

**Universiti Teknologi MARA**

**The Assessment of Particulate Matter  
10 (PM10) Functional Data at Klang  
and Shah Alam monitoring stations**

**Engku Nurul Adlin Binti Engku Zainal Abidin**

**Report submitted in fulfillment of the requirements  
for  
Bachelor of Science (Hons.) Management  
Mathematics  
Faculty of Computer and Mathematical Sciences**

**July 2020**

## **STUDENT'S DECLARATION**

I certify that this report and the research to which it refers are the product of my own work and that any ideas or quotation from the work of other people, published or otherwise are fully acknowledged in accordance with the standard referring practices of the discipline.



.....  
**ENGK NURUL ADLIN BINTI ENGKU ZAINAL ABIDIN**  
**2017987687**

**AUGUST 5, 2020**

## ABSTRACT

Air pollution in Malaysia is becoming a major environmental issue due to the rising number of vehicles, open burning, and the release of toxic chemicals.  $PM_{10}$  is a dangerous air pollutant with the capability to negatively affect human health and the surrounding environment.  $PM_{10}$  contamination is intangible and odourless, so its presence in the atmosphere is unnoticeable and often neglected. The occurrence of a high concentration of an abnormal pollutant is more simply known as an anomaly and may describe problems of air quality. Thus, this study was carried out to assess the functional curve of  $PM_{10}$  behaviour in Shah Alam and Klang and also to detect the anomaly in  $PM_{10}$  functional data. The statistical method used in this study is Functional Data Analysis combined with the robust Mahalanobis distance to detect the anomalies using air quality data recorded from 2014 to 2016 in an hourly interval. The Bayesian Information Criteria is used to determine the number of basis functions,  $K$ . Based on the anomalies that has been detected, it showed that Klang is more polluted compared to Shah Alam. The maximum level of anomalies was observed during a twenty-four hour period. In conclusion, the detection of anomaly was found useful in investigating air pollution in this study. The findings of this study imply that the location and background of a station play a significant role in influencing the anomalies of  $PM_{10}$ .

**Keywords:** Functional data analysis,  $PM_{10}$  functional data, basis function, anomaly detection.

# TABLE OF CONTENTS

<b>CONTENTS</b>	<b>PAGE</b>
<b>SUPERVISOR'S APPROVAL</b>	ii
<b>DECLARATION</b>	iii
<b>ACKNOWLEDGEMENT</b>	iv
<b>ABSTRACT</b>	v
<b>TABLE OF CONTENTS</b>	vi
<b>LIST OF FIGURES</b>	viii
<b>LIST OF TABLES</b>	ix
<b>LIST OF ABBREVIATIONS</b>	x
 <b>CHAPTER ONE: INTRODUCTION</b>	
1.1 Background of the Study	1
1.2 Problem Statement	2
1.3 Objective of the Study	3
1.4 Scope of the Study	3
1.5 Significance of the Study	4
 <b>CHAPTER TWO: LITERATURE REVIEW</b>	
2.1 Functional Data Analysis	5
2.2 Application of Functional Data Analysis	6
2.3 Functional Principal Component Analysis	7
2.4 Detection of Anomalies	7
2.5 Study about PM10 in Malaysia	8
2.6 Summary	8

## **CHAPTER THREE: RESEARH METHODOLOGY**

3.1	Method of Data Collection	9
3.2	Method of Data Analysis	9
3.3	Flowchart of the study	10
	3.3.1 Import Dataset	10
	3.3.2 Data Conversion	11
	3.3.2.1 Fourier Basis	12
	3.3.3 Anomaly Detection	13
	3.3.4 Functional Principal Component Analysis	15

## **CHAPTER FOUR: RESULTS AND DISCUSSIONS**

4.1	Result of Data Conversion	16
	4.1.1 Fourier Basis Transform	17
4.2	Anomaly Detection for Klang and Shah Alam	18
4.3	Result of Functional Principal Component Analysis	21

## **CHAPTER FIVE: CONCLUSION AND RECOMMENDATION**

5.1	Conclusion	22
5.2	Recommendation	23

<b>REFERENCES</b>	24
-------------------	----

## **APPENDICES**

APPENDIX A: DATASET FOR PARTICULATE MATTER 10 (PM10) IN KLANG AND SHAH ALAM MONITORING STATIONS	27
--	----

APPENDIX B: COMMAND R STUDIO PROGRAMMING	45
--	----