

## ABSTRACT

This study is on modelling circular data and proposed several methods of detecting outliers of the model considered in the study. In particular, the linear functional relationship model is considered where the maximum likelihood parameters and the covariance matrix are derived for the case where the error concentration parameters are unequal since previous study only considered the parameter estimation for the case of equal error concentration. The parameter estimate of  $x$  variable is improved by applying the iterative procedure in which the incidental parameter stops accumulating after the values converge to a finite number. The parameter estimate of the concentration parameter is estimated by using modified Bessel function in which it is expanded to become a cubic function. Monte Carlo simulation study shows that the proposed parameter estimations give a small bias with mean resultant length near to 1 and small estimated root mean square errors that indicate an adequacy of the estimation. Next, some methods to detect the presence of outlier in circular data are discussed. Previous studies only considered outlier detection methods for equal error concentration case. Therefore, in this thesis, all outlier detection methods take into account of both equal and unequal error concentration parameters in the model. The first proposed method in outlier detection is by using the determinantal equation of the covariance matrix, called *covratio*, in which the covariance matrix is based on the one derived in the first part of the study. Another two proposed methods used to detect the outlier are by using the difference mean circular errors. The two trigonometric functions are the *cosine* (*FDMCEC*) and *sine* (*FDMCES*) functions. The cut-off equations are derived based on the 5% upper percentile of the simulation study for each method

for 95% confident level. The feasibility of all of the methods is assessed by the power of performance in Monte Carlo simulation studies when outlier is planted in the data. The results from the simulation study suggest that the power of performance for all three outlier detection methods achieves the maximum percentage which is 100% as the level of contamination increases. Hence, this suggests the feasibility of the method used in outlier detection.

## ABSTRAK

Kajian ini adalah mengenai pemodelan data bulatan dan mencadangkan beberapa kaedah untuk mengesan titik terpencil. Secara khususnya, model hubungan fungsian linear dipertimbangkan di mana parameter kemungkinan maksimum dan matriks kovarians diperolehi untuk kes di mana parameter tumpuan ralat tidak sama memandangkan kajian sebelum ini hanya mempertimbangkan kes parameter tumpuan ralat adalah sama. Anggaran parameter bagi pembolehubah  $x$  ditingkatkan dengan menggunakan prosedur iteratif di mana parameter sampingan berhenti terkumpul setelah nilai-nilai itu berkisar kepada nombor terhingga. Parameter tumpuan dianggarkan dengan menggunakan fungsi Bessel yang diubahsuai di mana ia diperluas untuk menjadi fungsi paduan. Kajian simulasi Monte Carlo menunjukkan bahawa anggaran parameter yang dicadangkan memberi bias yang kecil dengan nilai min paduan dan penganggar ralat min kuasa dua yang kecil menunjukkan kecukupan anggaran. Seterusnya, beberapa kaedah untuk mengesan kehadiran titik terpencil dalam data bulatan akan dipertimbangkan di mana semua kaedah mengambil kira parameter kepekatan ralat sama rata dan tidak sama rata dalam model tersebut. Kaedah pertama yang dicadangkan dalam pengesanan titik terpencil adalah dengan menggunakan persamaan determinansi matriks kovarians, yang dipanggil *covratio*, di mana matriks kovarians berdasarkan pada yang diperolehi di bahagian pertama kajian. Dua lagi kaedah yang dicadangkan yang digunakan untuk mengesan titik terpencil adalah dengan menggunakan perbezaan min kesilapan pekeliling. Kedua-dua fungsi trigonometri adalah fungsi kosinus (*FDMCEC*) dan sinus (*FDMCES*). Persamaan pemotongan diperolehi berdasarkan

peratus atas 5% kajian simulasi untuk setiap kaedah untuk tahap yakin 95%. Kemungkinan semua kaedah dinilai oleh kuasa prestasi dalam kajian simulasi Monte Carlo apabila titik terencil dimasukkan ke dalam data. Hasil daripada kajian simulasi mencadangkan bahawa kekuatan prestasi untuk ketiga-tiga kaedah pengesanan terluar mencapai peratusan maksimum sebagai tahap peningkatan pencemaran dengan demikian mencadangkan kemungkinan kaedah tersebut.

## ACKNOWLEDGEMENT

Alhamdulillah, I am thankful to Allah the Most Merciful. He has helped me all the time during my thick and thin, provides me helps beyond my expectations and indeed, the best of all providers. I would like to express my sincere and deepest gratitude to Prof. Dr. Abdul Ghapor Hussin and Assoc Prof. Dr. Yong Zulina Zubairi for guiding me throughout this entire study. Their meticulous guidance, continuous encouragement and constant support from the beginning of this study have really helped me to make this study to the completion. I would like to thank Assoc Prof. Dr. Wan Rozita too.

I would like to convey a special appreciation to my parents, Mdm Kamariah Kassim and Mr Mokhtar Aman, my favourite aunty Ms Aminah Aman, my sister Nurul Adibah and my brother Affiq Ashraff for their love and prayers throughout my life. I am really thankful for their support, patience and understanding towards me.

Thirdly, I would like to exclusively acknowledge my research team especially Dr Rossita, Mdm Hafizah, Mdm Liyana, Mdm Azuraini, Dr Norli, Dr Adilah, Dr Siti Fatimah, Dr Syazwan, Dr Amel and Mr Iqbal for sharing me their help, supports and knowledge that guide me to make this thesis possible.

Last but not least, I would like to thank Pusat Pengurusan Penyelidikan & Inovasi UPNM for supporting me financially, Pusat Asasi Sains UM and Institut Sains Matematik UM for providing me the workplace and facilities to carry out my simulation during this study.

## **APPROVAL**

The Examination Committee has met on **2<sup>nd</sup> July 2019** to conduct the final examination of **Nurkhairany Amyra binti Mokhtar** on her degree thesis entitled ‘**Outlier Detection Methods of Unreplicated Linear Functional Relationship Model for Circular Variables**’.

The committee recommends that the student be awarded the **Doctor of Philosophy (Statistics)**. Members of the Examination Committee were as follows.

**Prof. Dr. Aidy bin Ali**

Faculty of Engineering

Universiti Pertahanan Nasional Malaysia

(Chairman)

**Assoc. Prof. Dr. Ummul Fahri binti Abdul Rauf**

Centre for Defence Foundation Studies

Universiti Pertahanan Nasional Malaysia

(Internal Examiner)

**Prof. Dr. Azami bin Zaharim**

Faculty of Engineering and Built Environment

Universiti Kebangsaan Malaysia

(External Examiner)

**Prof. Dr. Muhammad Qaiser Shahbaz**

Faculty of Sciences

King Abdul Aziz University

(External Examiner)

## **APPROVAL**

This thesis was submitted to the Senate of Universiti Pertahanan Nasional Malaysia and has been accepted as fulfilment of the requirements for the degree of Doctor of Philosophy (Statistics).

The members of the Supervisory Committee were as follows.

**Prof. Dr. Abdul Ghapor Hussin**

Faculty of Defence Science and Technology

Universiti Pertahanan Nasional Malaysia

(Main Supervisor)

**Assoc. Prof. Dr. Yong Zulina Zubairi**

Centre for Foundation Studies in Science

Universiti Malaya

(Co-Supervisor)

**Assoc. Prof. Dr. Wan Rozita Wan Din**

Centre for Defence Foundation Studies

Universiti Pertahanan Nasional Malaysia

(Co-Supervisor)

**UNIVERSITI PERTAHANAN NASIONAL MALAYSIA**

**DECLARATION OF THESIS**

Student's full name : Nurkhairany Amyra binti Mokhtar  
Date of birth : 18 August 1991  
Title : Outlier Detection Methods of Unreplicated Linear  
Functional Relationship Model for Circular Variables  
Academic Session : 2018/2019

I hereby declare that the work in this thesis is my own except for quotations and summaries which have been duly acknowledged.

I further declare that this thesis is classified as:

- CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)\*
- RESTRICTED** (Contains restricted information as specified by the organisation where research was done)\*
- OPEN ACCESS** I agree that my thesis to be published as online open access (full text)

I acknowledge that Universiti Pertahanan Nasional Malaysia reserves the right as follows.

1. The thesis is the property of Universiti Pertahanan Nasional Malaysia.
2. The library of Universiti Pertahanan Nasional Malaysia has the right to make copies for the purpose of research only.
3. The library has the right to make copies of the thesis for academic exchange.

\_\_\_\_\_  
Signature

\_\_\_\_\_  
\*\*Signature of Supervisor/ Dean of  
CGS/ Chief Librarian

910818-04-5356  
IC/Passport Number

Prof. Dr. Abdul Ghapor bin Hussin  
\*\*Name of Supervisor/ Dean of  
CGS/ Chief Librarian

Date: \_\_\_\_\_

Date: \_\_\_\_\_

Note : \* If the thesis is CONFIDENTIAL OR RESTRICTED, please attach the letter from the organisation stating the period and reasons for confidentiality and restriction.

\*\* Witness



# TABLE OF CONTENTS

	PAGE
ABSTRACT	ii
ABSTRAK	iv
ACKNOWLEDGEMENT	vi
APPROVAL	vii
DECLARATION OF THESIS	ix
TABLE OF CONTENTS	x
LIST OF TABLES	xvi
LIST OF FIGURES	xix
CHAPTER 1 INTRODUCTION	1
1.1 Research Background	1
1.2 Problem Statement	2
1.3 Research Objectives	3
1.4 Research Scope	4
1.5 Significance of Research	4
1.6 Limitation of study	5
1.7 Thesis Outline	5
CHAPTER 2 LITERATURE REVIEW	7
2.1 Introduction of Chapter 2	7
2.2 Circular Statistics	7
2.3 Circular Regression Models	10
2.4 Error-in-Variables Model	13

2.5 Outlier Detection	17
2.6 Summary of Reviews	20
CHAPTER 3 METHODOLOGY	27
3.1 Introduction of Chapter 3	27
3.2 Flowchart of Study	27
3.3 The Von Mises Distribution	29
3.4 Parameter Estimation using Maximum Likelihood Method	31
3.5 Derivation of Covariance Matrix	31
3.6 Outlier Detection Methods	32
3.6.1 <i>Covratio</i> Statistic for Outlier Detection	33
3.6.2 Functional Difference Mean Circular Error in Cosine Function for Outlier Detection	33
3.6.3 Functional Difference Mean Circular Error in Sine Function for Outlier Detection	34
3.7 Monte Carlo Simulation Study using TIBCO Splus Statistical Software	35
3.7.1 Performance Measures in Simulation Study	36
3.8 Practical Data	38
3.8.1 Wind Direction Data from Kuala Terengganu	38
3.8.2 Wind Direction data from Holderness Coastline	40
3.9 Summary of Methodology	42
CHAPTER 4 PARAMETER ESTIMATION FOR UNREPLICATED LINEAR FUNCTIONAL RELATIONSHIP MODEL FOR CIRCULAR VARIABLES	43

4.1 Introduction of Chapter 4	43
4.2 Unreplicated LFRM Assuming Equal Error Concentration Parameter	44
4.3 Unreplicated LFRM Assuming Unequal Error Concentration Parameters	46
4.3.1 Maximum Likelihood Estimation of $X_i$	46
4.3.2 Maximum Likelihood Estimation of Rotation Parameter	47
4.3.3 Maximum Likelihood Estimation of Concentration Parameters	49
4.3.4 Covariance Matrix of the Unreplicated LFRM	50
4.4 Simulation Study for Unreplicated LFRM Assuming Unequal Error Concentration Parameters	53
4.5 Simulation Result for Parameter Estimation	55
4.6 Application of Parameter Estimation	57
4.6.1 Application of Parameter Estimation on Real Wind Direction Data	58
4.6.2 Application of Parameter Estimation on Synthetic Data	59
4.7 Summary of Chapter 4	60
<b>CHAPTER 5 OUTLIER DETECTION METHOD USING COVRATIO STATISTICS IN THE UNREPLICATED LFRM</b>	<b>62</b>
5.1 Introduction of Chapter 5	62
5.2 Detecting Outlier Using <i>Covratio</i> Statistics	63
5.3 <i>Covratio</i> Statistics for the Unreplicated LFRM Assuming Equal Error Concentration Parameters	63
5.3.1 Determining the Cut-Off Equation of <i>Covratio</i> Statistics in Unreplicated LFRM Assuming Equal Error Concentration Parameters	66

5.3.2 The Power of Performance of <i>Covratio</i> Statistics in Unreplicated LFRM Assuming Equal Error Concentration Parameters	71
5.4 <i>Covratio</i> Statistics for the Unreplicated LFRM Assuming Unequal Error Concentration Parameters	76
5.4.1 Determining the Cut-Off Equation of <i>Covratio</i> Statistics in Unreplicated LFRM Assuming Unequal Error Concentration Parameters	77
5.4.2 The Power of Performance of <i>Covratio</i> Statistics in Unreplicated LFRM Assuming Unequal Error Concentration Parameters	81
5.5 Application of <i>Covratio</i> Statistics	85
5.5.1 Application of <i>Covratio</i> Statistics on Real Data	85
5.5.2 Application of <i>Covratio</i> Statistics on Synthetic Data	87
5.6 Summary of Chapter 5	88
CHAPTER 6 OUTLIER DETECTION METHOD USING FUNCTIONAL DIFFERENCE MEAN CIRCULAR ERROR IN THE UNREPLICATED LFRM	90
6.1 Introduction of Chapter 6	90
6.2 Detecting Outlier Using Functional Difference Mean Circular Error ( <i>FDMCE</i> ) Statistics	91
6.2.1 Functional Difference Mean Circular Error <i>Cosine</i> ( <i>FDMCEC</i> )	91
6.2.2 Functional Difference Mean Circular Error <i>Sine</i> ( <i>FDMCES</i> )	92
6.3 <i>FDMCE</i> Statistics for the Unreplicated LFRM Assuming Equal Error Concentration Parameters	93
6.3.1 Determining the Cut-Off Equation of <i>FDMCEC</i> and <i>FDMCES</i> Statistics in Unreplicated LFRM Assuming Equal Error Concentration Parameters	93

6.3.2 The Power of Performance of <i>FDMCE</i> Statistics in Unreplicated LFRM Assuming Equal Error Concentration Parameters	100
6.4 <i>FDMCE</i> Statistics for the Unreplicated LFRM Assuming Unequal Error Concentration Parameters	109
6.4.1 Determining the Cut-Off Equation of <i>FDMCEC</i> and <i>FDMCES</i> Statistics in Unreplicated LFRM Assuming Unequal Error Concentration Parameters	109
6.4.2 The Power of Performance of <i>FDMCE</i> Statistics in Unreplicated LFRM Assuming Unequal Error Concentration Parameters	115
6.5 Application of <i>FDMCE</i> statistics	122
6.5.1 Application of <i>FDMCE</i> statistics on Real Data	122
6.5.2 Application of <i>FDMCE</i> statistics on Synthetic Data	125
6.6 Summary of Chapter 6	126
CHAPTER 7 CONCLUSIONS AND RECOMMENDATIONS	128
7.1 Conclusions of Thesis	128
7.2 Contributions	130
7.3 Recommendations	130
REFERENCES	132
LIST OF PUBLICATIONS	141
LIST OF CONFERENCES/ SEMINARS	143
APPENDIX A	144
APPENDIX B	147
APPENDIX C	148
APPENDIX D	149

APPENDIX E	151
APPENDIX F	153
DETAILS OF STUDENT	154

## LIST OF TABLES

TABLE NO.	TITLE	PAGE
2.1	Summary of literature review on circular regression models	21
2.2	Summary of literature review on error-in-variables models	24
2.3	Summary of literature review of outlier detection methods	26
3.1	Summary statistics for wind direction data of Kuala Terengganu.	40
3.2	Summary statistics for wind direction data of Holderness coastline.	42
4.1	Values of $\kappa$ and $\nu$ for the corresponding $\lambda$	53
4.2	Biasness of $\hat{\alpha}$ in unreplicated LFRM	55
4.3	Biasness of $\tilde{\kappa}$ in unreplicated LFRM	56
4.4	Biasness of $\tilde{\nu}$ in unreplicated LFRM	57
4.5	Mean and variance for parameter estimates of real wind direction data for Kuala Terengganu with GAIC values.	59
4.6	Mean and variance for parameter estimates with GAIC value applied on synthetic data.	60
5.1	The values of 1% percentile of $ COVRATIO_{(-i)} - 1 $ for equal error concentration	68

5.2	The values of 5% percentile of $ COVRATIO_{(-i)} - 1 $ for equal error concentration	68
5.3	The values of 10% percentile of $ COVRATIO_{(-i)} - 1 $ for equal error concentration	68
5.4	The power of performance of <i>covratio</i> statistics in outlier detection for equal error concentration	73
5.5	The values of 1% percentile of $ COVRATIO_{(-i)} - 1 $ for unequal error concentration	79
5.6	The values of 5% percentile of $ COVRATIO_{(-i)} - 1 $ for unequal error concentration	79
5.7	The values of 10% percentile of $ COVRATIO_{(-i)} - 1 $ for unequal error concentration	79
5.8	The power of performance of <i>covratio</i> statistics in outlier detection for unequal error concentration	82
6.1	The values of 1% percentile of <i>FDMCEC</i> for equal error concentration	95
6.2	The values of 5% percentile of <i>FDMCEC</i> for equal error concentration	95
6.3	The values of 10% percentile of <i>FDMCEC</i> for equal error concentration	96
6.4	The values of 1% percentile of <i>FDMCES</i> for equal error concentration	98
6.5	The values of 5% percentile of <i>FDMCES</i> for equal error concentration	98



6.6	The values of 10% percentile of <i>FDMCES</i> for equal error concentration	98
6.7	The power of performance of <i>FDMCEC</i> in outlier detection for equal error concentration	102
6.8	The power of performance of <i>FDMCES</i> in outlier detection for equal error concentration	106
6.9	The values of 1% percentile of <i>FDMCEC</i> for unequal error concentration	110
6.10	The values of 5% percentile of <i>FDMCEC</i> for unequal error concentration	110
6.11	The values of 10% percentile of <i>FDMCEC</i> for unequal error concentration	110
6.12	The values of 1% percentile of <i>FDMCES</i> for unequal error concentration	112
6.13	The values of 5% percentile of <i>FDMCES</i> for unequal error concentration	113
6.14	The values of 10% percentile of <i>FDMCES</i> for unequal error concentration	113
6.15	The power of performance of <i>FDMCEC</i> in outlier detection for unequal error concentration	116
6.16	The power of performance of <i>FDMCES</i> in outlier detection for unequal error concentration	119

## LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
2.1	A graphical presentation of circular data.	8
3.1	An illustration of the distribution according to the value of the concentration parameter.	30
3.2	Rose diagram for wind direction data of Kuala Terengganu in 2014.	39
3.3	Rose diagram for wind direction data of Kuala Terengganu in 2015.	39
3.4	Rose diagram for wind direction measured by using high frequency radar system.	41
3.5	Rose diagram for wind direction measured by using anchored wave buoy.	41
5.1	The power series graph to determine the cut-off equation at 1% significance level for equal error concentration.	69
5.2	The power series graph to determine the cut-off equation at 5% significance level for equal error concentration.	70
5.3	The power series graph to determine the cut-off equation at 10% significance level for equal error concentration.	70
5.4	Power of performance for <i>covratio</i> statistics in detecting outlier for equal error concentration	74

	when $n = 30$ .	
5.5	Power of performance for <i>covratio</i> statistics in detecting outlier for equal error concentration when $n = 70$ .	74
5.6	Power of performance for <i>covratio</i> statistics in detecting outlier for equal error concentration when $n = 100$ .	75
5.7	Power of performance for <i>covratio</i> statistics in detecting outlier for equal error concentration when $n = 130$ .	75
5.8	The power series graph to determine the cut-off equation at 1% significance level for unequal error concentration	80
5.9	The power series graph to determine the cut-off equation at 5% significance level for unequal error concentration	80
5.10	The power series graph to determine the cut-off equation at 10% significance level for unequal error concentration	81
5.11	Power of performance for <i>covratio</i> statistics in detecting outliers for unequal error concentration when $n = 30$ .	83
5.12	Power of performance for <i>covratio</i> statistics in detecting outliers for unequal error concentration when $n = 70$ .	83

5.13	Power of performance for <i>covratio</i> statistics in detecting outliers for unequal error concentration when $n = 100$	84
5.14	Power of performance for <i>covratio</i> statistics in detecting outliers for unequal error concentration when $n = 130$ .	84
5.15	Values of $ COVRATIO_{(-i)} - 1 $ for all 129 observations of the UK Rutherford wind direction data.	86
5.16	Values of $ COVRATIO_{(-i)} - 1 $ for all 55 observations of Kuala Terengganu wind direction data.	87
5.17	Values of $ COVRATIO_{(-i)} - 1 $ for all 70 observations of synthetic data.	88
6.1	The power series graph for <i>FDMCEC</i> to determine the cut-off equation for $\kappa = 10$ at 5% significance level for equal error concentration.	96
6.2	The power series graph for <i>FDMCEC</i> to determine the cut-off equation for $\kappa = 15$ at 5% significance level for equal error concentration.	97
6.3	The power series graph for <i>FDMCEC</i> to determine the cut-off equation for $\kappa = 20$ at 5% significance level for equal error concentration.	97
6.4	The power series graph for <i>FDMCES</i> to determine	99

	the cut-off equation for $\kappa = 10$ at 5% significance level for equal error concentration.	
6.5	The power series graph for <i>FDMCES</i> to determine the cut-off equation for $\kappa = 15$ at 5% significance level for equal error concentration.	99
6.6	The power series graph for <i>FDMCES</i> to determine the cut-off equation for $\kappa = 20$ at 5% significance level for equal error concentration.	100
6.7	Power of performance for <i>FDMCEC</i> in detecting outliers for equal error concentration when $n = 30$ .	103
6.8	Power of performance for <i>FDMCEC</i> in detecting outliers for equal error concentration when $n = 70$ .	103
6.9	Power of performance for <i>FDMCEC</i> in detecting outliers for equal error concentration when $n = 100$ .	104
6.10	Power of performance for <i>FDMCEC</i> in detecting outliers for equal error concentration when $n = 130$ .	104
6.11	Power of performance for <i>FDMCEC</i> in detecting outliers for equal error concentration when $\kappa = 10$ .	105
6.12	Power of performance for <i>FDMCES</i> in detecting outliers for equal error concentration when $n = 30$ .	107
6.13	Power of performance for <i>FDMCES</i> in detecting outliers for equal error concentration when $n = 70$ .	107

6.14	Power of performance for <i>FDMCES</i> in detecting outliers for equal error concentration when $n = 100$ .	108
6.15	Power of performance for <i>FDMCES</i> in detecting outliers for equal error concentration when $n = 130$ .	108
6.16	The power series graph for <i>FDMCEC</i> to determine the cut-off equation for $\kappa = 10$ at 5% significance level for unequal error concentration	111
6.17	The power series graph for <i>FDMCEC</i> to determine the cut-off equation for $\kappa = 15$ at 5% significance level for unequal error concentration	111
6.18	The power series graph for <i>FDMCEC</i> to determine the cut-off equation for $\kappa = 20$ at 5% significance level for unequal error concentration.	112
6.19	The power series graph for <i>FDMCES</i> to determine the cut-off equation for $\kappa = 10$ at 5% significance level for unequal error concentration.	114
6.20	The power series graph for <i>FDMCES</i> to determine the cut-off equation for $\kappa = 15$ at 5% significance level for unequal error concentration.	114
6.21	The power series graph for <i>FDMCES</i> to determine the cut-off equation for $\kappa = 20$ at 5% significance level for unequal error concentration.	115
6.22	Power of performance for <i>FDMCEC</i> in detecting	117

	outliers for unequal error concentration when $n = 30$ .	
6.23	Power of performance for <i>FDMCEC</i> in detecting outliers for unequal error concentration when $n = 70$ .	117
6.24	Power of performance for <i>FDMCEC</i> in detecting outliers for unequal error concentration when $n = 100$ .	118
6.25	Power of performance for <i>FDMCEC</i> in detecting outliers for unequal error concentration when $n = 130$ .	118
6.26	Power of performance for <i>FDMCES</i> in detecting outliers for unequal error concentration when $n = 30$ .	120
6.27	Power of performance for <i>FDMCES</i> in detecting outliers for unequal error concentration when $n = 70$ .	120
6.28	Power of performance for <i>FDMCES</i> in detecting outliers for unequal error concentration when $n = 100$ .	121
6.29	Power of performance for <i>FDMCES</i> in detecting outliers for unequal error concentration when $n = 130$ .	121
6.30	Values of <i>FDMCEC</i> for all 129 observations of the Humberside Coast wind direction data.	122

6.31	Values of <i>FDMCES</i> for all 129 observations of the Humberside Coast wind direction data.	123
6.32	Values of <i>FDMCEC</i> for all 55 observations of Kuala Terengganu wind direction data.	124
6.33	Values of <i>FDMCES</i> for all 55 observations of Kuala Terengganu wind direction data.	124
6.34	Values of <i>FDMCEC</i> for all 70 observations of synthetic data.	125
6.35	Values of <i>FDMCES</i> for all 70 observations of synthetic data.	126



# CHAPTER 1

## INTRODUCTION

### 1.1 Research Background

Circular data cannot be analysed by standard linear statistical methods. This is because of the wrapped around nature of the data. However, there are some selected periodical and directional statistical distributions that can be used for modelling circular data such as uniform, cardioid, wrapped normal, and von Mises distribution (Heckenbergerova et al., 2015). Circular data happens in many evidences in our real life, in various scientific fields such as earth sciences, meteorology, biology, physics, psychology, and medicine (Pandolfo et al., 2018).

One of specific examples of the studies in circular data is when Ferguson et al. (1967) used circular data to investigate the homing ability of the northern cricket frog *Acris crepitans*, in which the orientation of the frog was recorded. Gareth (2007) built a time series model for circular data with applications to protein conformational angles. In 2015, Masseran identified a distribution of wind direction for Kudat station in Malaysia (Masseran, 2015).