SHORT-TERM PREDICTION OF PARTICULATE MATTER USING MULTIPLE LINEAR REGRESSION MODEL IN URBAN AREAS IN PENINSULAR MALAYSIA

NURHANIS AINA BINTI MOHD SHARIZAL

FACULTY OF CIVIL ENGINEERING TECHNOLOGY
UNIVERSITI MALAYSIA PERLIS
2022

SHORT-TERM PREDICTION OF PARTICULATE MATTER USING MULTIPLE LINEAR REGRESSION MODEL IN URBAN AREAS IN PENINSULAR MALAYSIA

by

NURHANIS AINA BINTI MOHD SHARIZAL

Report submitted in partial fulfillment
of the requirements for the degree
of Bachelor of Engineering

UNIVERSITI MALAYSIA PERLIS

JULY 2022

ACKNOWLEDGEMENT

Alhamdulillah, thanks to Allah the creator to all being up in the sky and on the earth because of His grace, I am still alive and have to do responsibilities given to me as a Muslim. First of all, I want to thank Dr. Norazrin binti Ramli for guiding me to do this research from start until finish, which is required of me as a student of the Faculty School of Civil Engineering to meet part of the requirement for final year project that relate to the field in environmental and evaluation in final semester and finally graduate in this course. I believe without her support and excellent supervision, this project never has been successful. I will use all the knowledge and experiences that I have gain during the way to finish this project for work and future. Without their advice and guidance, it would be difficult for me to complete my project.

Furthermore, a special thank is also given to the Department of Environment Malaysia who has given me a permission and providing the data needed to carry out this project. To my family especially my beloved parents Zunida binti Abd Rashid and Mohd Sharizal bin Mohd Shariff for always gave a positive morals support during my study and not to forget my siblings Nur Batrisyia and NurBalqis Damia because always provided me with full support and encouragement me along the journey of completing this final year project. Not forget a special to thank all my friends Nurin Haziqah, Fatin Atira and Luqman Jamil and all my friends from School of Environmental Engineering and all my housemates for their support and advice. Finally, nobody is perfect and as human I cannot run away from mistakes, I want to put ten fingers and apologies if there any mistakes during completing my final year project.

APPROVAL AND DECLARATION SHEET

This project report titled Short-Term Prediction of Particulate Matter using Multiple Linear Regression model in Urban Areas in Peninsular Malaysia was prepared and submitted by Nurhanis Aina binti Mohd Sharizal (Matrix Number:181130704) and has been found satisfactory in terms of scope, quality and presentation as partial fulfillment of the requirement for the Bachelor of Environmental Engineering with Honours in Universiti Malaysia Perlis (UniMAP).

Checked and Approved by

(DR. NORAZRIN BINTI RAMLI)
Project Supervisor

Faculty of Civil Engineering Technology
Universiti Malaysia Perlis

JULY 2022

MODEL PERAMALAN JANGKA PENDEK KEPEKATAN ZARAH TERAMPAI MENGGUNAKAN KAEDAH REGRASI LINEAR BERGANDA DI KAWASAN BANDAR SEMENANJUNG MALAYSIA

ABSTRAK

Zarah berdiameter aerodinamik kurang daripada 10µm (PM_{10}) dan zarah berdiameter aerodinamik kurang daripada 2.5µm ($PM_{2.5}$) adalah salah satu bahan pencemar udara yang boleh memberi kesan negatif kepada kesihatan terhadap manusia dan alam sekitar. Tujuan kajian ini adalah untuk meramalkan kepekatan bahan zarah untuk hari berikutnya dan hari berikutnya akan datang (PM_{10D1}, PM_{10D2} dan $PM_{2.5D1}$, $PM_{2.5D2}$) dengan menggunakan model Regrasi Linear Berganda (MLR). Parameter meteorologi dan gas yang digunakan dalam kajian ini adalah zarah terampai hari ini (PM_{10D0}), ($PM_{2.5D0}$), kelajuan angin (WS), suhu (TEMP), kelembapan relatif (RH), sulfur dioksida (SO_2), nitrogen dioksida (NO_2), ozon (O_3) dan karbon monoksida (CO. Data purata harian yang digunakan dalam kajian ini dibahagikan kepada data latihan (80 %) dan data pengesahan (20 %) dan digunakan dari tahun 2018 hingga 2020. Empat (4) stesen pemantuan kualiti udara telah dipilih dalam kajian ini untuk meramalkan kepekatan PM_{10} dan $PM_{2.5}$ untuk hari (PM_{10D1}, PM_{10D2} dan $PM_{2.5D1}$, $PM_{2.5D2}$) iaitu Klang, Bukit Rambai, Batu Pahat dan Kuala Terengganu. Hasil keseluruhan data yang diperolehi dari kajian ini menunjukkan bahawa stesen pemantauan Klang menyumbang kepekatan PM_{10} dan $PM_{2.5}$ paling tinggi berbanding stesen pemantauan yang lain. Ini menunjukkan bahawa Klang adalah kawasan yang lebih tercemar kerana ia dikenali sebagai kawasan yang sangat maju dan dipenuhi kawasan perindustrian. Hasilnya menunjukkan bahawa model Regrasi Linear Berganda (MLR) adalah model terbaik di stesen Klang dalam meramalkan kepekatan PM_{10} dan $PM_{2.5}$ untuk hari berikutnya.

SHORT-TERM PREDICTION OF PARTICULATE MATTER USING MULTIPLE LINEAR REGRESSION MODEL IN URBAN AREAS IN PENINSULAR MALAYSIA

ABSTRACT

Particulate matter with an aerodynamic diameter less than 10µm (PM_{10}) and particulate matter with an aerodynamic diameter less than 2.5µm ($PM_{2.5}$) is one of the most air pollutants that can give negative effect on human health and environment. The purpose of this research is to predict the particulate matter PM_{10} and $PM_{2.5}$ concentration for the next day (Day 1) and the next day (Day 2) by using Multiple Linear Regression (MLR) models. The meteorological and gaseous parameters that are used in this study are particulate matter (PM_{10D0}), particulate matter ($PM_{2.5D0}$), humidity, temperature, sulphur dioxide (SO_2), nitrogen dioxide (NO_2), ozone (O_3) carbon monoxide (CO). The daily mean data that are used in this study are divided into training data (80 %) and validation data (20 %) that applied from 2018 until 2020. Four (4) monitoring stations were selected in this study to predict the PM_{10} and $PM_{2.5}$ concentration for the next day (Day 1) and the next two days (Day 2) (PM_{10D1}, PM_{10D2} and $PM_{2.5D1}$, $PM_{2.5D2}$) which are Klang, Bukit Rambai, Batu Pahat and Kuala Terengganu. The results of overall data that are obtained from this study has shown that Klang monitoring stations contributed the highest mean value of PM_{10} and $PM_{2.5}$ concentration compared to the other monitoring stations. This indicated that Klang is a more polluted area as it is known as a highly industrialised area and urban area. The results shows that Multiple Linear Regression (MLR) is the best model in predicting PM_{10} and $PM_{2.5}$ concentration for the next day is significant at Klang monitoring station.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABREVIATIONS

API	Air Pollution Index
ANN	Artificial Neural Network
ASMA	Alam Sekitar Sdn, Bhd
CO	Carbon Monoxide
CV	Coefficient of Variation
CAA	Clean Air Act
CAQM	Continuous Air Quality Standard
COPD	Chronic Obstructive Pulmonary Disease
D_1	Day 1
D_2	Day 2
DOE	Department of Environment
EPA	Environmental Protection Agency
IA	Index Agreement
MLR	Multiple Linear Regression
MCO	Malaysian Control Order
NAE	Normalized Absolute Error
NAAQS	National Ambient Air Quality Standard
NO_2	Nitrogen Oxide
O_3	Ozone
OLS	Ordinary Least Squares
PA	Prediction Accuracy
PM	Particulate Matter
PM_{10}	Particulate Matter with aerodynamic diameter less than 10μm
$PM_{2.5}$	Particulate Matter with aerodynamic diameter less than 2.5μm
PSI	Pollutant Standard Index (IA)

RMSE	Root Mean Square Error
R^2	Coefficient of Determination
RH	Relative Humidity
RMSE	Root Mean Square Error
SD	Standard Deviation
SO_2	Sulphur Dioxide
Temp	Temperature
U.S.EPA	United State Environmental Protection Agency
VOC	Volatile Organic Compounds
WHO	World Health Organization
WS	Wind Speed

CHAPTER 1

INTRODUCTION

1.1 Background of The Study

Air pollution is a introduces various pollutants into the atmosphere that can causing damage to human beings, wildlife, natural environment and living orgasms (Kinney, 2008; Brauer et al., 2012; Kim et al., 2013). In order for air pollution to occur, there must be a significant quantity of undesired solid or gaseous particles in the air in sufficient numbers to be detrimental to human health and the environment. The health effects of air pollution, observed from both indoor and outdoor environments, have been of great concern due to the high exposure risk even at relatively low concentrations of air pollutants. Any chemical, physical, or biological element that alters the natural features of the atmosphere is considered air pollution. There are several causes of air pollution, including combustion devices in the home, automobiles and industry. Carbon monoxide and ozone are two of the most harmful pollutants to human health, as well as particulate matter and nitrogen dioxide. In addition to respiratory and other disorders, air pollution is a major source of morbidity and death in both indoor and outdoor environments (World Health Organization, 2018).

Research shown by Hao, Zhu and Fan (2014) that China and Southeast Asia have increasing air pollution significant challenge in developing countries due to the drastic increasing of economic growth in urban areas. These cities in China have been ranked as the most polluted in the world. The main problems that made China most polluted cities in the world due to the lack mandatory standards of indoor air quality (IAQ), lack of effective evaluation about air cleaning product, lack of proper maintenance of air cleaners, lack of labelling for the decorating materials (Hao et al., 2014). China always faced problems of air quality indoor (IAQ) because of extensive use of solids fuels for household heating and cooking.

While, in Thailand a study by Arrapongsanuwat and Meesad (2011) reported that major concern of high air pollution in Thailand is due the hazardous emissions from traffic vehicle and industry especially in Bangkok. The high emission of air pollution in Bangkok is due to the high contribution of the adverse effects variety of contaminants emitted into the atmosphere by natural and man-made processes such as industrial emissions, fixed combustions, and traffic vehicle. According to the results of monitoring the air quality surrounding Bangkok, the pollutants that were found exceed the requirements are ozone (O_3) and particle matters (PM) (Thailand Pollution Control Department (PCD), 2007).

In order to preserve the nation's health and the environment, the Clean Air Act (CAA) establishes the National Ambient Air Quality Standards (NAAQS) for each qualifying criteria of air pollutant. The Clean Air Act mandates that EPA has establish National Ambient Air Quality Standards (NAAQS) for pollutants that often found in outdoor air such as particulate matter (PM_{10} and $PM_{2.5}$), sulphur dioxide (SO_2), nitrogen oxide (NO_2), carbon monoxide (CO) and ozone (O_3)

While in Department of Environment Malaysia is the government agency in charge of air quality monitoring in Malaysia These days, there are 51 locations in urban, sub-urban, industrial, and background areas are constantly monitored by the Department of Environment (DOE) for detecting any significant of air quality status that may harm to human health and environmental surrounding. Malaysia and other nations they have their own standards for the maximum permissible concentrations of many types of air pollutants in atmosphere (Department of Environment Malaysia, 2018)

Regarding to DOE, there are 4 continuous air quality monitoring (CAQM) out of 51 station that establish in Malaysia to control air pollutants parameters in ambient air quality. These CAQM are designed to collect the continuously data in 24-hours a day. During the monitoring period, CAQM will include monitoring on gaseous pollutants and meteorology parameters, temperature control enclosures and the collected data will be kept for future used. (Depart of Environment Malaysia, 2018). In this context, Malaysia is using Air Pollution Index (API) to compare the air quality with other nation guidelines. According to the U.S. Environmental Protection Agency (U.S.EPA), Malaysia's API system closely matches the Pollutant Standard Index (PSI)

Malaysia has a long history of rapid growth of development which allowed the country to diversify its economy and become less reliant on agriculture and commodity exports in order to become a more diverse and open economy. The unbalanced urbanisation and quick industrial expansion led to the high levels of pollutants in Southern Peninsular Malaysia (Department of Environment Report, 2018). As an example, air pollution is a situation that has the potential to generate severe economic issues in nations like Malaysia, where tourism is the primary source of income. Urban regions have a dependence on road-based transportation due to the separation of work, commercial, housing, and leisure sectors that leads to high levels of urban air pollution and greenhouse gas emissions that are steadily impacting the quality of life for urban residents (World Health Organization, 2018). The most essential and crucial in urbanizations due to the high increasing number of economic growing in places such Penang, Kuala Lumpur, Shah Alam, Ipoh, and Johor Bharu, among others (Sahrir et al., 2019).

Table 1.1: Malaysia Air Pollution Index (API) (Source: Department of Environment Malaysia, 2021)

Malaysia's API	
Air Pollution Index (API)	Air Quality Category Status
0-50	Good
51-100	Moderate
101-200	Unhealthy
2001-300	Very Unhealthy
301+	Hazardous

Due the Movement Control Order (MCO) issued by the Malaysian Government in 2020 resulted in considerable improvements in air quality in Malaysia, as demonstrated by the excellent and moderate API levels achieved in 2020 compared to the previous year. During 2019, the API result remain mains same as previous years. In 2020, due to the less commercial, industrial, and social activity during the Malaysia Government Movement Control Order (MCO) period air quality in Malaysia is now better than it was in 2019. Malaysia's air quality be much better in 2020 compared with 2019 because of the dry weather and the absence of transboundary haze.

1.2 Problem Statement

In urban areas, particulate matter (PM) is one of the most significant air pollutants, contributing to deterioration in air quality and posing a hazard to human health. Most developing nations and megacities are grappling with rising levels of ambient particulate matter and are often in violation of international environmental treaties. (Azhari et al., 2021). PM emissions in the atmosphere are one of the most serious environmental issues among urban areas. Apart from the effects on air quality, PM has also been associated with mortality and respiratory diseases. Burdens from diseases such as strokes, heart disease, lung cancer, and both chronic and acute respiratory disorders including asthma can be significantly reduced by lowering air pollution levels and a reduction of PM pollution. Road traffic emissions have been one of the largest causes of environmental pollution, with a combination of different contaminants, including carbon monoxide (CO), nitrogen oxides (NO), volatile organic compounds (VOC) and particulate matter (World Health Organization Report, 2018). Particulate matter with an aerodynamic diameter of less than 10 μm (PM_{10}) has been found as a significant air contaminant in major cities across Peninsular Malaysia (Wen et al., 2016).

There has been a significant deal of worry about the health implications of air pollution because of the substantial exposure risk even at low levels of air pollutants. According to the World Health Organization, more than two million people die each year as a direct result of air pollution (Shah et al., 2013). Knowing that emissions from vehicles are the main cause of air pollution, those who stay in big urban areas are the ones who are most likely to be adversely impacted by it. If an industrial catastrophe occurs, the spread of hazardous fumes may be devastating to the residents of nearby towns. The dispersion of pollutants is affected by a wide range of parameters, including the consistency of the atmosphere and the velocity of the wind. (Roya & Parinaz, 2009).

Air pollution among urban areas can led to short-term effect for human health problem. Even though on low air pollution days, it still can a negative influence on health for those who are vulnerable and sensitive about their health conditions. Chronic Obstructive Pulmonary Illness (COPD) cough, shortness of breath, wheezing, asthma, respiratory disease, and high hospitalization rates are all linked to short-term exposure to air pollution (a measurement of morbidity) (Manisalidis et al., 2020). This health problem will include effects on human respiratory such as bronchonstriction and cardiovascular changes, as well as asthma symptoms to chronic effects like respiratory issues. In addition, being exposed to high levels of pollutants in a short period of time might cause early death as a result of respiratory and cardiovascular disorders.

Air pollution may have a number of negative effects on children's health. Perinatal impacts, infant mortality, respiratory problems, allergies, malignancies, cardiovascular disorders, increased oxidative stress, endothelial dysfunction, mental disorders, and vitamin D insufficiency are only a few of other serious consequences (Brunekreef, 2007). For old people, children and people with diabetes and predisposing heart or lung disease especially asthma need extra aware of health protection due to this air pollution that occurred due to the particulate emissions.

The intake of air pollution is strongly connected with the development of respiratory illnesses. These pollutants will enter the body via the airways and build up in the cells as a result. Damage to target cells should be proportional to the amount of each pollutant component present, as well as the source and dosage. The health consequences are also highly depending on the nation, the region, the season, and the time. In addition to the considerations listed above, prolonged exposure to the pollutant should increase the likelihood of long-term health consequences (Manisalidis et al., 2020).

When people are exposed to air pollution, they might experience a variety of cardiovascular side effects. The changes that occur in blood cells as a result of long-term exposure may have an impact on heart function. Long-term exposure to traffic emissions has been linked to coronary arteriosclerosis, but short-term exposure has been linked to hypertension, stroke, myocardial infarction, and heart insufficiency (Kathole & Couri, 2011)

The prediction models are an essential tool since the prediction model is built to decrease autocorrelation or inaccuracy in the model. Statistical modelling has high potential accuracy prediction of PM_{10} and $PM_{2.5}$ concentration. For traffic restriction control at least minimum one or two days are required for mitigation measure of traffic vehicles in order to reduce the possibility risk of critical concentration levels. Short-term forecasting platforms should be created and deployed as a quick alert system in order to inform the public of dangerous air pollution occurrences and to alter air pollution management efforts. The use of forecasting models might help authorities make judgments about enforcing laws when air pollution concentrations above of the set limits (Shahraiyni & Sodoudi, 2016).

A prediction PM_{10} concentration in urban area are more difficult with other pollutants such as nitrogen oxide (NO_2), it is due to the due to the complex interactions among a number of factors (Kukkonen et al., 2003). As a result, despite the fact that an accurate air quality forecasting model is required to notify the public and activate pollution control measures. However, there is no effective action generally taken under conditions of increased PM_{10} concentrations because forecasting models are missing or inadequate (Baklanov et al.,2007). There is a limitation of statistical models that it does not consider into physical aspects of the data and it not applicable to other sites.

Thus, this study is carried out to enhance the prediction of future PM_{10} and $PM_{2.5}$ concentrations in Peninsular Malaysia by using Multiple Linear Regression (MLR) for predicting the future next day (Day 1) and next two days (Day 2) of PM_{10} and $PM_{2.5}$ concentrations. Moreover, the aim of this study is to identify the good models for predicting PM_{10} and $PM_{2.5}$ concentrations in Peninsular Malaysia. Therefore, an effective prediction models at urban monitoring station need to predict the short-term concentration of PM_{10} and $PM_{2.5}$ to empowers authorities and public about preventative precautionary exposure to avoid unhealthy levels of air quality status and implement strategic initiatives that can made improvement of air quality status. Thus, there are few various methods that widely used by researcher for predicting the PM_{10} and $PM_{2.5}$ concentrations in Peninsular Malaysia.

Most of the researcher are found widely used statistical method such as Multiple Linear Regression (MLR), Bayesian Autoregressive, time series, and Markov Chain. In recent years, there has been an increasing interest in machine learning for predicting and forecast ambient air pollution. Several short- and long-term forecasting applications have benefited from machine learning techniques that have been effectively adopted. The Multilayer Perceptron (MLP) and Artificial Network Neural Network (ANN) are the most often used machine learning methods in the prediction of PM_{10} and $PM_{2.5}$ concentrations (ANN) (Nur Shaziayani et al., 2020).

A study by Sayegh, Munir, Habeebullah, (2014), demonstrated comparing statistical models for predictions. PM_{10} concentrations between Multiple Linear Regression Model (MLR) and Quantile Regression Model (QRM) in Makkah. The result showed that Quantile Regression Model (QRM) have capture a better model prediction compared to the Multiple Linear Regression. However, the models monitoring site in Makkah was limited and it can limit the performance comparison of the models.

Thus, this study is carried out to determine the PM_{10} and $PM_{2.5}$ concentrations in short period in Peninsular Malaysia. Since, many of other researchers from other countries this study provided an additional source of effectives the Multiple Linear Regression Model in urban areas and compare with PM concentrations of Malaysian Ambient Air Quality Standard 2020.

1.3 Research Objectives

The objectives of this study are:

i. To determine the descriptive analysis of PM_{10} and $PM_{2.5}$ concentrations and compare with PM concentrations of Malaysian Ambient Air Quality Standard 2020.

ii. To predict short-term PM concentrations of PM_{10} and $PM_{2.5}$ by using Multiple Linear Regression models.

iii. To validate the good model for the next day (Day 1) and the next two days (Day 2) of PM_{10} and $PM_{2.5}$ concentrations.

1.4 Scope of Research

In this research, the selected urban monitoring areas four stations in Peninsular Malaysia. The monitoring data was obtained from Department of Environment (DOE) Malaysia from January 2018 until December 2020. The data will be analysed by using Multiple Linear Regression models (MLR). Annual daily measurements for PM_{10} and $PM_{2.5}$ taken at four stations which Klang, Bukit Rambai, Batu Pahat, and Kuala

Terengganu. This research aim concentrate on two specific pollutants which is PM_{10} and $PM_{2.5}$.

There are 80 % of the monitoring data were utilised for training and 20 % were used for validation of the models according to the results of the study. This predictions of PM_{10} and $PM_{2.5}$ concentration from 2018 until 2020 can be observed through box plot graph. It can be drawn which are related to urban areas ambient air quality standards changes in Peninsular Malaysia. The major contributing factors to predicts PM_{10} and $PM_{2.5}$ can be achieved by determining the association between PM_{10} and $PM_{2.5}$ concentrations with meteorological and gaseous parameters.

Average and maximum daily data of PM_{10} and $PM_{2.5}$ concentrations were utilized to conduct trend analysis. In order to forecast short-term PM_{10} and $PM_{2.5}$ concentrations, the daily data average of PM_{10} and $PM_{2.5}$ concentrations was utilized for statistical models based on multiple linear regression (MLR).

Regarding to this study, the major dependent variables is particulate matter which is PM_{10} and $PM_{2.5}$ while the independent variables for this study are PM_{10} and $PM_{2.5}$ at day 0 (PM_{10} $\mu gm^3$and $PM_{2.5}\mu g\ m^3$), sulphur dioxide (SO_2 ; ppm), nitrogen oxide (NO_2 ; ppm), ozone (O_3 ; ppm), carbon monoxide (CO ; ppm), wind speed (WS ; ppm), relative humidity (RH ; %) and temperature (Temp ; °C).

In order to identify the good model, there are several performance indicators that will be used such as root mean squared error (RMSE), normalized absolute error (NAE), index of agreement (IA), prediction accuracy (PA) and coefficient of determination (R^2). Overall, this study carried out in term of identify the good model for predicting the PM_{10} and $PM_{2.5}$ concentrations with meteorological parameters and gaseous in Peninsular Malaysia for future use of prediction model's research.

1.5 Thesis Outline

Chapter 1 will explain about overview of air pollution, air quality index, ambient air pollutants and problem statement. While for Chapter 2 this chapter will be discussed about the literature review for particulate matter sources of PM_{10} and $PM_{2.5}$ concentration, effect of PM_{10} and $PM_{2.5}$ on human health effect, meteorological parameter influence on air pollution. For literature review, this chapter will explain detailed about model prediction that will be used in this study and other researcher for prediction of PM_{10} and $PM_{2.5}$ concentration.

In Chapter 3, This chapter will explain the procedure of applying this Multiple Linear Regression to predict PM_{10} and $PM_{2.5}$ concentrations. Other than that, performance indicator will be applied for this study research also will be discussed. While in Chapter 4, based on the study goals, results and discussions are discussed. For each monitoring site, the characteristics of PM_{10} and $PM_{2.5}$, meteorological, and gaseous parameters concentrations were reviewed in this chapter's first and second sections; Models for multiple linear regression were developed. In order to summarize the findings of this chapter, each model's performance indicator was determined and included. Lastly, in Chapter 5 also included a summary of the limitations that this research had, as well as recommendations for further progress to be made in the future.

CHAPTER 2

LITERATURE REVIEW

2.1 Particulate Matter

Particulate matter (PM) is a mixture of solids and water vapour found in ambient air. It is widely established that PM is prevalent in ambient air, both inside and outdoors. PM indoors may begin in one of two ways: they may be emitted from the outside or they may be formed organically in the interior environment (Santiasih & Hermana, 2017). Clean Air Act 1970 stated that there are six major pollutants known as the 'criteria' of air pollutants, which includes the particulate matter (PM), ozone (O_3), nitrogen oxides (NO), carbon monoxide (CO), sulphur dioxide (SO_2) and lead (Pb). It is commonly recognised that particulate matter (PM) may be found in the ambient air, as well as the interior and outdoor environments. PM indoors may come from the outside and be emitted into the interior environment, or they may be organically created in the indoor environment. Normally around 80% majority particulate emission occurred from indoor activities. The exposures occurred normally from indoor activities because of inside the house (cooking, cleaning, surface coating and material building), vehicle (combustion and air freshner) and inside an office (copying machine and printer) (Santiasih & Hermana, 2017).

Particulate matter (PM) is known as particle pollution. The phrase used to describe an airborne combination of solid particles and liquid droplets. The term for particulate matter such as dust, smoke, or dirt and it is sufficiently big and black to be visible with the naked eye. It will be identified using an electron microscope at the tiniest scale (United States Environmental Protection Agency, 2013). Particulate matter in the environment consists of a variety of particles of varied sizes and chemical compositions. This study is focused on PM_{10} and $PM_{2.5}$. While physical properties affect the transport and deposition of particles in the human respiratory system, chemical composition determine their impact on health (El-Fadel & Massoud, 2000). The particles of a diameter of less than 10 micrometers (microns) may enter the body via the nose and throat, where they can cause major health problems. Heart and lungs might be affected if the particles are inhaled. (Mohd Zahid et al., 2018).

World Health Organizations (WHO) stated that air pollution can kill around seven million people over the world. Based on WHO data shown that there are almost 99 % the global population breathe air that exceeds WHO guidelines limits. For fine particulate matter ($PM_{2.5}$) the annual mean is 5 μg/m^3 annual mean. Meanwhile for 24-hour mean is 15 μg/m^3. Coarse particulate matter (PM_{10}) the annual mean is 15 μg/m^3 annual mean and for 24-hour mean is 45 μg/m^3.

The United States Environmental Protection Agency (U.S.EPA, 2013) and Malaysia's Department of Environment have established air quality guidelines to assess the concentration of contaminated air that is considered hazardous to public health and the environment. DOE's Recommended Malaysian Air Quality Guidelines (RMAQG) were utilized to set the pollutant concentration in contaminated air in this analysis (Department of Environmental, 2016)

Table 2.1: Malaysia Ambient Air Quality Guidelines

Pollutant	Average Time	Concentration	Standard (2020)
Particulate matter (PM_{10})	1 year	(μg/m^3)	40
	24 hour		100
Particulate matter ($PM_{2.5}$)	1 year		15
	24 hour		35
Sulphur dioxide (SO_2)	1 year		250
	24 hour		80
Nitrogen dioxide (NO_2)	1 year		280
	24 hour		70
Ozone (O_3)	1 year		280
	8hour		180
Carbon monoxide (CO)	1 hour		100
	8 hour		10

This substance is mostly affected by the many sources of emissions, which are in turn affected by the production process, the type of raw material, and how the material is treated. Based on the size of the particles, they can be put into two groups: coarse particles and fine particles. Coarse particles are small pieces of matter that are bigger than 2 m in diameter. Fine particles are smaller than 2 m in diameter. There are two types of PM, called primary and secondary. Secondary particles are made by a chemical reaction between precursor gases and ozone. Primary particles are made by direct emission. (World Health Organization, 2018).

2.1.1 Primary and Secondary Particulate Matter (PM)

Particulate Matter (PM) is primarily formed by a variety of emission sources, such as combustions or mechanical activities that produce dusts and release directly into the atmosphere as solids or liquid droplets. Domestic coal-burning stoves and industrial activities like fossil-fuel-burning diesel engines and biodiesel engines are the chief indoor producers of primary PM. In addition to secondary PM, the printing machine also generates major PM. Black carbon, crustal dust, or soil dust are the most common major PM emissions, which often have a coarse mode. (Santiasih & Hermana, 2017).

As a result, secondary particles are released into the air as a result of chemical reactions that make substances with low volatility. It was then cooled and turned into a solid or liquid, forming PM. This secondary PM could quickly stick to existing particles and form new ones through a process called homogeneous nucleation. These new particles were left behind as a residue of cloud droplets that had dried up. If there were derived gases and ozone in the air, the chemical reaction took place. (Santiasih & Hermana, 2017).

2.1.2 PM_{10} and $PM_{2.5}$

Fine particulate matter, either solid or liquid, that has a diameter of 10 micrometres (μm) or less is called PM_{10}. The diameter of a human hair, on average, ranges from (50μm to 70μm). Floating dust and aerosols are two common terms for fine particles. Long-distance transport is possible because to fine particles that may stay suspended for days or even weeks. Due to precipitation and gravity, larger particles quickly return to the surface. This PM_{10} known as coarse particle due to the size of particulate matters (USEPA 2013).

$PM_{2.5}$ particles are defined as those with a diameter of 2.5 microns or less. Phytoplankton 2.5 ($PM_{2.5}$) particles are discrete that they can only be seen using an electron microscope. $PM_{2.5}$ is the most dangerous kind of air pollution, and it is the most harmful to human health. $PM_{2.5}$ is tiny enough to stay floating in the air for extended periods of time and can be absorbed deeply into the circulation when inhaled. $PM_{2.5}$ also known as fine particles due to the size diameter less than 2.5 μm in diameter or smaller. These $PM_{2.5}$ group of particles also encompasses ultrafine and nanoparticles which are generally classified as having diameters less than 0.1 μm (U.S.EPA, 2013).

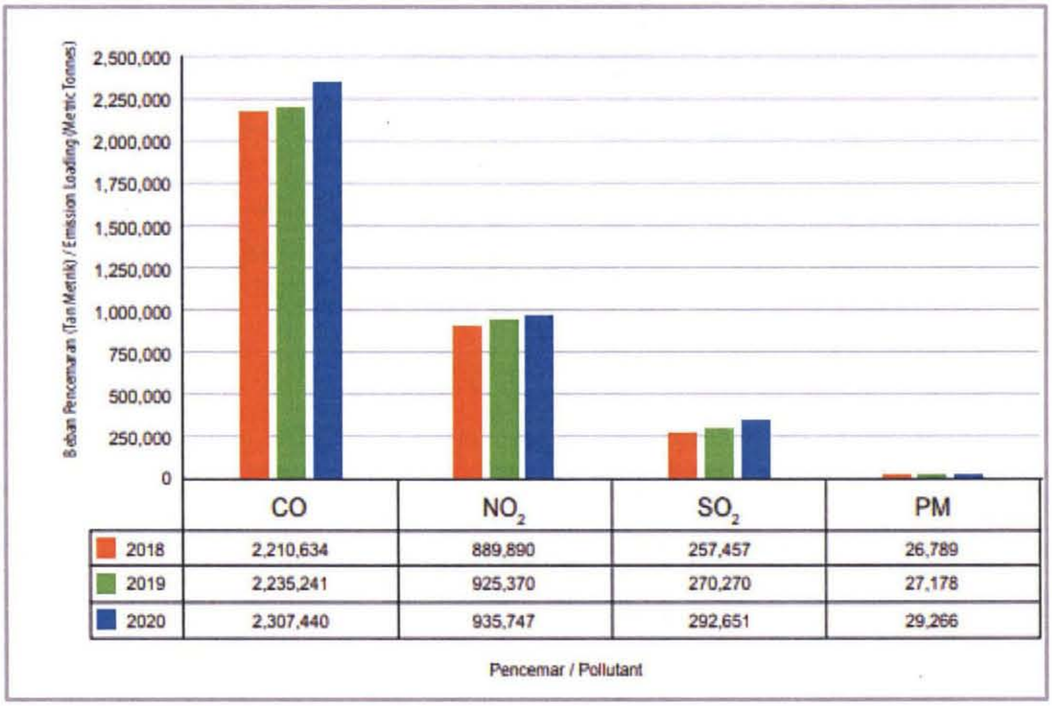
2.2 Sources of PM_{10} and $PM_{2.5}$ in Atmosphere

Both indoor and outdoor sources may produce PM_{10}, which is a pollutant. Primary human emissions, secondary atmospheric reactions, and natural sources are the three major source groups. The term "primary human sources" refers to situations in which airborne particles are generated primarily by human activities. A few examples are mining dust, slash and burn farming, road and construction dust, wood-burning stoves, and fossil fuel power plants. (Burns et al., 2020)

In the home, the PM_{10} comes from various sources such as stoves, space heaters, fireplaces and chimney and tobacco smoke. All these are producing carbon and its contained sulphur dioxide gases. The research, which was carried out in the South African, discovered that inhabitants were more at risk of PM_{10} exposure during the winter months because of home fuel burning. This coal is readily available and affordable, it is the most common source of energy for townships in South Africa. Cooking, smoking, dusting, and vacuuming may all contribute to particle pollution, which is especially prevalent in interior environments. Coal combustion produces fine particles more often than crustal (earth) or biological origins, which generate coarse particles more frequently than coal combustion (U.S.EPA,2013).

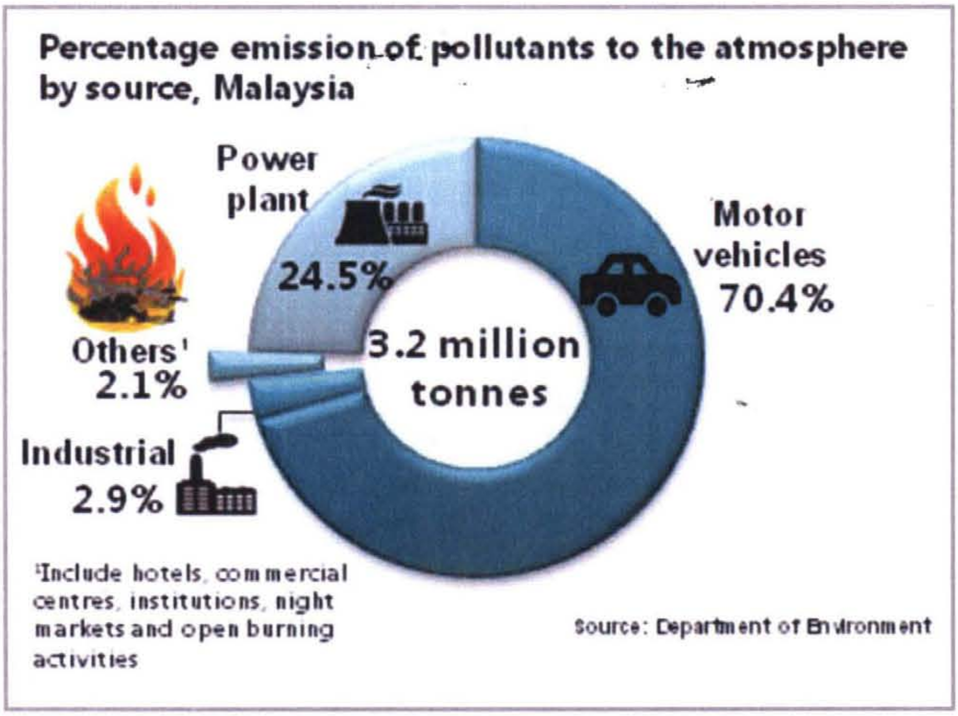
According to the Department of Environment (DOE) Malaysia, there are six main contributor of air pollution which are from power generation, industries, development activities, motor vehicles, open burning, forest fire and land clearing. In early 2020, World Health Organization (WHO) has declared a movement control order (MCO) was issued to limit the spread of the virus in Malaysia by closing down businesses and schools. Due to this MCO, the overall Air Pollution Index (API) 2020 showed a significant dedicated the good sign levels compared with the (API) in 2019 (Department of Environment Malaysia, 2018)

The limited of industrial, commercial social activity are main contributor to the improved of air quality in Malaysia during MCO period. Besides that, the wet days and the absence of transboundary haze in 2020 showed a better air quality condition in Malaysia compared with last few years. Figure 2.1 shown the air pollutant emission from all sources from 2018 to 2020.



	CO	NO_2	SO_2	PM
2018	2,210,634	889,890	257,457	26,789
2019	2,235,241	925,370	270,270	27,178
2020	2,307,440	935,747	292,651	29,266

Figures 2.1: Air Pollutant Emission Load from all sources in 2018 to 2020 (Source: Air Quality Report 2020, Department Environment Malaysia)



Figures 2.2: PM emission load sources from 2018 – 2020 (Source: Air Quality Report 2020, Department Environment Malaysia)

Based on Figure 2.1 in 2020, the total amount of carbon monoxide (CO), nitrogen dioxide (NO_2), sulphur dioxide (SO_2), and particulate matter (PM) emissions is expected to be (2,307,440) metric tons, according to the most recent estimates (PM). According to the Department of Statistics Malaysia, it shown in Figure 2.2 that 70 % of the motor vehicle emission are the highest contribution into air pollution occurred in 2017. This air pollutant sources are followed by power plant sources 24.5 %, industrial 2.9 % and others sources emissions 2.1 % respectively in Figure 2.2. Based on data reported by Department of Environment (DOE), the main contributor in air pollutant is motor vehicles in urban areas. The rising of motor vehicle number especially in urban areas are main contributor to main sources emission of PM_{10} and $PM_{2.5}$ concentration (Department of Statistics (2019); Abdullah et al., 2019).

2.2.1 Motor Vehicles

According to Department of Statistics Malaysia (2019), it showed the reported in Monthly Statistical Bulletin Malaysia, the number of new vehicles registered across 2017 until 2018 as reported in Table 2.2.

Table 2.2: Registration of motor vehicle in Malaysia from 2017-2018

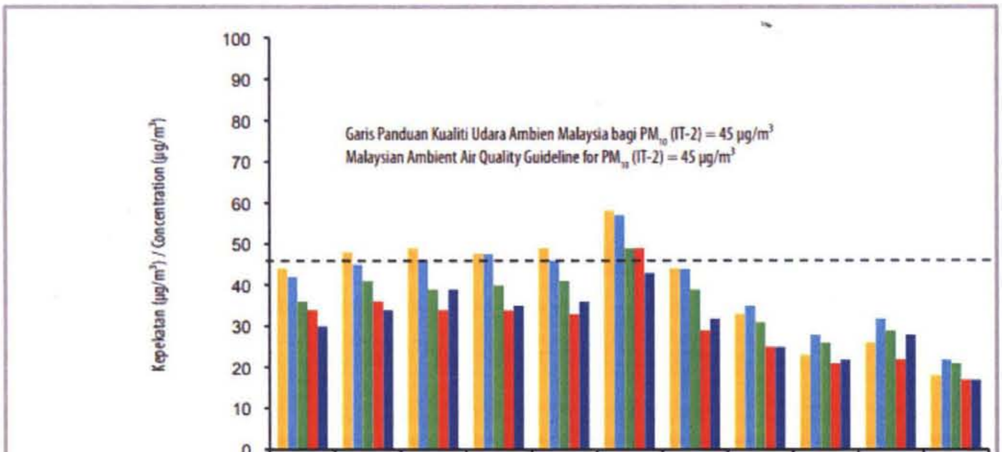
Period	Number					
	Number of vehicles registered through year	Motorcars	Motorcycle	Public Transport	Commercial vehicles	Other vehicles
2017	1,125,900	573,073	495,510	5,534	35,316	16,367
2018	1,218,662	600,313	552,741	8,191	40,715	16,702

The reported new vehicles registered showed that in 2018 the number of vehicles registered recorded an increasing 8.2 % compared with 2017 especially for motorcars which are 600,313 in 2018, followed by motorcycle 552,741, public transport 58,191, commercial vehicles 70,715 and other vehicles 16,702. The highest increasing was contributed by public transports 48.0 %, followed by commercial vehicle 15.3 % and motorcycle 11.5 % in 2018 respectively.

A study by Sahrir et al., (2019), stated that carbon monoxide is a dangerous gas which can pollute into atmosphere due to the formation of oxidation hydrocarbons with other organic compound as combustion process. Motor vehicles are the main contributor sources of carbon monoxide (CO) emission. Traffic vehicle exhaust is one commonly mentioned as sources of carbon monoxide (CO)

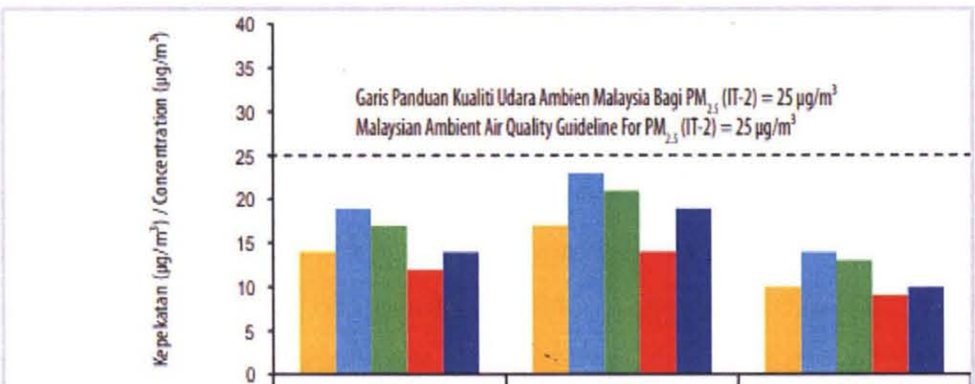
2.2.2 Industries and Power Plant Generation

According to the Department of Environment (2018), it shown that recorded in industrial areas there was high emission of PM_{10} and $PM_{2.5}$ concentration occurred across 2010-2020. Besides PM_{10} and $PM_{2.5}$ concentration, there is sulphur dioxide emission as a main sources of air pollution index increasing.



	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Industri / Industrial	44	48	49	48	49	58	44	33	23	26	18
Bandar / Urban	42	45	46	48	46	57	44	35	28	32	22
Pinggir Bandar / Suburban	36	41	39	40	41	49	39	31	26	29	21
Latarbelakang / Background	34	36	34	34	33	49	29	25	21	22	17
Pendalaman / Rural	30	34	39	35	36	43	32	25	22	28	17

Figure 2.3: Emission of PM_{10} concentration at all areas from 2010-2020



	2018	2019	2020
Industri / Industrial	14	17	10
Bandar / Urban	19	23	14
Pinggir Bandar / Suburban	17	21	13
Latarbelakang / Background	12	14	9
Pendalaman / Rural	14	19	10

Figure 2.4: Emission of $PM_{2.5}$ concentration at all areas from 2010-2020

A study by Sahrir et al., (2019), stated that the majority of the companies of a small or medium size do not have pollution control equipment installed. This results in a rise in the emission of pollutants, particularly in the industrial districts that can be the main sources contributor to the air pollution.

It has been claimed by Abdul Rahman (2017), that China is the most polluted state in the world. It is coal, the most polluting fossil fuel, that has driven China's fast industrial growth. Across the country, new power plants and factories are springing up using old, inefficient, and highly polluting technologies because it is cheaper and faster to develop them than the alternatives.

Industrial toxin emission produces an abundant discharge that affected biological environment that made air pollutant index become unstable due to the toxin emission sources from industries. In addition to fuel-fired boilers, internal combustion gas boilers, and gas stoves, a number of industrial sources are responsible for the production of carbon monoxide (CO). Besides that, carbon dioxide (CO_2) is being released into the atmosphere as a result of the combustion of fossil fuels. Sulphur dioxide (SO_2) is emitted from industrial stacks because fuels have a higher standard concentration of sulphur (Munsif et al., 2021).

2.2.3 Major Contaminants as Contributors to Air Pollution

According to the Department of Environment (DOE), PM_{10} and $PM_{2.5}$ is one of the major contributors that listed in the Recommended Malaysian Air Quality Guidelines (RMAQG). DOE stated that there are six major pollutants occurred in the air due to the emission from industries, motor vehicles and open burning and forest fire such as one (O_3), carbon monoxide (CO), nitrogen dioxide (NO_2), sulphur dioxide (SO_2), particulate matter of less than 10 microns in size (PM_{10}) and particulate matter of less than 2.5 microns in size ($PM_{2.5}$). Pollutant gases and particulate matter come from a variety of sources. Sulphurous and photochemical haze can be caused by six contaminants that have been identified.

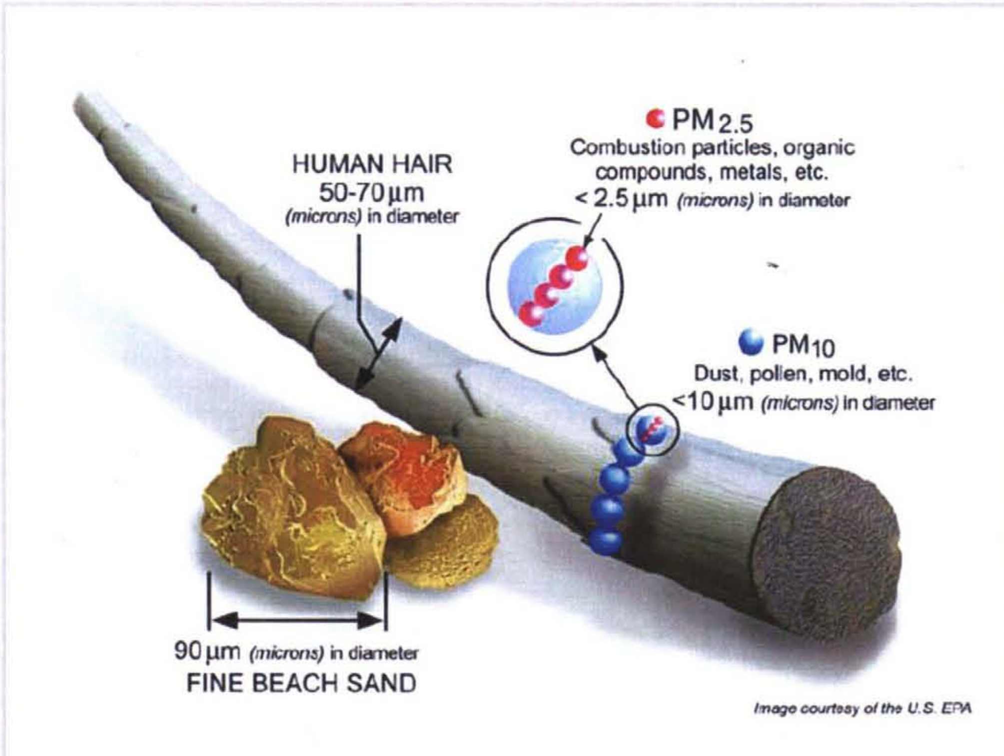
Table 2.3: Emission of pollutants to the atmosphere according to sources (Sources: Department of Environment, Malaysia (2018))

Pollutants	Sources
Ozone (O_3)	Human activity produces and accumulates ozone (O_3) in the atmosphere, which is a source of pollution. The interplay of sunlight, heat, nitrogen oxides, and volatile carbon-containing compounds creates this colorless gas in the troposphere, which has an unpleasant stench.
Carbon Monoxide, (CO)	Incomplete combustion of fuel produces carbon monoxide, an odorless and colorless gas. In the United States, vehicles and engines account for 80 % of CO emissions.
Nitrogen dioxide (NO_2)	Smog and acid deposition are caused in part by the reddish-brown gas NO_2, which is highly reactive and has a bad odor. The vast majority of NOx emissions come from the combustion of automobile engines, electrical utilities, and industrial processes.
Sulphur dioxide (SO_2)	It is a colorless gas that has a strong smell. The majority usage of SO_2 is from industry that resulted a combustion of coal for electricity generation. Only 5 % of the traffic on the roads is generated by automobiles.
Particulate matter (PM_{10} & $PM_{2.5}$)	Dust, soot, and sulphates and nitrates are examples of major contaminants in particulate matter. When breathing, particulate debris can harm the respiratory system. The primary source of PM_{10} pollution is airborne dust, whereas the primary source of $PM_{2.5}$ pollution is human-caused burning.

2.3 Health and Environmental Effect of PM_{10} and $PM_{2.5}$

As a general rule, particles smaller than 10 (μm) represent the greatest risk. In most cases, these smaller particles enter the lungs by the nose and throat. Persons with heart or lung illness, people with diabetes, the elderly, and children are among the most vulnerable to the impacts of these particles when they are breathed (up to 18 years of age). Even while they may irritate the eyes, nose, and throat, larger particles (greater than 10 μm) provide less of a health risk since they seldom penetrate the lungs (U.S.EPA, 2013).

Particulate matter can come from various sources like dust storms, plants, and human-made sources like industry and vehicle emissions. This can increase the number and severity of asthma attacks, as well as consider to bronchitis and other lung diseases worse (Leonardo et al., 2005). High levels of PM pollution were shown to increase the risk of allergies and bronchial hyperresponsiveness in children. It showed that respiratory symptoms in youngsters (with bronchial hyper-responsiveness and high serum total levels of PM) increased by as much as 139 % for every 100 g/m^3 rise in PM.



Figures 2.5: Comparison between human hair with particulate matter diameter. (Source: U.S.EPA 2013)

The composition of particulate particles changes constantly. Variations in composition can be found. Toxins found in particles include metals, organic compounds, dust, and microscopic fragments of living materials. According to the Figures 2.5, aerosol and sub-microscopic particles are both examples of particulate matter, which can have a wide range of sizes. Particulate matter 2.5 ($PM_{2.5}$) is less than the diameter of a human hair.).

Road traffic emissions have been one of the largest causes of environmental pollution, with a combination of different contaminants, including carbon monoxide (CO), nitrogen oxides (NO), volatile organic compounds (VOC) and $PM_{2.5}$. Vehicle emissions are becoming a major source of air pollution concerns as the number of people using the roads to enable urbanisation and industrialization processes continues to grow at an alarming rate. When it comes to measuring traffic-related air pollution in real time, it is both difficult and expensive to do so. As a result, traffic pollution models are used to calculate traffic emissions based on traffic data and vehicle-specific emission factors that vary depending on the type of vehicle and driving conditions. Particulate matter can be either emitted directly from man-made or natural sources, with man-made sources generally resulting in greater amount of $PM_{2.5}$ (Azhari et al., 2021)

Some of the most common man-made sources of $PM_{2.5}$ such as motor combustion, power plant combustion, smokes from fireworks. Moreover, this $PM_{2.5}$ can be occurred from natural sources such as dust soot and dirt. Season, weather, climate, stage of urbanization, nation, and region all have an impact on the major sources of $PM_{2.5}$. Journal of Environmental Science and Technology reported that 33 percent of Central Canada's $PM_{2.5}$ pollution came from the United States. Wildfires were shown to be the dominant source of $PM_{2.5}$ emissions from Central Canada, while household combustion was the primary source in Northern, Atlantic and Western Canada. (Patrick et al., 2003).

It has been well proven that exposure to ambient $PM_{2.5}$ is associated with a higher risk of death, a shorter life expectancy, and a variety of health impacts, including respiratory disorders and cardiovascular diseases, as well as low birth weight. $PM_{2.5}$ microscopic size increases its potential to be lodged deep into the respiratory tracts (Tecer et al., 2008). Pollutants smaller than 2.5 microns are capable of accessing the circulatory system and even the brain at this size. The throat and airways might become affected and breathing difficult as a result of exposure to high amounts of particulate matter. Children and adults and those with heart or pulmonary conditions are the most vulnerable to particle pollution exposure. $PM_{2.5}$ exposure has been linked to a wide range of major health issues, according to several research (Tecer et al., 2008)

2.3.1 Relationship Air Pollutant Index with Health Effect

Since 2018, the $PM_{2.5}$ metric has been thoroughly analysed. From 2010 to 2020, the average yearly air quality data from the monitoring sites was used to compute the trend in air quality, with reference to the Malaysian Ambient Air Quality Standard. At 12 micrograms per cubic meter per year, ambient air quality in 2020 will be lower below the Malaysian Ambient Air Quality Standard IT-2 Guideline value of 25 micrograms per cubic inch per year Due to humid weather conditions, a decrease in forest and bush fire instances in the country, and the lack of transboundary haze incidents in 2020, the $PM_{2.5}$ concentration in 2020 was significantly lower than in 2019. (Department of Environment Malaysia, 2021).

Table 2.4: Health effect based on Air Pollution Index Status

API	Status	Health Effect
0-50	Good	No health impact due to the low pollution expected
51-100	Moderate	Pollution at moderate level does not endanger human health.
101-200	Unhealthy	Heart and lung complication for those high-risk worsen health condition due to the high risk of emission of air pollutant.
201-300	Very Unhealthy	Those with heart and lung problems, physical activity might worsen their condition and reduce their tolerance for it.
>300	Hazardous	High risk of hazardous that can affect public health especially elderly children.
>500	Emergency	Hazardous to high risk to human it can affect respiratory disease such as asthma.

According to the Table 2.4, Air pollution index (API) is a measure of the level of pollution in a certain location. This API measure the average concentrations of SO_2, NO_2,

CO, O_3, $PM_{2.5}$, and PM_{10}. After completed one-hour cycle of data the readings of API status can be obtained. The data air quality was described in terms of the API status and health implication based on the API status.

2.4 Weather Parameter

Air pollution concentrations are affected by weather conditions, the source of pollutants and the area topography. However, the most affected the air quality ambient are meteorological factors. Meteorological factors are emitted in wide range of processes, including evaporation and condensation, vapor deposition, and atmospheric transport and chemical change (Dominick et al., 2012). In Malaysia, according to the Department of Environment the meteorological parameters such as relative humidity (RH), wind speed (WS) and temperature have major influence to the PM_{10} and $P_{2.5}$ concentration in atmosphere (Abdullah et al., 2017). In order to accurately predict PM_{10} concentrations, it is important to use models that take into account weather conditions.

The relationship between air pollution and meteorological parameters was found it is complicated due to the different factors of each season. Meteorological parameters such as air pressure, relative humidity, and wind speed all have a significant impact on PM concentrations (Tian et al., 2014). As a result, managing particulate matter is critical due to the unpredictable nature of weather, climate change, and air pollution. A study by Tian et al (2014), air quality in summer was the best can be affected on PM concentration. In general, air pressure, relative humidity, and wind speed were the most relevant meteorological parameters impacting PM_{10} concentrations.

2.5 Multiple Linear Regression for Predicting Particulate Matter

The Multiple Linear Regression (MLR) models is frequently used for prediction on past and previous data which can help for decision making. This model is based on a simple computation and straightforward implementation of the link between the dependent variable and multiple independent variables, such as weather conditions and gaseous contaminants (Mirsha et al., 2015).

Malaysian researchers have used the MLR model to estimate the concentration of PM_{10} and $PM_{2.5}$ on the East Coast Peninsular Malaysia, based on several site classifications of rural, suburban and urban areas and during different types of monsoons (Yuen et al, 2018). In addition, the MLR was used for forecasting PM_{10} concentration during haze transboundary in Malaysia based the Air Quality Monitoring Station (AQMS) from the Department of Environment (DOE).

A study by Lola et al., (2016), demonstrated improving the prediction model of multiple linear regression (MLR) by approach combining with artificial neural network (ANN) for better predict future) of PM_{10} concentrations. These studies are selected Kuala Terengganu as their monitoring station for this study. Due to the minor errors created, error reduction, and enhanced correlation coefficient for all parameters in MSE and MAE, respectively, the findings demonstrated that their suggested model has the potential to improve the performance of the model in comparison ANN and MLR. It became clear that the model had developed was superior than ANN and MLR in terms of its efficiency, precision, and near-perfect correlation.

Nowadays, there are various of statistical tools for predicting the PM_{10} and $PM_{2.5}$ concentrations have been used widely in Malaysia. Multiple Linear Regression (MLR) is used to determine more than one predictor for a certain situation and it can simultaneously predict using several independent variables. Many researchers use multiple regression analysis (MLR) as a way to represent how a dependent variable responds to many independent factors. A big data studies of air quality studies can be associated with weather parameters in terms of dependent and independent response via statistical tool of MLR. This MLR tools are extremely useful in providing information that allows the authority and community to take precautionary measures to avoid or limit their exposure of unhealthy level air quality and made improvements of air quality at specific locations. (Abdullah et al., 2017).

As mentioned by Ul-Saufie et al., (2012), Robust Regression Method was used. Robust Regression Method it is an excellent alternative approach because, when compared to the regular least square method, it has the capacity to reduce the effect of outliers in data sets to a significant extent. The Gamma distribution is the most appropriate for representing PM_{10} and $PM_{2.5}$ concentration in industrial regions. Robust regression works by assigning a weight to each data point. If a substantial change occurs in a tiny area of the data, the Robust method is less sensitive than ordinary least squares (OLS). (Ul-Saufie et al., 2012).

CHAPTER 3

METHODOLOGY

3.1 Research Flowchart

In this study, there are four stations in Peninsular Malaysia were selected as study areas as monitoring stations. This study used the daily average monitoring statistics from 2018 to 2020 that gained from Department of Environment Malaysia (DOE). The first objective of this study to identify descriptive analysis which is (mean, median, mode, standard deviation, skewness and kurtosis) of PM_{10} and $PM_{2.5}$, gaseous parameters (nitrogen dioxide NO_2, sulphur dioxide SO_2, and carbon monoxide CO) and meteorological parameters (relative humidity and ambient temperature).

The next stage is to evaluate data using MLR to forecast PM_{10} and $PM_{2.5}$ concentrations for the next day (Day 1) and next two days (Day 2) to obtain the good model. Five performance measures are R^2, IA, PA, NAE, and RMSE. Figure 3.1 shows flowchart of this study.

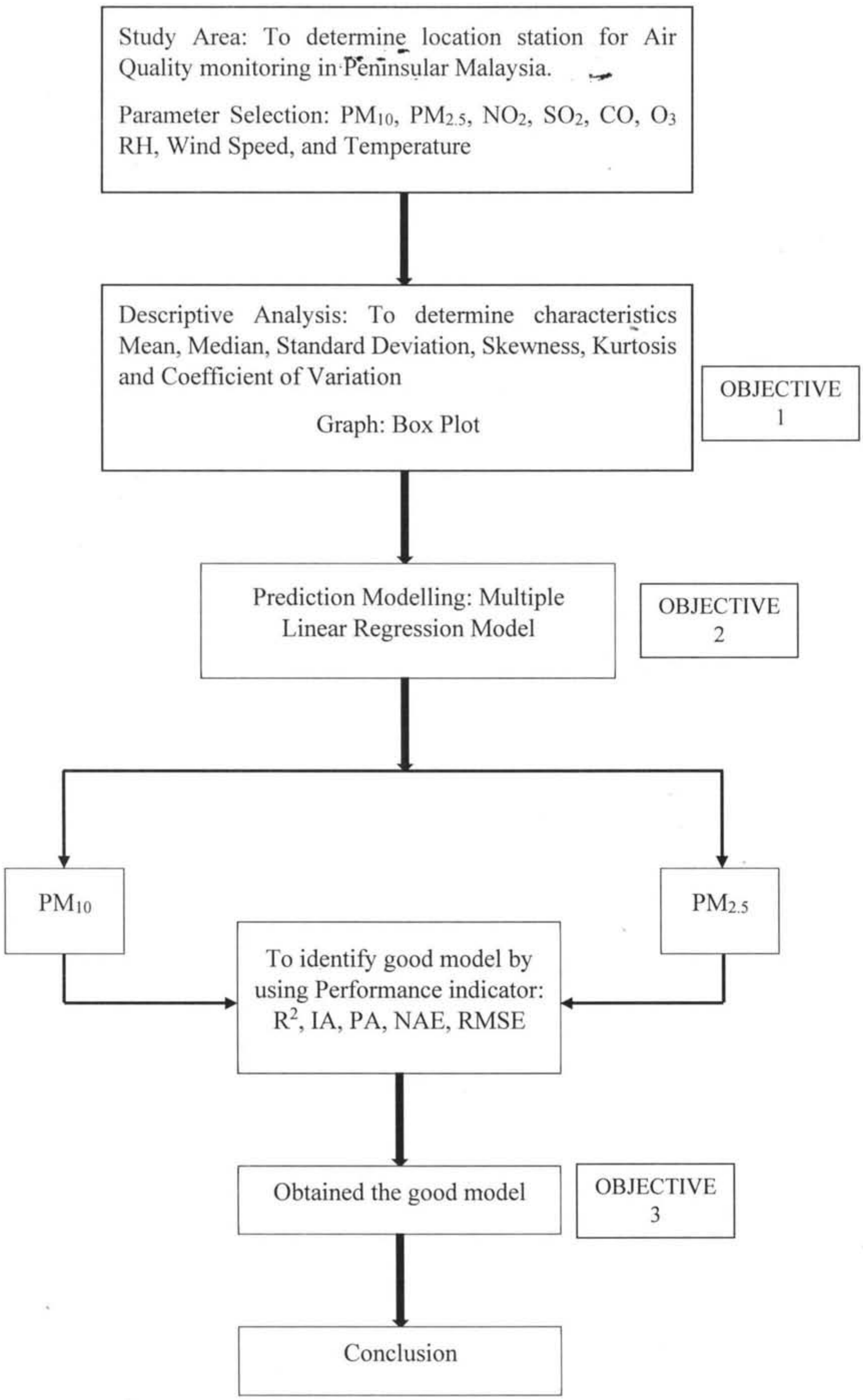


Figure 3.1: Research Flowchart

3.2 Study Area

There were four stations were used for predicting the PM_{10} and $PM_{2.5}$ concentrations. All these four stations are from urban areas which were (Selangor, Melaka Johor, and Terengganu). All these four stations are expected to be highly polluted due to rapid traffic vehicles in urbanization areas. The data was collected through a continuous monitoring by Alam Sekitar Sdn. Bhd. (ASMA) and on behalf control by Department of Environment Malaysia (DOE).

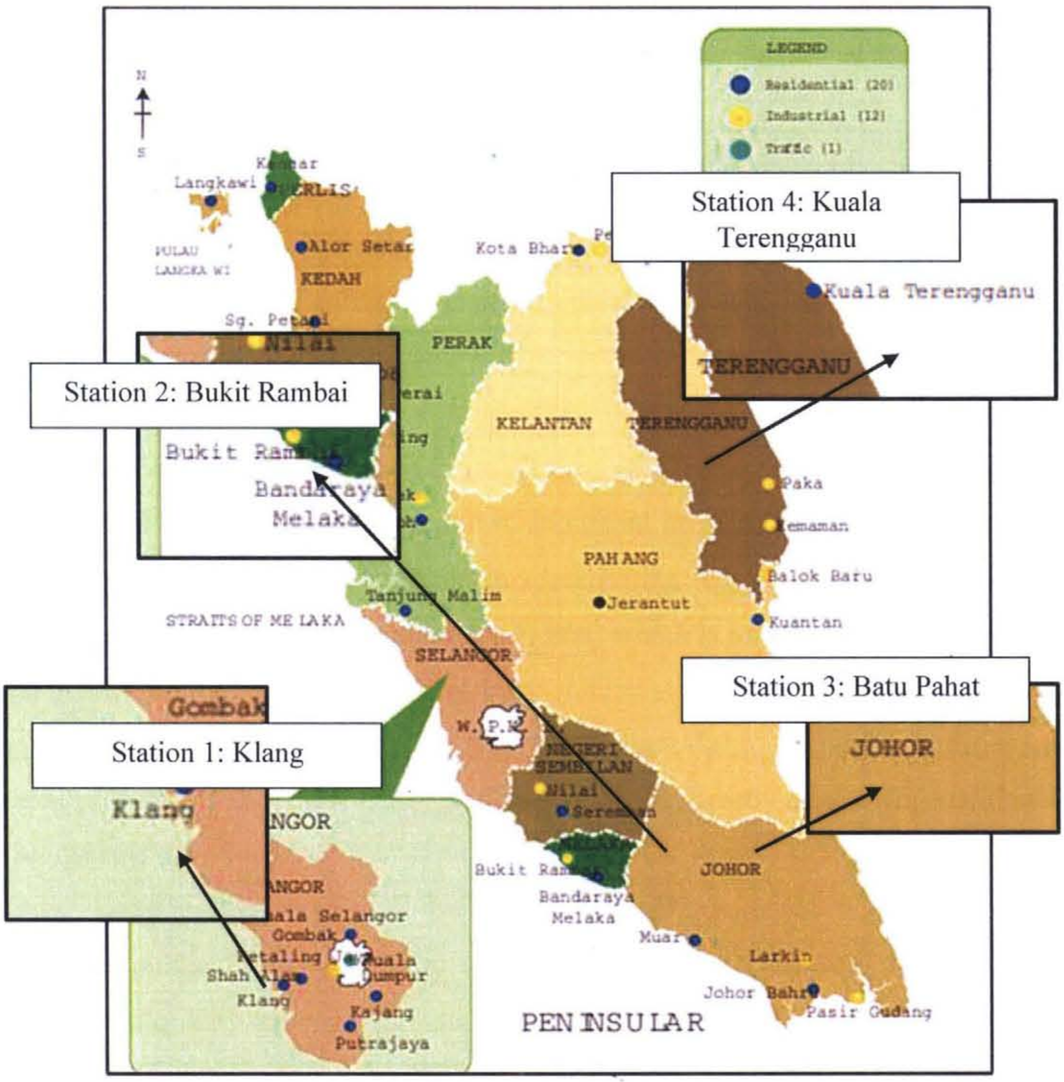


Figure 3.2: Air Monitoring Stations in Peninsular Malaysia (Source: Malaysia Environmental Quality 2018).

The monitoring station at Selangor is Klang. Klang is one of Malaysia's most important economic centres. The increasing urbanisation in Klang in last decades has led to a broad range of environmental problems which is air pollution issues. The main air pollution emission that occurs in Klang are from traffic vehicles, industries and shipping activity. In Malaysia, Klang known as industrial area this will be proven that Klang has high levels of unburned hydrocarbons emitted by motor vehicles and other oil and gas-related operations, as well as the high levels of SO_2 emitted by the high reliance on sulphur-based fuels for industrial production and electric power generation (Abdullah et al., 2012)

Furthermore, Klang is chosen as monitoring stations because it is located nearby housing area and Port Klang. Port Klang is one of Malaysia's most important industrial centres, with a large number of heavy vehicles passing through. Ahmad et al., (2012) reported that the number demand of cars or vehicles in Klang are increasing. In 2009, there were roughly 19.02 million registered vehicles in Malaysia, which nearly 68 % of Malaysians had at least one vehicle.

Batu Pahat, Johor is chosen as monitoring station because the concentration SO_2, NO_2 and CO is highly increasing due to industrial area and the emission from heavy vehicle from the industrial. Moreover, this monitoring station Batu Pahat is nearby with University Tun Hussein Onn Malaysia (UTHM) which is in student population area and commercial area that surrounded by busy vehicles and transportation. Batu Pahat is chosen as monitoring station because it is busiest place due to the public facilities such as banks, post office, shop lots, petrol station that located nearby monitoring station (Radin et al., 2016).

This public facility is nearby with UTHM campus. Therefore, it has increasing ambient air quality among staff and university student. Additionally, the station is surrounded by three to four-story buildings, which allows the particulate matter and black carbon occurred from motorised vehicles (Radin et al., 2016).

Bukit Rambai is chosen for this study because it is located nearby with industrial park area that contributed to high concentration of particulate matter from industry heavy vehicles and from the industry chimney. The Bukit Rambai station is surrounded with residential area and public facilities such as station bus. It also tends to present high of particulate matter emission due to the construction activities residential nearby, industrial emission and heavy traffic vehicles.

Kuala Terengganu known as capital city in Terengganu, there are a lot of school surrounded in Kuala Terengganu. Around 80 primary schools entire of Kuala Terengganu, industrial park and tourism centres. Therefore, Kuala Terengganu are selected as monitoring station due to high particulate matter occur from heavy traffic vehicles. This coastal urban region has been increasing fast over the last several years, and it reached the urbanization level of 59.1 % in the year 2010. The state of Terengganu has a total of three monitoring stations for the quality of the air, and the Department of Environment is in charge of overseeing all of them. Each station is situated in a separate district. The location of the air quality monitoring station that was selected for this research can be found in the state capital of Terengganu, which is known as Sekolah Men. Keb. Chabang Tiga (Abdullah et al., 2017).

3.3 Descriptive Analysis

The descriptive analysis is used to identify characteristics of the monitoring data extracted. It helps to describe and summarize data points in constructive ways such as showing it in graph plot. Thus, the descriptive analysis will consist tendency measure such as mean, median, mode and standard deviation are presented. Apart from descriptive analysis, the graph box plot and shape measure (skewness and kurtosis) were performed as well.

3.3.1 Box Plot

In order to determine the characteristics of PM_{10} and $PM_{2.5}$ concentrations, the box plot were plotted. In simplest form, the boxplot presents five sample statistics. It provides the lower, upper, maximum quartile, median and skewness. The length of the box indicates the range of the sample, while the line across the box indicates the sample's centre. The sample will symmetric or skewed either left or right by proving from positions of the box in its whisker and the line of the box. The following diagram shows a standard box plot and whisker plot. A box plot is a graph of statistical data that shows the second and third quartiles as two rectangles. The median is shown by the vertical line, while the lower quartile (LQ) and upper quartile (UQ) are shown by the sides of the rectangle (UQ). It is important to use a box plot to find outliers and extreme values where the monitoring record is larger than the whisker.

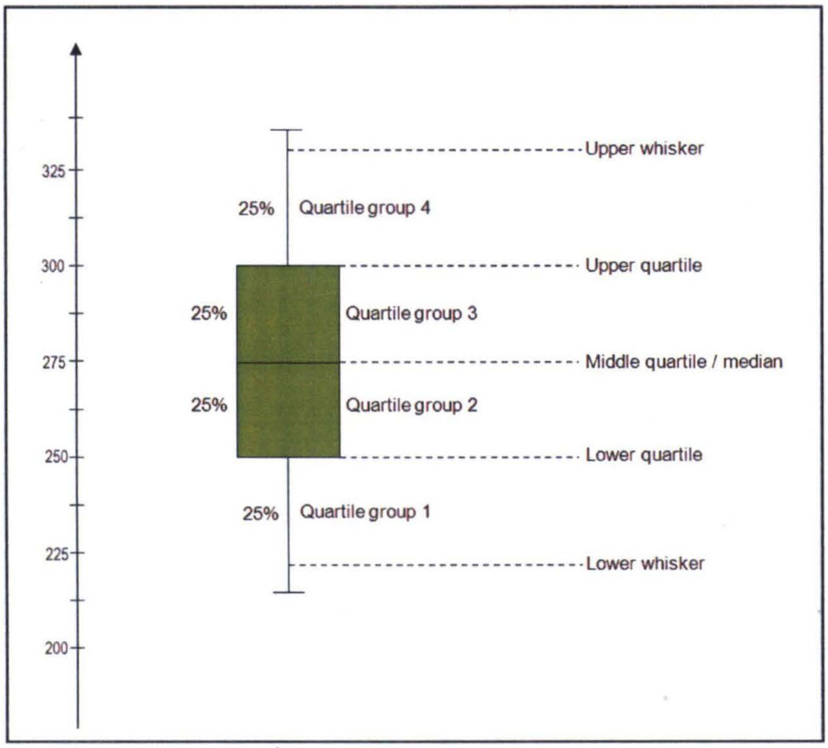


Figure 3.3: Standard box and whisker plot

3.3.2 Standard Deviation

The standard deviation is the dispersion of data in normal distribution. Normally, standard deviation known as indicates that represents accurately the mean of sample data. There are five general measures of dispersion which is variance, standard deviation, range, minimum and maximum

3.3.2 Skewness and Kurtosis

Skewness is a way to measure symmetry, or more specifically, how symmetrical the normal distribution is. Kurtosis is a way to measure how much a distribution has peaks. kurtosis is sometimes used to refer to the original kurtosis value (proper). If a distribution or a data set at the centre point, therefore the mean, median and mode are called symmetric distribution (skewness = 0, kurtosis = 0). A normal skewness and kurtosis if the data in range (-1 and +1) (Mishra et al., 2019).

3.4 Multiple Linear Regression (MLR) Model

Multiple Linear Regression (MLR) is used to create a linear model for any dataset by integrating linear regressions. In metropolitan regions, multiple linear regression (MLR) has been widely utilized to anticipate PM_{10} and $PM_{2.5}$concentrations. As a result, it suffers from a lack of precision due to the linear representation of non-linear systems as well as a lack of ability to capture extreme values (Shahraiyni and Sodoudi, 2016)

Multiple linear regression models are used when the research variable is dependent on more than one explanatory or independent variable. In two aspects, this model extends the standard linear regression model. While it does not allow for arbitrary forms, it permits the mean function to rely on more than one explanatory variable and have shapes other than straight lines.

$$PM_{10,\,D1} = \begin{array}{l} \beta_0 + \beta_{1PM10,\,D0} + \beta_{2PM2.5,\,D0} + \beta_3 SO_{2\,i} + \beta_4 NO_{2i} + \beta_5 O_{3i} + \\ \beta 6CO_i + \beta 7WS_i + \beta 8RH_i + \beta 9TEMP_i + \varepsilon_i \end{array} \tag{3.1}$$

$$i = 1, 2, \ldots, n$$

where;

- $\beta_0, \beta_1, \beta_2, \ldots\ldots \beta_9$ where it is coefficients of the regression model
- ε_i random error
- The independent parameters ($PM_{10,D0}$, $PM_{2.5,D0}$, SO_2, NO_2, O_3, CO, WS, RH and TEMP)
- The probability distribution of ε is normal
- Random errors are independent

Methods have been developed to enable regression models to approximate linear relationships and non-additivity in order to get better approximations. It is possible to make use of nonlinear regression models in certain cases, however doing so will increase the modelling effort. Linear components may be included in a regression model by an experienced user of multiple regression.

In this study, multiple linear regression model was used for determines the prediction of PM_{10} and $PM_{2.5}$ concentrations for the next day and next two days. Therefore, the parameters that will used in this model are temperature (T), relative humidity (RH), nitrogen dioxide (NO_2), sulphur dioxide (SO_2), carbon monoxide (CO), ozone (O_3), and particulate matter (PM_{10} and $PM_{2.5}$).

3.5 Performance Indicator

In order to identify the good model, the performance indicators will be utilised to predict the PM_{10} and $PM_{2.5}$ concentration. Among the performance indicator that has been used are mean absolute error (MAE), root mean square error (RMSE), and the one that measures the accuracy of the prediction (PA), the coefficient of determination (R^2) normalized absolute error (NAE) and index of agreement (IA) (Ul-Saufie et al., 2012).

Table 3.1: Performance Indicators (Ul-Saufie et al., 2013)

Performance Indicators	Equation	Explanation
Root mean square error (RMSE)	$\mathrm{RMSE} = \frac{1}{n-1} \sum_{i=1}^{n} (Pi - Oi)^2$	The best model has the values closer to zero (0) or the smallest values.
Normalized absolute error (NAE)	$\mathrm{NAE} = \frac{\sum_{i}^{n}=1 \lvert Pi-Oi \rvert}{\sum_{i}^{n}=1\, Oi}$	
Prediction Accuracy (PA)	$\mathrm{PA} = \left(\frac{\sum_{i}^{n}=1(P_i - \bar{O})^2}{\sum_{i}^{n}=1(Oi - \bar{O})^2} \right)$	The best model has the values closer to one (1)
Index of agreement (IA)	$\mathrm{IA} = 1 - \left(\frac{\sum_{i}^{n}=1(P_i - Oi)^2}{\sum_{i=1}^{n}(\lvert P-\bar{O} \rvert + \lvert Oi - \bar{O})^2} \right)$	
Coefficient of determination (R^2)	$R^2 = \left(\frac{\sum_{i=1}^{n}(Pi-P)(Oi-O)}{n.Spred.Sobs} \right)$	

n = Number of observed data, O_i = Observed data, $\bar{O}$ = Mean of observed data
P_i = Predicted data, $\overline{P}$ = Mean of predicted data,
S_{obs} = Standard deviation of observed data
S_{pred} = Standard deviation of predicted data

CHAPTER 4

RESULTS AND DISCUSSIONS

4.1 Introduction

This chapter will present a discussion of this study to achieve the objective. The first objective of this study, determining the PM_{10} and $PM_{2.5}$ characteristics by comparing them with the Malaysian Air Ambient Quality Guideline (MAAQS). The second objective was to apply Multiple Linear Regression models to identify the good model. These performance models obtained an equation to identify the best model for predicting PM_{10} and $PM_{2.5}$ concentrations. Based on the equations generated by Multiple Linear Regression of PM_{10} and $PM_{2.5}$ concentrations, a good model for four selected monitoring stations across three years 2018 to 2020 was determined. The performance indicators for multiple linear regression models (MLR) predictions of PM_{10} and $PM_{2.5}$ concentrations will be evaluated with the actual value in order to achieve a good model to forecast the short-term prediction PM_{10} and $PM_{2.5}$ concentrations.

4.2 Descriptive Analysis of PM_{10} and $PM_{2.5}$

The statistical analysis was carried out by applying it to the daily averages of the yearly concentration of PM_{10} and $PM_{2.5}$. It is important for determination of air pollution status and it also providing an explanation for distribution of variables. This study will provide a summary of PM_{10} and $PM_{2.5}$ concentration at four monitoring station in Peninsular Malaysia which are located at Klang, Bukit Rambai, Batu Pahat and Kuala Terengganu. These statistical analyses were used to compared with the relationships between PM_{10} and $PM_{2.5}$ concentration with Malaysian Ambient Air Quality Guidelines (MAAQG). Statistical package for social sciences (SPSS) was used to perform control measures and evaluations on the raw data in order to guarantee it had acceptable levels of data quality.

4.2.1 Descriptive Analysis of PM_{10}

As shown in Table 4.1, during the study period at all station the average concentration of PM_{10} was occurred at Klang station was the highest mean (135.10 µg/m^3) and followed in decreasing order by Bukit Rambai (87.44 µg/m^3), Batu Pahat (67.49 µg/m$^{3)}$ and Kuala Terengganu (62.07 µg/m3) respectively. However, Klang station daily PM_{10} concentration was exceed (100 µg/m^3) due to Malaysian Ambient Air Quality Guidelines.

Table 4.1: The descriptive analysis of PM_{10} concentration for 2018-2020

	Klang	Bukit Rambai	Batu Pahat	Kuala Terengganu
Mean	**38.83**	29.52	23.86	25.83
Median	**36.83**	25.44	21.65	24.29
SD	**16.38**	11.82	10.30	10.23
CV	0.41	0.38	**0.43**	0.39
Skewness	**1.83**	1.45	1.23	1.33
Kurtosis	**7.46**	4.07	2.55	4.61
Maximum	**135.10**	**87.44**	**67.49**	**62.07**

*SD: standard deviation CV: coefficient of variation

While for the kurtosis PM_{10} characteristics, the distribution for all location was skewed to the right. The kurtosis shown positive value for all monitoring station. Based in Table 4.1, Klang has the highest skewness among of other stations which was (1.83), followed by Bukit Rambai (1.45), Kuala Terengganu (1.33) and the lowest skewness was located at Bukit Rambai. Overall, all the monitoring station are indicating normal distribution because the skewness value is above one (1). A distribution is called approximate normal if skewness or kurtosis (excess) of the data are between (– 1) and (+ 1). (Gupta et al., 2019).

Furthermore, the coefficient of variation provides the result as a ratio of the standard deviation to the mean value, which is given as a percentage. CV = 100 Times (SD/mean) (Gupta et al., 2019). The coefficient of variance (CV) that showed in table above is to measure the dispersion of the data.

According to the CV values on Table 4.1, the dispersion of PM_{10} concentration at all monitoring stations it shown the greatest CV values was at Batu Pahat (0.43), followed by those in Klang (0.41), Kuala Terengganu (0.39), and Bukit Rambai (0.38). It was shown that Kuala Terengganu (0.39) and Bukit Rambai (0.38) roughly had the same variance for the PM_{10} concentrations.

As shown in Figure 4.1 – Figure 4.4, box plots of PM_{10} concentration for three years at all monitoring stations. Box plots were used to identify the concentration outliers. All of the monitoring stations have showed PM_{10} concentrations at each station across 2018 to 2020. The outlier is most occured in Klang, Bukit Rambai, Batu Pahat and Kuala Terengganu

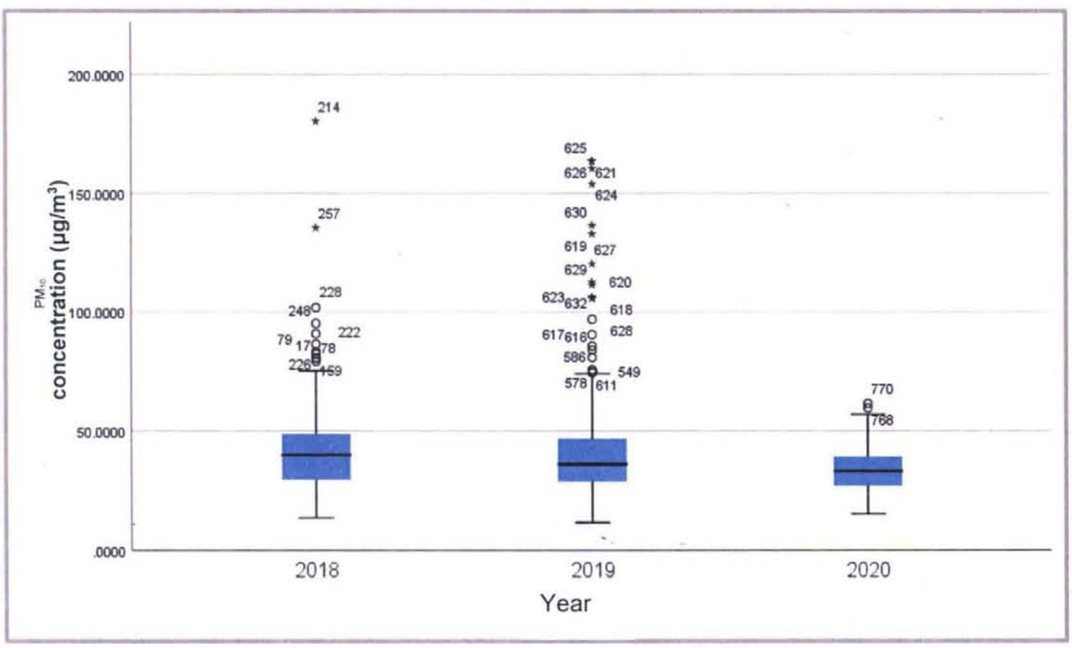


Figure 4.1: Box plot of PM_{10} at Klang Monitoring Station

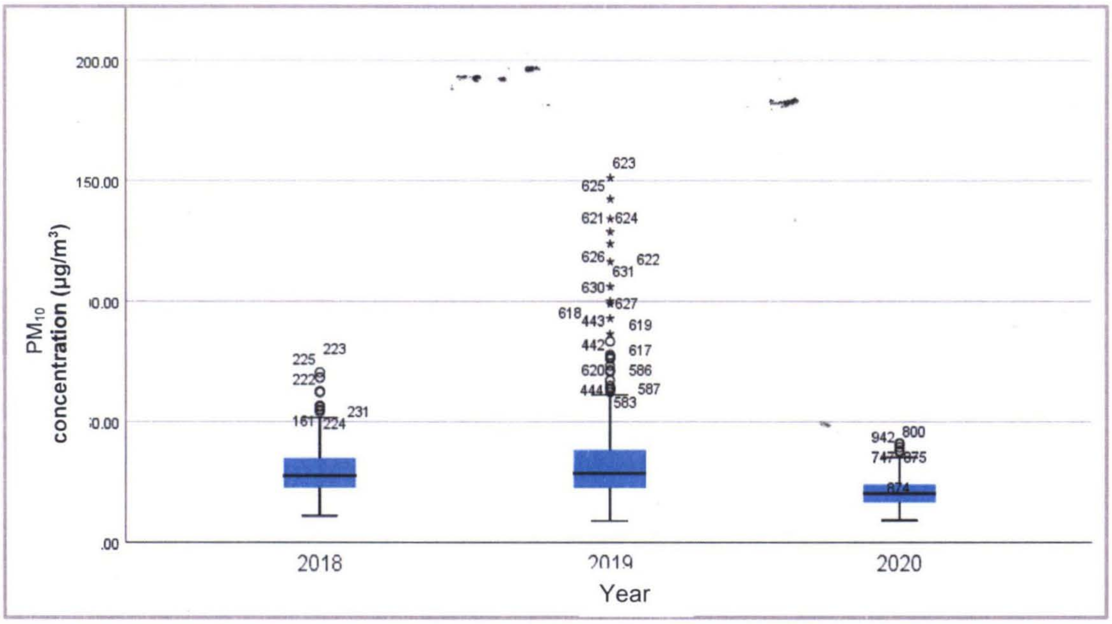


Figure 4.2: Box plot of PM_{10} at Bukit Rambai Monitoring Station

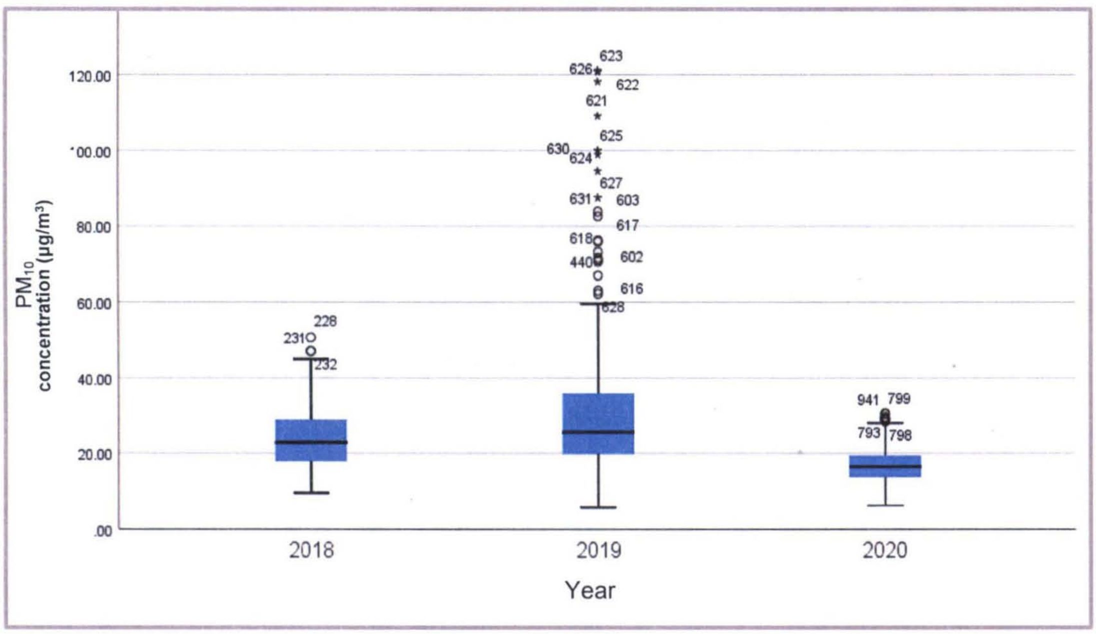


Figure 4.3: Box plot of PM_{10} at Batu Pahat Monitoring Station

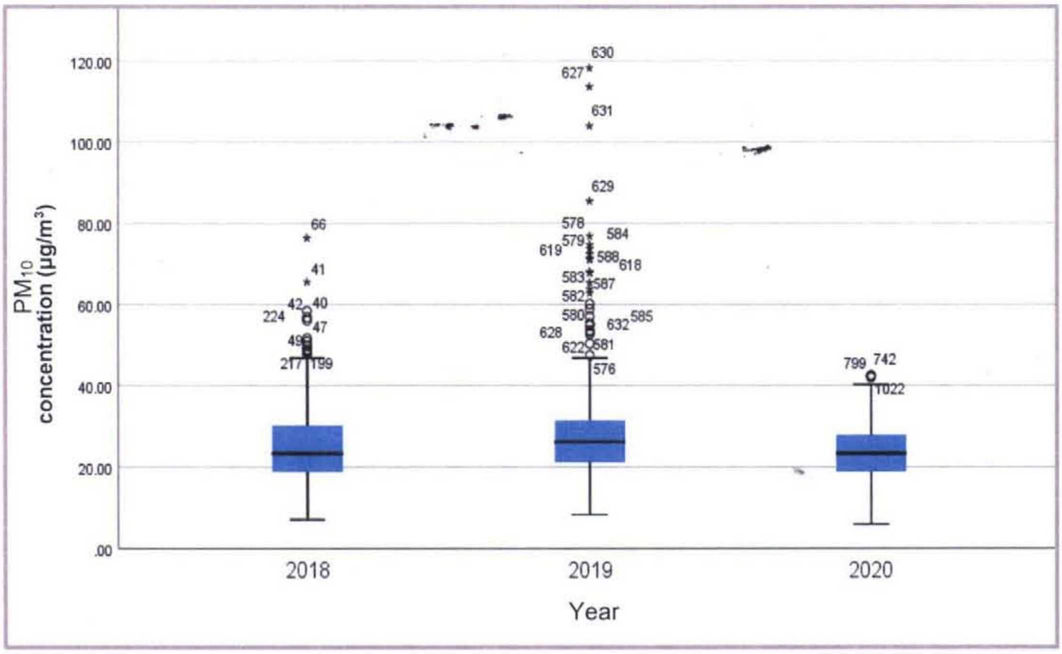


Figure 4.4: Box plot of PM_{10} at Kuala Terengganu Monitoring Station

4.2.2 Descriptive Analysis of $PM_{2.5}$

As shown in Table 4.2, all station the average concentration of $PM_{2.5}$ is occurred at Klang station was the highest mean (113.67 µg/m^3) and followed in decreasing order by Bukit Rambai 75.83 µg/m^3), Kuala Terengganu (62.07 µg/m$^{3)}$ and Batu Pahat (56.84 µg/m^3) respectively. In this study, the average concentration PM_{10} and $PM_{2.5}$ shown a similarity variance of average concentrations at Kuala Terengganu monitoring station.

Table 4.2: The descriptive analysis of $PM_{2.5}$ concentration for 2018-2020

	Klang	Bukit Rambai	Batu Pahat	Kuala Terengganu
Mean	28.14	22.10	16.45	17.19
Median	26.50	17.83	14.28	15.81
SD*	12.78	10.66	9.19	8.64
CV*	0.44	0.46	0.55	0.49
Skewness	2.51	1.54	1.26	1.37
Kurtosis	10.52	4.34	2.47	4.39
Maximum	**113.67**	**75.83**	**56.84**	**62.07**

*SD: standard deviation CV: coefficient of variation

While for the kurtosis $PM_{2.5}$ characteristics, the distribution for all location was skewed to the right. Based on Table 4.2, the kurtosis values remain highest skewness at Klang station among of other stations which was (10.52), followed by Kuala Terengganu (4.39), Bukit Rambai (4.34) and the lowest skewness was located at Batu Pahat (2.47). All the monitoring station are indicating normal distribution because the skewness value is above one (1).

According to the CV values on Table 4.2. the dispersion of $PM_{2.5}$ concentration at all monitoring stations it shown the greatest CV values was at Batu Pahat (0.55), followed by those in Kuala Terengganu (0.49), Bukit Rambai (0.46), and Klang (0.44). The result showed that Batu Pahat has the highest CV values among of other station.

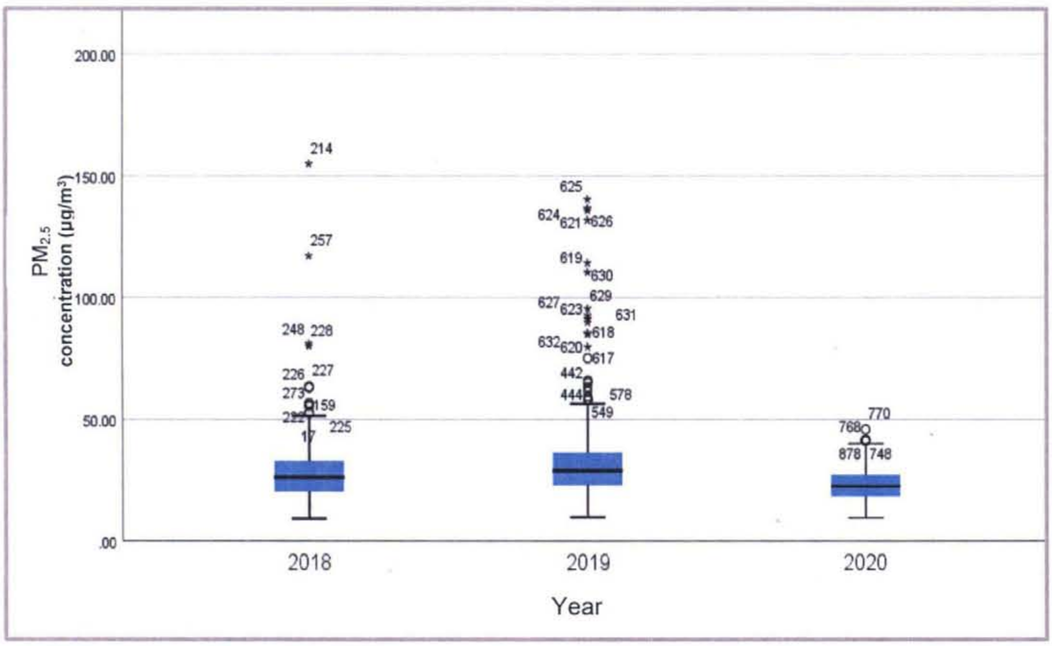


Figure 4.5 Box plot of $PM_{2.5}$:at Klang Monitoring Station

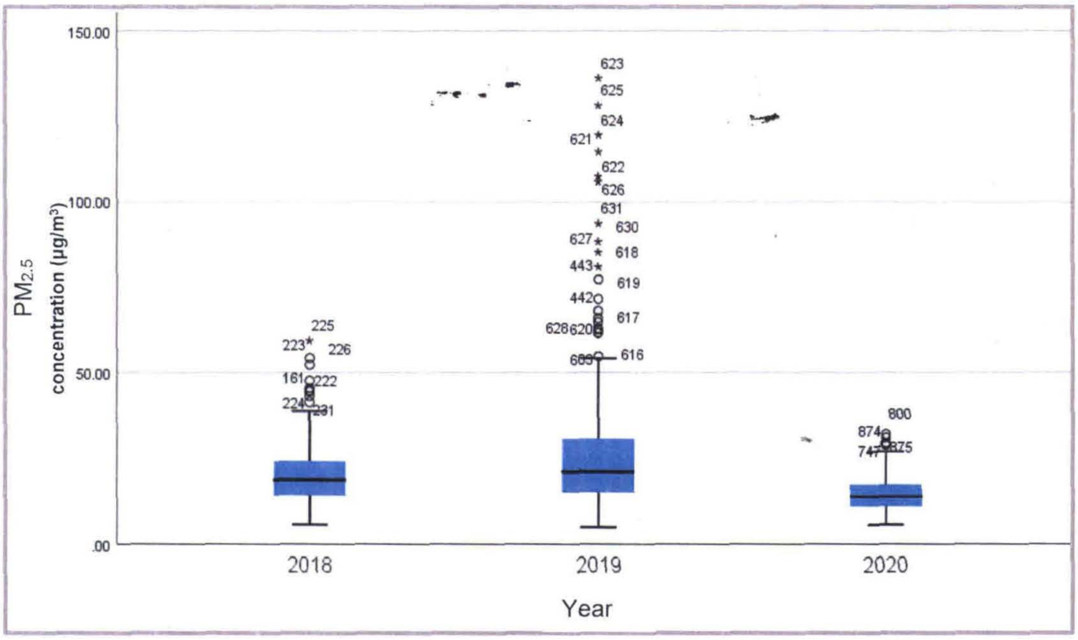


Figure 4.6: Box plot of $PM_{2.5}$ at Bukit Rambai Monitoring Station

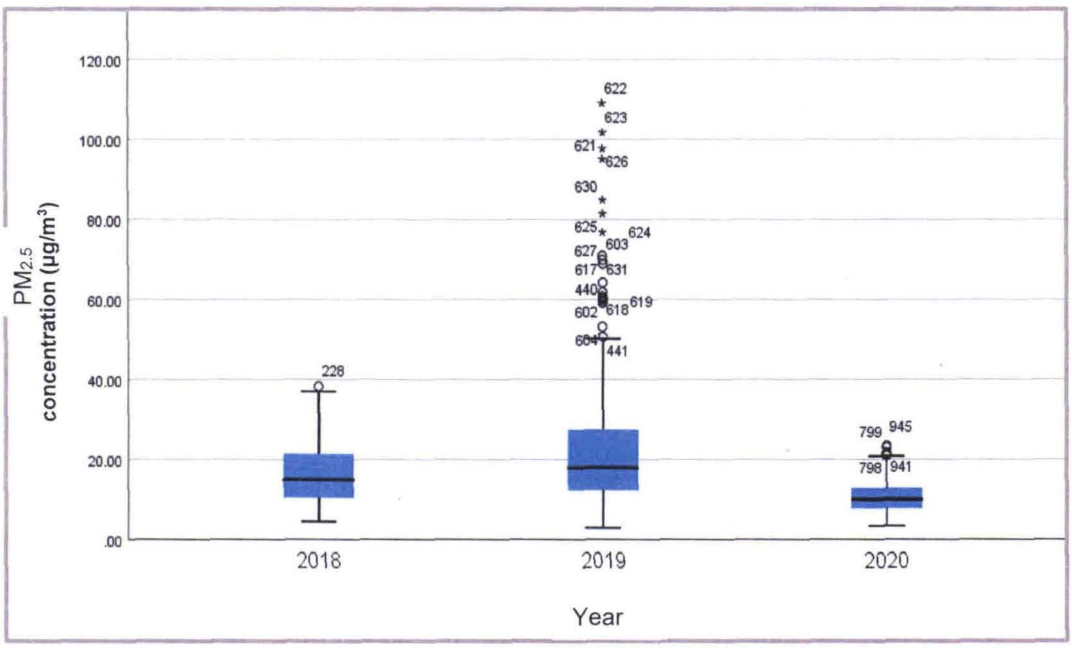


Figure 4.7: Box plot of $PM_{2.5}$ at Batu Pahat Monitoring Station

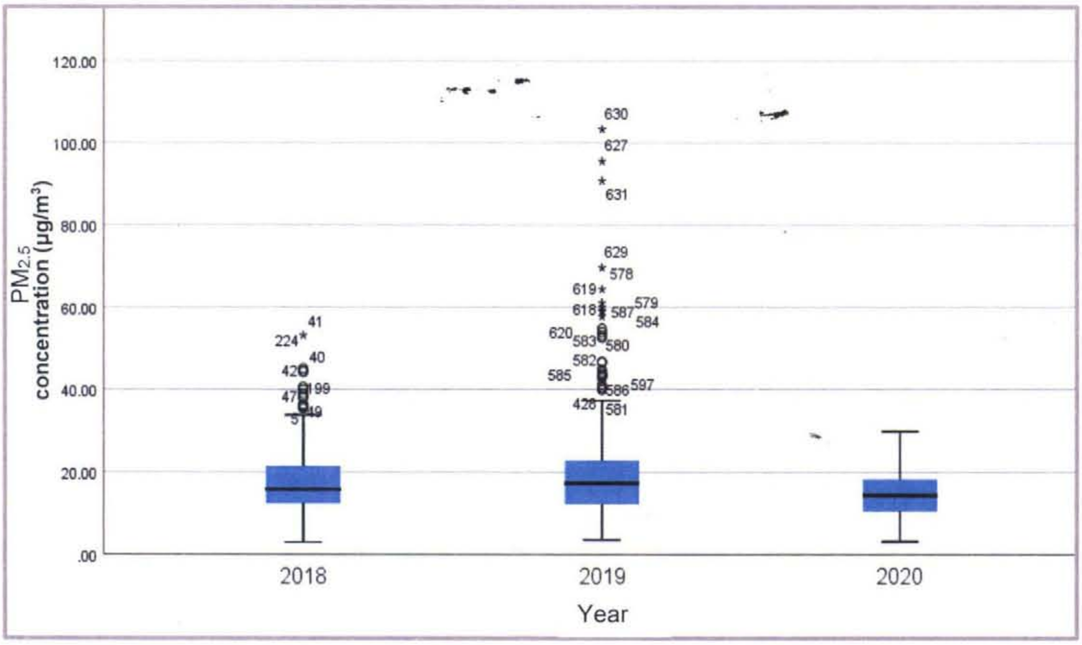


Figure 4.8: Box plot of $PM_{2.5}$ at Kuala Terengganu Monitoring Station

4.2.3 Descriptive Analysis of Ambient Air Quality

Based on the descriptive analysis trend in both table, the descriptive analysis summary of the occurrences between 2018 to 2020 at four monitoring sites located in Klang, Bukit Rambai, Batu Pahat, and Kuala Terengganu. The mean concentrations of PM_{10} that across all stations was below 40 μg/m^3 while for $PM_{2.5}$ concentrations was exceed 35 μg/m^3 on a yearly basis over the last three years. While, for the coefficient of variation (CV) for all station is small which it is indicates less dispersion of PM_{10} and $PM_{2.5}$ concentration data for all monitoring stations. Therefore, the lower the value of the coefficient of variation (CV), the more precise the prediction of PM_{10} and $PM_{2.5}$ concentration.

Overall, all the stations trend in descriptive analysis of annual average for PM_{10} and $PM_{2.5}$ concentration in 2018 to 2020 was within the limits of Malaysian Ambient Air Quality Standards (MAAQG) which is 50 μg/m^3 and 25μg/m^3 per year. However, at Klang monitoring station the concentration PM_{10} and $PM_{2.5}$ are highest than other stations. Klang is one of the most important economic areas in Malaysia. This was shown by the high number of unburned hydrocarbons that came from vehicles as well as the high amount of SO_2 coming from factories and power plants that use a lot of sulphur fuel.

Table 4.3: Descriptive Analysis for PM_{10} for all locations from 2018-2020

Malaysian Ambient Air Quality Guidelines (40 µg/m³ per year & 100 µg/m³ per day)			
Klang	**2018**	**2019**	**2020**
Mean	**41.30**	**41.42**	**33.78**
SD	17.33	22.58	9.22
CV	0.42	0.55	0.27
Maximum	**180.23**	**163.54**	**61.55**
Bukit Rambai	**2018**	**2019**	**2020**
Mean	**29.62**	**33.95**	**20.66**
SD	9.84	19.83	5.78508
CV	0.33	0.58	0.28
Maximum	**70.34**	**151.09**	**40.91**
Batu Pahat	2018	2019	2020
Mean	24.34	30.32	16.94
SD	8.55	17.44	4.90
CV	0.35	0.58	0.29
Maximum	**50.65**	**121.22**	**30.59**
Kuala Terengganu	**2018**	**2019**	**2020**
Mean	**25.33**	**28.86**	**23.31**
SD	9.76	14.07	6.85
CV	0.39	0.49	0.29
Maximum	**76.35**	**118.19**	**42.44**

*units = µg/m³

Table 4.4: Descriptive Analysis for $PM_{2.5}$ for all locations from 2018-2020

Malaysian Ambient Air Quality Guidelines (15 µg/m³ per year & 35 µg/m³ per day)			
Klang	**2018**	**2019**	**2020**
Mean	**28.14**	**33.04**	**23.24**
SD	13.05	18.521	6.78
CV	0.46	0.56	0.29
Maximum	**154.85**	**140.32**	**45.85**

Table 4.4: Descriptive Analysis for $PM_{2.5}$ for all locations from 2018-2020 (*continued*)

Bukit Rambai	**2018**	**2019**	**2020**
Mean	**20.14**	**25.95**	**14.39**
SD	8.58	18.48	4.92
CV	0.43	0.71	0.34
Maximum	**59.37**	**136.1**	**32.01**
Batu Pahat	**2018**	**2019**	**2020**
Mean	**16.53**	**22.15**	**10.66**
SD	7.65	15.80	4.13
CV	0.46	0.71	0.387
Maximum	**38.18**	**108.94**	**23.39**
Kuala Terengganu	**2018**	**2019**	**2020**
Mean	**17.39**	**19.74**	**14.43**
SD	7.73	12.84	5.36
CV	0.44	0.65	0.37
Maximum	**53.17**	**103.26**	**29.77**

*units = $\mu g/m^3$

4.2.4 Meteorological Parameters and Gaseous

Meteorological conditions play a big impact of the formation and elevated PM_{10} and $PM_{2.5}$ where it gives major influence on ambient air quality through the various in atmosphere weather directly and indirectly. The meteorological factors in this study were wind speed, relative humidity and ambient temperature, sulphur dioxide (SO_2), nitrogen dioxide (NO_2), ozone (O_3) and carbon monoxide (CO). A study prove that Multiple Linear Regression (MLR) models have been developed to predict the concentrations of PM_{10} and $PM_{2.5}$ based on a number of multivariate studies, including the construction of MLR models for the prediction on Day 1 and Day 2 of PM_{10} and $PM_{2.5}$ concentration (Abdullah et al., 2019). According to the study Ul-Saufie et al., (2012) results showed, the use of atmospheric parameters with gases as inputs was more effective than utilizing atmospheric characteristics without gases. Table 4.5 shows the descriptive analysis with meteorological and gaseous parameters.

Table 4.5: Summary of descriptive analysis for meteorological parameter and gaseous

Study Area	Parameter	Maximum	Minimum	Mean	SD
Klang	Relative humidity (%)	95.19	61.31	79.31	5.73
Bukit Rambai		95.41	65.38	81.66	5.03
Batu Pahat		96.81	72.91	85.00	4.28
Kuala Terengganu		97.96	72.78	83.10	4.38
Klang	Ambient Temp. (%)	31.00	24.46	28.44	1.13
Bukit Rambai		29.77	23.06	27.48	1.04
Batu Pahat		29.15	22.64	26.52	0.97
Kuala Terengganu		29.95	23.12	26.87	1.10
Klang	SO_2 (ppm)	0.0064	0.0004	0.0015	0.0009
Bukit Rambai		0.0036	0.0001	0.0014	0.0006
Batu Pahat		0.0029	0.0006	0.0016	0.0004
Kuala Terengganu		0.002	0.0002	0.0008	0.0003
Klang	NO_2 (ppm)	0.0333	0.0062	0.0178	0.0048
Bukit Rambai		0.0137	0.0013	0.0070	0.0024
Batu Pahat		0.0101	0.0018	0.0051	0.0017
Kuala Terengganu		0.0107	0.0012	0.0050	0.0015
Klang	O_3 (ppm)	0.037	0.003	0.016	0.006
Bukit Rambai		0.047	0.008	0.022	0.007
Batu Pahat		0.037	0.007	0.017	0.005
Kuala Terengganu		0.038	0.005	0.016	0.005
Klang	CO (ppm)	1.67	0.38	0.95	0.22
Bukit Rambai		1.32	0.19	0.65	0.21
Batu Pahat		1.17	0.19	0.59	0.16
Kuala Terengganu		0.99	0.26	0.57	0.18

*SD- Standard Deviation

Based on Table 4.5, it showed that the relative humidity is the main factor that influence in air pollutants which the maximum values of relative humidity was recorded at Kuala Terengganu (97.96), followed by Batu Pahat (96.81), Bukit Rambai (95.41) and Klang (95.19). This relative humidity factors are main contributed to the number of rainfalls. Therefore, it will increase the number of water vapour in atmosphere and relative humidity will decrease the PM_{10} and $PM_{2.5}$ concentrations in atmosphere (Abdullah, et al., (2017) and Azmi et al., (2009)) stated that high number of rainfalls will decrease PM_{10} and $PM_{2.5}$ concentrations by the wash-out process of the atmospheric aerosols in the atmosphere. The maximum daily temperature was recorded at Klang (31.00) followed by Bukit Rambai (29.77), Batu Pahat (29.15), and Kuala Terengganu (29.95) respectively.

4.3 Multiple Linear Regression (MLR)

Multiple Linear Regression (MLR) models have been developed for forecasting urban air quality. These models were used for the purpose of predicting the air quality in urban areas. Normally, this Multiple Linear Regression models applicable be used for predicting PM_{10} and $PM_{2.5}$ concentrations across urban areas. In air pollution forecasting, the multiple linear regression model has been frequently used for many years as an approach that relies on present and past data for predicting for future used (Abdullah et al., 2020). These models are established us for predicting the significant variability of air pollutant between dependant variables and independent variables (Abdullah et al., 2020). A study by Fong et al., (2018) show how multiple linear models regressions indicates information of the present day with meteorological forecasts of the next day can help forecasting daily PM_{10} concentrations for sites located Terengganu.

The independent variables in this study are particulate matter with aerodynamic diameter less than 10μm (PM_{10}) and less than 2.5μm ($PM_{2.5}$), sulphur dioxide (SO_2), nitrogen dioxide (SO_2), carbon monoxide (CO), wind speed, relative humidity and temperature. The standardized coefficient variable is applied in this model. Therefore, the higher the absolute value of standardized coefficient it can contribute a better contribution of the variable results.

4.3.1 Klang Monitoring Station

Klang is mainstream region for economic growth in Malaysia where is an urban commercial and industrial area. This Klang station was recorded due to high number of unhealth days in 2010, 2011 and 2012 (Elhadi et al., 2018) Based on the result in Table 4.6 and Table 4.7, predicting of PM_{10} for Day 1 concentrations is higher at PM_{10D0} (0.966) with R^2 value was (0.507). While for prediction Day 2 of PM_{10} concentrations the standardized coefficient (Beta) value for PM_{10D0} was (1.001) and it is slightly higher than previous Day 1 prediction and the R^2 value was (0.427) and the value slightly decrease from Day 1 prediction.

While, for predicting of $PM_{2.5}$ for Day 1 and Day 2 the value of beta and R^2 are decreasing compare with $PM_{2.5}$ concentrations predicting. For next day (Day 1) prediction the value was (0.575) with the R^2 value (0.451) and the lower R^2 value for the next two-day (Day 2) prediction which was R^2 value (0.372). Klang is totally affected by the main roads traffic vehicles of the industrial activities and there was a high population of residential area which affected the air quality status in Klang. This unhealthy air quality status will affect human health, animals and vegetation threatened due to the rise of economic industrial at Klang areas (Elhadi et al., 2018).

Table 4.6: The correlation coefficient between pollutants and meteorological parameters at Klang monitoring station for PM_{10} (Day 1 and Day 2)

Model	PM_{10D1}		PM_{10D2}	
	Beta	R^2	Beta	R^2
$PM_{10,D0}$	0.966	0.507	1.001	0.427
$PM_{2.5,D0}$	-0.347		-0.469	
SO_2	-830.453		-780.317	
NO_2	-203.024		-407.065	
O_3	-118.025		-173.862	
CO	-1.953		-0.358	
WS	0.264		2.411	
RH	-0.740		-0.733	
Temperature	-2.801		-3.199	

Table 4.7: The correlation coefficient between pollutants and meteorological parameters at Klang monitoring station for $PM_{2.5}$ (Day 1 and Day 2)

Model	$PM_{2.5D1}$		$PM_{2.5D2}$	
	Beta	R^2	Beta	R^2
$PM_{10,D0}$	0.071		0.148	
$PM_{2.5,D0}$	0.575		0.413	
SO_2	-584.391		-606.069	
NO_2	-187.129		-343.534	
O_3	-100.913	0.451	-156.351	0.372
CO	-0.488		0.763	
WS	0.211		2.264	
RH	-0.572		-0.565	
Temperature	-1.866		-2.062	

4.3.2 Bukit Rambai Monitoring Station

Based on Table 4.8 and Table 4.9, it showed the contribution coefficients of PM_{10} and $PM_{2.5}$ concentration at Bukit Rambai station are summarized. The contributor PM_{10} and $PM_{2.5}$ concentration value in this model are bold. Bukit Rambai is located under industrial areas. Therefore, a major of transportations of industry network around of Bukit Rambai monitoring station.

Ahmat et al., (2015) demonstrated that Bukit Rambai moderately several numbers of unhealthy days occurred at Bukit Rambai station. Even though, this Bukit Rambai station surrounded with industry areas it can remain the healthy air quality status was a good sign.

From the study result, the model indicates for predicting the next day (Day 1) and next two day (Day 2) of PM_{10} and $PM_{2.5}$ concentration across 2018 to 2020. The main contributors of this station are humidity, temperature, sulphur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO) and PM_{10} and $PM_{2.5}$ values that day which knows as (PM_{10D0} and $PM_{2.5D0}$) The highest R^2 value was PM_{10D1} (0.774) and $PM_{2.5D1}$ (0.795).

Table 4.8: The correlation coefficient between pollutants and meteorological parameters at Bukit Rambai monitoring station for PM_{10} (Day 1 and Day 2)

Model	PM_{10D1}		PM_{10D2}	
	Beta	R^2	Beta	R^2
$PM_{10,D0}$	**0.689**	0.774	0.487	0.622
$PM_{2.5,D0}$	0.201		**0.350**	
SO_2	311.388		**470.009**	
NO_2	**45.193**		28.157	
O_3	**-99.642**		-207.820	
CO	**0.296**		-1.236	
WS	-1.209		**-0.669**	
RH	-0.426		-0.456	
Temperature	-1.191		-1.201	

Table 4.9: The correlation coefficient between pollutants and meteorological parameters at Bukit Rambai monitoring station for $PM_{2.5}$ (Day 1 and Day 2)

Model	$PM_{2.5D1}$		$PM_{2.5D2}$	
	Beta	R^2	Beta	R^2
$PM_{10,D0}$	**0.013**	0.795	-0.015	0.646
$PM_{2.5,D0}$	**0.856**		0.821	
SO_2	202.126		**243.641**	
NO_2	**10.628**		-9.657	
O_3	-70.782		-153.442	
CO	**0.844**		-0.302	
WS	-1.764		-1.338	
RH	-0.429		-0.453	
Temperature	-1.001		-0.980	

4.3.3 Batu Pahat Monitoring Station

From Table 4.10 and Table 4.11, it is showed the value R^2 of PM_{10} and $PM_{2.5}$ for (Day 1) and (Day 2) are in the range (0.592 to 0.775). This is showed that the regression model fitting the good sign. The main contributors in this model's regression for predicting PM_{10} and $PM_{2.5}$ concentrations at Batu Pahat station are humidity, temperature, sulphur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO) and PM_{10} and $PM_{2.5}$ values that day which knows as (PM_{10D0} and $PM_{2.5D0}$.)

For both prediction value PM_{10} and $PM_{2.5}$, sulphur dioxide (SO_2) had indicated the highest values. High levels of concentrations (SO_2) can lead to worst human respiratory system. This (SO_2) gas is reacted with the small particle in ambient that can cause major health effect to human lifestyle (Radin et al., 2016).

Table 4.10 The correlation coefficient between pollutants and meteorological parameters at Batu Pahat monitoring station for PM_{10} (Day 1 and Day 2)

Model	PM_{10D1}		PM_{10D2}	
	Beta	R^2	Beta	R^2
$PM_{10,D0}$	0.622	0.757	0.548	0.592
$PM_{2.5,D0}$	0.198		0.152	
SO_2	-464.707		-1325.670	
NO_2	837.058		**1153.460**	
O_3	-115.809		-164.417	
CO	0.956		0.217	
WS	1.389		2.956	
RH	-0.262		-0.299	
Temperature	-0.756		-0.663	

Table 4.11: The correlation coefficient between pollutants and meteorological parameters at Batu Pahat monitoring station for $PM_{2.5}$ (Day 1 and Day 2)

Model	$PM_{2.5D1}$		$PM_{2.5D2}$	
	Beta	R^2	Beta	R^2
$PM_{10,D0}$	0.021		0.186	
$PM_{2.5,D0}$	0.769		0.467	
SO_2	-442.357		-1113.741	
NO_2	857.839		**1173.836**	
O_3	-99.313	**0.775**	-114.299	0.612
CO	0.296		-0.425	
WS	0.687		2.024	
RH	-0.315		-0.334	
Temperature	-0.605		-0.482	

4.3.4 Kuala Terengganu Monitoring Station

For predicting the next day (Day 1) and next two days (Day 2) of PM_{10} and $PM_{2.5}$ concentration across 2018 to 2020 at Kuala Terengganu was showed in Table 4.12 and Table 4.13. The main contributors of this station are humidity, temperature, sulphur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO) and PM_{10} and $PM_{2.5}$ values that day which knows as (PM_{10D0} and $PM_{2.5D0}$). The highest R^2 was for PM_{10D1} (0.609) and while for $PM_{2.5}$ the highest R^2 value $PM_{2.5D1}$ (0.638).

Result in Table 4.12 and Table 4.13 showed that there was highest of nitrogen dioxide (SO_2) concentration predicting values in range (436.081 to 453.462) for predicting the next day (Day 1) and next day (Day 2) of PM_{10} and $PM_{2.5}$ concentrations.

Table 4.12: The correlation coefficient between pollutants and meteorological parameters at Kuala Terengganu monitoring station for PM_{10} (Day 1 and Day 2)

Model	PM_{10D1}		PM_{10D2}	
	Beta	R^2	Beta	R^2
$PM_{10,D0}$	0.476	**0.609**	0.466	0.420
$PM_{2.5,D0}$	0.320		0.200	
SO_2	-55.555		-503.377	
NO_2	421.423		436.081	
O_3	168.862		183.336	
CO	-7.556		-9.133	
WS	1.552		0.729	
RH	-0.180		-0.047	
Temperature	-0.633		-0.494	

Table 4.13: The correlation coefficient between pollutants and meteorological parameters at Kuala Terengganu monitoring station for $PM_{2.5}$ (Day 1 and Day 2)

Model	$PM_{2.5D1}$		$PM_{2.5D2}$	
	Beta	R^2	Beta	R^2
$PM_{10,D0}$	0.027	**0.638**	0.128	0.439
$PM_{2.5,D0}$	0.713		0.478	
SO_2	-784.116		-1408.247	
NO_2	440.498		453.462	
O_3	104.806		129.275	
CO	-3.879		-5.493	
WS	0.607		0.031	
RH	-0.132		0.004	
Temperature	0.061		0.351	

4.4 Multiple Linear Model Regression for Urban Monitoring Station

Based on Table 4.14 shows the MLR models indicates for all station in Peninsular Malaysia. The multiple linear regression model (MLR) was applied to determine the variability of the proposed equation for predicting the values PM_{10} and $PM_{2.5}$ concentrations across the next day (Day 1) and the next two days (Day 2). The main contributor for indicates in (MLR) model was PM_{10} and $PM_{2.5}$ at that day which is known as PM_{10D0} and $PM_{2.5D0}$.

Table 4.14: Multiple Linear regression model of $PM_{10,D1}$, $PM_{10,D2}$ & $PM_{2.5,D1}$, $PM_{2.5,D2}$ at all monitoring station, 2018-2020

Station	Multiple Linear Regression (MLR) model
Klang	$PM_{10,D1}$: (158.149)+(0.966*PM_{10D0})-(0.347*$PM_{2.5D0}$)-(830.453*SO_2)-(203.024*NO_2)-(118.025*O_3)-(1.953*CO)+(0.264*WS)-(0.740*RH)-(2.801*TEMP)
	$PM_{10,D2}$: (171.144)+(1.001*PM_{10D0})-(0.469*$PM_{2.5D0}$)-(780.317*NO_2)-(407.065*SO_2)-(173.862*O_3)-(0.358*CO)+(2.411*WS)-(0.733*RH)-(3.199*TEMP)
	$PM_{2.5,D1}$: (114.016)+(0.071*PM_{10D0})+(0.575*$PM_{2.5D0}$)-(584.391*NO_2)-(187.129*SO_2)-(100.913*O_3)-(0.488*CO)+(0.211*WS)-(0.572*RH)-(1.866*TEMP)
	$PM_{2.5,D2}$: (120.538)+(0.148*PM_{10D0})+(0.413*$PM_{2.5D0}$)-(606.069*NO_2)-(343.534*SO_2)-(156.351*O_3)+(0.763*CO)+(2.264*WS)-(0.565*RH)-(2.062*TEMP)
Bukit Rambai	$PM_{10,D1}$: (76.032)+(0.689*PM_{10D0})+(0.201*$PM_{2.5DO}$)+(311.388*SO_2)+(45.193*NO_2)-(99.642*O_3)+(0.296*CO)-(1.209*WS)-(0.426*RH)-(1.191*TEMP)
	$PM_{10,D2}$: (83.806)+(0.487*PM_{10D0})+(0.350*$PM_{2.5D0}$)+(470.009*SO_2)+(28.157*NO_2)-(207.820*O_3)-(1.236*CO)-(0.669*WS)-(0.456*RH)-(1.201*TEMP)
	$PM_{2.5,D1}$: (69.200)+(0.013*PM10)+(0.856*PM2.5)+(202.126*SO2)+(10.628*NO2)-(70.782*O3)+(0.844*CO)-(1.764*WS)-(0.429*RH)-(1.001*TEMP)
	$PM_{2.5,D2}$: (74.023)-(0.015*PM_{10D0})+(0.821*$PM_{2.5}$)+(243.641*SO_2)-(9.657*NO_2)-(153.442*O_3)-(0.302*CO)-(1.338*WS)-(0.453*RH)-(0.980*TEMP)

(continued)

Table 4.14: Multiple Linear regression model of $PM_{10,D1}$, $PM_{10,D2}$ & $PM_{2.5,D1}$, $PM_{2.5,D2}$ at all monitoring station, 2018-2020 *(continued)*

Station	Multiple Linear Regression (MLR) model
Batu Pahat	$PM_{10,D1}$: (44.826)+(0.622*PM_{10D0})+(0.198*$PM_{2.5D0}$)-(464.707*SO_2)+(837.058*NO_2)-(115.809*O_3)+(0.956*CO)+(1.389*WS)-(0.262*RH)-(0.756*TEMP)
	$PM_{10,D2}$: (47.586)+(0.548*PM_{10D0})+(0.152*$PM_{2.5D0}$)-(1325.670*SO_2)+(1153.460*NO_2)+(164.417*O_3)+(0.217*CO)+(2.956*WS)-(0.299*RH)-(0.663*TEMP)
	$PM_{2.5,D1}$: (43.496)+(0.021*PM_{10D0})+(0.769*$PM_{2.5D0}$)-(442.357*SO_2)+(857.839*NO_2)-(99.313*O_3)+(0.296*CO)+(0.687*WS)-(0.315*RH)-(0.605*TEMP)
	$PM_{2.5,D2}$: (41.641)+(0.186*PM_{10D0})+(0.467*$PM_{2.5D0}$)-(1113.741*SO_2)+(1173.836*NO_2)-(114.299*O_3)-(0.425*CO)+(2.024*WS)-(0.334*RH)-(0.482*TEMP)
Kuala Terengganu	$PM_{10,D1}$: (37.641)+(0.476*PM_{10D0})+(0.320*$PM_{2.5D0}$)-(55.555*SO_2)+(421.423*NO_2)+(168.862*O_3)-(7.556*CO)+(1.552*WS)-(0.180*RH)-(0.633*TEMP)
	$PM_{10,D2}$: (27.165)+(0.466*PM_{10D0})+(0.200*$PM_{2.5D0}$)-(503.377*SO_2)+(436.081*NO_2)+(183.336*O_3)-(9.133*CO)+(0.729*WS)-(0.047*RH)-(0.494*TEMP)
	$PM_{2.5,D1}$: (11.818)+(0.027*PM_{10D0})+(0.713*$PM_{2.5D0}$)-(784.116*SO_2)+(440.498*NO_2)+(104.806*O_3)-(3.879*CO)+(0.607*WS)-(0.132*RH)+(0.061*TEMP)
	$PM_{2.5,D2}$: (-4.145)+(0.128*PM_{10D0})+(0.478*$PM_{2.5D0}$)-(1408.247*SO_2)+(453.462*NO_2)+(1229.275*O_3)-(5.493*CO)+(0.031*WS)+(0.004*RH)+(0.351*TEMP)

4.5 Performance Evaluation

In this study, performance indicators were used to determine the best model and distribution of PM_{10} and $PM_{2.5}$ concentrations for all stations. There are several performance indicators that are applied to a multiple linear regression model to obtain a good model, such as coefficient of determination (R^2), index of agreement (IA), prediction accuracy (PA), normalized absolute error (NAE), and root mean square

(RMSE). These performance indicators were used to evaluate the good fitted model for each station.

The error measures such as NAE show the best model if the evaluated values are closer to zero (0). While R^2, PA and IA values are known as accuracy measures for this model. These accuracy measures evaluated the values close to one (1), and it indicates a better model (Sansuddin et al., 2010). Performance indicators were used to determine the best model distribution to represent the PM_{10} and $PM_{2.5}$ concentrations.

Table 4.15: Performance indicator MLR for PM_{10} concentrations

Performance Indicator ($PM_{10,D1}$)				
Performance Indicator	Station			
	Klang	Bukit Rambai	Batu Pahat	Kuala Terengganu
R^2	**0.96**	0.78	0.76	0.60
IA	0.75	0.80	0.69	**0.98**
PA	**0.97**	0.55	0.44	0.52
NAE	0.21	0.21	0.23	**0.19**
RMSE	8.81	5.69	**4.90**	5.47
Performance Indicator ($PM_{10,D2}$)				
Performance Indicator	Station			
	Klang	Bukit Rambai	Batu Pahat	Kuala Terengganu
R^2	**0.93**	0.66	0.60	0.42
IA	0.70	0.73	0.43	**0.97**
PA	**0.88**	0.46	0.21	0.44
NAE	**0.22**	0.26	0.31	0.23
RMSE	7.94	6.54	**3.56**	**3.30**

Table 4.16: Performance indicator of MLR for$PM_{2.5}$ concentrations prediction

Performance Indicator ($PM_{2.5,D1}$)				
Performance Indicator	Station			
	Klang	Bukit Rambai	Batu Pahat	Kuala Terengganu
R^2	0.45	**0.80**	0.77	0.64
IA	**0.82**	0.79	0.72	0.56
PA	**0.68**	0.67	0.55	0.62
NAE	0.39	**0.20**	0.28	0.25
RMSE	**3.99**	4.83	4.21	7.57

Table 4.16: Performance indicator of MLR forPM2.5 concentrations prediction (*continued*)

Performance Indicator ($PM_{2.5,D2}$)				
Performance Indicator	Station			
	Klang	Bukit Rambai	Batu Pahat	Kuala Terengganu
R^2	0.37	**0.65**	0.61	0.44
IA	0.43	**0.75**	0.51	0.46
PA	**0.64**	0.51	0.50	0.34
NAE	**0.23**	0.30	0.36	0.27
RMSE	**3.72**	5.46	5.08	6.77

Based on Table 4.15 and Table 4.16 are summarized of result performance indicator. The result shows the model performance for predicting PM_{10} and $PM_{2.5}$ concentrations. In Table 4.15, showed that the model performance for prediction of PM_{10} concentrations next day (PM_{10D1}) indicates best model at Klang and Kuala Terengganu. Klang has showed the best model that the values of (R^2 = 0.96) and (PA = 0.97) which the best model that indicates values close to one (1). While for index of agreement, Kuala Terengganu indicate the best model which is the values closer to one (1) (IA = 0.98). The values of NAE and RMSE are indicate best model at Batu Pahat (NAE = 0.19) and Kuala Terengganu (RMSE = 4.90). Based on the Table 4.15, Klang and Kuala Terengganu indicate the best model prediction for next day (PM_{10D1}) and next two day (PM_{10D2}).

For Klang monitoring station, the results show in Table 4.16 MLR model for $PM_{2.5D1}$ predictions indicated the good performance with (IA = 0.82), (PA = 0.68) and (RMSE = 3.99). While for Bukit Rambai monitoring station, the model indicates good performance with (R^2 = 0.80) and (NAE = 0.20). The best model for $PM_{2.5D1}$ prediction is located at Klang and Bukit Rambai monitoring station. The next two day ($PM_{2.5D2}$), showed at Klang indicates good performance with (PA = 0.64), (NAE = 0.23), (RMSE = 3.72) that can be identify as good model prediction. At Bukit Rambai result showed the good performance with (R = 0.65) and (IA = 0.75).

By comparing with all monitoring station, the best model predictions to predict was at Klang Monitoring station. Thus, the predictions of next day (Day 1) and next two day (Day 2) of PM_{10} and $PM_{2.5}$ concentrations was at Klang monitoring station.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

A short-term prediction of PM_{10} and $PM_{2.5}$ concentration in Peninsular Malaysia monitoring stations of air monitoring data from 2018 to 2020 are successfully completed. The selected monitoring stations for this study are located at Klang, Bukit Rambai, Batu Pahat and Kuala Terengganu. Descriptive analysis of the mean, median, standard deviation, skewness and kurtosis of PM_{10} and $PM_{2.5}$ was carried out in comparison with meteorological and gaseous parameter in this study.

The main contributor parameters that applied in this model are relative humidity, temperature, windspeed, nitrogen dioxide (NO_2), sulphur dioxide (SO_2), ozone (O_3) and carbon monoxide (CO). The Multiple Linear Regression (MLR) was applied to indicates for predicting the next day (Day 1) and next two days (Day 2) of PM_{10} and $PM_{2.5}$ concentration across 2018 to 2020. The performance indicator is used to identify the best models for predicting PM_{10} and $PM_{2.5}$ concentration among four selected urban monitoring stations.

The prediction model multiple linear regression (MLR) models are applied to predict predicting the next day (Day 1) and next day (Day 2) of PM_{10} and $PM_{2.5}$ concentration. The prediction models were based on the performance indicator that identify a good model. The performance indicator that contributor in this prediction were (R^2, IA, PA, NAE, and RMSE).

The MLR models indicate most at monitoring stations, the PM_{10} concentrations and $PM_{2.5}$ at that day ($PM_{10, D0}$) and ($PM_{2.5, D0}$) was the major contribution in this MLR model prediction. The best MLR models are indicated good performance at Klang monitoring station for PM_{10D1} with $R^2 = 0.96$. While for $PM_{2.5D1}$, Klang monitoring stations also indicate good performance model prediction with $R^2 = 0.93$. Both of PM_{10} and $PM_{2.5}$ concentration showed Klang monitoring station was the good performances that indicates best model. Followed PM_{10D1} concentration by Bukit Rambai ($R^2 = 0.78$), Batu Pahat ($R^2 = 0.76$), and Kuala Terengganu ($R^2 = 0.60$). Comparing the performance indicator of all stations for PM_{10} and $PM_{2.5}$, showed that Klang is the best model from this study.

For $PM_{2.5}$ concentration prediction model was indicating good performance of Klang and Bukit Rambai monitoring station of $PM_{2.5D1}$. Bukit Rambai showed the good performances indicates of $R^2 = 0.80$ followed by Batu Pahat ($R^2 = 0.77$), Kuala Terengganu ($R^2 = 0.64$) and Klang ($R^2 = 0.45$). While for prediction next two day ($PM_{2.5D2}$) indicates Klang and Bukit Rambai has the good performances as identifying the best model predictions.

Overall, based on the descriptive analysis of PM_{10} summary results Klang monitoring stations were indicates the highest value (135.10 $\mu g/m^3$) while for the $PM_{2.5}$ descriptive analysis summary also located at Klang (113.67 $\mu g/m^3$). The lowest for PM_{10} concentrations descriptive analysis was located at Kuala Terengganu (62.07 $\mu g/m^3$) while for $PM_{2.5}$ concentrations at Batu Pahat (56.84 $\mu g/m^3$). All station of distribution PM_{10} and $PM_{2.5}$ concentrations are skewed to the right which is a good sign for air quality status.

In conclusion, the result performance indicator shown that Klang is the best model to predict the next day (Day 1) and next two day (Day 2) of PM_{10} and $PM_{2.5}$ concentrations that can be determined from this short-term prediction across three years of air monitoring data.

5.2 Recommendation

The goal of this study., the research result might be useful for future studies on air quality specially for models that predict concentrations of PM_{10} and $PM_{2.5}$ in Peninsular Malaysia.

This study helps to open the new findings studies to existing knowledge about predictions models of PM_{10} and $PM_{2.5}$ concentrations. This study also contributes to comparing the air particles concentrations such as PM_{10} and $PM_{2.5}$ with meteorology and gaseous parameters. The meteorology parameters will help to improve the accuracy research are monsoon, rainfall in a year and transboundary haze events in Malaysia.

In additional this study findings would improve the contribution to current approaches that dealing with air quality status prediction and management. Lastly, this study will help to provide a new source of studies research on short-term predicting of PM_{10} and $PM_{2.5}$ concentrations in Peninsular Malaysia.

Based on my studies, the using data period for my study was in short period was using in three years only (2018 – 2020). Therefore, to improve the accurancy of the predictions models by using bigger data such in long term duration. As example using minimum in five years duration data., the more longer period data be used, the more accurate predictions model performance can be resulted. Moreover, futured study should used more the model prediction and compared with MLR model to obtain the best models prediction in Peninsular Malaysia. Regarding on my study, only one models prediction was used. Therefore, by using more than one model predictions tool it can be improved the accuracy of predictions to indicates best model predicions. In a future study, different parameters such as wind direction, atmospheric pressure, and other air pollution data should be used to improve the MLR model's ability to predict PM_{10} and $PM_{2.5}$ concentrations.

REFERENCES

Abdul Rahman, S., Ismail, S.N.S., Sahani, M., Firoz, Rm, Latif, M., 2017. A case crossover analysis of primary air pollutants association on acute respiratory infection (ARI) among children in urban region of Klang valley, Malaysia. Ann. Trop. Med. Public Health 10 (Issue 1), 44. https://doi.org/10.4103/ATMPH.ATMPH_75_17. Medknow Publications

Abdullah M, A., Armi Abu Samah, M., & Yee Jun, T. (2012). An Overview of the Air Pollution Trend in Klang Valley, Malaysia. *Open Environmental Sciences*, *6*(1), 13–19. https://doi.org/10.2174/1876325101206010013

Abdullah, S., & Fong, S. (2017). Multiple Linear Regression (MLR) Model for Long Term PM10 concentration Forecasting during different monsoon seasons. *Journal of Sustainability Science and Management*, *12*, 60–69. Retrieved from https://jssm.umt.edu.my/wp-content/uploads/sites/51/2020/05/7-12.1.pdf

Abdullah, S., Ismail, M., Ahmed, A. N., & Abdullah, A. M. (2019). Forecasting Particulate Matter Concentration Using Linear and Non-Linear Approaches for Air Quality Decision Support. *Atmosphere*, *10*(11), 667. https://doi.org/10.3390/atmos10110667

Abdullah, S., Napi, N. N. L. M., Ahmed, A. N., Mansor, W. N. W., Mansor, A. A., Ismail, M., Ramly, Z. T. A. (2020). Development of Multiple Linear Regression for Particulate Matter (PM_{10}) Forecasting during Episodic Transboundary Haze Event in Malaysia. *Atmosphere*, *11*(3), 289. https://doi.org/10.3390/atmos11030289

Abdullah, S.; Ismail, M.; Fong, S.Y. (2017). Multiple Linear Regression (MLR) Models for Long Term PM_{10} Concentration Forecasting during Different Monsoon Seasons. J. Sustain. Sci. Manag. 2017, 12, 60–69. Published.

Abdullah, S.; Ismail, M.; Fong, S.Y.; Ahmed, A.N. Evaluation for Long Term PM_{10} Concentration Forecasting Using Multi Linear Regression (MLR) and Principal Component Regression (PCR) Models. *Environment Asia* 2016, 9, 101–110.

Abdul, S. A., Bakheit, C. S., & Al-Alawi, S. M. (2005). Principal Component and Multiple Regression Analysis in Modelling of Ground-level Ozone and Factors Affecting Its Concentrations. Environ. Model. Model. Software., 20: 1263-1271. Published. https://doi.org/10.1016/j.envsoft.2004.09.001

Ahmad A.M, Mohd Armi Abu Samah and Tham Yee Jun (2012). An Overview of the Air Pollution Trend in Klang Valley, Malaysia. Department of Environmental Science, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400

Serdang, Selangor, Malaysia. Published. https://doi.org/10.2174/1876325101206010013

Ahmat, H., Yahaya, A. S., & Ramli, N. A. (2015). PM_{10} Analysis for Three Industrialized Areas using Extreme Value. *Sains Malaysian*, *44*(2), 175–186. https://doi.org/10.17576/jsm-2015-4402-03

Arampongsanuwat, S., & Meesad, P. (2011). Prediction of PM $_{10}$ using Support Vector Regression. International Conference on Information and Electronics Engineering IPCSIT vol.6 (2011) © (2011) IACSIT Press, Singapore. Retrieved from International Conference on Information and Electronics Engineering IPCSIT vol.6 (2011) © (2011) IACSIT Press, Singapore website: http://www.ipcsit.com/vol6/24-E046.pdf

Azhari, A., Halim, N. D. A., Mohtar, A. A. A., Aiyub, K., Latif, M. T., & Ketzel, M. (2021). Evaluation and Prediction of PM_{10} and $PM_{2.5}$ from Road Source Emissions in Kuala Lumpur City Centre. *Sustainability*, *13*(10), 5402. https://doi.org/10.3390/su13105402

Azhari, A.; Halim, N.D.A.; Mohtar, A.A.A.; Aiyub, K.; Latif, M.T.; Ketzel, M. (2021). Evaluation and Prediction of PM_{10} and $PM_{2.5}$ from Road Source Emissions *in Kuala Lumpur City Centre.* Published. https://doi.org/10.3390/su13105402

Azmi, S. Z., Latif, M. T., Ismail, A. S., Juneng, L., & Jemain, A. A. (2009). Trend and status of air quality at three different monitoring stations in the Klang Valley, Malaysia. *Air Quality, Atmosphere & Health*, *3*(1), 53–64. https://doi.org/10.1007/s11869-009-0051-1

Baklanov, A., Hänninen, O., Slørdal, L. H., Kukkonen, J., Bjergene, N., Fay, B., … Ødegaard, V. (2007). Integrated systems for forecasting urban meteorology, air pollution and population exposure. *Atmospheric Chemistry and Physics*, *7*(3), 855–874. https://doi.org/10.5194/acp-7-855-2007

Brauer, M., Amann, M., Burnett, R. T., Cohen, A., Dentener, F., Ezzati, M., Henderson, S. B., Krzyzanowski, M., Martin, R. V., van Dingenen, R., van Donkelaar, A., & Thurston, G. D. (2012). Exposure Assessment for Estimation of the Global Burden of Disease Attributable to Outdoor Air Pollution. *Environmental Science & Technology*, *46*(2), 652–660. https://doi.org/10.1021/es2025752

Brunekreef, B. (2007). Health Effects of Air Pollution Observed in Cohort Studies in Europe. Published. https://doi.org/10.1038/sj.jes.7500628

Burns, J., Boogaard, H., Polus, S., Pfadenhauer, L. M., Rohwer, A. C., van Erp, A. M., Rehfuess, E. A. (2020). Interventions to reduce ambient air pollution and their effects on health: An abridged Cochrane systematic review. *Environment International*, *135*, 105400. https://doi.org/10.1016/j.envint.2019.105400

Department of Environment (DOE) Malaysia, 2016. Air Quality Report (online), (Accesses 5th June 2022), Available from World Wide Web: https://enviro2.doe.gov.my

Department of Environment (DOE) Malaysia, 2018. Air Quality Guideline (online), (Accesses 6th June 2022), Available from World Wide Web: https://enviro2.doe.gov.my

Department of Environment (DOE) Malaysia, 2020. Air Quality Report (online), (Accesses 5th June 2022), Available from World Wide Web: https://enviro2.doe.gov.my

Department of Statistics Malaysia 2017, Social Statistics Bulletin, ISSN (online), (Accessed 15th June 2022), Available from World Wide Web: https://www.dosm.gov.my

Department of Statistics Malaysia 2019, Compendium of Environment, ISSN (online), (Accessed 15th June 2022), Available from World Wide Web: https://www.dosm.gov.my

Department of Statistics Malaysia 2019, Social Statistics Bulletin, ISSN (online), (Accessed 15th June 2022), Available from World Wide Web: https://www.dosm.gov.my

Dominick, Doreena & Latif, Mohd Talib & Juahir, Hafizan & Aris, Ahmad Zaharin & Zain, Sharifuddin. (2012). An assessment of influence of meteorological factors on PM sub (10) and NO sub (2) at selected stations in Malaysia. Sustainable Environment Research. 22. 305-315.

El-Fadel, M., & Massoud, M. (2000). Particulate matter in urban areas: health-based economic assessment. *Science of the Total Environment*, *257*(2-3), 133–146. https://doi.org/10.1016/s0048-9697(00)00503-9

Elhadi, R. E., Abdullah, A. M., Abdullah, A. H., Ash'aari, Z. H., & Khan, M. F. (2018). Seasonal Variations of Atmospheric Particulate Matter and its Content of Heavy Metals in Klang Valley, Malaysia. *Aerosol and Air Quality Research*, *18*(5), 1148–1161. https://doi.org/10.4209/aaqr.2017.03.0113

Fernando, H. J. S., Mammarella, M. C., Grandoni, G., Fedele, P., Di Marco, R., Dimitrova, R., & Hyde, P. (2012). Forecasting PM10 in metropolitan areas: Efficacy of neural networks. *Environmental Pollution*, *163*, 62–67. https://doi.org/10.1016/j.envpol.2011.12.018

Fong, S., Abdullah, S., & Ismail, M. (2018). Forecasting of Particulate Matter (PM_{10}) Concentration Based on Gaseous Pollutant and Metrological Factors for Different Monsoons of Urban Coastal Area in Terengganu. *Journal of Sustainability Science and Management Special Issue Number*, *5*. Retrieved from https://jssm.umt.edu.my/wp-content/uploads/sites/51/2020/05/Bab-1.SI5_.pdf

Gupta, A., Mishra, P., Pandey, C., Singh, U., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia*, *22*(1), 67. https://doi.org/10.4103/aca.aca_157_18

Hao, J., Zhu, T., & Fan, X. (2014). Indoor Air Pollution and Its Control in China. *The Handbook of Environmental Chemistry*, 145–170. https://doi.org/10.1007/698_2014_257

Hashim, N.I.M.; Noor, N.M. Variations of Particulate Matter (PM_{10}) Concentration during Haze Episodes in Malaysia. In Proceedings of the 2017 Bangkok International Intellectual Property, Invention, Innovation and Technology Exposition, Bangkok, Thailand, 4 January 2017.

Katholi, R.E., & Couri, D.M. (2011). Left Ventricular Hypertrophy: Major Risk Factor in Patients with Hypertension: Update and Practical Clinical Applications. Int J Hypertense. (2011) *2011:495349.* Published. https://doi.org/10.4061/2011/495349103

Kelishadi, R., & Poursafa, P. A., and Bezirtzoglou E (2020) Environmental and Health Impacts of Air Pollution: *A Review. Front. Public Health 8:14*. Published. https://doi.org/10.3389/fpubh.2020.00014

Khairuddin,A., Norhayati Muhammad, Hasliza Y., Samsiah J., Azlan., Balkis A., Laily B. Din. (2017). Identification of most tolerant lichen species to vehicular traffic's pollutants at Batu Pahat area. Published. https://doi.org/10.1063/1.5005411

Kim K. H., Jahan S. A., Kabir. E. (2013). A Review on Human Health Perspective of Air Pollution with Respect to Allergies and Asthma. Published. https://doi.org/10.1016/j.envint.2013.05.007

Kinney P.L. (2008). Climate Change, Air Quality, and Human Health. Published. https://doi.org/10.1016/j.amepre.2008.08.025

Kukkonen, J., Partanen, L., Karppinen, A., Walden, J., Kartastenpää, R., Aarnio, P., … Berkowicz, R. (2003). Evaluation of the OSPM model combined with an urban background model against the data measured in 1997 in Runeberg Street, Helsinki. *Atmospheric Environment*, *37*(8), 1101–1112. https://doi.org/10.1016/s1352-2310(02)00957-3

Leonardo Trasande MD, MPP, George D. Thurston ScD (2005). The Role of Air Pollution in Asthma and Other Pediatric Morbidities. Published. https://doi.org/10.1016/j.jaci.2005.01.056

Lola, M. S., Ramlee, M. N. A., Gunalan, G. S., Zainuddin, N. H., Zakariya, R., Idris, M., & Khalil, I. (2016). Improved the Prediction of Multiple Linear Regression Model Performance Using the Hybrid Approach: A Case Study of Chlorophyll-a at the Offshore Kuala Terengganu, Terengganu. *Open Journal of Statistics*, *06*(05), 789–804. https://doi.org/10.4236/ojs.2016.65065

Maher Elbayoumi, Nor Azam Ramli, Noor Faizah Fitri Md Yusof. (2015). Development and Comparison of Regression Models and Feedforward Backpropagation Neural Network Models to Predict Seasonal Indoor $PM_{2.5-10}$ and $PM_{2.5}$ *Concentrations in Naturally Ventilated Schools.* Published. https://doi.org/10.1016/j.apr.2015.09.001

Manisalidis, Ioannis, Stavropoulou E., Stavropoulos A and Bezirtzoglou E. (2020). "Environmental and Health Impacts of Air Pollution: A Review." Frontiers in Public Health, vol. 8, no. 14, 20 Feb. 2020, 10.3389/fpubh.2020.00014. https://doi.org/10.3389/fpubh.2020.00014

Mishra D, Goyal P & Upadhyay A (2015). Artificial Intelligence Based Approach to Forecast PM2.5 during Haze Episodes: A Case Study of Delhi, India. *Atmospheric. Environment.* 2015, 102, 239–248. Published. https://doi.org/10.1016/j.atmosenv.2014.11.050

Mishra, Prabhaker & Pandey, ChandraM & Singh, Uttam & Gupta, Anshul & Sahu, Chinmoy & Keshri, Amit. (2019). Descriptive Statistics and Normality Tests for Statistical Data. Annals of Cardiac Anaesthesia. 22. 67-72. 10.4103/aca.ACA_157_18.

Mohammadyan, Mahmoud & Ghoochani, Mahboobeh & Kloog, Itai & Abdul-wahab, Sabah & Heibati, Behzad & Godri Pollitt, Krystal. (2017). Assessment of indoor and outdoor particulate air pollution at an urban background site in Iran. Environmental Monitoring and Assessment. 189. 10.1007/s10661-017-5951-1.

Mohd Zahid, A. Z., Abdul Malik, N. N. A., & Kassim, J. (2018). Particulate matter study at residential and educational areas in Shah Alam, Malaysia. *MATEC Web of Conferences*, *250*, 06010. https://doi.org/10.1051/matecconf/201825006010

Munsif, R., Zubair, M., Aziz, A., & Nadeem Zafar, M. (2021). Industrial Air Emission Pollution: Potential Sources and Sustainable Mitigation. *Environmental Emissions*. https://doi.org/10.5772/intechopen.93104

Nur Shaziayani, W., Zia Ul-Saufie, A., Libasin, Z., Norsyiha Ahmad Shukri, F., Sarimah Syed Abdullah, S., & Mohamed Noor, N. (2020). A Review of PM10 Concentrations Modelling in Malaysia. *IOP Conference Series: Earth and Environmental Science*, *616*(1), 012008. https://doi.org/10.1088/1755-1315/616/1/012008

Patrick K. H. Lee, Jeffrey R. Brook, Ewa Dabek-Zlotorzynska, and Scott A. Mabury. (2003). Identification of the Major Sources Contributing to $PM_{2.5}$ *Observed in Toronto. Environmental Science & Technology* **2003** *37* (21), 4831-4840. https://doi.org/10.1021/es026473i

Pires, J., Martins, F., Sousa, S., Alvim-Ferraz, M. and Pereira, M. (2008). Prediction of the Daily Mean PM10 Concentrations Using Linear Models. Am. J. Environ. Sci. 4: 445–453.

Pollution Control Department, Thailand (PCD), 2007. Thailand State of Pollution Report (online), (Accesses 29th May 2022), Available from World Wide Web: http://infofile.pcd.go.th/mgt/Report_Eng2550.pdf

Radin Mohamed, R. M. S., Hakim Rahim, A. F., & Mohd Kassim, A. H. (2016). A Monitoring of Air Pollutants (CO, $SO_2$and NO) in Ambient Air Near an Industrial

Area. *MATEC Web of Conferences*, *47*, 05022. https://doi.org/10.1051/matecconf/20164705022

Sansuddin, N., Ramli, N. A., Yahaya, A. S., Yusof, N. F. F. M., Ghazali, N. A., & Madhoun, W. A. A. (2010). Statistical analysis of PM10 concentrations at different locations in Malaysia. *Environmental Monitoring and Assessment*, *180*(1-4), 573–588. https://doi.org/10.1007/s10661-010-1806-8

Santiasih, I., & Hermana, J. (2017). A review: The Physicochemical characteristics of Indoor Particulate Matters in Relation to Human Health. Volume 12, No 6, March 2017. ARPN Journal of Engineering and Applied Sciences. (Santiasih and Hermana 2017)

Sayegh, A. S., Munir, S., & Habeebullah, T. M. (2014). Comparing the Performance of Statistical Models for Predicting PM10 Concentrations. *Aerosol and Air Quality Research*, *14*(3), 653–665. https://doi.org/10.4209/aaqr.2013.07.0259 (Sayegh, Munir, & Habeebullah, 2014)

Shah ASV, Langrish JP, Nair H, McAllister DA, Hunter AL, Donaldson K, et al. (2013). *Global* Association of Air Pollution and Heart Failure: A Systematic Review and Meta-Analysis. Published. https://doi.org/10.1016/S0140-6736(13)60898-3

Sahrir, Syazwani & Abdullah, Ahmad & Ponrahono, Zakiah & Sharaai, Amir Hamzah. (2019). Environmetric Study on Air Quality Pattern for Assessment in Klang Valley, Malaysia. 8. 17-24.

Shahraiyni, T H., & Sodoudi, S. (2016). Statistical Modeling Approaches for PM_{10} Prediction in Urban Areas; A Review of 21st Century Studies. Atmosphere, 7(2), 1-24. doi: https://doi.org/10.3390/atmos7020015

Tecer, L. H., Süren, P., Alagha, O., Karaca, F., & Tuncel, G. (2008). Effect of Meteorological Parameters on Fine and Coarse Particulate Matter Mass Concentration in a Coal-Mining Area in Zonguldak, Turkey. *Journal of the Air & Waste Management Association*, *58*(4), 543–552. https://doi.org/10.3155/1047-3289.58.4.543

Tecer, Alagha, Karaca, Tuncel, & Eldes, (2008). Particulate Matter ($PM_{2.5}$, $PM1_{0-2.5}$, and PM_{10}) and Children's Hospital Admissions for Asthma and Respiratory Diseases*: A Bidirectional Case-Crossover Study, Journal of Toxicology and Environmental Health,* https://doi.org/10.1080/15287390801907459

Tian, Guangjin & Qiao, Zhi & xu, Xinliang. (2014). Characteristics of particulate matter (PM_{10}) and its relationship with meteorological factors during 2001-2012 in Beijing. Environmental pollution (Barking, Essex : 1987). 192. 10.1016/j.envpol.2014.04.036.

Ul-Saufie, A.Z., A.S. Yahaya, N.A.Ramlia and H.A (2012). Robust regression models for predicting PM_{10} concentration in an industrial area. *Hamid International Journal of Engineering and Technology*, 2 (3) (2012).

Ul-Saufie, A.Z., Yahaya, A.S., Ramli, N.A. (2013), Future Daily PM_{10} Concentrations Prediction by Combining Regression Models and Feedforward Backpropagation Models with Principle Component Analysis (PCA), *Atmospheric Environment, 77, pp.621.-630. https://doi.org/10.1016/j.atmosenv.2013.05.017*

United State Environment Protection Agency (USEPA). 1994.Information about Air Quality Index (AQI), (Online), (Acessed 10th December 2021), Available from World Wide Web: http://www.air.dnr.state.ga.us./information/aqi.html

Wen, Yan, Nor, Ahmad, Nabila, Nurul & Sulaiman, Zulaikha. (2016). Transboundary Air Pollution in Malaysia: Impact and Perspective on Haze. *Nova Journal of Engineering and Applied Sciences. 5. 10.20286/Nova-Jeas-050103.* Published. https://doi.org/10.20286/nova-jeas-050101

World Health Organisation. 2018. Monitoring health for the SDGs, sustainable development goals. (online), (Access 13th January 2022), Available from World Wide Web: https://www.who.int

Yuen, F.S., Abdullah, S.; Ismail, M. (2018). Forecasting of Particulate Matter (PM_{10}) Concentration based on Gaseous Pollutants and Meteorological Factors for Different Monsoons of Urban Coastal Area in Terengganu. J. Sustain. Sci. Management. 2018, 13, 3–17. Published. https://doi.org/10.3390/atmos10110667

Zounemat Kermani, M. Hourly predictive Levenberg–Marquardt ANN and multi linear regression models for predicting of dew point temperature. *Meteorology Atmosphere Physical* **117,** 181–192 (2012). https://doi.org/10.1007/s00703-012-0192-x

APPENDICES

APPENDIX A: Multiple Linear Regression coefficients for Klang monitoring station

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	158.149	28.688		5.513	.000
	PM10	.966	.090	.966	10.758	.000
	PM2.5	-.347	.112	-.279	-3.088	.002
	NO2	-830.453	470.017	-.045	-1.767	.078
	SO2	-203.024	124.404	-.058	-1.632	.103
	O3	-118.025	84.644	-.039	-1.394	.164
	CO	-1.953	2.626	-.026	-.744	.457
	WS	.264	1.674	.005	.158	.875
	RH	-.740	.139	-.227	-5.328	.000
	TEMP	-2.801	.670	-.167	-4.179	.000

a. Dependent Variable: PM10D1

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	171.144	30.959		5.528	.000
	PM10	1.001	.097	1.001	10.323	.000
	PM2.5	-.469	.121	-.378	-3.873	.000
	NO2	-780.317	507.000	-.042	-1.539	.124
	SO2	-407.065	134.444	-.117	-3.028	.003
	O3	-173.862	91.294	-.058	-1.904	.057
	CO	-.358	2.833	-.005	-.126	.900
	WS	2.411	1.807	.043	1.334	.182
	RH	-.733	.150	-.224	-4.885	.000
	TEMP	-3.199	.723	-.191	-4.422	.000

a. Dependent Variable: PM10D2

Coefficients[a]						
		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	114.016	24.378		4.677	.000
	PM10	.071	.076	.088	.930	.353
	PM2.5	.575	.095	.575	6.025	.000
	NO2	-584.391	399.408	-.039	-1.463	.144
	SO2	-187.129	105.715	-.067	-1.770	.077
	O3	-100.913	71.928	-.041	-1.403	.161
	CO	-.488	2.232	-.008	-.219	.827
	WS	.211	1.423	.005	.148	.882
	RH	-.572	.118	-.217	-4.842	.000
	TEMP	-1.866	.570	-.138	-3.276	.001

a. Dependent Variable: PM2.5D1

Coefficients[a]						
		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	120.538	26.087		4.621	.000
	PM10	.148	.082	.183	1.806	.071
	PM2.5	.413	.102	.413	4.046	.000
	NO2	-606.069	427.207	-.041	-1.419	.156
	SO2	-343.534	113.285	-.122	-3.032	.002
	O3	-156.351	76.926	-.064	-2.032	.042
	CO	.763	2.387	.013	.320	.749
	WS	2.264	1.523	.051	1.487	.137
	RH	-.565	.126	-.215	-4.474	.000
	TEMP	-2.062	.610	-.153	-3.384	.001

a. Dependent Variable: PM2.5D2

APPENDIX B: Multiple Linear Regression coefficients for Bukit Rambai monitoring station

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	76.032	15.811		4.809	.000
	PM10	.689	.131	.689	5.263	.000
	PM2.5	.201	.144	.185	1.398	.162
	SO2	311.388	421.572	.013	.739	.460
	NO2	45.193	130.556	.008	.346	.729
	O3	-99.642	45.507	-.047	-2.190	.029
	CO	.296	1.821	.004	.162	.871
	WS	-1.209	.621	-.050	-1.946	.052
	RH	-.426	.088	-.145	-4.838	.000
	TEMP	-1.191	.341	-.082	-3.489	.001

a. Dependent Variable: PM10D1

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	83.806	20.444		4.099	.000
	PM10	.487	.169	.487	2.878	.004
	PM2.5	.350	.186	.322	1.881	.060
	SO2	470.009	546.552	.020	.860	.390
	NO2	28.157	169.133	.005	.166	.868
	O3	-207.820	58.849	-.097	-3.531	.000
	CO	-1.236	2.355	-.015	-.525	.600
	WS	-.669	.806	-.028	-.831	.406
	RH	-.456	.114	-.155	-4.000	.000
	TEMP	-1.201	.442	-.082	-2.720	.007

a. Dependent Variable: PM10D2

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	69.200	13.812		5.010	.000
	PM10	.013	.114	.014	.115	.909
	PM2.5	.856	.126	.856	6.802	.000
	SO2	202.126	368.270	.009	.549	.583
	NO2	10.628	114.049	.002	.093	.926
	O3	-70.782	39.753	-.036	-1.781	.075
	CO	.844	1.591	.011	.530	.596
	WS	-1.764	.543	-.080	-3.252	.001
	RH	-.429	.077	-.159	-5.577	.000
	TEMP	-1.001	.298	-.075	-3.356	.001

a. Dependent Variable: PM2.5D1

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	74.023	18.144		4.080	.000
	PM10	-.015	.150	-.016	-.099	.921
	PM2.5	.821	.165	.821	4.964	.000
	SO2	243.641	485.074	.011	.502	.616
	NO2	-9.657	150.108	-.002	-.064	.949
	O3	-153.442	52.229	-.078	-2.938	.003
	CO	-.302	2.090	-.004	-.145	.885
	WS	-1.338	.715	-.060	-1.871	.062
	RH	-.453	.101	-.168	-4.480	.000
	TEMP	-.980	.392	-.073	-2.500	.013

a. Dependent Variable: PM2.5D2

APPENDIX C: Multiple Linear Regression coefficients for Batu Pahat monitoring station

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	44.826	14.851		3.018	.003
	PM10	.622	.127	.623	4.898	.000
	PM2.5	.198	.143	.179	1.386	.166
	SO2	-464.707	592.451	-.015	-.784	.433
	NO2	837.058	180.850	.116	4.628	.000
	O3	-115.809	50.509	-.048	-2.293	.022
	CO	.956	1.576	.012	.607	.544
	WS	1.389	.886	.032	1.568	.117
	RH	-.262	.082	-.085	-3.191	.001
	TEMP	-.756	.334	-.054	-2.260	.024

a. Dependent Variable: PM10D1

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	47.586	19.242		2.473	.014
	PM10	.548	.165	.548	3.328	.001
	PM2.5	.152	.185	.138	.824	.410
	SO2	-1325.670	767.747	-.042	-1.727	.085
	NO2	1153.460	234.267	.159	4.924	.000
	O3	-164.417	65.417	-.069	-2.513	.012
	CO	.217	2.040	.003	.106	.915
	WS	2.956	1.148	.068	2.574	.010
	RH	-.299	.106	-.097	-2.812	.005
	TEMP	-.663	.433	-.048	-1.531	.126

a. Dependent Variable: PM10D2

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	43.496	12.931		3.364	.001
	PM10	.021	.111	.023	.188	.851
	PM2.5	.769	.124	.770	6.187	.000
	SO2	-442.357	515.865	-.015	-.858	.391
	NO2	857.839	157.471	.131	5.448	.000
	O3	-99.313	43.979	-.046	-2.258	.024
	CO	.296	1.372	.004	.216	.829
	WS	.687	.771	.018	.891	.373
	RH	-.315	.071	-.113	-4.416	.000
	TEMP	-.605	.291	-.048	-2.077	.038

a. Dependent Variable: PM2.5D1

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	41.641	16.972		2.454	.014
	PM10	.186	.145	.205	1.279	.201
	PM2.5	.467	.163	.467	2.864	.004
	SO2	-1113.741	677.159	-.039	-1.645	.100
	NO2	1173.836	206.625	.179	5.681	.000
	O3	-114.299	57.698	-.053	-1.981	.048
	CO	-.425	1.799	-.006	-.236	.813
	WS	2.024	1.013	.052	1.998	.046
	RH	-.334	.094	-.119	-3.559	.000
	TEMP	-.482	.382	-.038	-1.261	.208

a. Dependent Variable: PM2.5D2

APPENDIX D: Multiple Linear Regression coefficients for Kuala Terengganu monitoring station

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	37.641	11.758		3.201	.001
	PM10	.476	.065	.475	7.340	.000
	PM2.5	.320	.077	.277	4.142	.000
	SO2	-55.555	542.259	-.002	-.102	.918
	NO2	421.423	176.651	.060	2.386	.017
	O3	168.862	48.688	.088	3.468	.001
	CO	-7.556	2.203	-.090	-3.430	.001
	WS	1.552	.770	.046	2.014	.044
	RH	-.180	.070	-.074	-2.585	.010
	TEMP	-.633	.255	-.067	-2.481	.013

a. Dependent Variable: PM10D1

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	27.165	14.339		1.895	.058
	PM10	.466	.079	.465	5.897	.000
	PM2.5	.200	.094	.173	2.122	.034
	SO2	-503.377	662.042	-.018	-.760	.447
	NO2	436.081	215.465	.062	2.024	.043
	O3	183.336	59.367	.095	3.088	.002
	CO	-9.133	2.686	-.108	-3.401	.001
	WS	.729	.939	.022	.776	.438
	RH	-.047	.085	-.019	-.553	.580
	TEMP	-.494	.311	-.052	-1.588	.113

a. Dependent Variable: PM10D2

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	11.818	9.806		1.205	.228
	PM10	.027	.054	.031	.491	.623
	PM2.5	.713	.064	.712	11.072	.000
	SO2	-784.116	452.197	-.033	-1.734	.083
	NO2	440.498	147.311	.072	2.990	.003
	O3	104.806	40.602	.063	2.581	.010
	CO	-3.879	1.837	-.053	-2.112	.035
	WS	.607	.642	.021	.945	.345
	RH	-.132	.058	-.063	-2.273	.023
	TEMP	.061	.213	.007	.289	.773

a. Dependent Variable: PM2.5D1

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-4.145	12.220		-.339	.735
	PM10	.128	.067	.148	1.902	.057
	PM2.5	.478	.080	.478	5.958	.000
	SO2	-1408.247	564.232	-.060	-2.496	.013
	NO2	453.462	183.633	.075	2.469	.014
	O3	129.275	50.596	.077	2.555	.011
	CO	-5.493	2.289	-.075	-2.400	.017
	WS	.031	.801	.001	.038	.969
	RH	.004	.072	.002	.054	.957
	TEMP	.351	.265	.043	1.326	.185

a. Dependent Variable: PM2.5D2