pp)))
w111 <- pmin(vecone, w11)
cut<-list(pp,w111)
rr=lmrob(d1~(x1+x2)^2+I(x1^2)+I(x2^2),method              =
"MM")$residuals
res=abs(rr)
shat=lmrob(res~(x1+x2)^2+I(x1^2)+I(x2^2),method            =
"MM")$coeff
snhat=shat[1]+shat[2]*x1+shat[3]*x2+shat[4]*x3+shat[4]*x4+
shat[5]*x5
w1=1/(snhat^2)
wfinal<-w1*w12
fitMM=lmrob(d1~(x1+x2)^2+I(x1^2)+I(x2^2),weights=wfinal,me
thod = "MM")
print(fitMM)
}
```

Desirability Function

```
library(rrcov)
set.seed(123)
for(j in 1:100){
  x1=c(-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-
1,0,1)
  x2=c(-1,-1,-1,0,0,0,1,1,1,-1,-1,-1,0,0,0,1,1,1,-1,-1,-
1,0,0,0,1,1,1)
  x3=c(-1,-1,-1,-1,-1,-1,-1,-1,-
1,0,0,0,0,0,0,0,0,0,1,1,1,1,1,1,1,1,1)
  x4=c(-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-
1,0,1)
  x5=c(-1,-1,-1,0,0,0,1,1,1,-1,-1,-1,0,0,0,1,1,1,-1,-1,-
1,0,0,0,1,1,1)
  x6=c(-1,-1,-1,-1,-1,-1,-1,-1,-
1,0,0,0,0,0,0,0,0,0,1,1,1,1,1,1,1,1,1)
  x7=c(-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-
1,0,1)
  x8=c(-1,-1,-1,0,0,0,1,1,1,-1,-1,-1,0,0,0,1,1,1,-1,-1,-
1,0,0,0,1,1,1)
  x9=c(-1,-1,-1,-1,-1,-1,-1,-1,-
1,0,0,0,0,0,0,0,0,0,1,1,1,1,1,1,1,1,1)

  mu1=500+(x1+x2+x3)^2+x1+x2+x3
  sig1=sqrt(100+(x1+x2+x3)^2+x1+x2+x3)

  mu2=500+(x4+x5+x6)^2+x4+x5+x6
  sig2=sqrt(100+(x4+x5+x6)^2+x4+x5+x6)

  mu3=500+(x7+x8+x9)^2+x7+x8+x9
  sig3=sqrt(100+(x7+x8+x9)^2+x7+x8+x9)

  mu4=500+(x1+x2+x3)^2+x1+x2+x3
  sig4=sqrt(100+(x1+x2+x3)^2+x1+x2+x3)
```

```
 mu5=500+(x4+x5+x6)^2+x4+x5+x6
 sig5=sqrt(100+(x4+x5+x6)^2+x4+x5+x6)

 mu6=500+(x7+x8+x9)^2+x7+x8+x9
 sig6=sqrt(100+(x7+x8+x9)^2+x7+x8+x9)

 q1=matrix(0,nrow=27,ncol=6)
 for(i in 1:27){
      q1[i,]=rnorm(6,mu1[i],sig1[i])
      q1[27]=rnorm(1,50,10^2)

 }

q2=matrix(0,nrow=27,ncol=6)
 for(i in 1:27){
      q2[i,]=rnorm(6,mu2[i],sig2[i])
      q2[27]=rnorm(1,50,10^2)

 }
q3=matrix(0,nrow=27,ncol=6)
 for(i in 1:27){
      q3[i,]=rnorm(6,mu3[i],sig3[i])
       q3[27]=rnorm(1,50,10^2)

 }
q4=matrix(0,nrow=27,ncol=6)
 for(i in 1:27){
      q4[i,]=rnorm(6,mu4[i],sig4[i])

 }
q5=matrix(0,nrow=27,ncol=6)
 for(i in 1:27){
      q5[i,]=rnorm(6,mu5[i],sig5[i])

 }
q6=matrix(0,nrow=27,ncol=6)
 for(i in 1:27){
      q6[i,]=rnorm(6,mu6[i],sig6[i])

 }


d11=round(data.frame(q1),0)
d22=round(data.frame(q2),0)
d33=round(data.frame(q3),0)
d44=round(data.frame(q4),0)
d55=round(data.frame(q5),0)
d66=round(data.frame(q6),0)

d111=cbind(d11[,1],d22[,1],d33[,1],d44[,1],d55[,1],d66[,1])
d222=cbind(d11[,2],d22[,2],d33[,2],d44[,2],d55[,2],d66[,3])
d333=cbind(d11[,3],d22[,3],d33[,3],d44[,3],d55[,3],d66[,3])
d444=cbind(d11[,4],d22[,4],d33[,4],d44[,4],d55[,4],d66[,4])
d555=cbind(d11[,5],d22[,5],d33[,5],d44[,5],d55[,5],d66[,5])
d666=cbind(d11[,6],d22[,6],d33[,6],d44[,6],d55[,6],d66[,6])
```

```
ybar1=round(apply(d111,1,mean),1)
ybar2=round(apply(d222,1,mean),1)
ybar3=round(apply(d333,1,mean),1)
ybar4=round(apply(d444,1,mean),1)
ybar5=round(apply(d555,1,mean),1)
ybar6=round(apply(d666,1,mean),1)


s1=round(apply(d111,1,sd),1)
s2=round(apply(d222,1,sd),1)
s3=round(apply(d333,1,sd),1)
s4=round(apply(d444,1,sd),1)
s5=round(apply(d555,1,sd),1)
s6=round(apply(d666,1,sd),1)

newdata=data.frame(cbind(x1,x2,x3,ybar1,ybar2,ybar3,ybar4,ybar5,
ybar6,s1,s2,s3,s4,s5,s6))

cr.out1=lm(ybar1
~x1+x2+x3+x1:x2+x1:x3+x2:x3+I(x1^2)+I(x2^2)+I(x3^2),data=newdata
)
cr.out2=lm(ybar2~x1+x2+x3+x1:x2+x1:x3+x2:x3+I(x1^2)+I(x2^2)+I(x3
^2),data=newdata)
cr.out3=lm(ybar3~x1+x2+x3+x1:x2+x1:x3+x2:x3+I(x1^2)+I(x2^2)+I(x3
^2),data=newdata)
cr.out4=lm(ybar4
~x1+x2+x3+x1:x2+x1:x3+x2:x3+I(x1^2)+I(x2^2)+I(x3^2),data=newdata
)
cr.out5=lm(ybar5~x1+x2+x3+x1:x2+x1:x3+x2:x3+I(x1^2)+I(x2^2)+I(x3
^2),data=newdata)
cr.out6=lm(ybar6~x1+x2+x3+x1:x2+x1:x3+x2:x3+I(x1^2)+I(x2^2)+I(x3
^2),data=newdata)
cr.out7=lm(s1~x1+x2+x3+x1:x2+x1:x3+x2:x3+I(x1^2)+I(x2^2)+I(x3^2)
,data=newdata)
cr.out8=lm(s2~x1+x2+x3+x1:x2+x1:x3+x2:x3+I(x1^2)+I(x2^2)+I(x3^2)
,data=newdata)
cr.out9=lm(s3~x1+x2+x3+x1:x2+x1:x3+x2:x3+I(x1^2)+I(x2^2)+I(x3^2)
,data=newdata)
cr.out10=lm(s4~x1+x2+x3+x1:x2+x1:x3+x2:x3+I(x1^2)+I(x2^2)+I(x3^2
),data=newdata)
cr.out11=lm(s5~x1+x2+x3+x1:x2+x1:x3+x2:x3+I(x1^2)+I(x2^2)+I(x3^2
),data=newdata)
cr.out12=lm(s6~x1+x2+x3+x1:x2+x1:x3+x2:x3+I(x1^2)+I(x2^2)+I(x3^2
),data=newdata)



dat1=matrix(0,10,1)
for (i in 1:10)
{      dat1[i,]=cr.out1$coeff[i]
              }

dat2=matrix(0,10,1)
for (i in 1:10)
{      dat2[i,]=cr.out2$coeff[i]
              }
```

```
dat3=matrix(0,10,1)
for (i in 1:10)
{      dat3[i,]=cr.out3$coeff[i]
              }
dat4=matrix(0,10,1)
for (i in 1:10)
{      dat4[i,]=cr.out4$coeff[i]
              }

dat5=matrix(0,10,1)
for (i in 1:10)
{      dat5[i,]=cr.out5$coeff[i]
              }

dat6=matrix(0,10,1)
for (i in 1:10)
{      dat6[i,]=cr.out6$coeff[i]
              }

dat7=matrix(0,10,1)
for (i in 1:10)
{      dat7[i,]=cr.out7$coeff[i]
              }

dat8=matrix(0,10,1)
for (i in 1:10)
{      dat8[i,]=cr.out8$coeff[i]
              }

dat9=matrix(0,10,1)
for (i in 1:10)
{      dat9[i,]=cr.out9$coeff[i]
              }

dat10=matrix(0,10,1)
for (i in 1:10)
{      dat10[i,]=cr.out10$coeff[i]
              }

dat11=matrix(0,10,1)
for (i in 1:10)
{      dat11[i,]=cr.out11$coeff[i]
              }

dat12=matrix(0,10,1)
for (i in 1:10)
{      dat12[i,]=cr.out12$coeff[i]
              }


max=-1
min=0
data1=c()
data2=c()
data3=c()
str1=NULL
str2=NULL
```

```
n1=NULL

for(i in seq(from=-1,to=1,by=0.1))
{
  for(j in seq(from=-1,to=1,by=0.1))
  {
    for(k in seq(from=-1,to=1,by=0.1))
    {
v1=
dat1[1]+dat1[2]*i+dat1[3]*j+dat1[4]*k+dat1[5]*i^2+dat1[6]*j^2+da
t1[7]*k^2+dat1[8]*i*j+dat1[9]*i*k+dat1[10]*j*k

v2=
dat2[1]+dat2[2]*i+dat2[3]*j+dat2[4]*k+dat2[5]*i^2+dat2[6]*j^2+da
t2[7]*k^2+dat2[8]*i*j+dat2[9]*i*k+dat2[10]*j*k


v3=dat3[1]+dat3[2]*i+dat3[3]*j+dat3[4]*k+dat3[5]*i^2+dat3[6]*j^2
+dat3[7]*k^2+dat3[8]*i*j+dat3[9]*i*k+dat3[10]*j*k


v4=dat4[1]+dat4[2]*i+dat4[3]*j+dat4[4]*k+dat4[5]*i^2+dat4[6]*j^2
+dat4[7]*k^2+dat4[8]*i*j+dat4[9]*i*k+dat4[10]*j*k


v5=dat5[1]+dat5[2]*i+dat5[3]*j+dat5[4]*k+dat5[5]*i^2+dat5[6]*j^2
+dat5[7]*k^2+dat5[8]*i*j+dat5[9]*i*k+dat5[10]*j*k


v6=dat6[1]+dat6[2]*i+dat6[3]*j+dat6[4]*k+dat6[5]*i^2+dat6[6]*j^2
+dat6[7]*k^2+dat6[8]*i*j+dat6[9]*i*k+dat6[10]*j*k

v7=
dat7[1]+dat7[2]*i+dat7[3]*j+dat7[4]*k+dat7[5]*i^2+dat7[6]*j^2+da
t7[7]*k^2+dat7[8]*i*j+dat7[9]*i*k+dat7[10]*j*k

v8=
dat8[1]+dat8[2]*i+dat8[3]*j+dat8[4]*k+dat8[5]*i^2+dat8[6]*j^2+da
t8[7]*k^2+dat8[8]*i*j+dat8[9]*i*k+dat8[10]*j*k

v9=
dat9[1]+dat9[2]*i+dat9[3]*j+dat9[4]*k+dat9[5]*i^2+dat9[6]*j^2+da
t9[7]*k^2+dat9[8]*i*j+dat9[9]*i*k+dat9[10]*j*k

v10=
dat10[1]+dat10[2]*i+dat10[3]*j+dat10[4]*k+dat10[5]*i^2+dat10[6]*
j^2+dat10[7]*k^2+dat10[8]*i*j+dat10[9]*i*k+dat10[10]*j*k

v11=
dat11[1]+dat11[2]*i+dat11[3]*j+dat11[4]*k+dat11[5]*i^2+dat11[6]*
j^2+dat11[7]*k^2+dat11[8]*i*j+dat11[9]*i*k+dat11[10]*j*k

v12=
dat12[1]+dat12[2]*i+dat12[3]*j+dat12[4]*k+dat12[5]*i^2+dat12[6]*
j^2+dat12[7]*k^2+dat12[8]*i*j+dat12[9]*i*k+dat12[10]*j*k
```

```
#----------------------------------------------
if (d1>=0 & d1<=1 & d2>=0 & d2<=1 & d3>=0 & d3<=1 & d4>=0 & d4<=1
& d5>=0 & d5<=1 & d6>=0 & d6<=1 & d7>=0 & d7<=1& d8>=0 & d8<=1 &
d9>=0 & d9<=1 & d10>=0 & d10<=1 & d11>=0 & d11<=1 & d12>=0 &
d12<=1)
       {
D=mean(c(abs(log(d1)),abs(log(d2)),abs(log(d3)),abs(log(d4)),abs
(log(d5)),abs(log(d6)))
S=mean(c(abs(log(d7)),abs(log(d8)),abs(log(d9)),abs(log(d10)),ab
s(log(d11)),abs(log(d2))))
                    DS=exp(-((0.9*D)+(0.1*S)))
                    data1=c(data1,D)
                    data2=c(data2,S)
                    data3=c(data3,DS)
                    str1=rbind(str1,c(d1,d2,d3,d4,d5,d6))
                    str2=rbind(str2,c(d7,d8,d9,d10,d11,d12))
                    v111=v1
                    v22=v2
                    v33=v3
                    v44=v4
                    v55=v5
                    v66=v6
                    w11=v7
                    w22=v8
                    w33=v9
                    w44=v10
                    w55=v11
                    w66=v12
       n1=rbind(n1,c(i,j,k,v111,v22,v33,v44,v55,v66,w11,w22,w33,w
44,w55,w66))
       }
#-------------------------------------------------
       result=cbind(data1,data2,data3,str1,str2,n1)
#-------------------------------------------------
    }
  }
}

colnames(result, do.NULL = FALSE)
col2=as.matrix(result)
#max(data3)
col3=result[which.max(data3),]
#print(r)
n2=rbind(n2,c(col3))
}
```

BIODATA OF STUDENT

Nasuhar Ab. Aziz was born in Kota Bharu, Kelantan in the year of 1982. She received her Bachelor degree in Statistics from UiTM Shah Alam, Selangor in 2007. In 2012, she received her Master degree in Statistics from Universiti Malaya (UM). She is currently a lecturer at the Faculty of Computer & Mathematical Sciences, UiTM Kelantan and have more than 10 years experiences teaching statistics to statistics students. Her interest in Robust Statistics started to develop during her study in PhD at UPM. She has published and co-authored a number of papers in journals.

LIST OF PUBLICATIONS

Journal Papers

Nasuhar Ab. Aziz & Habshah Midi (2022). Penalty Function Optimization in Dual Response Surfaces Based on Decision Maker's Preference and Its Application to Real Data. *Symmetry*, *14*(3), 601.

Habshah Midi & **Nasuhar Ab. Aziz** (2019). High Breakdown Estimator for Dual Response Optimization in the Presence of Outliers. *Sains Malaysiana*, *48*(8), 1771-1776.

Habshah Midi & **Nasuhar Ab. Aziz** (2018). Augmented Desirability Function for Multiple Responses with Contaminated Data. *Journal of Engineering and Applied Sciences*, *13*(16), 6626-6633.

Conference Papers

Nasuhar, AA. and Habshah M. (2017). Simultaneous Optimizing Multiple Responses with Contaminated Data. *Proceeding of the 4th International Conference on Mathematical Sciences and Computer Engineering (ICMSCE)*, pp. 56 – 63, ISBN: 978-967-13798-9-9.

UNIVERSITI PUTRA MALAYSIA

STATUS CONFIRMATION FOR THESIS / PROJECT REPORT AND COPYRIGHT

ACADEMIC SESSION : Second Semester 2022/2023

TITLE OF THESIS / PROJECT REPORT :

ROBUST DIAGNOSTICS AND ESTIMATION FOR DUAL RESPONSE SURFACE FUNCTION WITH HETEROSCEDASTIC ERRORS USING NEW OPTIMIZATION TECHNIQUE IN THE PRESENCE OF OUTLIERS

NAME OF STUDENT: NASUHAR BINTI AB. AZIZ

I acknowledge that the copyright and other intellectual property in the thesis/project report belonged to Universiti Putra Malaysia and I agree to allow this thesis/project report to be placed at the library under the following terms:

1. This thesis/project report is the property of Universiti Putra Malaysia.
2. The library of Universiti Putra Malaysia has the right to make copies for educational purposes only.
3. The library of Universiti Putra Malaysia is allowed to make copies of this thesis for academic exchange.

I declare that this thesis is classified as :

*Please tick (√)

☐	**CONFIDENTIAL**	(Contain confidential information under Official Secret Act 1972).
☐	**RESTRICTED**	(Contains restricted information as specified by the organization/institution where research was done).
☐	**OPEN ACCESS**	I agree that my thesis/project report to be published as hard copy or online open access.

This thesis is submitted for :

☐	**PATENT**	Embargo from_____________ until _____________ (date) (date)

Approved by:

_____________________	_____________________
(Signature of Student) New IC No/ Passport No.:	(Signature of Chairman of Supervisory Committee) Name:
Date :	Date :

[Note : If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization/institution with period and reasons for confidentially or restricted.]