

**PROBABILISTIC AND DISTRIBUTION  
MODELLING FOR PREDICTING PM<sub>10</sub>  
CONCENTRATION IN MALAYSIA**

**HAZRUL BIN ABDUL HAMID**

**UNIVERSITI SAINS MALAYSIA**

**2013**

**PROBABILISTIC AND DISTRIBUTION MODELLING  
FOR PREDICTING PM<sub>10</sub> CONCENTRATION  
IN MALAYSIA**

**By**

**HAZRUL BIN ABDUL HAMID**

**Thesis submitted in fulfillment of the requirements  
for the degree of  
Doctor of Philosophy**

**June 2013**

## **ACKNOWLEDGEMENTS**

First and foremost, I would like to express my deepest praise to Allah who has given me strength, faith and determination to complete this research. Alhamdulillah.

I would like to express my sincere and deepest appreciation to Assoc Prof Ahmad Shukri Yahaya and Prof Dr Nor Azam Ramli. They are very understanding, patient and have given me an invaluable support and guidance. Thank you for understanding my situation as a part-time student.

My sincere thanks and deepest gratitude to my parent and family for their unfailing love, relentless encouragement, support and prayers that have contributed towards the accomplishment of this thesis. Special thanks to committee of Clean Air Research Group, Ahmad Zia, Zul Azmi, Dr. Nurul Izma, Norrimi, Maisarah, Salmi, Hasfazilah and Azian. Their friendship, cooperation and help are greatly appreciated. Last but not least, thank you to my colleague at Penang Matriculation College.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
TABLE OF CONTENTS	iii
LIST OF TABLES	vii
LIST OF FIGURES	xv
LIST OF ABBREVIATIONS	xix
ABSTRAK	xxi
ABSTRACT	xxiii
<b>CHAPTER 1 – INTRODUCTION</b>	
1.1 Air Pollution	1
1.2 Particulate Matter	15
1.3 Problem Statement	25
1.4 Objectives	25
1.5 Scope of Research	27
1.6 Thesis Outline	28
<b>CHAPTER 2 – LITERATURE REVIEW</b>	
2.1 Air Pollution	30
2.1.1 Characteristic and Behaviour of PM <sub>10</sub> Concentration	42
2.2 Statistical Distribution of Air Quality Concentration	43
2.2.1 Parameter Estimations	50
2.3 Time Series Modelling of Air Pollutants Concentration	51

## **CHAPTER 3 – METHODOLOGY**

3.1 Introduction	55
3.2 Study Area and Data	56
3.3 The Descriptive Statistics Analysis	57
3.3.1 Skewness	57
3.3.2 Kurtosis	58
3.3.3 Graphical Representation	59
3.4 The t-test	60
3.5 Analysis of Variance (ANOVA)	62
3.5.1 Duncan’s Multiple Range Test	63
3.6 Distributions and Parameter Estimation	64
3.6.1 Two-parameter Lognormal Distribution	64
3.6.2 Three Parameter Lognormal Distribution	66
3.6.3 Two Parameter Weibull Distribution	69
3.6.4 Three Parameter Weibull Distribution	71
3.6.5 Two Parameter Gamma Distribution	73
3.6.6 Three Parameter Gamma Distribution	76
3.7 Performance Indicators	77
3.7.1 Root Mean Square Error (RMSE)	78
3.7.2 Normalized Absolute Error (NAE)	78
3.7.3 Mean Absolute Percentage Error (MAPE)	79
3.7.4 Index of Agreement (IA)	80
3.7.5 Coefficient of Determination ( $R^2$ )	80
3.7.6 Prediction Accuracy (PA)	81

3.8 Procedure of Time Series Analysis	81
3.8.1 Autoregressive (AR) processes	82
3.8.2 Moving Average (MA) processes	83
3.8.3 Autoregressive Moving Average (ARMA) processes	84
3.8.4 Autoregressive Integrated Moving Average (ARIMA) processes	84
3.8.5 Seasonal ARIMA $(p, d, q) \times (P, D, Q)_s$	85
3.8.6 Data Collection and Examination	86
3.8.7 Stationarity and Seasonality of Time Series	87
3.8.8 Model Identification and Estimation	88
3.8.9 Validation	89
<b>CHAPTER 4 – RESULTS</b>	
4.1 Introduction	90
4.2 General Approach of PM <sub>10</sub> Concentration	90
4.2.1 Descriptive Statistics	90
4.2.2 Analysis of Variance	102
4.2.3 Comparison of PM <sub>10</sub> Concentration using bar chart, ANOVA and t-test	109
4.3 The Probability Distribution of PM <sub>10</sub> Concentrations	132
4.4 Development of Time Series Models	161
4.5 Validation of the models	192
<b>CHAPTER 5 – CONCLUSIONS</b>	
5.1 Conclusion of the research	203
5.2 Contributions, limitations and recommendations	206

<b>REFERENCES</b>	207
<b>APPENDICES</b>	220
<b>LIST OF PUBLICATIONS</b>	238

## LIST OF TABLES

	<b>PAGE</b>	
Table 1.1	The Malaysia air pollution index	7
Table 1.2	Specific air pollutants and associated health effects	14
Table 1.3	Malaysia ambient air quality guideline	15
Table 1.4	Comparisons of fine and coarse particle matter	18
Table 1.5	World Health Organization air quality guidelines and interim targets for particulate matter : annual mean concentrations	19
Table 1.6	World Health Organization air quality guidelines and interim targets for particulate matter : 24-hour concentrations	20
Table 2.1	The probability density functions that are useful in representing air pollution concentrations	44
Table 2.2	Summary of the latest research on statistical distribution of air quality concentration	49
Table 2.3	Summary of the latest research of time series analysis for air quality concentration	54
Table 3.1	Monitoring sites, type of station and coordinate of location	56
Table 3.2	An example of ANOVA table	62
Table 4.1	Analysis of variance for Nilai, Negeri Sembilan	103
Table 4.2	Duncan multiple range test for Nilai, Negeri Sembilan	103
Table 4.3	Analysis of variance for Kuching, Sarawak	104
Table 4.4	Duncan multiple range test for Kuching, Sarawak	104
Table 4.5	Analysis of variance for Perai, Pulau Pinang	105
Table 4.6	Duncan multiple range test for Perai, Pulau Pinang	105



Table 4.7	Analysis of variance for Kuala Terengganu, Terengganu	106
Table 4.8	Duncan multiple range test for Kuala Terengganu, Terengganu	106
Table 4.9	Analysis of variance for Bachang, Melaka	106
Table 4.10	Duncan multiple range test for Bachang, Melaka	107
Table 4.11	Analysis of variance for Seberang Jaya, Pulau Pinang	107
Table 4.12	Duncan multiple range test for Seberang Jaya, Pulau Pinang	108
Table 4.13	Analysis of variance for Jerantut, Pahang	108
Table 4.14	Duncan multiple range test for Jerantut, Pahang	108
Table 4.15	Analysis of variance to compare the mean of PM <sub>10</sub> concentration for each hour in Nilai	110
Table 4.16	Duncan multiple range test to compare the mean of PM <sub>10</sub> concentration by hour for Nilai from 2003 to 2009	111
Table 4.17	Analysis of variance to compare the mean of PM <sub>10</sub> concentration for each hour in Kuching	113
Table 4.18	Duncan multiple range test to compare the mean of PM <sub>10</sub> concentration by hour for Kuching from 2003 to 2009	114
Table 4.19	Analysis of variance to compare the mean of PM <sub>10</sub> concentration for each hour in Perai	116
Table 4.20	Duncan multiple range test to compare the mean of PM <sub>10</sub> concentration by hour for Perai from 2003 to 2009	117
Table 4.21	Analysis of variance to compare the mean of PM <sub>10</sub> concentration for each hour in Kuala Terengganu	118
Table 4.22	Duncan multiple range test to compare the mean of PM <sub>10</sub> concentration by hour for Kuala Terengganu from 2003 to 2009	119
Table 4.23	Analysis of variance to compare the mean of PM <sub>10</sub> concentration for each hour in Bachang, Melaka	120

Table 4.24	Duncan multiple range test to compare the mean of PM <sub>10</sub> concentration by hour for Bachang Melaka from 2003 to 2009	121
Table 4.25	Analysis of variance to compare the mean of PM <sub>10</sub> concentration for each hour in Seberang Jaya	122
Table 4.26	Duncan multiple range test to compare the mean of PM <sub>10</sub> concentration by hour for Seberang Jaya from 2003 to 2009	123
Table 4.27	Analysis of variance to compare the mean of PM <sub>10</sub> concentration for each hour in Jerantut	124
Table 4.28	Duncan multiple range test to compare the mean of PM <sub>10</sub> concentration by hour for Jerantut from 2003 to 2009	125
Table 4.29	The t-test results to compare the mean of day and night for Nilai, Negeri Sembilan	127
Table 4.30	The t-test results to compare the mean of day and night for Kuching, Sarawak	128
Table 4.31	The t-test results to compare the mean of day and night for Perai, Pulau Pinang	129
Table 4.32	The t-test results to compare the mean of day and night for Kuala Terengganu	130
Table 4.33	The t-test results to compare the mean of day and night for Bachang, Melaka	130
Table 4.34	The t-test results to compare the mean of day and night for Seberang Jaya, Pulau Pinang	131
Table 4.35	The t-test results to compare the mean of day and night for Jerantut Pahang	131
Table 4.36	Performance indicator and the best parameter estimator for Nilai by using gamma distribution	134
Table 4.37	Performance indicator and the best parameter estimator for Nilai by using lognormal distribution	135
Table 4.38	Performance indicator and the best parameter estimator for Nilai by using Weibull distribution	136

Table 4.39	Best distribution for Nilai, Negeri Sembilan	137
Table 4.40	Parameter estimated based on the best distribution for Nilai, Negeri Sembilan	137
Table 4.41	The predicted and actual exceedences for Nilai, Negeri Sembilan	140
Table 4.42	Best distribution for Kuching, Sarawak	142
Table 4.43	Parameter estimated based on the best distribution for Kuching, Sarawak	142
Table 4.44	The predicted and actual exceedences for Kuching, Sarawak	144
Table 4.45	Best distribution for Perai, Pulau Pinang	146
Table 4.46	Parameter estimated based on the best distribution for Perai, Pulau Pinang	146
Table 4.47	The predicted and actual exceedences for Perai, Pulau Pinang	148
Table 4.48	Best distribution for Kuala Terengganu, Terengganu	149
Table 4.49	Parameter estimated based on the best distribution for Kuala Terengganu, Terengganu	149
Table 4.50	The predicted and actual exceedences for Kuala Terengganu, Terengganu	151
Table 4.51	Best distribution for Bachang, Melaka	151
Table 4.52	Parameter estimated based on the best distribution for Bachang, Melaka	152
Table 4.53	The predicted and actual exceedences for Bachang, Melaka	154
Table 4.54	Best distribution for Seberang Jaya, Pulau Pinang	155
Table 4.55	Parameter estimated based on the best distribution for Seberang Jaya, Pulau Pinang	155
Table 4.56	The predicted and actual exceedences for Seberang Jaya, Pulau Pinang	157
Table 4.57	Best distribution for Jerantut, Pahang	158

Table 4.58	Parameter estimated based on the best distribution for Jerantut, Pahang	158
Table 4.59	The predicted and actual exceedences for Jerantut, Pahang	160
Table 4.60	ADF statistic and $p$ -value for Nilai monitoring station	169
Table 4.61	The best time series model for PM <sub>10</sub> concentration for Nilai, Negeri Sembilan	170
Table 4.62	Time series equation based on the best time series model for Nilai	171
Table 4.63	ADF statistic and $p$ -value for complete set of monitoring records of PM <sub>10</sub> concentration in Nilai for composite years from April 2003 to March 2009	172
Table 4.64	The best time series model in Nilai for composite years from April 2003 to March 2009	172
Table 4.65	Time series equation based on the best time series model in Nilai for composite years from April 2003 to March 2009	173
Table 4.66	ADF statistic and $p$ -value for Kuching monitoring station	173
Table 4.67	The best time series model for PM <sub>10</sub> concentration for Kuching	174
Table 4.68	Time series equation based on the best time series model for Kuching	175
Table 4.69	ADF statistic and $p$ -value for complete set of monitoring records of PM <sub>10</sub> concentration in Kuching for composite years from April 2003 to March 2009	176
Table 4.70	The best time series model in Kuching for composite years from April 2003 to March 2009	176
Table 4.71	Time series equation based on the best time series model in Kuching for composite years from April 2003 to March 2009	176
Table 4.72	ADF statistic and $p$ -value for Perai monitoring station	177
Table 4.73	The best time series model for PM <sub>10</sub> concentration for Perai	177

Table 4.74	Time series equation based on the best time series model for Perai	178
Table 4.75	ADF statistic and $p$ -value for complete set of monitoring records of PM <sub>10</sub> concentration in Perai for composite years from April 2003 to March 2009	179
Table 4.76	The best time series model in Perai for composite years from April 2003 to March 2009	179
Table 4.77	Time series equation based on the best time series model in Perai for composite years from April 2003 to March 2009	179
Table 4.78	ADF statistic and $p$ -value for Kuala Terengganu monitoring station	180
Table 4.79	The best time series model for PM <sub>10</sub> concentration for Kuala Terengganu	180
Table 4.80	Time series equation based on the best time series model for Kuala Terengganu	181
Table 4.81	ADF statistic and $p$ -value for complete set of monitoring records of PM <sub>10</sub> concentration in Kuala Terengganu for composite years from April 2003 to March 2009	182
Table 4.82	The best time series model in Kuala Terengganu for composite years from April 2003 to March 2009	182
Table 4.83	Time series equation based on the best time series model in Kuala Terengganu for composite years from April 2003 to March 2009	182
Table 4.84	ADF statistic and $p$ -value for Bachang monitoring station	183
Table 4.85	The best time series model for PM <sub>10</sub> concentration for Bachang	183
Table 4.86	Time series equation based on the best time series model for Bachang	184
Table 4.87	ADF statistic and $p$ -value for complete set of monitoring records of PM <sub>10</sub> concentration in Bachang for composite years from April 2003 to March 2009	185

Table 4.88	The best time series model in Bachang for composite years from April 2003 to March 2009	185
Table 4.89	Time series equation based on the best time series model in Bachang for composite years from April 2003 to March 2009	185
Table 4.90	ADF statistic and $p$ -value for Seberang Jaya monitoring station	186
Table 4.91	The best time series model for PM <sub>10</sub> concentration for Seberang Jaya	186
Table 4.92	Time series equation based on the best time series model for Seberang Jaya	187
Table 4.93	ADF statistic and $p$ -value for complete set of monitoring records of PM <sub>10</sub> concentration in Seberang Jaya for composite years from April 2003 to March 2009	188
Table 4.94	The best time series model in Seberang Jaya for composite years from April 2003 to March 2009	188
Table 4.95	Time series equation based on the best time series model in Seberang Jaya for composite years from April 2003 to March 2009	188
Table 4.96	ADF statistic and $p$ -value for Jerantut monitoring station	189
Table 4.97	The best time series model for PM <sub>10</sub> concentration for Jerantut	189
Table 4.98	Time series equation based on the best time series model for Jerantut	190
Table 4.99	ADF statistic and $p$ -value for complete set of monitoring records of PM <sub>10</sub> concentration in Jerantut for composite years from April 2003 to March 2009	191
Table 4.100	The best time series model in Jerantut for composite years from April 2003 to March 2009	191
Table 4.101	Time series equation based on the best time series model in Jerantut for composite years from April 2003 to March 2009	191
Table 4.102	Model performance indicator for each set of monitoring records	193

Table 4.103 Model performance indicator for composite years from  
April 2003 to March 2009

201

## LIST OF FIGURES

	<b>PAGE</b>	
Figure 1.1	Monitoring stations in Peninsular Malaysia	2
Figure 1.2	Monitoring stations in East Malaysia	3
Figure 1.3	Number of registered vehicle in 2008 and 2009	8
Figure 1.4	Number of registered vehicle in 2010 and 2011	8
Figure 1.5	Number of inuse vehicle in 2008 and 2009	9
Figure 1.6	Number of inuse vehicle in 2010 and 2011	9
Figure 1.7	Number of industrial air pollution sources in 2011	10
Figure 1.8	Air pollution index status in 2011 in several areas in west coast of Peninsular Malaysia	11
Figure 1.9	Air pollution index status in 2011 in several areas in east coast of Peninsular Malaysia	12
Figure 1.10	Air pollution index status in 2011 in several areas in Sabah and Sarawak	13
Figure 1.11	Annual average concentration of PM <sub>10</sub> in Malaysia from 1999 to 2011	22
Figure 1.12	Annual average concentration of PM <sub>10</sub> by land type	23
Figure 1.13	Particulate matter emission load by sources in Malaysia in 2011	24
Figure 3.1	Research flow	55
Figure 3.2	Location of monitoring stations	57
Figure 3.3	The standard box plot	60
Figure 3.4	Procedure of Duncan's multiple range test	64
Figure 4.1	Box plot and descriptive statistics for Nilai, Negeri Sembilan	91
Figure 4.2	Box plot and descriptive statistics for Kuching, Sarawak	93



Figure 4.3	Box plot and descriptive statistics for Seberang Perai Pulau Pinang	96
Figure 4.4	Box plot and descriptive statistics for Kuala Terengganu Terengganu	97
Figure 4.5	Box plot and descriptive statistics for Bachang, Melaka	98
Figure 4.6	Box plot and descriptive statistics for Seberang Jaya Pulau Pinang	100
Figure 4.7	Box plot and descriptive statistics for Jerantut, Pahang	101
Figure 4.8	Bar chart of average PM <sub>10</sub> concentration for Nilai	110
Figure 4.9	Bar chart of average PM <sub>10</sub> concentration for Kuching	112
Figure 4.10	Bar chart of average PM <sub>10</sub> concentration for Seberang Perai	115
Figure 4.11	Bar chart of average PM <sub>10</sub> concentration for Kuala Terengganu	118
Figure 4.12	Bar chart of average PM <sub>10</sub> concentration for Bachang Melaka	120
Figure 4.13	Bar chart of average PM <sub>10</sub> concentration for Seberang Jaya	122
Figure 4.14	Bar chart of average PM <sub>10</sub> concentration for Jerantut	124
Figure 4.15	The probability density function plot of the selected distribution for PM <sub>10</sub> concentration in Nilai	139
Figure 4.16	The cumulative distribution function plot of the selected distribution for PM <sub>10</sub> concentration in Nilai	139
Figure 4.17	The probability density function plot of the selected distribution for PM <sub>10</sub> concentration in Kuching	143
Figure 4.18	The cumulative distribution function plot of the selected distribution for PM <sub>10</sub> concentration in Kuching	144
Figure 4.19	The probability density function plot of the selected distribution for PM <sub>10</sub> concentration in Perai	147
Figure 4.20	The cumulative distribution function plot of the selected distribution for PM <sub>10</sub> concentration in Perai	147

Figure 4.21	The probability density function plot of the selected distribution for PM <sub>10</sub> concentration in Kuala Terengganu	150
Figure 4.22	The cumulative distribution function plot of the selected distribution for PM <sub>10</sub> concentration in Kuala Terengganu	150
Figure 4.23	The probability density function plot of the selected distribution for PM <sub>10</sub> concentration in Bachang	153
Figure 4.24	The cumulative distribution function plot of the selected distribution for PM <sub>10</sub> concentration in Bachang	153
Figure 4.25	The probability density function plot of the selected distribution for PM <sub>10</sub> concentration in Seberang Jaya	156
Figure 4.26	The cumulative distribution function plot of the selected distribution for PM <sub>10</sub> concentration in Seberang Jaya	156
Figure 4.27	The probability density function plot of the selected distribution for PM <sub>10</sub> concentration in Jerantut	159
Figure 4.28	The cumulative distribution function plot of the selected distribution for PM <sub>10</sub> concentration in Jerantut	159
Figure 4.29	Time series plot of PM <sub>10</sub> concentration for Nilai, Negeri Sembilan	162
Figure 4.30	Time series plot of PM <sub>10</sub> concentration for Kuching Sarawak	163
Figure 4.31	Time series plot of PM <sub>10</sub> concentration for Perai, Pulau Pinang	164
Figure 4.32	Time series plot of PM <sub>10</sub> concentration for Kuala Terengganu	165
Figure 4.33	Time series plot of PM <sub>10</sub> concentration for Bachang, Melaka	166
Figure 4.34	Time series plot of PM <sub>10</sub> concentration for Seberang Jaya, Pulau Pinang	167
Figure 4.35	Time series plot of PM <sub>10</sub> concentration for Jerantut, Pahang	168
Figure 4.36	Observed and predicted of PM <sub>10</sub> concentration using the best time series model for Nilai	193
Figure 4.37	Observed and predicted of PM <sub>10</sub> concentration using the best time series model for Kuching	194

Figure 4.38	Observed and predicted of PM <sub>10</sub> concentration using the best time series model for Perai	195
Figure 4.39	Observed and predicted of PM <sub>10</sub> concentration using the best time series model for Kuala Terengganu	196
Figure 4.40	Observed and predicted of PM <sub>10</sub> concentration using the best time series model for Bachang, Melaka	197
Figure 4.41	Observed and predicted of PM <sub>10</sub> concentration using the best time series model for Seberang Jaya	198
Figure 4.42	Observed and predicted of PM <sub>10</sub> concentration using the best time series model for Jerantut	199
Figure 4.43	Observed and predicted of PM <sub>10</sub> concentration for composite Years from April 2003 to March 2008	201

## LIST OF ABBREVIATIONS

ACF	Auto-correlation function
ADF	Augmented Dickey-Fuller
AIC	Akaike information criterion
ANOVA	Analysis of Variance
API	Air Pollutant Index
AR	Autoregressive
ARIMA	Autoregressive integration moving average
ASMA	Alam Sekitar Malaysia Berhad
cdf	Cumulative distribution function
CO	Carbon monoxide
DoE	Department of Environment (Malaysia)
IA	Index of Agreement
ILP	Institut Latihan Perindustrian
MA	Moving average
MAAQG	Malaysian Ambient Air Quality Guidelines
MAPE	Mean absolute percentage error
MoM	Method of moments
NAE	Normalized Absolute Error
NO <sub>2</sub>	Nitrogen Dioxide
PACF	Partial auto-correlation function
pdf	Probability density function

PM <sub>10</sub>	Particulate matter less than 10 μm
PWM	Probability weighted moments
R <sup>2</sup>	Coefficient of Determination
RH	Relative Humidity
RMSE	Root Mean Square Error
SE-1	Simple explicit1
SE-2	Simple explicit 2
SO <sub>2</sub>	Sulphur Dioxide
USEPA	United States Environmental Protection Agency
WHO	World Health Organization

**PERMODELAN KEBARANGKALIAN DAN TABURAN UNTUK  
MERAMALKAN KEPEKATAN PM<sub>10</sub> DI MALAYSIA**

**ABSTRAK**

Penurunan tahap kualiti udara boleh memberi kesan yang signifikan terutamanya kepada kesihatan manusia. Warga tua, kanak-kanak dan pesakit asma merupakan antara golongan yang paling terkesan jika berhadapan dengan kualiti udara yang rendah. Di Malaysia, Jabatan Alam Sekitar bertanggungjawab untuk memantau dan merekodkan kualiti udara dan pada ketika ini, terdapat sebanyak 52 stesen pemantauan kualiti udara yang dikendalikan oleh Alam Sekitar Malaysia Berhad (ASMA). Antara parameter yang diukur ialah partikel udara bersaiz kurang daripada 10  $\mu\text{m}$ , sulfur dioksida, nitrogen dioksida, karbon monoksida, suhu persekitaran, tahap kelembapan dan juga kelajuan angin. Adalah penting untuk melakukan peramalan pencemaran udara bagi memberi amaran awal kepada golongan berkaitan selain memudahkan pengurusan kualiti udara oleh pihak berkuasa tempatan. Oleh kerana model peramalan yang berkaitan data pencemar udara yang telah dibangunkan sebelum ini hanya tertumpu kepada model taburan statistik dengan dua parameter bagi peramalan jangka panjang, kajian ini membandingkan taburan statistik dengan tiga parameter kerana pencarian model terbaik bagi meningkatkan ketepatan ramalan adalah masih diperlukan. Tujuh stesen pemantauan kualiti udara dipilih dalam kajian ini bagi mewakili kawasan perindustrian (Nilai, Kuching dan Seberang Perai), kawasan bandar dan luar Bandar (Bachang, Kuala Terengganu dan Seberang Jaya) serta satu stesen dikategorikan sebagai stesen rujukan iaitu Jerantut. Data PM<sub>10</sub> bagi setiap jam dari tahun 2003 hingga 2009 telah digunakan bagi membanding dan menilai perlakuan pencemar udara PM<sub>10</sub> bagi ketujuh-tujuh stesen

yang terlibat. Data ini juga digunakan bagi mendapatkan model taburan statistik terbaik. Tiga taburan statistik yang digunakan secara meluas dalam bidang kejuruteraan alam sekitar iaitu taburan lognormal, taburan gamma dan taburan Weibull terlibat dalam kajian ini. Berdasarkan petunjuk prestasi, taburan statistik dengan tiga parameter berjaya memberikan ramalan yang lebih baik jika dibandingkan dengan taburan statistik dengan dua parameter bagi kebanyakan stesen pemantauan yang terlibat. Taburan terbaik bagi mewakili kepekatan PM<sub>10</sub> di kawasan industri, bandar, luar bandar dan stesen rujukan adalah taburan lognormal kecuali bagi kawasan Perai dan Bachang pada tahun 2005 serta Jerantut pada tahun 2009 dimana taburan Weibull adalah lebih baik berbanding taburan lognormal. Berdasarkan keputusan yang diperolehi, secara amnya taburan dengan tiga parameter adalah taburan terbaik bagi kawasan industri. Kepekatan PM<sub>10</sub> yang diramal yang melebihi tahap piawaian dalam unit hari juga dianggarkan dengan menggunakan taburan terbaik. Sebagai tambahan, iaitu bagi mendapatkan peramalan jangka pendek, purata harian bacaan kepekatan PM<sub>10</sub> digunakan bagi mendapatkan model siri masa terbaik. Dapatan kajian menunjukkan bahawa model ARIMA bermusim sesuai digunakan di kesemua tujuh stesen pemantauan kualiti udara yang terlibat.

# **PROBABILISTIC AND DISTRIBUTION MODELLING FOR PREDICTING PM<sub>10</sub> CONCENTRATION IN MALAYSIA**

## **ABSTRACT**

The decline in air quality can have a significant impact, particularly on human health. The elderly, children and people with asthma are among the most affected if faced with low air quality level. In Malaysia, the Department of Environment is responsible for monitoring and recording air quality and at this point, there are a total of 52 air quality monitoring stations operated by Alam Sekitar Malaysia Berhad (ASMA). The parameters monitored are airborne particles smaller than 10 µm, sulphur dioxide, nitrogen dioxide, carbon monoxide, ambient temperature, relative humidity and wind speed. It is important to do air pollution forecasting to give an early warning to people in addition to help the management of air quality by the local authorities. Since the forecasting model for air pollutants previously developed only focused on statistical distribution model with two parameters for long-term forecasting, this research compares the three parameter statistical distributions technique to search for the best models for improving the prediction accuracy. Seven air quality monitoring stations chosen in this study represent the industrial area (Nilai, Kuching and Seberang Perai), urban and sub-urban areas (Bachang, Kuala Terengganu and Seberang Jaya) and one station categorized as the background station which is Jerantut. Hourly PM<sub>10</sub> concentrations from 2003 to 2009 are used to assess and compare the behaviour of PM<sub>10</sub> concentrations at the selected monitoring stations. These monitoring records are also used to obtain the best statistical distribution models. The three statistical distributions techniques that are widely used in the field of environmental engineering namely



lognormal distribution, the gamma distribution and Weibull distribution were used to fit hourly averages of  $PM_{10}$  concentrations. Based on the performance indicators, the three parameter statistical distributions models provide the best long term prediction compared to the two parameter models at most monitoring stations. Predicted  $PM_{10}$  concentrations that exceed the permissible guidelines in the unit of day were also estimated using the best distribution. The best distribution to represent the  $PM_{10}$  concentration in industrial, urban, sub urban and background area is lognormal distribution except for Perai and Bachang in 2005 and Jerantut in 2009 where Weibull distribution performs better compared to lognormal distribution. In general, three-parameter distribution is the best distribution for industrial monitoring site based on the result obtained. The short-term forecasting, the annual daily particulate matter ( $PM_{10}$ ) was used to get the best time series model. The results showed that seasonal ARIMA is the most suitable time series model to forecast the  $PM_{10}$  concentrations for all seven sites involved in this study.

# CHAPTER 1

## INTRODUCTION

### 1.1 Air Pollution

Air pollution emissions degrade air quality whether in urban or rural settings. An issue of great concern has been the detrimental effect of low air quality onto human health (Kampa and Castanas, 2008). Understanding the behaviour of air pollution statistically would allow predictions to be made accurately and more pressing are needs to understand the distribution that fits the collected data which can further be used for predictions of exceedences. In the early days of abundant resources and minimal development pressures, little attention was paid to growing environmental concern in Malaysia (Afroz et al., 2003). The Department of Environment is major institution in Malaysia that is responsible for monitoring the status of air quality throughout the country to perceive any significant change which may cause harm to human health and environment.

In Malaysia, there are 52 monitoring locations throughout the country that belong to the Department of Environment (Department of Environment, Malaysia, 2011). The parameters monitored include Total Suspended Particulates, Particulate Matter (PM<sub>10</sub>), Sulphur Dioxide (SO<sub>2</sub>) and several airborne heavy metals. This 52 monitoring stations are categorized into five categories which is industrial, urban, sub-urban, rural and

background station. Figure 1.1 and Figure 1.2 shows the location of the monitoring stations in Peninsular Malaysia and East Malaysia respectively.



Figure 1.1 : Monitoring stations in Peninsular Malaysia.  
(Source : Department of Environment, 2011)

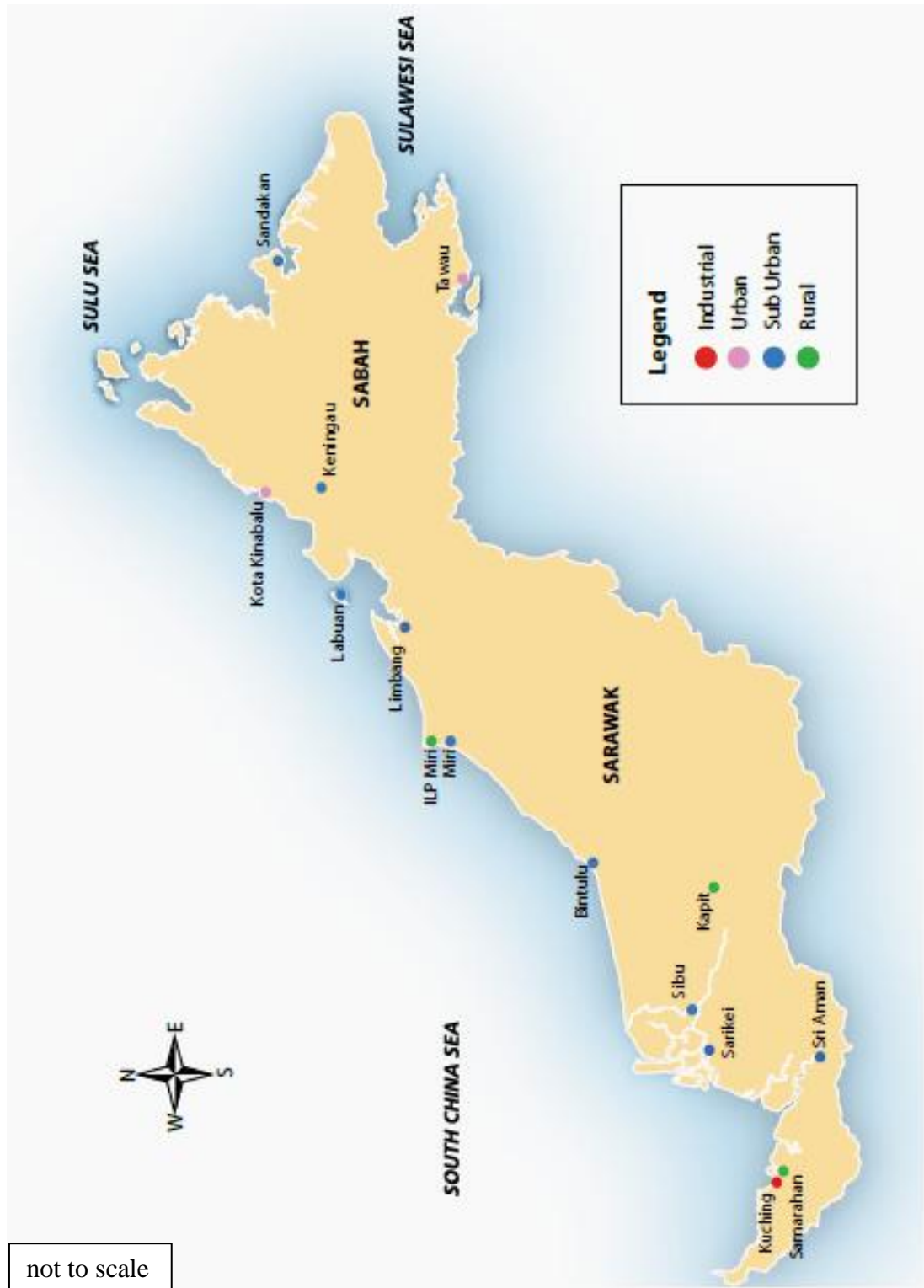


Figure 1.2 : Monitoring stations in East Malaysia.  
 (Source : Department of Environment, 2011)

Nevers (2000) defines air pollution as the presence of undesirable material in air, in quantities large enough to produce harmful effects to human health, vegetation, human property, or the global environment as well as create aesthetic insults in the form of brown or hazy air or unpleasant smells.

Air pollutants can be divided into two categories which are primary pollutants and secondary pollutants (Stander, 2000). Primary pollutants are those that have the same form (state and chemical composition) in the ambient atmosphere as when emitted from the sources. Secondary pollutants are those that have changed in form after leaving the source due to oxidation or decay or reaction with other primary pollutants (Stander, 2000). There are many sources of air pollution such as mobile sources, stationary sources and open burning sources (Afroz et al., 2003). Mobile sources include personal vehicles, commercial vehicles and motorcycles. Stationary sources refer to factory and industry, power stations, industrial fuel burning processes, and domestic fuel burning while open burning sources refer to burning of solid wastes and forest fires. However, the major contributor for this unhealthy event is peat swamp and forest fires (Department of Environment, Malaysia, 2002).

In Malaysia, three major sources of air pollution are mobile sources, stationary sources and open burning sources. However, for the past few years, emissions from mobile sources have been the major sources of air pollution which contribute around 70% to 75% of total air pollution in Malaysia (Afroz et al., 2003). The number of vehicle increased significantly each year and that is the reason why mobile sources contribute

the highest percentage of air pollutant. The federal territory Kuala Lumpur has the highest vehicle population followed by Johor, Selangor, Perak and Pulau Pinang (Department of Environment, 2006).

Stationary sources are also a major contributor to air pollution in Malaysia and stationary sources refer to power plant and industrial pollution. In 2010, industrial pollution increase by 71.44% compared to 2009 (Department of Environment, 2012). The third major contributor of air pollution in Malaysia is open burning and commonly caused by transboundary pollution, it is unexpected and also out of control.

The ambient air quality measurement in Malaysia is described in term of air pollution index (API). Air pollution index is obtained from the measurement of particles below 10 micrometre in aerodynamic diameter and several gaseous such as carbon monoxide, sulphur dioxide, ozone and nitrogen dioxide. To determine the API for a given period of time, the monitoring records for all five pollutants included in the API system were calculated based on the average concentration. The maximum index between the five pollutants was selected as the API (Department of Environment, 2000). Table 1.1 shows the air pollution index for Malaysia. The Malaysia air pollution index has six categories to categorize the level of air pollution namely good, moderate, unhealthy, very unhealthy, hazardous and emergency. Details for each category is also explained in Table 1.1.

Figure 1.3 and Figure 1.4 shows the number of registered vehicle in 2008 and 2009 and the number of registered vehicle in 2010 and 2011 respectively. As a comparison, the numbers of passenger cars, motorcycles, buses and taxis have increased every year. For sure, the increment of vehicle will contribute to the air pollution. Figure 1.5 and Figure 1.6 shows the number of in use vehicle in 2008 and 2009 and the number of in use vehicle in 2010 and 2011 respectively.

Table 1.1: The Malaysia air pollution index

Air Pollution Index	Diagnosis	Level of Pollution	Health Measures
0 – 50	Good	Pollution low and has no ill effects on health.	<ul style="list-style-type: none"> <li>- no restriction of activities for all group of people</li> <li>- to practice healthy lifestyle e.g. not to smoke, exercise regularly and not to observe proper nutrition.</li> </ul>
50 – 100	Moderate	Moderate pollution and has no ill effects on health.	<ul style="list-style-type: none"> <li>- no restriction of activities for all group of people</li> <li>- to practice healthy lifestyle e.g. not to smoke, exercise regularly and not to observe proper nutrition.</li> </ul>
101 – 200	Unhealthy	Mild aggravation of symptoms among high risk person e.g. those with heart or lung disease.	<ul style="list-style-type: none"> <li>- restriction of outdoor activities for high risk person</li> <li>- general population should reduce vigorous outdoor activities.</li> </ul>
201 – 300	Very Unhealthy	Significant aggravation of symptoms and decreased exercise tolerance in person with heart or lung disease.	<ul style="list-style-type: none"> <li>- elderly and person with known heart or lung disease should stay indoor and reduce physical activity</li> <li>- general population should reduce vigorous outdoor activities</li> <li>- those with any health problems to consult doctor.</li> </ul>
301 – 500	Hazardous	Severe aggravation of symptoms and endangers health.	<ul style="list-style-type: none"> <li>- elderly and person with existing heart or lung disease should stay indoor and reduce physical activity</li> <li>- general population should reduce vigorous outdoor activities</li> </ul>
Above 500	Emergency	Severe aggravation of symptoms and endangers health	<ul style="list-style-type: none"> <li>- general population advised to follow the orders of the national Security Council and always to follow the announcements through the mass media</li> </ul>

Source: Department of Environment, Malaysia (2009)



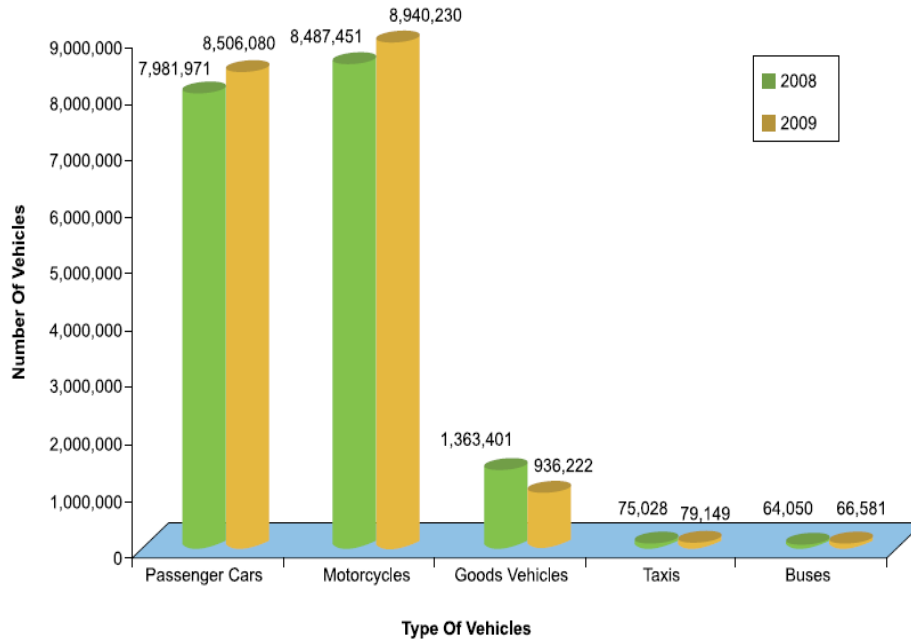


Figure 1.3 : Number of registered vehicle in 2008 and 2009  
(Source : Department of Environment, 2009)

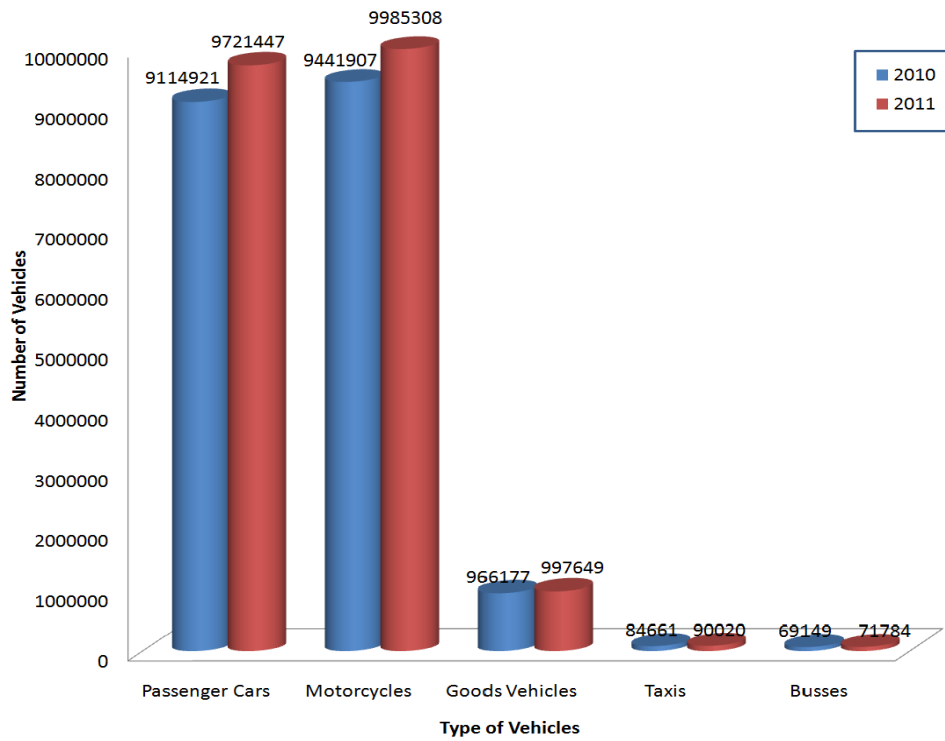


Figure 1.4 : Number of registered vehicle in 2010 and 2011  
(Source : Department of Environment, 2011)

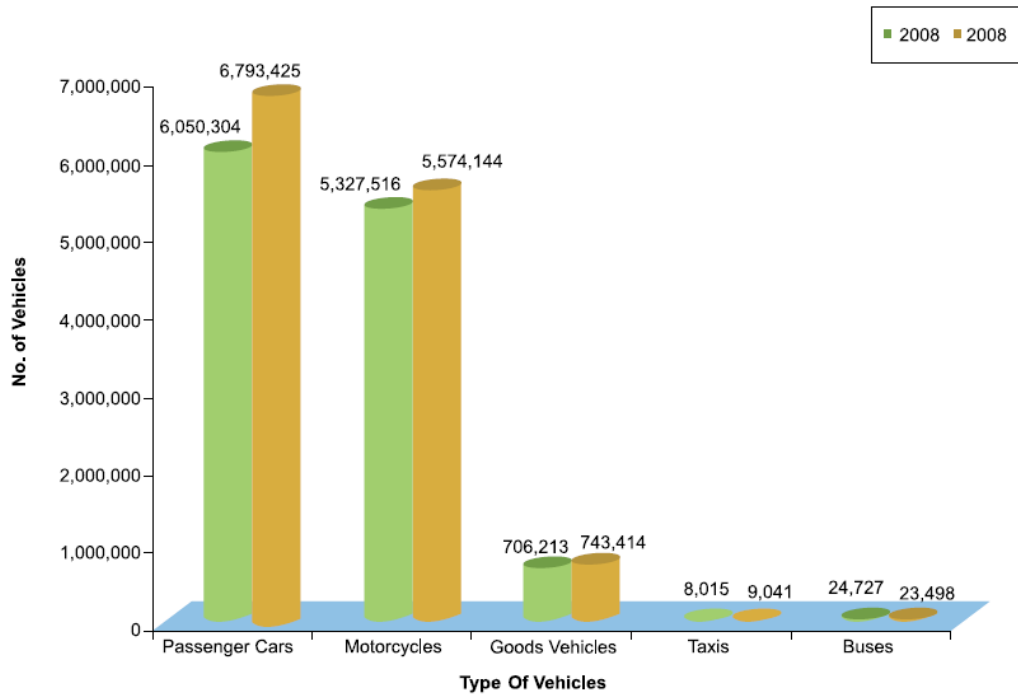


Figure 1.5 : Number of in use vehicle in 2008 and 2009  
(Source : Department of Environment, 2009)

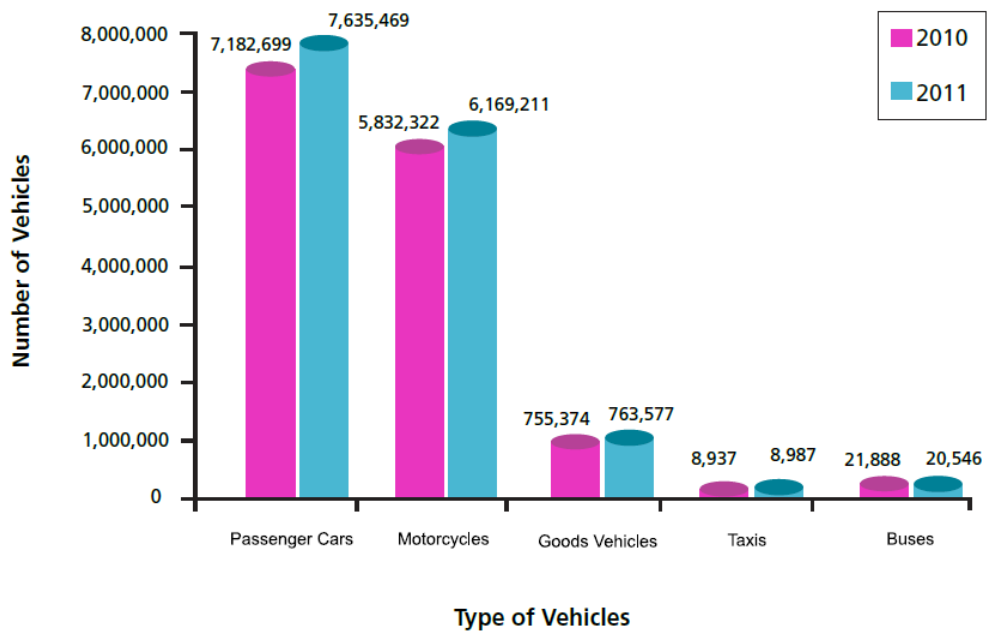


Figure 1.6: Number of in use vehicle in 2010 and 2011  
(Source : Department of Environment, 2011)

The number of industrial air pollution sources in Malaysia in 2011 is shown in Figure 1.7. Johor recorded the highest number of industrial sources followed by Selangor. The lowest number of industrial sources was recorded in Wilayah Persekutuan Labuan and Perlis.

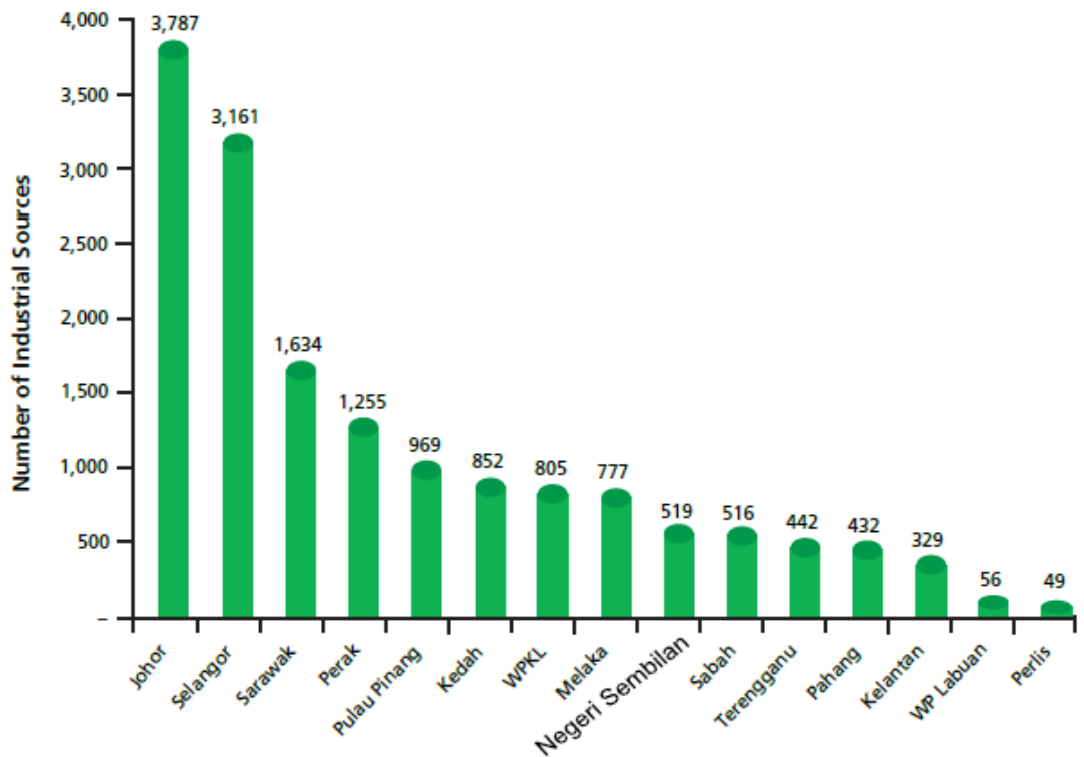


Figure 1.7 : Number of industrial air pollution sources in 2011  
(Source : Department of Environment, 2011)

Figure 1.8 shows the Air Pollution Index status in 2011 in several areas in the west coast of Peninsular Malaysia. There are several unhealthy days recorded in west coast of Peninsular Malaysia. Tanjung Malim in Perak recorded the highest number of unhealthy days followed by Alor Setar, Nilai and Port Dickson. However, Bukit Rambai recorded the lowest number of healthy days with only 16 days in 2011 followed by Nilai with

only 64 healthy days. Figure 1.9 shows the Air Pollution Index status in 2011 in several areas in east coast of peninsular Malaysia. There are no unhealthy days recorded at all monitoring stations in east coast of peninsular Malaysia except Kota Tinggi which recorded one unhealthy day in 2011 (Department of Environment, 2011).

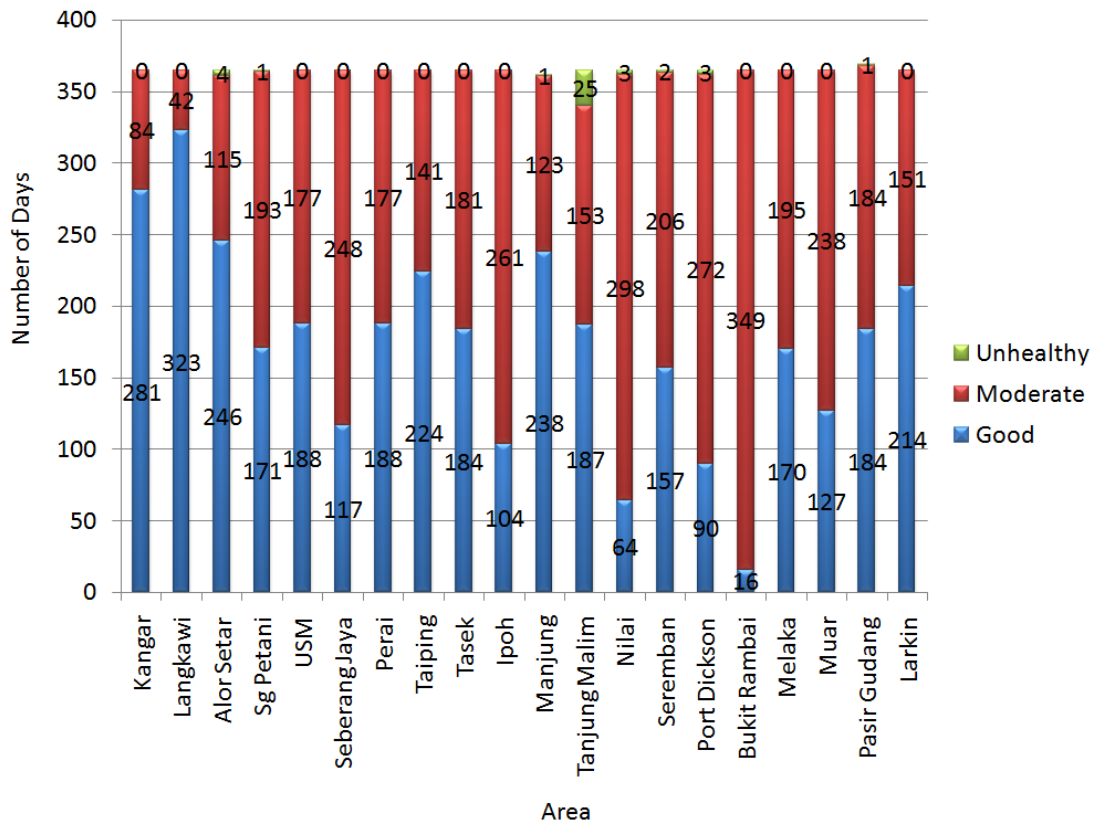


Figure 1.8 : Air Pollution Index status in 2011 in several areas in west coast of Peninsular Malaysia  
(Source : Department of Environment, 2011)

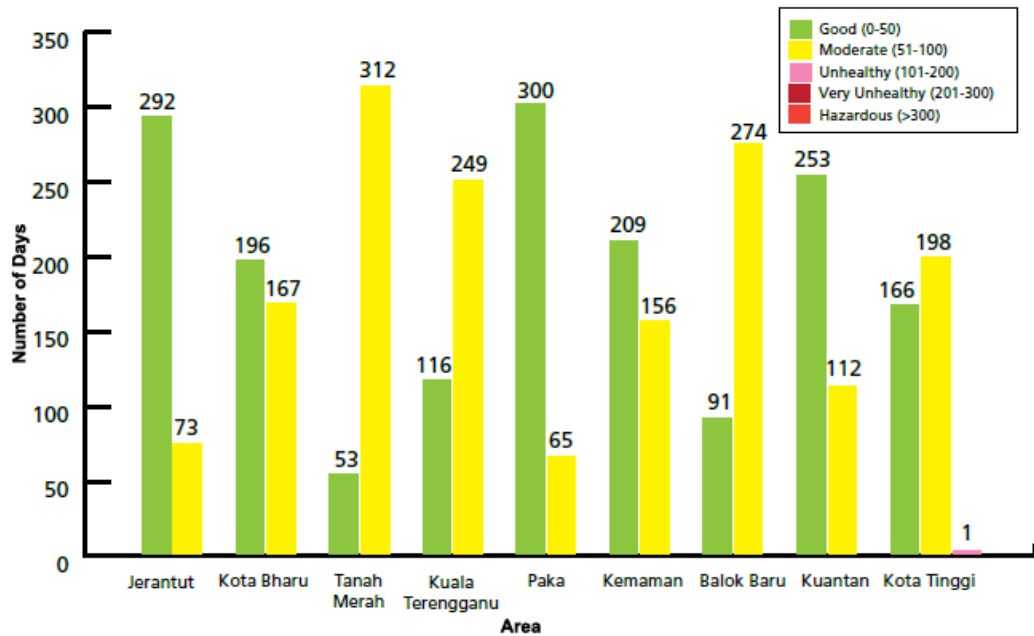


Figure 1.9 : Air Pollution Index status in 2011 in several areas in east coast of peninsular Malaysia  
(Source : Department of Environment, 2011)

The Air Pollution Index status in several monitoring stations in Sabah and Sarawak in 2011 is shown in Figure 1.10. All 15 monitoring stations in Sabah and Sarawak recorded the majority of good air pollution index except for Bintulu where the moderate air pollution index is higher than good. Only Sri Aman, Sibul and ILP Miri recorded unhealthy days in 2011.

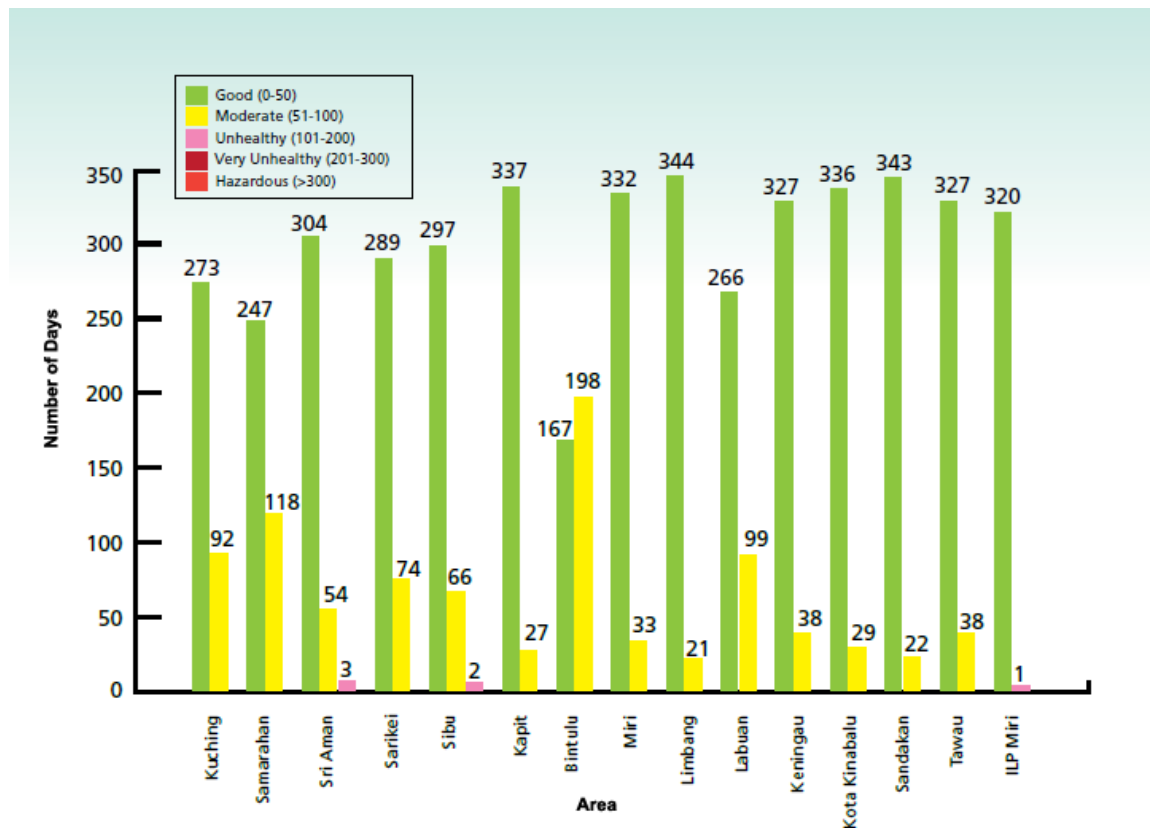


Figure 1.10 : Air Pollution Index status in 2011 in several areas in Sabah and Sarawak (Source : Department of Environment, 2011)

The impact of air pollution is noticeable, especially for human being where it can cause several significant effects including carcinogenic effects. Such health problems include cardiac arrhythmias, reducing lung function, asthma, chronic bronchitis and increasing respiratory symptoms, such as sinusitis, sore throat, dry and wet cough, and hay fever (World Health Organization, 1998). There are possible short-term and long term health effects of exposure to air pollution. In the short-term, high levels of air pollution lead to an acute condition. In addition, blockage of sunlight by air pollutants may promote the spread of harmful bacteria and viruses that would otherwise be killed by ultraviolet B

light (Beardsley et al., 1997). Table 1.2 give the specific air pollutant and associated health effects and Table 1.3 give the detail of the Malaysia Ambient Air Quality Guideline (MAAQG). The MAAQG was calculated based on the averaging time for each pollutant concentration. For example, the guideline for daily average of  $PM_{10}$  concentration is  $150 \mu\text{g}/\text{m}^3$  and the annual average of  $PM_{10}$  concentration is  $50 \mu\text{g}/\text{m}^3$  (Department of Environment, 2011).

Table 1.2 : Specific Air Pollutants and Associated Health Effects

Pollutants	Health
CO	<ul style="list-style-type: none"> <li>• Reduction in ability of the circulatory system to transport oxygen</li> <li>• Impairment of performance on task requiring vigilance</li> <li>• Aggravation of cardiovascular disease</li> </ul>
NO <sub>2</sub>	<ul style="list-style-type: none"> <li>• Increased susceptibility to respiratory pathogens</li> </ul>
O <sub>3</sub>	<ul style="list-style-type: none"> <li>• reduction in pulmonary function</li> <li>• Coughing</li> <li>• Chest discomfort</li> <li>• Increased asthma attacks</li> </ul>
Peroxyacetyl nitrate, aldehydes	<ul style="list-style-type: none"> <li>• Eye irritation</li> </ul>
SO <sub>2</sub> / Particulates	<ul style="list-style-type: none"> <li>• Increased prevalence of chronic respiratory diseases</li> <li>• Increased risk of acute respiratory diseases</li> </ul>

Source : Stern (1984)

Table 1.3 : Malaysia Ambient Air Quality Guideline

Pollutant Type	Averaging time	Malaysia Guidelines	
		ppm	$\mu\text{g} / \text{m}^3$
Ozone	1 hour	0.10	200
	8 hour	0.06	120
Carbon Monoxide	1 hour	30.0	35
	8 hour	9.0	10
Nitrogen Dioxide	1 hour	0.17	320
	24 hour	0.04	
Sulphur Dioxide	1 hour	0.13	350
	24 hour	0.04	105
Particulate Matter (PM <sub>10</sub> )	24 hour		150
	12 month		50
Total Suspended Particulate	24 hour		260
	12 month		90
Lead	3 month		1.5

Source : Department of Environment Malaysia (2011)

## 1.2 Particulate Matter

Particulate matter (PM) is the term for particles found in the air, including dust, dirt, soot, smoke and liquid droplets. Particulate matter is made up of a number of components, including acids such as nitrates and sulphates, organic chemicals, metals



and soil or dust particles. Particles can be suspended in the air for long periods of time. In the simplest terms, particulate matter is anything solid or liquid suspended in the air (Dockery, 2009). Some particles are large and dark enough to be seen as soot or smoke. Others are so small that they can be detected individually only with specialized microscopes (World Health Organization, 2002).

Some particles are directly emitted into the air from a variety of sources such as vehicles, factories, construction sites, farms and unpaved roads. Other particles may be formed in the air when gases from burning fuels chemically react with sunlight and water vapour. These can result from fuel combustion in motor vehicles, at oil fields and refineries, at power plants and in other industrial processes.

The size of particles is directly linked to their potential for causing health problems. The presence of particulate matter in atmosphere can cause severe health impacts to human health that range from minor nose and throat irritations to more severe impacts such as hospital admissions and premature mortality (Koenig, 2001). Particulate matter is also a useful indicator of several sources of outdoor pollution, such as fossil-fuel combustion (Kunzli et al., 2000).

Particulate matter can be categorized into two categories. The first category is coarse particles such as those found near roadways and dusty industries, ranging in size from 2.5 to 10 micrometres in aerodynamic diameter. The second category is fine particles such as those found in smoke and haze, have aerodynamic diameters smaller than 2.5

micrometres (World Health Organization, 2006). They differ from PM<sub>10</sub> in origin and chemistry. These particles can be directly emitted from source such as forest fires or can be form when gases emitted from power plants, industries and automobiles react in the air. Table 1.4 shows the comparisons between fine and coarse particulate matter.

World Health Organization (2006) has detailed the air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulphur dioxide. Table 1.5 shows the World Health Organization Air Quality Guidelines and Interim Targets for Particulate Matter for Annual Mean Concentrations and Table 1.6 shows the World Health Organization Air Quality Guidelines and Interim Targets for Particulate Matter for 24-hour concentrations. For the annual mean of PM<sub>10</sub> concentration, Malaysia chose to use the interim target-2 which is 50 µg/m<sup>3</sup> and for the 24-hour concentrations, Malaysia chose to use the interim target-1 which is 150 µg/m<sup>3</sup> (Department of Environment, 2009).

Table 1.4 : Comparisons of fine and coarse particle matter

	Fine Particles	Coarse particles
Formed from	Gaseous	Large solids and liquid droplets
Formed by	Chemical reaction ; nucleation; condensation; coagulation; evaporation of fog and cloud droplets in which gases have dissolved and reacted	Mechanical disruption (e.g. crushing, grinding, abrasion of surfaces); evaporation of sprays; suspension of dusts
Composed of	Sulphate; nitrate; ammonium; hydrogen ion; elemental carbon; organic compounds ; particle bound water	Resuspended dust (e.g. soil dust, street dust); coal and oil fly ash; metal dioxides of crustal elements ; sea salt; pollen, mould spores; plant/animal fragments
Solubility	Largely soluble, hygroscopic and deliquescent	Largely insoluble and non-hygroscopic
Sources	Combustion of coal, oil, gasoline, diesel, wood; atmospheric transformation products of NO <sub>x</sub> , SO <sub>2</sub> and organic compounds including biogenic species (e.g. terpenes); high temperature processes, smelters, steel mills, etc	Re-suspension of industrial dust and soil tracked onto roads; suspension from disturbed soil (e.g. farming, mining, unpaved roads) ; biological sources; construction and demolition; coal and oil combustion; ocean spray.
Lifetimes	Days to weeks	Minutes to hours
Travel distance	100 to 1000 kilometres	Less than 1 to 10 of kilometres.

Source : United State Environmental Protection Agency (1996)

Table 1.5 : World Health Organization Air Quality Guidelines and Interim Targets for Particulate Matter : Annual Mean Concentrations

	PM <sub>10</sub>	PM <sub>2.5</sub>	Basis for the selected level
Interim target-1 (IT-1)	70	35	These levels are associated with about a 15% higher long-term mortality risk relative to the WHO Air Quality Guidelines.
Interim target-2 (IT-2)	50	25	In addition to other health benefits, these levels lower the risk of premature mortality by approximately 6% [2 – 11%] relative to the IT-1 level.
Interim target-3 (IT-3)	30	15	In addition to other health benefits, these levels reduce the mortality risk by approximately 6% [2 – 11%] relative to the IT-2 level.
Air Quality Guideline (AQG)	20	10	These are the lowest levels at which total, cardiopulmonary and lung cancer mortality have been shown to increase with more than 95% confidence in response to long-term exposure to PM <sub>2.5</sub> .

Source : World Health Organisation (2006)

Table 1.6 : World Health Organization Air Quality Guidelines and Interim Targets for Particulate Matter : 24-hour Concentrations

	PM <sub>10</sub>	PM <sub>2.5</sub>	Basis for the selected level
Interim target-1 (IT-1)	150	75	Based on published risk coefficients from multi-centre studies and meta-analyses (about 5% increase of short-term mortality over the AQG value).
Interim target-2 (IT-2)	100	50	Based on published risk coefficients from multi-centre studies and meta-analyses (about 2.5% increase of short-term mortality over the AQG value).
Interim target-3 (IT-3)	75	37.5	Based on published risk coefficients from multi-centre studies and meta-analyses (about 1.2% increase in short-term mortality over the AQG value).
Air Quality Guideline (AQG)	50	25	Based on relationship between 24-hour and annual PM levels.

Source : World Health Organization (2006)

Firstly, this research will be concerned with coarse particles that are ten micrometre in diameter (PM<sub>10</sub>) or smaller. Previous researches have shown that particles larger than 10 µm in aerodynamic diameter did not penetrate the body's defences in nose, mouth, and upper airways so it is unlikely to cause respiratory effects (Dockery, 2009). PM<sub>10</sub> generally pass through the throat and nose and enter the lungs. Once inhaled, these particles can affect the heart and lungs and cause serious health effects (Dockery, 2009). Large particles can be trapped in nose and throat and are removed when cough or sneeze. In some areas, particulate matter can be very serious because of high levels of industrial activity.

PM<sub>10</sub> is also a useful indicator of several sources of outdoor air pollution, such as fossil-fuel combustion (Kunzli et al., 2000). Increase in ambient PM<sub>10</sub> concentration can lead to significant impacts on the respiratory health of children, the elderly and susceptible individuals, which are normally associated with reduced lung function, asthma, pneumonia, bronchitis and emphysema (Shamsul, 2002). At extremely high levels and long term exposure, it may even cause death.

Secondly, this research will focus on the short-term prediction of PM<sub>10</sub> concentration, gaseous pollutants and meteorological parameters. The simultaneous time series model will be used in this part. As Malaysia has two seasons which is wet season and dry season by referring to the south west monsoon and north east monsoon (Md Yusof et al., 2010 ; Oliver et al., 2011), the time series modelling will also consider the seasonality.

In Malaysia, all monitoring stations by the Department of Environment are used to observe PM<sub>10</sub> concentration, gaseous pollutants and also meteorological parameters. Figure 1.11 shows the annual average concentration of PM<sub>10</sub> in Malaysia from 1999 to 2011. The Malaysia Ambient Air Quality Guideline (MAAQG) for the annual mean of PM<sub>10</sub> concentration is 50 µg/m<sup>3</sup>. In 1999, there are only 45 monitoring stations but increase to 51 stations in 2003. Another monitoring station was added bringing the total numbers of monitoring stations are 52. The annual average of PM<sub>10</sub> concentration between 1999 and 2011 complied with the MAAQG. Even though the annual average of PM<sub>10</sub> concentration did not exceed the limit, but in 2002, the annual mean was 50 µg/m<sup>3</sup>.

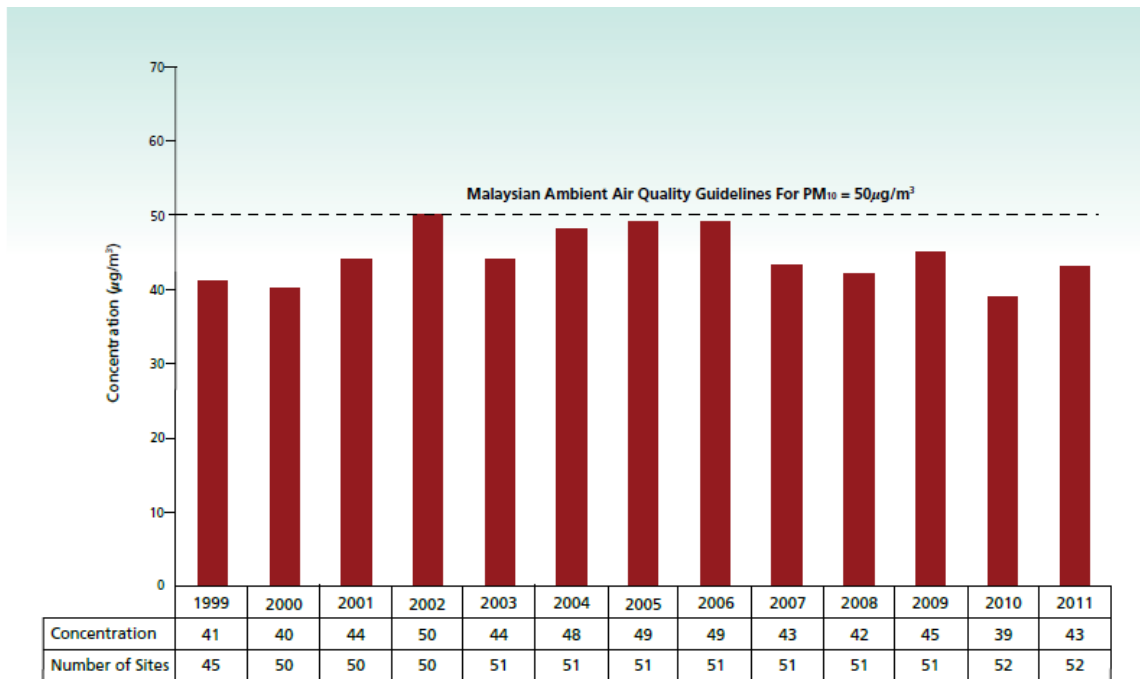


Figure 1.11 : Annual Average Concentration of PM<sub>10</sub> in Malaysia from 1999 to 2011  
Source : Department of Environment (2011)

Figure 1.12 give the detail of the annual average concentration of PM<sub>10</sub> by land type from 1999 to 2011. Almost every year, the urban area recorded the highest mean of PM<sub>10</sub> concentration. In 2002, 2004 and 2005, the annual mean of PM<sub>10</sub> concentration in urban area exceeded the Malaysia Ambient Air Quality Guideline.

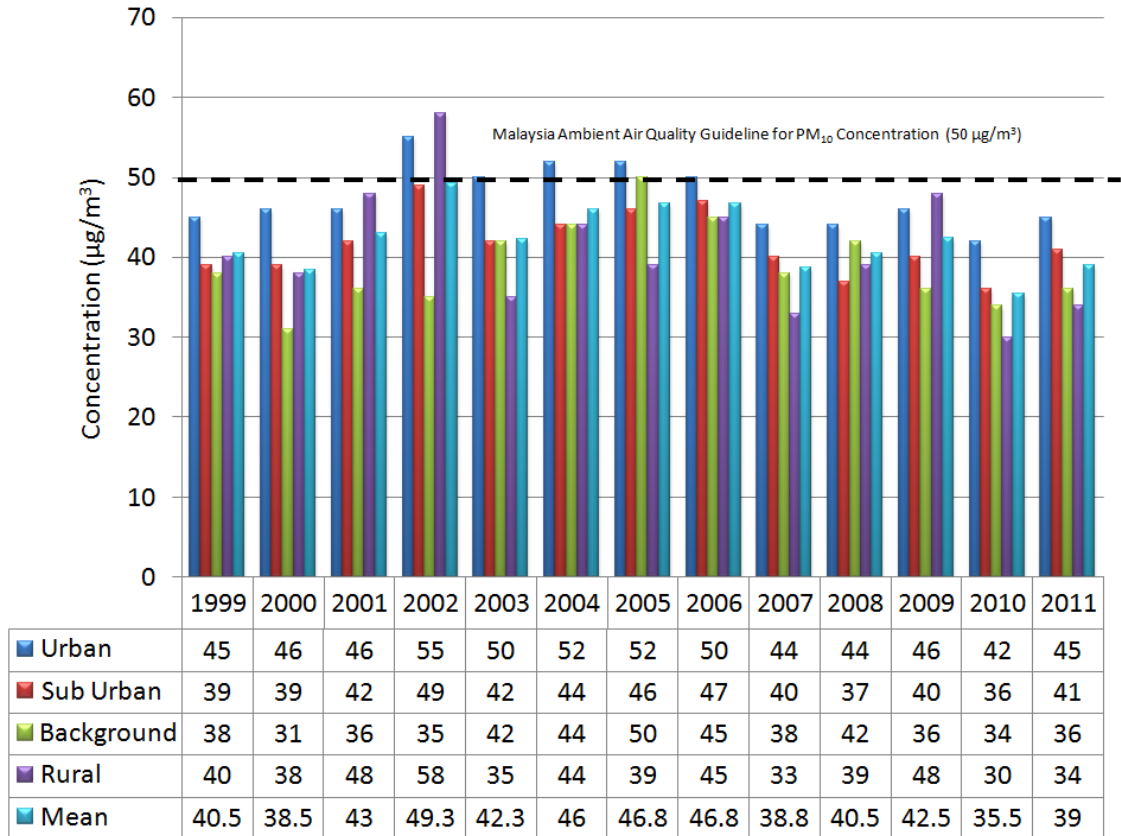


Figure 1.12 : Annual average concentration of PM<sub>10</sub> by land type  
Source : Department of Environment (2011)

The sources of particulate matter in 2011 are given in Figure 1.13. The highest contributor of particulate matter emission was from industrial sources which is 42%. At second place with a total contribution of 25% is from power plant. Motor vehicles were also one of the main contributors with a total contribution of 17% and the remaining from other sources. From Figure 1.13, industrial sources is a major contributor to the particulate matter pollution in Malaysia.



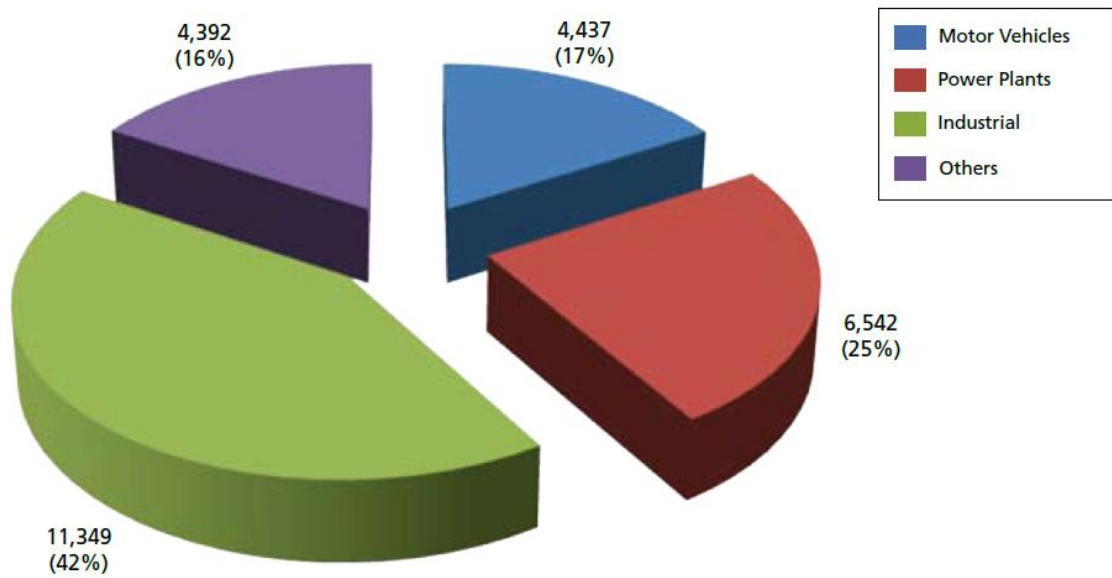


Figure 13: Particulate Matter emission load by sources in Malaysia in 2011  
(Source : Department of Environment Malaysia, 2011)