



UNIVERSITI PUTRA MALAYSIA

***ROBUST DETECTION MEASURES AND ROBUST PARAMETER
ESTIMATION METHODS IN CIRCULAR UNIVARIATE AND SIMPLE
CIRCULAR REGRESSION MODEL***

EHAB ABDULSALAM MAHMOOD

FS 2018 9



**ROBUST DETECTION MEASURES AND ROBUST PARAMETER
ESTIMATION METHODS IN CIRCULAR UNIVARIATE AND SIMPLE
CIRCULAR REGRESSION MODEL**

By

EHAB ABDUSALAM MAHMOOD

**Thesis Submitted to the School of Graduated Studies, Universiti Putra Malaysia,
in Fulfillment of the Requirements for the Degree of
Doctor of Philosophy**

December 2017

COPYRIGHT

All materials contained within this thesis, including without limitation text, logos, icons, photographs and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright © Universiti Putra Malaysia



DEDICATIONS

- * *To my parents who teach me so many original principles and support me to obtain the highest certificate by trust of God and self-confidence.*
- * *My wife, who was away from her family to realize our common dream of obtaining PhD. She always prays to satisfy this dream.*
- * *To the best gift of God to me, my kids, Hawraa, Zahraa, Hussein and Afnan, who are all my life ... I can not imagine the life without seeing their innocent eyes.*
- * *To my sister and brother, who are partners of my childhood and toys. They always encourage and support me to do the best.*



Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirements for the degree of Doctor of Philosophy

ROBUST DETECTION MEASURES AND ROBUST PARAMETER ESTIMATION METHODS IN CIRCULAR UNIVARIATE AND SIMPLE CIRCULAR REGRESSION MODEL

By

EHAB ABDUSALAM MAHMOOD

December 2017

Chairman : Professor Habshah Midi, PhD

Faculty : Science

The univariate and the simple circular regression model can be used in many scientific fields. There is evidence that the classical methods to estimate the parameters are adversely affected by outliers. Hence, it is very crucial to detect outliers in circular data. Some existing methods such as *Mardia*, *M*, *A*, and *Chord* are developed in this regard. Unfortunately, these methods are formulated to identify only a single outlier. Hence, we propose robust circular distance (*RCDu*) statistic to identify a single and multiple outliers in the univariate circular data. The results of the study indicate that the *RCDu* statistic is successful in detecting outliers with smaller masking and swamping rates.

Not much research is focused on the robust estimation of univariate circular distribution when the circular data have outliers. Thus, robust methods are proposed to estimate the circular location parameter, circular variance and mean resultant length of von Mises distribution. The findings signify that the two proposed methods have done a credible job compared to other methods in this study.

This thesis also addresses the issue of existing outliers in the simple circular regression model. Not much consideration is given to investigate the identification methods of outliers in such model. Hence, we propose robust circular distance (*RCDy*) statistic to detect outliers in the response variable of the simple circular regression model. The results of the study indicate that the *RCDy* has the highest proportion of detection outliers with the lowest rate of masking.

To the best of our knowledge, no research is focused on the detection of outliers in the response and the explanatory variables of a simple circular regression model.

Hence, robust circular distance (RCD_{xy}) statistic is formulated to detect outliers in the response and the explanatory variables. The results show that the RCD_{xy} statistic is very successful to detect outliers with low rates of masking and swamping.

The maximum likelihood estimator (MLE) is the commonly used method to estimate model parameters of the simple circular regression model. However, the MLE is inefficient if the circular data have outliers. To the best of our knowledge, no work has been done to propose robust method to estimate parameters of the simple circular regression model when the response variable has outliers. Therefore, the robust MWLE 1 and MWLE 2 are developed. The findings indicate that the MWLE2 and the MWLE1 are more efficient than the MLE.

To date, there is no robust parameters estimation method of a simple circular regression model is developed when outliers are present in the response and the explanatory variables. Therefore, two robust estimators namely MWLE1 and MWLE2 are established. The results show that the performance of the MWLE2 and the MWLE1 are more efficient than the MLE when outliers are present in both X and Y directions.

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

**SUKATAN PENGESANAN TEGUH DAN KAEDAH PENGANGGARAN
PARAMETER TEGUH DALAM UNIVARIAT SIRKULER DAN MODEL
REGRESI SIRKULAR MUDAH**

Oleh

EHAB ABDUSALAM MAHMOOD

Disember 2017

Pengerusi : Profesor Habshah Midi, PhD

Fakulti : Sains

Univariat dan model regresi sirkuler mudah boleh digunakan dalam pelbagai bidang saintifik. Terdapat bukti bahawa kaedah klasik untuk menganggar parameter model sirkuler terjejas teruk oleh pencilan. Oleh itu, adalah sangat penting untuk mengesan pencilan dalam data sirkuler. Beberapa kaedah sedia ada seperti Mardia, M, A, dan Chord telah dibangunkan dalam hal ini. Malangnya, semua kaedah ini dirumuskan hanya untuk mengesan pencilan tunggal. Oleh itu, kami mencadangkan statistik sirkuler jarak teguh (RCDu) untuk mengesan pencilan tunggal dan berganda dalam data sirkuler univariat. Hasil kajian menunjukkan bahawa statistik RCDu berhasil dalam mengesan pencilan dengan kadar litupan dan limpahan yang lebih kecil.

Tidak banyak kajian difokuskan ke atas penganggaran teguh taburan sirkuler univariat bila data sirkuler mempunyai pencilan. Oleh yang demikian, kaedah teguh dicadangkan untuk menganggar parameter lokasi sirkuler, varians sirkuler dan purata hasil panjang taburan von Mises. Penemuan menunjukkan bahawa kedua kaedah yang dicadangkan telah menghasilkan keputusan yang dipercayai berbanding kaedah lain.

Tesis ini juga mengetengahkan isu kewujudan pencilan dalam model regresi sirkuler mudah. Tidak banyak pertimbangan yang diberikan untuk menyelidik kaedah pengesanan pencilan dalam model tersebut. Oleh yang demikian, kami mencadangkan statistik jarak sirkuler teguh (RCDy) untuk mengesan pencilan dalam pembolehubah sambutan model regresi sirkuler mudah. Keputusan kajian menunjukkan bahawa RCDy mempunyai kadar pengesanan pencilan paling tinggi dengan kadar litupan paling rendah.

Sepanjang pengetahuan kami, belum ada kajian difokuskan ke atas pengesanan pencilan dalam pembolehubah sambutan dan penerang bagi model regresi sirkuler mudah. Oleh yang demikian, statistik sirkuler jarak teguh (RCDxy) diformulasi untuk mengesan pencilan dalam pembolehubah sambutan dan penerang. Keputusan menunjukkan bahawa statistik RCDxy sangat berhasil mengesan pencilan dengan kadar litupan dan limpahan yang rendah.

Penganggar kebolehdajadian maksimum (MLE) merupakan kaedah umum digunakan untuk menganggar parameter model regresi sirkuler mudah. Namun begitu, penganggar MLE tidak cekap sekiranya data sirkuler mempunyai pencilan. Sepanjang pengetahuan kami, belum ada kajian dibuat bagi mencadangkan kaedah teguh untuk menganggar parameter bagi model regresi sirkuler mudah bila pembolehubah sambutan mempunyai pencilan. Oleh yang demikian, penganggar teguh MWLE 1 dan MWLE 2 dibangunkan. Penemuan menunjukkan bahawa MWLE 1 dan MWLE 2 lebih cekap berbanding MLE.

Sehingga kini, belum ada kaedah penganggaran parameter teguh bagi model regresi sirkuler mudah dibangunkan bila pencilan wujud dalam pembolehubah sambutan dan penerang. Oleh yang demikian, dua penganggar teguh iaitu MWLE 1 dan MWLE 2 diwujudkan. Keputusan menunjukkan bahawa prestasi MWLE 1 dan MWLE 2 lebih cekap berbanding MLE bila pencilan wujud dalam kedua arah X dan Y.

ACKNOWLEDGEMENTS

To begin with, I would like to say thanks to my God, who has provided me health and grace throughout my all life.

I would like to express my deepest gratitude and thankfulness to my dear supervisor Prof. Dr. Habshah Midi for being immensely generous and supportive throughout my doctoral study. I am deeply honored to have the opportunity to complete my degree under her supervision and be one of her students. It has been a great pleasure working with Prof. Dr. Habshah Midi and no word of thanks are enough for her constant encouragement.

Special thanks and deepest gratitude to my previous supervisor Dr. Sohel Rana. I am very grateful to him. He supported me to do the best for the first two years of my doctoral study.

A big word of thanks to my supervisory committee member Prof. Dr. Abdul Ghapor Hussin for all his support, and helping me significantly with my research.

I would like to express my thanks to my supervisory committee member Dr. Jayanthi a/p Arasan.

A special word of thanks to my family, for giving me the happiness and for standing by with me patiently with their never ending encouragement, prayers and support throughout my doctoral pursue.

Finally, I would like to extend my thanks to my dear frinds and colleagues Dr. Mohanad Ali, Dr. Mohammed Alguraibawi, Dr. Ali Kahleel, Nasuha and Shelan, who have supportd and helped me during my study.

This thesis was submitted to the Senate of 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 Bt Midi, PhD

Professor
Faculty of Science
Universiti Putra Malaysia
(Chairman)

Jayanthi A/P Arasan, PhD

Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Member)

Abdul Ghapor Hussin, PhD

Professor
Faculty of Defence Science and Technology
Universiti Pertahanan Nasional Malaysia
(Member)

ROBIAH BINTI YUNUS, PhD

Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by 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 other 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 the 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 software.

Signature: _____

Date: _____

Name and Matric No.: Ehab Abdulsalam Mahmood, GS40619

Declaration by Members of 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 Midi

Signature : _____
Name of
Member of
Supervisory
Committee : Associate Professor Dr. Jayanthi A/P Arasan

Signature : _____
Name of
Member of
Supervisory
Committee : Professor Dr. Abdul Ghapor Hussin

TABLE OF CONTENTS

	Page
ABSTRACT	i
ABSTRAK	iii
ACKNOWLEDGEMENTS	v
APPROVAL	vi
DECLARATION	viii
LIST OF TABLES	xiv
LIST OF FIGURES	xx
LIST OF ABBREVIATIONS	xxi
 CHAPTER	
1 INTRODUCTION	1
1.1 Introduction and Background of the Study	1
1.2 Importance and Motivation of the Study	3
1.3 Research Objectives	5
1.4 Scope and Limitation of Study	5
1.5 Overview of the Thesis	6
2 LITERATURE REVIEW	9
2.1 Introduction	9
2.2 Background	9
2.2.1 Univariate Circular Data	9
2.2.2 Circular Regression Model	12
2.3 Circular Data	14
2.4 Circular Distance	18
2.5 Graphical Representation of Circular Data :	19
2.5.1 P-P (Probability-Probability) Plot	19
2.5.2 Plot Circular Data	20
2.5.3 Rose Diagram	20
2.6 Outliers	21
2.7 Bootstrapping	22
3 DETECTION OF OUTLIERS IN UNIVARIATE CIRCULAR DATA	23
3.1 Introduction	23
3.2 The von Mises Distribution	24
3.2.1 Properties of von Mises Distribution	24
3.2.2 Maximum Likelihood Estimator of von Mises Distribution	25
3.3 Detection of Outlier by Existing Methods	27
3.3.1 Mardia Statistic	27
3.3.2 M Statistic	28
3.3.3 A Statistic	28
3.3.4 Chord Statistic	30

3.4	The Proposed Robust Circular Distance Denoted as <i>RCDu</i> Statistic	30
3.5	Cut-Off Points of The <i>RCDu</i> Statistic	32
3.6	Performance of <i>RCDu</i> Statistic	35
	3.6.1 Single Outlier	35
	3.6.2 Multiple Outliers	40
3.7	Numerical Example	44
3.8	Conclusion	48
4	ROBUST ESTIMATION OF CIRCULAR PARAMETERS	49
4.1	Introduction	49
4.2	Mean Direction	50
4.3	Robust Methods to Estimate Circular Location Parameter	50
	4.3.1 Median Direction	50
	4.3.2 <i>M</i> -Estimation for Circular Data	51
	4.3.3 Circular Hodges-Lehmann (CHL)	53
	4.3.4 Trimmed Mean Direction (TMD)	53
	4.3.4.1 Trimmed Mean Direction 1 (TMD 1)	54
	4.3.4.2 Trimmed Mean Direction 2 (TMD 2)	55
4.4	Comparing Measures of Circular Location Parameter	55
4.5	Classical Measures of Concentration and Dispersion	70
	4.5.1 Mean Resultant Length	70
	4.5.2 Circular Variance	71
4.6	Trimmed Circular Variance (TCV)	71
	4.6.1 Trimmed Circular Variance 1 (TCV1)	72
	4.6.2 Trimmed Circular Variance 2 (TCV 2)	72
4.7	Comparing Measures of Circular Variance	73
4.8	Trimmed Mean Resultant Length (TMRL)	81
4.9	Numerical Example	84
4.10	Conclusion	85
5	DETECTION OF OUTLIERS IN THE RESPONSE VARIABLE OF A SIMPLE CIRCULAR REGRESSION MODEL	86
5.1	Introduction	86
5.2	Simple Circular Regression Model	87
	5.2.1 Assumptions of The Simple Circular Regression Model	87
	5.2.2 Maximum Likelihood Estimator (MLE) of The Simple Circular Regression Model	88
5.3	Conventional Methods for The Detection of Outliers in a Simple Circular Regression Model	90
	5.3.1 <i>COVRATIO</i> Statistic	90
	5.3.2 Mean Circular Error Statistic	91
	5.3.2.1 <i>DMCEs</i> Statistic	91
	5.3.2.2 <i>DMCEc</i> Statistic	92
5.4	The Proposed Robust Circular Distance <i>RCDy</i> Statistic	92
5.5	Cut-Off Points of The <i>RCDy</i> Statistic	93
5.6	Performance of <i>RCDy</i> Statistic	95

5.7	Numerical Examples	105
5.8	Conclusion	107
6	DETECTION OF OUTLIERS IN THE RESPONSE AND THE EXPLANATORY VARIABLES OF A SIMPLE CIRCULAR REGRESSION MODEL	108
6.1	Introduction	108
6.2	The Proposed Robust Circular Distance RCD_{xy} Statistic	109
6.3	Cut-Off Points of The RCD_{xy} Statistic	110
6.4	The Performance of RCD_{xy} Statistic	113
6.5	Numerical Examples	122
6.6	Conclusion	123
7	ROBUST PARAMETER ESTIMATION OF A SIMPLE CIRCULAR REGRESSION MODEL WHEN OUTLIERS ARE IN THE RESPONSE VARIABLE	125
7.1	Introduction	125
7.2	Maximum Weighted Likelihood Estimator (MWLE)	125
7.3	Simulation Study	129
7.4	Numerical Examples	140
7.5	Conclusion	141
8	ROBUST PARAMETER ESTIMATION OF A SIMPLE CIRCULAR REGRESSION MODEL WHEN OUTLIERS ARE IN THE RESPONSE AND THE EXPLANATORY VARIABLES	143
8.1	Introduction	143
8.2	The Proposed Robust Estimation Method	143
8.3	Simulation Study	144
8.4	Numerical Examples	155
8.5	Conclusion	156
9	SUMMARY, CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER STUDIES	157
9.1	Introduction	157
9.2	Summary	157
9.2.1	Detection of Outliers in Univariate Circular Data	158
9.2.2	Robust Estimation of Circular Parameters	158
9.2.3	Detection of Outliers in The Response Variable of a Simple Circular Regression Model	159
9.2.4	Detection of Outliers in the Response and the Explanatory Variables of a Simple Circular Regression Model	159
9.2.5	Robust Parameter Estimation of A Simple Circular Regression Model When Outliers are in The Response Variable	159
9.2.6	Robust Parameter Estimation of a Simple Circular Regression Model When Outliers are in The Response and The Explanatory Variables	160

9.3	Conclusion	160
9.4	Areas of Future Studies	161

REFERENCES	163
APPENDICES	174
BIODATA OF STUDENT	235
LIST OF PUBLICATIONS	236



LIST OF TABLES

Table		Page
3.1	The 5% upper percentile of the null distribution of <i>RCDu</i>	33
3.2	The 10% upper percentile of the null distribution of <i>RCDu</i>	34
3.3	Proportions of detection a single outlier and rates of masking and swamping for the statistics (n=20)	37
3.4	Proportions of detection a single outlier and rates of masking and swamping for the statistics (n=60)	39
3.5	Proportions of detection of outliers and rates of masking and swamping of <i>RCDu</i> statistic with 5%, 10%, 15% and 20% contaminated data (n=20)	41
3.6	Proportions of detection of outliers and rates of masking and swamping of <i>RCDu</i> statistic with 5%, 10%, 15% and 20% contaminated data (n=60)	43
3.7	Number of outliers detected of masking and swamping for all <i>Mar</i> , <i>M</i> , <i>A</i> , <i>Chord</i> and <i>RCDu</i> statistics (bedded sandstone layers data)	45
3.8	Number of outliers detected of masking and swamping for 5%, 10%, 15% and 20% for all statistics (bedded sandstone layers data)	46
4.1	The Bias of μ , <i>med</i> , <i>M</i> , CHL and TMD2 for clean data	56
4.2	The CMD of μ , <i>med</i> , <i>M</i> , CHL and TMD2 for clean data	57
4.3	The bias of Trims1 to 5 for clean data (TMD 1)	58
4.4	The CMD of Trims1 to 5 for clean data (TMD 1)	59
4.5	The bias of Trims1 to 5 for contaminated data (5%, 10%, 15% and 20%), n=20	60
4.6	The bias of Trims1 to 5 for contaminated data (5%, 10%, 15% and 20%), n=60	61
4.7	The CMD of Trims1 to 5 for contaminated data (5%, 10%, 15% and 20%), n=20	62
4.8	The CMD of Trims1 to 5 for contaminated data (5%, 10%, 15% and 20%), n=60	63

4.9	The Bias of μ , med , M , CHL, TMD 1 and TMD2 for contaminated data (5%, 10%, 15% and 20%), n=20	65
4.10	The Bias of μ , med , M , CHL, TMD 1 and TMD2 for contaminated data (5%, 10%, 15% and 20%), n=60	66
4.11	The CMD of μ , med , M , CHL, TMD 1 and TMD2 for contaminated data (5%, 10%, 15% and 20%), n=20	67
4.12	The CMD of μ , med , M , CHL, TMD 1 and TMD2 for contaminated data (5%, 10%, 15% and 20%), n=60	68
4.13	The circular variance (Classical, M and TCV2) for clean data	74
4.14	The circular variances Trims1 to 5 for contaminated data (5%, 10%, 15% and 20%), n=20	75
4.15	The circular variances Trims1 to 5 for contaminated data (5%, 10%, 15% and 20%), n=60	76
4.16	The circular variances Classic M , TCV 1 and TCV2 for contaminated data (5%, 10%, 15% and 20%), n=20	78
4.17	The circular variances Classic M , TCV 1 and TCV2 for contaminated data (5%, 10%, 15% and 20%), n=60	79
4.18	The mean resultant length computed by Classic M , TMRL 1 and TMRL 2 for clean data	81
4.19	The mean resultant length computed by Classic M , TMRL 1 and TMRL 2 for contaminated data (5%, 10%, 15% and 20%), n=20	82
4.20	The mean resultant length computed by Classic M , TMRL 1 and TMRL 2 for contaminated data (5%, 10%, 15% and 20%), n=60	83
4.21	The estimation and CMD of μ , med , M , CHL, TMD 1 and TMD 2 for original and modified frogs data	84
4.22	The circular variance and the mean resultant length of Classic, M and the two trimmed methods for original and modified frogs data	85
5.1	The cut-off points of RCD_y with 5% upper percentile	94
5.2	The cut-off points of RCD_y with 10% upper percentile	95
5.3	Proportions of detection outliers and rates of masking and swamping, (Y) 5% contamination (n=60)	97

5.4	Proportions of detection outliers and rates of masking and swamping, (Y) 10% contamination (n=60)	98
5.5	Proportions of detection outliers and rates of masking and swamping, (Y) 15% contamination (n=60)	99
5.6	Proportions of detection outliers and rates of masking and swamping, (Y) 20% contamination (n=60)	100
5.7	Proportions of detection outliers and rates of masking and swamping, (Y) 5% contamination (n=100)	101
5.8	Proportions of detection outliers and rates of masking and swamping, (Y) 10% contamination (n=100)	102
5.9	Proportions of detection outliers and rates of masking and swamping, (Y) 15% contamination (n=100)	103
5.10	Proportions of detection outliers and rates of masking and swamping, (Y) 20% contamination (n=100)	104
6.1	The cut-off points of RCD_{xy} with 5% upper percentile	111
6.2	The cut-off points of RCD_{xy} with 10% upper percentile	112
6.3	Proportions of detection of outliers and rates of masking and swamping, (X,Y) 5% contamination (n=60)	114
6.4	Proportions of detection of outliers and rates of masking and swamping, (X,Y) 10% contamination (n=60)	115
6.5	Proportions of detection of outliers and rates of masking and swamping, (X,Y) 15% contamination (n=60)	116
6.6	Proportions of detection of outliers and rates of masking and swamping, (X,Y) 20% contamination (n=60)	117
6.7	Proportions of detection of outliers and rates of masking and swamping, (X,Y) 5% contamination (n=100)	118
6.8	Proportions of detection of outliers and rates of masking and swamping, (X,Y) 10% contamination (n=100)	119
6.9	Proportions of detection of outliers and rates of masking and swamping, (X,Y) 15% contamination (n=100)	120
6.10	Proportions of detection of outliers and rates of masking and swamping, (X,Y) 20% contamination (n=100)	121

7.1	Biases, variance and MSE of α and β of clean data (Y), ($n=60$)	130
7.2	Biases, variance and MSE of α and β at 5% contaminated data (Y), ($n=60$)	131
7.3	Biases, variance and MSE of α and β at 10% contaminated data (Y), ($n=60$)	132
7.4	Biases, variance and MSE of α and β at 15% contaminated data (Y), ($n=60$)	133
7.5	Biases, variance and MSE of α and β at 20% contaminated data (Y), ($n=60$)	134
7.6	Biases, variance and MSE of α and β of clean data (Y), ($n=100$)	135
7.7	Biases, variance and MSE of α and β at 5% contaminated data (Y), ($n=100$)	136
7.8	Biases, variance and MSE of α and β at 10% contaminated data (Y), ($n=100$)	137
7.9	Biases, variance and MSE of α and β at 15% contaminated data (Y), ($n=100$)	138
7.10	Biases, variance and MSE of α and β at 20% contaminated data (Y), ($n=100$)	139
7.11	Estimates and variance of α and β for the methods MLE, MWLE1 and MWLE2 ($n=129$)	141
7.12	Estimates and variance of α and β for the methods MLE, MWLE1 and MWLE2 ($n=78$)	141
8.1	Biases, variance and MSE of α and β of clean data (x and y) for sample size 60	145
8.2	Biases, variance and MSE of α and β at 5% contaminated (x and y) for sample size 60	146
8.3	Biases, variance and MSE of α and β at 10% contaminated data (x and y) for sample size 60	147
8.4	Biases, variance and MSE of α and β at 15% contaminated data (x and y) for sample size 60	148

8.5	Biases, variance and MSE of α and β at 20% contaminated data (x and y) for sample size 60	149
8.6	Biases, variance and MSE of α and β of clean data (x and y) for sample size 100	150
8.7	Biases, variance and MSE of α and β at 5% contaminated data (x and y) for sample size 100	151
8.8	Biases, variance and MSE of α and β at 10% contaminated data (x and y) for sample size 100	152
8.9	Biases, variance and MSE of α and β at 15% contaminated data (x and y) for sample size 100	153
8.10	Biases, variance and MSE of α and β at 20% contaminated data (x and y) for sample size 100	154
8.11	Estimates and variances of MLE, MWLE1, and MWLE2 (n=129)	155
8.12	Estimates and variances of MLE, MWLE1, and MWLE2 (n=78)	156
A1.1	Proportions of detection outliers and rates of masking and swamping, (Y) 5% contamination (n=200)	174
A1.2	Proportions of detection outliers and rates of masking and swamping, (Y) 10% contamination (n=200)	175
A1.3	Proportions of detection outliers and rates of masking and swamping, (Y) 15% contamination (n=200)	176
A1.4	Proportions of detection outliers and rates of masking and swamping, (Y) 20% contamination (n=200)	177
A1.5	Proportions of detection of outliers and rates of masking and swamping, (X,Y) 5% contamination (n=200)	178
A1.6	Proportions of detection of outliers and rates of masking and swamping, (X,Y) 10% contamination (n=200)	179
A1.7	Proportions of detection of outliers and rates of masking and swamping, (X,Y) 15% contamination (n=200)	180
A1.8	Proportions of detection of outliers and rates of masking and swamping, (X,Y) 20% contamination (n=200)	181
A1.9	Biases, variance and MSE of $\hat{\alpha}$ and $\hat{\beta}$ of clean data (Y), (n=200)	182

A1.10	Biases, variance and MSE of $\hat{\alpha}$ and $\hat{\beta}$ at 5% contaminated data(Y), (n=200)	183
A1.11	Biases, variance and MSE of $\hat{\alpha}$ and $\hat{\beta}$ at 10% contaminated data (Y), (n=200)	184
A1.12	Biases, variance and MSE of $\hat{\alpha}$ and $\hat{\beta}$ at 15% contaminated data (Y), (n=200)	185
A1.13	Biases, variance and MSE of $\hat{\alpha}$ and $\hat{\beta}$ at 20% contaminated data (Y), (n=200)	186
A1.14	Biases, variance and MSE of $\hat{\alpha}$ and $\hat{\beta}$ of clean data (x and y) for sample size 200	187
A1.15	Biases, variance and MSE of $\hat{\alpha}$ and $\hat{\beta}$ at 5% contaminated data (x and y) for sample size 200	188
A1.16	Biases, variance and MSE of $\hat{\alpha}$ and $\hat{\beta}$ at 10% contaminated data (x and y) for sample size 200	189
A1.17	Biases, variance and MSE of $\hat{\alpha}$ and $\hat{\beta}$ at 15% contaminated data (X and Y) for sample size 200	190
A1.18	Biases, variance and MSE of $\hat{\alpha}$ and $\hat{\beta}$ at 20% contaminated data (X and Y) for sample size 200.	191

LIST OF FIGURES

Figure	Page
2.1 Representation of the angle ϑ	15
2.2 Circular data points	15
2.3 Difference values of ϑ	16
2.4 Unimodal of orientation of termite mounds in Cape York Peninsula, North Queensland	17
2.5 Biomodal of directions of female turtles	18
2.6 Circular distance of the arc length ab	18
2.7 P-P plot of orientation of termite mounds in Cape York Peninsula, North Queensland	20
2.8 Rose diagram for the data on orientations of 76 turtles after laying eggs	21
3.1 RCD_u statistic for bedded sandstone layers data	45
3.2 RCD_u statistic for bedded sandstone layers data for 5%, 10%, 15% and 20% contaminated data	47
4.1 The bias of μ , med , M , CHL, TMD 1 and TMD2 for contaminated data (5%, 10%, 15% and 20%), $n=60$	69
4.2 The CMD of μ , med , M , CHL, TMD 1 and TMD2 for contaminated data (5%, 10%, 15% and 20%), $n=60$	70
4.3 The circular variance clean, Classic, M , TCV 1 and TCV 2 for contaminated data (5%, 10%, 15% and 20%), $n=60$	80
5.1 $[RCD_i]_y$ statistic of the wind direction data ($n=129$)	106
5.2 RCD_y statistic of the wind direction data ($n=78$)	107
6.1 RCD_{xy} statistic of the wind direction data ($n=129$)	122
6.2 RCD_{xy} statistic of the wind direction data ($n=78$)	123

LIST OF ABBREVIATIONS

TMD	Trimmed Mean Direction
(P-P) Plot	(Probability-Probability) Plot
dist	Circular Distance
$\hat{\mu}$	Mean Direction
disp	Dispersion Measure
R	Resultant Length
\bar{R}	Mean Resultant Length
CV	Circular Variance
k	Concentration Parameter
I_0	Modified Bessel function of the first kind and order zero
MLE	Maximum Likelihood Estimator
r_c	Circular Correlation Coefficient
RCD	Robust Circular Distance Statistic
Mar	Mardia Statistic
M	M statistic
A	A Statistic
<i>Chord</i>	Chord Statistic
<i>med</i>	Median Direction
TCV	Trimmed Circular Variance
TMRL	Trimmed Mean Resultant Length
CHL	Circular Hodge-Lehmann
CMD	Circular Mean Deviation
MCE	Mean Circular Error
$mean_t$	Trimmed Mean
MWLE	Maximum Weighted Likelihood Estimator
<i>var</i>	Variance
MSE	Mean Square Error

CHAPTER 1

INTRODUCTION

1.1 Introduction and Background of the Study

Statistical data are classified according to their distributional topologies into two sets : Linear data, which they can be represented on the straight line. Second, Circular data, which they can be represented on the circumference of unit circle.

Circular data can be measured by degree and distributed within $[0^\circ - 360^\circ)$. However, it is sometimes useful to measure by radians within $[0 - 2\pi)$. There are two main ways to represent typical circular observations: (i) The Compass, which is used to measure typical circular observations such as wind directions and the directions of migrating birds; and (ii) The Clock, which is used to measure typical circular observations such as arrival times (on a 24-hour clock) (Mardia and Jupp, 2000).

Circular data are used in contrastive scientific fields such as :

i. Meteorology

The circular data are used in the meteorological studies such as the wind directions (Gatto and Jammalamadaka, 2007; Johnson and Wehrly, 1977). The circular data also include the times of day at which thunderstorms occur and the frequencies of heavy rain in a year (Mardia and Jupp, 2000).

ii. Biology

The circular data are applied for animal navigation. For example, the direction of birds migpercentagen (Batschelet, 1981; Schmidt-Koenig, 1963, 1965). The circular data may measure the spawning times of a particular fish (Lund, 1999).

iii. Physics

In physics, the circular motion is defined as a movement of an object along the circumference of a circle, it is measured by angles (Knudsen and Hjorth, 2002). Another example is the source of signals in the case of airplane crashes (Lenth, 1981).

iv. Psychology

The circular data may be used in the studies of mental maps to represent the surroundings (Gordon et al. 1989; Rustler, 2012).

v. Medicine

The medical professionals are shown that many fields of medicine may be used the circular data such as chronobiology, chronotherapy and the study of biological clock (Mahesh, 2011). The angle of knee flexion is used as a measure of recovery of orthopaedic patients (Jammalamadaka et al. 1986).

vi. Geology

The circular data may be used in several fields of geology. For example, Geologists interest is to find out the direction of flow of rivers in the past (Sengupta and Rao, 1966) and to find out the direction of earthquake displacement in terms of the direction steepest decent (Rivest, 1997).

vii. Political Science

Gill and Hangartner (2010) used circular regression model in the study of domestic terrorism analysis.

viii. Demography

The circular data are used in the demography studies such as geographic marital patterns (Coleman and Haskey, 1986).

In general, circular data can be found whenever periodic phenomena occur. However, one of vital problems which may occur in the statistical data is the existence of outliers. In real-life applications, samples from any field might include noise, or outliers. The outliers cause a huge interpretative problem, misleading of statistical analysis and incorrect of parameters estimation. This problem is common and there are methods to detect them in the linear data. Researchers are interested in improving the methods of detecting outliers in statistical data. Many researchers have proposed methods to identify the outliers and proposed robust methods to estimate model parameters of linear data. As well as, statistical software packages, such as SPSS and Minitab and R program also provide a variety of methods for identifying outliers and robust methods for univariate and linear regression model. However, there are few methods in the literature that can detect the outliers in univariate circular data. These methods are proposed to identify a single outlier (Abuzaid et al. 2009; Abuzaid, 2010; Collett, 1980; Mardia, 1975).

In the literature, there is lack of researches that are interested to propose robust methods to estimate circular parameters of circular univariate (Ackermann, 1997; Ducharme and Milasevic, 1987; He and Simpson, 1992; Ko and Guttorp, 1988; Kutil, 2012; Laha et al. 2013; Laha and Mahesh, 2011; Lenth, 1981; Otieno and Anderson-Cook, 2006; Wehrly and Shine, 1981). The word "robust" is loaded with many-sometimes inconsistent-notations. In statistic science, it is used for the purposes of robustness signifying insensitivity to small deviations from the assumptions, where the robust methods are more resistance than classical methods. The robust methods are used if the assumptions of estimation are not satisfied or the statistical have outliers (Huber and Ronchetti, 2009) .

The problem of the existing outliers in the response variable of the simple circular regression model has not been remedied adequately (Abuzaid et al. 2013; Abuzaid, et al., 2011; Abuzaid, 2010) . Furthermore, no work has been done to propose methods to identify outliers in the response and the explanatory variables of the simple circular regression model.

Another serious problem, there is no robust method that has been proposed to estimate parameters of simple circular regression model when the response variable has outliers. Moreover, no work has been done to propose robust method to estimate model parameters of the simple circular regression model when the response and the explanatory variables have outliers.

1.2 Importance and Motivation of the Study

Circular data are used in many scientific fields. They might be represented by univariate or bivariate. The simple circular regression model is one of the important models to represent the relationship between two circular variables. Efficiency and accuracy of estimation of the model parameters depend on the suitability of data that is fitted to the circular regression model. However, circular data may have some inconsistent observations with the majority of the circular data, which are called outliers. The classical methods that applied to estimate model parameters are successful under some conditions. One of these conditions is that the circular data is free of outliers. Researchers have suggested either to identify outliers and then remedy them or apply some robust methods to estimate model parameters. In the literature, there are many methods to detect outliers in univariate and bivariate linear data , as well as robust methods to estimate model parameters. However, the methods that are suggested for linear data cannot be used for circular data because of the circular geometry theory. To overcome the problem of existing outlier in the univariate circular data, Abuzaid et al. (2009), Abuzaid (2010), Collett (1980) and Mardia (1975) proposed methods to identify outliers. However, these methods can detect a single outlier point but they are not successful to identify multiple outliers. The problem of existing outliers in the univariate circular data has not received enough consideration. This motivated us to propose a statistical method to identify outliers in the univariate circular data in the presence of multiple outliers. The proposed statistic $RCDu$ is

expected to show higher proportion of detection of outliers with smaller masking and swamping rates.

The mean direction is used to estimate circular location parameter of univariate circular data. However, this estimation gives incorrect estimation if the circular data have outliers. To overcome this problem, Lenth (1981) proposed M estimator to estimate circular location parameter when the circular data have outliers. Ackermann (1997), Ducharme and Milasevic (1987), He and Simpson (1992), Ko, and Guttorp (1988) and Wehrly and Shine (1981) proposed to use the median direction if the circular data have outliers. Otieno and Anderson-Cook (2006) extended Hodges-Lehmann method to estimate circular location parameter when the circular data have outliers. Kutil (2012) explained that the mean resultant length is bias. Laha et al. (2013) and Laha and Mahesh (2011) explained that the mean direction is not SB-robust but the trimmed mean direction (TMD) is SB-robust whereby SB-robust is called standardized bias robust (Ko and Guttorp, 1988). However, Laha et al. (2013) and Laha and Mahesh (2011) did not propose method for trimming in their algorithm of circular location parameter estimation even though it is now evident that TMD is SB-robust. This inspired us to extend two methods for trimming to estimate circular location parameter, circular variance and the mean resultant length.

This thesis is also concerned in the detection of outliers in the response variable of the simple circular regression model. This issue has been addressed by Abuzaid et al. (2013), Abuzaid et al. (2011) and Abuzaid (2010). However, these methods have a low proportion of detection of outliers and high rate of masking especially if the response variable has a high percentage of contamination. The weakness of these methods has motivated us to propose a new statistic that can identify outliers with higher proportion of detection and lower rate of masking.

Many methods have been proposed to detect outliers in the response and the explanatory variables of linear regression model. To the best of our knowledge, no work has been done to propose method to detect outliers in the response and the explanatory variables of the simple circular regression model. This inspired us to propose a statistic to identify outliers in the response and the explanatory variables.

This thesis also addresses the issue of robust estimation of the parameters of the simple circular regression model. For the linear regression model, this fact is pointed out by many standard books, articles and researchers. However, to date, no robust estimation approach has been proposed to estimate the model parameters of the simple circular regression model. This motivate us to propose two methods by extending maximum weighted likelihood estimator MWLE. First, to estimate parameters when the response variable has outliers. Second, when the response and the explanatory variables have outliers.

1.3 Research Objectives

The main goal is to investigate the existing of outliers problem in the univariate and bivariate circular data. The simple circular regression model is an important model to represent the relationship of bivariate circular data. The classical methods are used to estimate parameters of circular univariate and the simple circular regression model. There is evidence that the classical methods of estimation are affected by the outliers. Therefore, The foremost objectives of our research can be outlined systematically as follows:

- i. To formulate a new robust statistic to detect a single and multi-outliers in the univariate circular data.
- ii. To extend robust method to estimate circular location parameter, circular variance and mean resultant length when the circular data have outliers.
- iii. To formulate a new robust statistic to detect outliers in the response variable of a simple circular regression model.
- iv. To formulate a new robust statistic to detect outliers in the response and the explanatory variables of the simple circular regression model.
- v. To extend a robust method to estimate the parameters of the simple circular regression model when the response variable has outliers.
- vi. To extend a robust method to estimate the parameters of the simple circular regression model when the response and the explanatory variables have outliers.

1.4 Scope and Limitation of Study

Circular data are widely used in many scientific fields such as meteorology, biology, physics, psychology, medicine, geology, political science and demography.

In linear data, the Normal distribution has many desirable properties so all the classical statistical analysis methods have been applied under Normal distribution. Similarly, in circular data, the von Mises distribution has many desirable properties so it is considered in this study.

The classical methods are used to estimate parameters of the circular univariate such as the mean direction, the circular variance and the mean resultant length. Nonetheless, these methods are affected by the existence of outliers in the dataset. Hence, researchers proposed methods to identify outlier. Nevertheless, these methods are not enough to resolve the problem of presence outliers in the circular data and it still exists. Very few researchers have focused to propose robust methods to estimate the parameters.

The linear relationship between bivariate circular data can be represented by the simple circular regression model. The maximum likelihood estimator is used to estimate the model parameters. However, it has some assumption to apply, one of them that the circular variables are free of outliers. Abuzaid et al. (2013), Abuzaid et al. (2011) and Abuzaid (2010) proposed methods to identify outliers in the response variable of a simple circular regression model. Nonetheless, they have not proposed method to detect outliers in the explanatory variable. In addition, no work has been done to propose method to detect outliers in the response and the explanatory variables of the simple circular regression model. Furthermore, to date, no robust method has been proposed to estimate parameters of the simple circular regression model when outliers are present in a data set.

Since the robust statistic is relatively new technique in the circular data, there is no algorithm and statistical software and less references and data related to this area. Hence, it is so difficult to extend the robust methods that apply in the linear regression model such as least median of squares, least trimmed squares, S-estimator, M-estimator, MM-estimator and GM-estimator.

In the literatures, not many outlying data sets are available. Hence, the same data sets were used repeatedly for different objectives of this study.

1.5 Overview of the Thesis

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

Chapter Two: This chapter briefly presents the literature review of the univariate circular data and circular regression model. The definition of the circular data, circular distance and graphical representation are illustrated. The definition of outliers, masking and swamping are given. Finally, bootstrapping methods are also briefly discussed.

Chapter Three: This chapter discusses the existing of outliers in univariate circular data. The proposed robust statistic to detect a single and multi-outliers is presented. It depends on the circular distance between the observations and the median direction as a measure for detection. Finally, numerical examples and a Monte Carlo simulation study are presented.

Chapter Four: In this chapter, a robust method to estimate the mean direction, circular variance and mean resultant length is proposed. The proposed robust method is based on extending the trimmed procedure. Two methods of trimming are proposed. First, it depends on the circular distance between observations and the median

direction as a measure for trimming. Five circular distances away from the median direction are tested. Second, it depends on the method that is proposed in Chapter 3 as a measure for trimming. A real data analysis and a Monte Carlo simulation study are carried out to assess the performance of our proposed method.

Chapter Five: This chapter deals with the existing of outliers in the response variable of the simple circular regression model. A robust statistic is proposed to identify outliers. The proposed statistic depends on the circular distance between circular residuals and the median of the circular residuals as a measure for detection. A Monte Carlo simulation study and a numerical example are carried out to assess the performance of the proposed method.

Chapter Six: In this chapter, a proposed robust statistic to detect outliers in the response and the explanatory variables of the simple circular regression model is presented. The proposed method depends on the circular distance between observations of the response and the explanatory variables as a measure for detection. A Monte Carlo simulation study and a numerical example are presented to assess the performance of the proposed method.

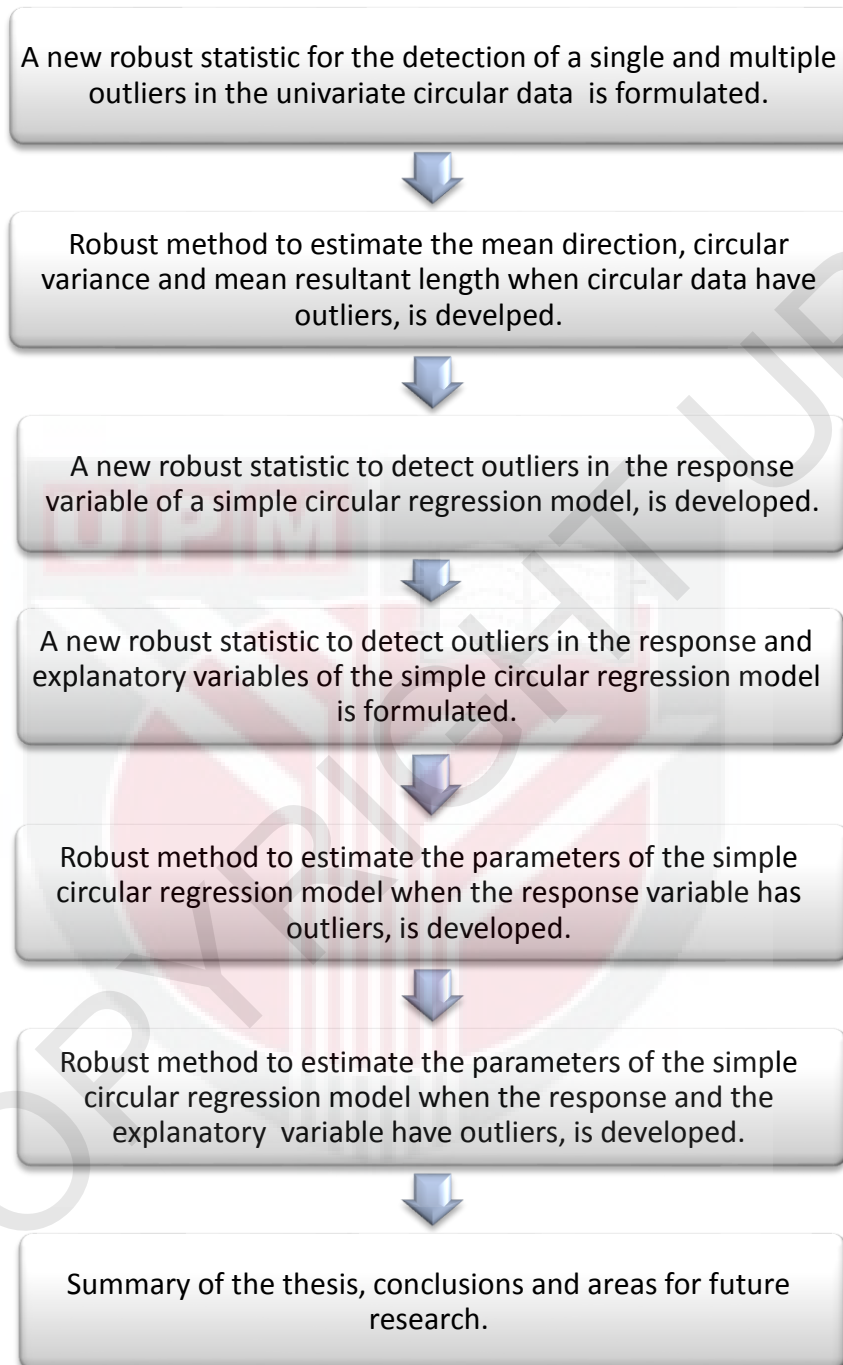
Chapter Seven: In this chapter, we propose a robust method to estimate the parameters of the simple circular regression model when the response variable has outliers. The proposed method extends the maximum weighted likelihood estimator. Two weight functions are proposed to apply the maximum weighted likelihood estimator. A numerical example and a Monte Carlo simulation study are applied to evaluate the performance of the proposed method.

Chapter Eight: This chapter deals with the existing of outliers in the response and the explanatory variables of the simple circular regression model. A robust method to estimate the model parameters is proposed. The maximum weighted likelihood estimator is extended to estimate the parameters. For this method, two weight functions are proposed. A numerical example and a Monte Carlo simulation study are applied to assess the performance of the proposed method.

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

The flow of the thesis is summarised by the following chart.

Chart to show the flow of the thesis



REFERENCES

- Abuzaid, A. H., Hussin, A. G. and Mohamed, I. B. (2008). Outlier labeling via circular boxplot. In *IASC2008*. Japan.
- Abuzaid, A. H., Mohamed, I. B. and Hussin, A. G. (2014). Procedures for outlier detection in circular time series models. *Environ Ecol Stat*, 21(4), 793–809.
- Abuzaid, A H, Hussin, A G, Rambli, A and Mohamed, I. (a). (2012). Statistics for a new test of discordance in circular data. *Communications in Statistics - Simulation and Computation*, 41, 1882–1890.
- Abuzaid, Ali H, Mohamed, I. B. H. and A. G. (b). (2012). Boxplot for circular variables. *Computational Statistics*, 27, 381–392.
- Abuzaid, A. H. (2010). *Some Problems Of Outliers In Circular Data Some Problems Of Outliers In Circular Data*. Unpublished PhD. thesis, University of Malaya, Malaysia.
- Abuzaid, A. H. (2012). Analysis of mother's day celebration via circular statistics. *The Philippine Statistician*, 61(2), 39–52.
- Abuzaid, A. H. (2013). On the influential points in the functional circular relationship models. *Pak.J.Stat.Oper.Res.*, IX(3), 333–342.
- Abuzaid, A. H., & Hussin, A. G. and Mohamed, I. B. (2013). Detection of outliers in simple circular regression models using the mean circular error statistic. *Journal of Statistical Computation and Simulation*, 83(2), 269–277.
- Abuzaid, A. H., Mohamed, I. B., & Hussin, A. G. (2009). A new test of discordancy in circular data. *Communications in Statistics - Simulation and Computation*, 38(4), 682–691.
- Abuzaid, A., Mohamed, I., Hussin, A. G., & Rambli, A. (2011). COVRATIO statistic for simple circular regression model. *Chiang Mai Journal of Science*, 38(3), 321–330.
- Ackermann, H. (1997). A note on circular nonparametrical classification. *Biometrical Journal*, 39(5), 577–587.
- Ahmed, Ejaz S., Volodin, Andrei I. and Hussein, A. A. (2005). Robust weighted likelihood estimation of exponential parameters. *IEEE Transactions on Reliability*, 54(3), 389–395.
- Alguraibawi, Mohammed, Midi, Habshah and Imon, A. H. M. R. (2015). A new robust diagnostic plot for classifying good and bad high leverage points in a multiple linear regression model. *Mathematical Problems in Engineering*, 2015, 1–12.

- Andrews, D. F., Bickel, P. J., Hampel, F. R., Huber, P. J., Rogers, W. H. and Tukey, J. W. (1972). *Robust Estimates of Location: Survey and Advances*. New Jersey: Princeton University Press.
- Andrews, D. F. (1974). A robust method for multiple linear regression. *Technometrics*, 16(4), 523–531.
- Atkinson, A. C. (1981). Two graphical displays for outlying and influential observations in regression. *Biometrika*, 68(1), 13–20.
- Atkinson, A. C. (1994). Fast very robust methods for the detection of multiple outliers. *Journal of the American Statistical Association*, 89(428), 1329–1339.
- Bagchi, Parthasarathy and Guttman, I. (1990). Spuriousity and outliers in directional data. *Journal of Applied Statistics*, 17(3), 341–350.
- Bagheri, Arezoo and Midi, H. (2009). Robust estimations as a remedy for multicollinearity caused by multiple high leverage points. *Journal of Mathematics and Statistics*, 5(4), 311.
- Bagheri, Arezoo and Midi, H. (2015). Diagnostic plot for the identification of high leverage collinearity-influential observations. *Statistics and Operations Research Transactions*, 39(1), 51–70.
- Barnett, Vic and Lewis, T. (1978). *Outlier in Statistical Data*. (1st edition). England: John Wiley & Sons.
- Barnett, Vic and Lewis, T. (1984). *Outliers in Statistical Data*. (2nd edition). England: John Wiley & Sons.
- Barnett, Vic and Lewis, T. (1994). *Outliers in Statistical Data*. (3rd edition). England: John Wiley & Sons.
- Batschelet, E. (1981). *Circular Statistics in Biology*. London: Academic Press.
- Beckman, R J and Cook, R. D. (1983). Outliers. *Technometrics*, 25(2), 119–149. <https://doi.org/10.1080/00401706.1983.10487840>
- Bentley, J. (2006). *Modelling circular data using a mixture of von Mises and uniform distributions*. Unpublished Master thesis, Simon Fraser University, Canada.
- Caires, S and Wyatt, L. R. (2003). A linear functional relationship model for circular data with an application to the assessment of ocean wave measurements. *Journal of Agricultural, Biological, and Environmental Statistics*, 8(2), 153–169.
- Chatterjee, S., & Hadi, A. S. (1986). Influential observations, high leverage points, and outliers in linear regression. *Statistical Science*, 1(3), 379–393.

- Chernick, Michael, R. and LaBudde, R. A. (2011). *An Introduction to Bootstrap Methods with Applications to R*. USA: John Wiley & Sons.
- Chernick, Michael, R. (2008). *Bootstrap Methods: A Guide for Practitioners and Researchers* (2nd editio). USA: John Wiley & Sons.
- Chernov, N. (2011). *Circular and Linear Regression: Fitting Circles and Lines by Least Squares*. Boca Raton, USA: Taylor and Francis Group.
- Coleman, D. A. and Haskey, J. C. (1986). Marital distance and its geographical orientation in England and Wales, 1979. *Transactions of the Institute of British Geographers*, 11(3), 337–355.
- Collett, D. and Lewis, T. (1981). Discriminating between the von Mises and Wrapped Normal distributions. *Australian & New Zealand Journal of Statistics*, 23(1), 73–79.
- Collett, D. (1980). Outliers in circular data. *Applied Statistics*, 29(1), 50–57.
- Croux, Christophe, Haesbroeck, Gentiane and Ruwet, C. (2013). Robust estimation for ordinal regression. *Journal of Statistical Planning and Inference*, 143(9), 1486–1499.
- Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap Methods and Their Application*. USA: Cambridge University Press.
- Di Marzio, Marco, Panzera, Agnese and Taylor, C. C. (2009). Local polynomial regression for circular predictors. *Statistics and Probability Letters*, 79(19), 2066–2075.
- Downs, T. D. and Mardia, K. V. (2002). Circular regression. *Biometrika*, 89(3), 683–697.
- Ducharme, Gilles R. and Milasevic, P. (1987). Some asymptotic properties of the circular median. *Communications in Statistics-Theory and Methods*, 16(3), 659–664.
- Efron, Bradley and Tibshirani, R. J. (1993). *An Introduction to The Bootstrap*. USA: Chapman & Hall, Inc.
- Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, 7, 1–26.
- Fan, Jianqing, Farmen, Mark and Gijbels, I. (1998). Local maximum likelihood estimation and inference. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(3), 591–608.

- Fernandez-Duran, J. J. (2007). Models for circular – linear and circular – circular Data constructed from circular distributions based on nonnegative trigonometric sums. *Biometrics*, 63, 579–585.
- Fisher, N I, Lewis, T and Willcox, M. E. (1981). Tests of discordancy for samples from Fisher’s distribution on the sphere. *Applied Statistiss*, 30, 230–237.
- Fisher, Nick I and Lee, A. J. (1983). A correlation coefficient for circular data. *Biometrika*, 70(2), 327–332.
- Fisher, N. I. (1993). *Statistical Analysis of Circular Data*. Great Britain: Cambridge University Press.
- Fitrianto, Anwar and Midi, H. (2010). Diagnostic-robust generalized potentials for identifying high leverage points in mediation analysis. *World Applied Sciences Journal*, 11(8), 979–987.
- Gatto, Riccardo and Jammalamadaka, S. R. (2007). The generalized von Mises distribution. *Statistical Methadology*, 4, 341–353.
- Gill, Jeff and Hangartner, D. (2010). Circular data in political science and how to handle it. *Political Analysis*, 18(3), 316–336.
- Gordon, A D, Jupp, P E and Byrne, R. W. (1989). The construction and assessment of mental maps. *British Journal of Mathematical and Statistical Psychology*, 42(2), 169–182.
- Gould, A. L. (1969). A regression technique for angular variates. *Biometrics*, 25, 683–700.
- Hadi, Ali S. and Luceno, A. (1997). Maximum trimmed likelihood estimators : a unified approach , examples , and algorithms. *Computational Statistics and Data Analysis*, 25, 251–272.
- Hadi, A. S. (1992). A new measure of overall potential influence in linear regression. *Computational Statistics and Data Analysis*, 14(1), 1–27.
- Hampel, Frank R., Ronchetti, Elvezio M., Rousseeuw, Peter J. and Stahel, W. A. (1986). *Robust Statistics : The Approach Based on Influence Functions*. USA: John Wiley & Sons.
- Hassan, S F, Zubairi, Y Z and Hussin, A. G. (b). (2010). Analysis of missing values in simultaneous linear functional relationship model for circular variables. *Scientific Research and Essays*, 5(12), 1483–1491.
- Hassan, Siti F, Hussin, Abdul G and Zubairi, Y. Z. (a). (2010). Estimation of functional relationship model for circular variables and its application in measurement problems. *Chiang Mai Journal of Science*, 37(2), 195–205.

- Hassan, S. F. (2009). *Development of Simultaneous Linear Circular Functional Relationship Model and Analysis of Missing Values*. Unpublished Master thesis, University of Malaya, Malaysia.
- He, Xuming and Simpson, D. G. (1992). Robust direction estimation. *The Annals of Statistics*, 20(1), 351–369.
- Hodges, J. L. and Lehmann, E. L. (1963). Estimates of location based on rank tests. *Annals of Mathematical Statistics*, 34(2), 598–611.
- Huber, Peter J. and Ronchetti, E. M. (2009). *Robust Statistics* (2nd editio). USA: Wiley.
- Huber, P. J. (1964). Robust estimation of a location: parameter. *The Annals of Mathematical Statistics*, 35(1), 73–101.
- Huber, P. J. (1981). *Robust Statistics*. USA: John Wiley & Sons.
- Huber, P. J. (1996). *Robust Statistical Procedures* (2nd editio). USA: Society for Industrial and Applied Mathematics.
- Hussin, A. G., Abuzaid, A.H., Ibrahim, A. I. N. and Rambli, A. (2013). Detection of outliers in the complex linear regression model. *Sains Malaysiana*, 42(6), 869–874.
- Hussin, A. G. and Abuzaid, A. (2012). Detection of outliers in functional relationship model for circular variables via complex form. *Pakistan Journal of Statistics*, 28(2), 205–216.
- Hussin, A G, Fieller, N. R. J. and Stillman, E. C. (2004). Linear regression model for circular variables with application to directional data. *Journal of Applied Science and Technology*, 8(1*2), 1–6.
- Hussin, Abdul G., Abuzaid, Ali, Zulkifili, Faiz and Mohamed, I. (2010). Asymptotic covariance and detection of influential observations in a linear functional relationship model for circular data with application to the measurements of wind directions. *ScienceAsia*, 36, 249–253.
- Hussin, Abdul Ghapor, Abdullah, N. A. and M. I. (a). (2010). A Complex linear regression model. *Sains Malaysiana*, 39(3), 491–494.
- Hussin, Abdul Ghapor, Abuzaid, Ali and Mohamed, I. (2009). Detection of outliers in the unreplicated linear circular functional relationship model via functional form. In *International Conference on Nonparametric Methods for Measurement Error Models and Related Topics*. Canada.

- Hussin, A. G. (1997). *Pseudo-Replication in Functional Relationships With Environmental Applications*. Unpublished PhD. thesis, University of Sheffield, UK.
- Ibrahim, S., Rambli, A., Hussin, A. G. and Mohamed, I. (2013). Outlier detection in a circular regression model using COVRATIO statistic. *Communications in Statistics - Simulation and Computation*, 42, 2272–2280.
- Ibrahim, S. (2013). *Some Outlier Problems in a Circular Regression Model*. Unpublished PhD. thesis, University of Malaya, Malaysia.
- Imon, A. H. M. R. (2002). Identifying multiple high leverage points in linear regression. *Journal of Statistical Studies*, 3, 207–218.
- Jammalamadaka, S. Rao, Bhadra, N., Chaturvedi, D., Kutty, T. K., Majumdar, P. P. and Poduval, G. (1986). Functional assessment of knee and ankle during level walking. *Data Analysis in Life Science*, 21–54.
- Jammalamadaka, S. Rao and SenGupta, A. (2001). *Topics in Circular Statistics*. Singapore: World Scientific.
- Johnson, Richard A. and Wehrly, T. (1977). Measures and models for angular correlation and angular-linear correlation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(2), 222–229.
- Jupp, P E and Mardia, K. V. (1980). A general correlation coefficient for directional data and related regression problems. *Biometrika*, 67(1), 163–173.
- Jureckova, Jana, Kumar Sen, Pranab and Picek, J. (2013). *Methodology in Robust and Nonparametric Statistics*. USA: Taylor & Francis Group.
- Jureckova, Jana and Picek, J. (2006). *Robust Statistical Methods with R*. USA: Taylor and Francis Group.
- Kamisan, Nur Arina Basilah, Hussin, Abdul Ghapor and Zubairi, Y. Z. (2010). Finding the best circular distribution for southwesterly monsoon wind direction in Malaysia. *Sains Malaysiana*, 39(3), 387–393.
- Kato, Shogo, Shimizu, Kunio and Shieh, G. S. (2008). A circular – circular regression model. *Statistica Sinica*, 18, 633–645.
- Kato, Shogo and Eguchi, S. (2016). Robust estimation of location and concentration parameters for the von Mises – Fisher distribution. *Statistical Papers*, 57, 205–234.
- Kempter, Richard, Leibold, Christian, Buzsáki, György, Diba, Kamran and Schmidt, R. (2012). Quantifying circular – linear associations: Hippocampal phase precession. *Journal of Neuroscience Methods*, 207, 113–124.

- Khokan, Mahfuzur Rahman, Bari, Wasimul and Khan, J. A. (2013). Weighted maximum likelihood approach for robust estimation : Weibull model. *Dhaka Univ. J.Sci.*, 61(2), 153–156.
- Kimber, A. C. (1985). A note on the detection and accommodation of outliers relative to Fisher's distribution on the sphere. *Applied Statistics*, 34, 169–172.
- Knudsen, Jens M. and Hjorth, P. G. (2002). *Elements of Newtonian Mechanics: Including Nonlinear Dynamics* (3rd editio). Berlin: Springer.
- Ko, Daijin and Guttorp, P. (1988). Robustness of estimators for directional data. *The Annals of Statistics*, 16(2), 609–618.
- Kutil, R. (2012). Biased and unbiased estimation of the circular mean resultant length and its variance. *Statistcs*, 46(4), 549–561.
- Laha, Arnab Kumar, Mahesh, K. C. and Ghosh, D. K. (2013). SB-robust estimators of the parameters of the wrapped normal distribution. *Communications in Statistics - Theory and Methods*, 42, 660–672.
- Laha, Arnab Kumar and Mahesh, K. C. (2011). SB-robustness of directional mean for circular distributions. *Journal of Statistical Planning and Inference*, 141, 1269–1276.
- Laha, Arnab Kumar and Mahesh, K. C. (2015). Robustness of tests for directional mean. *Statistcs*, 49(3), 522–536.
- Lenth, R. V. (1981). Robust measures of location for directional data robust measures of location for directional data. *Technometrics*, 23(1), 77–81.
- Liebscher, S, Kirschstein, T, Pandolfo, G, Porzio, G C and Ragozini, G. (2016). A robust estimator for the mean direction of the von Mises-Fisher distribution. In *International Conference on Robust Statistics*. Geneva.
- Lim, Hock Ann and Midi, H. (2016). Diagnostic robust generalized potential based on index set equality (DRGP (ISE)) for the identification of high leverage points in linear model. *Computational Statistics*, 3(31), 859–877.
- Lund, U. (1999). Least circular distance regression for directional data. *Journal of Applied Statistics*, 26(6), 723–733.
- Mahesh, K. C. (2011). *Robustness of Estimators and Tests with Directional Data*. Unpublished PhD. thesis, Saurashtra University, India.
- Mardia, Kanti V and Jupp, P. E. (2000). *Directional Statistics*. England: Wiley.
- Mardia, K. V. (1972). *Statistic of Directional Data*. London: Academic Press INC.

- Mardia, K. V. (1975). Statistics of directional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 37(3), 349–393.
- Markatou, M. (2000). Mixture models, robustness, and the weighted likelihood methodology. *Biometrics*, 56(2), 483–486.
- Maronna, R. A., & Martin R, R Douglas and Yohai, V. J. (2006). *Robust Statistics : Theory and Methods*. England: John Wiley & Sons.
- McKeague, C. P. (2011). *Elementary Algebra* (9th editio). USA: Cengage Learning.
- Midi, Habshah, Norazan, M. R. and Imon, A. H. M. R. (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.
- Midi, Habshah and Mohammed, A. M. (2015). Identification of good and bad high leverage points in multiple linear regression model. In *Proceedings of the 17th International Conference on Mathematical Methods, Computational Techniques and Intelligent Systems (Mamectis' 15)* (pp. 147–153). Spain.
- Midi, H. (1999). Preliminary estimators for robust non-linear regression estimation. *Journal of Applied Statistics*, 26(5), 591–600.
- Neykov, Neyko, Filzmoser, Peter, Dimova, R. and Neytchev, P. (2007). Robust fitting of mixtures using the trimmed likelihood estimator. *Computational Statistics & Data Analysis*, 52(1), 299–308.
- Otieno, B. Sango and Anderson-Cook, C. M. (2006). Measures of preferred direction for environmental and ecological circular data. *Environ Ecol Stat*, 13, 311–324.
- Pearson, E.S. and Hartley, H. O. (1966). Tables for statisticians. *Biometrika*, 1(3).
- Peker, K Özgür and Bacanlı, S. (2015). A group sequential test of circular data using the von Mises distribution. *Hacettepe Journal of Mathematics and Statistics*, 44(6), 1569–1578.
- Peng, Loo Yee, Midi, Habshah, Rana, Soheli and Fitrianto, A. (2016). Identification of multiple outliers in a generalized linear model with continuous variables. *Mathematical Problems in Engineering*, 2016.
- Perez, A. G. (2016). A von Mises approximation to the small sample distribution of the trimmed mean. *Metrika*, 79(4), 369–388.
- Pewsey, Arthur, Neuhauser, Markus and Ruxton, G. D. (2013). *Circular Statistics in R*. UK: Oxford University Press.

- Qin, Xu, Zhang, Jiang-she and Yan, X.-D. (a). (2011). A nonparametric circular-linear multivariate regression model with its application to wind energy. *International Mathematical Forum*, 6(53), 2615–2624.
- Rambli, Adzhar, Abuzaid, Ali H M, Mohamed, Ibrahim and Hussin, A. G. (2016). Procedure for detecting outliers in a circular regression model. *Plos One*, 11(4), 1–10.
- Rambli, Adzhar, Ibrahim, Safwati, Abdullah, Mohd Ikhwan, Hussin, Abdul Ghapor and Mohamed, I. (2012). On discordance test for the Wrapped Normal data. *Sains Malaysiana*, 41(6), 769–778.
- Rambli, Adzhar, Mohamed, Ibrahim, Abuzaid, Ali H M and Hussin, A. G. (2010). Identification of influential observations in circular regression model. In *Proceedings of The 10th Regional Conference on Statistical Sciences* (pp. 195–203).
- Rambli, Adzhar, Yunus, Rossita Mohamad, Mohamed, Ibrahim and Hussin, A. G. (2015). Outlier detection in a circular regression model. *Sains Malaysiana*, 44(7), 1027–1032.
- Rambli, A. (2011). *Outlier Detection in Circular Data and Circular-Circular Regression Model*. Unpublished Master thesis, University of Malaya, Malaysia.
- Rao, J. S. (1969). *Some Contributions to The Analysis of Circular Data*. Unpublished PhD. thesis, Indian Statistical Institute, India.
- Riazoshams, Hossein, Midi, Habshah and Sharipov, O. (2010). The performance of robust two-stage estimator in nonlinear regression with autocorrelated error. *Communications in Statistics-Simulation and Computation*, 39(6), 1251–1268.
- Riazoshams, Hossein and Midi, H. (2016). The performance of a robust multistage estimator in nonlinear regression with heteroscedastic errors. *Communications in Statistics - Simulation and Computation*, 45, 3394–3415.
- Rivest, L.-P. (1997). A decentred predictor for circular-circular regression. *Biometrika*, 84(3), 717–726.
- Rousseeuw, Peter J. and Leroy, A. M. (1987). *Robust Regression and Outlier Detection*. USA: John Wiley & Sons.
- Rustler, F. (2012). *Mind Mapping for Dummies*. England: Wiley.
- Sahinler, Suat and Topuz, D. (2007). Bootstrap and jackknife resampling algorithms for estimation of regression parameters. *Journal of Applied Quantitative Methods*, 2(2), 188–199.

- Sarma, Y. R. and Jammalamadaka, S. R. (1993). Circular Regression. In *Statistical Sciences and Data Analysis: Proceedings of the Third Pacific Area Statistical Conference*. Netherlands.
- Schmidt-Koenig, K. (1963). On the role of the loft, the distance and site of release in pigeon homing (the“ cross-loft experiment”). *The Biological Bulletin*, 125(1), 154–164.
- Schmidt-Koenig, K. (1965). Current problems in bird orientation. *Advances in the Study of Behavior*, 1, 217–278.
- Sengupta, Supriya and Rao, J. S. (1966). Statistical analysis of cross-bedding azimuths from the Kamthi formation around Bheemaram, Pranhita: Godavari Valley. *Sankhyā: The Indian Journal of Statistics, Series B*, 28(1/2), 165–174.
- Simeckova, M. (2005). Maximum weighted likelihood estimator in logistic regression. In *WDS'05 Proceedings of Contributed Papers* (pp. 144–148).
- Stigler, S. M. (1976). The anonymous professor Gergonne. *Historia Mathematica*, 3(1), 71–74.
- Tiku, Moti L. and Akkaya, A. D. (2004). *Robust Estimation and Hypothesis Testing*. New Delhi: New Age International.
- Tovar, G. Arnaiz and Rivas, R. (1986). Outliers in circular data: a Bayesian approach. *Qüestiió: Quaderns D'estadística I Investigació Operativa*, 10(1), 1–6.
- Vandev, D L and Neykov, N. M. (1998). About regression estimators with high breakdown point. *Statistics: A Journal of Theoretical and Applied Statistics*, 32(2), 111–129.
- von Mises, R. (1918). Über die 'Ganzzahligkeit' der Atomgewicht und verwandte Fragen. *Phys. Z.*, 19, 490–500.
- Wang, Xiaogang, Eeden, Constance Van and Zidek, J. V. (2004). Asymptotic properties of maximum weighted likelihood estimators. *Journal of Statistical Planning and Inference*, 119, 37–54.
- Warm, T. A. (1989). Weighted Likelihood Estimation of Ability in Item Response Theory. *Psychometrika*, 54(3), 427–450.
- Watson, G. S. (1986). Some estimation theory on the sphere. *Annals of the Institute of Statistical Mathematics*, 38(1), 263–275.
- Wehrly, Thomas E. and Shine, E. P. (1981). Influence curves of estimators for directional data. *Biometrika*, 68(1), 334–335.

Wilcoxon, R. (2012). *Introduction to Robust Estimation and Hypothesis Testing* (3rd editio). USA: Elsevier.

Wilcoxon, R. R. (2005). *Introduction to Robust Estimation and Hypothesis Testing* (2nd editio). USA: Elsevier.

