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 $(\pm 1, \pm 1, \dots, \pm 1)$ (Montgomery et al., 2021). For example, for $k = 2$;

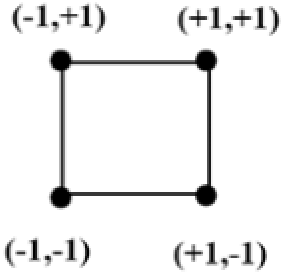

Figure 1.1 : The CCD with 2^k corner points

(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 $(\pm\alpha, 0, \dots, 0), (0, \pm\alpha, \dots, 0)$, etc. For example, for $k = 2$;

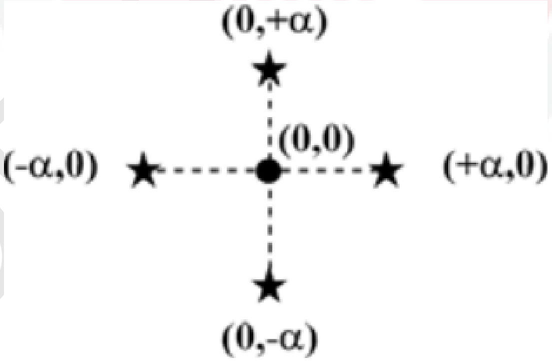

Figure 1.2 : The CCD with axial points $\alpha > 1.0$

(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 $(0,0, \dots, 0)$. For example, for $k = 2$;

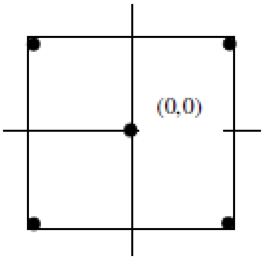

Figure 1.3 : The CCD with corner, axial and center points

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.

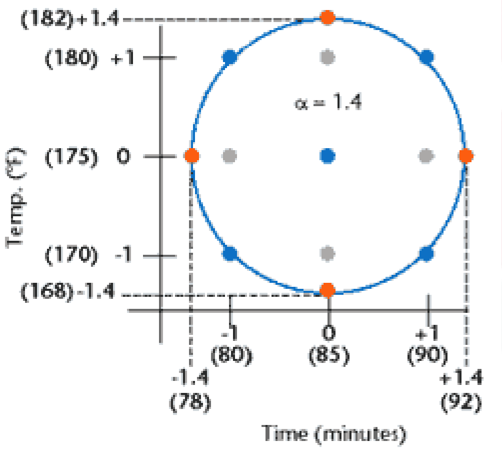

Figure 1.4 : Two variables CCC design

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.

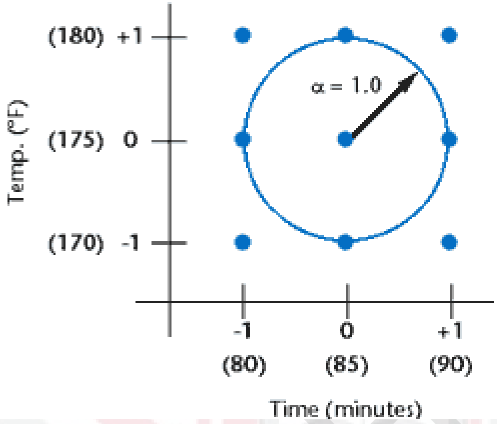

Figure 1.5 : Two variables CCF design

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).

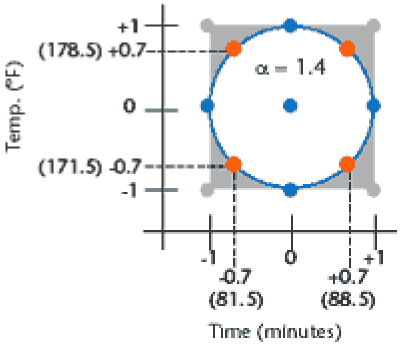

Figure 1.6 : Two variables CCI design

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, $\mathrm{sd}(\hat{y}_l)$. The overall desirability functions for the predicted responses, $\hat{y}_l$ and the standard deviation of the predicted responses, $\mathrm{sd}(\hat{y}_l)$ 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, $\mathrm{sd}(\hat{y}_l)$ (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, $\mathrm{sd}(\hat{y}_l)$ 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 $x_{ij} = (x_{i1}, x_{i2}, \dots, x_{ip})$ is written as $x_i = (0,0, \dots 0)$. 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.

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.