Estimation of Ground Level $PM_{2.5}$ Concentration Using Aerosol Optical Thickness from Modis Images in Peninsular Malaysia

Khaled Ali Ahmed Ben Youssef

Doctor of Philosophy
Universiti Putra Malaysia
2019

ESTIMATION OF GROUND LEVEL $PM_{2.5}$ CONCENTRATION USING AEROSOL OPTICAL THICKNESS FROM MODIS IMAGES IN PENINSULAR MALAYSIA

By

KHALED ALI AHMED BEN YOUSSEF

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

October 2019

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

ESTIMATION OF GROUND LEVEL $PM_{2.5}$ CONCENTRATION USING AEROSOL OPTICAL THICKNESS FROM MODIS IMAGES IN PENINSULAR MALAYSIA

By

KHALED ALI AHMED BEN YOUSSEF

October 2019

Chairman : Professor Ahmad Makmom, PhD
Faculty : Environmental Studies

Fine particulate matter is particulate matter lower in diameter than 2.5 μm ($PM_{2.5}$). It affects the public health, economic development, and the regional climate. Governments worldwide have been concerned about the levels of the $PM_{2.5}$ in the atmosphere and long ago began monitoring their levels continuously using air quality monitoring stations. The Malaysian government invested much in building ground monitoring stations. Most of these stations have been located in urban areas. In this context, use of remote sensing (RS) techniques and the geographic information system (GIS) in estimating the levels of the ambient $PM_{2.5}$ become more widespread. The Moderate Resolution Imaging Spectroradiometer (MODIS) Aerosol Optical thickness (AOT) product of the Terra Satellite can be used to estimate the $PM_{2.5}$ levels with an accuracy that depends on the statistical relationship between $PM_{2.5}$ and AOT. This study aimed at estimating the $PM_{2.5}$ mass concentration in Peninsular Malaysia by linking the MODIS sensor data with the measured $PM_{2.5}$ concentrations and meteorological parameters in the year 2013. The first objective of this study was to validate MODIS-AOT retrievals with ground AERONET AOT data after extracting the AOT from MODIS images. The second objective of the study was to correlate the ground $PM_{2.5}$ levels with the validated MODIS AOT after identifying the spatial and temporal AOT distributions by using Multiple Linear Regression Analysis (MLRA) and Geographically-Weighted Regression models (GWRMs) for Peninsular Malaysia. The third objective of the study was to accurately evaluate the accuracy of estimates of the relationship between the $PM_{2.5}$ levels and the corresponding MODIS AOT data in different pixel size groups. The geographic domain of this study was Peninsular Malaysia, which has an area of about 131,598 km^2, covering 40% of the land area of Malaysia and hosting approximately 80% of its population and economic activities. The methodology of the study consisted of three stages. First, the values of AOT were extracted from MODIS images and the AOT retrievals were validated with ground AERONET- AOT data. Second, the MLRA and GWR modeling were attempted to

spatially and temporally correlate the reported ground $PM_{2.5}$ levels and meteorological data with the validated MODIS AOT data after identification of the spatial and temporal AOT distributions in Peninsular Malaysia in the year 2013. Subsequently, a comparison of strengths and weaknesses was held between the various generated models. Lastly, an assessment of the accuracy of $PM_{2.5}$ estimation has been conducted on the MODIS spatial models.

The results showed that the MODIS AOT retrievals have a good correlation with the ground observations derived from AERONET as indicated by the values of the coefficient of determination (R^2) for the linear regression models, which were 0.87 and 0.78 for the daily average and the hourly average (± 30 min) data, respectively. The map of distribution of AOT indicated that the AOT concentrated in the western coast of Peninsular Malaysia. Mostly, the spatial and temporal AOT values in the southwest monsoon were higher than in the northeast monsoon throughout the study period. The MLRA and GWRM both gave almost identical estimates of the $PM_{2.5}$ concentrations. The analysis outcomes revealed that the R^2 value for the hourly $PM_{2.5}$ regression model (0.66) was higher than that for the daily $PM_{2.5}$ (R^2 = 0.53). Comparison with the literature uncovers that results of estimation of $PM_{2.5}$ using AOT from MODIS for Peninsular Malaysia are similar to the results of other studies in other parts of the world. Furthermore, assessment of the accuracies of the hourly and daily estimates of $PM_{2.5}$ disclosed that the 5 x 5 pixel size model had the lowest values of the mean-squared error (MSE), root mean-squared error (RMSE), and relative root mean-squared error (rRMSE). The relatively low error values associated with the 5 x 5 pixel size model indicate the accuracy of this model in Peninsular Malaysia.

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

ANGGARAN KEPEKATAN $PM_{2.5}$ PARAS BUMI MENGGUNAKAN KETEBALAN OPTIK AEROSOL BERDASARKAN IMEJ DARIPADA MODIS DI SEMENANJUNG MALAYSIA

Oleh

KHALED ALI AHMED BEN YOUSSEF

Oktober 2019

Pengerusi : Profesor Ahmad Makmom, PhD
Fakulti : Pengajian Alam Sekitar

Zarah terampai halus adalah zarah terampai berdiameter kurang daripada 2.5 μm ($PM_{2.5}$). Ia memberi kesan kepada kesihatan awam, pembangunan ekonomi, dan iklim kawasan. Kerajaan-kerajaan dunia amat mementingkan tahap $PM_{2.5}$ dalam atmosfera dan telah mengawasi tahapnya secara berterusan melalui stesen pengawasan kualiti udara. Kerajaan Malaysia telah menyalurkan peruntukan yang tinggi untuk membina stesen pemantauan atas tanah ini. Kebanyakan stesen ini berada di kawasan bandar. Dalam konteks ini, penggunaan teknik penderiaan jauh (RS) dan Sistem Maklumat Geografi (GIS) untuk menganggar tahap ambien $PM_{2.5}$ semakin meluas. Produk Ketebalan Optik Aerosol (AOT) Spektroradiometer Pengimejan Resolusi Sederhana (MODIS) daripada Satelit Terra boleh digunakan untuk menganggar tahap $PM_{2.5}$ dengan ketepatan yang bergantung kepada hubungan statistik $PM_{2.5}$ dan AOT. Kajian ini bertujuan untuk menganggar kepekatan jisim $PM_{2.5}$ di Semenanjung Malaysia melalui hubungan diantara data sensor MODIS dengan kepekatan $PM_{2.5}$ yang telah diukur dan parameter meterologi bagi tahun 2013. Objektif pertama kajian adalah mengesahkan dapatan AOT-MODIS dengan data AOT AERONET selepas mengekstrak AOT daripada imej MODIS. Objektif kedua kajian adalah mengkorelasikan tahap $PM_{2.5}$ paras bumi dengan AOT-MODIS yang telah disahkan selepas mengenal pasti taburan ruangan dan masa AOT menggunakan Analisis Regresi Linear Berbilang (MLRA) dan Model Regresi berpemberat Geografi (GWRMs) untuk Semenanjung Malaysia. Objektif ketiga kajian adalah menilai ketepatan anggaran hubungan antara tahap $PM_{2.5}$ dan data AOT-MODIS yang berkaitan untuk kumpulan saiz piksel berbeza. Domain geografi kajian ini adalah Semenanjung Malaysia yang mempunyai kawasan seluas 131,598 km^2, merangkumi 40% kawasan tanah Malaysia dan kira-kira 80% populasi dan aktiviti ekonomi negara. Metodologi kajian terdiri daripada tiga fasa. Pertama, nilai AOT yang diesktrak daripada imej MODIS dan dapatan AOT disahkan dengan data AOT AERONET bumi. Kedua, pemodelan MLRA dan GWR dijalankan untuk membina korelasi ruangan dan masa antara tahap $PM_{2.5}$ paras bumi dan data meterologi yang dilaporkan

dengan data AOT-MODIS yang disahkan selepasa taburan ruangan dan masa AOT di Semenanjung Malaysia bagi tahun 2013 dikenal pasti. Selepas itu, kekuatan dan kelemahan model-model yang dibangunkan dibandingkan. Akhirnya, ketepatan model ruangan MODIS dalam mengaggar $PM_{2.5}$ dinilai.

Keputusan menunjukkan bahawa dapatan AOT-MODIS mempunyai korelasi baik dengan pemerhatian paras-bumi yang diambil daripada AERONET seperti yang ditunjukkan oleh nilai pekali penentuan (R^2) untuk model regresi linear iaitu 0.87 dan 0.78 untuk purata data harian dan purata data setiapjam (± 30 min) masing-masing. Peta taburan AOT menunjukkan bahawa AOT tertumpu di kawasan pinggir laut Barat Semenanjung Malaysia. Nilai ruangan dan masa AOT kebanyakannya lebih tinggi semasa musim Monsun Barat Daya berbanding Monsun Timur Laut sepanjang tempoh kajian. Kedua-dua MLRA dan GWRM memberikan anggaran kepekatan $PM_{2.5}$ yang hampir serupa. Keputusan analisis menunjukkan bahawa nilai R^2 untuk model regresi $PM_{2.5}$ setiap jam (0.66) adalah lebih tinggi daripada model regresi $PM_{2.5}$ data harian (R^2 = 0.53). Perbandingan dengan literatur menunjukkan bahawa keputusan anggaran PM2.5 menggunakan AOT daripada MODIS untuk Semenanjung Malaysia adalah lebih kurang sama dengan keputusan kajian yang dijalankan di kawasan-kawasan lain di dunia. Selain itu, penilaian ketepatan anggaran $PM_{2.5}$ harian dan setiap jam menunjukkan bahawa model saiz piksel 5 × 5 mempunyai ralat min kuasa dua (MSE) dan ralat min punca kuasa dua (rRMSE) yang paling rendah. Nilai ralat yang rendah yang dikaitkan dengan model piksel bersaiz 5 × 5 menunjukkan ketepatan model untuk Semenanjung Malaysia yang dibangunkan dalam kajian ini.

ACKNOWLEDGEMENTS

I express my thanks to all the people who made it possible for me to complete this research and thesis. I express my sincere gratitude to all of them with deep acknowledgment.

I am highly grateful to Professor Ahmad Makmom Bin Abdullah, my supervisor, who gave me the opportunity to work under his guidance. His support and encouragement made this work possible. Additionally, his thoughtful discussions on the theoretical and technical aspects of aerosol research and approach to problem-solving helped me much in developing a reasonable scientific understanding of the various issues related to my research and a problem-solving approach to research.

I express my gratitude to the members of my supervisory committee, Associate Professor Helmi Zulhaidi Bin Mohd Shafri and Dr. Zulfa Hanan Binti Ashaari for providing me with insightful and thoughtful suggestions for my research work and dissertation.

I acknowledge the scientists of the MODIS mission and the associated NASA professionals for generation of the data that I employed in this study. I would also like to thank the AERONET program for the ground level AOT measurements. In addition, I gratefully acknowledge the Department of Environmental Sciences in Universiti Putra Malaysia for providing me with fine particulate matter and metrological data. As well, I sincerely thank my wife for her support and my friends for assisting me indirectly or directly throughout my study and research work.

I certify that a Thesis Examination Committee has met on 30 October 2019 to conduct the final examination of Khaled Ali Ahmed Ben Youssef on his thesis entitled "Estimation of Ground Level $PM_{2.5}$ Concentration Using Aerosol Optical Thickness from Modis Images in Peninsular Malaysia" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Thesis Examination Committee were as follows:

Normala bt Halimoon, PhD
Senior Lecturer
Faculty of Environmental Studies
Universiti Putra Malaysia
(Chairman)

Mohammad Firuz bin Ramli, PhD
Associate Professor
Faculty of Environmental Studies
Universiti Putra Malaysia
(Internal Examiner)

Mohd Hasmadi bin Ismail, PhD
Associate Professor
Faculty of Forestry
Universiti Putra Malaysia
(Internal Examiner)

Lalit Kumar, PhD
Professor
School of Environmental and Rural Science
University of New Zealand
Australia
(External Examiner)

ZURIATI AHMAD ZUKARNAIN, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 2 January 2020

This thesis was submitted to the Senate of the Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Ahmed Makmom Abdullah, PhD
Professor
Faculty of Environmental Studies
Universiti Putra Malaysia
(Chairman)

Helmi Zulhaidi Mohd Shafri, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Zulfa Hanan Asha'ari, PhD
Senior Lecturer
Faculty of Environmental Studies
Universiti Putra Malaysia
(Member)

ZALILAH MOHD SHARIFF, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 09 JAN 2020

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software

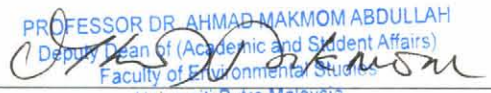
Signature: _____ Date: 17/01/2020

Name and Matric No: Khaled Ali Ahmed Ben Youssef, GS37721

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) were adhered to.

Signature: 

Name of Chairman of Supervisory Committee: Professor Dr. Ahmed Makmom Abdullah

Signature: 

Name of Member of Supervisory Committee: Associate Professor Dr. Helmi Zulhaidi Mohd Shafri

Signature: 

Name of Member of Supervisory Committee: Dr. Zulfa Hanan Asha'ari

TABLE OF CONTENTS

LISTS OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AERONET	Aerosol Robotic Network
AOD	Aerosol optical depth
AOT	Aerosol Optical Thickness
BHL	Boundary Layer Height
r	Correlation coefficient
R^2	Coefficient of determination
DB	Deep Blue
DEM	Digital Elevation Model
DT	Dark Target
EPA	Environmental Protection Agency
ESA	European Space Agency
GEOS-Chem	Goddard Earth Observing System Atmospheric Chemistry Transport
GIS	Geographic Information Systems
GWR	Geographically Weighted Regression
HDF	Hierarchical Data Format
LUTs	Look-up tables
MERIS	Medium Resolution Imaging Spectrometer
MISR	Multi-angle Imaging SpectroRadiometer
MLR	Multiple Linear Regression
MODIS	Moderate Resolution Imaging Spectro-radiometer
NASA	National Aeronautics and Space Administration
PM	Particulate Matter
PM_{10}	Particulate Matter (diameter <=10 μm)
$PM_{2.5}$	Particulate Matter (diameter <=2.5μm)

RH	Relatively humidity
RMSE	Root-Mean-Square Error
RS	Remote Sensing
SeaWiFS	Sea-viewing Wide Field-of-view Sensor
WHO	World Health Organization

CHAPTER 1

INTRODUCTION

1.1 Background of the Study

Evaluation of the air quality is a topical issue in atmospheric studies. Deterioration of local, regional, and globe air quality depends on the aerosol concentration in the atmosphere. The main pollutants of the air encircle particulate matter (PM), carbon monoxide, nitrogen oxides, sulphur oxides, and the ground level ozone. The Global Burden of Disease (GBD) project of the World Health Organization (WHO) ascribed more than 3.2 million premature deaths worldwide in 2010 to the ambient PM pollution. Moreover, in 2010, the ambient PM pollution ranked fourth in the list of health risk factors in East Asia (Gao et al., 2015). On the other hand, most of the countries in Asia have experienced fast economic development during the past decade. The increased urbanization, industrialization, and use of vehicles in cities in these countries, in addition to the trans-boundary haze pollution and the Asian dust phenomenon, have contributed to the rise in the concentrations of PM in Asian cities (Engel-Cox et al., 2004; Liu et al., 2017; Tahir et al., 2013).

In general, the coarse particles comprise alumino silicates and oxides of crystal elements and the principal sources of these particles encompass dust originating from the industry, roads, construction operations and material, agricultural activities, and construction. On other hand, the fine particulate matter ($PM_{2.5}$), also known as the respirable particulate material, is usually related to anthropogenic pollution (Mao et al., 2017). The $PM_{2.5}$ consists of varying amounts of nitrate, sulfate, and ammonium ions; water; organic compounds; elemental carbon, and small amounts of trace elements and soil dust. For comparison purposes, while the coarse-mode particles have an atmospheric half-life of minutes to hours, the half-life of $PM_{2.5}$ ranges from days to weeks. Furthermore, the $PM_{2.5}$ can travel hundreds to thousands of kilometers while the coarse particles can only travel one kilometer to tens of kilometers (Wilson and Suh, 1997).

The increase in the atmospheric levels of $PM_{2.5}$ has serious negative effects on the weather and contributes to climate change, besides negatively impacting public health and economic development. Effects of the $PM_{2.5}$ on climate change can be classified into direct and indirect effects. The direct effects include interactions of $PM_{2.5}$ with radiation, such as absorption and scattering of the solar radiation and terrestrial surface radiation, which influence temperature and the radiation budget balance (Wang, 2003; Xin et al., 2014). The indirect effects of $PM_{2.5}$ are mainly embodied in its influence on the density and chemical composition of the atmosphere, which, in turn, influence the climate. The indirect effects of $PM_{2.5}$ also include altering the characteristics of clouds, and even the precipitation, since the formation of clouds depends on the atmospheric composition and dynamics, among other factors (Yu et al., 2015).

Due to their very small sizes, the $PM_{2.5}$ can be deeply breathed and they find their way to the lungs from which they never go out (Sun et al., 2015). The WHO estimates that exposure to the outdoor air pollution in 2012 led to 3.7 million premature deaths (Ford and Heald, 2016). From a global perspective, Asia consists of countries with high population densities and a multitude of highly-polluted areas. For instance, it was reported that the levels of $PM_{2.5}$ in India are five times higher than the corresponding WHO standards (Figure 1.1).

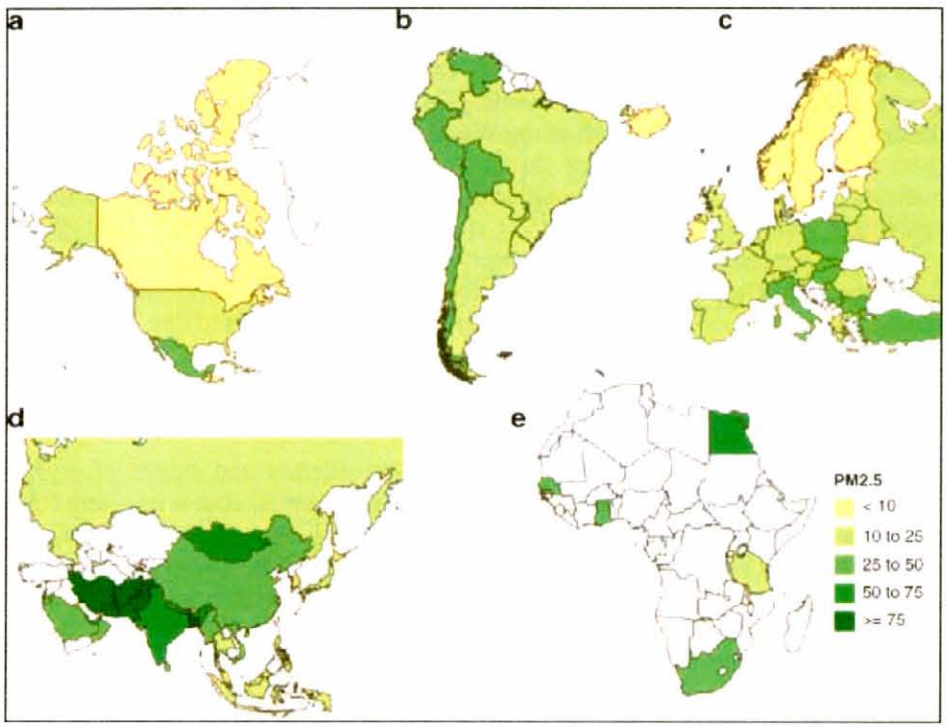


Figure 1.1 : Global Distribution of the $PM_{2.5}$ Concentration ($\mu g/m^3$) in Different Continents: (a) North America; (b) South America; (c) Europe; (d) Asia; and (e) Africa. Source: (Mukherjee and Agrawal, 2017)

Many epidemiological studies have shown that a rise in the risk of mortality from lung cancer and cardiovascular and respiratory diseases is associated with short-, and long-term exposure to $PM_{2.5}$, with the underlying presumption that causal relation exists between health outcomes and PM (Aurela et al., 2015; Changqing Lin et al., 2016; Wang et al., 2016). Besides its effect on human health and climate change, the $PM_{2.5}$ also results in economic losses. For example, the World Bank (2007) estimated the health cost of the urban PM pollution in China in 2003 at 157 billion Chinese Yuan by use of an adjusted human capital approach and at 520 billion Chinese Yuan by use of the value of the statistical life method (Kan et al., 2009). In Malaysia, the concentration of the $PM_{2.5}$ increases during the haze. As an example, the highest 24-h average $PM_{2.5}$ concentration was 136 $\mu g/m^3$ during the haze episode of 2015. For

comparison purposes, the corresponding concentration during non-haze times ranges from 14.3 $\mu g/m^3$ to 24.5 $\mu g/m^3$ (Latif et al., 2018).

So far, only few studies have examined the health effects of aerosols in Malaysia. The few available studies mainly concentrated on the health impacts of the aerosols during the haze crisis of the year 1997. Those studies found evidence on positive relation between number of the cases of diseases like conjunctivitis, asthma, and the upper respiratory tract infections and the Air Pollution Index (API). The high health damage cost, which was estimated at approximately RM129 million, was associated with the long duration of the 1997 haze episode (Kanniah et al., 2016).

Complaints about air pollution, which have been raised since 13th century when coal was first used in London (Bell and Davis, 2001), urged and pushed governments and public organizations to pay more attention than before to air pollution and, in consequence, relevant policies and standards were formulated. In the Air Quality Guidelines of the WHO, the yearly and daily (24 h) standards for $PM_{2.5}$ were set at 10 $\mu g/m^3$ and 25 $\mu g/m^3$, respectively (Organization, 2006). This guideline is widely used by decision makers across the world as a reference for setting air quality management goals and standards (Consultation, 2018).

In general, the ground-based air quality monitoring data are considered as accurate measurements of air quality. However, they usually represent the PM concentrations in relatively small regions (Sorek-Hamer et al., 2013). In addition, the number of ground stations is often limited and those stations are quite often sparse and unbalanced. This makes continual spatial monitoring of air quality difficult (Hu et al., 2013). Apart from spatial coverage and resolution, temporal coverage of the ground-level PM monitoring stations, which depends on the operation period and functionality of the instruments, varies highly (Karimian et al., 2016). Further, construction and maintenance of the ground stations are costly and time-, and labor-consuming.

In other respects, and in view of the drawbacks of ground monitoring of PM over vast areas, the remote sensing (RS) technology is the most important alternative for continuous exploration of PM over large areas (Kumar et al., 2008). This technology has various advantages. First, the images derived from the satellites can provide general and thorough information on air quality anywhere in the world (Chu et al., 2003). Second, the satellites provide an opportunity for acquiring global air quality data. Hence, they make it possible to discover the sources of the urban air pollutants and even global transport of those pollutants (Gupta et al., 2006). Furthermore, the RS technology corresponds to a low cost alternative for air quality monitoring for the developing countries which suffer from severe air pollution and yet lack ground-level monitoring stations. In this respect, previous studies reported significant correlations between the ground-level $PM_{2.5}$ concentrations and the satellite-derived Aerosol Optical Thickness (AOT) values (Chu et al., 2003; Van Donkelaar et al., 2016; Wang, 2003).

Monitoring mass concentrations of the PM at the ground level in association with the AOD is a challenging task. Numerous studies have been carried out to investigate the AOD-PM relation. The satellite-estimated PM concentrations are particularly essential for regions lacking ground data (Karimian et al., 2016). Even though the AOT describes polarized measurement of PM with no information on its vertical distribution, it, nonetheless, remains a reasonable and effective proxy for $PM_{2.5}$ prediction. Within this context, it is worth noting that AOD and $PM_{2.5}$ represent measures of two differing atmospheric pollutants (Gupta et al., 2006). The $PM_{2.5}$ concentration expresses the 'point' mass concentrations of particles near the surface while the AOD denotes the total columnar optical properties over an 'area' that is related to the spatial resolution of the measuring instrument. Several AOD retrieval algorithms have offered differing AOD products by using other devices such as the Moderate Resolution Imaging Spectro-radiometer (MODIS).

Few satellite sensors are available for AOD observations, including AVHRR, Multi angle Imaging Spectro-Radiometer (MISR), MODIS, and TOMS. Of these, the MODIS, which is boarded on Aqua and Terra satellites, is the most frequently used satellite sensor. Additionally, the user can use the 10-km MODIS AOD products that are provided by the National Aeronautics and Space Administration (NASA), rather than retrieving the AOD from the satellite images themselves. This product has been provided in 2000. It is based on the Dark Target (DT) algorithm and the Deep Blue (DB) algorithm. Various statistical models are used, including the simple linear regression model (SLRM), the multiple linear regression model (MLRM), and the geographically-weighted regression model (GWRM) and the artificial neural network (ANN) algorithms (Reid et al., 2013; Song et al., 2014).

Air pollution in Malaysia originates from mobile and stationary sources of pollution, e.g., open-burning activities. The stationary sources include emissions of dust from quarries and urban construction works, incinerators, energy power plants, and the industries. The open sources result in local pollution and encircle vegetation burning and trans-boundary pollution, especially during forest fires in neighboring countries due, mainly, to peat combustion. The mobile sources, on the other hand, are more related to traffic emissions (Abdul Halim et al., 2018). The increasing population will affect more resource consumption and the release of more air pollutants. This means that health of many more people will be affected by air pollution. Thus, monitoring and regulating local pollutant emission will be a challenging task in the following years. In other respects, studying $PM_{2.5}$ concentrations and emissions may help in reducing the international disputes. For example, in Klang Valley in Malaysia, haze has caused the PM_{10} mass concentration to increase more than four times and the levels of the AOT to increase more than three times over their levels before the haze episode (Amanollahi et al., 2011).

Numerous studies have examined the AOT over Malaysia (Amanollahi et al., 2011; Jamil et al., 2011; Kanniah et al., 2016; Kanniah et al., 2014), but there is limited research to date on the relationship between AOT and ground $PM_{2.5}$ concentration, and on how to utilize this information in $PM_{2.5}$ exposure modelling. This study

investigated the relationship between the $PM_{2.5}$ concentration and satellite-retrieved AOT using one-year (2013) data for Peninsular Malaysia.

Direct, though not perfect, correlation exists between the load of aerosol found in the atmosphere and AOD. Development of algorithms for retrieval of the MODIS-AOD irradiance observations turned out to be significant in the process of creating accurate estimates of the ground-level $PM_{2.5}$ concentrations (Lee et al., 2012). Other methods include the SLRMs of the relationship between AOD and measured $PM_{2.5}$, which has been developed for major cities worldwide to estimate the ground-level $PM_{2.5}$ concentrations with the aid of the satellite-derived MODIS-AOD (Evans et al., 2013; Martin, 2008). Though, light extinction around the total integrated atmospheric column is measured with MODIS whereas the aerosol amount is measured at the surface level only. Because of this, the correlation between these two variables is controlled by the vertical distribution of the aerosols and other meteorological parameters which influence the aerosol extinction coefficient like the mixing layer height, relative humidity (RH), and temperature (Li et al., 2009).

Even though satellite-based assessment of air quality is promising, it faces a number of challenges. Various factors can influence the relation between $PM_{2.5}$ and AOT. For instance, the satellite-derived parameter values provide columnar information about the ambient conditions while the $PM_{2.5}$ measurements represent the near-surface, dry mass concentrations. Moreover, the satellite data represent wide spatial areas and are prone to cloud interception (Zhang et al., 2005). But the daily data for the given area which these instruments provide at specific wavelengths make them great tools for assessment of pollutants in large areas.

Recent policy and research emphasis on the regional and intercontinental transport of varied air pollutants like $PM_{2.5}$ has spotlighted the need for additional sources of reliable data to augment the ground-based data for monitoring the air pollution, which, in effect, varies temporally and spatially. The satellite data can provide thorough information and visualization of the ground-based measurements of air quality and reliable data for air quality modelling. Numerous studies have examined the AOT over Malaysia (Amanollahi et al., 2011; Jamil et al., 2011; Kanniah et al., 2016; Kanniah et al., 2014). However, there is limited research thus far on the relationship between AOT and ground $PM_{2.5}$ levels in Malaysia, and on how to utilize this information in $PM_{2.5}$ exposure modelling. In response to this knowledge gap, this study investigated the relationships among the satellite-retrieved AOT values and the $PM_{2.5}$ mass concentrations using one-year data for Peninsular Malaysia.

Peninsular Malaysia lies in an area affected by high concentrations of $PM_{2.5}$ (Figure 1.1). Therefore, the present study focused on three issues: (i) characterizing the distribution of the MODIS pixels that cover the study area, taking into account seasonal variations; (ii) developing MLRMs of the relations among meteorological parameters and AOT data derived from the MODIS sensor at the local scale to estimate $PM_{2.5}$ concentration; and (iii) performing the analyses for individual pixels and array

of pixel (1 x 1, 3 x 3, and 5 x 5 pixels) for thorough assessment of effectiveness of AOT in estimation of the concentrations of $PM_{2.5}$ at different spatial resolutions.

1.2 Problem Statement

Particulate matter is an aggregation of solid and liquid particulates with different sizes and compositions. It is particles in suspension in the atmosphere whose concentration tells much about the urban air quality. For decades, air quality and pollution has raised high public health concern all over the world (Van Donkelaar et al., 2010) due to the ever growing urban and industrial developments. In particular, East Asia is the major source of aerosols globally, where anthropogenic pollutant particles, besides dust and sea salt, are abundant (Dai et al., 2014; Kim et al., 2014). Aerosol studies received high attention in southeast Asia and in Malaysia perhaps due to the extensive local generation of aerosols in the region, mainly including emissions from factories, automobiles, and open burning, including burning of wood, fossil fuels, trash, and plant leaves), in addition to the substantial amounts of trans- boundary air pollutants (Kanniah et al., 2014).

Remote sensing (RS) is an effective tool in the Earth science in particular and the climate studies in general for understanding the impacts of aerosols in the air on human health and environmental safety. Indeed, it is the only observational tool available for many parts of the world, including Southeast Asia. Owing to that the satellites have sensors on geostationary and polar platforms, remote sensing does actually provide systematic data over vast areas with high frequency. Additionally, movement and functioning of the satellites is not affected by the political and administrative boundaries. As such, the data they provide constitute the backbone of models. The satellites provide voluminous data on many parameters of the land, ocean, and atmospheric systems (Reid et al., 2013). However, with proliferation of population and urbanization, there are growing concerns about the urban air quality, which led to emergence of numerous studies on urban air turbidity during the last few decade to identify its sources and quantify the emissions of urban pollutants to the air (Clerbaux et al., 2010). Therefore, use of satellite RS allows for the collection of global or regional data at high resolution on the spatial and temporal distributions of most pollutants and contributes to our understanding of the air quality and changes in it over time.

Many researchers attempted to estimate the ground levels of $PM_{2.5}$ using satellite-based AOT data (Wang, 2003). The AOT is the extinction coefficient of aerosols of the points accumulating in the vertical direction (Hoek et al., 2013; Tao et al., 2013). The AOD research started in the mid 1970s. In 2003, Wang (2003) initiated use of the MODIS AOD in prediction of the ground-level $PM_{2.5}$ concentrations by linear regression analysis (LRA). Liu et al. (2004) developed a Chemical Transport Model (CTM) and (Lee et al., 2011) employed MODIS AOD data to develop day-specific Mixed-Effect Model (MEM). In recent years, the levels of $PM_{2.5}$ were estimated using various types of satellite sensors that included the Multi-Angle Imaging Spectrometer (J. Li et al., 2015)); MODIS (Geng et al., 2015; Zheng et al., 2016)); Polarization of

the Earth's Reflectance and Directionality (Léon et al., 2010); the Geostationary Operational Environment Satellite (Liu et al., 2009); the Cloud-Aerosol Lidar with Orthogonal Polarization (Toth et al., 2014); the Ozone Monitoring Instrument (J. Li et al., 2015); and the Sea-viewing Wide Field-of-view Sensor (Van Donkelaar et al., 2015a). Though studies of this sort are becoming more popular, their prediction results are in general unstable and varying from a region to another (Hu, 2009).

Peninsular Malaysia is an area with rapid development and population growth. Its weather is hot and humid with uniform temperatures throughout the year. Particulate matter (PM) is persistently creating an atmospheric problem in Malaysia, particularly since it has exceeded the standards of acceptable urban limits. For example, Jaafar et al. (2018) found that the concentrations of $PM_{2.5}$ in Malaysia far surpassed the USEPA air quality standard for $PM_{2.5}$ (a mean exposure rate of 35 $\mu g/m^3$ for 24 h) and the 2005 air quality guidelines of the WHO, that is, a mean exposure rate of 25 $\mu g/m^3$ for 24 h. The levels of $PM_{2.5}$ in the ambient air during the haze episodes are extremely high, ranging from 14.5 $\mu g/m^3$ to 160.9 $\mu g/m^3$. The highest $PM_{2.5}$ levels were reported in the third and fourth days of the haze episode, 146.2 $\mu g/m^3$ and 160.9 $\mu g/m^3$, respectively.

The PM has been categorized as one of the major air pollutants during the haze episodes. When its diameters fall within the respirable range, it can contribute to serious illness and, even, mortality because the fine particulates can enter the respiratory system easily through inhalation (Latif et al., 2018). The health effects of haze on the respiratory system and mortality and their effects on certain age groups were studied in Malaysia. For example, (Yaacob et al., 2016) conducted a study of the effect of haze on the peak expiratory flow rate (PEFR) of school children and found that there was a 15.0% reduction in the PEFR. Twenty two of the sample children had cough and headache. As well, mucus and throat symptoms were reported at high rates. Furthermore, Sahani et al. (2014) analyzed the health effects of haze in Klang Valley, Malaysia, from 2000 to 2007. Analysis uncovered that haze contributed a 19.0% increase in the respiratory mortalities and that there was a 41.4% growth in the delayed impacts of haze on the natural mortality of the children and a 66.0% growth in the respiratory mortalities of the adult females. In another example from Malaysia, Othman et al. (2014) reported that over the years of 2005, 2006, 2008, and 2009, there were 19 days on average of transboundary haze episodes during which the values of the API were equal to, or higher than, 76, that is, within the low to moderate air pollution hazard categories. On the average, every haze episode brought about an annual increase of 2.4% in the inpatient cases per 10,000 capita. Further, the study found that the marginal influence of haze on the inpatient rates was the highest for the children, followed by the young adults, the senior adults, and the infants.

Of all the constituents of the atmosphere, the aerosols stand as a highly-variable influential factor of solar radiation attenuation. It affects the climate, weather, and many weather-related phenomena and processes. In this context, the AOT is a spectral variable that reflects the amount of aerosols in a vertical column in the atmosphere from the viewpoint of their potential radiative effects. It is for this reason why having

a suitable representation of the AOD is of paramount importance for successful irradiance modeling (Bright and Gueymard, 2019).

In assessing urban aerosol content, satellite RS is a conventional tool that helps in measurement and prediction of ground-level PM concentrations. The satellites have the ability to monitor vast space, especially those areas in which ground monitoring stations are lacking. Past studies in Malaysia concentrated on the use of satellite data for measuring and monitoring the levels of PM_{10}. Only few studies addressed the levels of $PM_{2.5}$ in Malaysia. Thus, the use of RS data to estimate the levels of $PM_{2.5}$ in Peninsular Malaysia provides information that was not provided by the majority of previous air pollution studies in the country, which focused on the levels of air pollution in only certain parts of the country.

1.3 Significance of the Research

Atmospheric aerosols are among the most critical classes of the atmospheric pollutants. Various published works confirm that particles with small sizes have serious effects on human health, increasing the cardiovascular and respiratory diseases and reducing the life expectancy. This study underscores the benefits of using RS and geographic information system (GIS) data integrated with quantitative methods for air pollution studies. It uses satellite-derived AOT data for Malaysia, which is the main parameter employed in assessment of air pollution, particularly for quantifying the relationship between satellite data and measurements of ground air pollution monitoring stations. The study was carried out for Peninsular Malaysia, which has an area of almost 131,598 km^2 and hosts the majority of the population and economic activity in Malaysia.

In this study, firstly, data on the ground-level AOT were used to validate the latest MODIS data (collection 6 (C6)) available for Malaysia, which correspond to the period January 2012 to December 2013. Secondly, the AOT data used were Terra satellite data based on the MODIS sensor. They were employed to estimate the levels of $PM_{2.5}$ in Peninsular Malaysia in the year 2013. To the researcher's best knowledge, this study is the first study that used data derived from MODIS sensor to estimate the mass concentrations of $PM_{2.5}$ in Peninsular Malaysia.

The ground-based measurements of $PM_{2.5}$ are limited in Peninsular Malaysia. Data from only seven stations and one year (2013) were available for the study area. An alternative way was to combine the two datasets (the Satellite AOT and ground $PM_{2.5}$ data) to provide more robust information about the temporal and spatial distribution of $PM_{2.5}$ in Malaysia. The findings of this research can benefit decision making regarding air quality forecasts. Additionally, this study is useful for highlighting the health risks associated with exposure to PM and, ultimately for promulgating effective control strategies to protect the public health.

1.4 Research Questions

Since the $PM_{2.5}$ mass is measured from the ground irrespective of cloud cover while satellite data only provide AOT information during cloud-free and favorable retrieval conditions, so the researcher addressed the following questions:

1. What is the relation between the satellite and ground aerosol optical thickness (AOT) values?
2. How was the spatiotemporal distribution of AOT in Peninsular Malaysia in 2013?
3. How can satellite data on AOT be employed to determine the distribution of the fine particulate matter ($PM_{2.5}$)?
4. Are there relations among AOT and the meteorological parameters of temperature (T), relative humidity (RH), wind speed (WS), and wind direction (WD)?
5. Does the different array of 10km pixels affect the accuracy of $PM_{2.5}$ estimation?

1.5 Research Objectives

The overall purpose of this study was to estimate the mass concentrations of the $PM_{2.5}$ at the ground level in Peninsular Malaysia and explore the relationship between MODIS AOT and ground $PM_{2.5}$ levels. This purpose could be achieved by meeting the following objectives:

1. To validate MODIS-AOT retrievals extracted from MODIS images with ground AERONET-AOT data.
2. To correlate the ground $PM_{2.5}$ levels with the concomitant validated MODIS AOT data by using MLR and GWR modeling after identification of the spatiotemporal AOT distribution in Peninsular Malaysia.
3. To evaluate the relationship between $PM_{2.5}$ level and MODIS AOT data under different array of pixels using accuracy assessment.

1.6 Study Scope and Limitations

This research processed sets of quantitative and qualitative data in the GIS environment in an effort to estimate the levels of $PM_{2.5}$ in Peninsular Malaysia in the year 2013. The temporal and spatial distributions of $PM_{2.5}$ in the study area were also determined. The study used the latest Collection 6 (C6) of the AOT data that were obtained at the wavelength of 550 nm from the MODIS instrument aboard the Terra spacecraft, which travels across the equator from north to south in the morning. The MODIS sensor was selected because it provides daily data and the images it takes cover the entire study region. The hourly and daily ground concentrations of $PM_{2.5}$ ($\mu g/m^3$) were obtained from seven stations in Peninsular Malaysia.

Validation of the satellite data against ground measurements is challenging due to a number of reasons like different measured variables, lack of coinciding measurements, and problems in integration due to differences between the satellites readings and the point-based measurements of the ground instruments. Thus, validation of the reliability and accuracy of the AOT data derived from MODIS at the wavelength of 550 nm is required and is achieved by comparisons with independent data. In this context, the current study employed AERONET Level 2 AOT data as independent data for comparison purposes. Four-year historic data were used in the validation.

The period of study was between 2012 to 2015 in Peninsular Malaysia, there was only one AERONET station in the study area (USM_Penang; Peninsular Malaysia; Start running in November 2011), during that period the $PM_{2.5}$ data was limited, the researcher choose the year 2013 for estimate $PM_{2.5}$ because in that both the number of stations and the measured $PM_{2.5}$ data is relatively more obtainable. Effort has been made by the researcher to acquire $PM_{2.5}$ data in 2015 using handheld device (APECJ651), but due limited nature of that data it cannot be compared to the vast study area. However, after the period of study, the Department of Environment Malaysia has expanded the number of $PM_{2.5}$ monitoring stations over Malaysia (http://apims.doe.gov.my/public_v2/announcement.html). Additionally, the team of MODIS in 2015 developed a new spatial resolution MODIS Level-2 AOT with the spatial resolution 3×3 km^2 at nadir. Therefore, a study of this nature can be an opening for future research particularly regarding the test and constrains in the implementation of a new models for the estimation of $PM_{2.5}$ especially in the local scope considering the high resolution of the sensor (3×3 km^2).

The use of satellites for acquisition of surface data is an economical and reliable practice because satellite data can be obtained at low cost. So, local financial limitations will not be an obstacle to acquisition of such data. Thus, as a source of data for aerosol studies, satellites are the best option for less-developed nations, especially since the budgets for installation and maintenance of continuous air quality monitoring stations (CAQMS) is a significant constraint. However, one of the main limitations of AOT is that it is affected by clouds. Hence, it is missing when there is/was cloud coverage, which is the reason why quite often AOT data are missing at the point station in most MODIS pixels. Another limitation to this study is that some important covariates, e.g., planetary boundary layer (PBL) height and land use information, were missing. A widespread series of biomass burning events associated with forest fires, particularly in Sumatra and Kalimantan, Indonesia, caused a thick, smoky haze over a large portion of Southeast Asia. A huge amount of PM from this biomass burning was transported to Peninsular Malaysia. So, the concentrations of PM will be higher during hazy days than during non-haze days.

1.7 Organization of the Thesis

This thesis is organized in five chapters as follows:

Chapter 1 introduces to this study. It discusses sources and impacts of $PM_{2.5}$ and presents the objectives and significance of the study. Chapter 2 reviews the trends and present status of $PM_{2.5}$ and the implications of its levels on the population and the climate. As well, in this chapter, the researcher reviewed the measuring techniques in order to identify the most suitable of which to be adopted in this study. Chapter 3 outlines the research methodology, including the experimental design and structure, and the techniques used for determination of the $PM_{2.5}$ concentrations. Chapter 4 presents and discusses the research results. Chapter 5, then, lists the conclusions drawn from the research findings and gives recommendations for future research.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Large progress was observed in the last few years in the use of satellites for understanding optical properties of the atmospheric aerosols (Wang and Martin, 2007). Both natural and anthropogenic aerosols bring about lung and respiratory diseases and, eventually, lead to premature death (Gupta et al., 2006). Thus, monitoring the ground-level concentrations of $PM_{2.5}$ is critical for effective assessment of the air quality and evaluation of the potential related health risks (Bartell et al., 2013). Though, assessment of health risks of PM is quite challenging because their sizes matter a lot when it comes to their health implications.

The ground-based instruments are the instruments most widely used for measuring the concentration of ambient $PM_{2.5}$, both in the rural and urban areas of Asia, Australia, Europe, and the United States of America (USA). However, the ground-based instruments represent point measurements only. They do not have the spatial coverage that is required to map the local, regional, and global aerosol distributions (Chu et al., 2003; Wang, 2003).

Numerous factors influence the relationship between $PM_{2.5}$ concentration and AOT. For instance, the satellite data only provide columnar information about the ambient conditions while the $PM_{2.5}$ measurements represent the near-surface dry mass $PM_{2.5}$ concentrations. Moreover, the satellite data represent vast spatial areas and are, thus, prone to disruption by the clouds (Zheng et al., 2017) .

Although there are a number of satellite sensors, data from MODIS and MISR have high potential for mapping the global distribution of the aerosols and identifying their properties. This helps in obtaining indirect estimates of the PM, especially the $PM_{2.5}$, owing to that the ground measurements of the $PM_{2.5}$ are scanty in many regions in the world (Gupta et al., 2006).

2.2 Satellite Remote Sensing (RS) for Air Quality Analysis

The greenhouse gases, pollutant trace gases, and particles play major roles in air quality deterioration and negatively impact the climate and public health (Fiore et al., 2015; Xiong et al., 2015). The study of air quality in the site entails consideration of regional and global transportation of pollutants (Pinto and Grant, 1999). Identification of the of air pollutants and tracking their transport can be enhanced via satellite data in combination with other sources of data and modeling, both in the general and the specific event analysis (Engel-Cox et al., 2005). A substantial number of the satellite

instruments that are used in data collection for air quality applications are passive. These instruments detect electromagnetic radiation from the Sun after the processes of radiation absorption, reemitting, reflection, and scattering have taken place within the atmosphere and the Earth (Duncan et al., 2014; Martin, 2008).

Remote sensing of the atmosphere for climate and weather studies was initiated in the early 1960's and 1970's (Davis, 2007). Since the beginning of the implementation of sophisticated algorithms that make use of multiple wavelengths in aerosol RS, studies of the quality of satellite sensors has been greatly improved (Torres et al., 2002). The AOT data derived from setellite sensors are useful for air quality research aiming at monitoring and tracking aerosols over wide spatial domains (Wang, 2003). For instance, MODIS instrument has been providing aerosol data with reasonable accuracy since 2000. Consequently, a global model and satellite data together can extend the limited regional measurements to a much larger spatial scale (Chin et al., 2004).

An evaluation of the abilities of the main satellite instruments that have been designed for RS of aerosols was conducted by Martin (2008). In contrast, some researchers evaluated the prospects and limits of satellite sensors, their data results, and their application in air quality monitoring (Engel-Cox et al., 2004; Hoff and Christopher, 2012). This is necessary for air quality assessment in relation to aerosols and gases from the perspective of instruments orbiting beyond the atmosphere. Those researchers underlined several critical limitations of the satellite instruments in measuring atmospheric properties. The limitations they pinpointed include atmospheric transparency, orbit, the wavelength at which observations are taken, scattering, molecular spectroscopy, and absorption, which affect the measurement validity (Gutierrez, 2010).

A number of studies advocate the utility of satellites in assessment of air pollutants and weather parameters. As an example, Rhee and Im (2014) examined the daily and monthly estimates of the maximum and minimum air temperatures based on the MODIS sensor. Pour Biazar et al. (2011) presented the debut results of using the ozone profiles obtained from the Ozone Monitoring Instrument (OMI) and the AOD values obtained from MODIS to provide values of initial and lateral boundary parameters for regional air quality models above the USA. In another example, Zoogman et al. (2011) discovered that an Observing System Simulation Experiment (OSSE) is suitable for determination of the instrument requirements for geostationary satellite observations of ozone air quality in the USA.

Pope et al. (2014) reported that the quality of the OMI NO_2 tropospheric column product is good for detecting the synoptic influences of meteorology on the NO_2 tropospheric columns over the United Kingdom (UK). Data on AOT derived from satellite are valuable for studying air quality above local and large spatial domains to trail and monitor aerosols for estimation of the particulate matter (Kloog et al., 2015; Kong et al., 2016; Wang, 2003). Guo et al. (2016) assessed performance of using the PARASOL level 2 AOD only to forecast the ground-level $PM_{2.5}$ concentrations in

China. They used four empirical models: the quadratic regression model, linear regression model (LRM), logarithmic regression model, and power model. A comparison of the modeling results with the ground $PM_{2.5}$ concentrations revealed good match.

The satellites most frequently used for RS of air quality are the polar-orbiting, Sun-synchronous, low Earth orbit (LEO) satellites (Reid et al., 2013). These satellites can be utilized for viewing only the poles or observing similar places on the Earth simultaneously 24 h a day. These satellites are usually found at an altitude of around 800 km. A Sun-synchronous orbit is a special example of a polar orbit that passes over the equator at equal times in each orbit. Nevertheless, such satellites usually provide higher horizontal and vertical resolution than the geostationary (GEO) satellites due to closeness of the LEOs to the surface of the Earth (Vijayaraghavan et al., 2008). However, to simulate the aerosol vertical profile, the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite are the most suitable (Winker et al., 2007). Many studies revealed that vertical satellite data can be a good replacement for ground measurements as long as suitable adjustments are made to convert the columnar quantities to surface values.compositions

2.2.1 Satellite Spatial Resolution and Data Levels

The satellite data have numerous levels, according to the processing degree. The raw data acquired from the instrument are Level 0 (L0) and are subsequently processed to Level 1 (L1). The L1 data are formed by instrument pre-, and post-launch calibrations to yield radiances and geolocations of the data. Meanwhile, Level 2 (L2) and Level 3 (L3) data are treated from L1 to a geophysical factor, like AOD. The L2 and L3 data are the two most important levels for air quality (AQ) applications. The key dissimilarity between L2 and L3 data is that L2 data are the original observations and are not spatially gridded while the L3 data are mapped to a regular spatial grid. Their average over time is calculated based on day or month (Duncan et al., 2014) .

2.2.2 Satellites Temporal Resolution

Determination of the temporal resolution of satellite data is a function of several factors, including, for example, high temporal resolution global coverage (one day to two days revisit) of the MODIS (King et al., 1992). Most of the data are drawn from polar-orbiting satellite instruments (e.g., Terra, Aqua, and Aura) having 90-minute sequential orbits, hence attaining worldwide coverage in one or two orbits (Duncan et al., 2014). This indicates the importance of high spatial-temporal resolution in air monitoring studies, especially in the existing and potential megacities which are linked to air pollution predicament owing to industrial development and population explosion (Lili Li et al., 2015).

2.3 Particulate Matter (PM)

The quality of air is often assumed as the weather; it fluctuates, and, sometimes, in some days, it is better than in others. Changes in air quality are largely attributed to a specific meteorological and atmospheric thermodynamic environment lasting for certain days and resulting in gradual accumulations of air constituent concentrations. For example, Fiore et al. (2015) studied the quality of air and climate circumstances indicating an extensive array of surface levels of O_3 and $PM_{2.5}$ in the USA through the use of model projections as an alternative. Other studies have shown variations in the amounts of $PM_{2.5}$ throughout the season (Hu et al., 2009; Yadav et al., 2015).

Particulate matter (PM) is a combination of liquid and solid particles hanging in the air. It is usually classified as fine PM ($PM_{2.5}$; $d < 2.5$ mm) and coarse PM (PM_{10}; $2.5 < d < 10$ mm), where d denotes the aerodynamic diameter (Gupta and Christopher, 2008a). Greater toxicity has been ascribed by several studies to $PM_{2.5}$ than to PM_{10}, i.e., the coarse particulate matter (Lopez-Villarrubia et al., 2012; Polichetti et al., 2009). Moreover, the $PM_{2.5}$ can be transported to long distances (Radojevic, 2003). The bigger particles fall down to the ground surface, thereby polluting the soil or water bodies while the smaller particles are transported to far distances and continually mixed in the air (Anenberg et al., 2014).

A number of studies confirmed statistically-significant, high effects of $PM_{10\text{-}2.5}$ on the cardiac and cardiovascular systems. In addition, seven European studies reported that there was a significant high link between $PM_{2.5}$ and cardiovascular mortality (Pascal et al., 2014). Hansen et al. (2012) found that $PM_{2.5}$ pollution was more harmful than PM_{10} and that exposure during the cold season posed a greater risk to cardiovascular health than during the warm season. Correspondingly, national-scale, high-resolution maps of the daily levels of $PM_{2.5}$ are useful for identification of severe air pollution episodes and for health risk assessments (W. You et al., 2016a).

Though $PM_{2.5}$ ground monitoring sites give correct measurements, their spatial coverage is inadequate. This leads to deficiency in capturing the spatial variability of the levels of $PM_{2.5}$ (Chudnovsky et al., 2014). Zifeng Wang et al. (2010) stressed the importance of the vertical distributions of aerosols and RH, which show the critical effects of distributions of these two variables on the relationship between AOT and PM based on the concepts of atmospheric radiative transfer and optical properties of the aerosols. As a result, the numerous current advances and new sensors will boost the utilization of space-based aerosol measurements for forecasting the air quality (Al-Saadi et al., 2005).

The complex, multi-component mixtures of PM derived from natural and anthropogenic sources include different chemical materials, including metals, inorganic compounds, and organic compounds. In other respects, there are lots of air quality monitoring programs worldwide (Yadav and Satsangi, 2013) that led to the development of several national and international air quality control organizations to

regulate and monitor the ground-level concentrations of PM (Z. Ma et al., 2016). The Clean Air Act (CAA) of the USA mandates the Environmental Protection Agency (EPA) to create and control National Ambient Air Quality Standards (NAAQS) (Mcclellan, 2002). The Air Quality Guidelines of the WHO (Global Update 2005) set the annual and daily (24 h) standard for $PM_{2.5}$ at 10 $\mu g/m^3$ and 25 $\mu g/m^3$, respectively. This guideline is usually used for referencing during decision making throughout the world for the establishment of air quality management goals and standards (http://www.who.int/mediacentre/factsheets/fs313/en/).

2.4 Aerosol Optical Thickness (AOT)

An important element of the atmosphere of the Earth is the aerosols, which are made up of liquid material and solid particles with different sizes and chemical complexities (Prather et al., 2008). In general, the most common atmospheric parameter obtained from satellite observations is the AOT, which is a measure of disappearance of the electromagnetic radiation at given wavelength due to presence of aerosols in the atmospheric column (Chudnovsky et al., 2014; Wang et al., 2011).

A number of studies attempted to estimate the ground levels of $PM_{2.5}$ using satellite-derived AOT data (Chu et al., 2016; Van Donkelaar et al., 2006; Zheng et al., 2017). Yang et al. (2015) investigated the influence of variations in meteorological factors and pollutant emissions on decadal trends and inter-annual differences in losses of aerosols (sulphate, ammonium, nitrate, organic carbon, black carbon, and $PM_{2.5}$) from East Asia. The study was based on the Goddard Earth Observing System of chemical transportation model simulations of aerosols from 1986 to 2006 as determined by the assimilated NASA/GEOS-4 meteorological fields. The effects of variations in meteorological factors and emissions were separately carried out with sensitivity studies. Luan and Jaeglé (2013) employed the MODIS observations of AOD in combination with the GEOS-Chem chemical transport model to study the discharge of aerosols from East Asia and North America in the period 2004-2010. They discovered differences between the East Asian aerosol outflow and the North American aerosol outflow in diverse facets that included season, dust emission, and climate.

Satellite measurements show a trend of increase in AOD over East Asia (Choi et al., 2016). Numerous ground monitoring networks are currently in operation for monitoring aerosols in terms of AOT, which is a dimensionless variable that describes the electromagnetic radiation extinction at a particular wavelength (Khoshsima et al., 2014) and is used for diverse purposes. Van Donkelaar et al. (2010) described the Interagency Monitoring of Protected Visual Environment (IMPROVE) network, and the observation locations that form the State and Local Air Monitoring Stations (SLAMS) of EPA were described by Choi et al. (2009). (Aaltonen et al., 2012) gave an account on the Global Atmosphere Watch-Precision Filter Radiometer (GAW-PFR). The Chinese Sun Hazemeter Network (CSHNET) was described by L. L. Wang et al. (2010) and the German Weather Service (DWD) was described by Von Hoyningen-Huene et al. (2003), while the Australia Bureau of Meteorology (BOM)

was described by Cleugh et al. (2007). Mano et al. (2009) reported on Japan Meteorological Agency and Bouya et al. (2010) reported on the Atmospheric Radiation Measurement (ARM).

The properties of atmospheric aerosols having a high temporal resolution (15 min) are determined with AERONET, which is a network of ground-based Sun photometers that provides excellent screening for its distributed data (Maiersperger et al., 2013). The Sun-sky radiometer instruments that are located at the AERONET stations offer highly-accurate measurements of the aerosol optical properties from the ultraviolet to the near-infrared wavelength ranges (Holben et al., 1998). The AERONET stations are commonly employed in the validation of global or local AOT retrievals from the satellite sensor systems. Levy et al. (2010) found that the Sun photometer (AERONET) data can be used for ground truthing purposes. They used them to evaluate nine-year (2000-2008) MODIS-derived data of total AOD and aerosol size yields that were obtained from the Aqua and Terra sensors. The technique can evaluate the capacities of the aerosol models themselves without being affected by the performance of the SeaWiF and the atmospheric correction algorithm. Likewise, using this evaluation technique and the data obtained from 11 AERONET stations that are distant from the mainland in models disclosed good unison with the AERONET-measured values; having correlation coefficient (r) values greater than 0.86 (He et al., 2011). In the same manner, the AOD values drawn from MODIS, OMI, MISR, and CALIPSO satellite-borne instruments combined with those from four AERONET ground-based sites were used for validation of the satellite retrievals alongside ground measurements. In general, the MODIS and AERONET retrievals had excellent agreement with each other (Bibi et al., 2015).

2.5 Remote Sensing (RS) for AOT Retrieval

The use of satellite sensor in Earth observations is likely to be valuable practice for monitoring and mapping air pollutants as it provides thorough and synoptic views of vast areas in one snap shot (Themistocleous et al., 2012). Satellite RS can be a perfect substitute for the ground-level measurement. Based on capacity of the satellite observational systems to identify the importance of aerosols in the international climate change and local environment research, successions of satellite observational systems were developed at the end of the last century in an effort to achieve universal distribution of the aerosol optical features and the seasonal and annual differences of indirect and direct radiative forcing of the aerosols (Li et al., 2005).

The earliest sensor used in studies of AOD was the Advanced Very High Resolution Radiometer (AVHRR). In 1978, the AVHRR was on board the TIROS-N. Ever since, forty-year AVHRR datasets have been gathered, hence making it an exceptional repository for incessant long-term data of aerosol properties of land and ocean (Lovallo et al., 2009). Conversely, recovering AOD from AVHRR images is an arduous work due to band setting restriction and only a single wavelength is visible (Mei et al., 2014). Kaufman and Sendra (1988) recovered AOD from AVHRR data in a dark, dense vegetation region. In 2012, Li et al. (2013) proposed the Land Aerosol

property and Bidirectional reflectance Inversion using Time Series technique (LABITS) for retrieval of aerosol from NOAA AVHRR data. Though AVHRR AOD can help in the study of aerosol climatology and in the climate change research, even after the 1980s factors such as snow, ice, and cloud cover are still hindering AVHRR from retrieving AOD. Moreover, the monitor is very sensitive for the clouds, thus yielding a systematic and random error range that has fluctuations (Liu et al., 2004).

Akin to AVHRR, the Total Ozone Mapping Spectrometer (TOMS) is used from 1978 to the present for providing AOD data. The TOMS appended to Meteor-3 was terminated in December 1994, and subsequently replaced with TOMS, which is appended to the Earth Probe, and which started its operation in late July 1996. However, this created a gap in the data availability. This dataset comprises daily data for ozone and monthly AOD data at a resolution of 1°. As a backscattered close to ultraviolet radiation satellite, the TOMS possesses strong potential for identifying aerosol types and sizes (Torres et al., 2002).

From September 1997, the SeaWiFS regularly provides ocean color and atmospheric products (Menghua Wang et al., 2005). At the frequently-used wavelength of 550 nm, the over-land SeaWiFS aerosol dataset supplies retrieved AOD data, while the SeaWiFS bands are centered close to 490 nm and 670 nm. For some surface types, the band is centered near 412 nm. The principal quantity of interest is the AOD at the wavelength of 550 nm (Sayer et al., 2012). The SeaWiFS supplies daily images with pixels having a nominal spatial resolution of 1.1 km. Data of the SeaWiFS were temporally traversed between 14 September 1997 and 1 January 2010 (Schaeffer et al., 2012).

The MODIS possesses 36 channels cutting across the spectral range of 410 to 14,400 nm and signifying three spatial resolutions: 250 m, 500 m, and 1 km for two channels, five channels, and 29 channels, respectively. Aerosol retrieval uses seven of these channels (470-2,130 nm) to retrieve the aerosol characteristics (Chu et al., 2003; Remer et al., 2005). The MODIS acquires an AOD product (Terra: MOD04_L2; Aqua: MYD04_L2) at 10-kilometer resolution (Savtchenko et al., 2004).

The MODIS makes use of 3 channels: Band 3 channel at a wavelength of 470 nm (blue), Band 1 channel at a wavelength of 660 nm (red), and Band 7 channel at a wavelength of 2,130 nm (Mid-Infrared) to add an AOD at 550 nm (green). The wavelength of 2,130 nm is utilized due to the fact that the aerosols are almost transparent at this particular wavelength. In consequence, they can be employed in estimation of the surface albedo. So, the scattered and absorbed wavelength falls within the visible range of 470 nm (blue) to 660 nm (red) and is used for the calculation of the light extinction concentration ascribed to aerosols and a total AOD value at 550 nm is interpolated (Lee et al., 2011; Remer et al., 2013).

The retrieved AOD data are recorded at a resolution of 10 km^2, whereas MODIS originally uses a higher resolution of 0.5 km^2 for measuring reflectance values. Hence, all the 400 pixels (20 x 20) in the range of an area of 10 km^2 are inspected for cloud adulteration by a cloud-screening algorithm (Gupta and Christopher, 2008b). After removal of the cloudy pixels, more than 12 cloud-free pixels are still left, and the average reflectance of the residual pixels is used for retrieving AOD. The average reflectance measured is harmonized in a lookup table (LUT) having the pre-determined reflectance for several aerosol models in order to find the parameters that are highly related to the observed spectral reflectance for retrieving aerosol properties such as AOD (Schaap et al., 2009; Tian and Chen, 2010a).

Diner et al. (1998) reported that the algorithm of MISR for standard aerosol retrieval operates in a functioning, fully automatic mode. Moreover, the MISR gives reports on AOT and the type of aerosol at a resolution of 17.6 km through the analysis of MISR top-of-atmosphere radiances, starting from 16 x 16 pixel patches of a resolution of 1.1 km (Kahn et al., 2005). This quality results in a great increment in the ability to significantly evade cloud disturbance (Hoff and Christopher, 2012). Nevertheless, because of the features of the orbit of MISR, the observing monitor goes through some points close to the equator three or four times in the month. There is a great concern about a long time break as this creates difficulties in aerosols research in the areas where ground monitoring is absent (Kahn et al., 2005).

In addition to MISR, Advanced Along-Track Scanning Radiometer (AATSR) also gives double-angle observation (55° and 90°). Like AATSR, the Medium Resolution Imaging Spectrometer (MERIS) commenced on Envisat (Environmental Satellite) in 2002. The error produced by atmospheric scattering and single angle monitor absorption is reduced by the angle observation of MISR. Higher accuracy is achieved in desert zones with the AOD derived from AASTR than with AOD derived from MODIS. However, lower accuracy is obtained in the land in other regions than that obtained from the AOD derived from MISR (Bevan et al., 2012). In the same manner, the AATSR and MERIS data demonstrate good accuracy in arid areas. Yet, the AOD retrieved from MERIS data above the water surface can not produce an ideal accuracy (Benas et al., 2013b).

Many algorithms have been established for retrieving AOD from AATSR and MERIS. For instance, North et al. (2008) suggested the MERIS/AATSR synergy algorithm and Lyapustin et al. (2011) proposed the Multi-Angle Implementation of Atmospheric Correction (MAIAC) algorithm while Beloconi et al. (2016) reported aerosol observation for the estimation of urban PM_{10} and $PM_{2.5}$ concentrations in Greater London Area in 2016 based on synergistic MERIS/AATSR.

Data obtained from images of the Landsat 8 satellite were employed for retrieving AOD in Beijing through the MODIS DT method in combination with the visible near-infrared (VNIR) atmospheric correction method (ACM), whereby verification of accuracy is done by observing data from AERONET. The results revealed that the DT and VNIR method can be successfully used for inverting AOD in Beijing using data

from Landsat 8 satellite. Though, greater accuracy was obtained from the DT method than from the VNIR method, which had root mean-squared error (RMSE) values of 0.195 and 0.282, respectively (Ou et al., 2017).

He et al. (2015) used HJ-1 charge-coupled device (CCD) data, MODIS surface reflectance data, and the bidirectional reflectance distribution function (BRDF) data to investigate the practicability of satellite data in retrieving atmospheric aerosols for evaluating a delta region of the Yangtze River that was facing terrible atmospheric pollution. Because of absence of the 2.1 μm band that is not exposed to the effect of atmospheric aerosols from the HJ-1 CCD data, the MODIS surface reflectance and BRDF data products are required to retrieve the surface reflectance data (He et al., 2015). Li et al. (2012) recommended a novel aerosol retrieving algorithm that combines HJ-1 CCD and MODIS data. With this algorithm, high resolution (e.g., 100 m x 100 m) AOD can be acquired. Ground measurements with good precision were utilized in validation of the retrieved information. A high resolution is one of the advantages in comparison with the MODIS aerosol products because retrieval is of high resolution (100 m x 100 m) and numerous textural details and properties of the spatial distribution of aerosols can be acquired.

Numerous factors determine the retrieval accuracy and determine how good the different retrievals are. Retrieval limitations or weaknesses and instrumental designs are among the crucial factors for determining the quality of the RS aerosol data. Instrument calibration, availability of spectral channels for retrieving aerosol with weak overlap, non-homogenously distributed gases with potential for cloud screening, and spatial resolution are essential elements. Likewise, there are numerous features that affect the quality of the retrieval like the selected code for radiative transfer, surface reflectance treatment, procedure for cloud screening, and aerosol microphysical model. The potential of the aerosol retrieval microphysical model for successful retrieval relies largely on instrumental design. However, quite few user specifications are essential. This is, for instance, relating to the single scattering albedo of the aerosols (Myhre et al., 2005).

The images derived from multi-spectral satellite give radiance measurements that need to be converted to AOD, giving rise to an inverse problem. To overcome this problem, cloud identification is firstly carried out and the identified clouds are deleted from the satellite images to acquire the surface reflectance. The LUTs from which AOD is retrieved are compiled using the radiative transfer algorithms with pre-determined surface, atmospheric properties, and aerosol properties. In the case of the sub-micron aerosols, the AOD is retrieved over oceans with varying wavelengths that range from 0.47 to 2.1 μm through characterization of the ocean surface which serves as a function for solar angle, roughness, wind speed, and other factors. In the case of the over land aerosols, the background reflectance beneath the aerosol layer is necessary, so approximations are essential (Charlson et al., 1969; Kaufman et al., 2002). No association between the 0.55 μm and 2.13 μm wavelengths could be established for estimation of surface reflectance. Thus, MODIS independently retrieves the AOD at the wavelengths of 0.470 and 0.660 μm and interpolates these values to the wavelength

of 0.550 μm. Accordingly, for every pixel recognized as aerosols by the algorithm and for a specific Sun satellite observing geometry, the multispectral reflectance values produced by the satellites are utilized for obtaining AOD from the LUTs (Hsu et al., 2004; Levy et al., 2007).

Detection of aerosols by MISR on a different terrain reflectivity requires the aerosol phase function. With nine cameras pointed before and after, the MISR has revealed that the oblique views of the aerosols observed over a scene give a greater accuracy than that obtained from MODIS in AOD retrieval (Kahn et al., 2007) . Yang Liu et al. (2007) discovered the limitation of the MISR instrument, which is the narrow swath (360 km in comparison with the 2,400 km for MODIS) in which the cameras are actually viewed with a week, long repeat cycle close to the equator prior to survey of a region. Thus, the MISR has lower capacity than MODIS to occupy the spaces between the ground sensors.

Other measures than AOD are available that are beneficial in confirming the inherent characteristics of the aerosols. The aerosol absorption optical depth (AAOD) is measured using the GOME and OMI instruments. The sensitivity of the AAOD to the surface is less than that of OMI AOD or the aerosol index (AI) but the AAOD is highly sensitive to black carbon (BC) in aerosols and has been utilized in the detection of smoke aerosol plumes. Though this product is valuable for the delineation of plume and transport, sensitivity of the AAOD can still be lower due to the atmospheric opacity because of Rayleigh scattering. Hence, additional testing is required to confirm its usefulness for ground-based PM assessment (De Graaf et al., 2005; Torres et al., 2007).

The active RS CALIPSO (swath 64 km) views are much better when there is no solar background than they are during the day. Meanwhile, when compared with the AOD retrievals from the passive RS MODIS (swath 2,400 km) instrument, the RS CALIPSO requires daytime data (Torres et al., 2007).

2.6 Algorithm for AOT Retrieval

Determination of AOT from satellite sensors is important for assessment of the global climate, visibility, air quality, radiative forcing, and effect of the aerosols on the trace gases. Moreover, the satellite methods used in cloud retrieval centered on space-based observations (Kokhanovsky et al., 2010). Based on this, researches have been testing the AOT retrieval algorithms on the datasets of the satellites. Examples of these algorithms are the Multi-channel Reflectance algorithm (Levy et al., 2006)), the Contrast Reduction (CR) algorithm (e.g., Grosso and Paronis (2012)), the Retrieval Algorithm (SARA) and the DT algorithm (e.g., Bilal et al. (2014)), the Simplified Aerosol Deep Blue (DB) algorithm (e.g., Hsu et al. (2013)), and the Multi-channel Reflectance algorithm (Zhang et al., 2009).

The Polarization and Directionality of the Earth's Reflectances (POLDER) coupled with Polarization and Directionality algorithm was developed by Masuda et al. (1999) to be jointly applied polarization, and is known as bidirectional reflectance. The estimation of AOT depends on backward scattering polarization; a characterization carried out with the algorithm. Leroy et al. (1997) reported that the aerosol of pixels mixed with water over the land can be retrieved using this algorithm. Nevertheless, the polarization and directionality algorithm requires a Bidirectional Reflectance Distribution Function (BRDF).

Kaufman and Sendra (1988) used the Dark Object algorithm to approximate AOD through slight reflectance of dark vegetation in the visible spectrum. Ever since then, this technique has been generally employed in the generation of satellite-retrieved AOD, like the MODIS AOD data. Reflectance of the red and blue band for a dark object (i.e., pixel covered with vegetation) is negligible. Therefore, the reflectance received from the satellite can be taken as the atmospheric contribution. Identification of a dark object is normalized with the normal difference vegetation index (NDVI).

The increment of accuracy in AOD retrieval across open area such as the desert or bare land suggested a novel algorithm (the deep blue (DB) algorithm) as a result of the blue band. Low surface reflectance is observed at these particular blue wavelengths. Hence, changes in aerosols can be identified by total reflectance changes and improved spectral contraction (Ginoux et al., 2012; Hsu et al., 2004). Compared with earlier algorithms that primarily use bands that are longer than 600 nm (red), the deep blue algorithm can be implemented by any satellite monitors having blue band (Hsu et al., 2004). Though, initial monitors like AVHRR and TOMS lack the blue band. Therefore, the DB algorithm may not be applicable when carrying out historical AOD research.

2.7 Remote Sensing Accuracy Assessment

Accuracy expresses correctness. It measures degree of agreement between a standard value that is presumed to be correct and classified image of unknown quality. If the classification of the image matches with the standard value, then the image is described as accurate. In the statistical context, a high accuracy implies a low error and a low bias, that is, the estimated values are consistently close to the accepted reference value. High precision means that variability of the estimates (independent of their bias) is low (Andria et al., 2000).

2.7.1 Measurement of Accuracy of Maps

The task of accuracy assessment can be defined as a task of comparing two sources of information, one based on the analysis of remotely-sensed data (namely, the map) and another that is based on a differing source of information that is posited to be accurate, i.e., the reference data. The reference data are of obvious importance; if they are in error, then the attempt to assess the accuracy will be in error. Webster and Beckett

(1968) clarified that high accuracy in the statistical context means that the bias is low. In other words, the estimated values are consistently close to reference value. On the other hand, precision means that variability of the estimates (independent of their bias) is low. Usefulness of the map depends not only on its correctness but also on the precision with which the users can make statements about certain points that are depicted on the map.

Foody (2009) highlighted that one of the principal objectives of accuracy assessment is typically comparison, whether among classification results or against a reference standard. Statistically valid comparisons must have a sample size appropriate to the type of test, the effect size, power, significance level, and confidence limits of the comparison. In many studies, there are too few samples to permit statistically robust comparisons. Careful planning of the accuracy assessment and the comparison to hold can help in reducing use of designs that have limited chance of producing significant outcomes. In fact, as the resources are frequently limited and there is common desire for parsimony, the careful planning will additionally help in producing the most relevant design for the available resources.

2.7.2 Problems in Accuracy Assessment

Foody (2002) maintains that the fundamental approaches to accuracy assessment seem to be relatively straightforward; various problems are frequently encountered when assessing classification of an image. Those problems range from the issues related to failure to meet the fundamental underpinning presumptions to a limited amount of the information on quality of the map quality that is, in effect, conveyed by basic accuracy assessment. Various somehow interrelated problems limit quantification of the classification accuracy and, in consequence, use of the land cover maps that are derived from RS. These encircle types of errors, accuracy measure and reporting, spatial distribution of the error, sampling issues, employment of confusion matrix, accuracy of the reference or ground data and the land cover change product, use of the confusion, and the magnitude of error.

2.7.3 Accuracy Assessment MODIS Aerosol Retrieval

The underlying surface reflectance and aerosol models are the main factors which impact the accuracy of performance of the MODIS retrieval performance. Actually, the algorithm consists of three separate algorithms that are intended to derive AOT data over differing sorts of surfaces. The aerosol algorithm has been designed to draw global aerosol characteristics from window-band observations in the VNIR and the shortwave-infrared region. The dark target part denotes the separate retrievals over the dark (non-glint) deep ocean zones and dark (e.g., vegetated) land (Levy et al., 2009; Strandgren et al., 2014).

The DT approach is limited over bright surfaces for accurate AOD retrieval due to the differences in the spectral reflectivity ratio between dark and bright surfaces (Huang et al., 2018). To increase accuracy of AOD retrieval over bright areas like deserts or bare land, Hsu et al. (2004) proposed a novel algorithm (the DB algorithm) based on the blue band. Surface reflectance is low at these blue wavelengths so that change in aerosol can be detected with the change in the total reflectance and an enhanced spectral contrast.

Zifeng Wang et al. (2010) found that retrieving MODIS AOT over the land is practiced during daytime only and for the dark and cloud-free pixels. Therefore, the AOT retrievals are limited to the conditions prevailing over moderately-bright surfaces where the reflectance measured at 2.13 μm lies within the range of 0.15 to 0.25. The global products of the MODIS AOT have been offered on the website of NASA. Extensive effort of validation that used more than 8,000 MODIS retrievals collected by AERONET AOT measurement uncovered that, globally, the MODIS products possess an accuracy equivalent to ($\pm 0.05 \pm 0.15\ \tau$) over the land and to ($\pm 0.03 \pm 0.05\ \tau$) over the ocean. However, those estimates of accuracy apply to only the cloud-free pixels declared by AERONET (Baum et al., 2007).

Correlation with the Sun photometer data is a choice. It is a statistical indicator to employ for quantifying quality of the MODIS satellite product (Bréon et al., 2011). The statistical technique which was employed for quantifying the precision and accuracy of the differing MODIS algorithms in retrieval of the AOT is the orthogonal regression method ($Y = a + bX$) that is employed to estimate slope (b) and intercept (a) of the relation between AERONET AOT (X) and MODIS AOT (Y). The intercept serves as indicator of the uncertainty of surface reflectance estimation while the slope is linked with the error in the presumptions of the aerosol model. Further, the r, MAE, RMSE, the expected error (EE), and the relative mean bias (RMB) have been employed to assess the retrieval accuracy (Cheng et al., 2012).

2.8 Satellite Sensor for Prediction of Ground $PM_{2.5}$ Concentrations

Investigations of the empirical relationship between $PM_{2.5}$ concentration and the observational AOD are almost confined to China and the USA, with limited number of studies in Europe. Very few investigations were conducted over India (Sreekanth et al., 2017). Retrieval of the $PM_{2.5}$ concentration on the basis of satellite measurements of AOD can offer global coverage of the assessments of $PM_{2.5}$, which fills the gaps between the stations where surface measurements of $PM_{2.5}$ concentrations are lacking (Li et al., 2016). The levels of $PM_{2.5}$ can not be measured directly with satellite instruments. Nonetheless, they can be observed using AOT. If the $PM_{2.5}$ concentration is thoroughly combined in the boundary layer and the clouds are absent from the sky, then it is possible to express the AOT-$PM_{2.5}$ relation as:

$$AOT = PM_{2.5} x H x f(RH) x \frac{3Q_{ext,dry}}{4\rho r_{eff}} \quad (2.1)$$

where H is height of the boundary layer, $f(RH)$ is the ratio of the ambient to the dry extinction coefficients, ρ is the mass density of the aerosol, $Q_{ext,dry}$ is Mie extinction efficiency, and r_{eff} is the effective radius of the particle (Duncan et al., 2014).on

2.8.1 Factors Used in $PM_{2.5}$ Concentration Prediction Models

The inputs for model perfection, like visibility, commonly utilize natural meteorological parameters W. You et al. (2016a) such as the RH (Van Donkelaar et al. (2006); precipitation (Zheng et al. (2016); wind direction and speed and the dew point (Kumar et al. (2013); the boundary layer height (Liu et al. (2005); elevation (Paciorek and Liu (2009); pressure (Ma et al. (2014); and temperature (Liu et al., 2012).

Natural factors and meteorological variables affect $PM_{2.5}$ formation and dispersion. Certain weather conditions (e.g., visibility) are also affected by $PM_{2.5}$. Other factors are socio-economy, population, and road density (Paciorek and Liu (2009) and land use (Kloog et al., 2012).

2.8.2 $PM_{2.5}$ Concentration Prediction Models

The ground-level PM concentration is linked with the AOD derived from the satellite and can be empirically transformed into PM mass (Chu et al., 2003; Eck et al., 2005; Koelemeijer et al., 2006). Consequently, many empirical models have been established for prediction of the ground-level $PM_{2.5}$ or PM_{10} concentration from numerous satellite-derived AOD values (Table 2.1). The methods most commonly used for estimation of PM are simulation-based and observation-based methods.

Table 2.1 : Advantages and limitations of various models for estimation of the particulate matter (PM) concentration

Author	Place	Model	Advantages	limitations
(Beckerman et al., 2013; Mao et al., 2012)	USA	Land Use Regression	-Effective model in prediction of the long-term, intra-urban variations in air pollution. -Promising for air pollution exposure measurement in epidemiological studies, especially for the chronic health effects in large populations.	The variations in the levels of $PM_{2.5}$ are affected by height of the planetary boundary layer, the vertical aerosol structure, and the relative humidity, besides other meteorological variables.
(Stieb et al., 2015)	Canada			
(Di Nicolantonio et al., 2007)	Italy	Multivariate Regression	The multivariate regression model was relatively more appropriate than the linear correlation model and the confounding bias could be prevented by inclusion of suitable covariates in the model.	There are various limitations. Certain important covariates like seasonal variations in aerosol, regional variations, and land use information were missing from the models. Moreover, resolution and accuracy of the satellite-derived AOD and the meteorological data were low.
(Gupta and Christopher, 2009)	USA			
(Engel-Cox et al., 2004)	USA			
(Wang et al., 2013)	China			
(Zhang et al., 2009)	USA			
(Tsai et al., 2011)	China			
(Y. Liu et al., 2007)	USA			
(Y. Bai et al., 2016; Wei You et al., 2016; Zou et al., 2016)	China	Geographically Weighted Regression (GWR)	The GWRM can work with the daily, monthly, or yearly average AOD data alone and $PM_{2.5}$ data alone.	Because model construction is dependent on ground monitoring data, performance of the model may be of low reliability in areas that lack ground monitoring stations.
(Zifeng Wang et al., 2010)	China	Linear Correlations	The satellite-measured AOT can be empirically transformed into $PM_{2.5}$ mass by using the optical properties along height of the aerosol layer in the atmosphere.	-The correlation strength is affected by differences in terrain. In general, the low correlations can be found in the highly arid areas. -Cloud contamination reduces the number of the single pixels needed to provide a
(Chudnovsky et al., 2013; Engel-Cox et al., 2004; Wang, 2003)	USA			
(Koelemeijer et al., 2006)	Europe			

				strong positive correlation between AOT and $PM_{2.5}$ concentration.
(Geng et al., 2015) (Van Donkelaar et al., 2015b) (Kim et al., 2014)	China USA Korea	Chemical Transport	The model can forecast the $PM_{2.5}$ concentrations at the ground level without $PM_{2.5}$ data from the ground monitors and it takes into account the component of the AOD and effects of other pollutants.	The prediction impact was relatively variant and low among the different areas. Considering poor performance of the CTM, this model will consume energy, time, and financial resources to acquire the necessary physical and chemical information on $PM_{2.5}$ because of lack of data on types of pollutant emissions and emission listing in developed countries.
(Kloog et al., 2011; H. J Lee et al., 2016) (Xie et al., 2015; Zheng et al., 2016) (Just et al., 2015)	USA China Mexico	Mixed-Effect	This model can be applied widely to prediction of the $PM_{2.5}$ levels at the regional level by using varying meteorological and land use variables for model calibration. It can also be employed to predict the daily concentrations of $PM_{2.5}$. The model can also be employed to explore differences between the satellite-derived AOD-based $PM_{2.5}$ levels and the ground-based $PM_{2.5}$ levels in health effects research.	-Because of lack of the ground-level $PM_{2.5}$ monitoring data in some areas, the $PM_{2.5}$ monitoring data can not satisfy the requirements of the Kriging in MEM, which influences accuracy of the results. -Correlation between $PM_{2.5}$ concentration and AOD can drop when only the total AOD is used. It is not obvious which aerosols influencing AOD play the main role in total AOD or to what extent the other air pollutants impact this correlation. -Information on traffic pollution and land use is difficult to collect.
(Liu et al., 2012; Song et al., 2015) (Liu et al., 2009)	China USA	Generalized Additive (GA)	The model can predict distinctly-differing spatial patterns of the concentrations of $PM_{2.5}$.	The GAM can not account for the changing spatial patterns of $PM_{2.5}$ with time.

2.8.2.1 Observation-Based Methods

Empirical observation-based models depend on statistical regressions between the AOD and the *in situ* measured PM concentration. They encompass SLRM (Gupta et al., 2006; Wang, 2003); MLRMs, having more factors, for instance, T, AOD, BLH, WS, and RH (Benas et al., 2013a; Dimitriou and Kassomenos, 2014; Emili et al., 2010; Karimian et al., 2016); non-linear models (e.g., Baker and Foley (2011); land use regression (e.g.,(H. J Lee et al., 2016); and the GWR (Gupta and Christopher, 2009; Hu et al., 2014; Ma et al., 2014; Z. Ma et al., 2016). Furthermore, consideration of non-linear relationships among the dependent and independent variables can be assimilated by the GAM (Liu et al., 2009; Nakhle et al., 2015; Paciorek et al., 2008) and the Artificial Neural Network (ANN) algorithms (Gupta and Christopher, 2009).

2.8.2.2 Simulation-Based Methods

Recently Recently, the use of air quality simulation models grew into a principal practice in assessing the impacts of air pollution on the environment and individuals (Fox, 1981). The impacts of meteorology (such as the RH) and the vertical structure and properties of the aerosol like aerosol type and size distribution are taken into consideration by the simulation-based methods which are simulated with regional or global chemical transport models (Drury et al., 2010; Liu et al., 2004). Modern, three-dimensional air quality tools for modeling have always been underestimated for simulations of elevated PM_{10} and $PM_{2.5}$ concentrations. As well, three-dimensional air quality models have been developed in the last 15–20 years and considerable progress has been taking place in this research area (San José et al., 2008).

Philip et al. (2014) described the use of GEOS-Chem global chemical transport model (CTM) for calculating the native conversion parameters corresponding to each satellite observation. The temporal and three-dimensional spatial distributions are stimulated by this model in several aerosol components and gases by using integrated meteorology and emission inventories as the main inputs. A global simulation at $2^\circ \times 2.5^\circ$ spatial resolution was conducted along with three nested regional simulations at $0.5^\circ \times 0.667^\circ$ resolution from 2004 to 2008 via meteorological data assimilated from the Goddard Earth Observing System (GEOS-5) at the Assimilation Office of NASA Global Modeling (GMAO). The results of this global simulation are emphatically written with nested regional simulations across North America, East Asia, and Europe. The results indicated that the long-term average obtained from GEOS-Chem CTM simulated the long-term average ratio of $PM_{2.5}$ components to AOD.

The integrated process rate analysis used in the CMAQ model was employed to acquire quantitative data about the diverse processes in the atmosphere that influence the formation of the fine particles. The $PM_{2.5}$ process analysis shows that the principal fine particle emissions are the major sources of the high surface concentrations of $PM_{2.5}$ in three cities; Hangzhou, Suzhou, and Shanghai. The surface mass concentrations of $PM_{2.5}$ in Hangzhou and Suzhou used to have a horizontal transportation increment, which signifies the regional transport of PM in other region.

The aerosol transport process is the most influential process that brought about high $PM_{2.5}$ concentrations in Zhoushan because of the downwind location and its meteorological conditions conducive to formation of SOA such as high humidity. There can be a reduction in the $PM_{2.5}$ concentrations because of vertical advection and diffusion from a lower level to the upper air in urban centers. The planetary boundary layer (PBL) layer is the layer where the aerosol transport process takes place, making the $PM_{2.5}$ that exists in the urban areas undergo vertical transport from upper layers to the surface layer (Li et al., 2014).

The Weather Research and Forecasting (WRF) model is a non-hydrostatic model type, having some dynamic cores and several diverse varieties for physical parameterizations intended to denote the processes which the model can not resolve. This permits the application of the model to various dissimilar scales (Grell et al., 2005). The WRF-Chem model has been utilized by several researchers in the PM and meteorological field for simultaneous simulation of trace gases and aerosol in various studies like the studies of (Yang Bai et al., 2016; Gao et al., 2014; Jin et al., 2011; Sadavarte et al., 2016; Schwartz et al., 2012).

The three-dimensional atmospheric chemical conditions in July 2004 were estimated for the examination of the Community Multiscale Air Quality (CMAQ) model with meteorological fields Eta model. Moreover, the Eta-CMAQ system had also been positioned across the eastern part of the USA in the summer of 2004 for the provision of air quality forecast (AQF) guidance. The Eta-CMAQ AQF system comprised three principal components: (i) the Eta meteorological model for simulating the dynamic conditions of the atmosphere for the forecast period; (ii) the CMAQ model for simulation of transport, chemical evolution, and atmospheric pollutant deposition; and (iii) interface, PREMAQ, that is used for facilitation of the transformation of the Eta-derived meteorological fields to adapt with the CMAQ coordinate system, input data format, and grid structure (Mathur, 2008). A thorough assessment of the real-time forecast of $PM_{2.5}$, its chemical constituents, and its associated precursors was performed using the Eta-CMAQ model over the eastern USA by comparing model results. The daily variations of the observed $PM_{2.5}$ levels were reproduced and the majority (greater than 70%) of the observed $PM_{2.5}$ levels were captured (Yu et al., 2008).

For improvement of the AOT forecasts, global climate models are now utilized to simulate the standard satellite products. The foremost advantage of these models is their ability to simulate the factors (like size and chemical composition of the particles) that have effects on the relation between AOT and AOT. But the underlying principles and the operations are the main deficiencies between these models.

2.8.3 The Moderate Resolution Imaging Spectroradiometer (MODIS)

The The twin MODIS was launched aboard the Terra and Aqua satellites under the umbrella of the NASA Earth Observing System (EOS) program to acquire day-to-day remote measurements of aerosols across the land and ocean, and to have in-depth understanding of the global aerosol budget. Estimation of the contribution of the anthropogenic emissions to the aerosol budget and to radiative forcing is carried out using the data. The date 18 December 1999 marked the first day for the lunching of the Terra satellite. Its observation work commenced that very morning when it traveled across the equator during the day time at about 10:30 a.m local time (Román et al., 2014). On 4 May 2002, the Aqua satellite was launched for noon day observations. It traveled across the equator during the daytime at about 1:30 p.m local time (H. J Lee et al., 2016). The first group of confirmed products came up in Collection 2 (C2), but these were later swiftly replaced by C3 within the first two years of launch of Terra while C4 was the first steady, broadly-utilized and accurately-documented set of MODIS aerosol products (Remer et al., 2005). Nevertheless, an unacceptable bias level is inherent in the C4 aerosol product over land (Levy et al., 2005). The development and implementation of C5 was carried out using a second-generation land algorithm coupled with the implementation of other changes (Levy et al., 2007).

Levy et al. (2013) described the algorithms of the recent Collection 6 (C6) which retrieve AOD and aerosol size from the MODIS spectral reflectance. While there are no major developments over Collection 5 (C5), there are wide changes with pronounced impacts on the products and their interpretation. Three different retrieval algorithms were created from the C6 aerosol data sets which function over diverse surface types. The two DT algorithms use (i) dark in visible and longer wavelengths (over ocean) and (ii) over vegetated or dark-soiled land (dark in the visible), including the DB algorithm, initially developed for retrieving bright in the visible wavelength (over desert/arid land). A global DT-land and DT-ocean aerosol product will contain C6 at three-kilometer resolution. The novel C6 calibration approach eliminates key calibration patterns from the MODIS Level 1B data (Lyapustin et al., 2014). The provision of improved monitoring for the land, atmosphere, and ocean is one of the major factors to be considered when designing the MODIS (Justice et al., 1998; L. L. Wang et al., 2010).

2.8.3.1 The MODIS Sensor Data

To obtain the MODIS data, the EOS Pathfinder is used. The users prefer an advanced order of the data products, which is the major advantage of the Pathfinder. The challenge of data processing for the user is overcome by atmospherically-corrected, standard, temporal composites, or spatially-accumulated data sets, hence permitting the resources of each scientist to be dedicated to answering the science questions instead of processing data (Justice et al., 1998).

2.8.3.2 Validation of the MODIS AOT Retrievals

The Validation of the MODIS aerosol products is ongoing worldwide. The generated daily data are validated for MODIS and AERONET data in a very easy ASCII format (Ichoku et al., 2001). The development of the spatiotemporal approach has permitted the objective and quick satellite aerosol retrievals from the MODIS coupled with the retrievals of ground aerosol from AERONET (Ichoku et al., 2002). Use of the MODIS sensors started in 24 February 2000. The equipment is fortified with 36 different spectral channels ranging in wavelengths from 0.14-14 μm, with spatial resolutions of 250 m, 500 m, and 1.0 km, and provides worldwide aerosol retrieval coverage at NADIR 10×10 km grid (Chu et al., 2003). Retrieval errors of ($\Delta\tau a = \pm 0.05 \pm 0.2\ \tau a$) over the land were recorded in the first validation of MODIS-derived AODs. For comparison purposes, a somewhat lower error ($\Delta\tau a = \pm\ 0.03 \pm 0.05\ \tau a$) over the ocean was recorded for the AERONET Sunphotometer (Remer et al., 2002).

The accuracy of the MODIS global land surface albedo product (MOD43B3) was investigated by Wang et al. (2004), who employed ground measurements of the albedo for the period January 2001 to July 2003. The accuracy of 0.02 required by the climate models was met by the MODIS-derived albedo data. There are no clear differences between the MODIS albedo and the mean ground albedo measurements. The AOD derived from MODIS compares well with the AERONET data, where the aerosol observed is dominated by fine mode aerosol. This suggests that the single scattering albedo, which is the albedo assigned by the algorithm of the model is appropriate. These results were achieved in Western Europe, over the East Coast of the USA, and some parts of Asia (Levy et al., 2010).

Bibi et al. (2015) compared the values of AOD obtained from MODIS (DB), MISR, CALIPSO, OMI, and MODIS (Standard) satellite-borne instruments, including values obtained from four AERONET ground-based sites for the goal of validating the satellite retrievals against the ground measurements. Generally, comparing the MODIS (Standard) retrievals with those of the AERONET disclosed a high degree of correlation for the four investigated sites. However, the comparisons of MODISD retrievals with those of AERONET revealed higher correlations. At the same time, the comparisons between the retrievals of MISR and AERONET uncovered strong correlations as observed in the OMI-AERONET comparisons. Meanwhile, poor correlations were observed in the comparison of CALIPSO-AERONET retrievals due to the limited availability of point data.

2.8.3.3 The MODIS Sensor Estimate Ground-level $PM_{2.5}$

The relationship between AOD and $PM_{2.5}$ depends on many parameters such as particle optical properties, RH, the $PM_{2.5}$ vertical and diurnal concentration profiles, and surface reflectance (Chudnovsky et al., 2014). Use of satellite-derived AOT values for estimation of the levels of PM was first reported by Chu et al. (2003). These researchers concluded that MODIS aerosol retrievals give valuable views on the levels of air pollution worldwide. The MODIS has the capacity to monitor local and regional

air pollution, which is a monitoring that is realizable in any part of the globe using direct broadcasting of Terra and Aqua MODIS with data processing times that are usually less than an hour, that is, a monitoring that is close to the real time (twice a day). Generally, the results of comparison in the spatial patterns of the hourly $PM_{2.5}$ concentrations of the surface monitor with the derived-τa coincident MODIS patterns in the mid-Western USA and central Ontario disclosed comparable patterns (Kittaka et al., 2004).

2.8.3.4 Spatial Prediction of Particulate Matter (PM)

Van Donkelaar et al. (2010) compiled an international combination of AODs from MISR and MODIS (Figure 2.1) for improvement of the strength of correlation between the AOD data and the ground-based $PM_{2.5}$ measurements. They reported that extension of the satellite data for six or more years reduced the sampling errors. The unique global spatial resolution of $0.1° \times 0.1°$ holds relevant variation for population distribution. The application of GEOS-Chem is justified for tracking the aerosol vertical distribution. Van Donkelaar et al. (2010) discovered spatial agreement (r = 0.77) between the average ground-based $PM_{2.5}$ levels and the concomitant satellite-derived $PM_{2.5}$ data for North America

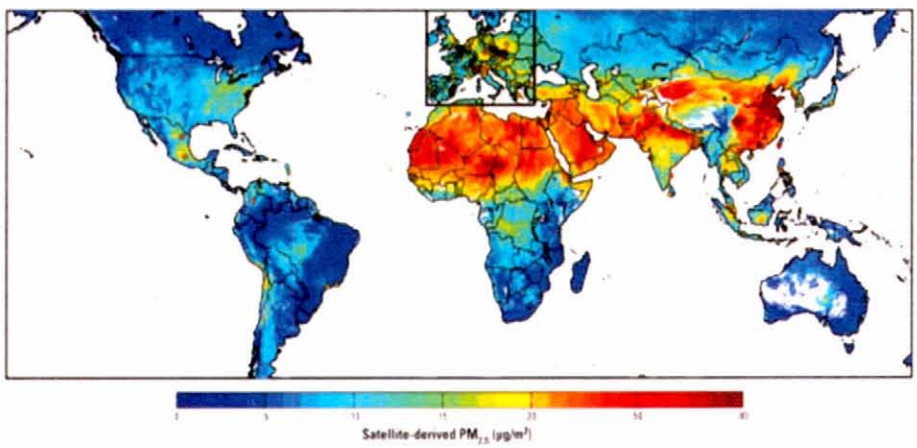


Figure 2.1 : Global, Satellite-based $PM_{2.5}$ Levels Averaged over Six Years (2001–2006)
The white space refers to water or areas with < 50 measurements.
The circles express the locations and values of the comparison sites outside the USA and Canada.
The black box denotes European sites
(Source: Van Donkelaar et al., 2010)

Likewise, Boys et al. (2014) used AOD data retrieved from SeaWiFS and MISR for production of an integrated, fifteen-year (1998 to 2012) world time series of ground-level $PM_{2.5}$ concentrations at a resolution of 1° x 1°(Figure 2.2). The relationship between ground-level $PM_{2.5}$ concentration and individual AOD retrieval was explored using GEOS-Chem. The highest mean annual $PM_{2.5}$ levels were found in East Asia. The major problem in the evaluation of the satellite-derived $PM_{2.5}$ levels worldwide is the lack of ground-based measurements (Van Donkelaar et al., 2015a).

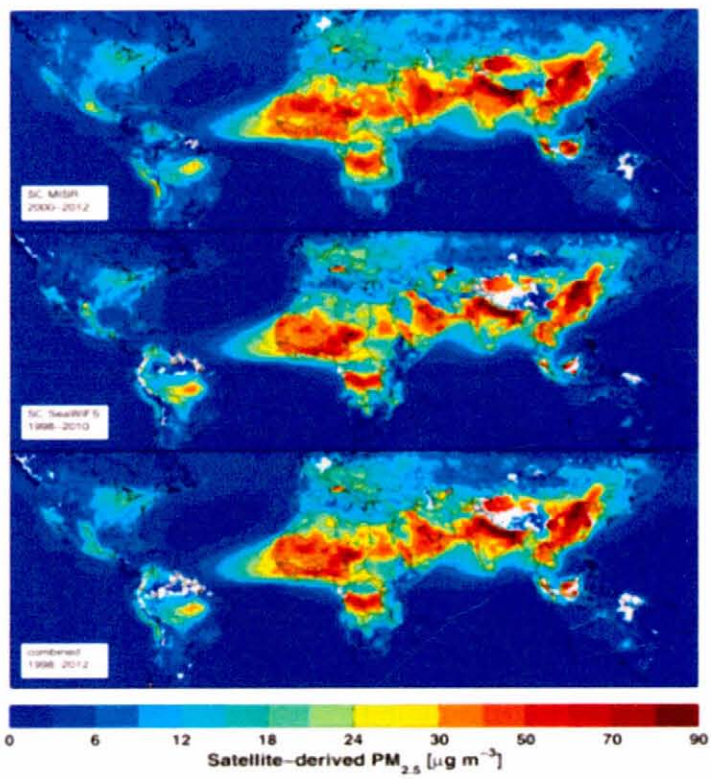


Figure 2.2 : Sampling-corrected 24h $PM_{2.5}$ Levels Drawn from MISR (Average of the Years 2000–2012), SeaWiFS (Average of the Years 1998–2010), and a MISR-SeaWiFS Combination (Average of the Years 1998–2012)
The grey color indicates missing data.
(Source: Boys et al., 2014)

A global dataset consisting of 210 ground-based literature values were employed by Van Donkelaar et al. (2015a) for evaluation of global satellite $PM_{2.5}$ assessments in several places in China and India. A high correlation ($r = 0.81$) was found despite the fact that the satellite $PM_{2.5}$ observations were considerably lower than the ground-based measurements (slope = 0.68). Van Donkelaar et al. (2016) presented a universally-applicable technique that combined geophysically-piloted simulations, the ground-based measurements, and satellite retrievals for improving the $PM_{2.5}$

representation at spatial scales proportionate to population density. Availability of higher temporal global $PM_{2.5}$ monitors will permit good representation by the GWRMs for seasonal-driven preference like the association with mineral dust and the burning of biomass. Areas subjected to heavy emissions of mineral dust constitute a problem for the satellite retrieval, the simulations, and the ground measurements.

Both at the regional and local levels, various research contributions have been made to elucidation of the correlations among the ground-level $PM_{2.5}$ concentration and MODIS data at the continent scale (Van Donkelaar et al., 2006)), the regional scale (Di Nicolantonio et al., 2007; Green et al., 2009; Guo et al., 2013; Hutchison, 2003; Kim et al., 2013; Kittaka et al., 2004; Y. Liu et al., 2007; Tian and Chen, 2007; Yao et al., 2012; W. You et al., 2016a), and even at the local scale, with a focus on group of specific cities (Guo et al., 2014; Gupta et al., 2006; Hutchison et al., 2005; Kong et al., 2016; H. J Lee et al., 2016; Y. Liu et al., 2007; Zhang and Li, 2015). Albeit it is agreed upon that MODIS AOD has a large potential for the estimation of the $PM_{2.5}$ concentrations, unbalanced data were detected based on the strength of correlation and their spatiotemporal variations.

2.8.3.5 Simple Linear Regression Models (SLRMs)

Chu et al. (2003) provided evidence on potential use of MODIS in monitoring of international, regional, and local air quality. They studied three areas simultaneously; India and eastern China, eastern US/Canada, and Western Europe. The selected areas represent the most industrialized and populous areas on Earth. Three urban regions were also included in the study: Greater Los Angeles, Beijing (China), and Northern Italy. The study had two central aims: (i) finding a way to generate AOD measurements from MODIS images with the same accuracy of the ground-based LIDAR measurements and (ii) determining if AOD data were good enough for estimating PM at multiple levels. To correlate LIDAR-derived AOD measurements with ground-based measurements of PM_{10}, regression analysis was applied and data from AOD and MODIS were compared with the LIDAR-derived AOD for validation. The coefficient of correlation (r) between the AOD and MODIS data was higher than 0.90, showing high agreement between MODIS AOD and the ground-based LIDAR measurements of AOD.

Correlation of MODIS AOD with surface PM levels was reported by Chu et al. (2003). They reported absence of significant correlation between MODIS AOD data and the ground-level PM measurements. This happened in extreme conditions, where smoke layers were extending from the surface to a height of 5 km. An r value of 0.82 was obtained in northern Italy, thus revealing the potential for using MODIS AOD for estimating and forecasting the regional air quality. The period of the study was July 2000 to May 2001. Seasonal variations were observed. The highest AOD values were reported in summer and spring while the lowest AOD values were recorded in winter. Different procedures were applied to each individual site and each step was not sufficiently accounted for in their study. For instance, meteorological conditions like high atmospheric temperature and strong winds were compared quantitatively with

AOD levels but were not integrated statistically as factors in the analysis. The scope and level of the paper is ambitious. However, it could have been clearer had it targeted MODIS AOD for accuracy testing or for the estimation of the regional levels of aerosols.

Comparing the actual colors of the images produced by AOD with the ground-based EPA monitoring networks uncovered that both give appreciable information for the determination of the urban air quality (Engel-Cox et al., 2004). The concepts presented by Chu et al. (2003) were expanded through the incorporation of MODIS-derived AOD data with ground-based $PM_{2.5}$ and PM_{10} measurements via regression analysis. Regional differences exist in the relationship between the ground measurements and the satellite-based measurements, most probably because the AOD-$PM_{2.5}$ relationship is influenced by certain variables like weather, climate, and the local terrain (Engel-Cox et al., 2004).

Further investigation of the variations in both the seasonal and geographical correlations among the $PM_{2.5}$ levels and AOD across the USA was carried out by Zhang et al. (2009). The relationship between AOD and $PM_{2.5}$ was determined by matching the observations of the Aqua and Terra satellites with the ground measurements at 10 EPA areas for two years (2005-2006) using MODIS AOD data. Their findings revealed the effects of seasonal and geographical factors on strength of the relationship of $PM_{2.5}$ with AOD across the eastern USA. High effects were detected in fall and summer. The highest R^2 value was recorded in the South-eastern USA but the relationship between these variables in the southwest areas (region 9), which includes California, had the lowest R^2 and r values ($r = 0.26$ and $R^2 = 0.17$). It is worth highlighting that this study did not include meteorological factors.

Contrary to the large-scale investigations, Gupta and Christopher (2008b) used MODIS AOD data and $PM_{2.5}$ ground measurements taken at a particular site in Birmingham and Alabama for seven years to determine how effective the MODIS data are in air quality monitoring and develop an insight into the monthly, seasonal, and inter-annual relationships between AOD and $PM_{2.5}$. The relationship between the average, day-to-day $PM_{2.5}$ concentration and AOD was characterized by a correlation coefficient of 0.52. In the case of the hourly measurements, the value of r was 0.62.

Natunen et al. (2010) explored the relation between $PM_{2.5}$ concentration and AOD at a small regional scale in Finland using seven-year data (2000-2006) representing ground-measurements of $PM_{2.5}$ produced by four stations in Helsinki. The researchers used MODIS AOD data derived from the Aqua and Terra satellites to examine how the correlation between AOD and $PM_{2.5}$ affects the mean of the temporal $PM_{2.5}$ levels and their seasonality. They found that the time mean strengthens the association between these two variables above and beyond the hourly $PM_{2.5}$. The closest satellite overpass was paired with the hourly $PM_{2.5}$ measurements. Additional 24-hour data of hourly measurements at both sides were averaged and the resultant mean value was correlated with AOD. The mean time having the highest correlation differed between

the four sites. The best results were associated with the 19-, 15-, 5-, and 24-hr average values. Overall, the *r* values ranged from 0.57 to 0.91.

Similarly, research carried out by Tian and Chen (2010b) considered the effects of mean average time on the AOD-$PM_{2.5}$ relationship across southern Ontario in Canada. To obtain the best general correlation, day-to-day and three-hour average values of ground-level $PM_{2.5}$ measurements were compared with each other. The best *r* value obtained (0.593) was associated with the three-hour window. Tian and Chen (2010b) reported aggregated values over 3 x 3 pixels with an improved *r* value of 0.03 over the single-center pixel values using Terra and Aqua satellites.

2.8.3.6 Multivariate Models

Successive research developed quite complex empirical models that integrate environmental parameters, including geographical and meteorological data, in multiple regression analyses at urban and regional scales (Kumar et al., 2008; Li et al., 2009; Mao et al., 2012; Wallace and Kanaroglou, 2007). The height of the PBL is a key environmental factor to take into account when assessing the quality of air via satellites due to its effects on the vertical distribution of aerosols. The PBL largely develops during the day and demonstrates a strong inversion at the top. In this layer, confinement and mixing of anthropogenic aerosol is pronounced. This is vital because the aerosol extinction measures the AOD for the whole column in the atmosphere from the ground till the satellite sensor. The measured AOD value is constant irrespective of the development of the boundary layer. In a situation of high boundary layer having fragile inversion, correlation between the ground-based measurements of $PM_{2.5}$ and AOD may not be strong (Gupta and Christopher, 2008b).

The AOD-$PM_{2.5}$ relationship depends largely on the height of the PBL. This, combined with meteorological factors that affect particle formation and growth like the RH, is the reason behind inadequacy of the simple two-variable, regression analyses linking $PM_{2.5}$ and AOD for ground-level $PM_{2.5}$ estimation. Few studies have integrated the height of the PBL in multiple regression analyses (Glantz et al., 2009; Hoff and Christopher, 2012; Y. Liu et al., 2007; Tsai et al., 2011; Zifeng Wang et al., 2010).

Based on the atmospheric radiative transfer theories and the characteristics of optical aerosols, the vertical distribution of aerosol was determined and it was found that the RH possesses high effects on the relationship between PM concentrations and AOT (Zifeng Wang et al., 2010; Zhang and Li, 2015). There is an indirect relation between ground-level $PM_{2.5}$ and AOT, with a strong correlation during the day by effect of the low cloud cover, good vertical mixing, and low RH inside the atmospheric column (Al-Hamdan et al., 2009). The AOD-$PM_{2.5}$ relationship can be affected by the RH through the change it induces in the optical parameters of the aerosols (Zhang et al., 2009). The highest correlation between the $PM_{2.5}$ mass and MODIS AOT was reported

by Gupta et al. (2006) in a condition of clear cloud, a RH of 40-50%, and an atmospheric mixing altitude varying in the range of 100-200m.

One of the premier studies to apply multiple linear regression analysis (MLRA) of PM to the meteorology research is the study of Gupta et al. (2006). These researchers assessed the air quality in 26 localities in Delhi, New York, Sydney, Switzerland, and Hong Kong using one-year AOD data derived from MODIS aboard the Terra and Aqua satellites besides ground-measurements of $PM_{2.5}$. An r value of 0.96 characterized the relationship between the bin-average daily AOD mean and the ground-level $PM_{2.5}$ measurements upon inclusion of the wind speed and the boundary layer height in the model. The best results were obtained when height of the boundary layer was within the range of 100 to 200 m and the cloud cover was 25%.

Multiple regression analysis was applied by Tian and Chen (2010a) to predict the hourly concentrations of $PM_{2.5}$ in southern Ontario, Canada. They introduced an enhancement to the previous results on the relation between the ground level AOD and $PM_{2.5}$ concentrations through the addition of boundary layer and meteorological factors. The goal of this study was develop a low-cost method for supplementing the ground monitoring stations. Data for the year 2004 were used in development and validation of the model. The model predictions correlated highly with the ground-based observations ($R^2 = 0.64$).

Di Nicolantonio et al. (2009) used MODIS AOD and climate model simulation to predict the ground-level $PM_{2.5}$ concentrations in Po River valley in northern Italy. The study covered three summer months (May to July) in 2007 and three winter months (January to March) in 2008. There was a good agreement in their predictive model between the *in situ* measurements of $PM_{2.5}$, where the R^2 values were 0.68 for MODIS/Terra and 0.59 for MODIS/Aqua. The levels of $PM_{2.5}$ were found to be greater in the winter months than in the summer months but the satellite-based concentrations of $PM_{2.5}$ seemed to have underestimated the values by 20%. The likely reason(s) for the seasonal fluctuations discovered in this study were not discussed by these researchers.

Zifeng Wang et al. (2010) examined the potential role of inclusion of RH data drawn from meteorological stations and boundary layer height derived from LIDAR in improving the strength of the correlation between $PM_{2.5}$ and AOD in Beijing, China, for a period of 15 months, extending from July 2007 through October 2008. Incorporation of the boundary layer height and the RH in the model resulted in improvement of the predictive power of the model by 14%, where the R^2 value was 0.48 in the case of the two-variable model and 0.62 in the case of the multivariate model.

A method for calibration of MODIS AOD data for accurate prediction of the ground-level concentrations of $PM_{2.5}$ was suggested by Lee et al. (2011). Their study was the first to investigate the relationship between $PM_{2.5}$ concentration and AOD on a day-

to-day basis. The AOD data were retrieved using MODIS on board of the Aqua and Terra satellites. The area of the study included sections of Rhode Island and Massachusetts. It was represented by 387 grids of 10 km^2 area, each, in the ArcGIS 9.3 environment. Using MLRA, the average $PM_{2.5}$ concentration was statistically estimated and predicted. A higher correlation ($r = 0.79$, $R^2 = 0.62$) was indicated by the MLRM than that produced by a SLRM of the same data ($r = 0.5$, $R^2 = 0.26$). Consequently, it was concluded that using multivariate model for predicting surface-level $PM_{2.5}$ concentrations has advantages over using simple, two variable model and that multivariate regression modeling will be perfect for the studies of time series and cross-sectional health effects.

Mao et al. (2012) developed and tested an enhanced land use regression (LUR) model that integrates monthly AOD data from MODIS Terra. Usually, variables like land use or land cover, point-source emission, traffic information, and population are utilized by LUR models for predicting long-term intra-urban air pollutant distribution. One major drawback of this technique is the limited available information about the temporal variability. In most of the cases, this deficiency limits the representation of the spatial variability of pollutants. There was no integration of meteorological data in the model. The concentrations predicted were observed to be greater in summer/fall than in other seasons. However, they agree with past studies in the eastern USA (Chu et al., 2003). This is possibly a result of increase in temperature, high humidity, and strong insulation that lead to atmospheric ions reacting and forming aerosol products.

The seasonal and spatial variations in aerosol concentrations were studied by (Mansha and Ghauri, 2011) across Karachi, Pakistan. The research data were obtained from MODIS AOD using the Sun photometers Aqua and Terra satellites and ground-based $PM_{2.5}$ measurements for the year 2008. Higher concentrations of $PM_{2.5}$ were observed in winter through the summer season due to the low boundary layer heights in winter and an increase in the wind speed in the summer. Variations in such parameters as temperature, the RH, and wind speed were predicted by the MODIS AOD. The study found that the developed regression models had R^2 values ranging from 0.47 to 0.67.

Similarly, a study on the surface PM and MODIS AOD retrievals was carried out by Tsai et al. (2011). The study aimed at assessing the capacity of satellite for air quality monitoring in Taiwan. The independent variables included the boundary layer heights and the RH. Validation of the MODIS AOD retrievals (the R^2 values were 0.82 for Terra and 0.69 for Aqua) was performed using two Sun photometers. The AOD-$PM_{2.5}$ relationship was found to be crucially dependent on height of the boundary layer as evidenced by the higher R^2 values (0.86 vs. 0.77).

Observations of PM are scanty around the globe because of problems in the related sampling methods and high cost of sampling. Large-scale spatial coverage is provided by satellite observation at a low cost and, hence, it has been utilized in several studies in the last 10 years for analyzing the relation between AOD and the surface $PM_{2.5}$ concentration in diverse parts of the world (Chen et al., 2016).

Many former studies were based on MODIS AOT products. The daily products of the MODIS sensor have been broadly employed to evaluate the relation between air quality and the $PM_{2.5}$ concentration. In recent years, water and air pollution in Malaysia, especially with industrial wastes, has become serious. Air pollution is the main factor that has been impacting the human health, forest species, agricultural crops, and the varied ecosystems. Monitoring studies of the ambient air quality and the generated data indicate that the concentrations of some air pollutants are increasing with time in several big cities and are not all the time within the acceptable limits according to the national ambient air quality standards. However, data on air pollution and case studies are much limited in Malaysia (Afroz et al., 2003). In recent times, satellite RS is broadly used to monitor quality. Because of lack of such studies in this field in Malaysia, the current study fills a gap in the literature by using data obtained from the MODIS sensor to estimate the $PM_{2.5}$ concentrations in Peninsular Malaysia.

In tropical countries like Malaysia, the cloud pattern can be highly variable as a result of the high humidity and the unpredictable weather, particularly during the monsoons. The clouds have a high effect on the MODIS AOT retrievals. For this reason, this study investigated the effects of three sizes of MODIS pixels (1×1, 3×3, and 5×5 pixels) on estimates of ground level mass concentrations of $PM_{2.5}$.

The MODIS sensor on board the Terra satellite overpasses the study area; Peninsular Malaysia, once per day at approximately 10:30 a.m local time. Instead of the $PM_{2.5}$ concentrations that can be obtained from 24-hour samplers in the ground stations, satellite data were used and statistical analysis was conducted in this study by regression modeling to explore the relations among AOT (once a day), meteorological parameters (24 records per day), and the dependent variables (hourly and daily $PM_{2.5}$ concentrations). The hourly data were synchronized with the satellite passage and the daily data are the 24-hr average $PM_{2.5}$ concentrations. The R^2 value was employed as a measure of the predictive powers of the regression models.

Peninsular Malaysia is one of the zones in the world that are most affected by the haze episodes in east Asia that are frequently triggered by biomass burning in Sumatra, Indonesia. For this reason, satellite RS (MODIS) can provide aerosol data at sufficient spatial (10 km) and temporal (daily) coverage for this region to allow for studying phenomena such as patterns of aerosols and their impact on air quality.

The foregoing review of previous studies brings to surface that MODIS data were used in different climates and geographical locations. Furthermore, they were applied at global, regional, and local scales (Table 2.2). This study follows a physical approach that takes advantage of satellite coverage at a local scale to estimate the ambient $PM_{2.5}$ mass concentration. Overall, results of previous studies pinpoint high correlations between MODIS AOT and ground $PM_{2.5}$ measurements in different periods. In the light of these results, the researcher posits that there is a positive correlation between satellite and ground measurements of $PM_{2.5}$ in Peninsular Malaysia.

Table 2.2 : Estimation of $PM_{2.5}$ levels in different countries using data drawn from the MODIS sensor and different models

Place	Study period	Model [1]	R^2 Value [2]	Reference
USA	2002	LC	0.960	(Wang, 2003)
USA	2002	LC	0.185	(Engel-Cox et al., 2004)
USA	2004	MLR	0.423	(Engel-Cox et al., 2006)
Europe	2003	LC	0.360	(Koelemeijer et al., 2006)
India	2003	MLR	0.700	(Kumar et al., 2007)
Canada	2005	MLR	0.760	(Wallace and Kanaroglou, 2007)
USA	2004	GAM	0.360	(Paciorek et al., 2008)
USA	2004-2006	MLR	0.365	(Gupta and Christopher, 2009)
Global	2001-2006	CTM	0.593-0.689	(Van Donkelaar et al., 2010)
China	2007-2008	LC	0.470	(Zifeng Wang et al., 2010)
USA	2003	MEM	0.970	(Lee et al., 2011)
USA	2000-2008	MEM	0.850	(Kloog et al., 2012)
USA	2003	LC	0.470	(Chudnovsky et al., 2013)
USA	2003-2007	MLR	0.760	(Saunders et al., 2014)
Global	2012	MLR	0.850	(Lai et al., 2014)
China	2012-2013	GWR	0.710	(Ma et al., 2014)
China	2012-2013	GWR	0.738	(Song et al., 2014)
Mexico	2004-2014	MEM	0.724	(Just et al., 2015)
USA	2007-2011	MEM	0.770	(Mihye Lee et al., 2016)
China	2013	GAM	0.691	(You et al., 2015)
China	2013	MLR	0.462	(Zhang and Li, 2015)
USA	206-2012	MEM	0.666	(H. J Lee et al., 2016)
China	2000-2014	LC	0.672	(C. Lin et al., 2016)
USA	2003-2008	MEM	0.870	(Shi et al., 2015)
Global	1998-2014	GWR	0.810	(Van Donkelaar et al., 2016)
China	2014	GWR	0.810	(W. You et al., 2016b)
China	2013	GWR	0.750	(Zou et al., 2016)

(1) LC: Linear Correlation; MLR: Multiple Linear Regression; MEM: Mixed-Effect Model; CTM: Chemical Transport Model; GWR: Geographically-Weighted Regression.
(2) Model performance was assessed in terms of the corresponding R2 value.

2.9 Chapter Summary

Measurements of $PM_{2.5}$ can be acquired from the air quality monitoring networks and programs that are operating in many countries. However, the ground $PM_{2.5}$ measurements only represent the $PM_{2.5}$ levels in a limited area around the monitoring sites. Under these circumstances, the satellite measurements can be used to acquire $PM_{2.5}$ data with greater spatial coverage. A strong relation between $PM_{2.5}$ measurements and the satellite AOT has been reported by previous research. Numerous satellites and algorithms have been employed for AOT retrieval. In previous studies, the MODIS was the satellite most often used. It provides the users with AOT data that are generated using the DP and DT vegetation algorithms to users directly so that they will not have to retrieve AOT from the RS data by themselves. In 2014, NASA released the MODIS Collection 6 product. Usually, ground-level AOT measurements, such as the AOT provided by AERONET, can be used to calibrate the

satellite AOT data. The current models for AOT-$PM_{2.5}$ can be categorized into observation-based and simulation-based methods. In the observation-based models, linear regression modeling is most often used and spatial regression can consider the spatial autocorrelation. For improving the AOT-$PM_{2.5}$ estimation model, other factors (e.g., meteorological factors) are also taken into consideration in more complex models. The $PM_{2.5}$ distribution estimated from satellite data can contribute to pronounced enhancement of information about the concentrations of $PM_{2.5}$ at the ground surface level in a wide spatial range. To sum up, the previous studies summarized in Table 2.2 used MODIS data for global, regional, and local areas to estimate the levels of $PM_{2.5}$.

CHAPTER 3

METHODOLOGY

3.1 Overview of the Methodology

This chapter defines and describes the study area as well as the satellite and ground AOT data which were used to estimate the ground $PM_{2.5}$ concentrations. This is followed by description of the instruments, the spectroradiometer MODIS and Sun photometer retrieval of AERONET station, and the data collection. In addition, this chapter provides explanation of the image and data analysis and the proposed methods and approaches. Figure 3.1 displays the methodological framework.

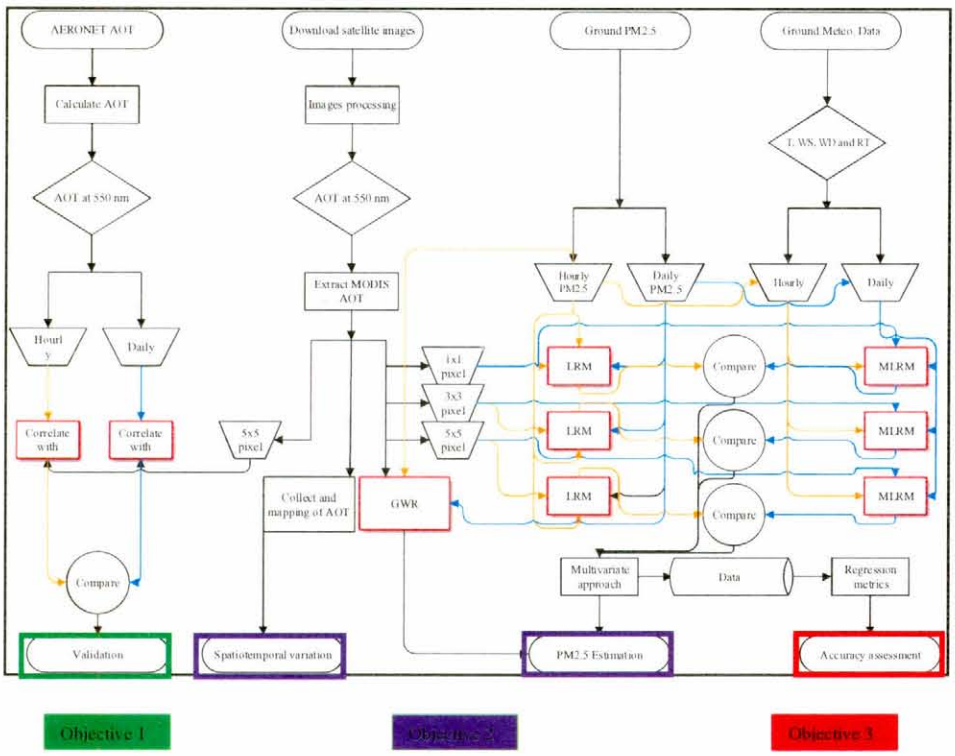


Figure 3.1 : Methodological Framework

The ground monitoring data are considered as accurate measurements. However, they only represent the concentrations in a relatively small area. Furthermore, the number of ground stations is usually limited and the distributions of those stations are unbalanced. This makes continual spatial monitoring difficult. Apart from resolution

and spatial coverage, temporal coverage of the ground-level PM monitoring stations varies highly as it depends on functionality and the operation period of the instrument. Moreover, construction and maintenance of the ground stations are time-consuming, costly, and labor-consuming.

The aforementioned drawbacks of ground monitoring of PM prompted continuous exploration of alternative means of PM estimation with RS techniques, which has several advantages. First, the image derived from the satellite can provide general and complete information on air quality anywhere in the world. Second, owing to that the satellite provides a chance for acquiring information about the global air quality, it enables discovery of the sources of the urban air pollutants and even global transportation of those pollutants. Additionally, satellite makes it less costly to monitor the air quality for the developing countries that suffer from severe air pollution but lack ground stations. Previous studies using different models reported a positive relationship between ground-level $PM_{2.5}$ concentrations and the satellite-derived AOT.

There are limited data on levels of $PM_{2.5}$ in Malaysia. The WHO confirmed that PM pollution has health effects, even at very low concentrations. There is significant relationship between exposure to high concentrations of the small PM (PM_{10} and $PM_{2.5}$) and morbidity and mortality, both daily and over time. When the concentrations of the small and fine PM are reduced, mortality drops. This allows the policy makers to project the improvements in the public health if particulate air pollution is reduced. The concentrations of PM are often the highest in the urban areas of the low-, and middle-income countries.

Satellite RS provides data on AOT, which is a measure of the light extinction (that is, light scattering and absorption) by the atmospheric aerosols. The values of AOT reflect abundance of particles in the atmospheric column. Accordingly, they have been employed in statistical models to predict the ground $PM_{2.5}$ concentrations. In this context, the data of the MODIS sensor are the AOT products most broadly employed worldwide for $PM_{2.5}$ estimation. Hence, Collection 6 (C6) of the MODIS sensor aboard Terra satellite was used in this study. These data, which have a resolution of 10 km, were obtained from the website of the Atmospheric Archive and Distribution System (http://ladsweb.nascom.nasa.gov/ data/search.html). Due to the lack of countrywide monitoring networks of $PM_{2.5}$ in Malaysia, this study uses $PM_{2.5}$ data for only one year (2013). These data (Table 3.1) were obtained from the Department of Environmental (DoE), Malaysia. The MODIS AOT data and the *in situ* ground $PM_{2.5}$ observations during the year 2013 were employed in modeling.

Aerosol Optical Thickness data drawn from AERONET were employed in the present study to calibrate the MODIS data. The AERONET data are broadly employed for validating the satellite-derived AOT because of their high accuracy. They are available online in near real time. This study used Level 2.0 AERONET data since this data level was generated by the cloud-screened procedure. The AERONET site (called USM_Penang) is the only station in Peninsular Malaysia. It started its operation at the

end of the year 2011. For additional data for validation, the sources of data that are necessary for this study were extended to include the nearest AERONET stations (Figure 3.2) as will be discussed later in this chapter.

Table 3.1 : Description of the ground and satellite data employed in the current study

Parameter	Frequency	Source	Year
$PM_{2.5}(\mu m/m^3)$	Hourly (synchronized with satellite passing) Daily (average)	Monitoring stations of the Department of Environment (DoE), Malaysia	2013
Meteorological parameters	Hourly (synchronized with satellite passing) Daily (average)		
Satellite AOT	Once a day (Image)	MODIS Terra satellite	2012 -2015
AERONET AOT	Hourly average ($\pm$ 30 min) synchronized with satellite passing Daily average	AERONET station (Figure 3.1)	2012 -2015

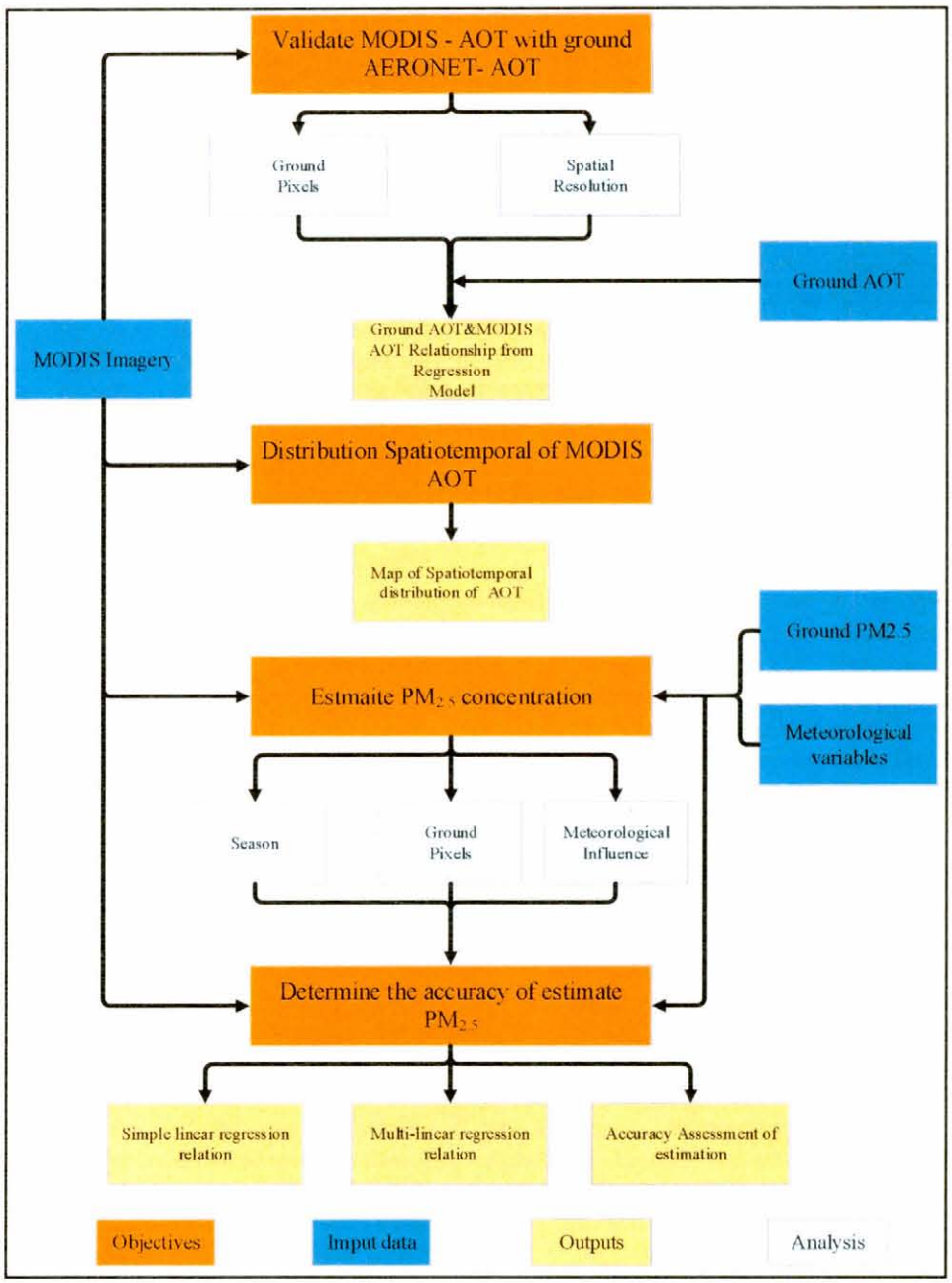


Figure 3.2 : Flow Chart of the Data Processing

3.2 The Area of Interest

In this study, the researcher extends the work to include responses of ground instruments (Sun photometer) to AOT, which are especially important for MODIS instrument validation.

3.2.1 Regional Extent

Data drawn from AERONET, a global Sun photometer network, are commonly employed as ground truth data in validation of the satellite retrievals. There are hundreds of sites for this network around the world. Besides the only station in the study area (USM_Penang; Peninsular Malaysia), one other AERONET station existed in Kuching, east of Malaysia. However, there are three other AERONET stations within the region that are located in neighbouring countries. These are Songkhla_Met_Sta in Thailand, Pontianak in Indonesia, and Singapore; in Singapore. Taken together, these stations lie in an area that has the coordinates of 10° S - 22° N latitude and 95° E - 140° E longitude (Figure 3.3).

Direct Sun measurements are taken by AERONET automatically every 15 min by Sun, and sky-scanning CIMEL Sun photometers during the daylight hours. Standard processing does automatically generate AOT from the direct transmission, calibrated using Langley method in bands centred at approximately 340, 380, 440, 500, 675, 870, and 1,020 nm (Holben et al., 1998).

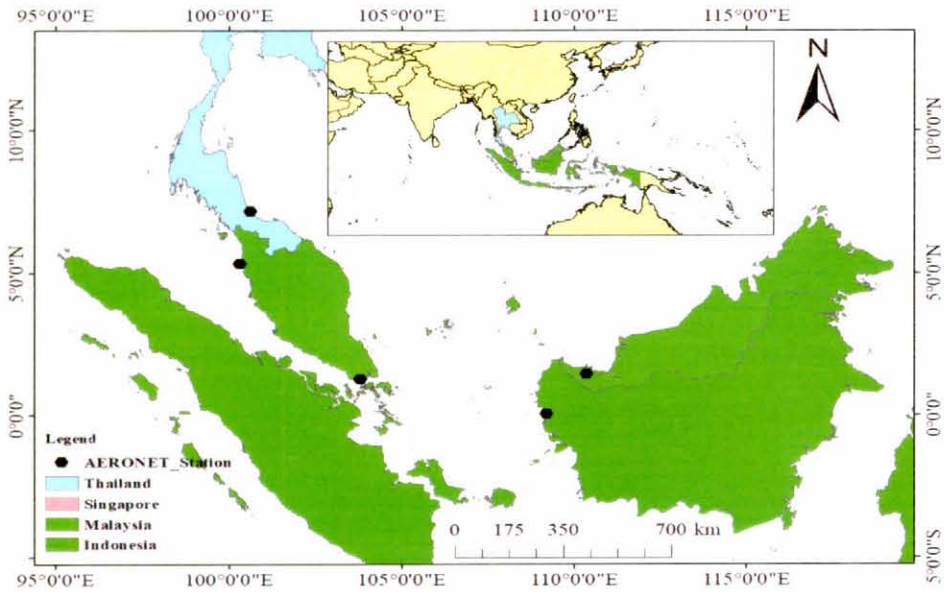


Figure 3.3 : Locations of five AERONET stations in the region of the study area

3.2.2 The Study Area

The study area is Peninsular Malaysia. It has an area of 131,598 km^2 and is located between the latitudes of 1° and 7° N and the longitudes of 99° and 105° E, north of Singapore and east south of Thailand (Figure 3.4).

This area belongs to the humid tropical climate zone, where the weather is humid and hot over the whole year, with a period of two to three months of dry weather that is experienced in most of the parts of the study area between June and November. Its weather is affected by the monsoons, that is, the northeast monsoon, which prevails from October of a year to March of the next year, and the southwest monsoon, which extends from April to September. So, Peninsular Malaysia is characterized by humid tropical climate all over the year; the weather is humid and warm, and the temperature ranges from 20 °C to 32 °C. The mountainous topography and complex land-sea interactions in the study area affect the tropical climate greatly. April to May and July to August are the periods with the highest average temperatures. The lowest average temperatures are encountered from November to January. The average monthly humidity ranges from 70% to 90% and varies by month and location (Tan et al., 2017).

Peninsular Malaysia extends over nearly 40% of the land area of Malaysia and host about 80.0% of its population and economic activities (Kanniah et al., 2014). The recent, rapid economic growth and development, especially in the industrial and transportation domains, has initiated varying environmental problems in certain large cities in Peninsular Malaysia (Kanniah et al., 2014; Rajab et al., 2013; Semire et al., 2012; Suhaila et al., 2010).

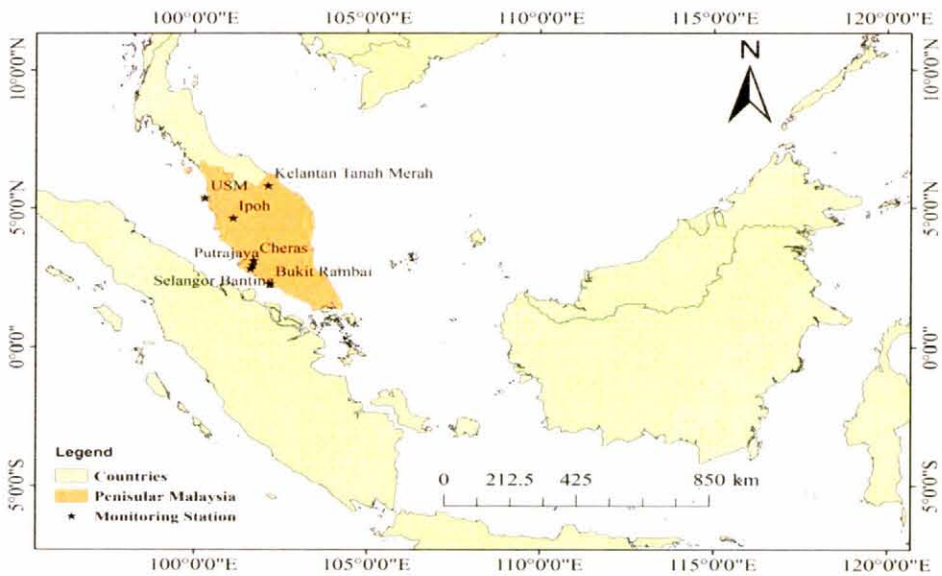


Figure 3.4 : Locations of Seven Air Quality Monitoring Stations in Peninsular Malaysia

3.3 The Moderate Resolution Imaging Spectroradiometer (MODIS)

The MODIS instruments are on board of the Earth Observation System (EOS) Terra satellite since 1999 and aboard Aqua satellite since 2002 at an orbit height of 705 km. These two satellites are polar-orbiting. However, while Terra satellite orbits at descending node (southward) and crosses the equator at nearly 10:30 a.m local time, Aqua satellite overpasses the equator at 13:30 p.m local time. Because of a $\pm 55^{\circ}$ cross-track scan, the MODIS swath is almost 2,330 km wide. Therefore, it can provide global coverage every one to two days.

3.3.1 Instrument Distribution

MODIS is a passive RS instrument that explores various components of the Earth and of its atmosphere. It allows for determination of the temporal and spatial distributions of the global aerosol and cloud fields, detection of forest fires, and monitoring of the vegetation index. It obtains data using 36 spectral bands that range from 0.4 μm to 14.4 μm. Depending on the band, three horizontal resolutions are possible; 250- meter, 500- meter, and 1,000-meter resolutions.

In the current study, the Terra MODIS AOT data were Level 2 (L2) data (MOD04 - Collection 6) with a full resolution of 500 m. The MODIS imagery and the AOT data are relevant for ambient air quality monitoring. The ability to visualize regional-scale events with the L2 data can be utilized effectively in gaining further understanding of the extent of the regional haze events. This derives importance from the need for in-depth understanding of the effects of large events on the pollutant levels at the local scale.

3.3.2 Aerosol Retrieval over Land

For the region of interest, the aerosol land algorithm was used. This algorithm is based on three channels; two visible channels (470 nm and 660 nm) and the 2.12 μm (NIR) channel. The NIR channel contains information on large particles and is referred to as the coarse model. This algorithm analyzes data of a nominal region of 10 km × 10 km area, which is equivalent to 20 × 20 pixels at a resolution of 500 m. All the pixels that contain clouds are omitted. For meteorological applications, the AOT readings taken at the wavelength of 550 nm are usually used as a standard parameter. Thus, it is computed by interpolation from the AOT readings taken at the wavelengths of 660 nm and 470 nm according to the Ångström law.

3.3.3 The Research Data

For validation purposes, this study used AOD data drawn from the NASA Terra MODIS sensor (Collection 6; Level 2; 10km resolution), which is available for Malaysia for the years 2012-2015. MODIS observes a swath that is approximately

2,330 km wide (Figure 3.5). The same MODIS images for the year 2013 were used for $PM_{2.5}$ estimation. It is worth mentioning that the MODIS aerosol products are available on the website of NASA (https://ladsweb.modaps.eosdis.nasa.gov). The values of AOT were extracted at the wavelength of 550 nm (MODIS parameter name: Optical_Depth_Land_And_Ocean). Only the highest quality AOT (QAC flag = 3) has been included over land for this parameter. The MODIS was chosen as it covers Peninsular Malaysia and due to its high frequency, temporal resolution, and wide swath width. It provides bulky data overall to correlate with the data of the ground monitoring stations and the satellite-derived AOT values.

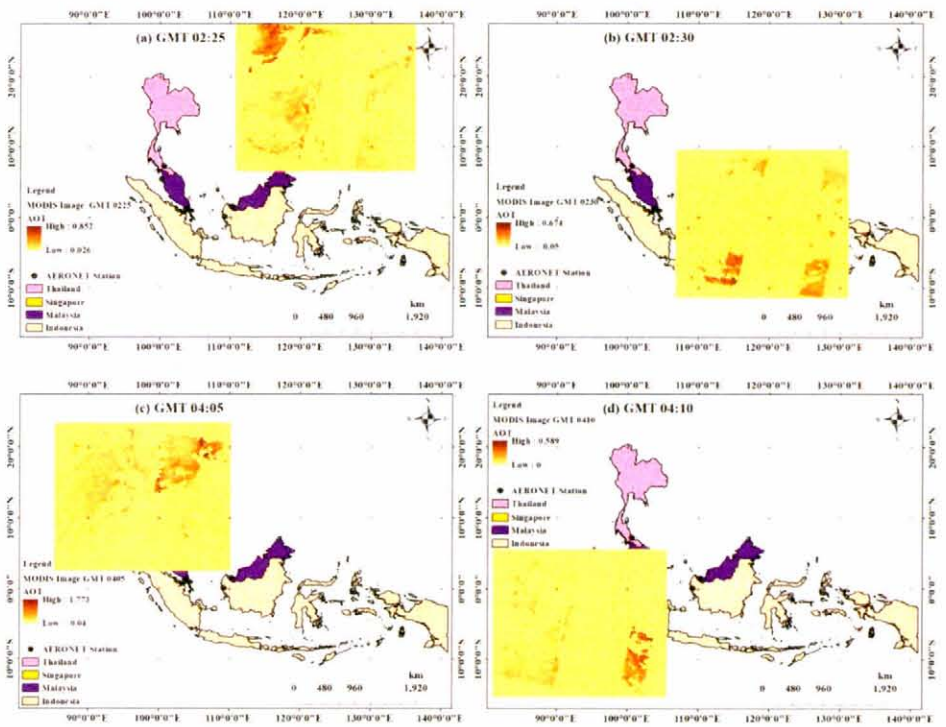


Figure 3.5 : Terra MODIS AOT Image for the Region on 1 November 2012, (a) 02:25 GMT, (b) 02:30 GMT, (c) 04:05 GMT, (d) 04:10 GMT

The majority of the previous studies only used the MODIS AOT to derive surface $PM_{2.5}$ concentrations with a spatial resolution of 10 km (1 x 1 pixel ~ 10 km^2). A serious problem inherent in the satellite data is lack of consistency in their availability. In aerosol investigations, for instance, the cloudy days are usually 'no value' days. Actually, this is one of the factors that make the regulators avoid using satellite data to help in monitoring the air quality in the highly cloudy countries like Malaysia. For this reason, this study employed three MODIS pixel size array (1x1, 3x3, and 5 x 5 pixels) in models to estimate ground-level concentrations of $PM_{2.5}$ in Peninsular Malaysia to determine the pixel group that produces the most accurate quantification of the relationship between the ground levels of $PM_{2.5}$ concentration and MODIS

AOT. The reason behind creating squares around the $PM_{2.5}$ monitors which perform point measurements and not around the AOT pixel centre is that the missing retrievals which are caused by the cloud cover, surface reflectivity, and satellite orbit patterns can improve the quantity and quality of the data.

3.3.4 Extraction of AOT

The image-processing software ENvironment for Visualizing Images (ENVI), which is written in Interactive data language (IDL), is one of the most broadly employed image-processing and analysis software. It is also used actively in integrated vector GIS analysis. By using the user function capability of ENVI, algorithms written in the IDL can be integrated with the menus of ENVI (interface). This software provides a library of procedures and programming tools for the user functions that enable for, and facilitate, handling the input and output, plotting, preparing reports, and managing files. Accordingly, the ENVI handles all the pre-, and post-processing analysis steps, including file management after implementation of the user function. It can open and process most MODIS HDF files. In this study, the researcher employed the ENVI 5.3 Classic with the extension MODIS Conversion Tool Kit to import and convert projection of the daily MODIS scenes and products.

Determination of the AOT from satellite image data can be practiced as a means of evaluating the air pollution in any area of interest. This study followed an exact pixel-to-point match procedure to extract AOT values, where each point represents a ground-level station. The number of stations used in this study to validate the MODIS AOT data (2012 -2015; 1,461 days for each station) was four while the number of stations used to estimate the $PM_{2.5}$ levels was seven (2013; 365 days for each station). The stations were overlaid on the MODIS images to derive the AOT values for each pixel, where a monitoring station is defined by its latitude and longitude. In this study, the average 5 x 5 pixel size group was employed in validation of the MODIS AOT while different arry of pixel were used for estimation of the daily average and hourly $PM_{2.5}$ concentrations.

Figure 3.6 is a plot of an example on outcomes of spatial distribution of AOT derived from MODIS at the wavelength of 550 nm on 7 May 2013 (11:50 a.m local time) at a monitoring station in Ipoh. The considered pixel groups include one pixel and the average of 3 x 3 and 5 x 5 pixels. The pixel inspector tool in ArcGIS tool was used to view the MODIS pixel values (AOT) of a raster dataset. All the values within the square of different pixels in the ground site that are marked were extracted.

In this study, the MODIS Level-2 AOT is used with the spatial resolution 10×10 km^2 at nadir. The spatiotemporal subset approach for the estimate is used here, comparing a spatial subset from the satellite data (1x1, 3x3 and 5 x 5 pixels). It justifies the size of these subsets based upon an estimate of average aerosol front speed, the requirement to obtain a statistically-significant sample size, and the observation that larger window sizes could introduce errors from cloudy pixels and changing topography. Moreover,

hourly $PM_{2.5}$ estimation can help improve our understanding of how the column AOT and surface $PM_{2.5}$ vary during the day for practical air quality applications. Based on known daily guideline of standard of the $PM_{2.5}$ (ex. $PM_{2.5}$ mass concentration daily subject to World Health Organization (WHO-2006) and the United States Environmental Protection Agency (US EPA-2013) guideline, the $PM_{2.5}$ standard 25 and 35 $\mu g/m^3$ 24-hour mean respectively.), the 24-hours mean data was needed to estimate $PM_{2.5}$ pollution in the study area

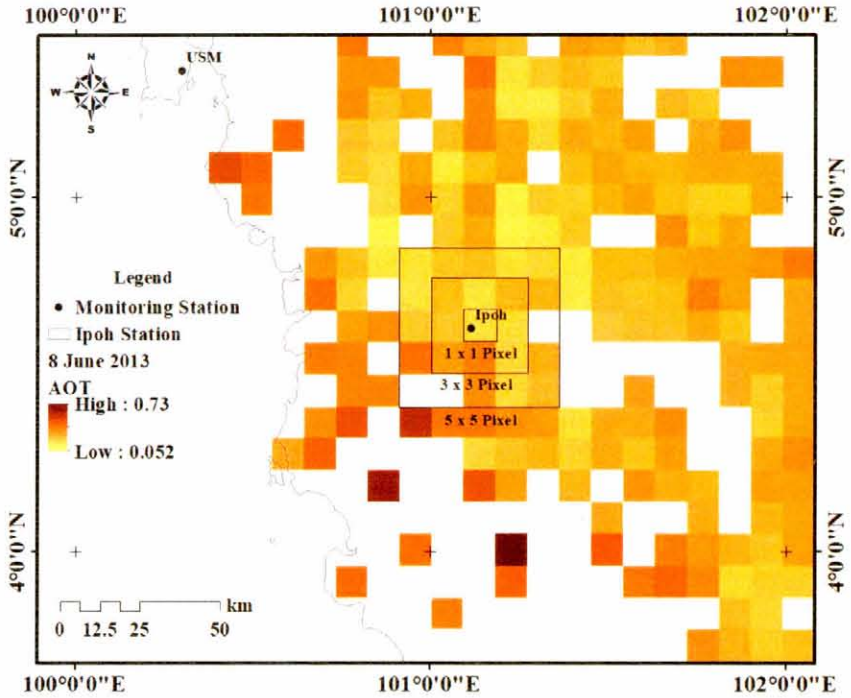


Figure 3.6 : Spatial Distribution of the AOT Derived from MODIS at the Wavelength of 550 nm on 8 June 2013 (11:50 am Local Time) in Ipoh (Malaysia)

3.4 Aerosol Robotic Network (AERONET) Sunphotometer

The sun photometer used in this study was designed and manufactured by the AERONET project, which is a group of ground-based RS aerosol networks launched by NASA and PHOTONS (PHOtométrie pour le Traitement Opérationnel de Normalisation Satellitaire; University of Lille 1). It was expanded greatly by networks and collaborators from national agencies, individual scientists, universities, institutes, and partners. As Figure 3.3 shows, the AERONET stations that are the closest to the study area are USM_Penang and Kuching, which are both located in Malaysia; Songkhla_Met_Sta, which is placed in Thailand; Pontianak, which is located in Indonesia; and Singapore, which is placed in Singapore.

3.4.1 Instrument

The AERONET Sunphotometer measures attenuation of the solar radiation in the atmosphere by the aerosols in the cloud-free condition and provides direct, total, columnar AOT measurements at the wavelengths of 0.34-2.1 mm. An example of AERONET AOD assessments for Penang at eight wavelengths over one day (17 March 2013) and the monthly trend are displayed in Figure 3.7. The data were taken in 2013 at USM_Penang station. The AOT values retrieved from the Sun photometer are the most accurate RS aerosol data. They were, accordingly, employed in this study to verify the satellite RS AOT data. Most Sun photometers take measurements at several wavelengths every 15 min during the day, from which AOT values are drawn with a precision in the order of 0.01– 0.02.

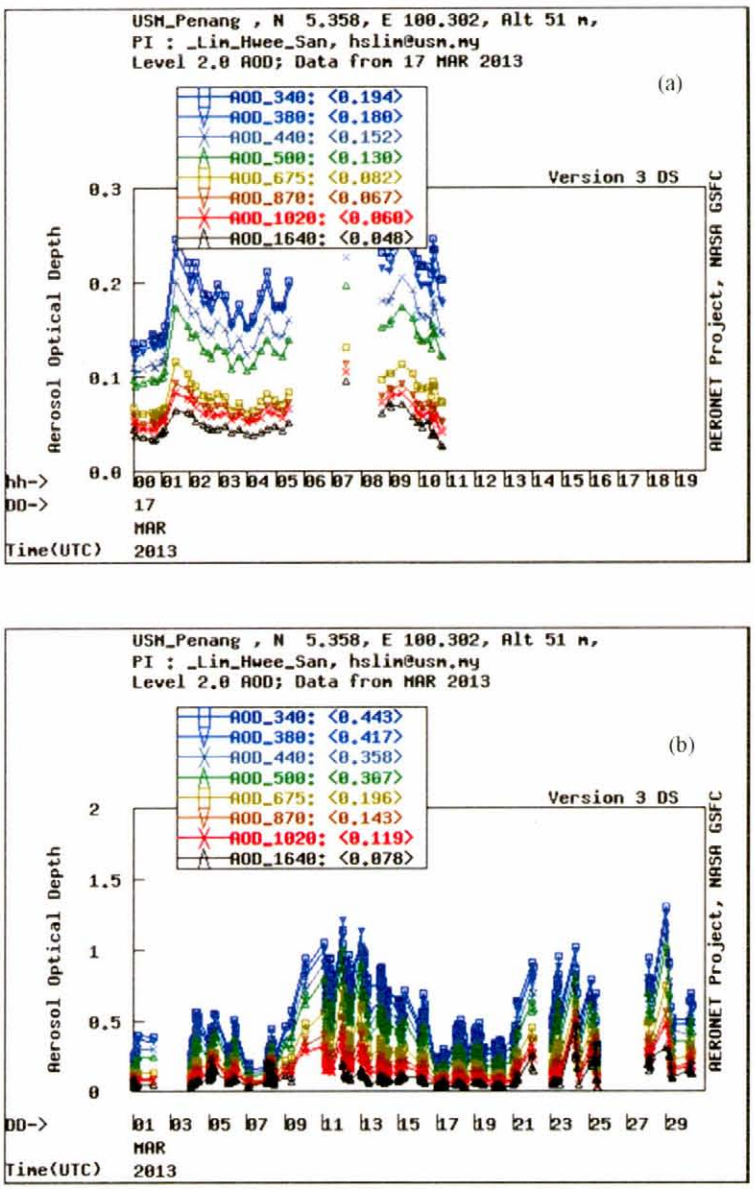


Figure 3.7 : Estimates of AOD for Penang (Malaysia) Derived from AERONET at Eight Wavelengths on 17 March 2013: (a) Daily Estimates and (b) the Monthly Trend (Source AERONET Homepage)

3.4.2 The Research Data

The Sun photometer is used to measure the spectral Sun irradiance and the sky radiances, from which the AOT can be estimated at the wavelength of 550 nm. The AOT data were processed into three levels of quality: Level 1.0 (unscreened data), Level 1.5 (cloud-screened data), and Level 2.0 (cloud-screened and quality-assured data). The data obtained from the stations were for four years (2012-2015). Ground AOT measurements at the wavelength of 550 nm were obtained from AERONET sites in the region and used to validate the MODIS AOT estimates at the same wavelength. The Sunphotometer AOTs were averaged within ± 30 min from satellite passage. The spatially-averaged MODIS AOTs were compared with the temporally-averaged Sunphotometer estimates. Level 2.0 AERONET AOT data were the most reliable data to use in validation of the MODIS aerosol product.

The retrieval wavelength of 550 nm characteristic of the MODIS AOT retrievals does not exist within the AERONET wavelengths (340, 380, 440, 500, 675, 870, 1,020, and 1,648 nm) as shown in Figure 3.7. To obtain AOT estimates from the AERONET data of the MODIS sensor at the wavelength of 550 nm, the following equation should be used:

$$\tau_{550} = Exp\left[ln\tau_{500} - ln(500nm / 550nm)\alpha\right] \tag{3.1}$$

Where τ_{550} is the AERONET AOT value at the wavelength of 550 nm and α is a standard Ångström exponent.

The data employed in this study were cloud-screened and quality-assured Level 2 data of Version 2 direct Sun algorithm, which are available at the website of AERONET (http://aeronet.gsfc.nasa.gov/new_web/index.html). Bearing in mind that AERONET does not take measurements at the wavelength of 550 nm, the data were interpolated to the wavelength of 550 nm by using the standard Ångström exponent, α, as defined in Equation (3.2):

$$\alpha = \frac{ln(\tau_1 / \tau_2)}{ln(\lambda_1 / \lambda_2)} \tag{3.2}$$

where τ_1 *and* σ_2 are the AOT at the wavelengths λ_1 (500 nm) and λ_2 (675 nm), respectively. These two wavelengths are the nearest available pair of bounding wavelengths from AERONET. Accordingly, Equation 2 assumes the form:

$$\alpha = \frac{ln(\tau_{500} / \tau_{675})}{ln(\lambda_{675} / \lambda_{500})}$$

In view of knowledge of the aerosol transport rates and the empirical tests performed, the AOT retrievals were obtained in this study in square boxe 50 km x 50 km (MODIS pixel) centred over AERONET sites using the validation method based on the correlation between the AERONET AOT and the MODIS AOT data. This selection scheme is limited by the fact that most of the validation sites are located on land. This suggests that the satellite pixels do not match the measurement site exactly. One underlying presumption is that the satellite value, calculated for a substantial number of neighboring pixels that are very close to the site, is representative of the area and, thus, can be compared with the field value.

3.5 The $PM_{2.5}$ and Meteorological Data

The PM is a mixture of liquid droplets and solid particles that are suspended in the air. They are often classified as coarse PM (PM_{10}; $2.5 < d < 10$ μm) and fine PM ($PM_{2.5}$; $d < 2.5$ μm), where d is the aerodynamic diameter of the particle. The present study is mainly concerned with the $PM_{2.5}$. Those particulates can originate from varying sources like vehicle emissions, dust, industrial emissions, and agricultural and forest fires. In terms of $PM_{2.5}$, the air quality is ever degrading all over the world because of the increasing pressure of urbanization that has serious consequences on the climate, visibility, human health, and the economy. Generally, there are limited ground measurements, which is a limitation that poses a challenge for air pollution studies and monitoring. With launch of the Aqua and Terra polar-orbiting satellites, there has been growing emphasis on use of satellite data for studying $PM_{2.5}$ concentrations in an effort to overcome some of the obstacles relating to unavailability of ground measurements.

The atmospheric variables were hourly and daily average $PM_{2.5}$ mass concentrations (in the unit of $\mu g/m^3$). The meteorological variables were the hourly and daily average RH (%), T (°C), WD (°), and WS (m/s). All data were for the period January to December, 2013. The $PM_{2.5}$ and meteorological data were acquired from seven CAQMSs (Table 3.2), which are Banting, Bukit Rambai, Cheras, Ipoh, Kelantan Tanah Merah, Putrajaya, and USM stations (Figure 3.4). These CAQMSs are operated by Alam Sekitar Malaysia Sdn. Bhd (ASMA), which is the subsidiary company of the DoE, Malaysia. The instruments used in all stations were fully automated. They provided hourly measurements of PM and meteorological data.

Table 3.2 : Coordinates of the air quality monitoring stations in Malaysia

Code of station	Site of station	Name	Latitude	Longitude
CA0006	Sek. Men. Keb. Bukit Rambai	Bukit Rambai	2.265400	102.175900
CA0054	Sek.Men.Keb.Seri Permaisuri, Cheras	Cheras	3.106267	101.717867
CA0059	SMK. Tanah Merah	Kelantan Tanah Merah	5.811183	102.133333
CA0008	Sek. Men. Jalan Tasek, Ipoh	Ipoh	4.629683	101.116067
CA0053	Sek. Keb. Putrajaya 8(2), Jln P8/E2, Presint 8, Putrajaya	Putrajaya	2.931917	101.681817
CA0060	Kolej MARA, Banting	Banting	2.816683	101.623017
CA0038	Universiti Sains Malaysia, Pulau Pinang	USM	5.358800	100.297733

This study goes beyond establishment of the AOT-$PM_{2.5}$ relationship to assessment of air quality in terms of the levels of PM. It also addresses factors that affect meteorological parameters and which govern this relationship. One of the main factors that impact quantification of the AOT-$PM_{2.5}$ relation is the lack of information on vertical distribution of the aerosols and height of the PBL. However, the use of meteorological factors only is not sufficient to explain the $PM_{2.5}$ concentrations because the human activities too produce a substantial amount of $PM_{2.5}$. Road traffic, aviation and marine transportation, household activities (cooling and heating), construction, and industry (especially the heavy industry) are all sources of emission of $PM_{2.5}$. When data on these factors are missing, accurate estimation of the relationship between ground $PM_{2.5}$ measurements and MODIS AOT data is negatively impacted. Table 3.3 lists the months for which $PM_{2.5}$ and meteorological data were available in the study period.

Table 3.3 : The months of the study year (2013) for which $PM_{2.5}$ and meteorological data are available at each air quality monitoring station

Station	Month											
Bukit Rambai	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.			
Cheras	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Kelantan Tanah Merah	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.			
Ipoh	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.		
Putrajaya	Jan.			Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Banting	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.			
USM	Jan.	Feb.	Mar.	Apr.	May	Jun.		Aug.	Sep.	Oct.	Nov.	Dec.

3.6 Statistical Models

Statistical analysis in this study aimed at exploring the statistical relations of satellite-retrieved AOT (550 nm) with daily average and hourly average $PM_{2.5}$ concentrations. The hourly average was estimated ± 30 min, depending on the time of satellite (AERONET AOT) passage. Linear regression modeling was performed using the XLSTAT statistical software. In like manner, the daily and hourly averages of the ground mass concentrations of $PM_{2.5}$ ($\mu g/m^3$) were correlated, each separately, with MODIS AOT (Independent variable) to predict the $PM_{2.5}$ concentration (dependent variable). As well, multiple linear regression analysis was performed to analyze the effects of meteorological variables (RH, T, WS, and WD) on the concentration of $PM_{2.5}$.

3.6.1 Multiple Linear Regression Modeling

Multiple linear regression modeling has been used since 2005 to estimate the $PM_{2.5}$ concentrations from satellite AOT data. One of the advantages of multiple linear regression analysis (MLRA) is that the confounding bias can be avoided by inclusion of appropriate covariates in the model. In the MLRMs in the current study, the $PM_{2.5}$ concentration measured at the ground level served as the dependent variable while AOT was the independent variable. Other factors were incorporated in the model as covariates to determine whether or not they can improve the predictive powers of the regression models (Figure 3.8) predictive of the $PM_{2.5}$ concentration. The other factors than AOT that were included in the MLRMs were T, WD, WS, height of the boundary layer, and aerosol type.

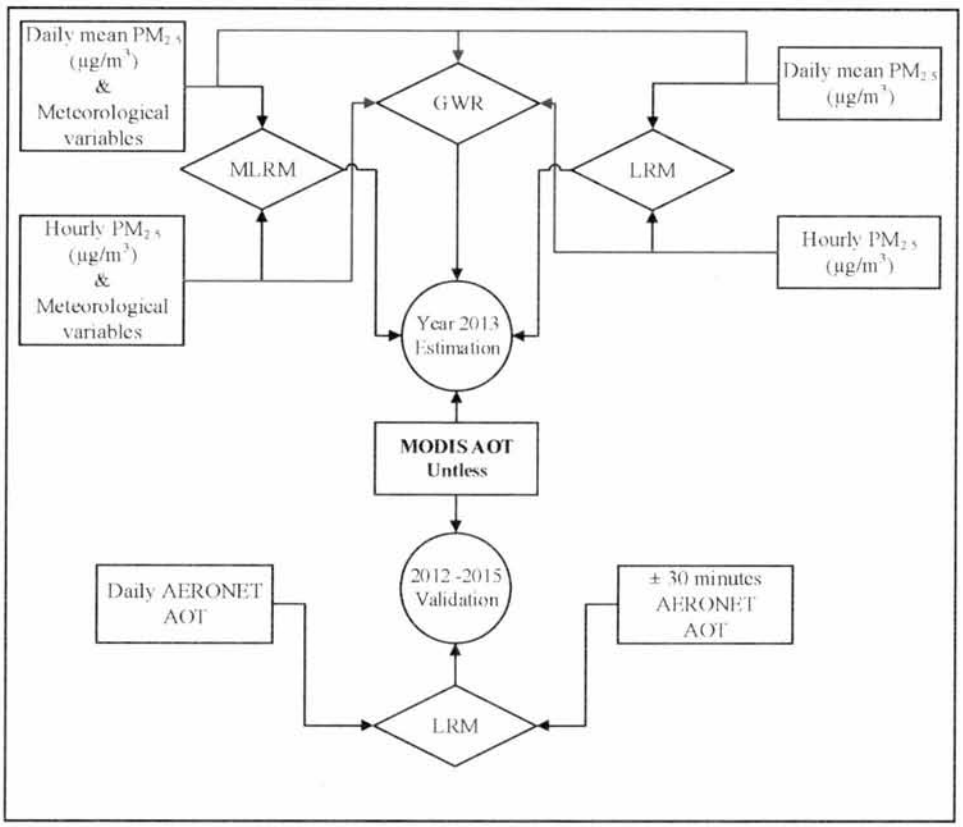


Figure 3.8 : Flow of Steps for Building the Geographically-Weighted Regression Model (GWRM), Linear Regression Model (LRM), and Multiple Linear Regression Model (MLRM) for Estimation of the Concentrations of $PM_{2.5}$ and Validation of the MODIS AOT Estimates

Equation (3) describes the simple linear regression model (SLRM):

$$PM_{2.5} = \alpha_0 + \alpha_{aot}(AOT) \tag{3.3}$$

where $PM_{2.5}$ is the dependent variable ($PM_{2.5}$ concentration), α_0 is the intercept, and α_{AOT} is the slope of the SLRM.

To incorporate meteorological and AOT data, multiple linear regression models were developed. The equation for these models takes the general form:

$$PM_{2.5} = \alpha o + \alpha_{AOT}(AOT) + \alpha_{WS}(WS) + \alpha_{WD}(WD) + \alpha_{RH}(RH) + \alpha_T(T) \tag{3.4}$$

where $PM_{2.5}$ is the dependent variable ($PM_{2.5}$ concentration); α_0 is the intercept; α_{AOT}, α_{AOT}, α_{WS}, α_{WD}, α_{RH}, and α_T are the coefficients of the independent variables; WS is wind speed (m/s); WD is wind direction (°); and T is temperature (°C).

3.6.2 Geographically-Weighted Regression Model (GWRM)

The GWRM produces continuous surface of parameters by taking into consideration the spatial variations of the parameters instead of presuming a globally-constant coefficient. On a daily basis, the traditional GWRM can be written as shown in Eq. (3.5):

$$PM_{2.5(i,j,d)} = \alpha_{1(i,j,d)} + \alpha_{AOT(i,j,d)}.AOT_{(i,j,d)} + \alpha_{WS(i,j,d)}.WS_{(i,j,d)} + \alpha_{WD(i,j,d)}.WD_{(i,j,d)} + \alpha_{RH(i,j,d)}.RH_{(i,j,d)} + \alpha T_{(i,j,d)}.T_{(i,j,d)} \quad (3.5)$$

where $PM_{2.5(i,j,d)}$ is the ground-level $PM_{2.5}$ at site (i, j); $\alpha 1_{(i,j,d)}$ is the intercept of the regression equation; $\alpha AOT_{(i,j,d)}$, $\alpha WS_{(i,j,d)}$, $\alpha WD_{(i,j,d)}$, $\alpha RH_{(i,j,d)}$, and $\alpha T(i,j,d)$ are the slopes of the corresponding predictor variables; $AOT_{(i,j,d)}$ is the MODIS AOT value at site (i, j) in day d; $WS_{(i,j,d)}$ is the wind speed at site (i, j) in day d; $WD_{(i,j,d)}$ is the wind direction at site (i, j) in day d; $RH_{(i,j,d)}$ is the relative humidity at site (i, j) in day d; $T_{(i,j,d)}$ is the temperature at site (i, j) in day d.

The GWRM was calculated in the current study using the GWR4 software, which is a Microsoft Windows-based application. The Gaussian GWRM was selected as the model type of choice because it is suitable for modeling numerical responses. The GWRM is suitable for modeling count and binary responses. In addition, the adaptive type of kernel was selected because of the uneven distribution of the $PM_{2.5}$ monitoring sites. Moreover, the 'Golden Section Search' was employed to automatically specify the optimal bandwidth size while the Akaike Information Criterion (AIC) was employed for bandwidth selection. In general, the AIC can be employed to identify the model that is closest to reality and the optimum model ought to have the lowest AIC value.

Afterwards, the residual of the models was assessed via spatial autocorrelation analysis. In this research, spatial autocorrelation was performed by computing Moran's I values using ArcGIS 10.5. Detailed account on Moran's I values was provided by (Hu et al., 2013). In general, Moran's I values range from -1 to +1 (Hystad et al., 2015). When the Moran's I value is higher than zero, then there is positive spatial autocorrelation in the data. Negative values, on the other hand, denote negative autocorrelation. If Moran's I value is close to zero, then there is no spatial autocorrelation (Hu et al., 2013). The typical GWRM have residuals with no significant spatial autocorrelations (Quan Wang et al., 2005; Zhao et al., 2010). In other words, Moran's I values should be close to zero (Hu et al., 2013).

3.6.3 Model validation

In this study, ten-fold cross validation was performed. The dataset was partitioned into ten parts, each consisting of nearly 10.0% of the data points. Afterwards, nine parts were employed in training the model and the remaining part was employed in validation. The same process was followed for every fold.

3.7 Accuracy Assessment

The latest versions of the algorithms were used to generate aerosol products from the whole time series of the spectral observations of MODIS. Global validation attempts indicated potential uncertainties in AOT estimates of ($\Delta\tau = \pm 0.05 \pm 0.15\tau_{\text{AERONET}}$) over land and ($\Delta\tau = \pm 0.03 \pm 0.05\tau_{\text{AERONET}}$) over the ocean at the wavelength of 550 nm. These potential uncertainties encircle uncertainties in the radiative transfer through the atmosphere, potential errors in the computations, and the assumptions about the optical properties of the aerosol. In addition to this, the total uncertainty encompasses uncertainties in the presumed optical properties of the land surface and uncertainties related to pixel selection, cloud masking, and instrument accuracy and calibration.

This study applied spatiotemporal analysis, which produces grid of 5x5 MODIS aerosol retrieval pixels with the AERONET station in the middle. As every MODIS aerosol pixel represents about ten-kilometer area, the 5×5 pixel represents an area of nearly 50 km by 50 km. Spatial statistics for the MODIS subset were computed and compared with the temporal statistics (the average daily and hourly average statistics) of the AERONET observations taken within ± 30 min of the MODIS overpass.

3.7.1 Regression Metrics

Assessment of the accuracies of the developed models was carried out using the validated MODIS AOT and the $PM_{2.5}$ concentrations estimated in different pixel groups (1 x 1, 3 x 3, and 5 x 5). Evaluation of the levels of performance of the various models was based on the following four metrics: R^2, the RMSE, the relative RMSE rRMSE, and the Mean Absolute Percentage Error (MAPE).

3.7.1.1 The Coefficient of Determination (R^2)

In the case of simple regression analysis, the coefficient of determination is the square of Pearson's product-moment correlation coefficient (r). It expresses the amount of variance between two variables that is described by the linear fit. Correlation, usually expressed by a correlation coefficient (r), indicates the direction and strength of a linear relation between two variables (e.g., observed values and the corresponding model estimates). Several coefficients are usually employed for different situations. If we have a series of n observations and n model values, then the Pearson product-

moment correlation coefficient can be employed to express the correlation between the observations and the concomitant model estimates as follows (Equation 3.6):

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})\cdot(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{3.6}$$

3.7.1.2 The Root Mean-Square Error (RMSE)

The RMSE is a commonly-employed measure of the difference between the values predicted by a model and the values observed actually and which were modeled. Those individual differences are also described as residuals. The RMSE serves to compile them into one measure of predictive power. This error measure of model predictions is defined with regard to the estimated values (X_{model}) as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}} \tag{3.7}$$

where X_{obs} is the observed values of the variable of interest and X_{model} is the model estimates of those values at place or time i.

3.7.1.3 Relative RMSE (rRMSE)

The rRMSE is calculated using the following formula:

$$rRMSE = \frac{RMSE}{\bar{y}} x100\% \tag{3.8}$$

where the RMSE is the root mean-squared error and $\bar{y}$ is the mean of the values of the observed variable.

3.7.1.4 Mean Absolute Percentage Error (MAPE)

The mean absolute percentage error (MAPE) is given by the equation:

$$MAPE = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{yi - fi}{yi}\right| \tag{3.9}$$

where y_i is the value of the observed variable, f_i is the estimated value of the variable, and N is the number of observations.

3.8 Chapter Summary

This chapter provided an account on the methods used in this study to process the research data. The method consisted of four main processes: AOT validation, data pre-processing, model construction, and model assessment. Firstly, AOT validation was intended to validate the MODIS AOD product with the ground-level AOT data. Data pre-processing aimed at preparing the satellite data and the meteorological data for use in model construction. In the next step, the MLR and GWR models were built to estimate $PM_{2.5}$ concentrations and investigate its spatial distribution. Lastly, the model estimates of $PM_{2.5}$ in different groups of MODIS pixels were compared.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

The MODIS aerosol products have been used to analyze the spatial and temporal variability of aerosols for various applications and fields like calculation of the radiation budget and studies of cloud microphysics (Bellouin et al., 2008). Fine particulate matter plays important role in formation of the haze (Qiao et al., 2016; Zhang et al., 2013). Haze pollution has become nearly annual event in Southeast Asia since 1982, with the toughest episodes taking place in 1997 to 1998 and in 2006 to 2007. Since 1990, Malaysia is experiencing episodes of haze during the long dry weather period, usually ascribed to forest fires (Hod, 2016). The usefulness of satellite measurements of AOD for $PM_{2.5}$ monitoring was established in 2000 with increasing number of studies in this field (Hutchison, 2003; Wang, 2003).

4.2 Validation MODIS Sensor Data

The AOT retrievals are obtained at different places and times, depending on the progressive orbits whereas the AERONET Sunphotometer observations are made at differing time intervals. For taking the temporal and spatial variabilities into consideration, these data need to be collocated. The AERONET observations can be processed into daily average and hourly average (± 30 min). The hourly average corresponds to the reading taken when the satellite (Terra) overpasses the ground station. Table 4.1 shows the number of matches between the AERONET AOT and MODIS AOT stations. The percentages of the matching MODIS AOT and AERONET AOT data are listed in Figure 4.1. Generally, the match between the AERONET AOT and MODIS AOT data was low. It was less than 30% during the period 2012 to 2015 (1,416 days). The highest match in data (20.73%) was observed in USM station while the lowest match (9.38%) was observed in Kuching station (Figure 4.1). Based on this information, the researcher infers that temporal match in data between MODIS AOT and AERONET AOT depends on the season, ground data measurement, and cloud cover in the study area. This suggests that for increasing the strength of correlation between satellite and ground data, the period of study is an important parameter to take into consideration in future studies.

Table 4.1 : Numbers of matching AERONET station measurements and MODIS sensor data points

Station	Period[1]	Number of matching data points
Kuching		137
Pontianak		192
Singapore	2012-2015	160
Songkhla_Mat_Sta		187
USM_Penang		303

(1) This period corresponds to 1,416 days.

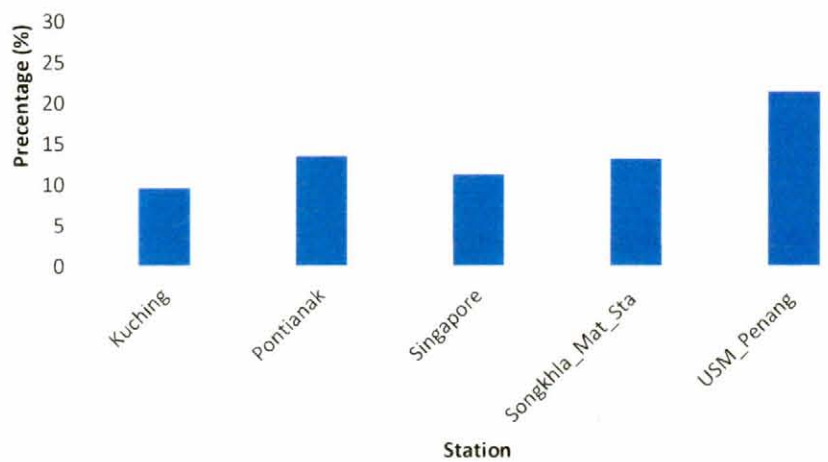


Figure 4.1 : Percentages of Matching MODIS Data and AERONET Station Measurements during the Period 2012 -2015 (1,461 Days)

4.2.1 Temporal Trend in the MODIS AOT Values

In order to compute the important basic statistical descriptions, statistical analysis was applied to all data from 2012 to 2015, that is, values of AOT, specifically, the MODIS AOT values taken at the wavelength of 550 nm (daily average), AERONET AOT values taken at the wavelength of 550 nm, and AERONET AOT values taken at the wavelength of 550 nm within ± 30 min, i.e., the hourly average (Table 4.2). The basic statistical descriptors of the data are the minimum (Min), maximum (Mix), and mean values, in addition to the standard deviation (S.D).

Table 4.2 reveals that the mean daily average of the ground AOT measurements of the AERONET stations was close to the AOT value derived from the MODIS AOT. That is, the mean daily values of the ground measurements (AERONET AOT) are almost similar to those of the satellite values (MODIS AOT), while the hourly average (± 30

min) values are lower. This means that the average AOT values were lower than the other averages in all stations (Table 4.2).

Table 4.2 : Summary statistics of the daily average AERONET AOT values, the hourly average AERONET AOT values, and MODIS AOT observations at the AERONET stations

Station	Variable[1]	N[2]	Min.	Max.	Mean	S.D
Kuching	AERONET Daily average	137	0.041	2.734	0.344	0.453
	AERONET Hourly average	137	0.057	2.136	0.291	0.363
	MODIS AOT	137	0.030	4.841	0.484	0.783
Pontianak	AERONET Daily average	192	0.032	4.785	0.503	0.766
	AERONET Hourly average	192	0.013	4.285	0.402	0.627
	MODIS AOT	192	0.073	4.823	0.564	0.801
Singapore	AERONET Daily average	160	0.104	1.876	0.394	0.269
	AERONET Hourly average	160	0.057	1.105	0.294	0.208
	MODIS AOT	160	0.123	1.638	0.454	0.304
Songkhla_Mat_Sta	AERONET Daily average	187	0.048	5.025	0.291	0.428
	AERONET Hourly average	187	0.035	5.562	0.279	0.589
	MODIS AOT	187	0.069	4.436	0.357	0.437
USM_Penang	AERONET Daily average	303	0.064	3.506	0.389	0.375
	AERONET Hourly average	303	0.035	3.706	0.305	0.367
	MODIS AOT	303	0.059	4.055	0.421	0.408

(1) All tabulated values were taken at the wavelength of 550 nm.
(2) N: Number of observations.

The results (Table 4.2) show the values of AERONET AOT, hourly average AERONET AOT, and MODIS AOT had similar trend in the period 2012 to 2015 in all stations; Kuching (Figure 4.2), Pontianak (Figure 4.3), Singapore (Figure 4.4), Songkhla_Mat_Sta (Figure 4.5), and USM_Penang (Figure 4.6). Therefore, good correlation is noticed between the ground and satellite measurements during the study period. The highest AOT values were reported in 2015. A possible explanation for this finding is haze since a haze episode took place in 2015, which was connected to the long dry conditions caused by the El Nino effect. These findings (Table 4.2; Figures 4.2-4.6) agree with findings of a study conducted in Indonesian provinces, where a state of emergency was announced by the government of Indonesia, considering the tremendous effect of haze, which also affected Malaysia and Singapore (Ho et al., 2018).

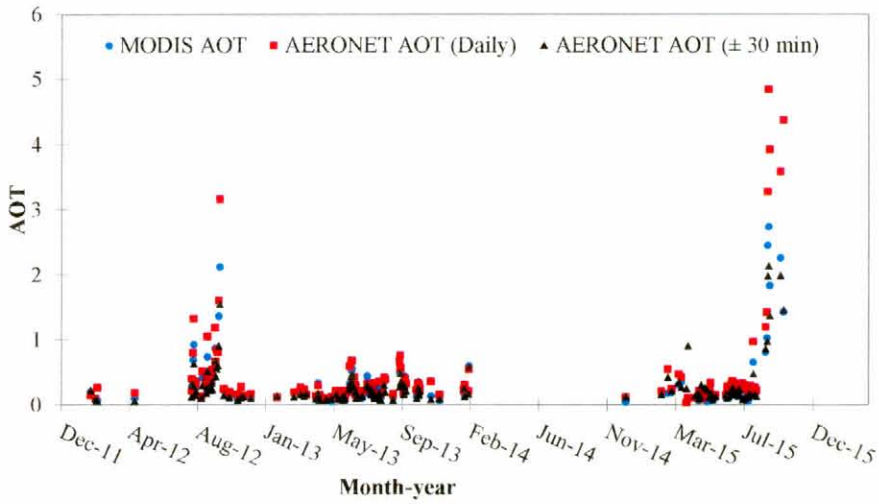


Figure 4.2 : Time Series of AERONET AOT Daily Average, AERONET AOT Hourly Average, and MODIS AOT in Kuching Station

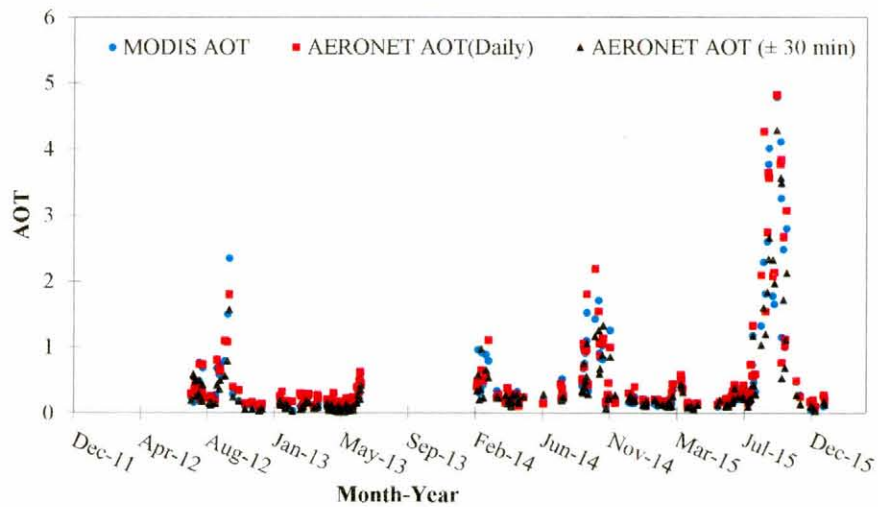


Figure 4.3 : Time series of AERONET AOT Daily average, AERONET AOT Hourly average at 550 nm and MODIS AOT 550 nm in Pontianak station

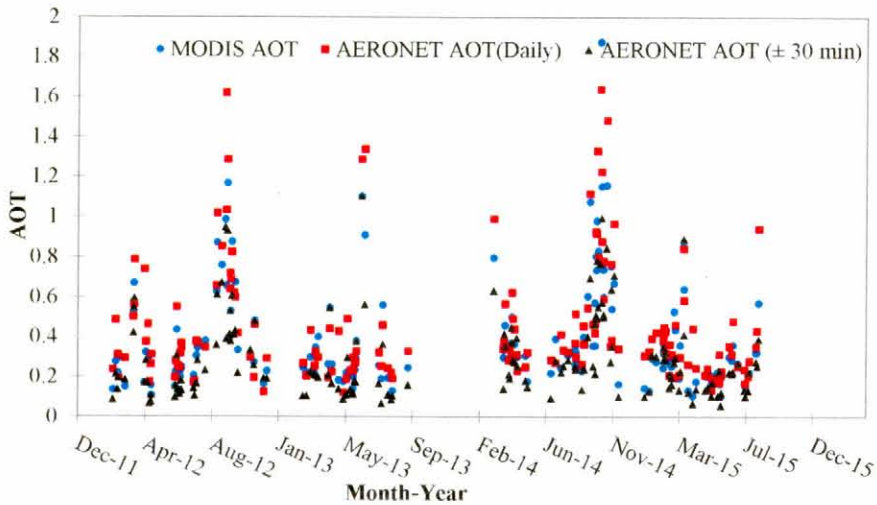


Figure 4.4 : Time series of AERONET AOT Daily average, AERONET AOT Hourly average at 550 nm and MODIS AOT 550 nm in Singapore station

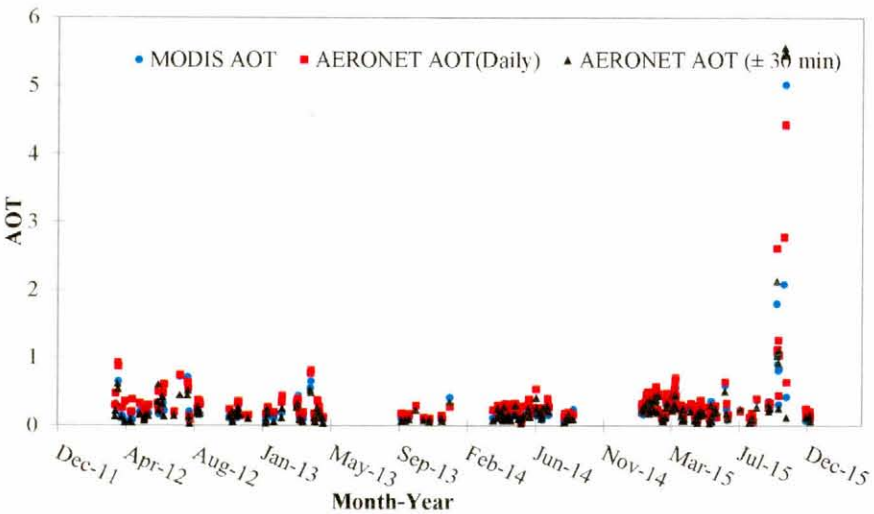


Figure 4.5 : Time series of AERONET AOT Daily average, AERONET AOT Hourly average at 550 nm and MODIS AOT 550 nm in Songkhla_Mat_Sta station

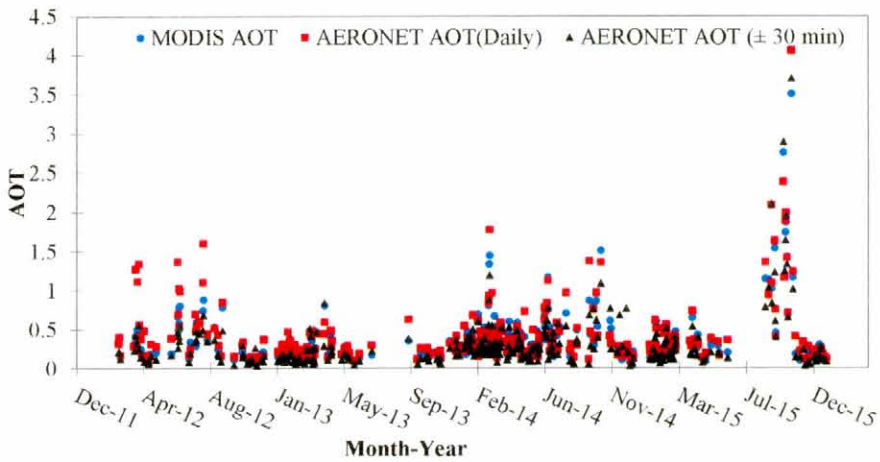


Figure 4.6 : Time series of AERONET AOT Daily average, AERONET AOT Hourly average at 550 nm and MODIS AOT 550 nm in USM_Penang station

4.2.2 Validtion of the MODIS AOT Retrievals with AERONET AOT Data

Table 4.3 lists the values of the seasonal average mean AERONET AOT (0.387 ± 0.49), hourly average AERONET AOT (0.315 ± 0.46), and MODIS AOT (0.451 ± 0.56) in the period 2012 to 2015. Although the sources of the AOT measurements are different, the seasonal trends in AOT over the region are comparable with annual variation. Overall, the values of R^2 for the regression models of daily average AERONET AOT, hourly average AERONET AOT, and MODIS AOT were higher than 0.75 during the study period in all sites (Table 4.4). These results present additional support for the practice of using AERONET data to validate satellite-derived AOD values because of their quite high accuracy (Ma et al., 2014).

Table 4.3 : Summary statistics of the daily average AERONET AOT, hourly average AERONET AOT (± 30 min), and MODIS AOT for all stations

Variable	Observations	Minimum	Maximum	Mean	S.D
Daily average AERONET AOT	979	0.032	5.025	0.387	0.488
Hourly average AERONET AOT (± 30 min)	979	0.013	5.562	0.315	0.460
MODIS AOT	979	0.030	4.841	0.451	0.564

Table 4.4 : Coefficients of correlations among daily average AERONET AOT, hourly average AERONET AOT, and MODIS AOT[1], [2]

Variables	MODIS AOT	AERONET AOT Daily average	AERONET AOT Hourly average
MODIS AOT	1	0.871	0.779
Daily average AERONET AOT	0.871	1	0.851
Hourly average AERONET (± 30 min)	0.779	0.851	1

(1) All values were taken at the wavelength of 550 nm
(2) p values were lower than 0.0001

There are noticeable seasonal variations in the ground and satellite AOT values at the background regional site. The seasonal and monthly variations in the daily average AERONET AOT, hourly average AERONET AOT, and MODIS AOT were highly consistent (Figure 4.7). These results coincide with findings of some previous studies. For example, Tan et al. (2015) found that the AOD always shifts to a higher AOD during the northeast and southwest monsoons in Penang and Kuching than during the inter–monsoon period. Despite differences in the study periods between the current study and the study of Tan et al. (2015), the relative frequency distributions of Angstrom at Penang and Kuching have the same pattern.

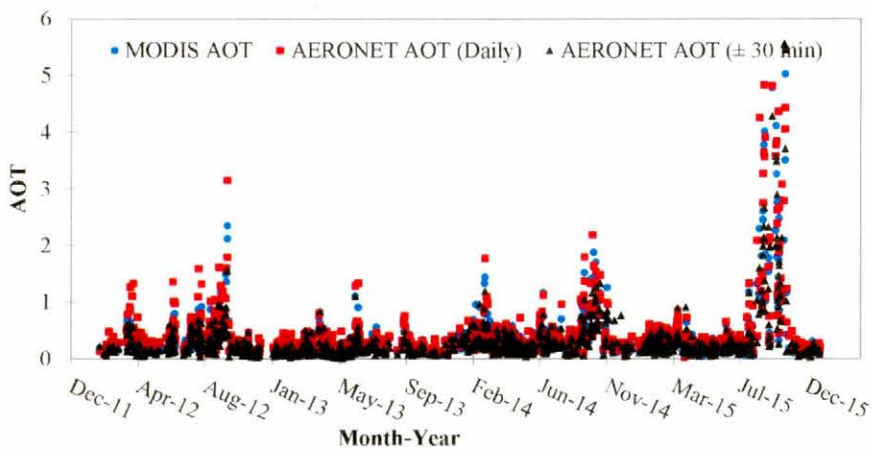


Figure 4.7 : Time Series of the Daily Average AERONET AOT, Hourly Average AERONET AOT, and MODIS AOT from January 2012 to December 2015

The MODIS AOD compares better with ground-based measurements at Songkhla station that at the other stations (Jantarach et al., 2012). More data were collected during the dry period than during the other periods, which is a period known for high biomass burning activities. Particular emphasis was laid on identification of the aerosols originating from biomass burning, which peak during those months (Salinas

et al., 2013b). On account of these findings (Figure 4.7), the researcher concludes that the weather conditions and haze played substantial role in exacerbating the distribution of air pollutants and the amount of PM during the southwest (dry season). The most noticeable finding to emerge from the analysis is that there is a response to the low and high AOT readings (Figure 4.7). In light of these findings, it can be inferred that AOT retrieved from AERONET is broadly applied for validation of the satellite-derived AOT values.

In this study, five AERONET stations were selected for validation of the MODIS sensor data. Figure 4.8 presents a scatter plot of validation of the MODIS AOT retrievals. The R^2 values for both the hourly average (Figure 4.8) and the daily average (Figure 4.9) were higher than 0.7, indicating that the correlation between the measurements and observations is significant, positive, and strong. It is found that the daily ground AOT values obtained from AERONET better explain the AOT characteristics than the hourly AOT values (± 30 min) and that they can be more suitable for the validation of the MODIS sensor data. As can be noticed in Figure 4.9, agreement between the MODIS AOD retrievals and the daily average AERONET AOT observations ($R^2 = 0.87$) was higher than agreement between the MODIS AOD retrievals and the hourly average AERONET AOT observations ($R^2 = 0.77$). However, the findings of the current study of strong relations between the AERONET measurements and MODIS AOT flow in harmony with findings of previous research revealing that the MODIS sensor data tend to have high relationship with the AERONET observations over land, where values of R^2 ranged from 0.69 to 0.99 (Huang et al., 2015; Jing-Mei et al., 2010; Levy et al., 2010; Z. Ma et al., 2016; Misra et al., 2008; Pan et al., 2010).

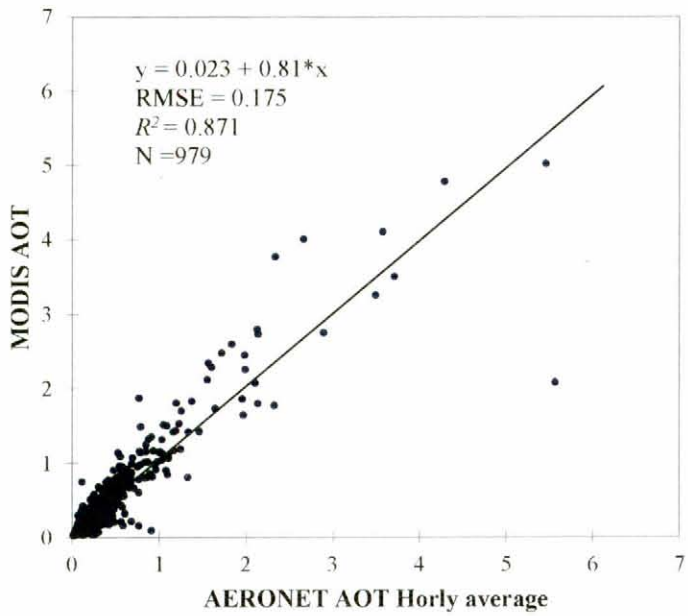


Figure 4.8 : Scatter Plot of MODIS AOD Retrievals against the Hourly Average AERONET AOT Observations

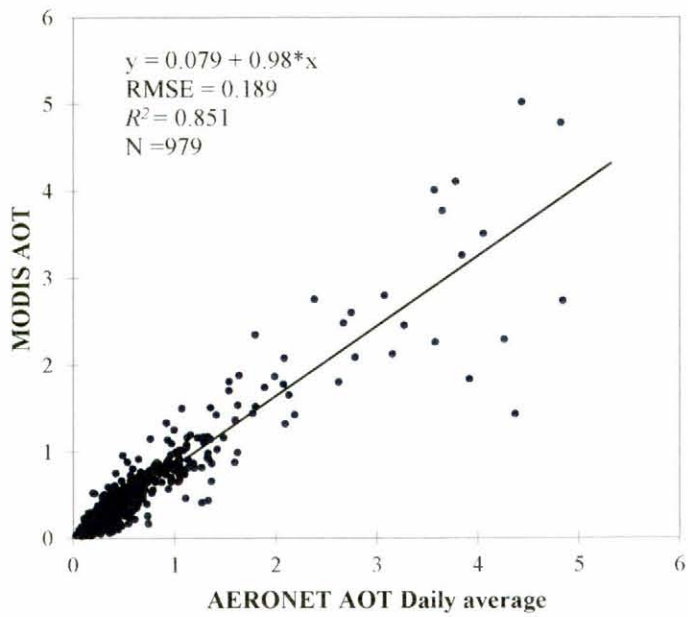


Figure 4.9 : Scatter plot of MODIS AOD Retrievals against the Daily Average AERONET AOT Observations

Results of the present study bear resemblance with those of some previous studies, where MODIS AOT observations were validated by holding comparisons with ground-based AERONET observations at the wavelength of 550 nm. Guo et al. (2014) reported a good correlation between the mean hourly AERONET measurements and MODIS AOD retrievals at the wavelength of 550 nm, with r values of 0.895 for the Aqua satellite and 0.950 for the Terra satellite. Zheng et al. (2017) too reported a high correlation between AERONET and MODIS AOT ($R^2 = 0.85$). The results of the study of Alam et al. (2011) suggest that the value of the coefficient of correlation between AERONET and MODIS ($r = 0.72$) is indicative of better AOT estimates that are closer to the terrestrial measurements than what can be obtained based on the relationship between AERONET and MISR ($r = 0.11$). In summary, the results prove that MODIS AOD products correlate well with the AERONET AOD observations. Although some bias still exists, the validation results are satisfactory overall.

Generally, validation of the satellite data helps in reducing the various uncertainties for determining the ground-level aerosol concentrations, simulating the aerosol-climate interactions, and forecasting the air quality, both at the regional and local scales. Of the various sensors that continuously provide aerosol columnar data throughout the world, MODIS on board of the Aqua and Terra satellites is recognized as the sensor most widely resorted to for relevant data (Mhawish et al., 2017).

4.3 Correlating the Ground AOT with Validated MODIS AOT

This section presents and discusses the results of correlating the ground AOT with the validated MODIS AOT values. The main aim of this step was to identify the temporal and spatial variations in the AOT over Peninsular Malaysia using the data obtained from the MODIS sensor, then to estimate the levels of the $PM_{2.5}$ using these data.

Peninsular Malaysia has an area of 131,598 km^2. This area is covered with approximately 1,314 MODIS pixels (one pixel represents an area of 10 x 10 km^2). Table 4.5 shows that in only 10 days during the study period (2013) did MODIS cover more than 50% of the overall area covered with the pixels. The highest number of pixels covered by MODIS was 994 of 1,314 in 18 May 2013, while in the rest of the days the percentage of pixels covered by MODIS was less than 50%.

Table 4.5 : Percentages of pixels covered by MODIS during the study year (2013)

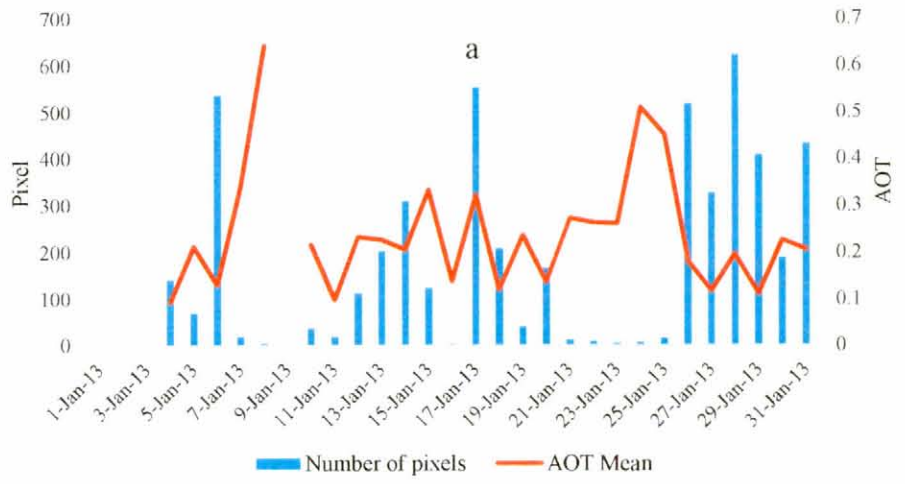
No.	Date	Number of pixels out of 1,314 pixels	Date	The tenth high max. AOT	Date	The tenth high mean of AOT
1	18-May-13	994	2-Nov-13	6.000	2-Nov-13	2.142
2	27-May-13	962	24-Jun-13	4.956	24-Jun-13	1.699
3	19-Jun-13	895	25-Jun-13	4.576	25-Jun-13	1.640
4	3-Jun-13	888	19-Jun-13	4.571	24-Jul-13	1.326
5	24-Jun-13	887	23-Jun-13	4.456	27-Sep-13	1.072
6	16-Apr-13	816	21-Jun-13	4.420	30-Aug-13	0.995
7	17-Aug-13	789	20-Jun-13	3.075	9-Apr-13	0.9178
8	24-Aug-13	772	29-Aug-13	2.820	23-Jul-13	0.9107
9	8-Jun-13	750	24-Jul-13	2.582	26-Sep-13	0.9106
10	7-May-13	703	22-Jul-13	2.327	27-Aug-13	0.8956

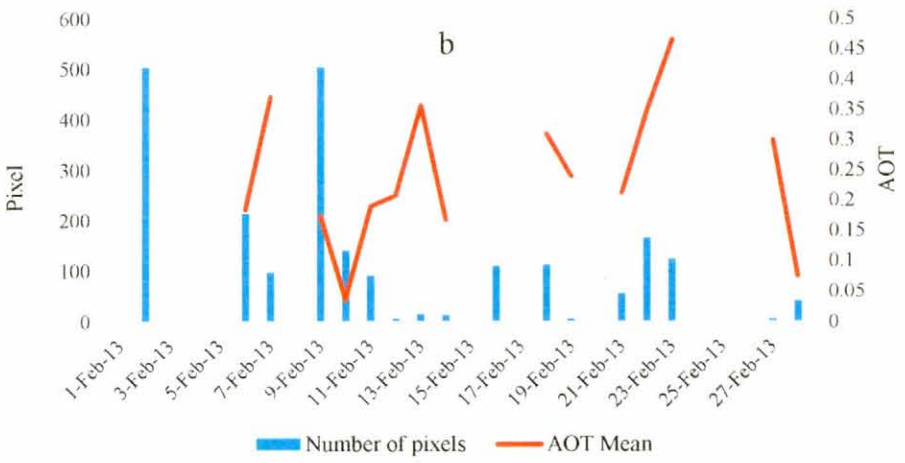
4.3.1 Temporal Distribution of AOT

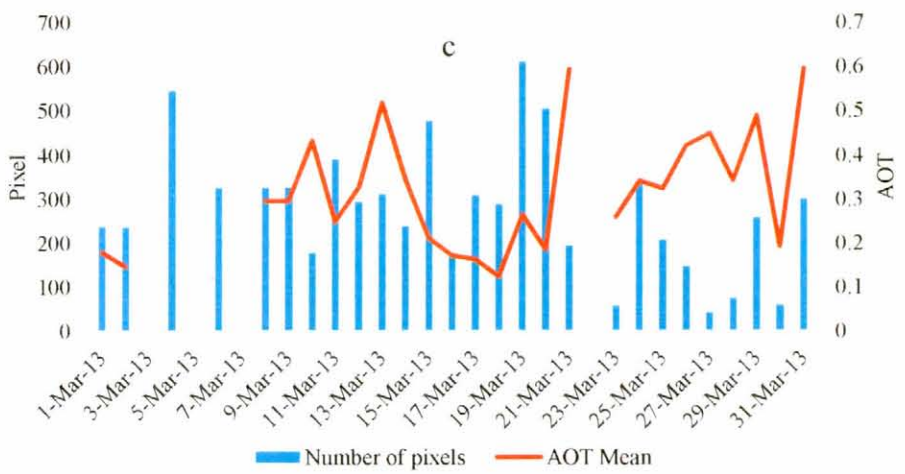
Table 4.6 and Figure 4.10 show the results of MODIS AOT retrievals by month over Peninsular Malaysia in the year 2013. The results indicate that massing data appear in each month. However, the highest number of missing data points characterized November (13 days) and February (11 days). A possible explanation for this is the cloud cover which is a serious limitation to satellite AOT observation. The range of maximum AOT values was 0.57 in December to 2.14 in November, while the lowest value of AOT (0.036) was reported in February. The temporal and spatial variations in aerosols and their impact on radiative transfer over Malaysia were studied by (Kanniah et al., 2016) via the integrated approach based on columnar aerosol disappearance and the corresponding AOT over a given area. It can, thus, be highlighted that the MODIS images captured AOT in a variety of values. The highest mean average monthly AOT value (0.512) was recorded in June. In general, the months of the southwest monsoon recorded high mean and median AOT values.

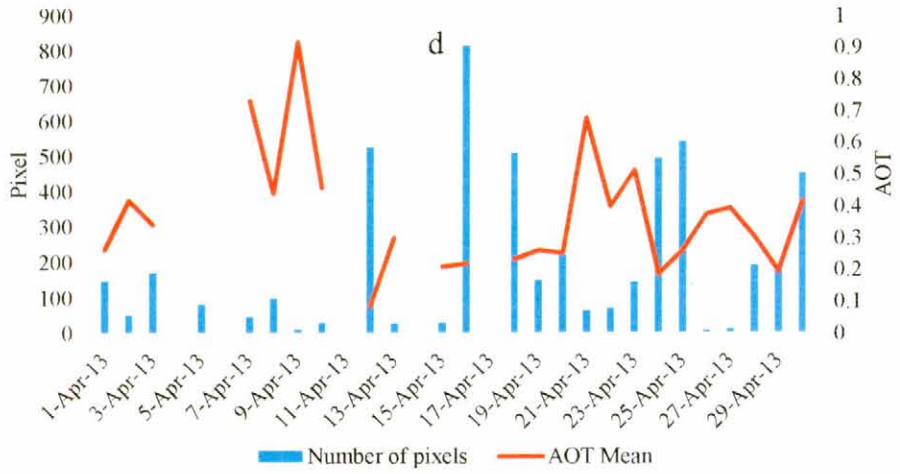
Table 4.6 : The available and missing MODIS AOT pixels by month over Peninsular Malaysia in 2013

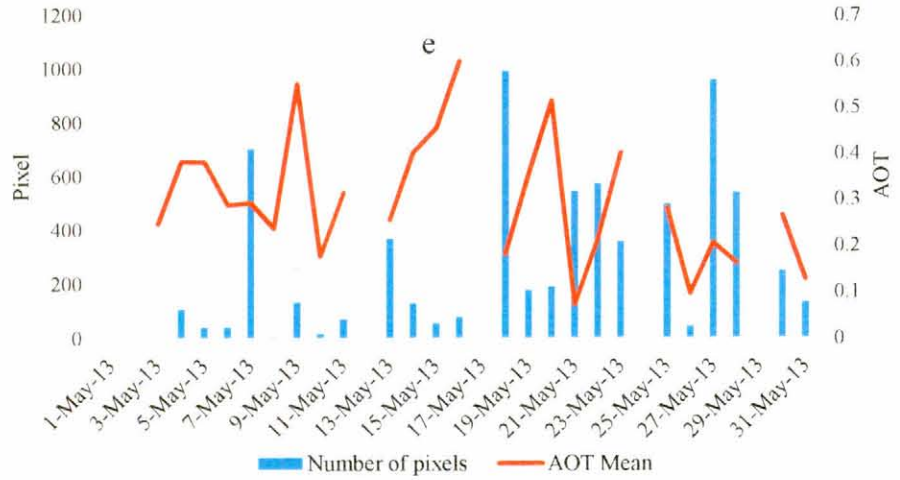
Month	Observation	Missing	Min	Max.	Mean	Median
January	31	4	0.124	0.596	0.313	0.296
February	28	11	0.036	0.466	0.257	0.241
March	31	4	0.124	0.596	0.314	0.296
April	30	5	0.086	0.918	0.385	0.342
May	31	6	0.074	0.601	0.299	0.282
June	30	5	0.144	1.699	0.512	0.334
July	31	7	0.149	1.326	0.484	0.455
August	31	4	0.044	0.995	0.430	0.386
September	30	9	0.076	1.072	0.439	0.356
October	30	8	0.059	0.607	0.293	0.266
November	30	13	0.121	2.142	0.479	0.329
December	31	15	0.107	0.570	0.307	0.288

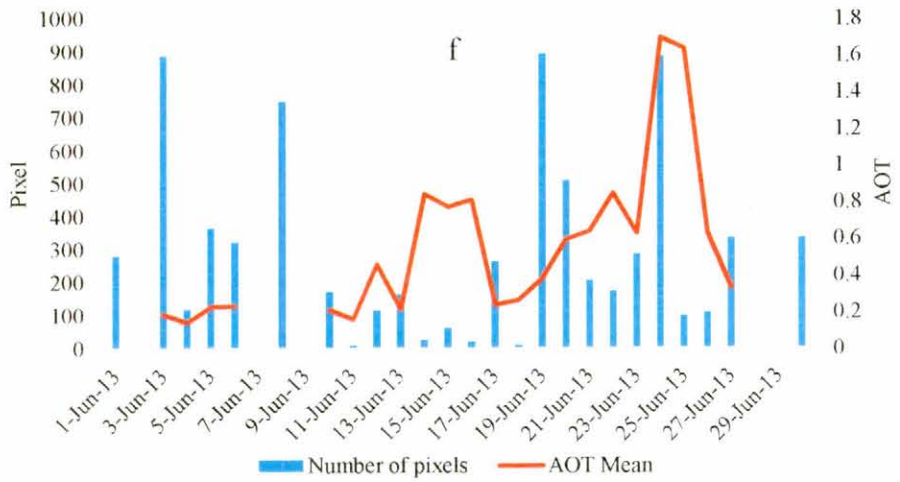

a
Pixel
AOT
700
600
500
400
300
200
100
0
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1-Jan-13
3-Jan-13
5-Jan-13
7-Jan-13
9-Jan-13
11-Jan-13
13-Jan-13
15-Jan-13
17-Jan-13
19-Jan-13
21-Jan-13
23-Jan-13
25-Jan-13
27-Jan-13
29-Jan-13
31-Jan-13
Number of pixels
AOT Mean

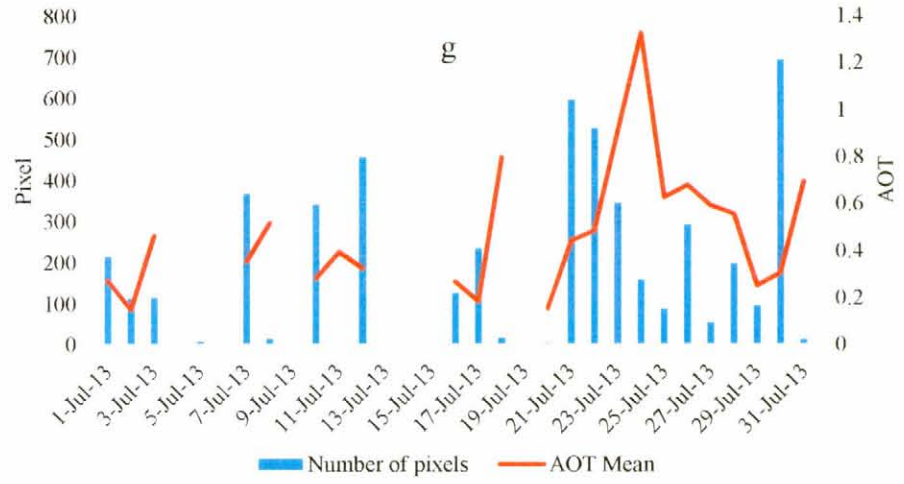

b
Pixel
AOT
600
500
400
300
200
100
0
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1-Feb-13
3-Feb-13
5-Feb-13
7-Feb-13
9-Feb-13
11-Feb-13
13-Feb-13
15-Feb-13
17-Feb-13
19-Feb-13
21-Feb-13
23-Feb-13
25-Feb-13
27-Feb-13
Number of pixels
AOT Mean

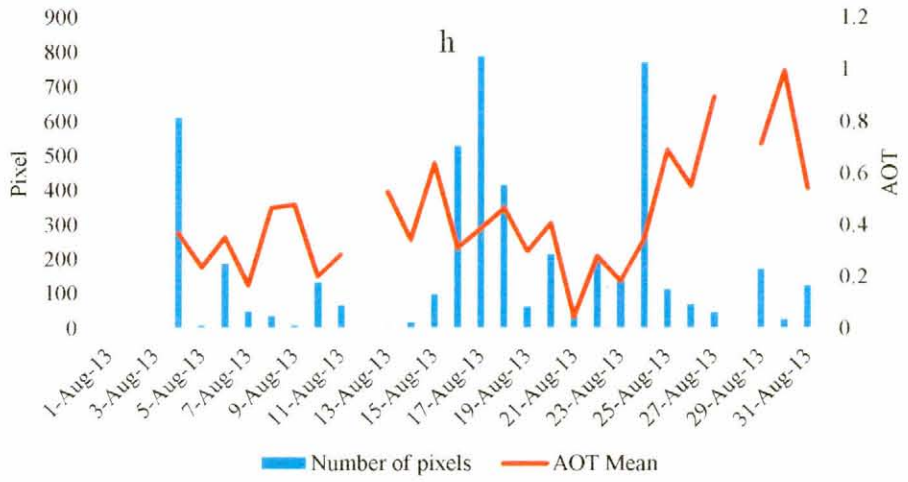

c
Pixel
AOT
700
600
500
400
300
200
100
0
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1-Mar-13
3-Mar-13
5-Mar-13
7-Mar-13
9-Mar-13
11-Mar-13
13-Mar-13
15-Mar-13
17-Mar-13
19-Mar-13
21-Mar-13
23-Mar-13
25-Mar-13
27-Mar-13
29-Mar-13
31-Mar-13
Number of pixels
AOT Mean

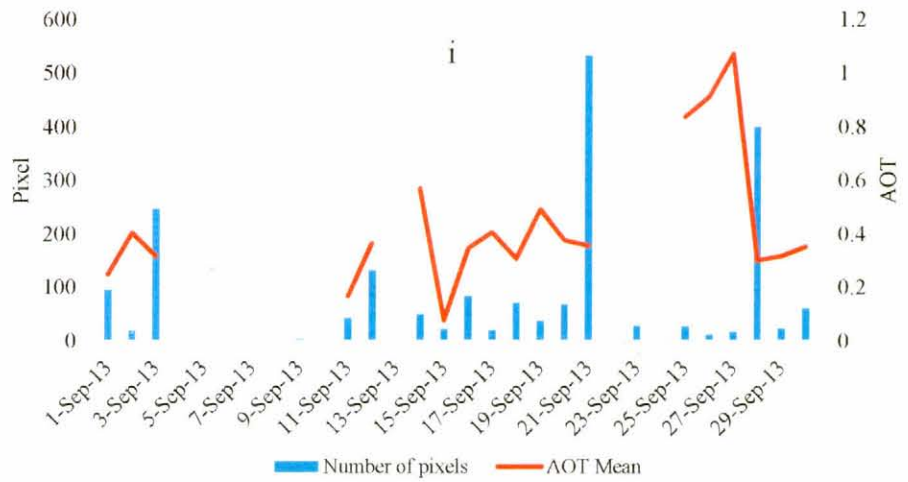

d
Pixel
AOT
900
800
700
600
500
400
300
200
100
0
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1-Apr-13
3-Apr-13
5-Apr-13
7-Apr-13
9-Apr-13
11-Apr-13
13-Apr-13
15-Apr-13
17-Apr-13
19-Apr-13
21-Apr-13
23-Apr-13
25-Apr-13
27-Apr-13
29-Apr-13
Number of pixels
AOT Mean


e
Pixel
AOT
1200
1000
800
600
400
200
0
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1-May-13
3-May-13
5-May-13
7-May-13
9-May-13
11-May-13
13-May-13
15-May-13
17-May-13
19-May-13
21-May-13
23-May-13
25-May-13
27-May-13
29-May-13
31-May-13
Number of pixels
AOT Mean


f
Pixel
AOT
1000
900
800
700
600
500
400
300
200
100
0
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1-Jun-13
3-Jun-13
5-Jun-13
7-Jun-13
9-Jun-13
11-Jun-13
13-Jun-13
15-Jun-13
17-Jun-13
19-Jun-13
21-Jun-13
23-Jun-13
25-Jun-13
27-Jun-13
29-Jun-13
Number of pixels
AOT Mean


g
Pixel
AOT
800
700
600
500
400
300
200
100
0
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1-Jul-13
3-Jul-13
5-Jul-13
7-Jul-13
9-Jul-13
11-Jul-13
13-Jul-13
15-Jul-13
17-Jul-13
19-Jul-13
21-Jul-13
23-Jul-13
25-Jul-13
27-Jul-13
29-Jul-13
31-Jul-13
Number of pixels
AOT Mean


h
Pixel
AOT
900
800
700
600
500
400
300
200
100
0
1.2
1
0.8
0.6
0.4
0.2
0
1-Aug-13
3-Aug-13
5-Aug-13
7-Aug-13
9-Aug-13
11-Aug-13
13-Aug-13
15-Aug-13
17-Aug-13
19-Aug-13
21-Aug-13
23-Aug-13
25-Aug-13
27-Aug-13
29-Aug-13
31-Aug-13
Number of pixels
AOT Mean


i
Pixel
AOT
600
500
400
300
200
100
0
1.2
1
0.8
0.6
0.4
0.2
0
1-Sep-13
3-Sep-13
5-Sep-13
7-Sep-13
9-Sep-13
11-Sep-13
13-Sep-13
15-Sep-13
17-Sep-13
19-Sep-13
21-Sep-13
23-Sep-13
25-Sep-13
27-Sep-13
29-Sep-13
Number of pixels
AOT Mean

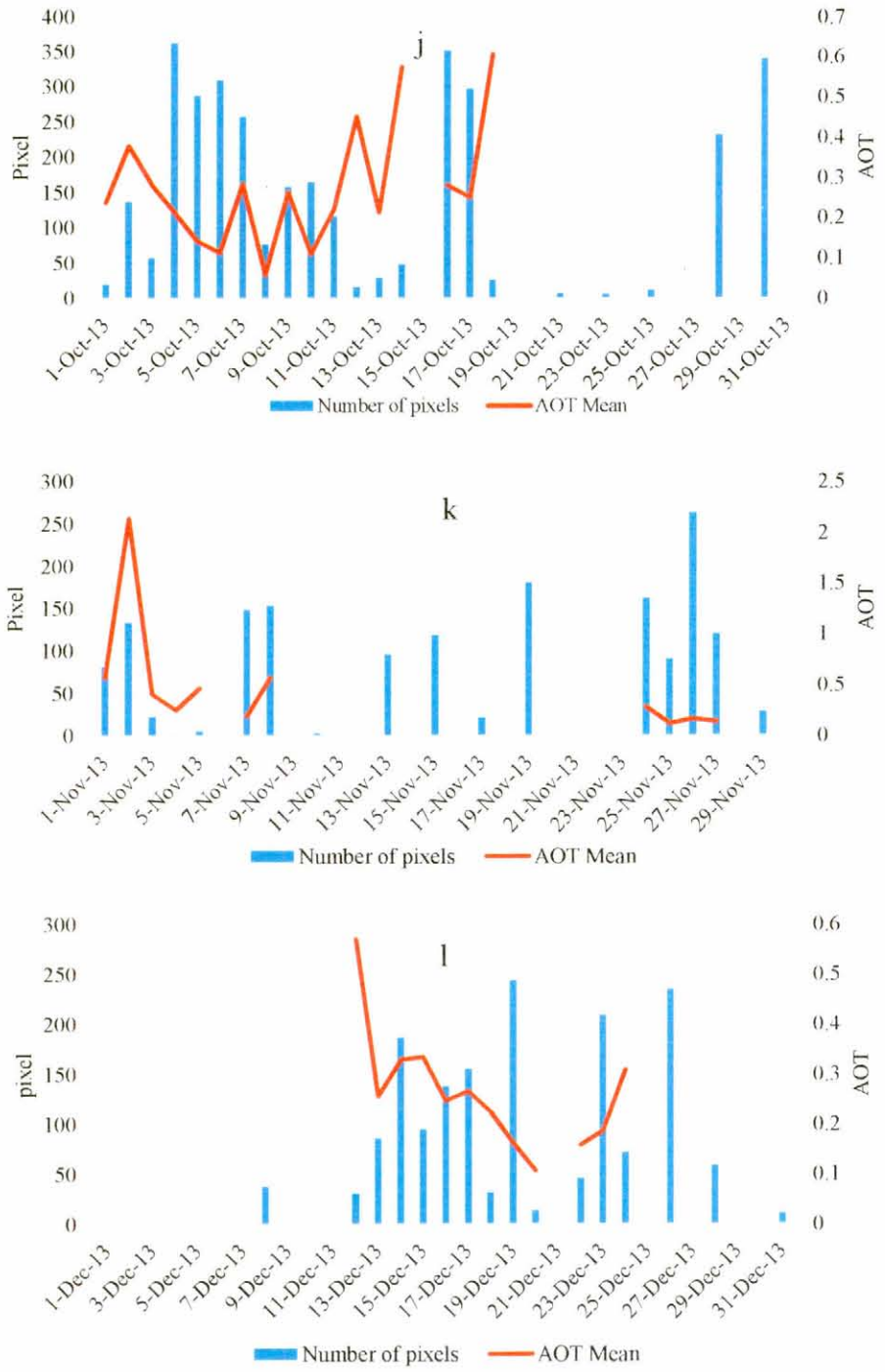
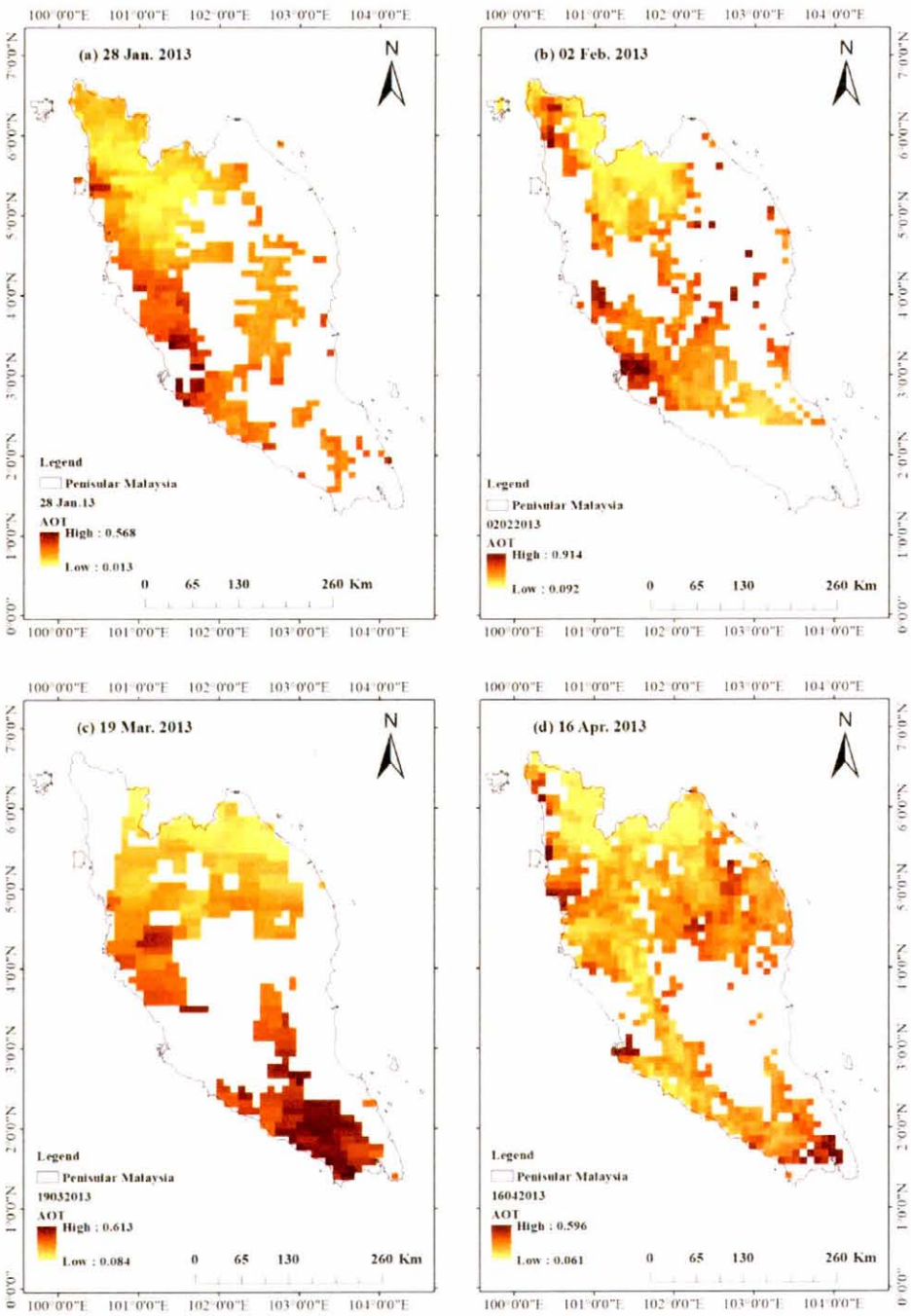
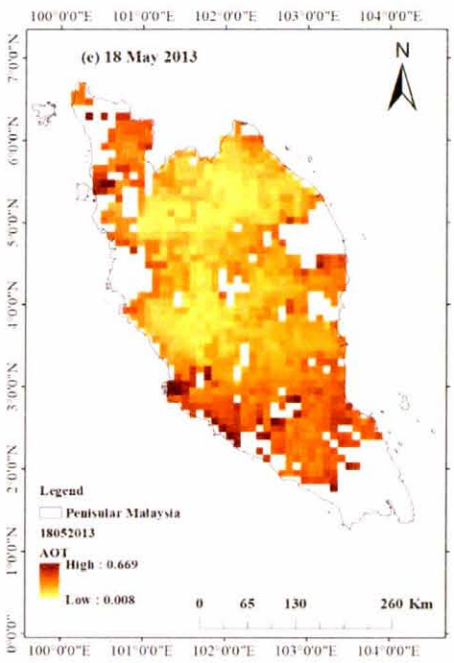


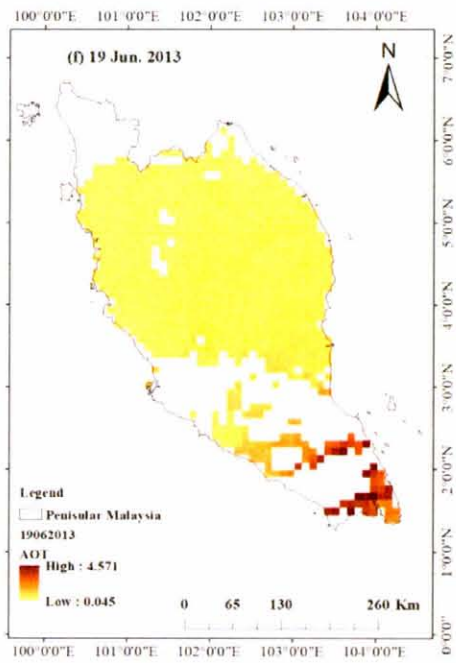
Figure 4.10 : The Average Daily AOT Values Observed Monthly over Peninsular Malaysia in 2013

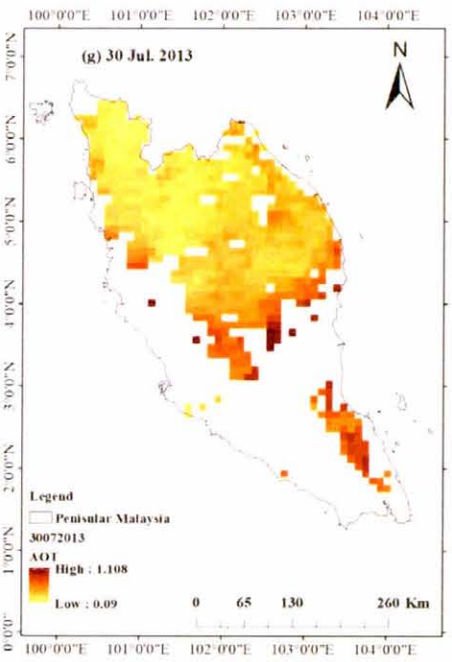
4.3.2 Spatial Distribution of AOT

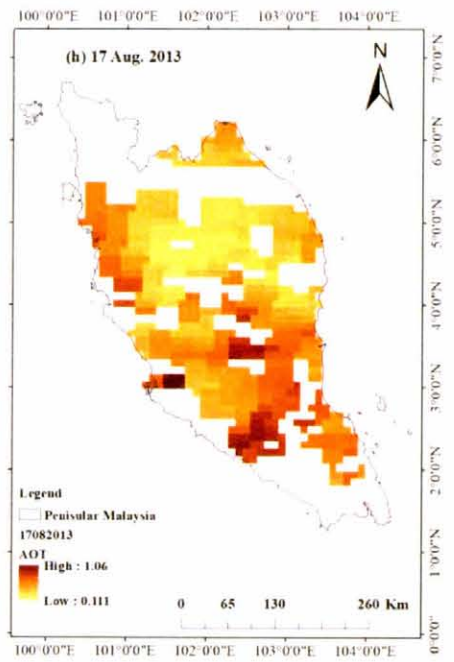
Figure 4.11 uncovers that there are many missing data in the study area, most probably due to a cloud cover. On the other hand, the dark orange color, expressing high values of AOT, concentrated on the west coast of peninsular Malaysia. It is highly expected that the values of AOT were influenced by the season. This area is always affected by haze due to biomass burning, which increases the concentration of AOT. Many previous studies (e.g., Othman et al. (2014)) reported that Malaysia is exposed to the haze and that it was exposed to haze during 2005 and 2006.


(a) 28 Jan. 2013
N
Legend
Penisular Malaysia
28 Jan.13
AOT
High : 0.568
Low : 0.013
0 65 130 260 Km
(b) 02 Feb. 2013
N
Legend
Penisular Malaysia
02022013
AOT
High : 0.914
Low : 0.092
0 65 130 260 Km
(c) 19 Mar. 2013
N
Legend
Penisular Malaysia
19032013
AOT
High : 0.613
Low : 0.084
0 65 130 260 Km
(d) 16 Apr. 2013
N
Legend
Penisular Malaysia
16042013
AOT
High : 0.596
Low : 0.061
0 65 130 260 Km


(e) 18 May 2013
N
Legend
Penisular Malaysia
18052013
AOT
High : 0.669
Low : 0.008
0 65 130 260 Km


(f) 19 Jun. 2013
N
Legend
Penisular Malaysia
19062013
AOT
High : 4.571
Low : 0.045
0 65 130 260 Km


(g) 30 Jul. 2013
N
Legend
Penisular Malaysia
30072013
AOT
High : 1.108
Low : 0.09
0 65 130 260 Km


(h) 17 Aug. 2013
N
Legend
Penisular Malaysia
17082013
AOT
High : 1.06
Low : 0.111
0 65 130 260 Km

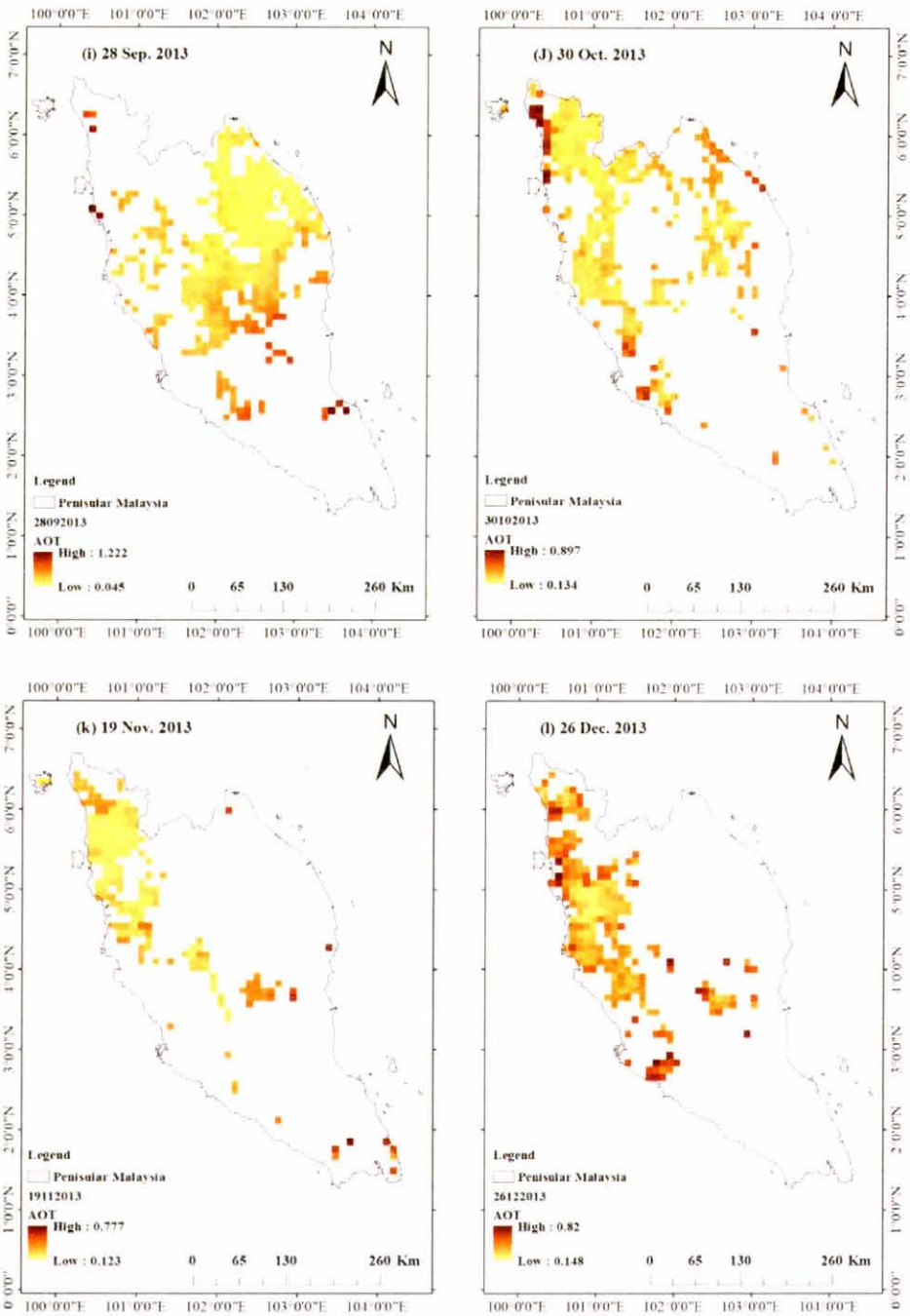


Figure 4.11 : Spatial Distribution of the MODIS-AOT Retrievals in Peninsular Malaysia in 2013

Other studies of the relationship between haze and AOT (e.g., Tsai et al. (2011)) employed MODIS-derived AOD data for air quality monitoring. These studies found that the height of the haze layer has higher influence on correlation than the BLH as a result of abundance of aerosols above the boundary layer. Additionally, Salinas et al. (2013a) found that the fine mode AOD indicated high levels of aerosol loading and that there were similar high values of the Angstrom exponent number, consistent with presence of the fine mode PM. Nevertheless, these results suggest that data obtained using the MODIS sensor to retrieve AOT provide further information for assessing the effect of haze and for estimating the concentrations of $PM_{2.5}$.

It is encouraging to compare these figures with findings of Amanollahi et al. (2011) who found that change in the concentration of PM_{10} and value of AOT were affected by haze in Klang Valley during and after the severe haze episode, which extended from July to August 2005. It is seen that the high aerosol loadings were usually located in Klang Valley in Peninsular Malaysia, which is the most developed area in Malaysia and which is more exposed to air pollution than other areas in Malaysia due to its geographical location, establishment of large-scale commercial and industrial activities, abundance of densely-populated areas, and the high vehicular traffic (Sahani et al., 2014). Overall, these results compare with those of the study of (Shon, 2015) which reported high effect of biomass burning on the optical properties of aerosols in Southeast Asia.

4.3.3 Annual Means and Seasonal Variations in MODIS AOT Retrievals in Monitoring Stations

The seasonal variations in AOT were tracked using MODIS data for one year; 2013 (Table 4.7). The seasonal and annual AOT means and associated standard deviations are listed in Table 4.7. It is found that the highest AOT values were recorded during the southwest monsoon, especially over Cheras (0.726 ± 0.600) and Putrajaya (0.716 ± 0.342). The results also show that there are increases in the AOT values in the southwest monsoon at all stations. Meanwhile, no clear trend was observed at Kelantan Tanah Merah station, where the mean annual AOT value was 0.327 ± 0.171, the mean southwest AOT value was 0.324 ± 0.190, and the mean northeast AOT value was 0.338 ± 0.160. These AOT values are almost identical. This may suggest that the station on the east coast was affected by small quantities of AOT. Tan et al. (2014) found that the amounts of the aerosols resulting from biomass burning increased during the southwest monsoon due to the active open-burning activities in local areas and neighboring countries as identified by the current study.

Table 4.7 : Annual and seasonal MODIS AOT means and standard deviations at different stations in Peninsular Malaysia in 2013

Station	Annual and seasonal means and standard deviations of MODIS AOD		
	Annual mean	Southwest monsoon	Northeast monsoon
Banting	0.451 ± 0.241	0.534 ± 0.291	0.394 ± 0.217
Bukit Rambai	0.387 ± 0.059	0.431 ± 0.310	0.369 ± 0.206
Cheras	0.509 ± 0.408	0.726 ± 0.600	0.407 ± 0.243
Ipoh	0.375 ± 0.210	0.448 ± 0.234	0.328 ± 0.194
Putrajaya	0.545 ± 0.291	0.716 ± 0.342	0.449 ± 0.203
Kelantan Tanah Merah	0.327 ± 0.171	0.324 ± 0.190	0.338 ± 0.160
USM	0.393 ± 0.298	0.527 ± 0.506	0.342 ± 0.150

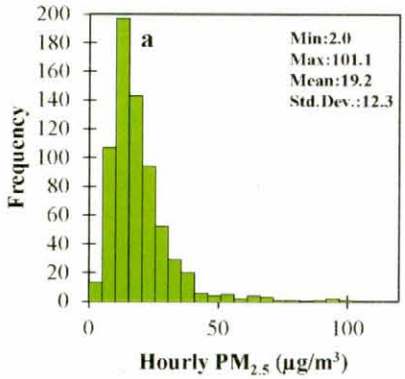
4.3.4 The MODIS AOT Retrievals and their Relation to Ground-Level $PM_{2.5}$ Concentrations

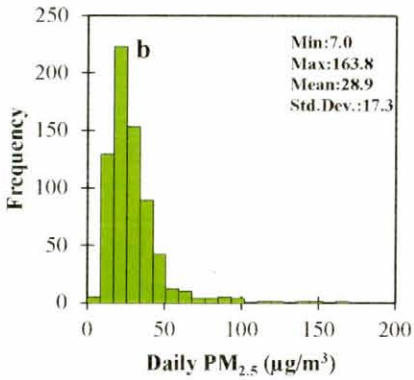
Accuracy of the AOT values drawn from satellite does not only depend on quality of the satellite data (e.g., radiometric calibration) but also on the algorithm employed in converting the satellite data to AOT data. Since the AOT retrieval process is carried out only during the daytime and for the cloud-free pixels, the most important step is applying the appropriate cloud mask. The smaller the size of the pixel, the more efficient is the process of removal of the cloudy pixels. Stated otherwise, the AOT represents the column-integrated value while the PM is measured at the ground level. In consequence, correlation between these two variables may depend on vertical distribution of the aerosols. As a result, estimation of the PM concentrations from the MODIS AOT data may result in remarkable errors. To enhance the correlation between the $PM_{2.5}$ and AOT in Peninsular Malaysia, this study included several meteorological parameters in the regression models.

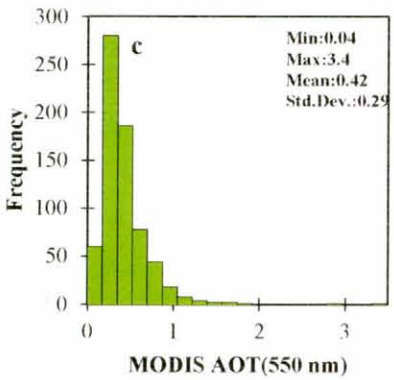
4.3.4.1 Descriptive Statistics

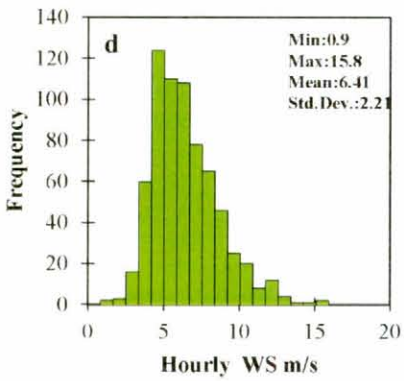
Daily and hourly meteorological and $PM_{2.5}$ data, besides the AOT data, were used for parameterization of the SLRM and the MLRM. The modeling results are summarized in Figure 4.12 using histograms and descriptive statistics. The sensor data were collected between 11:00 a.m and 01:00 p.m, Malaysia local time. The MODIS instruments are on board of the Terra satellite, which crosses the equator around 10:30 a.m local solar time (Wei You et al., 2016). The mean value of the AOT (independent and unitless variable) used in the SLRM and MLRM was 0.42, while the standard deviation was 0.29. The hourly and daily ranges for the corresponding ground-level $PM_{2.5}$ concentrations (dependent variable measured in unit $\mu g/m^3$ unit) are 2.00 $\mu g/m^3$ to 101.00 $\mu g/m^3$ and 7.00 $\mu g/m^3$ to 163.79 $\mu g/m^3$, respectively.

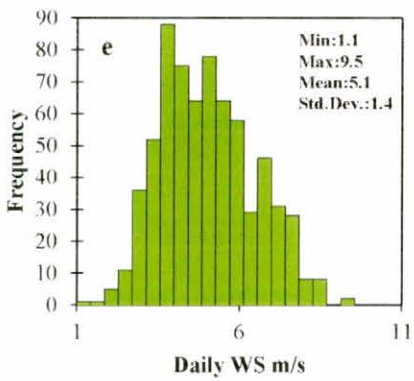
The ground-based meteorological variables used in the hourly and daily MLRMs were T with hourly and daily mean values of 30.10 °C and 27.50 °C, respectively. The mean hourly and daily values of WD were 145.88° and 1,151.1°, respectively, and the hourly and daily RH values were found to be exhibiting a similar frequency distribution with an hourly mean of 68.91% and daily mean of 77.68%. The maximum hourly and daily values of WS reached 15.80 m/s and 9.45 m/s, respectively, with standard deviations of 2.21 and 1.44, respectively (Figure 4.12).


a
Min:2.0
Max:101.1
Mean:19.2
Std.Dev.:12.3
Frequency
Hourly $PM_{2.5}$ ($\mu g/m^3$)


b
Min:7.0
Max:163.8
Mean:28.9
Std.Dev.:17.3
Frequency
Daily $PM_{2.5}$ ($\mu g/m^3$)


c
Min:0.04
Max:3.4
Mean:0.42
Std.Dev.:0.29
Frequency
MODIS AOT(550 nm)


d
Min:0.9
Max:15.8
Mean:6.41
Std.Dev.:2.21
Frequency
Hourly WS m/s


e
Min:1.1
Max:9.5
Mean:5.1
Std.Dev.:1.4
Frequency
Daily WS m/s

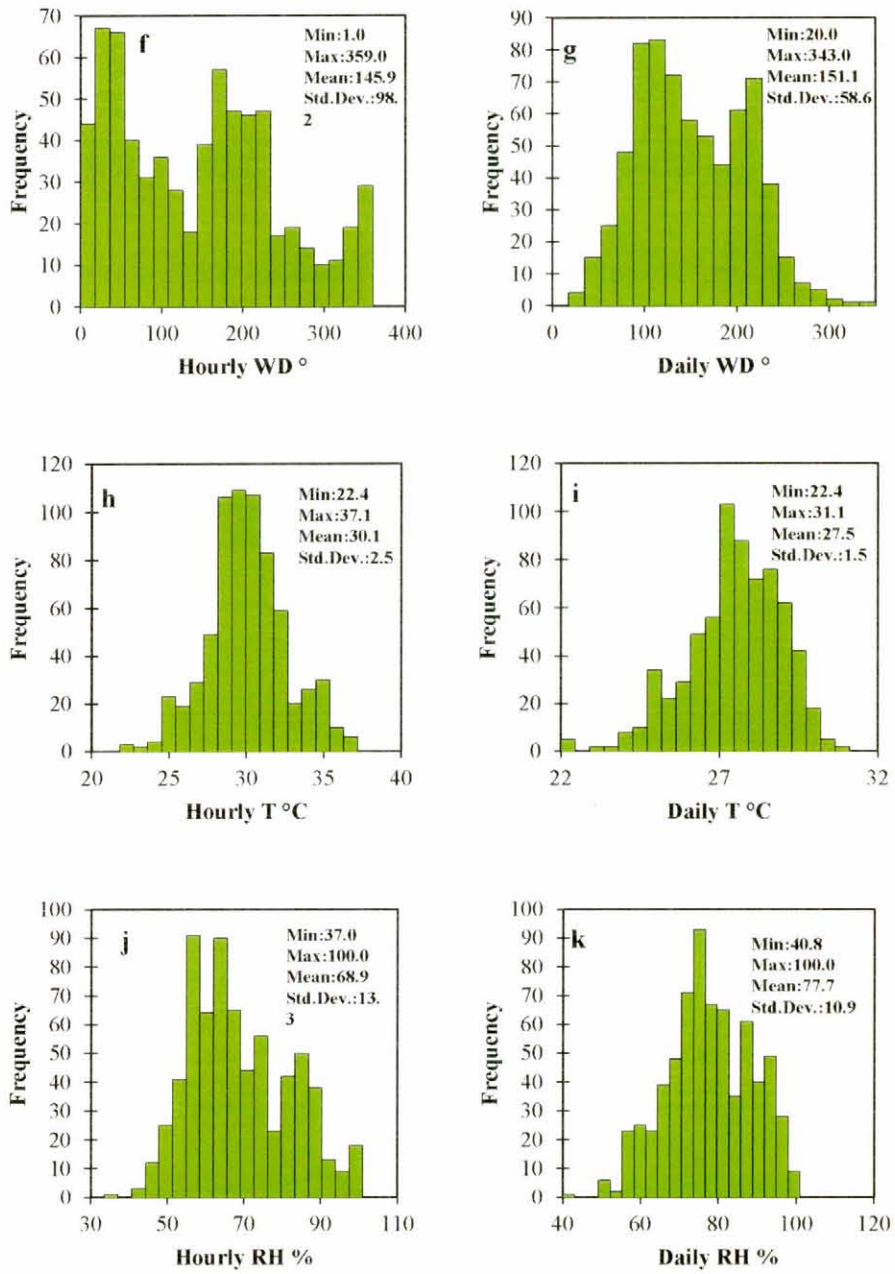


Figure 4.12 : Histograms and Descriptive Statistics of the Regression Model Variables: (a) Hourly PM2.5; (b) Daily PM2.5; (c) AOT; (d) Hourly Wind Speed; (e) Daily Wind Speed; (f) Hourly Wind Direction; (k) Daily Wind Direction

4.3.4.2 Observations of $PM_{2.5}$

Box plots of the hourly data (synchronized with satellite passing passage) are depicted in Figure 4.13 whereas box plots of the daily average $PM_{2.5}$ concentrations ($\mu g/m^3$) for the study year (2013) are given by Figure 4.14. These figures show that the daily mean concentration was higher than the hourly concentration at all stations. Banting, Cheras, and Putrajaya stations, which are located in Klang Valley, recorded the highest daily and hourly $PM_{2.5}$ concentrations. For instance, the maximum hourly $PM_{2.5}$ concentration was 101 $\mu g/m^3$ in Bukit Rembai (Figure 4.13) while the maximum daily $PM_{2.5}$ concentration was 163.7 $\mu g/m^3$ in Putrajaya (Figure 4.14). A possible explanation for this can be that the station is located in Klang Valley, which is one of the leading economic zones in Malaysia (Haberle et al., 2001; Khan et al., 2016). This means that the industrial and motor vehicle emissions are the principal sources of pollutants in this area. Additionally, Malaysia is exposed to trans-boundary pollution caused by aerosols from biomass burning in neighboring countries, particularly Indonesia. Southwesterly winds transport the PM and gases from the forest fires in Indonesia, particularly Sumatra, between July and September, which are the dry season months (Kanniah and Yaso, 2010).

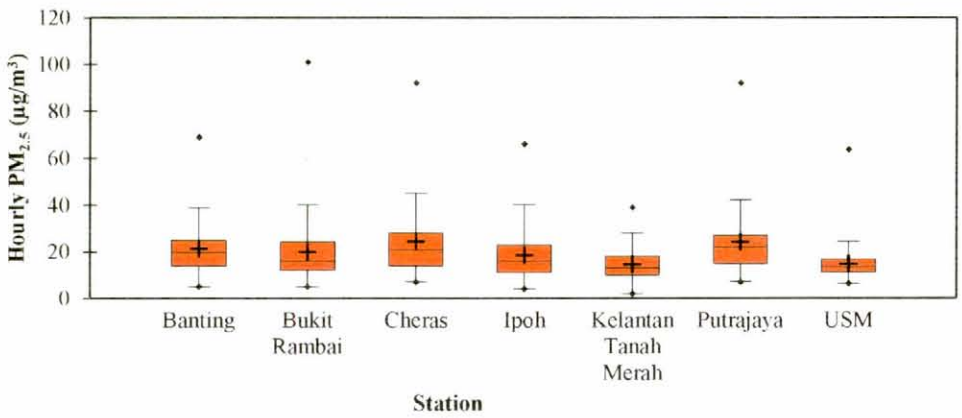


Figure 4.13 : Box Plots of the Hourly $PM_{2.5}$ Concentrations ($\mu g/m^3$) in Peninsular Malaysia in 2013

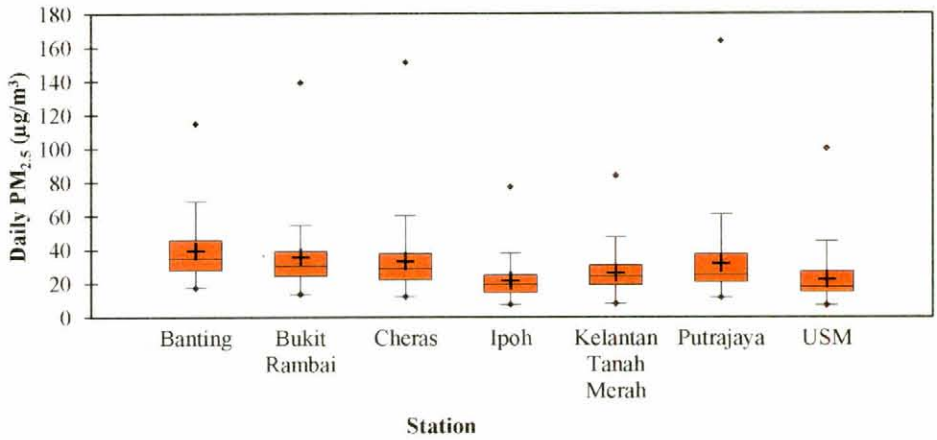


Figure 4.14 : Box Plots of the Daily Concentrations of $PM_{2.5}$ ($\mu g/m^3$) in Peninsular Malaysia in 2013

4.3.5 Regression Models for Estimation of the $PM_{2.5}$ Concentrations

4.3.5.1 Simple Linear Regression Analysis (SLRA)

This study was interested in quantifying the association between MODIS AOT and ground $PM_{2.5}$ concentrations. To this end, the researcher explored the correlations between the daily mean $PM_{2.5}$ concentrations and the ground-observed AOT levels in the year 2013. The relationship between the hourly $PM_{2.5}$ concentrations and the MODIS AOT values retrieved for Peninsular Malaysia at the wavelength of 550 nm were used in the regression modeling.

Figure 4.15 is a scatter plot of MODIS AOT values and the corresponding daily ground-observed concentrations. The figure indicates performance of the SLRM predictive of the daily ground $PM_{2.5}$ concentrations from the MODIS AOT values. In the case of the daily data, the linear regression equation is:

$$y = 24.28x + 11.09;\ R^2 = 0.49 \qquad (4.1)$$

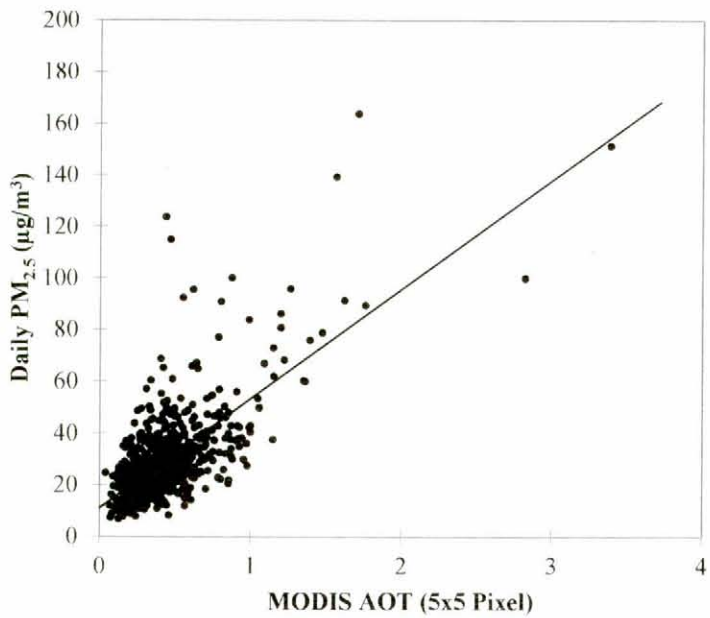


Figure 4.15 : Regression of the Daily $PM_{2.5}$ Concentrations (µg/m³) on MODIS AOT (5x5 Pixels)

In the case of the hourly data (Figure 4.16), the linear regression function takes the form:

$$y = 34.68x + 4.62;\ R^2 = 0.66 \quad (4.2)$$

Various factors can contribute to the AOT–PM relation. Most important of which is the retrieval algorithm employed, which necessitates that the aerosol type and the surface characteristics have to be very close to the real-time circumstances. This is so owing to that the presumptions on the meteorological conditions, surface characteristics, and aerosol properties in the retrieval process can bring about inaccurate MODIS AOD values. Additionally, aerosol retrievals are only possible under the cloud-free sky condition (Zhang and Li, 2015).

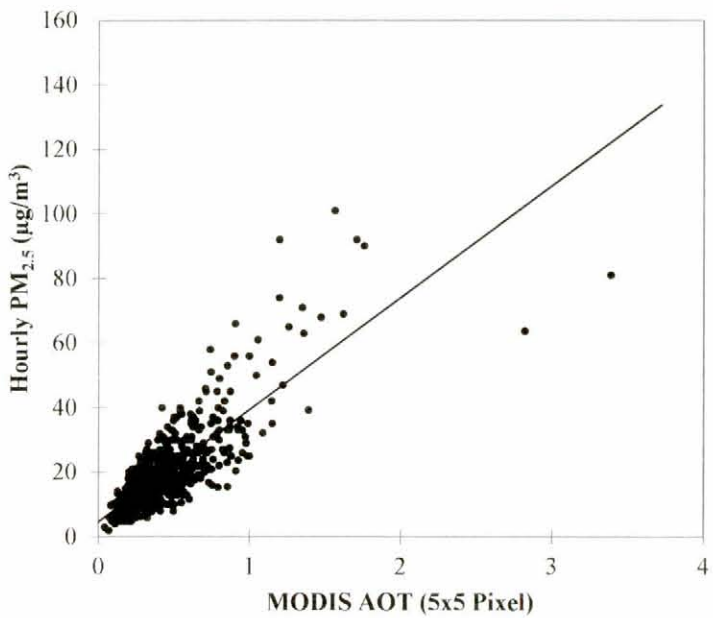


Figure 4.16 : Regression of the Hourly $PM_{2.5}$ Concentrations (µg/m³) on MODIS AOT (5x5 Pixels)

These findings have important implications for development of the regression model predictive of the ground $PM_{2.5}$ concentrations for Peninsular Malaysia. The R^2 values of the daily and hourly MLRMs are 0.49 and 0.66, respectively, which indicate that these two MLRMs can explain 49% and 66% of the variations in the daily and hourly $PM_{2.5}$ concentrations, respectively. However, because of the observed differences in the values of the correlation coefficients, it was concluded that other factors like the RH, vertical structure of the aerosols, meteorological conditions, and temporal and spatial distributions have impacts on the relationship between the $PM_{2.5}$ concentration and the AOT.

4.3.5.2 Multiple Linear Regression Analysis (MLRA)

While most previous studies have focused on exploring the relationships between collected hourly or daily AOT and $PM_{2.5}$ measurements individually, this study took into consideration the daily and hourly $PM_{2.5}$ concentrations to determine the variability between satellite AOT and $PM_{2.5}$ concentrations using SLRMs and MLRMs. Tables 4.8 and 4.9 show summary statistics of the daily and hourly variables, respectively, employed in regression analysis in this study. The MODIS AOT (5x5 pixels) data were used as independent variable in both the SLRM and MLRM because there was one MODIS AOT data record only per day. The mean and standard deviation of MODIS AOT were 0.421 and 0.285, respectively. The large difference between the maximum (3.387) and minimum (0.039) MODIS AOT values compromises sensitivity of the MODIS sensor to measure AOT in terms of intensity. The average

daily $PM_{2.5}$ concentration and associated standard deviation (Table 4.8) and the average hourly $PM_{2.5}$ concentration and associated standard deviation (Table 4.9) in the year 2013 in Peninsular Malaysia were 28.907 ± 17.337 ($\mu g/m^3$) and 19.229 ± 12.247 ($\mu g/m^3$), respectively. These mean concentrations are less than the standard average daily concentration of $PM_{2.5}$ (35 $\mu g/m^3$) stipulated by US EPA.

Table 4.8 : Summary statistics of the daily wind direction (°), wind speed (m/s), relative humidity (%), and temperature (°C)

Statistic	Daily $PM_{2.5}$ (µg/m3)	Daily WS (m/s)	Daily WD (°)	Daily T (°C)	Daily RH (%)	MODIS AOT (5x5 Pixels)
Observations	685	685	685	685	685	685
Minimum	7.00	1.080	20.000	22.380	40.810	0.039
Maximum	163.79	9.450	343.000	31.070	100.00	3.387
Median	25.00	4.880	142.420	27.550	77.580	0.350
Mean	28.907	5.019	151.082	27.486	77.672	0.421
Standard deviation	17.337	1.443	58.526	1.478	10.857	0.285

Table 4.9 : Summary statistics of the hourly wind direction (°), wind speed (m/s), relative humidity (%), and temperature (°C)

Statistic	Hourly $PM_{2.5}$ (µg/m3)	Hourly WS (m/s)	Hourly WD (°)	Hourly T (°C)	Hourly RH (%)	MODIS AOT (5x5 Pixels)
Observations	685	685	685	685	685	685
Minimum	2.0	0.90	1.00	22.380	37.00	0.039
Maximum	101.00	15.800	359.000	37.100	100.00	3.387
Median	16.00	6.100	148.880	29.900	66.370	0.350
Mean	19.228	6.428	145.854	30.046	68.912	0.421
Standard deviation	12.247	2.205	98.159	2.546	13.251	0.285

The hourly (Table 4.9) and average daily (Table 4.8) values of meteorological variables were used to improve the ability of the MLRMs to estimate the mass concentrations of the $PM_{2.5}$ near the surface. Several notable developments over previous works characterized this study and resulted in obtaining an improved representation of the relationship between AOT and the ground-based $PM_{2.5}$ measurements. Besides the meteorological variables, parameters influencing the satellite AOD retrievals can be incorporated into MLRA to improve the correlation coefficient (Gupta et al., 2006).

Table 4.10 and Table 4.11 show the values of the coefficients of correlation among the daily and hourly, respectively, values of MODIS AOT, $PM_{2.5}$, and several meteorological variables. A positive correlation was found between $PM_{2.5}$ and AOT. The hourly data (Table 4.11) manifested a higher correlation (0.808) than the daily data (0.696). These differences may be due to the effects of meteorological and climate

factors. For instance, the estimation of particle mass can be influenced by increase in the RH owing to that the particle size and, thus, mass inflate dramatically as the RH increases (Boyouk et al., 2010; Kumar et al., 2007). Comparatively, the daily data in Table 4.10 unveils that there is a similar pattern in the relationships amongst the daily $PM_{2.5}$ concentration, MODIS AOT, and the studied meteorological variables.

Table 4.10 : Coefficients of correlation among the daily values of MODIS AOT, $PM_{2.5}$, and meteorological variables

Variable	$PM_{2.5}$	WS	WD	T	RH	MODIS AOT
	($\mu g/m^3$)	(m/s)	(°)	(°)	(%)	(5x5 pixels)
Daily $PM_{2.5}$ ($\mu g/ m^3$)	1	-0.258	0.226	-0.138	0.073	0.696
Daily WS (m/s)	-0.258	1	-0.406	0.321	-0.529	-0.193
Daily WD (°)	0.226	-0.406	1	-0.096	0.384	0.113
Daily T (°C)	-0.138	0.321	-0.096	1	-0.257	-0.102
Daily RH (%)	0.073	-0.529	0.384	-0.257	1	0.121
MODIS AOT (5x5 pixels)	0.696	-0.193	0.113	-0.102	0.121	1

Table 4.11 : Hourly Correlation matrix data of MODIS AOT, $PM_{2.5}$ and meteorological variables for study area, Peninsular Malaysia

Variable	$PM_{2.5}$	WS	WD	T	RH	MODIS AOT
	($\mu g/m^3$)	(m/s)	(°)	(°)	(%)	(5x5 pixels)
Hourly $PM_{2.5}$ ($\mu g/ m^3$)	1	-0.113	0.035	0.031	0.134	0.808
Hourly WS (m/s)	-0.113	1	-0.005	-0.054	-0.108	-0.086
Hourly WD (°)	0.035	-0.005	1	-0.101	0.032	0.006
Hourly T (°C)	0.031	-0.054	-0.101	1	-0.209	-0.001
Hourly RH (%)	0.134	-0.108	0.032	-0.209	1	0.135
MODIS AOT (5x5 pixels)	0.808	-0.086	0.006	-0.001	0.135	1

The $PM_{2.5}$ concentration was estimated by first setting it as the dependent variable in the initial modeling process, whose main goal was to define the model structure on daily (Table 4.10) and hourly (Table 4.11) bases. Besides, average data were used in development of various regression models, whose estimates were then compared with ground measurements of the $PM_{2.5}$ concentration. In view of the finding of a high degree of correlation between AOT and $PM_{2.5}$ in the current study, the researcher concludes that reasonably powerful models can be obtained for estimation of the ground $PM_{2.5}$ concentrations.

The AOT data and meteorological parameters were integrated into MLRMs to improve model performance. In this study, the first step was analyzing the linear relation between AOT and $PM_{2.5}$ concentration. Therefore, simple linear regression analysis (SLRA) was employed, once using the daily data records (Figure 4.15) and

once using the hourly data rerecords (Figure 4.16). Afterwards, effects of four meteorological parameters (WS, WD, RH, and T) were investigated using multiple linear regression analysis (MLRA). In the case of the daily data, the generated MLRM has the following equation (Table 4.12):

Daily $PM_{2.5}$ ($\mu g/m^3$) = 48.023+0.49*MODIS AOT-1.56*Daily WS (m/s) + 0.044*Daily (°)-0.62*Daily T (°C) - 0.23*Daily RH (%) (4.3)

However, the MLRM obtained for the hourly records has the form (Table 4.13):

Hourly $PM_{2.5}$ ($\mu g/m^3$) = -1.89+34.37*MODIS AOT - 0.16*Hourly WS (m/s) + 0.004*Hourly WD (°) + 0.19*Hourly T (C°) + 0.026*Hourly RH (%) (4.3)

Table 4.12 : Values of parameters of the multiple linear regression model predictive of the daily $PM_{2.5}$ concentrations ($\mu g/m^3$)

Source	Value	Standard error	t	Pr > |t|	Lower bound (95%)	Upper bound (95%)
Intercept	48.023	10.452	4.595	< 0.0001	27.501	68.545
Daily WS (m/s)	-1.554	0.401	-3.876	0.000	-2.342	-0.767
Daily WD (°)	0.044	0.009	5.026	< 0.0001	0.027	0.061
Daily T (°C)	-0.615	0.330	-1.865	0.063	-1.262	0.032
Daily RH (%)	-0.234	0.051	-4.562	< 0.0001	-0.334	-0.133
MODIS AOT (5x5 Pixels)	40.488	1.636	24.755	< 0.0001	37.277	43.700

Table 4.13 : Values of parameters of the multiple linear regression model predictive of the hourly $PM_{2.5}$ concentrations ($\mu g/m^3$)

Source	Value	Standard error	t	Pr > |t|	Lower bound (95%)	Upper bound (95%)
Intercept	-1.855	4.169	-0.445	0.656	-10.040	6.330
Hourly WS (m/s)	-0.216	0.126	-1.715	0.087	-0.464	0.031
Hourly WD (°)	0.004	0.003	1.468	0.143	-0.001	0.010
Hourly T (°C)	0.187	0.111	1.678	0.094	-0.032	0.405
Hourly RH (%)	0.026	0.022	1.200	0.231	-0.016	0.068
MODIS AOT (5x5 Pixels)	34.363	0.975	35.247	< 0.0001	32.449	36.278

These findings accord with the findings of a number of previous studies which provided evidence on that the meteorological variables that strongly affect the $PM_{2.5}$ mass concentration encircle change in sunlight by effect of season and clouds; vertical

mixing of the air pollutants in the boundary layer of the atmosphere; moisture; temperature; and short-, and long-range transport by winds; and recirculation of the air mass by the local winds (Gupta and Christopher, 2008b). Thereupon, integrating ground-based PM observations and relevant meteorological variables with MODIS (Terra) AOT is vital to PM estimation from space using regression models.

4.3.5.3 Ten-Fold Cross Validation(CV)

To assess whether the models were well specified or not and to ensure that they were not over-fitting, ten-fold CV was conducted after construction of the annual MLRM. The results of CV are shown in Figure 4.17. In CV, daily and hourly models had similar performance. The values of R^2 and RMSE of the model using daily data (Figure 7.17a) were 0.70 and 10.02 ($\mu g/m^3$), respectively, and the values of R^2 and RMSE of the model using hourly data (Figure 4.17b) were 0.79 and 7.16 ($\mu g/m^3$), respectively. However, the MAPE was 25.68% in CV of the hourly model and 31.26% in CV of the daily model.

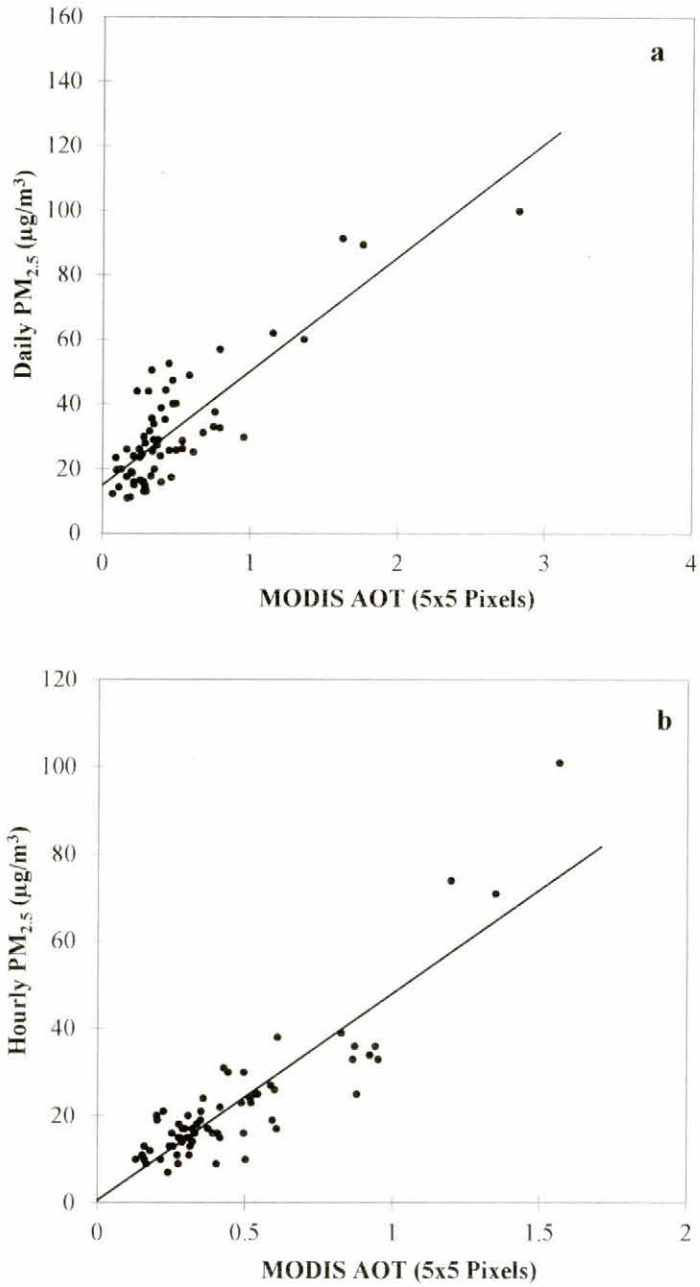


Figure 4.17 : Scatter Plot of the $PM_{2.5}$ Levels against the MODIS AOD Observations in Ten-fold Cross Validation of Regression Models: (a) Daily Data and (b) Hourly Data

According to results of the case study, it is found that the MLRM employing hourly data manifested good performance in estimation of the $PM_{2.5}$ concentrations using satellite-based AOT. In CV of the annual estimates, the R^2 value was 0.79 and the values of RMSE and MAPE were 10.02 $\mu g/m^3$ and 25.68%, respectively. Liu et al. (2019) used the model they developed to assess the hourly surface concentrations of $PM_{2.5}$ at the level of the Advanced Himawari Imager (AHI) pixel. In CV, the values of R^2, RMSE, and the mean prediction error (i.e., error in prediction of $PM_{2.5}$ concentration from a reference $PM_{2.5}$ value) were 0.86, 17.2 $\mu g/m^3$, and 10.3 $\mu g/m^3$ at an hourly scale. Consequently, ten-fold CV of models built using ground-based PM observations and MODIS (Terra) AOT appears to be a satisfactory means of assessing performance of regression models in PM estimation from the space.

4.3.6 Geographically -Weighted Regression Model (GWRM)

The Values of performance metrics (R^2, AIC, and Moran's I) for the GWRM and the global regression model are summarized by Table 4.14. In this study, the GWRM was based on the annual average data of all monitoring stations. Thus, the number of training samples for the GWRM and the global regression model was all of the 685 data records. One commonly employed index to evaluate the spatial autocorrelation is Moran's I statistic. values of Moran's I higher than zero point to positive spatial autocorrelation while values lower than zero indicate negative spatial autocorrelation and values close to zero mean absence of spatial autocorrelation (Hu et al., 2013). In the current study, the values of Moran' I were higher than zero, both in the hourly model (0.104) and the daily model (0.174). The two models were significant (p = 0.000). This indicates presence of positive spatial autocorrelation in each of these two models.

Table 4.14 : Performance metrics for the geographically-weighted regression model (GWRM) and the global regression model

Description	Model	N	AIC	R^2	Moran's I
Hourly	GWRM	685	4559.720	0.70	0.104
	Global Regression Model	685	4656.837	0.65	
Daily	GWRM	685	5304.802	0.59	0.174
	Global Regression Model	685	5404.347	0.48	

4.3.7 Descriptive Statistics for the Seasonal Data of the Study Variables

Table 4.15 summarizes the effect of seasonal variation on $PM_{2.5}$, AOT, WS, WD, RH, and T. The AOT (unitless) and $PM_{2.5}$ mass concentration ($\mu g/m^3$) showed high seasonal variability. The mean AOT value during the southwest monsoon was higher (0.52) than that during the northeast monsoon (0.37). The maximum AOT value

during the southwest monsoon was 3.39 while the maximum AOT value during the northeast monsoon was 1.36. On the other hand, the highest daily and hourly mean mass concentrations of $PM_{2.5}$ (24.32 μg/m^3 and 37.67 μg/m^3, respectively) were recorded during the southwest monsoon. The mean hourly $PM_{2.5}$ mass concentration during the northeast monsoon was 16.82 μg/m^3 and the daily mean $PM_{2.5}$ mass concentration was 25.27 μg/m^3. These findings accord with results obtained by Kanniah et al. (2014), who researched into the spatial distribution of AOT over Peninsular Malaysia on an annual basis from 2000–2009 and reported that the maximum values lied in the range of 1.6–2.7 during the dry season.

Table 4.15 : Descriptive statistics for the seasonal data of the study variables

Northeast monsoon	Description	Mean	Max.	Min.	Median	SD
$PM_{2.5}$ (μg/m^3)	Hourly	16.82	63.00	3.00	15.00	8.04
	Daily	25.27	60.11	7.00	24.40	9.22
RH (%)	Hourly	65.76	99.00	37.00	63.00	11.83
	Daily	77.30	99.29	55.33	75.50	10.38
WD (°)	Hourly	125.56	359.00	1.00	96.50	95.65
	Daily	146.33	343.00	33.33	135.83	61.32
WS (m/s)	Hourly	5.9971	13.20	1.80	5.99	1.99
	Daily	4.919	8.13	2.13	4.94	1.42
Temp. (°C)	Hourly	30.33	36.30	22.40	30.00	2.39
	Daily	27.54	30.44	22.40	27.55	1.19
AOT (unitless)	Once a day	0.37	1.36	0.04	0.31	0.20
Southwest monsoon	**Description**	**Mean**	**Max.**	**Min.**	**Median**	**SD**
$PM_{2.5}$ (μg/m^3)	Hourly	24.32	101.00	2.00	19.00	17.39
	Daily	37.67	163.79	7.55	31.39	25.28
RH (%)	Hourly	70.02	100.00	44.00	67.00	12.48
	Daily	76.13	100.00	50.10	76.60	11.35
WD (°)	Hourly	162.56	358.00	4.00	171.00	94.78
	Daily	159.73	299.29	38.75	156.83	56.15
WS (m/s)	Hourly	6.78	15.80	0.90	6.50	2.410
	Daily	5.16	9.45	1.08	4.92	1.53
Temp. (°C)	Hourly	29.77	36.30	23.40	29.55	2.09
	Daily	27.40	30.21	22.38	27.53	1.60
AOT (unitless)	Once a day	0.52	3.39	0.07	0.42	0.40

Meteorologically, the hourly and daily temperatures were found to be undergoing low seasonal variations and there were no noticeable differences between the WSs during the southwest and the northeast monsoons. The hourly WS was 15.80 m/s, which was the maximum value during the year. During the observation period (2013), the RH was relatively high, with hourly and daily mean values of 67.76% and 77.30% during the northeast monsoon, respectively. During the southwest season, the mean hourly and daily values were 70.02% and 76.13%, respectively (Table 4.15).

4.3.7.1 Temporal and Seasonal Variability of $PM_{2.5}$ and AOT

Variations in the one-year MODIS AOT observations and the daily average and hourly $PM_{2.5}$ concentrations (observations and predictions) over Peninsular Malaysia from January 2013 to December 2013 are depicted in figures 4.18-4.24 for Banting, Putrajaya, Bukit Rambai, Cheras, Ipoh, Kelantan Tanah Merah, and USM stations, respectively. Annually, the highest AOT value (> 3.0) was recorded in Cheras station (Figure 4.21). It is possible that the area around this station was exposed to a higher aerosol loading during June than during other months due to the impact of haze during biomass burning. It is noticed that the AOT values and the daily and hourly $PM_{2.5}$ concentrations peaked at all stations during June. It seems possible that these results are due to the southwest monsoon, which prevails from June to September every year and during which haze episodes take place in Peninsular Malaysia due to biomass burning in Sumatra, Indonesia (Azmi et al., 2010).

The observations and the prediction results followed the same pattern of seasonal variability during the period of study at all stations. It was found that there is agreement between the hourly and daily $PM_{2.5}$ concentrations at Kelatan Tanah Marh station (Figure 4.23). Likewise, there was match between the observations and predictions at USM station (Figure 4.24). There were differences between the hourly and daily concentrations of $PM_{2.5}$ at Cheras (Figure 4.21) and Ipoh stations (Figure 4.22). However, the observed differences were only slight. Figures 4.18-4.24 disclose that the high concentrations were recorded during the southwest monsoon months.

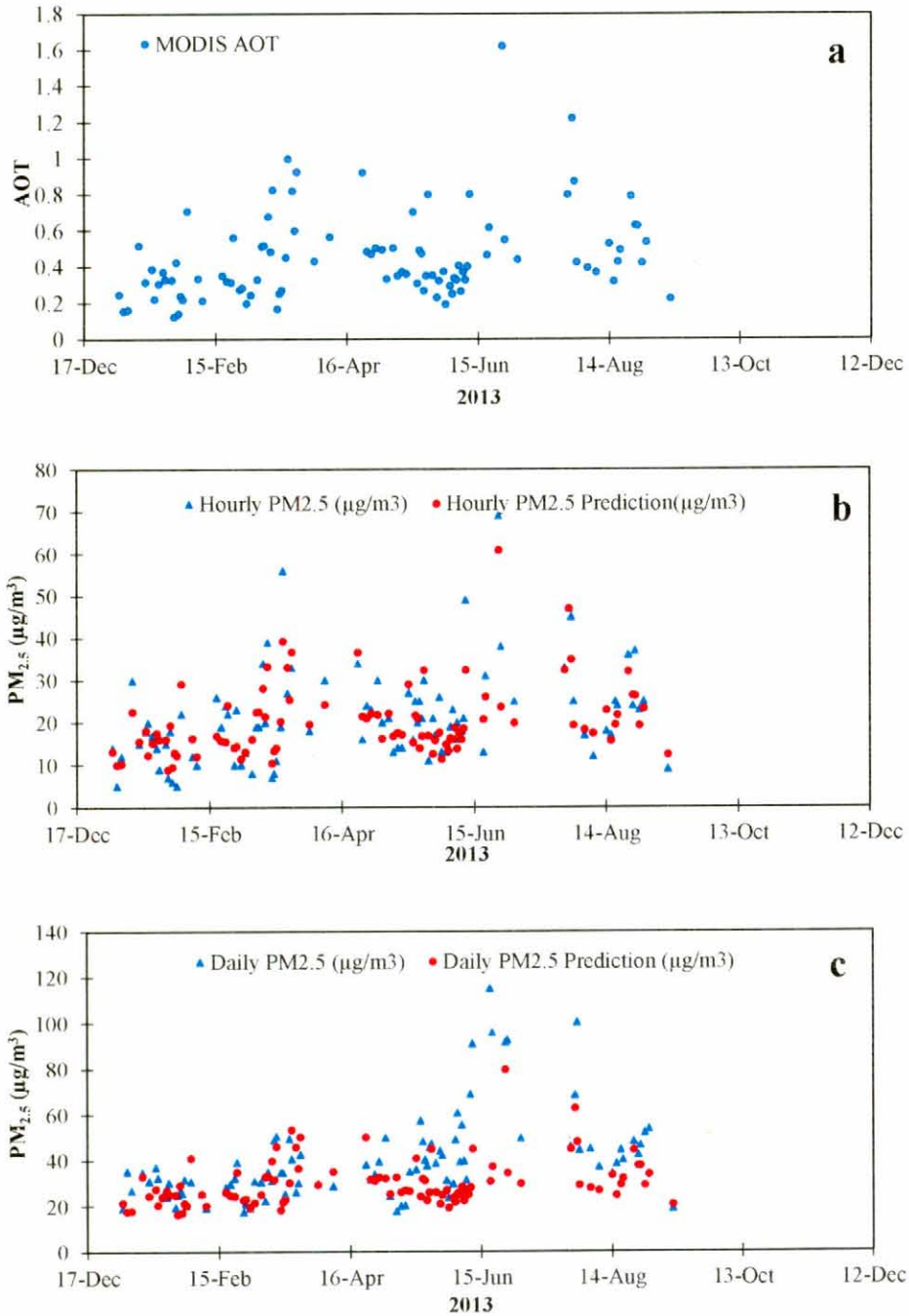


Figure 4.18 : Temporal Variations in AOT and $PM_{2.5}$ Concentrations at Banting Station in 2013: (a) Temporal Variations in AOT; (b) Hourly Observations and Predictions of $PM_{2.5}$ Concentrations; (c) Daily Observations and Predictions of $PM_{2.5}$ Concentrations

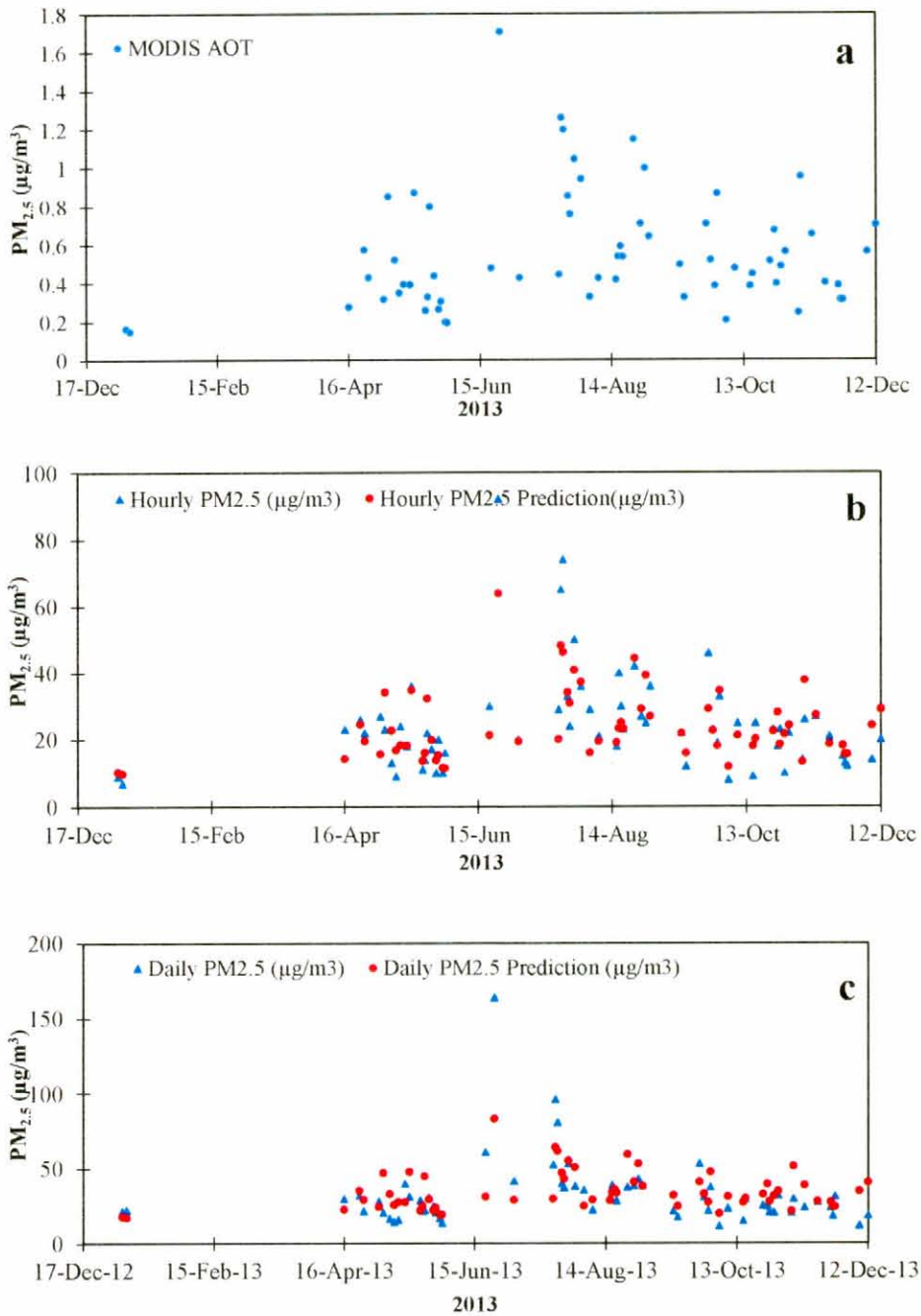


Figure 4.19 : Temporal Variations in AOT and $PM_{2.5}$ Concentrations at Purajaya Station in 2013: (a) Temporal Variations in AOT; (b) Hourly Observations and Predictions of $PM_{2.5}$ Concentrations; (c) Daily Observations and Predictions of $PM_{2.5}$ Concentrations

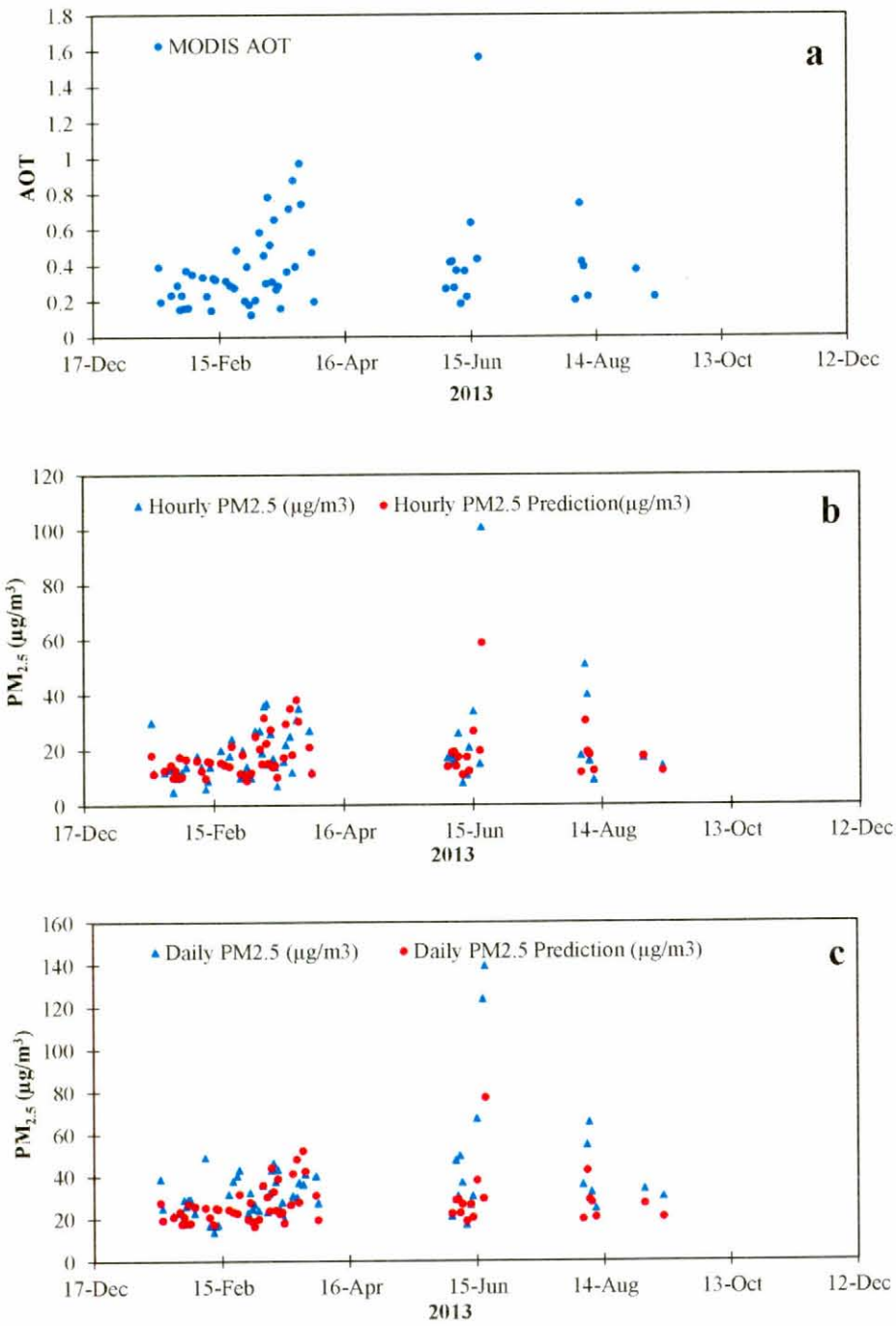


Figure 4.20 : Temporal Variations in AOT and $PM_{2.5}$ Concentrations at Bukit Rambai Station in 2013: (a) Temporal Variations in AOT; (b) Hourly Observations and Predictions of $PM_{2.5}$ Concentrations; (c) Daily Observations and Predictions of $PM_{2.5}$ Concentrations

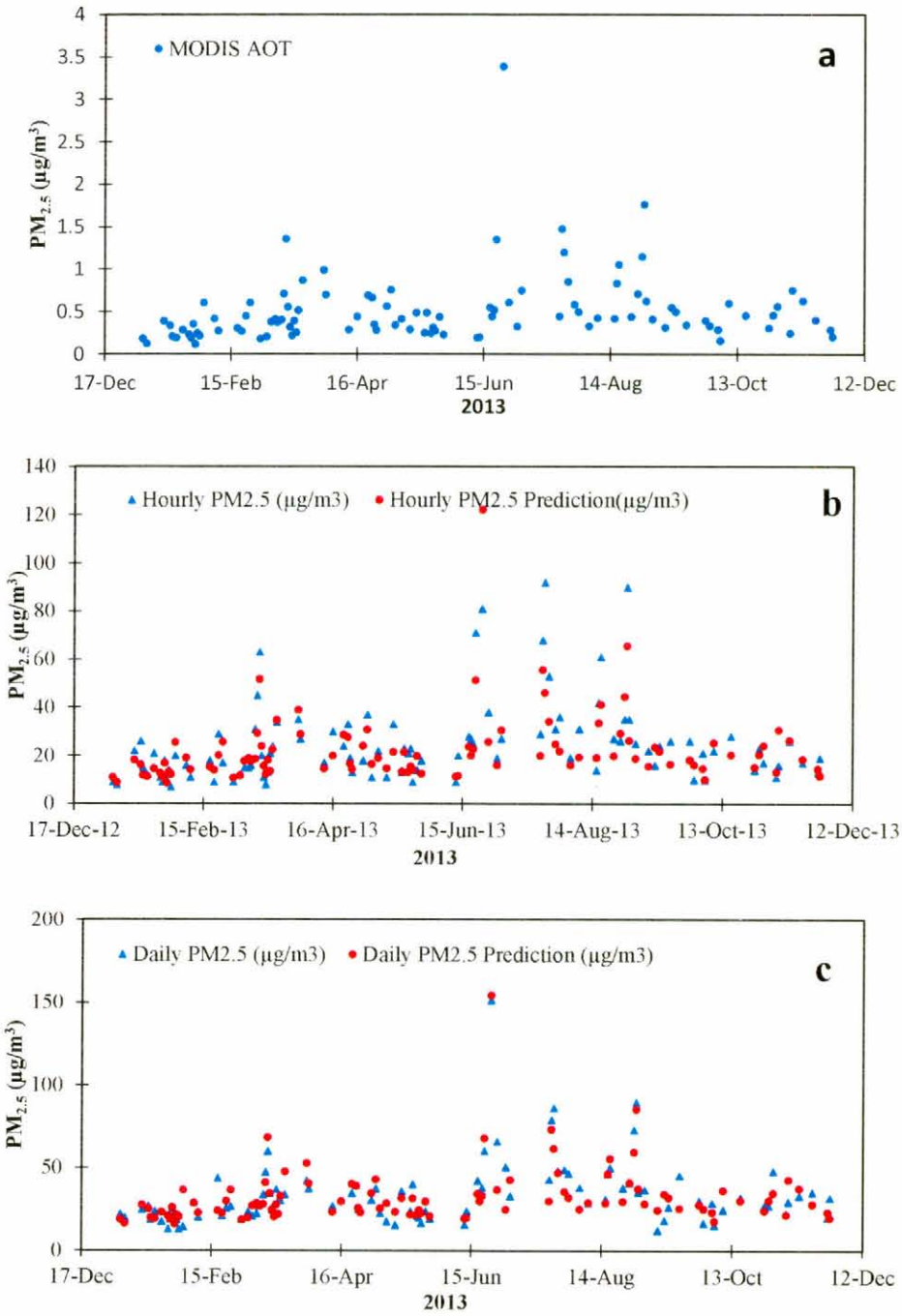


Figure 4.21 : Temporal Variations in AOT and $PM_{2.5}$ Concentrations at Cheras Station in 2013: (a) Temporal Variations in AOT; (b) Hourly Observations and Predictions of $PM_{2.5}$ Concentrations; (c) Daily Observations and Predictions of $PM_{2.5}$ Concentrations

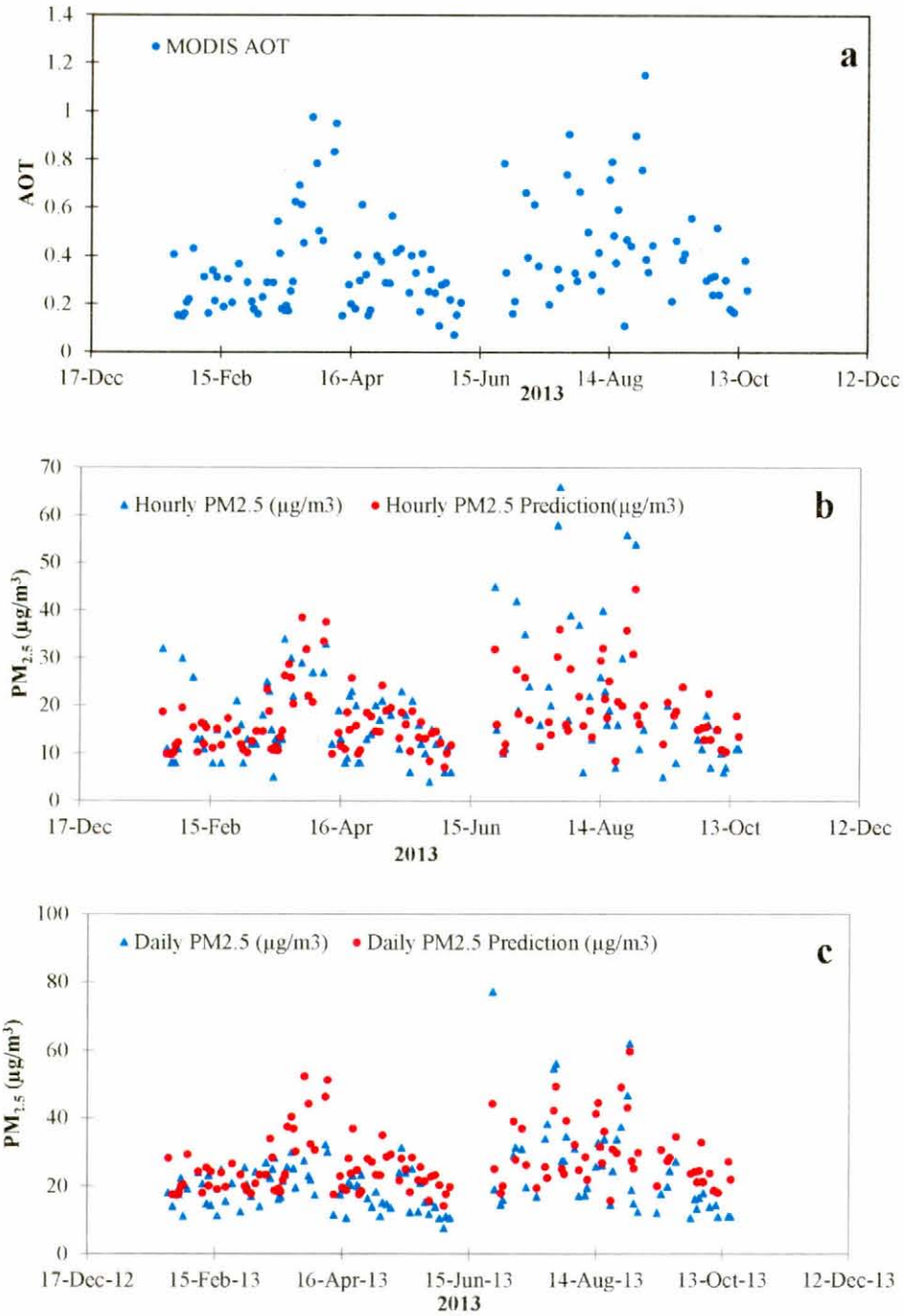


Figure 4.22 : Temporal Variations in AOT and $PM_{2.5}$ Concentrations at Ipoh Station in 2013: (a) Temporal Variations in AOT; (b) Hourly Observations and Predictions of $PM_{2.5}$ Concentrations; (c) Daily Observations and Predictions of $PM_{2.5}$ Concentrations

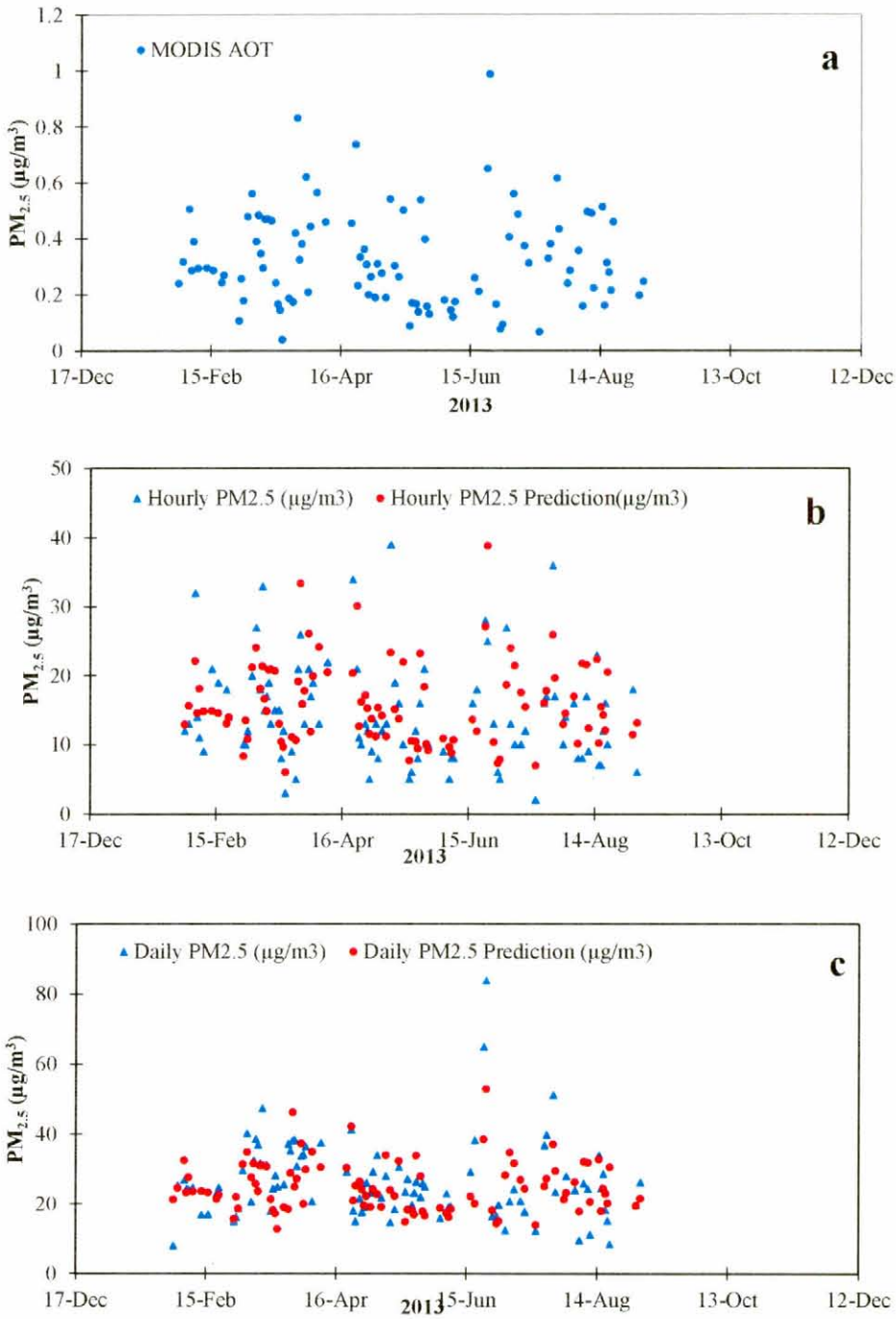


Figure 4.23 : Temporal Variations in AOT and $PM_{2.5}$ Concentrations at Kelantan Tanah Merah Station in 2013: (a) Temporal Variations in AOT; (b) Hourly Observations and Predictions of $PM_{2.5}$ Concentrations; (c) Daily Observations and Predictions of $PM_{2.5}$ Concentrations

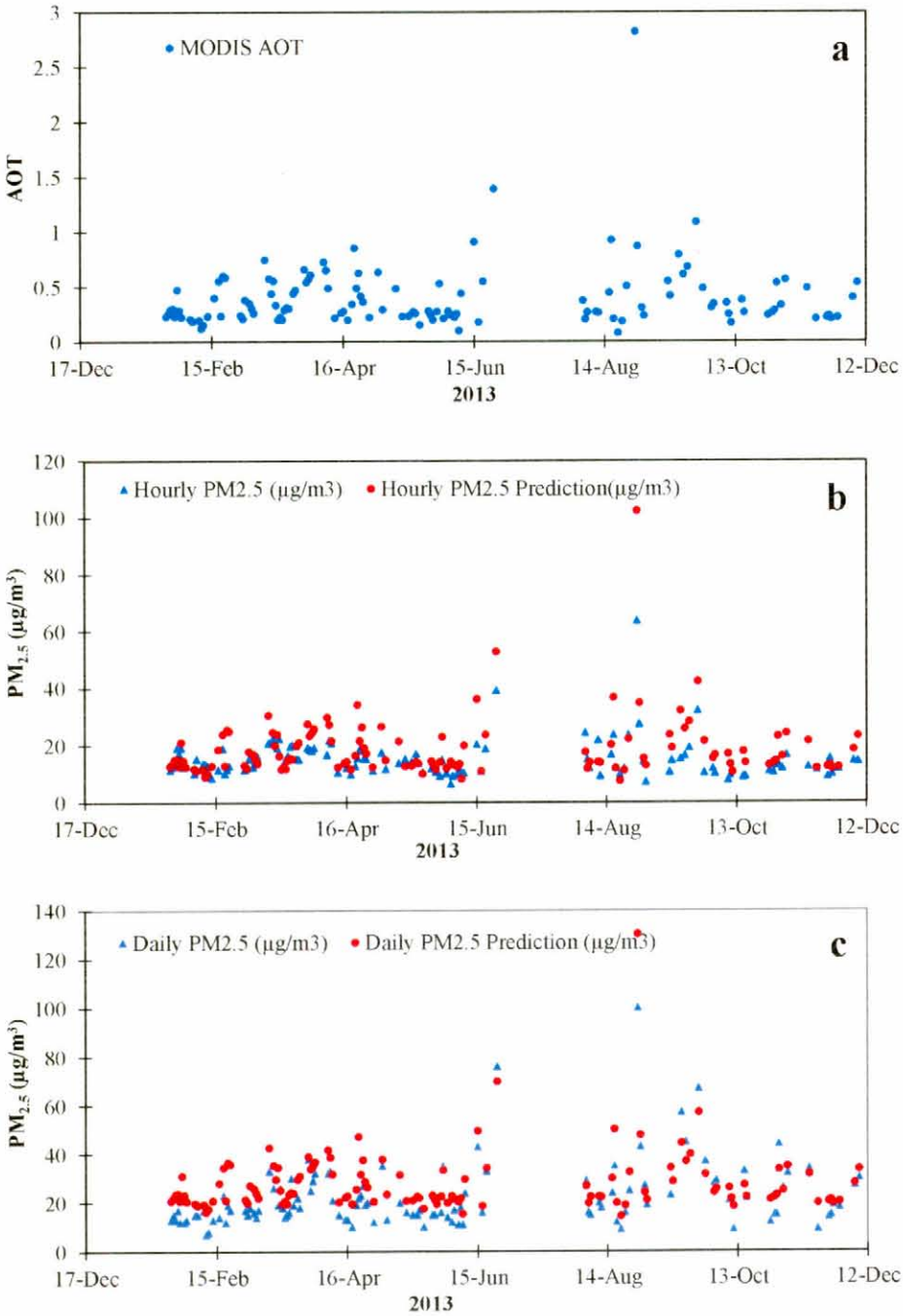


Figure 4.24 : Temporal Variations in AOT and $PM_{2.5}$ Concentrations at USM Station in 2013: (a) Temporal Variations in AOT; (b) Hourly Observations and Predictions of $PM_{2.5}$ Concentrations; (c) Daily Observations and Predictions of $PM_{2.5}$ Concentrations

A possible explanation for the increase in the $PM_{2.5}$ concentrations in June at most stations is the smoke plumes of fires in Riau Province, Indonesia, which traveled towards Singapore and southern Malaysia causing heavy pollution. Satellite datasets were highly useful in indicating the temporal and spatial variations in the aerosols and carbon monoxide (CO) concentration. In general, most of the winds in the region are westerly during June but in June 2013 they blew eastwards (Vadrevu et al., 2014).

In addition, the shape of the temporal AOT curve for the MODIS matches that of the observed and predicted concentrations of $PM_{2.5}$ extremely well. Table 4.16 lists the percentage error values in predictions of the hourly and daily $PM_{2.5}$ concentrations for all studied monitoring stations. The lowest hourly mean percent error was 2.83% in Petrajaya whereas the highest percentage error (23.30%) was reported for USM station. The lowest daily mean percent error (1.6%) was reported for Cheras station while the highest percentage error (24.97%) was reported for USM station.

Table 4.16 : Percentage error values in predictions of the hourly and daily $PM_{2.5}$ concentrations at all stations

Station	Statistic	Hourly $PM_{2.5}$ (µg/m3)	Hourly $PM_{2.5}$ Prediction	Percent Error %	Daily $PM_{2.5}$ (µg/m3)	Daily $PM_{2.5}$ Prediction	Percent Error %
Banting	No. of observations	96	96		96	96	
	Minimum	5.00	8.91	87.20	17.43	16.32	6.37
	Maximum	69.00	60.80	11.88	115.00	79.58	30.80
	Range	64.00	51.89		97.57	63.26	
	Median	19.50	17.83		34.83	27.20	
	Mean	21.31	20.27	4.88	39.48	30.17	23.58
Putrajaya	No. of observations	69	69		69	69	
	Minimum	7.00	9.84	40.57	11.68	17.45	49.40
	Maximum	92.00	63.85	30.59	163.79	83.30	49.14
	Range	85.00	54.01		152.11	65.85	
	Median	22.00	21.16		25.29	31.26	
	Mean	24.20	23.52	2.83	31.77	34.13	7.43
Ipoh	No. of observations	125	125		125	125	
	Minimum	4.00	7.14	78.50	7.55	14.16	87.55
	Maximum	66.00	44.54	32.51	77.20	59.75	22.60
	Range	62.00	37.40		69.65	45.60	
	Median	16.00	15.61		19.50	24.49	
	Mean	18.52	17.63	4.81	21.69	26.95	24.25
Bukit Rambai	No. of observations	60	60		60	60	
	Minimum	5.00	8.96	79.20	13.75	16.38	19.13
	Maximum	101.00	58.86		139.33	77.22	44.58
	Range	96.00	49.90		125.58	60.84	
	Median	16.00	15.94		30.61	24.89	
	Mean	19.90	18.05	9.29	35.61	27.46	22.89
Cheras	No. of observations	105	105		105	105	
	Minimum	7.00	8.66	23.71	12.24	16.02	30.88
	Maximum	92.00	122.08	32.69	151.38	154.29	1.92
	Range	85.00	113.42		139.14	138.27	
	Median	21.00	19.00		28.92	28.62	
	Mean	24.48	22.28	8.99	33.16	32.61	1.65
Kelantan Tanah Merah	No. of observations	97.00	97.00		97.00	97.00	
	Minimum	2.00	5.99	199.50	8.00	12.76	59.50
	Maximum	39.00	38.87	0.33	83.92	52.84	37.04
	Range	37.00	32.88		75.92	40.08	
	Median	13.00	14.87		24.29	23.58	
	Mean	14.51	15.97	10.06	26.17	24.93	4.73
USM	No. of observations	133	133		133	133	
	Minimum	6.26	7.43	18.69	7.00	14.51	107.28
	Maximum	63.68	102.42	60.83	100.00	130.32	30.32
	Range	57.42	94.99		93.00	115.80	
	Median	13.50	14.92		18.00	23.65	
	Mean	14.81	18.26	23.30	22.18	27.72	24.97

These results indicate that there are further variations in the AOT-$PM_{2.5}$ relation that can not be explained by the atmospheric structure. One reasoning is that the $PM_{2.5}$ concentration at the surface may not be representative of the whole PBL column. Other factors like aerosol composition and aerosol size distributions also cause variability in the mass-optical extinction ratio. Likewise, the temporal variations in AOT and $PM_{2.5}$ concentrations are strongly affected by the season as shown in figures 4.18-4.24, where the highest values were recorded in June.

4.3.7.2 Spatial and Seasonal Distributions of $PM_{2.5}$ Concentrations

Figures 4.25, 4.26, and 4.27 represent all the data, data for the southwest monsoon, and data for the northeast monsoon, respectively. The results of estimation by using Spatial Kriging (SK) appear to be unrealistic and seem to be heavily dependent on the interpolation scheme. The smooth changes over rather long distances and the massive presence of high values (the dark orange areas) reflect the instrumental character of SK. Due to the sparsely-located monitoring stations, the ground $PM_{2.5}$ concentrations at many locations had to be estimated by interpolating the values of the stations that are spatially far apart. Such estimates can hardly represent the real distribution of the ground $PM_{2.5}$ concentrations across space.

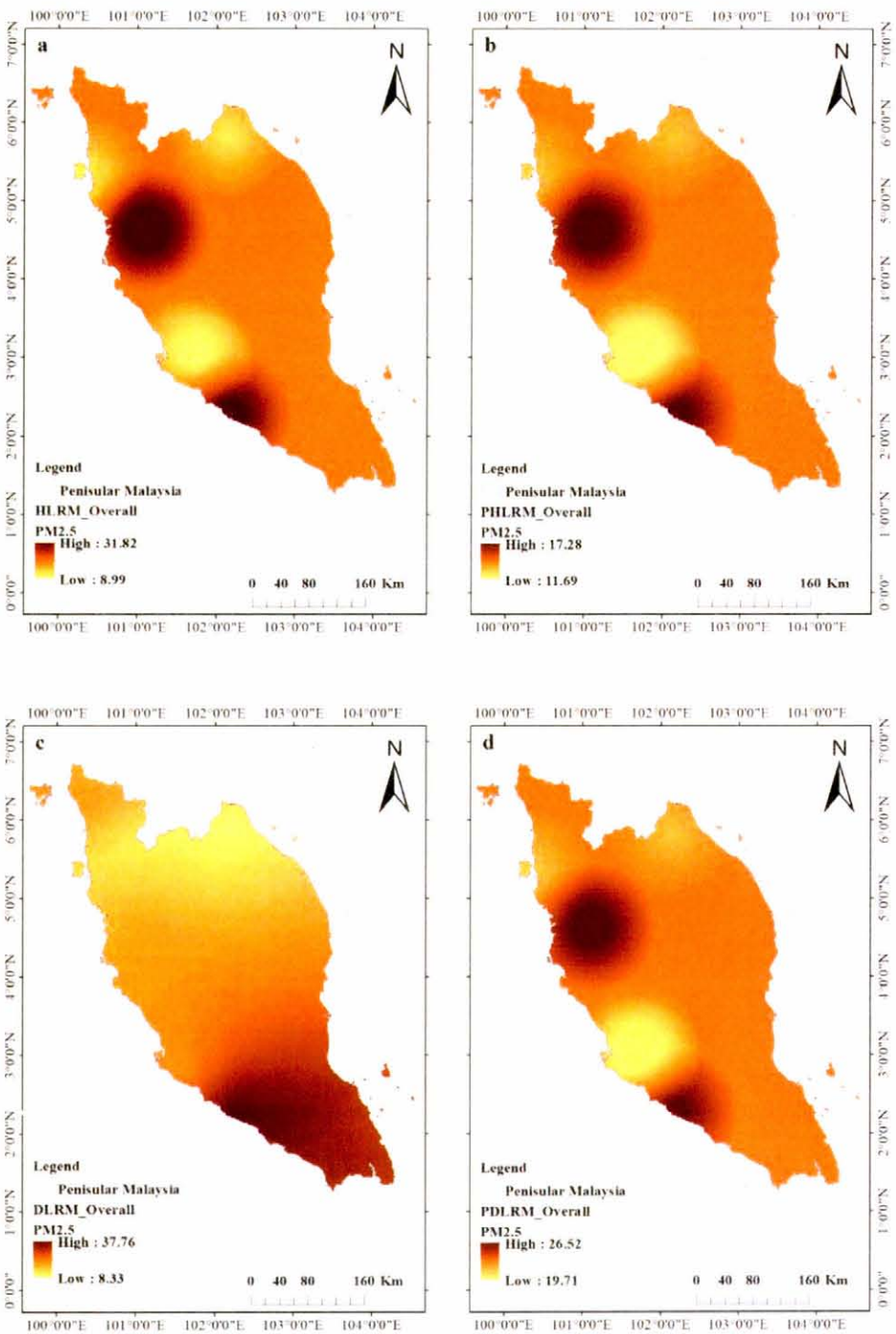


Figure 4.25 : Maps of Ground-level Distribution of $PM_{2.5}$ during the Study Period (2013): (a) Hourly Observations; (b) Hourly Prediction; (c) Daily Average of Observations; (d) Daily Average of Predictions Using Spatial Kriging

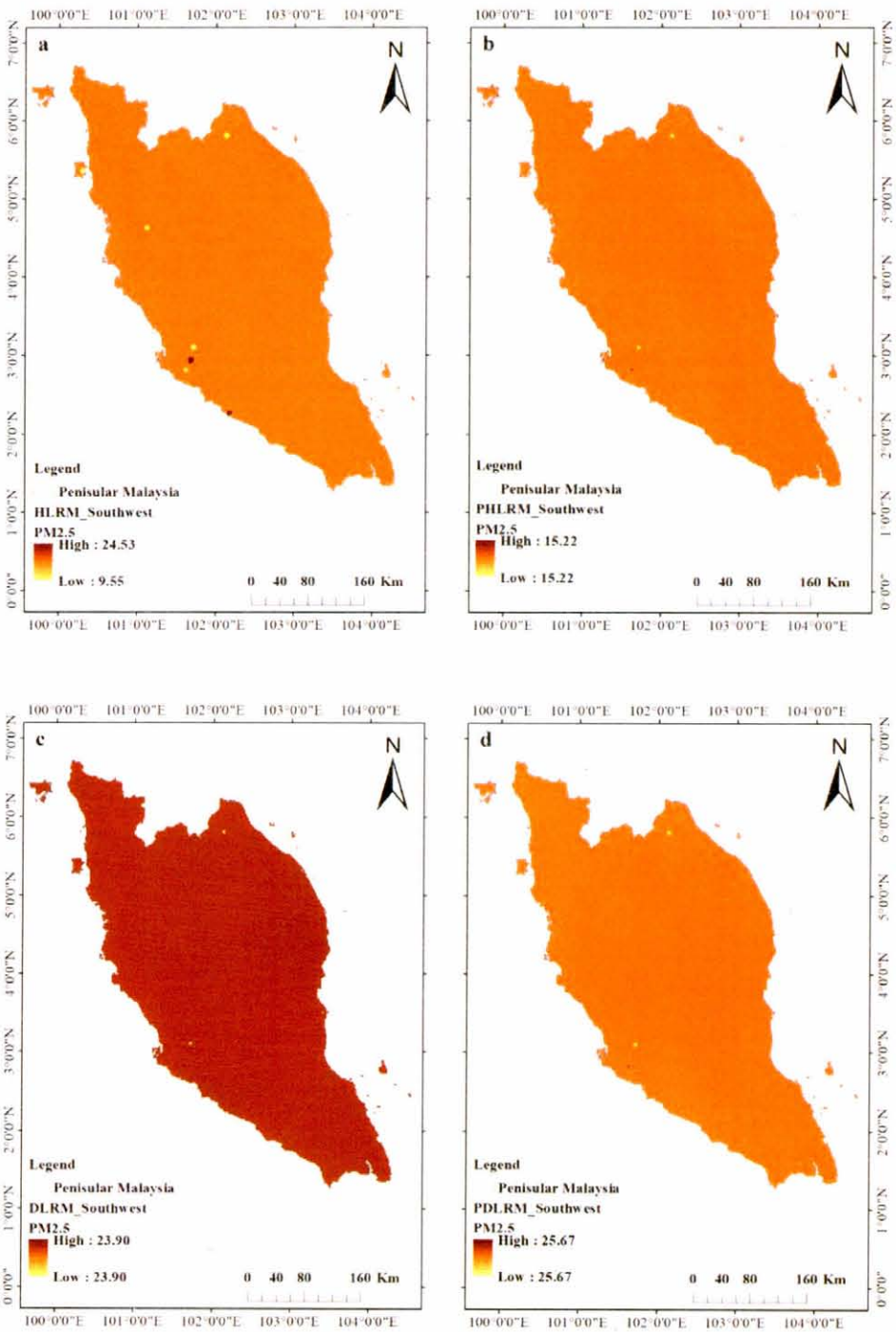


Figure 4.26 : Maps of Ground-level Distribution of $PM_{2.5}$ during the Southwest Monsoon: (a) Hourly Observations; (b) Hourly Prediction; (c) Daily Average of Observation); (d) Daily Average of Predictions Using Spatial Kriging

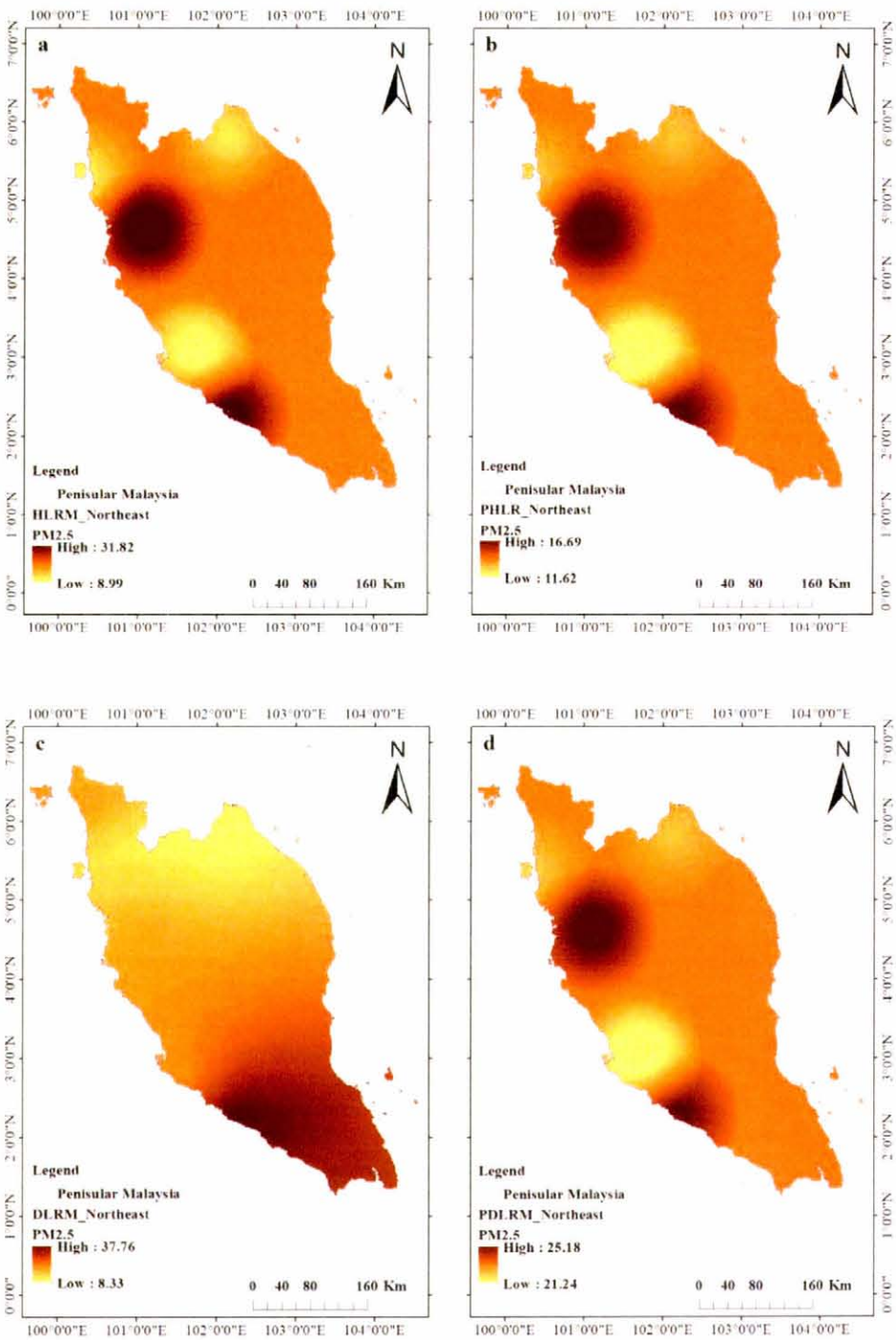


Figure 4.27 : Maps of Ground-level Distribution of $PM_{2.5}$ during the Northeast Monsoon: (a) Hourly Observations; (b) Hourly Prediction; (c) Daily Average of Observation; (d) Daily Average of Predictions Using Spatial Kriging

Figure 4.26 points out that the predictions of the LRM do not match the actual concentrations in spatial distribution in the southwest monsoon. To the contrary, figures 4.27a and 4.27b underline that the predictions of $PM_{2.5}$ were identical for the daily average observations and hourly average predictions, respectively. These differences can be explained based on the values of R^2, MSE, RMSE, and MAPE (Table 4.17). Spatial Kriging was found to be relatively strong estimator with a very low RMSE of 5.01 $\mu g/m^3$. Further, its estimates had low MSE and MAPE (Table 4.17).

Table 4.17 : Values of measures of model performance for the northeast monsoon, southwest monsoon, and overall year (2013) data

Accuracy measures	HLRM Southwest	DLRM Southwest	HLRM Northeast	DLRM Northeast	HLRM Overall	DLRM Overall
Observations	222	222	276	276	685	685
R^2	0.66	0.55	0.61	0.28	0.65	0.48
MSE	102.11	291.25	25.10	61.22	52.17	155.37
RMSE	10.11	17.07	5.01	7.82	7.22	12.46
MAPE	36.17	37.62	24.55	29.08	28.97	32.89

The results of this study unveil that the spatial distributions of the hourly and daily $PM_{2.5}$ concentrations are high on the southwest coast in both monsoons. These results agree with findings of Kanniah et al. (2014) who reported that the AOT assumes its maximal values during the dry season, especially along the west coast while during the wet season the AOT takes its lowest levels. The seasonal spatial characteristics of distribution of the $PM_{2.5}$ concentrations in Peninsular Malaysia can be explained as the combined impact of climatic conditions and local emissions.

4.3.7.3 The MODIS AOT and Quality Contrast

Besides MODO4-L2 collection 6, the AOT was used to predict $PM_{2.5}$ concentrations based on hourly data and linear regression modeling (Figure 4.16). The images depicting the regression modeling results of this study show predictions of $PM_{2.5}$ concentrations for few days of the year 2013 (Figures 4.28-4.35). The figures show missing AOT ground pixel data around cloudy areas, which indicates that $PM_{2.5}$ measurements can be taken in the clear sky conditions during satellite overpass. The satellite images taken in 24 June 2013 (Figure 4.31) show highly different colors that distinguish heavy haze and cloud. Generally, clouds are white and haze appears gray.

On 24 June 2013, the average $PM_{2.5}$ concentration ranged from 181.45 $\mu g/m^3$ in Banting to 39.23 $\mu g/m^3$ in USM, which are higher concentrations than those in non-haze days (Table 4.18). The high concentrations of $PM_{2.5}$ in that day are likely related to locations of the stations (Figure 3.2), such as Banting, Putrajaya, and Cheras being located in the areas vulnerable to haze (Figure 4.31a). In that particular day (24 June 2013), though there is clear indication of high haze density, the AOT is not recognizable, indicating the limited ability of the MODIS sensor to retrieve AOT

during dense haze days. But MODIS can still assist via true color image in visualizing the haze density which proved to be impossible via MODIS AOT (Figure 4.31a).

Table 4.18 lists the ground average daily and hourly measurements of $PM_{2.5}$ concentrations at the monitoring stations under study in Peninsular Malaysia. The figures in this table for a given day can be compared with the data in Figures 4.28–4.35, which present the ground predictions of distribution of $PM_{2.5}$. The daily and hourly observations were within the bounds of the concomitant predictions that were obtained from LRM. For example, in 23 May 2013, the hourly $PM_{2.5}$ concentration ranged from 8.00 $\mu g/m^3$ in Bukit Rambii and Cheras to 18.00 $\mu g/m^3$ in USM. Meantime, the daily $PM_{2.5}$ concentration fell in the range of 16.33 $\mu g/m^3$ in USM to 47.04 $\mu g/m^3$ in Banting. These figures comply well with the predictions of the $PM_{2.5}$ concentrations summarized by Figure 4.30. These calculation outcomes underscore the potentials of the suggested models for predicting the patterns of distribution of $PM_{2.5}$ concentrations in Peninsular Malaysia.

The MODIS-Terra true-color image and AOT data for 24 June 2013 were used to obtain the spatial extent and intensity of the haze episode. The true-color image in Figure 4.31 shows that gray haze plumes appeared over the west coast and obscured the surface of most of central Peninsular Malaysia. The MODIS AOT data appearing in Figure 4.31 uncover that strong aerosol loadings, with high AOT values greater than 4.0, were associated with most of the haze plumes. However, the results of AOT are recorded at the boundary of haze plume. The AOT data disappear in the mid of the haze plume. It is difficult to provide an explanation for this result, but it may be related to sensitivity of MODIS to retrieve AOT in the area affected by the cloud and high intensity of haze.

There are medium to high correlations between the ground $PM_{2.5}$ concentrations and the AOT values drawn from RS tools like MISR and MODIS. The so derived AOT values have appreciable ability to predict the ground $PM_{2.5}$ concentrations (You et al., 2015). These correlations exist in spite of the vertical mismatch between the total column aerosol, as measured by AOD, and the ground $PM_{2.5}$ concentration; the level of interest to the health studies, and the temporal mismatch between the 24-hour average $PM_{2.5}$ concentration and the daytime (usually a single snapshot) AOT (Paciorek and Liu, 2009).

One of the major criticisms of the ability of the satellite-based methods to steadily provide reliable AOD and, therefore, $PM_{2.5}$ estimations is their weak ability to deal with the challenge posed on them by the clouds (Christopher and Gupta, 2010). Some studies (e.g., Engel-Cox et al. (2004)) used MOD021KM-Level-1B to obtain red-green-blue (RGB) true color images. On the other hand, it is argued that $PM_{2.5}$ plays a big role in the formation of haze (Watson, 2002). The most recent haze that struck most of the parts of Malaysia took place in 2013. It diminished horizontal visibility across the country because the thick smoke blanketed Peninsular Malaysia (Chang, 2016).

Table 4.18 : $PM_{2.5}$ concentration (ug/m^3) hourly and daily at the station used to estimation ground-level of PM2.5

Date (2013)	Hourly (H) & Daily (D)	$PM_{2.5}$ ($\mu g/m^3$) at the stations						
		Banting	Bukit Rambii	Cheras	Ipoh	Putrajaya	Kelantan Tanah Merah	USM
28 Jan.	H	16.12	13.00	9.00	3.00		15.00	9.00
	D	25.96	28.91	25.74	18.27		56.71	19.05
20 Mar.	H	5.32	4.00	24.00	9.00		16.00	7.00
	D	40.47	30.95	30.48	23.30		24.87	16.33
23 May	H	14.00	8.00	8.00	11.00	18.00	16.00	18.00
	D	47.08	23.62	16.52	20.91	22.30	26.26	16.33
24 Jun.	H	113.00		81.00	93.00	92.00	59.00	30.00
	D	181.45		151.38	77.20	163.79	83.92	39.23
18 Aug.	H	25.00	41.00	61.00	14.00	21.00	12.00	19.00
	D	44.45	41.00	50.00	33.79	36.64	18.19	12.15
28 Sep.	H	11.00		33.00	13.00	23.00		7.00
	D	18.81		30.21	10.58	30.92		10.37
21 Oct.	H			5.00		16.00		25.00
	D			8.78		11.53		9.80
26 Dec.	H			37.00		35.00		13.00
	D			30.38		20.63		12.92

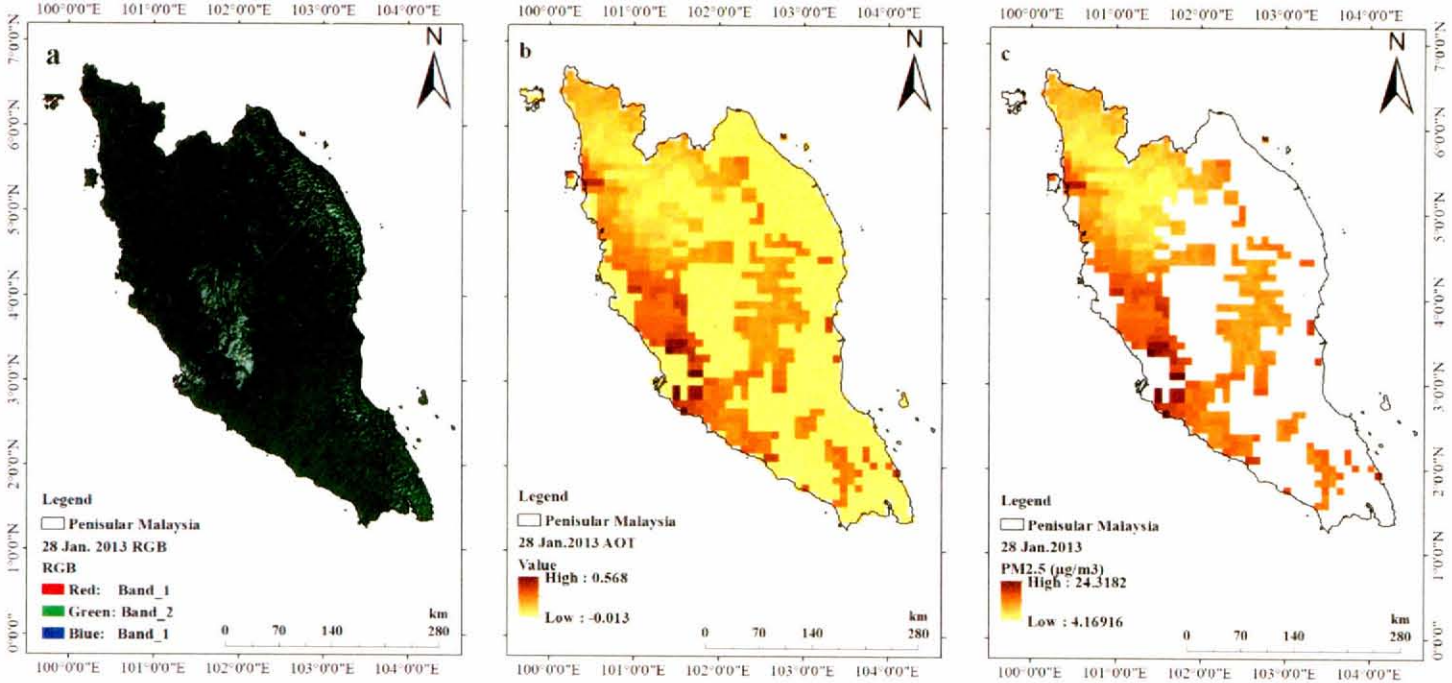


Figure 4.28 : Color Images: (a) True Color Image; (b) Retrieval of AOT from MODIS; (c) Linear Regression Model Predictions of $PM_{2.5}$ Concentrations on 28 January 2013

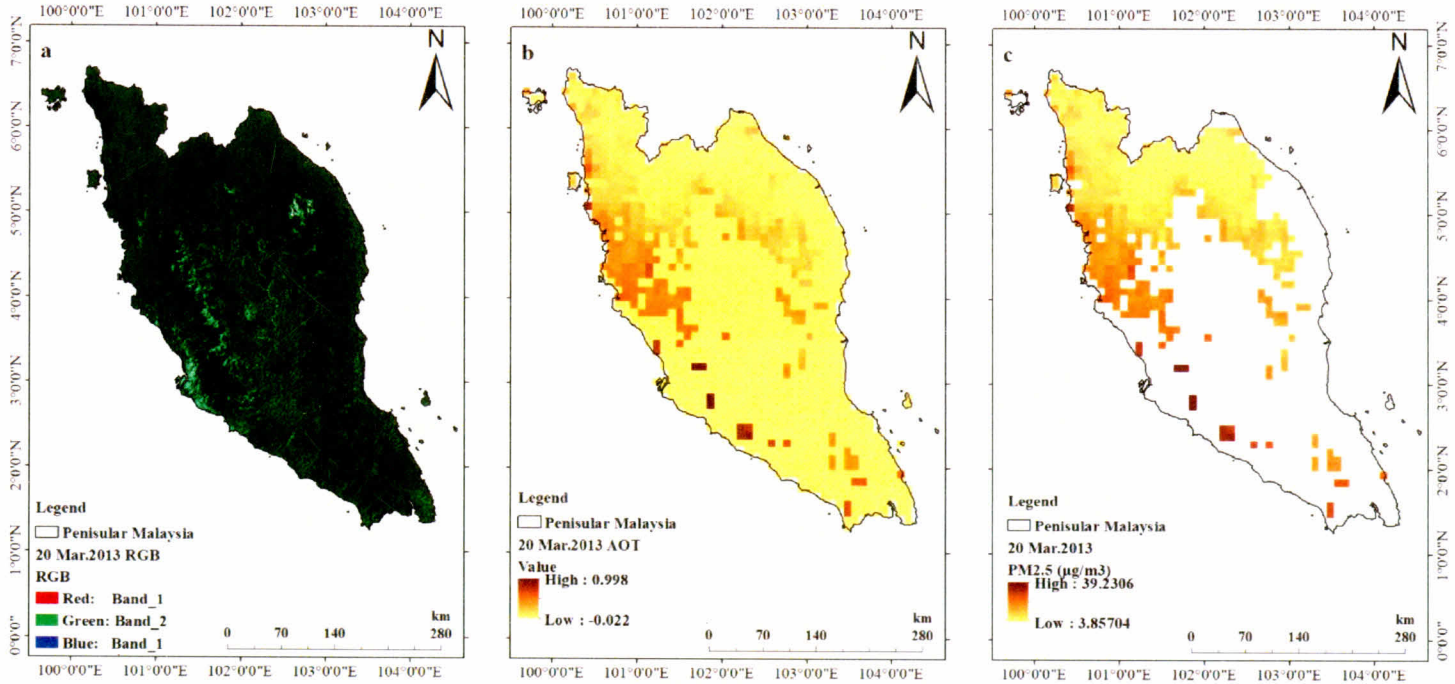


Figure 4.29 : Color Images: (a) True Color Image; (b) Retrieval of AOT from MODIS; (c) Linear Regression Model Predictions of $PM_{2.5}$ Concentrations on 20 March 2013

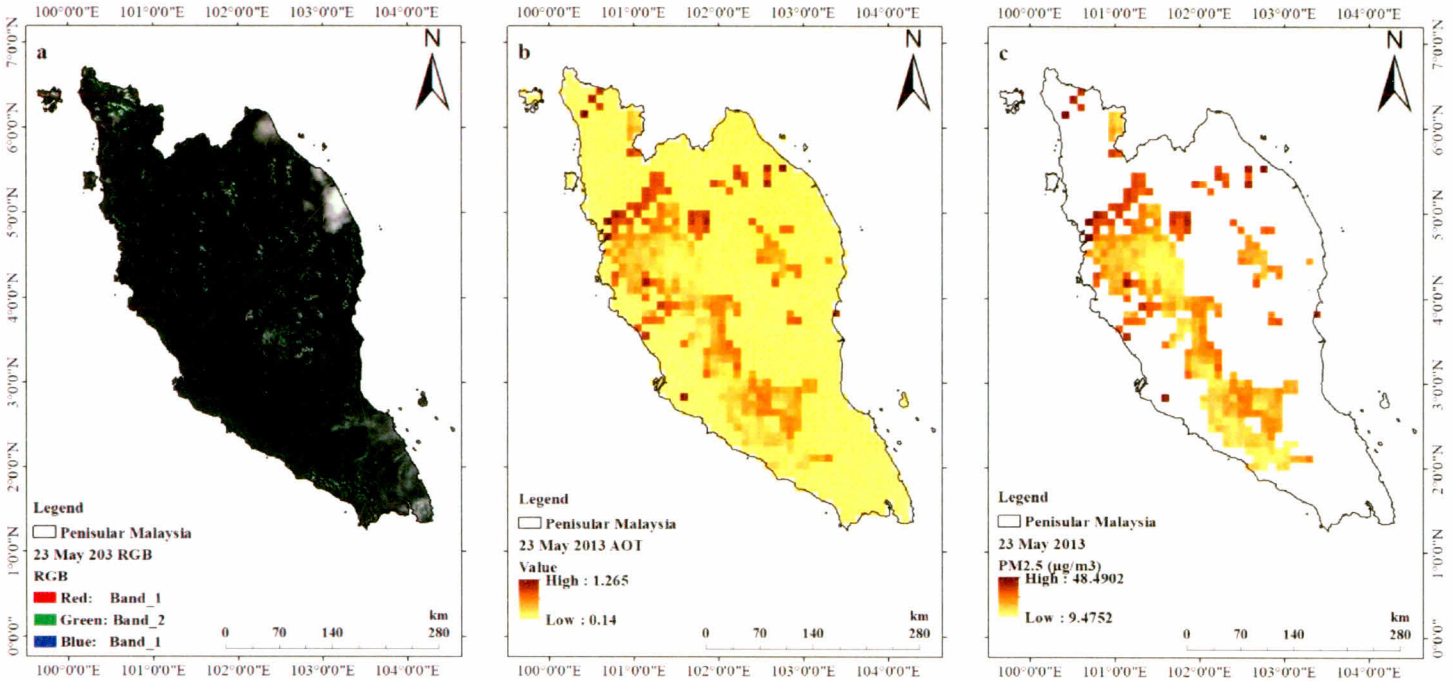


Figure 4.30 : Color Images: (a) True Color Image; (b) Retrieval of AOT from MODIS; (c) Linear Regression Model Predictions of $PM_{2.5}$ Concentrations on 23 May 2013

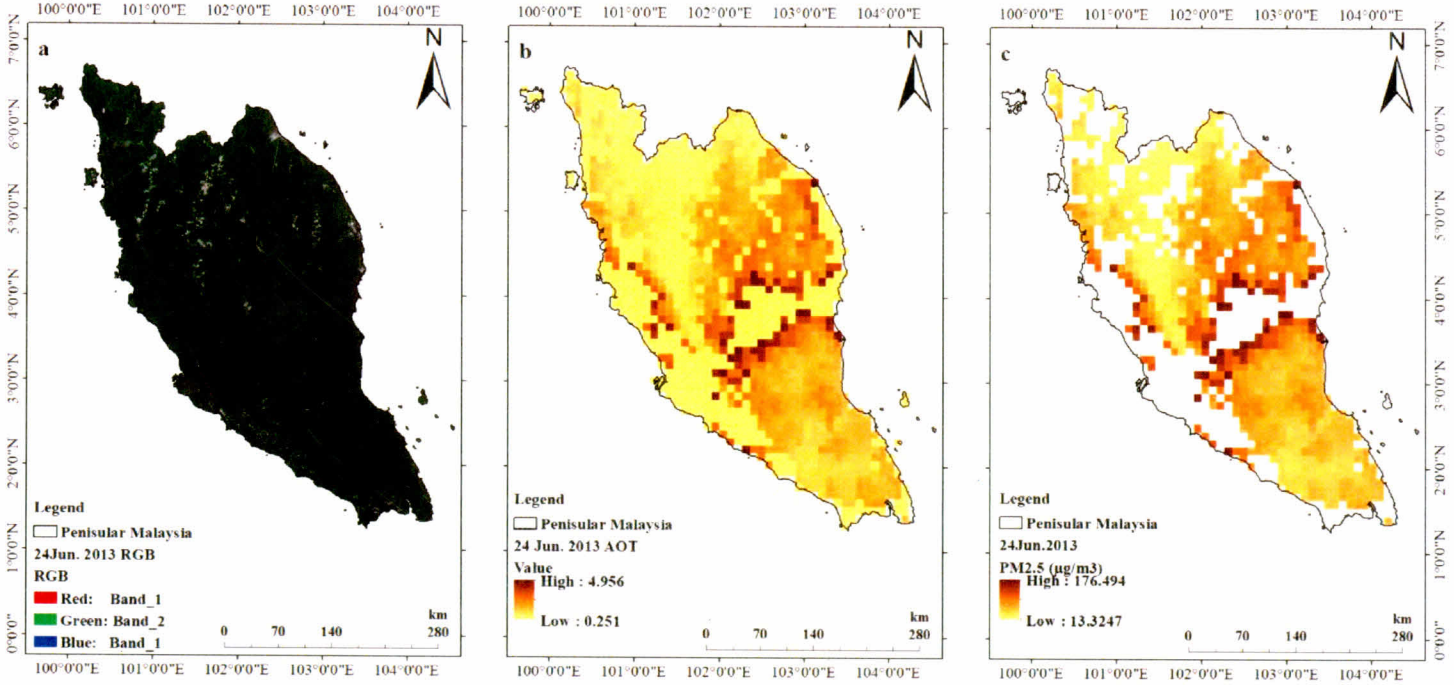


Figure 4.31 : Color Images: (a) True Color Image; (b) Retrieval of AOT from MODIS; (c) Linear Regression Model Predictions of $PM_{2.5}$ Concentrations on 24 June 2013

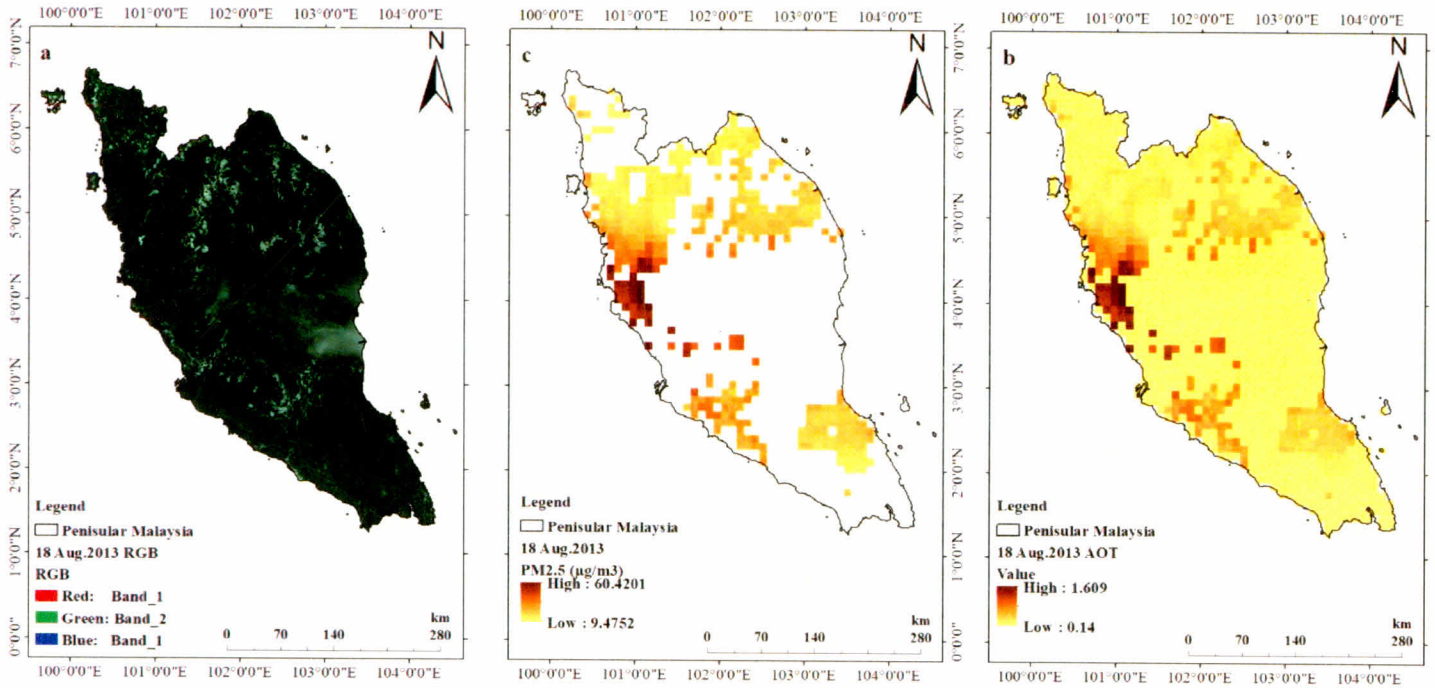


Figure 4.32 : Color Images: (a) True Color Image; (b) Retrieval of AOT from MODIS; (c) Linear Regression Model Predictions of $PM_{2.5}$ Concentrations on 18 August 2013

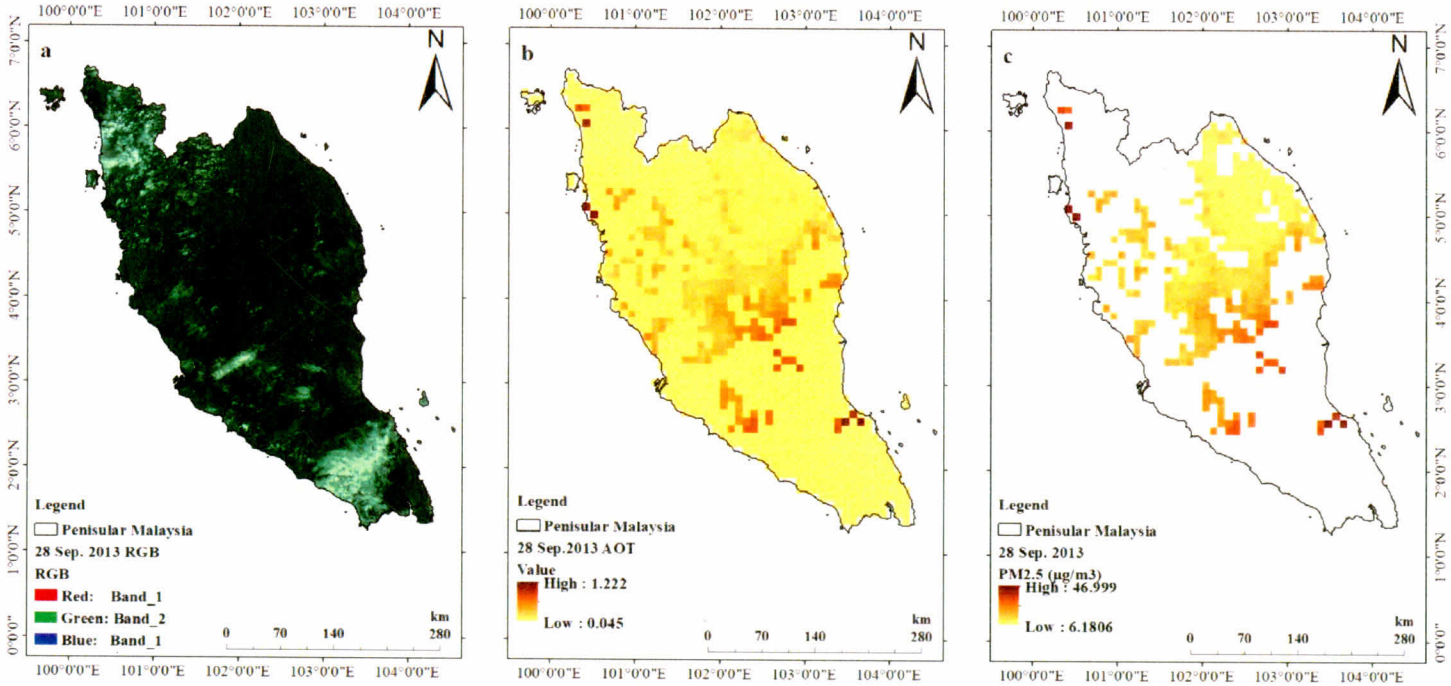


Figure 4.33 : Color Images: (a) True Color Image; (b) Retrieval of AOT from MODIS; (c) Linear Regression Model Predictions of $PM_{2.5}$ Concentrations on 28 September 2013

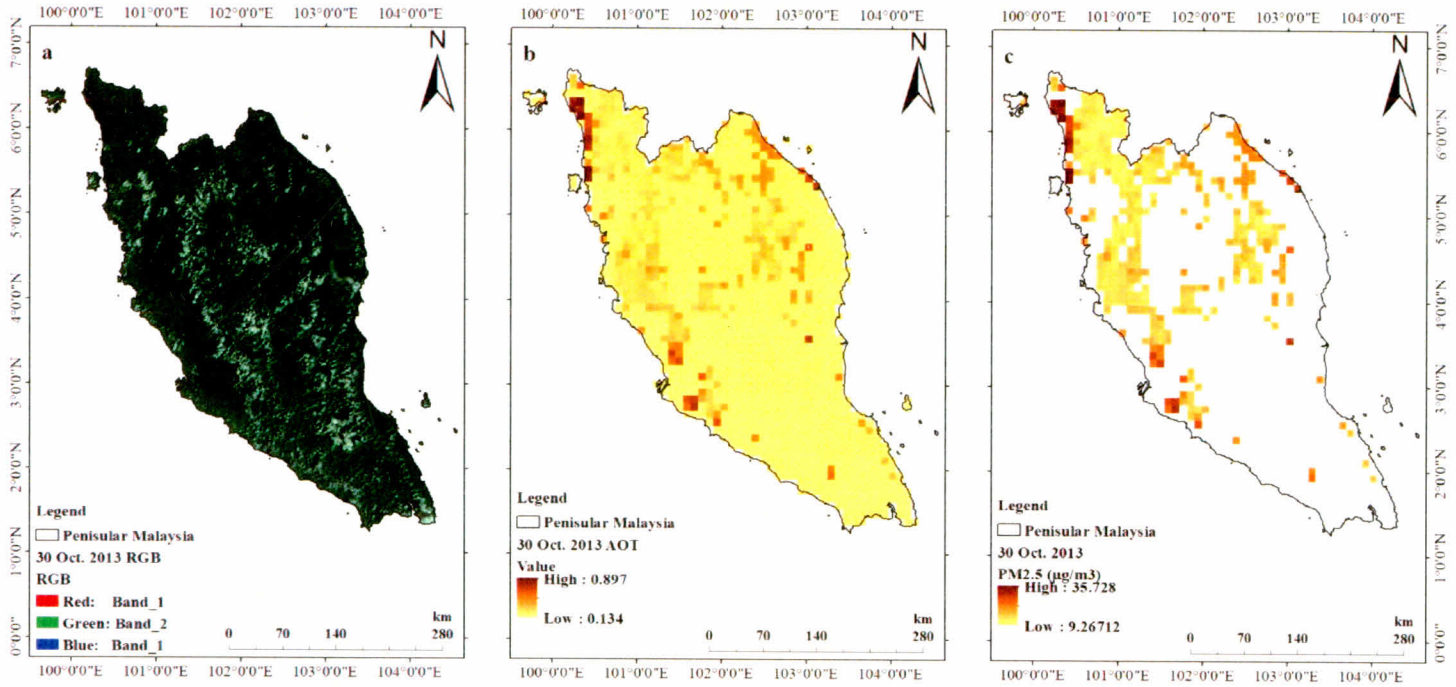


Figure 4.34 : Color Images: (a) True Color Image; (b) Retrieval of AOT from MODIS; (c) Linear Regression Model Predictions of $PM_{2.5}$ Concentrations on 30 October 2013

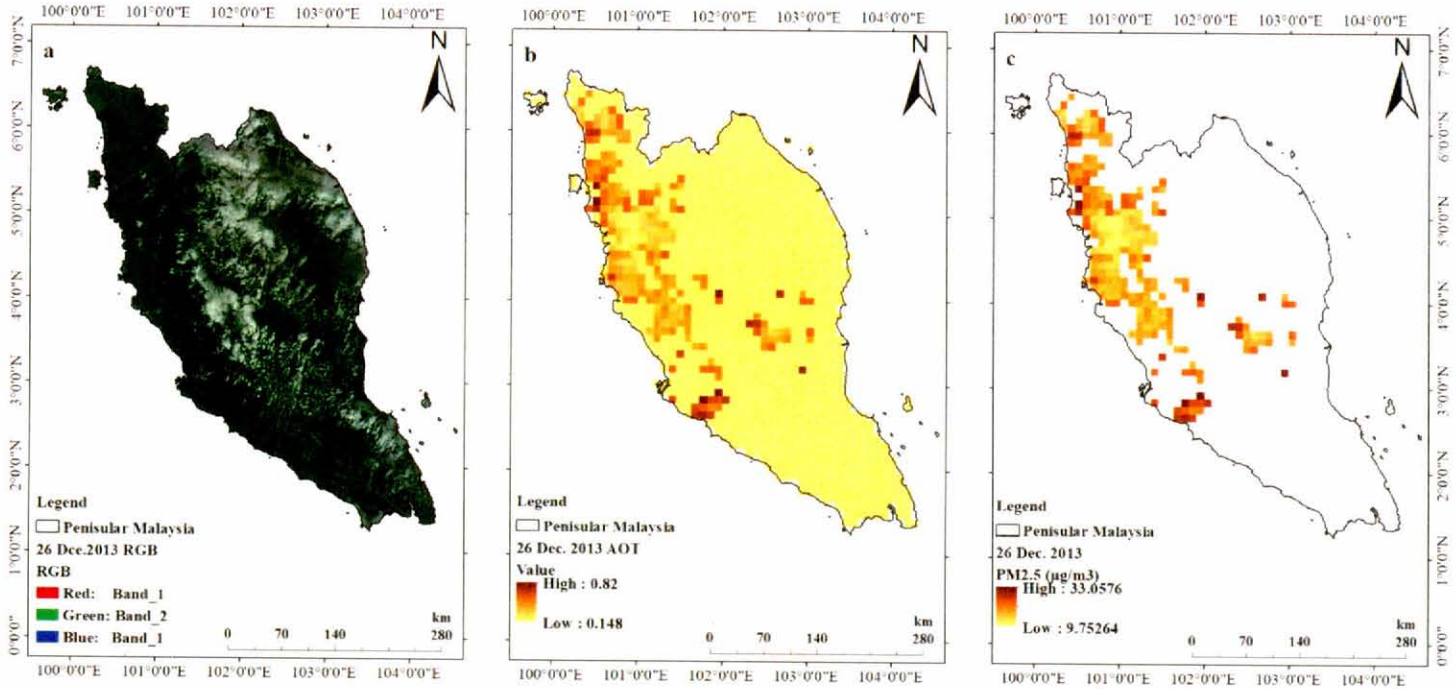


Figure 4.35 : Color Images: (a) True Color Image; (b) Retrieval of AOT from MODIS; (c) Linear Regression Model Predictions of $PM_{2.5}$ Concentrations on 26 December 2013

4.4 Assessment of the $PM_{2.5}$ Estimates for Peninsular Malaysia

This study compared the AERONET AOT values with the MODIS AOT values at the wavelength of 550 nm in an effort to evaluate relevance of the satellite data for the study area. Spatial statistics for the MODIS data and corresponding temporal AERONET data were first obtained. The 50 km × 50 km window of MODIS AOD matches ground-based data. The correlation between AERONET and MODIS AOT was good. In estimation of the ground-level $PM_{2.5}$ levels at the monitoring stations based on the MODIS sensor, MODIS AOT values were extracted at different pixel sizes (1x1, 3x3, and 5x5 pixels) and regression modeling was applied to assess the effect of MODIS AOT on the $PM_{2.5}$ levels for each pixel size group.

4.4.1 Correlations among Ground $PM_{2.5}$ Concentrations and MODIS AOT Values for the Different Pixel Size Groups

Table 4.19 and Figure 4.36 indicate the matching ground-level $PM_{2.5}$ concentrations at each station for different spatial MODIS pixel size groups. More matching data points were found in the 5x5 pixel group than in the 3x3 pixel group and the lowest number of matching data points was found in the 1x1 pixel group. For example, at the USM station, which had the highest number of matching data points in the 5x5 pixel group, the number of matching data points was 10 times the number of matching data points in the 1x1 pixel group (Figure 4.36). According to these findings, the researcher concludes that the values of MODIS AOT in the 5x5 pixel size group can minimize the temporal noise in the AOT data.

Table 4.19 : Numbers of matching PM2.5 concentration and MODIS AOT data points in each pixel size group for each ground monitoring station in 2013

	Station	The year 2013(365 days) MODIS AOT (1x1) Pixel	MODIS AOT (3x3) Pixels	MODIS AOT (5x5) Pixels
1	Banting	25	32	96
2	Bukit rambai	17	24	60
3	Cheras	27	37	105
4	Ipoh	50	59	125
5	KelantanTanah Merah	31	40	97
6	Putrajaya	18	26	69
7	USM	12	20	133
Overall		180/365	238/365	685/365

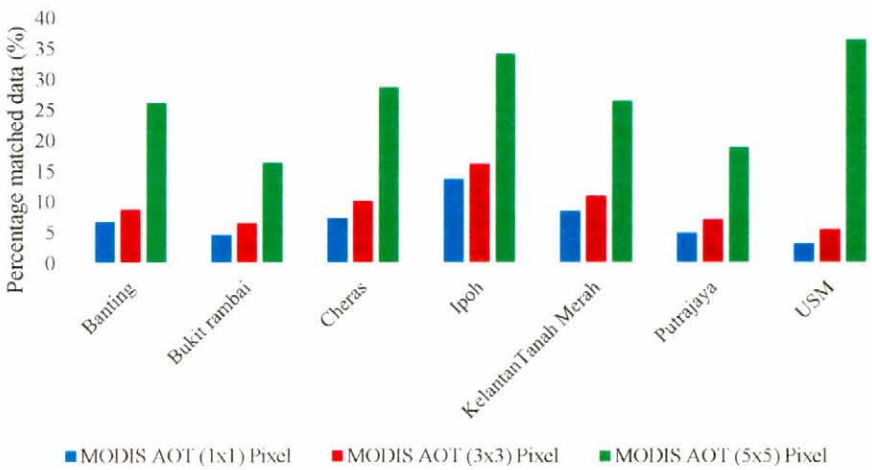


Figure 4.36 : Percentage Matching MODIS AOT and Difference MODIS AOT Data Points in each Pixel Size Group for each Ground Monitoring Station in 2013

Table 4.21 lists the values of the coefficients of correlations among the daily and hourly ground $PM_{2.5}$ concentrations and the MODIS AOT values for different pixel size groups. The results disclose that there were positive correlations between ground $PM_{2.5}$ concentrations and the MODIS AOT values in the different pixel groups. The highest hourly correlation (0.97) was noticed in the 5 x 5 pixel group and the lowest correlation was found in the 1 x 1 pixel group (0.29) at Bukit Rambi station. On the other hand, the highest daily correlation (0.80) was noticed at Cheras (5 x 5 pixels) and KelantanTanah Merah (1x1 pixel) while the lowest value was found in the 1x1 pixel size group (0.13) at Bukit Rambai station.

Table 4.20 : Coefficients of correlation among the annual hourly and daily ground $PM_{2.5}$ concentrations ($\mu g/m^3$) and MODIS AOT in the 1 x 1, 3 x 3, and 5 x 5 pixel groups for the study stations

	Station	$PM_{2.5}$ ($\mu g/m^3$)	MODIS AOT (1x1 Pixel)	MODIS AOT (3x3 Pixels)	MODIS AOT (5x5 Pixels)
1	Banting	Hourly	0.543	0.601	0.708
		Daily	0.146	0.194	0.405
2	Bukit Rambai	Hourly	0.288	0.364	0.966
		Daily	0.138	0.217	0.759
3	Cheras	Hourly	0.826	0.870	0.906
		Daily	0.720	0.792	0.799
4	Ipoh	Hourly	0.643	0.728	0.797
		Daily	0.375	0.434	0.496
5	KelantanTanah Merah	Hourly	0.675	0.716	0.719
		Daily	0.801	0.787	0.761
6	Putrajaya	Hourly	0.850	0.831	0.866
		Daily	0.680	0.633	0.682
7	USM	Hourly	0.458	0.655	0.567
		Daily	0.503	0.651	0.764
Overall		Hourly	0.603	0.656	0.846
		Daily	0.315	0.384	0.618

Overall, the results (Table 4.20) spotlight moderate to strong associations between MODIS AOT and the $PM_{2.5}$ concentration. For the hourly data, the r values were 0.60, 0.66, and 0.85 for the 1×1, 3×3, and 5×5 pixel size groups, respectively. Meantime, for the daily data, the r values were 0.32, 0.38, and 0.62 for the 1×1, 3×3, and 5×5 pixel size groups, respectively. These results point out that higher correlation between MODIS AOT and the $PM_{2.5}$ concentration can be obtained by increasing the pixel size.

4.4.2 Values of the Coefficient of Determination (R^2) at Different Temporal Scales

The MODIS helped many researchers discover a significant correlation between total-column AOT and ground-level $PM_{2.5}$ measurements (H. J Lee et al., 2016; Wang and Chen, 2016; You et al., 2015; Zheng et al., 2017). Table 4.21 presents the R^2 values for the regression models based on hourly and daily mean $PM_{2.5}$ measurements and MODIS AOT values for the seven monitoring stations considered in this study. The annual variations in the values of R^2 for the regression models of AOT and $PM_{2.5}$ were reported (Table 4.21). For the hourly LRM, the R^2 values ranged from 0.446 in Cheras to 0.772 in Banting station. In the case of the MLRM, the R^2 values ranged from 0.459 in Kelantan Tanah Merah to 0.794 at Bukit Rambai station. For the daily LRM, the R^2 values ranged from 0.186 in Cheras to 0.781 in USM. In the case of MLRM, the R^2 values ranged from 0.334 in Banting to 0.813 at USM station (Table 4.21). These results demonstrate that the R^2 values produced by the MLRMs were in general higher than those generated by the SLRMs. Viewed from other angle, the results of the LRMs and MLRMs unveil effect of the meteorological factors on the association between AOT and $PM_{2.5}$ concentration. The fact that the R^2 values associated with the MLRMs

were higher than those associated with the SLRMs suggests that incorporation of meteorological factor in models predictive of ground $PM_{2.5}$ concentrations from AOT improves the predictive powers of those models.

Table 4.21 : Values of the coefficient of determination associated with regression models based on daily and hourly AOT and $PM_{2.5}$ values for the air quality monitoring stations under study

Station	Model	Hourly	Daily
		R^2	R^2
Banting	LRM	0.772	0.248
	MLRM	0.789	0.334
Bukit Rambai	LRM	0.746	0.397
	MLRM	0.794	0.510
Cheras	LRM	0.446	0.186
	MLRM	0.475	0.790
Ipoh	LRM	0.713	0.473
	MLRM	0.731	0.503
Kelantan Tanah Merah	LRM	0.449	0.371
	MLRM	0.459	0.491
Putrajaya	LRM	0.701	0.532
	MLRM	0.727	0.561
USM	LRM	0.729	0.781
	MLRM	0.747	0.813

Results of assessment of the predictive powers of the regression models predictive of ground $PM_{2.5}$ concentrations from AOT values on a seasonal basis using hourly and daily data are summarized by Table 4.22. The regression results underscore that correlation between these two variables is higher in the southwest monsoon than in the northeast monsoon. On an hourly basis, the R^2 values associated with the MLRMs were higher in the southwest monsoon than in the northeast monsoon. For instance, in Bukit Rambai, the value of R^2 was 0.945 in the northeast monsoon and 0.763 in the southwest monsoon (Table 4.22). These findings support that the ground-level $PM_{2.5}$ concentrations and MODIS AOT values are sensitive to time. One probable explanation for this time sensitivity can be the timing of biomass burning activities in Sumatra (Indonesia) as was previously highlighted. However, the overriding effects of the southwest monsoonal winds may have contributed to the daily concentrations of $PM_{2.5}$ in this area.

Table 4.22 : Values of the coefficient of determination associated with regression models using daily and hourly AOT and $PM_{2.5}$ values in the two monsoons and the air quality monitoring stations under study

Stations	Southwest				Northeast			
	Daily (R^2)		Hourly(R^2)		Daily(R^2)		Hourly(R^2)	
	LRM	MLRM	LRM	MLRM	LRM	MLRM	LRM	MLRM
Banting	0.274	0.334	0.839	0.864	0.251	0.308	0.743	0.775
Bukit Rambai	0.591	0.667	0.894	0.945	0.278	0.431	0.666	0.763
Cheras	0.837	0.859	0.646	0.738	0.480	0.548	0.690	0.659
Ipoh	0.625	0.649	0.772	0.837	0.223	0.272	0.689	0.698
Putrajaya	0.629	0.798	0.749	0.782	0.021	0.352	0.552	0.628
Kelantan Tanah Merah	0.520	0.710	0.456	0.527	0.280	0.552	0.546	0.599
USM	0.872	0.892	0.853	0.868	0.549	0.667	0.307	0.438

According to the foregoing results (Table 4.22), the researcher infers that satellite RS data can provide valuable information on the area distribution of aerosols over a wide geographical range. These data can be valuable source of information for studies of air pollution that can compensate for the shortage of ground monitoring stations (Wang et al., 2013). In this regard, linear regression can be employed to convert the satellite AOT observations to $PM_{2.5}$ mass concentrations, which are two of the most useful indices of air quality (Gupta and Christopher, 2009).

4.4.3 Meteorological Parameters Influencing Prediction of $PM_{2.5}$ Concentrations

Multiple linear regression analysis following the 'Stepwise' variable entry method was employed in this study to assess the comparative effects of meteorological parameters on predictions of the ground $PM_{2.5}$ concentrations. No significant increases in the values of R^2 were observed with the inclusion of T (°C), WS (s/m), WD (°), and RH (%) with MODIS AOT in the regression models (Table 4.23). However, increases of 0.77% and 8.25% were observed in the regression models using hourly and daily data, respectively. The R^2 values of the regression models using hourly and daily data were improved by addition of meteorological parameters (Table 4.23). These results match with those of Liu et al. (2005); Paciorek et al. (2008) who also reported that including meteorological variables with RS data in a MLRM increases the R^2 value in $PM_{2.5}$ concentration prediction. These factors contribute to accuracy of the regression model. You et al. (2015) conducted a study to identify influential meteorological factors for prediction of ground-level $PM_{2.5}$ concentrations like visibility, surface T, surface WS, and surface RH. Stepwise regression analysis showed that there was significant increase in the R^2 value of the model upon inclusion of the foregoing factors in the models.

Table 4.23 : MODIS model linear coefficient of determination PM2.5 concentration prediction as a function of different independent variables

Model	Independent variable	Hourly		Daily	
		R^2	Change in R^2 (%)	R^2	Change in R^2 (%)
1	AOT	0.653	-	0.485	-
2	AOT, WS	0.655	0.31	0.501	3.30
3	AOT, WD	0.654	0.15	0.507	4.54
4	AOT, RH	0.654	0.15	0.485	-
5	AOT, T	0.654	0.15	0.489	0.82
6	AOT, WS, WD, RH and T	0.658	0.77	0.528	8.26

4.4.4 Assessment of Model Accuracy

Evaluation of the models developed in this study was carried out using the testing and validation data subsets and five performance evaluation metrics: R^2, MSE, RMSE, rRMSE, and MAPE. An account on the validation results is provided in the following sub-sections.

4.4.4.1 MODIS AOT and AERONET AOT Validation

Table 4.24 presents the results of validation of the MODIS AOT values against the average hourly and daily AERONET AOT data for the period 2012-2015. As can be observed in Table 4.24, there are no noticeable differences in the values of the performance metrics between the models employing hourly and daily data during the studied period (2015-2015). Accordingly, the researcher reaches to the conclusion that availability of long time series data of AOT from the Aerosol Robotic Network enables satisfactory assessment of MODIS products.

Table 4.24 : Results of validation of the MODIS AOT values against the average hourly and daily AERONET AOT data for the period 2012-2015

Description	Hourly average	Daily average
Observations	979	979
R^2	0.851	0.870
MSE	0.036	0.030
RMSE	0.189	0.175
MAPE (%)	35.829	32.101
rRMSE (%)	48.84	45.21
Slop	0.978	0.079
Intercept	0.807	0.023

4.4.4.2 Estimation of $PM_{2.5}$

Tables 4.25 and 4.26 present the results of assessment of accuracy of the LRMs using hourly and daily data, respectively. The results indicate that the MODIS model explained 36.3%, 43.7%, and 65.3% of the variations in the hourly $PM_{2.5}$ concentrations in the cases of the 1x1, 3x3, and 5x5 pixel size groups (Table 4.25), respectively, and 9.9%, 12.9%, and 48.5% of the variations in the daily $PM_{2.5}$ oncentrations in the cases of the 1x1, 3x3, and 5x5 pixel size groups, respectively (Table 4.26).

Table 4.25 : Results of assessment of accuracy of the linear regression models estimating hourly $PM_{2.5}$ concentrations in 2013

Description	1 x 1 Pixel	3 x 3 Pixels	5 x 5 Pixels
Observations	180	238	685
R^2	0.363	0.437	0.653
MSE ($\mu g/m^3$)	93.266	74.247	52.171
RMSE ($\mu g/m^3$)	9.657	8.617	7.223
rRMSE (%)	55.82	48.92	37.56
MAPE (%)	36.306	32.476	28.970
Slope	31.368	34.586	34.676
Intercept	5.989	4.967	4.619

Table 4.26 : Results of assessment of accuracy of the linear regression models estimating daily $PM_{2.5}$ concentrations in 2013

Description	1 x 1 Pixel	3 x 3 Pixels	5 x 5 Pixels
Observations	180	238	685
R^2	0.099	0.129	0.485
MSE ($\mu g/m^3$)	270.117	220.255	155.368
RMSE ($\mu g/m^3$)	16.435	14.841	12.465
rRMSE (%)	59.75	53.91	43.12
MAPE (%)	38.763	36.960	32.888
Slope	2.424	25.990	42.278
Intercept	19.183	18.019	11.095

Levels of performance of the various regression models developed in this study were evaluated for different pixel size groups with reference to the values of R^2, MSE, RMSE, rRMSE, and MAPE. The results related to the regression models based on the hourly (Table 4.25) and daily (Table 4.26) data demonstrate that the 5 x 5 pixel group model had the lowest values of MSR, RMSE, rRMSE, and MAPE, which pinpoints

superiority of this model over the models generated for the 1x1 and 3x3 pixel size groups. These findings provide additional evidence on importance of spatio-temporal resolution in the relation between ground $PM_{2.5}$ concentrations and MODIS AOT. Additionally, the accuracy of the MODIS-derived AOT data was low, which can result in information bias since certain important covariates like land uses, regional variations in the aerosols, and seasonal variations in the aerosols were missing from the designed models due to lack of information on them.

4.4.5 Comparing of R^2 Values Between Studies

A Comparison of values of R^2 between studies to evaluate satellite-based prediction of $PM_{2.5}$ concent rations for Peninsular Malaysia is provided by Table 4.27. Review of this table discloses that the values of R^2 which this study obtained are consistent with literature values, where the model using hourly data has an R^2 value of 0.66 while the model based on daily data has an R^2 value of 0.45.

Table 4.27 : Comparison of values of the coefficient of determination between studies to evaluate satellite-based prediction of $PM_{2.5}$ concentrations for Peninsular Malaysia

Location	R^2 value		Period	reference
	Hourly	Daily		
USA		0.90	2008 - 2013	(Zhang et al., 2018)
China		0.71	April 2014 to April 2015	(Yao et al., 2018)
China		0.24	May to June 2016	(Yan et al., 2017)
USA		0.60	2003	(Hu et al., 2013)
China	0.25		From October to December 2013	(Zhang and Li, 2015)
USA		0.89	2003	(Chudnovsky et al., 2014)
China		0.70	2013	(Zongwei Ma et al., 2016)
China	0.29	0.18	2015	(Guo et al., 2017)
China		Ranged from 0.65 to 0.72	2009 -2010	(Kong et al., 2016)
China		0.82	October 2004	(Wang et al., 2013)
Malaysia	0.66	0.49	2013	This study

CHAPTER 5

SUMMARY AND CONCLUSIONS

5.1 Summary

This study was based on one-year (2013), hourly and daily MODIS-AOT data obtained from the Terra satellite and $PM_{2.5}$ data measured at seven ground stations in Peninsular Malaysia. The data processing results demonstrate that the MODIS AOT values have medium to high relationships with the measured levels of $PM_{2.5}$, where the r values characteristic of the relationships of MODIS AOT values with the hourly and the daily mean $PM_{2.5}$ concentrations were 0.67 and 0.76, respectively. These results lead to the conclusion that the AOT values retrieved from the satellite constitute a useful tool for the air quality studies over vast spatial domains aiming at tracking and monitoring particulates.

Until now, the $PM_{2.5}$ informational collections have broad temporal and spatial resolutions but are yet inadequate. A purposeful, keen effort is needed to compile a compendium database of temporal and spatial resolutions of the $PM_{2.5}$ concentrations in Malaysia to more fully investigate the relationship between the ground $PM_{2.5}$ measurements and the satellite-determined levels of particulates.

The results of the present study emphasize that estimation of the ground-level $PM_{2.5}$ concentrations can be enhanced by including in future models meteorological variables that are directly affecting the levels of PM in the atmosphere. This study developed SLRMs and MLRMs that predict the concentrations of ground $PM_{2.5}$. In the MLRMs, the meteorological predictor variables improved the estimation of the ground-level $PM_{2.5}$ substantially.

5.2 Conclusion

The present study sheds light on the fact that satellite imagery and data can provide highly useful input to air quality policies. Researchers have been conducting quantitative analyses that compare satellite sensor data with ground air quality measurements. By far, there is consensus among those researchers on strong relationships among the $PM_{2.5}$ concentrations, MODIS AOT data, and other ground air quality measures.

Our understanding of the sources and transport of air pollutants can be improved much by using satellite data in combination with other relevant data from different sources and suitable models. Regression modeling can be used to determine the concentration of PM using MODIS sensor data which cover broad geospatial areas and are available

at least once a day. When continuously monitored, the MODIS sensor will produce valid AOT data that can be employed to build a satisfactory database of reliable data.

The MODIS AOT sensor was validated with the AERONET AOT. A strong correlation was found between the AERONET AOT and MODIS AOT observations. It was also found that the MODIS AOT products tended to overestimate the ground-level AOT values. Nonetheless, most of the values fell within a reasonable range of AERONET AOT values. The results of this study point out that the products of the MODIS AOT sensor are suitable enough to be used for $PM_{2.5}$ estimation. For regions that lack AEORNET sites, the MODIS AOT products can be considered as a reasonable substitute.

Simple linear regression models, MLRMs, and GWRMs were developed to estimate the ground $PM_{2.5}$ concentrations. The modeling outcomes uncovered that there are statistically-significant, positive correlations amongst all the variables included in the analysis. The prediction capacities of the LRMs, in terms of the value of R^2, were noticeably higher when taking into account the meteorological data, especially when meteorological data are incorporated in the same model with the hourly $PM_{2.5}$ concentrations. In addition, ten-fold cross validation was performed to validate the models. The results disclosed similar R^2 values for the LRMs using daily and hourly data. By comparing the spatial distributions of the $PM_{2.5}$ in different seasons, a better match between the predicted and actual $PM_{2.5}$ levels was found in the northeast monsoon than in the southwest monsoon.

Accuracies of the various models were assessed using training and validation data partitions and four performance evaluation metrics: R^2, MSE, RMSE, and rRMSE. These metrics were used to assess the levels of performance of the LRMs predictive of ground $PM_{2.5}$ concentrations from AOT. In this regard, various factors can explain poor correlation between AOT and the ground concentrations of $PM_{2.5}$. One factor is that these two variables measure different particle categories.

The relationship between AOT and $PM_{2.5}$ concentration may be affected by the atmospheric conditions during the study period. Thereupon, studies of long periods are much likely to obtain highly robust results since influence of the spatial conditions on the investigation results can be at a minimum. The current study used twelve-month data and its findings coincided with the results of previous studies employing longer time series. Consequently, continuous monitoring of air quality is mandatory to obtain accurate assessment of the AOT-$PM_{2.5}$ relationship. In view of these findings, the researcher concludes that the MODIS-$PM_{2.5}$ models are capable of estimating the air pollution levels at different spatio-temporal scales. In other respects, it appears that the study period does not affect the strength of the AOT-$PM_{2.5}$ relationship.

There were few limitations to this study. Information about influential covariates like the seasonal variations in aerosols, land use information, and pixel sizes was unavailable for the researcher. Moreover, the cloud mask and ephemeral water body

tests reduce the amount of retrievable data. Another limitation is the short study period, which was governed by the availability of relevant data. Within this context, (Wang et al., 2017) reported that long-term matching satellite and ground $PM_{2.5}$ concentration data will improve the correlation between AOT and $PM_{2.5}$ concentration.

Another limitation is the lack of ground $PM_{2.5}$ data. In addition, the sampling frequencies were different as $PM_{2.5}$ represents the 24-hour average whereas AOT corresponds to snap-shot of the daytime PM level. Furthermore, the ground monitoring stations in the study area did not have adequate coverage of the PM concentrations and the necessary meteorological factors over Peninsular Malaysia. Lastly, different factors have differing effects on the accuracies of the ground and space measurements.

5.3 Scope for Future

1. Predictions of the ground $PM_{2.5}$ levels can be much improved by inclusion of meteorological variables related to PM in the models.
2. It is important to study variations in the fluxes of PM and the related parameters of the surface boundary layer during the different seasons, with particular emphasis on the monsoon periods.
3. Some studies found that there is statistically-significant relationship between mortality and the $PM_{2.5}$ concentrations in the atmosphere. Research should be conducted using RS data to contribute refined knowledge to this issue.
4. Long-term monitoring of ground-level concentrations of $PM_{2.5}$ is needed to develop estimation accuracy of the satellites.
5. The vertical structure of the PM and its components should be researched using more RS tools like Lidar.
6. The AOT products of other sensors such as Aqua, MISR, and SeaWIFS can be combined to increase the number of the available AOT retrieval days for more accurate $PM_{2.5}$ predictions.

REFERENCES

Aaltonen, V., Rodriguez, E., Kazadzis, S., Arola, A., Amiridis, V., Lihavainen, H., and De Leeuw, G. (2012). On the variation of aerosol properties over Finland based on the optical columnar measurements. *Atmospheric Research, 116*, 46-55.

Abdul Halim, N. D., Latif, M. T., Ahamad, F., Dominick, D., Chung, J. X., Juneng, L., and Khan, M. F. (2018). The long-term assessment of air quality on an island in Malaysia. *Heliyon, 4*(12), e01054. doi:10.1016/j.heliyon.2018.e01054

Afroz, R., Hassan, M. N., and Ibrahim, N. A. (2003). Review of air pollution and health impacts in Malaysia. *Environmental Research, 92*(2), 71-77. doi:10.1016/s0013-9351(02)00059-2

Al-Hamdan, M. Z., Crosson, W. L., Limaye, A. S., Rickman, D. L., Quattrochi, D. A., Estes, M. G., Qualters, J. R., Sinclair, A. H., Tolsma, D. D., Adeniyi, K. A., and Niskar, A. S. (2009). Methods for Characterizing Fine Particulate Matter Using Ground Observations and Remotely Sensed Data: Potential Use for Environmental Public Health Surveillance. *Journal of the Air & Waste Management Association, 59*(7), 865-881. doi:10.3155/1047-3289.59.7.865

Al-Saadi, J., Szykman, J., Pierce, R. B., Kittaka, C., Neil, D., Chu, D. A., Remer, L., Gumley, L., Prins, E., Weinstock, L., MacDonald, C., Wayland, R., Dimmick, F., and Fishman, J. (2005). Improving national air quality forecasts with satellite aerosol observations. *Bulletin of the American Meteorological Society, 86*(9), 1249-1261. doi:10.1175/BAMS-86-9-1249

Alam, K., Qureshi, S., and Blaschke, T. (2011). Monitoring spatio-temporal aerosol patterns over Pakistan based on MODIS, TOMS and MISR satellite data and a HYSPLIT model. *Atmospheric Environment, 45*(27), 4641-4651.

Amanollahi, J., Abdullah, A. M., Ramli, M. F., and Pirasteh, S. (2011). Real time assessment of haze and pm 10 aided by modis aerosol optical thickness over klang valley, Malaysia. *World Applied Sciences Journal, 14*(SPL ISS 1), 08-13. doi:10.1029/2001GL013206

Andria, G., D'orazio, A., Ekuakille, A. L., Moretti, M., Pierl, P., Tralli, F., and Tropeano, M. (2000). *Accuracy assessment in photo interpretation of remote sensing ERS-2 SAR images.* Paper presented at the Proceedings of the 17th IEEE Instrumentation and Measurement Technology Conference [Cat. No. 00CH37066].

Anenberg, S. C., West, J. J., Yu, H., Chin, M., Schulz, M., Bergmann, D., Bey, I., Bian, H., Diehl, T., Fiore, A., Hess, P., Marmer, E., Montanaro, V., Park, R., Shindell, D., Takemura, T., and Dentener, F. (2014). Impacts of intercontinental transport of anthropogenic fine particulate matter on human

mortality. *Air Quality, Atmosphere and Health, 7*(3), 369-379. doi:10.1007/s11869-014-0248-9

Aurela, M., Saarikoski, S., Niemi, J. V., Canonaco, F., Prevot, A. S., Frey, A., Carbone, S., Kousa, A., and Hillamo, R. (2015). Chemical and source characterization of submicron particles at residential and traffic sites in the Helsinki Metropolitan area, Finland. *Aerosol Air Qual. Res, 15*, 1213-1226.

Azmi, S. Z., Latif, M. T., Ismail, A. S., Juneng, L., and Jemain, A. A. (2010). Trend and status of air quality at three different monitoring stations in the Klang Valley, Malaysia. *Air Quality Atmosphere and Health, 3*(1), 53-64. doi:10.1007/s11869-009-0051-1

Bai, Y., Wu, L., Qin, K., Zhang, Y., Shen, Y., and Zhou, Y. (2016). A geographically and temporally weighted regression model for ground-level PM2. 5 estimation from satellite-derived 500 m resolution AOD. *Remote Sensing, 8*(3), 262.

Bai, Y., Wu, L. X., Qin, K., Zhang, Y. F., Shen, Y. Y., and Zhou, Y. (2016). A Geographically and Temporally Weighted Regression Model for Ground-Level PM2.5 Estimation from Satellite-Derived 500 m Resolution AOD. *Remote Sensing, 8*(3), 21. doi:10.3390/rs8030262

Baker, K. R., and Foley, K. M. (2011). A nonlinear regression model estimating single source concentrations of primary and secondarily formed PM2.5. *Atmospheric Environment, 45*(22), 3758-3767. doi:10.1016/j.atmosenv.2011.03.074

Bartell, S. M., Longhurst, J., Tjoa, T., Sioutas, C., and Delfino, R. J. (2013). Particulate air pollution, ambulatory heart rate variability, and cardiac arrhythmia in retirement community residents with coronary artery disease. *Environmental Health Perspectives, 121*(10), 1135.

Baum, B. A., Holz, R. E., Huang, H. L., Lee, Y. K., Yang, P., Nasiri, S. L., King, M. D., and Platnick, S. (2007). *Inference and validation of cloud phase from MODIS, AIRS and CALIPSO Data.* Paper presented at the Hyperspectral Imaging and Sounding of the Environment, HISE 2007, Santa Fe, NM.

Beckerman, B. S., Jerrett, M., Martin, R. V., van Donkelaar, A., Ross, Z., and Burnett, R. T. (2013). Application of the deletion/substitution/addition algorithm to selecting land use regression models for interpolating air pollution measurements in California. *Atmospheric Environment, 77*, 172-177.

Bell, M. L., and Davis, D. L. (2001). Reassessment of the lethal London fog of 1952: novel indicators of acute and chronic consequences of acute exposure to air pollution. *Environmental Health Perspectives, 109*(suppl 3), 389-394.

Bellouin, N., Jones, A., Haywood, J., and Christopher, S. A. (2008). Updated estimate of aerosol direct radiative forcing from satellite observations and comparison against the Hadley Centre climate model. *Journal of Geophysical Research: Atmospheres, 113*(D10).

Beloconi, A., Kamarianakis, Y., and Chrysoulakis, N. (2016). Estimating urban PM10 and PM2. 5 concentrations, based on synergistic MERIS/AATSR aerosol observations, land cover and morphology data. *Remote Sensing of Environment, 172*, 148-164.

Benas, N., Beloconi, A., and Chrysoulakis, N. (2013a). Estimation of urban PM10 concentration, based on MODIS and MERIS/AATSR synergistic observations. *Atmospheric Environment, 79*, 448-454.

Benas, N., Chrysoulakis, N., and Giannakopoulou, G. (2013b). Validation of MERIS/AATSR synergy algorithm for aerosol retrieval against globally distributed AERONET observations and comparison with MODIS aerosol product. *Atmospheric Research, 132*, 102-113.

Bevan, S. L., North, P. R., Los, S. O., and Grey, W. M. (2012). A global dataset of atmospheric aerosol optical depth and surface reflectance from AATSR. *Remote Sensing of Environment, 116*, 199-210.

Bibi, H., Alam, K., Chishtie, F., Bibi, S., Shahid, I., and Blaschke, T. (2015). Intercomparison of MODIS, MISR, OMI, and CALIPSO aerosol optical depth retrievals for four locations on the Indo-Gangetic plains and validation against AERONET data. *Atmospheric Environment, 111*, 113-126. doi:10.1016/j.atmosenv.2015.04.013

Bilal, M., Nichol, J. E., and Chan, P. W. (2014). Validation and accuracy assessment of a Simplified Aerosol Retrieval Algorithm (SARA) over Beijing under low and high aerosol loadings and dust storms. *Remote Sensing of Environment, 153*, 50-60. doi:10.1016/j.rse.2014.07.015

Bouya, Z., Box, G. P., and Box, M. A. (2010). Seasonal variability of aerosol optical properties in Darwin, Australia. *Journal of Atmospheric and Solar-Terrestrial Physics, 72*(9), 726-739.

Boyouk, N., Léon, J.-F., Delbarre, H., Podvin, T., and Deroo, C. (2010). Impact of the mixing boundary layer on the relationship between PM2. 5 and aerosol optical thickness. *Atmospheric Environment, 44*(2), 271-277.

Boys, B. L., Martin, R. V., Van Donkelaar, A., MacDonell, R. J., Hsu, N. C., Cooper, M. J., Yantosca, R. M., Lu, Z., Streets, D. G., Zhang, Q., and Wang, S. W. (2014). Fifteen-year global time series of satellite-derived fine particulate matter. *Environmental Science and Technology, 48*(19), 11109-11118. doi:10.1021/es502113p

Bréon, F.-M., Vermeulen, A., and Descloitres, J. (2011). An evaluation of satellite aerosol products against sunphotometer measurements. *Remote Sensing of Environment, 115*(12), 3102-3111. doi:10.1016/j.rse.2011.06.017

Bright, J. M., and Gueymard, C. A. (2019). Climate-specific and global validation of MODIS Aqua and Terra aerosol optical depth at 452 AERONET stations. *Solar Energy, 183*, 594-605.

Chang, S.-C. (2016). Atmospheric impacts of Indonesian fire emissions: Assessing remote sensing data and air quality during 2013 Malaysian haze. *Procedia Environmental Sciences, 36*, 176-179.

Charlson, R. J., Ahlquist, N., Selvidge, H., and MacCready Jr, P. (1969). Monitoring of atmospheric aerosol parameters with the integrating nephelometer. *Journal of the Air Pollution Control Association, 19*(12), 937-942.

Chen, M., Boyle, E. A., Switzer, A. D., and Gouramanis, C. (2016). A century long sedimentary record of anthropogenic lead (Pb), Pb isotopes and other trace metals in Singapore. *Environmental Pollution, 213*, 446-459.

Cheng, T., Chen, H., Gu, X., Yu, T., Guo, J., and Guo, H. (2012). The inter-comparison of MODIS, MISR and GOCART aerosol products against AERONET data over China. *Journal of Quantitative Spectroscopy and Radiative Transfer, 113*(16), 2135-2145. doi:10.1016/j.jqsrt.2012.06.016

Chin, M., Chu, A., Levy, R., Remer, L., Kaufman, Y., Holben, B., Eck, T., Ginoux, P., and Gao, Q. (2004). Aerosol distribution in the Northern Hemisphere during ACE-Asia: Results from global model, satellite observations, and Sun photometer measurements. *Journal of Geophysical Research D: Atmospheres, 109*(23), 1-15. doi:10.1029/2004JD004829

Choi, J., Fuentes, M., and Reich, B. J. (2009). Spatial-temporal association between fine particulate matter and daily mortality. *Computational Statistics and Data Analysis, 53*(8), 2989-3000. doi:10.1016/j.csda.2008.05.018

Choi, M., Kim, J., Lee, J., Kim, M., Park, Y. J., Jeong, U., Kim, W., Hong, H., Holben, B., Eck, T. F., Song, C. H., Lim, J. H., and Song, C. K. (2016). GOCI Yonsei Aerosol Retrieval (YAER) algorithm and validation during the DRAGON-NE Asia 2012 campaign. *Atmospheric Measurement Techniques, 9*(3), 1377-1398. doi:10.5194/amt-9-1377-2016

Christopher, S. A., and Gupta, P. (2010). Satellite remote sensing of particulate matter air quality: The cloud-cover problem. *Journal of the Air and Waste Management Association, 60*(5), 596-602. doi:doi: 10.1029/ 2003JD004248

Chu, D. A., Kaufman, Y., Zibordi, G., Chern, J., Mao, J., Li, C., and Holben, B. (2003). Global monitoring of air pollution over land from the Earth Observing System□Terra Moderate Resolution Imaging Spectroradiometer (MODIS). *Journal of Geophysical Research: Atmospheres, 108*(D21).

Chu, Y., Liu, Y., Li, X., Liu, Z., Lu, H., Lu, Y., Mao, Z., Chen, X., Li, N., and Ren, M. (2016). A Review on Predicting Ground PM2. 5 Concentration Using Satellite Aerosol Optical Depth. *Atmosphere, 7*(10), 129.

Chudnovsky, A. A., Kostinski, A., Lyapustin, A., and Koutrakis, P. (2013). Spatial scales of pollution from variable resolution satellite imaging. *Environmental Pollution, 172*, 131-138.

Chudnovsky, A. A., Koutrakis, P., Kloog, I., Melly, S., Nordio, F., Lyapustin, A., Wang, Y., and Schwartz, J. (2014). Fine particulate matter predictions using high resolution Aerosol Optical Depth (AOD) retrievals. *Atmospheric Environment, 89*, 189-198. doi:10.1016/j.atmosenv.2014.02.019

Clerbaux, C., Turquety, S., and Coheur, P. (2010). Infrared remote sensing of atmospheric composition and air quality: Towards operational applications. *Comptes Rendus Geoscience, 342*(4-5), 349-356. doi:10.1016/j.crte.2009.09.010

Cleugh, H. A., Leuning, R., Mu, Q., and Running, S. W. (2007). Regional evaporation estimates from flux tower and MODIS satellite data. *Remote Sensing of Environment, 106*(3), 285-304.

Consultation, W. E. (2018). Available evidence for the future update of the WHO Global Air Quality Guidelines (AQGs). In.

Dai, T., Schutgens, N. A. J., Goto, D., Shi, G. Y., and Nakajima, T. (2014). Improvement of aerosol optical properties modeling over Eastern Asia with MODIS AOD assimilation in a global non-hydrostatic icosahedral aerosol transport model. *Environmental Pollution, 195*, 319-329. doi:10.1016/j.envpol.2014.06.021

Davis, G. (2007). History of the NOAA satellite program. *Journal of Applied Remote Sensing, 1*(1), 012504-012504-012518.

De Graaf, M., Stammes, P., Torres, O., and Koelemeijer, R. (2005). Absorbing Aerosol Index: Sensitivity analysis, application to GOME and comparison with TOMS. *Journal of Geophysical Research: Atmospheres, 110*(D1).

Di Nicolantonio, W., Cacciari, A., Bolzacchini, F., Ferrero, L., Volta, M., and Pisoni, E. (2007). *MODIS aerosol optical properties over North Italy for estimating surface-level PM2. 5*. Paper presented at the Proceedings of Envisat Symposium.

Di Nicolantonio, W., Cacciari, A., and Tomasi, C. (2009). Particulate matter at surface: Northern Italy monitoring based on satellite remote sensing, meteorological fields, and in-situ samplings. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2*(4), 284-292.

Dimitriou, K., and Kassomenos, P. (2014). A study on the reconstitution of daily PM10 and PM2.5 levels in Paris with a multivariate linear regression model. *Atmospheric Environment, 98*, 648-654. doi:10.1016/j.atmosenv.2014.09.047

Diner, D. J., Beckert, J. C., Reilly, T. H., Bruegge, C. J., Conel, J. E., Kahn, R. A., Martonchik, J. V., Ackerman, T. P., Davies, R., and Gerstl, S. A. (1998). Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview. *Ieee Transactions on Geoscience and Remote Sensing, 36*(4), 1072-1087.

Drury, E., Jacob, D. J., Spurr, R. J., Wang, J., Shinozuka, Y., Anderson, B. E., Clarke, A. D., Dibb, J., McNaughton, C., and Weber, R. (2010). Synthesis of satellite (MODIS), aircraft (ICARTT), and surface (IMPROVE, EPA□AQS, AERONET) aerosol observations over eastern North America to improve MODIS aerosol retrievals and constrain surface aerosol concentrations and sources. *Journal of Geophysical Research: Atmospheres, 115*(D14).

Duncan, B. N., Prados, A. I., Lamsal, L. N., Liu, Y., Streets, D. G., Gupta, P., Hilsenrath, E., Kahn, R. A., Nielsen, J. E., and Beyersdorf, A. J. (2014). Satellite data of atmospheric pollution for US air quality applications: Examples of applications, summary of data end-user resources, answers to FAQs, and common mistakes to avoid. *Atmospheric Environment, 94*, 647-662.

Eck, T., Holben, B., Dubovik, O., Smirnov, A., Goloub, P., Chen, H., Chatenet, B., Gomes, L., Zhang, X. Y., and Tsay, S. C. (2005). Columnar aerosol optical properties at AERONET sites in central eastern Asia and aerosol transport to the tropical mid□Pacific. *Journal of Geophysical Research: Atmospheres, 110*(D6).

Emili, E., Popp, C., Petitta, M., Riffler, M., Wunderle, S., and Zebisch, M. (2010). PM 10 remote sensing from geostationary SEVIRI and polar-orbiting MODIS sensors over the complex terrain of the European Alpine region. *Remote Sensing of Environment, 114*(11), 2485-2499.

Engel-Cox, J. A., Hoff, R. M., Rogers, R., Dimmick, F., Rush, A. C., Szykman, J. J., Al-Saadi, J., Chu, D. A., and Zell, E. R. (2006). Integrating lidar and satellite optical depth with ambient monitoring for 3-dimensional particulate characterization. *Atmospheric Environment, 40*(40), 8056-8067.

Engel-Cox, J. A., Holloman, C. H., Coutant, B. W., and Hoff, R. M. (2004). Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality. *Atmospheric Environment, 38*(16), 2495-2509. doi:10.1016/j.atmosenv.2004.01.039

Engel-Cox, J. A., Young, G. S., and Hoff, R. M. (2005). Application of satellite remote-sensing data for source analysis of fine particulate matter transport events. *Journal of the Air and Waste Management Association, 55*(9), 1389-1397.

Evans, J., van Donkelaar, A., Martin, R. V., Burnett, R., Rainham, D. G., Birkett, N. J., and Krewski, D. (2013). Estimates of global mortality attributable to particulate air pollution using satellite imagery. *Environ Res, 120*, 33-42. doi:10.1016/j.envres.2012.08.005

Fiore, A. M., Naik, V., and Leibensperger, E. M. (2015). Air quality and climate connections. *Journal of the Air & Waste Management Association (1995), 65*(6), 645-685. doi:10.1080/10962247.2015.1040526

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment, 80*(1), 185-201.

Foody, G. M. (2009). Sample size determination for image classification accuracy assessment and comparison. *International Journal of Remote Sensing, 30*(20), 5273-5291.

Ford, B., and Heald, C. L. (2016). Exploring the uncertainty associated with satellite-based estimates of premature mortality due to exposure to fine particulate matter. *Atmospheric Chemistry and Physics, 16*(5), 3499-3523.

Fox, D. G. (1981). Judging air quality model performance. *Bulletin of the American Meteorological Society, 62*(5), 599-609.

Gao, M., Guttikunda, S. K., Carmichael, G. R., Wang, Y., Liu, Z., Stanier, C. O., Saide, P. E., and Yu, M. (2015). Health impacts and economic losses assessment of the 2013 severe haze event in Beijing area. *Science of the Total Environment, 511*, 553-561.

Gao, Y., Zhao, C., Liu, X. H., Zhang, M. G., and Leung, L. R. (2014). WRF-Chem simulations of aerosols and anthropogenic aerosol radiative forcing in East Asia. *Atmospheric Environment, 92*, 250-266. doi:10.1016/j.atmosenv.2014.04.038

Geng, G. N., Zhang, Q., Martin, R. V., van Donkelaar, A., Huo, H., Che, H. Z., Lin, J. T., and He, K. B. (2015). Estimating long-term PM2.5 concentrations in China using satellite-based aerosol optical depth and a chemical transport model. *Remote Sensing of Environment, 166*, 262-270. doi:10.1016/j.rse.2015.05.016

Ginoux, P., Prospero, J. M., Gill, T. E., Hsu, N. C., and Zhao, M. (2012). Global□ scale attribution of anthropogenic and natural dust sources and their emission rates based on MODIS Deep Blue aerosol products. *Reviews of Geophysics, 50*(3).

Glantz, P., Kokhanovsky, A., von Hoyningen-Huene, W., and Johansson, C. (2009). Estimating PM2.5 over southern Sweden using space-borne optical measurements. *Atmospheric Environment, 43*(36), 5838-5846. doi:10.1016/j.atmosenv.2009.05.017

Green, M., Kondragunta, S., Ciren, P., and Xu, C. (2009). Comparison of GOES and MODIS Aerosol Optical Depth (AOD) to Aerosol Robotic Network (AERONET) AOD and IMPROVE PM2.5 mass at Bondville, Illinois. *Journal of the Air and Waste Management Association, 59*(9), 1082-1091. doi:10.3155/1047-3289.59.9.1082

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B. (2005). Fully coupled "online" chemistry within the WRF model. *Atmospheric Environment, 39*(37), 6957-6975.

Grosso, N., and Paronis, D. (2012). Comparison of contrast reduction based MODIS AOT estimates with AERONET measurements. *Atmospheric Research, 116*, 33-45. doi:10.1016/j.atmosres.2011.09.008

Guo, H., Cheng, T., Gu, X., Chen, H., Wang, Y., Zheng, F., and Xiang, K. (2016). Comparison of four ground-level PM2.5 estimation models using PARASOL aerosol optical depth data from China. *International Journal of Environmental Research and Public Health, 13*(2). doi:10.3390/ijerph13020180

Guo, J., Xia, F., Zhang, Y., Liu, H., Li, J., Lou, M., He, J., Yan, Y., Wang, F., Min, M., and Zhai, P. (2017). Impact of diurnal variability and meteorological factors on the PM2.5 - AOD relationship: Implications for PM2.5 remote sensing. *Environ Pollut, 221*, 94-104. doi:10.1016/j.envpol.2016.11.043

Guo, J. P., Wu, Y. R., Zhang, X. Y., and Li, X. W. (2013). Estimation of PM2.5 over eastern China from MODIS aerosol optical depth using the back propagation neural network. *Huanjing Kexue/Environmental Science, 34*(3), 817-825. doi:10.1029/2002JD003179

Guo, Y., Feng, N., Christopher, S. A., Kang, P., Zhan, F. B., and Hong, S. (2014). Satellite remote sensing of fine particulate matter (PM2.5) air quality over Beijing using MODIS. *International Journal of Remote Sensing, 35*(17), 6522-6544. doi:10.1080/01431161.2014.958245

Gupta, P., and Christopher, S. A. (2008a). An evaluation of Terra-MODIS sampling for monthly and annual particulate matter air quality assessment over the Southeastern United States. *Atmospheric Environment, 42*(26), 6465-6471. doi:10.1016/j.atmosenv.2008.04.044

Gupta, P., and Christopher, S. A. (2008b). Seven year particulate matter air quality assessment from surface and satellite measurements. *Atmospheric Chemistry and Physics, 8*(12), 3311-3324. doi:10.5194/acp-8-3311-2008

Gupta, P., and Christopher, S. A. (2009). Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: Multiple regression approach. *Journal of Geophysical Research Atmospheres, 114*(14). doi:10.1029/2008JD011496

Gupta, P., Christopher, S. A., Wang, J., Gehrig, R., Lee, Y., and Kumar, N. (2006). Satellite remote sensing of particulate matter and air quality assessment over global cities. *Atmospheric Environment, 40*(30), 5880-5892. doi:10.1016/j.atmosenv.2006.03.016

Gutierrez, E. (2010). Using satellite imagery to measure the relationship between air quality and infant mortality: An empirical study for Mexico. *Population and Environment, 31*(4), 203-222. doi:10.1007/s11111-009-0096-y

Haberle, S. G., Hope, G. S., and van der Kaars, S. (2001). Biomass burning in Indonesia and Papua New Guinea: natural and human induced fire events in

the fossil record. *Palaeogeography, Palaeoclimatology, Palaeoecology, 171*(3), 259-268.

Hansen, A., Bi, P., Nitschke, M., Pisaniello, D., Ryan, P., Sullivan, T., and Barnett, A. G. (2012). Particulate air pollution and cardiorespiratory hospital admissions in a temperate Australian city: a case-crossover analysis. *Science of the Total Environment, 416*, 48-52.

He, J., Zha, Y., Zhang, J., Gao, J., Li, Y., and Chen, X. (2015). Retrieval of aerosol optical thickness from HJ-1 CCD data based on MODIS-derived surface reflectance. *International Journal of Remote Sensing, 36*(3), 882-898.

He, X., Pan, D., Bai, Y., Zhu, Q., and Gong, F. (2011). Evaluation of the aerosol models for SeaWiFS and MODIS by AERONET data over open oceans. *Applied Optics, 50*(22), 4353-4364.

Ho, A. F. W., Wah, W., Earnest, A., Ng, Y. Y., Xie, Z., Shahidah, N., Yap, S., Pek, P. P., Liu, N., Lam, S. S. W., Ong, M. E. H., and Singapore, P. I. (2018). Health impacts of the Southeast Asian haze problem - A time-stratified case crossover study of the relationship between ambient air pollution and sudden cardiac deaths in Singapore. *Int J Cardiol, 271*, 352-358. doi:10.1016/j.ijcard.2018.04.070

Hod, R. (2016). The Impact of Air Pollution and Haze on Hospital Admission for Cardiovascular and Respiratory Diseases. *International Journal of Public Health Research, 6*(1), 707-712.

Hoek, G., Krishnan, R. M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B., and Kaufman, J. D. (2013). Long-term air pollution exposure and cardio- respiratory mortality: a review. *Environmental Health, 12*(1), 43.

Hoff, R. M., and Christopher, S. A. (2012). Remote Sensing of Particulate Pollution from Space: Have We Reached the Promised Land? *Journal of the Air & Waste Management Association, 59*(6), 645-675. doi:10.3155/1047-3289.59.6.645

Holben, B. N., Eck, T., Slutsker, I., Tanre, D., Buis, J., Setzer, A., Vermote, E., Reagan, J., Kaufman, Y., and Nakajima, T. (1998). AERONET—A federated instrument network and data archive for aerosol characterization. *Remote Sensing of Environment, 66*(1), 1-16.

Hsu, N., Jeong, M. J., Bettenhausen, C., Sayer, A., Hansell, R., Seftor, C., Huang, J., and Tsay, S. C. (2013). Enhanced Deep Blue aerosol retrieval algorithm: The second generation. *Journal of Geophysical Research: Atmospheres, 118*(16), 9296-9315.

Hsu, N. C., Tsay, S.-C., King, M. D., and Herman, J. R. (2004). Aerosol properties over bright-reflecting source regions. *Ieee Transactions on Geoscience and Remote Sensing, 42*(3), 557-569.

Hu, R. M., Sokhi, R. S., and Fisher, B. E. A. (2009). New algorithms and their application for satellite remote sensing of surface PM2.5 and aerosol absorption. *Journal of Aerosol Science, 40*(5), 394-402. doi:10.1016/j.jaerosci.2009.01.005

Hu, X., Waller, L. A., Al-Hamdan, M. Z., Crosson, W. L., Estes, M. G., Jr., Estes, S. M., Quattrochi, D. A., Sarnat, J. A., and Liu, Y. (2013). Estimating ground-level PM(2.5) concentrations in the southeastern U.S. using geographically weighted regression. *Environ Res, 121*, 1-10. doi:10.1016/j.envres.2012.11.003

Hu, X., Waller, L. A., Lyapustin, A., Wang, Y., and Liu, Y. (2014). Improving satellite-driven PM2.5 models with Moderate Resolution Imaging Spectroradiometer fire counts in the southeastern U.S. *Journal of Geophysical Research: Atmospheres, 119*(19), 11375-11386. doi:10.1002/2014JD021920

Hu, Z. (2009). Spatial analysis of MODIS aerosol optical depth, PM2.5, and chronic coronary heart disease. *International Journal of Health Geographics, 8*(1). doi:10.1186/1476-072X-8-27

Huang, C.-Y., Ho, H.-C., and Lin, T.-H. (2018). Improving the image fusion procedure for high-spatiotemporal aerosol optical depth retrieval: a case study of urban area in Taiwan. *Journal of Applied Remote Sensing, 12*(4), 042605.

Huang, G. H., Huang, C. L., Li, Z. Q., and Chen, H. (2015). Development and Validation of a Robust Algorithm for Retrieving Aerosol Optical Depth Over Land From MODIS Data. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8*(3), 1152-1166. doi:10.1109/jstars.2015.2396491

Hutchison, K. D. (2003). Applications of MODIS satellite data and products for monitoring air quality in the state of Texas. *Atmospheric Environment, 37*(17), 2403-2412. doi:10.1016/S1352-2310(03)00128-6

Hutchison, K. D., Smith, S., and Faruqui, S. J. (2005). Correlating MODIS aerosol optical thickness data with ground-based PM2.5 observations across Texas for use in a real-time air quality prediction system. *Atmospheric Environment, 39*(37), 7190-7203. doi:10.1016/j.atmosenv.2005.08.036

Hystad, P., Setton, E., Cervantes, A., Poplawski, K., Deschenes, S., Brauer, M., van Donkelaar, A., Lamsal, L., Martin, R., and Jerrett, M. (2015). *Creating national air pollution models for population exposure assessment in Canada.* University of British Columbia,

Ichoku, C., Allen Chu, D., Mattoo, S., Kaufman, Y. J., Remer, L. A., Tanré, D., Slutsker, I., and Holben, B. N. (2002). A spatio-temporal approach for global validation and analysis of MODIS aerosol products. *Geophysical Research Letters, 29*(12), 1-1 - 1-4.

Ichoku, C., Chu, D. A., Mattoo, S., Kaufman, Y. J., Remer, L. A., Tanre, D., Slutsker, I., and Holben, B. N. (2001). *Techniques of global validation of aerosol retrievals from MODIS.* Paper presented at the Geoscience and Remote Sensing Symposium, 2001. IGARSS'01. IEEE 2001 International.

Jaafar, S. A., Latif, M. T., Razak, I. S., Wahid, N. B. A., Khan, M. F., and Srithawirat, T. (2018). Composition of carbohydrates, surfactants, major elements and anions in PM 2.5 during the 2013 Southeast Asia high pollution episode in Malaysia. *Particuology, 37*, 119-126. doi:10.1016/j.partic.2017.04.012

Jamil, A., Ahmad, M. A., Raheleh, F., Muhamad, F. R., and Saeid, P. (2011). PM10 distribution using remotely sensed data and GIS techniques; Klang Valley, Malaysia. *Environment Asia, 4*(1), 47-52.

Jantarach, T., Masiri, I., and Janjai, S. (2012). Comparison of MODIS aerosol optical depth retrievals with ground-based measurements in the tropics. *Procedia Engineering, 32*, 392-398.

Jin, M., Shepherd, J. M., and Zheng, W. (2011). Urban surface temperature reduction via the urban aerosol direct effect: A remote sensing and WRF model sensitivity study. *Advances in Meteorology, 2010*.

Jing-Mei, Y., Jin-Huan, Q., and Yan-Liang, Z. (2010). Validation of aerosol optical depth from Terra and Aqua MODIS retrievals over a tropical coastal site in China. *Atmospheric and Oceanic Science Letters, 3*(1), 36-39.

Just, A. C., Wright, R. O., Schwartz, J., Coull, B. A., Baccarelli, A. A., Tellez-Rojo, M. M., Moody, E., Wang, Y. J., Lyapustin, A., and Kloog, I. (2015). Using High-Resolution Satellite Aerosol Optical Depth To Estimate Daily PM2.5 Geographical Distribution in Mexico City. *Environmental Science & Technology, 49*(14), 8576-8584. doi:10.1021/acs.est.5b00859

Justice, C. O., Vermote, E., Townshend, J. R., Defries, R., Roy, D. P., Hall, D. K., Salomonson, V. V., Privette, J. L., Riggs, G., and Strahler, A. (1998). The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *Ieee Transactions on Geoscience and Remote Sensing, 36*(4), 1228-1249.

Kahn, R. A., Gaitley, B. J., Martonchik, J. V., Diner, D. J., Crean, K. A., and Holben, B. (2005). Multiangle Imaging Spectroradiometer (MISR) global aerosol optical depth validation based on 2 years of coincident Aerosol Robotic Network (AERONET) observations. *Journal of Geophysical Research: Atmospheres, 110*(D10).

Kahn, R. A., Li, W. H., Moroney, C., Diner, D. J., Martonchik, J. V., and Fishbein, E. (2007). Aerosol source plume physical characteristics from space□based multiangle imaging. *Journal of Geophysical Research: Atmospheres, 112*(D11).

Kan, H., Chen, B., and Hong, C. (2009). Health impact of outdoor air pollution in China: current knowledge and future research needs. In: National Institute of Environmental Health Sciences.

Kanniah, K. D., Kaskaoutis, D. G., San Lim, H., Latif, M. T., Kamarul Zaman, N. A. F., and Liew, J. (2016). Overview of atmospheric aerosol studies in Malaysia: Known and unknown. *Atmospheric Research, 182*, 302-318. doi:10.1016/j.atmosres.2016.08.002

Kanniah, K. D., Lim, H. Q., Kaskaoutis, D. G., and Cracknell, A. P. (2014). Investigating aerosol properties in Peninsular Malaysia via the synergy of satellite remote sensing and ground-based measurements. *Atmospheric Research, 138*, 223-239. doi:10.1016/j.atmosres.2013.11.018

Kanniah, K. D., and Yaso, N. (2010). *Preliminary analysis of the spatial and temporal patterns of aerosols and their impact on climate in Malaysia using MODIS satellite data.* Paper presented at the ISPRS Technical Commission VIII Symposium on Networking the World with Remote Sensing.

Karimian, H., Li, Q., Li, C., Jin, L., Fan, J., and Li, Y. (2016). An improved method for monitoring fine particulate matter mass concentrations via satellite remote sensing. *Aerosol and Air Quality Research, 16*(4), 1081-1092. doi:10.4209/aaqr.2015.06.0424

Kaufman, Y. J., Gobron, N., Pinty, B., Widlowski, J. L., and Verstraete, M. M. (2002). Relationship between surface reflectance in the visible and mid□IR used in MODIS aerosol algorithm□theory. *Geophysical Research Letters, 29*(23).

Kaufman, Y. J., and Sendra, C. (1988). Algorithm for automatic atmospheric corrections to visible and near-IR satellite imagery. *International Journal of Remote Sensing, 9*(8), 1357-1381.

Khan, M. F., Sulong, N. A., Latif, M. T., Nadzir, M. S. M., Amil, N., Hussain, D. F. M., Lee, V., Hosaini, P. N., Shaharom, S., and Yusoff, N. A. Y. M. (2016). Comprehensive assessment of PM2. 5 physicochemical properties during the Southeast Asia dry season (southwest monsoon). *Journal of Geophysical Research: Atmospheres, 121*(24).

Khoshsima, M., Ahmadi-Givi, F., Bidokhti, A. A., and Sabetghadam, S. (2014). Impact of meteorological parameters on relation between aerosol optical indices and air pollution in a sub-urban area. *Journal of Aerosol Science, 68*, 46-57. doi:10.1016/j.jaerosci.2013.10.008

Kim, H. S., Chung, Y. S., and Kim, J. T. (2014). Spatio-temporal variations of optical properties of aerosols in East Asia measured by MODIS and relation to the ground-based mass concentrations observed in central Korea during 2001 similar to 2010. *Asia-Pacific Journal of Atmospheric Sciences, 50*(2), 191-200. doi:10.1007/s13143-014-0007-8

Kim, M., Zhang, X., Holt, J. B., and Liu, Y. (2013). Spatio-Temporal Variations in the Associations between Hourly PM2.5 and Aerosol Optical Depth (AOD) from MODIS Sensors on Terra and Aqua. *Health (Irvine Calif), 5*(10A2), 8-13. doi:10.4236/health.2013.510A2002

King, M. D., Kaufman, Y. J., Menzel, W. P., and Tanre, D. (1992). Remote sensing of cloud, aerosol, and water vapor properties from the Moderate Resolution Imaging Spectrometer (MODIS). *Ieee Transactions on Geoscience and Remote Sensing, 30*(1), 2-27.

Kittaka, C., Szykman, J., Pierce, B., Al-Saadi, J., Neil, D., Chu, A., Remer, L., Prins, E., and Holdzkom, J. (2004). Utilizing MODIS satellite observations to monitor and analyze fine particulate matter, PM2.5, transport event. *Combined Preprints: 84th American Meteorological Society (AMS) Annual Meeting*, 6373-6377.

Kloog, I., Koutrakis, P., Coull, B. A., Lee, H. J., and Schwartz, J. (2011). Assessing temporally and spatially resolved PM2.5 exposures for epidemiological studies using satellite aerosol optical depth measurements. *Atmospheric Environment, 45*(35), 6267-6275. doi:10.1016/j.atmosenv.2011.08.066

Kloog, I., Nordio, F., Coull, B. A., and Schwartz, J. (2012). Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal PM2.5 exposures in the mid-atlantic states. *Environmental Science and Technology, 46*(21), 11913-11921. doi:10.1021/es302673e

Kloog, I., Sorek-Hamer, M., Lyapustin, A., Coull, B., Wang, Y., Just, A. C., Schwartz, J., and Broday, D. M. (2015). Estimating daily PM2.5 and PM10 across the complex geo-climate region of Israel using MAIAC satellite-based AOD data. *Atmos Environ (1994), 122*, 409-416. doi:10.1016/j.atmosenv.2015.10.004

Koelemeijer, R. B. A., Homan, C. D., and Matthijsen, J. (2006). Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe. *Atmospheric Environment, 40*(27), 5304-5315. doi:10.1016/j.atmosenv.2006.04.044

Kokhanovsky, A. A., Deuzé, J., Diner, D., Dubovik, O., Ducos, F., Emde, C., Garay, M., Grainger, R., Heckel, A., and Herman, M. (2010). The inter-comparison of major satellite aerosol retrieval algorithms using simulated intensity and polarization characteristics of reflected light. *Atmospheric Measurement Techniques, 3*(4), 909.

Kong, L., Xin, J., Zhang, W., and Wang, Y. (2016). The empirical correlations between PM2.5, PM10 and AOD in the Beijing metropolitan region and the PM2.5, PM10 distributions retrieved by MODIS. *Environmental Pollution, 216*, 350-360. doi:10.1016/j.envpol.2016.05.085

Kumar, N., Chu, A., and Foster, A. (2007). An empirical relationship between PM 2.5 and aerosol optical depth in Delhi Metropolitan. *Atmospheric Environment, 41*(21), 4492-4503.

Kumar, N., Chu, A., and Foster, A. (2008). Remote sensing of ambient particles in Delhi and its environs: estimation and validation. *International Journal of Remote Sensing, 29*(12), 3383-3405. doi:10.1080/01431160701474545

Kumar, N., Liang, D., Comellas, A., Chu, A. D., and Abrams, T. (2013). Satellite-based PM concentrations and their application to COPD in Cleveland, OH. *Journal of Exposure Science and Environmental Epidemiology, 23*(6), 637-646.

Lai, H. K., Tsang, H., Thach, T. Q., and Wong, C. M. (2014). Health impact assessment of exposure to fine particulate matter based on satellite and meteorological information. *Environmental Sciences: Processes and Impacts, 16*(2), 239-246. doi:10.1039/c3em00357d

Latif, M. T., Othman, M., Idris, N., Juneng, L., Abdullah, A. M., Hamzah, W. P., Khan, M. F., Nik Sulaiman, N. M., Jewaratnam, J., Aghamohammadi, N., Sahani, M., Xiang, C. J., Ahamad, F., Amil, N., Darus, M., Varkkey, H., Tangang, F., and Jaafar, A. B. (2018). Impact of regional haze towards air quality in Malaysia: A review. *Atmospheric Environment, 177*, 28-44. doi:10.1016/j.atmosenv.2018.01.002

Lee, H. J., Chatfield, R. B., and Strawa, A. W. (2016). Enhancing the Applicability of Satellite Remote Sensing for PM2.5 Estimation Using MODIS Deep Blue AOD and Land Use Regression in California, United States. *Environmental Science and Technology, 50*(12), 6546-6555. doi:10.1021/acs.est.6b01438

Lee, H. J., Coull, B. A., Bell, M. L., and Koutrakis, P. (2012). Use of satellite-based aerosol optical depth and spatial clustering to predict ambient PM2.5 concentrations. *Environmental Research, 118*, 8-15. doi:10.1016/j.envres.2012.06.011

Lee, H. J., Liu, Y., Coull, B. A., Schwartz, J., and Koutrakis, P. (2011). A novel calibration approach of MODIS AOD data to predict PM2.5 concentrations. *Atmospheric Chemistry and Physics, 11*(15), 7991-8002. doi:10.5194/acp-11-7991-2011

Lee, M., Koutrakis, P., Coull, B., Kloog, I., and Schwartz, J. (2016). Acute effect of fine particulate matter on mortality in three Southeastern states from 2007–2011. *Journal of Exposure Science and Environmental Epidemiology, 26*(2), 173.

Léon, J.-F., Liousse, C., Galy-Lacaux, C., Doumbia, T., and Cachier, H. (2010). *Monitoring of ambient fine particulate matter concentrations from space: application to European and African cities.* Paper presented at the Remote Sensing.

Leroy, M., Deuze, J., Bréon, F., Hautecoeur, O., Herman, M., Buriez, J., Tanré, D., Bouffies, S., Chazette, P., and Roujean, J. (1997). Retrieval of atmospheric properties and surface bidirectional reflectances over land from

POLDER/ADEOS. *Journal of Geophysical Research: Atmospheres, 102*(D14), 17023-17037.

Levy, R., Mattoo, S., Munchak, L., Remer, L., Sayer, A., Patadia, F., and Hsu, N. (2013). The Collection 6 MODIS aerosol products over land and ocean. *Atmos. Meas. Tech, 6*(11), 2989-3034.

Levy, R. C., Remer, L., Martins, J., Kaufman, Y., Plana-Fattori, A., Redemann, J., and Wenny, B. (2005). Evaluation of the MODIS aerosol retrievals over ocean and land during CLAMS. *Journal of the Atmospheric Sciences, 62*(4), 974-992.

Levy, R. C., Remer, L. A., and Dubovik, O. (2007). Global aerosol optical properties and application to Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land. *Journal of Geophysical Research: Atmospheres, 112*(D13).

Levy, R. C., Remer, L. A., Kleidman, R. G., Mattoo, S., Ichoku, C., Kahn, R., and Eck, T. F. (2010). Global evaluation of the Collection 5 MODIS dark-target aerosol products over land. *Atmospheric Chemistry and Physics, 10*(21), 10399-10420. doi:10.5194/acp-10-10399-2010

Levy, R. C., Remer, L. A., Mattoo, S., Vermote, E. F., and Kaufman, Y. J. (2006). A new algorithm for retrieving aerosol properties over land from MODIS spectral reflectance.

Levy, R. C., Remer, L. A., Tanré, D., Mattoo, S., and Kaufman, Y. J. (2009). Algorithm for remote sensing of tropospheric aerosol over dark targets from MODIS: collections 005 and 051: Revision 2; Feb 2009. *Download from http://modisatmos.gsfc.nasa.gov/_docs/ATBD_MOD04_C005_rev2.pdf.*

Li, C., Mao, J., Lau, A. K., Yuan, Z., Wang, M., and Liu, X. (2005). Application of MODIS satellite products to the air pollution research in Beijing. *Science in China Series D(Earth Sciences), 48*, 209-219.

Li, H., Faruque, F., Williams, W., Al-Hamdan, M., Luvall, J., Crosson, W., Rickman, D., and Limaye, A. (2009). Optimal temporal scale for the correlation of AOD and ground measurements of PM2.5 in a real-time air quality estimation system. *Atmospheric Environment, 43*(28), 4303-4310. doi:10.1016/j.atmosenv.2009.06.004

Li, J., Carlson, B. E., and Lacis, A. A. (2015). How well do satellite AOD observations represent the spatial and temporal variability of PM2.5 concentration for the United States? *Atmospheric Environment, 102*, 260-273. doi:10.1016/j.atmosenv.2014.12.010

Li, L., Huang, C., Huang, H. Y., Wang, Y. J., Yan, R. S., Zhang, G. F., Zhou, M., Lou, S. R., Tao, S. K., Wang, H. L., Qiao, L. P., Chen, C. H., Streets, D. G., and Fu, J. S. (2014). An integrated process rate analysis of a regional fine particulate matter episode over yangtze river delta in 2010. *Atmospheric Environment, 91*, 60-70. doi:10.1016/j.atmosenv.2014.03.053

Li, L., Yang, J., and Wang, Y. (2015). Retrieval of high-resolution atmospheric particulate matter concentrations from satellite-based aerosol optical thickness over the Pearl River Delta Area, China. *Remote Sensing, 7*(6), 7914-7937.

Li, S., Joseph, E., and Min, Q. (2016). Remote sensing of ground-level PM2.5 combining {AOD} and backscattering profile. *Remote Sensing of Environment, 183*, 120-128. doi:https://doi.org/10.1016/j.rse.2016.05.025

Li, Y., Xue, Y., de Leeuw, G., Li, C., Yang, L., Hou, T., and Marir, F. (2013). Retrieval of aerosol optical depth and surface reflectance over land from NOAA AVHRR data. *Remote Sensing of Environment, 133*, 1-20.

Li, Y., Xue, Y., He, X., and Guang, J. (2012). High-resolution aerosol remote sensing retrieval over urban areas by synergetic use of HJ-1 CCD and MODIS data. *Atmospheric Environment, 46*, 173-180.

Lin, C., Li, Y., Lau, A. K., Deng, X., Tim, K., Fung, J. C., Li, C., Li, Z., Lu, X., and Zhang, X. (2016). Estimation of long-term population exposure to PM2. 5 for dense urban areas using 1-km MODIS data. *Remote Sensing of Environment, 179*, 13-22.

Lin, C., Li, Y., Lau, A. K. H., Deng, X., Tse, T. K. T., Fung, J. C. H., Li, C., Li, Z., Lu, X., Zhang, X., and Yu, Q. (2016). Estimation of long-term population exposure to PM2.5 for dense urban areas using 1-km MODIS data. *Remote Sensing of Environment, 179*, 13-22. doi:10.1016/j.rse.2016.03.023

Liu, H., Fang, C., Zhang, X., Wang, Z., Bao, C., and Li, F. (2017). The effect of natural and anthropogenic factors on haze pollution in Chinese cities: A spatial econometrics approach. *Journal of Cleaner Production, 165*, 323-333. doi:10.1016/j.jclepro.2017.07.127

Liu, J., Weng, F., and Li, Z. (2019). Satellite-based PM2. 5 estimation directly from reflectance at the top of the atmosphere using a machine learning algorithm. *Atmospheric Environment, 208*, 113-122.

Liu, Y., Franklin, M., Kahn, R., and Koutrakis, P. (2007). Using aerosol optical thickness to predict ground-level PM2.5 concentrations in the St. Louis area: A comparison between MISR and MODIS. *Remote Sensing of Environment, 107*(1-2), 33-44. doi:10.1016/j.rse.2006.05.022

Liu, Y., He, K., Li, S., Wang, Z., Christiani, D. C., and Koutrakis, P. (2012). A statistical model to evaluate the effectiveness of PM 2.5 emissions control during the Beijing 2008 Olympic Games. *Environment International, 44*, 100-105.

Liu, Y., Koutrakis, P., and Kahn, R. (2007). Estimating fine particulate matter component concentrations and size distributions using satellite-retrieved fractional aerosol optical depth: Part 1—Method development. *Journal of the Air & Waste Management Association, 57*(11), 1351-1359.

Liu, Y., Paciorek, C. J., and Koutrakis, P. (2009). Estimating regional spatial and temporal variability of PM2.5 concentrations using satellite data, meteorology, and land use information. *Environmental Health Perspectives, 117*(6), 886-892. doi:10.1289/ehp.0800123

Liu, Y., Park, R. J., Jacob, D. J., Li, Q., Kilaru, V., and Sarnat, J. A. (2004). Mapping annual mean ground□level PM2. 5 concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States. *Journal of Geophysical Research: Atmospheres, 109*(D22).

Liu, Y., Sarnat, J. A., Kilaru, V., Jacob, D. J., and Koutrakis, P. (2005). Estimating ground-level PM2.5 in the eastern United States using satellite remote sensing. *Environmental Science and Technology, 39*(9), 3269-3278. doi:10.1021/es049352m

Lopez-Villarrubia, E., Iniguez, C., Peral, N., Garcia, M. D., and Ballester, F. (2012). Characterizing mortality effects of particulate matter size fractions in the two capital cities of the Canary Islands. *Environ Res, 112*, 129-138. doi:10.1016/j.envres.2011.10.005

Lovallo, M., Marchese, F., Pergola, N., and Telesca, L. (2009). Fisher information measure of temporal fluctuations in satellite advanced very high resolution radiometer (AVHRR) thermal signals recorded in the volcanic area of Etna (Italy). *Communications in Nonlinear Science and Numerical Simulation, 14*(1), 174-181.

Luan, Y., and Jaeglé, L. (2013). Composite study of aerosol export events from East Asia and North America. *Atmospheric Chemistry and Physics, 13*(3), 1221-1242.

Lyapustin, A., Wang, Y., Laszlo, I., Kahn, R., Korkin, S., Remer, L., Levy, R., and Reid, J. (2011). Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm. *Journal of Geophysical Research: Atmospheres, 116*(D3).

Lyapustin, A., Wang, Y., Xiong, X., Meister, G., Platnick, S., Levy, R., Franz, B., Korkin, S., Hilker, T., and Tucker, J. (2014). Scientific impact of MODIS C5 calibration degradation and C6+ improvements.

Ma, Z., Hu, X., Huang, L., Bi, J., and Liu, Y. (2014). Estimating ground-level PM2. 5 in China using satellite remote sensing. *Environmental Science & Technology, 48*(13), 7436-7444.

Ma, Z., Hu, X., Sayer, A. M., Levy, R., Zhang, Q., Xue, Y., Tong, S., Bi, J., Huang, L., and Liu, Y. (2016). Satellite-Based Spatiotemporal Trends in PM2.5 Concentrations: China, 2004-2013. *Environ Health Perspect, 124*(2), 184-192. doi:10.1289/ehp.1409481

Ma, Z., Liu, Y., Zhao, Q., Liu, M., Zhou, Y., and Bi, J. (2016). Satellite-derived high resolution PM2.5 concentrations in Yangtze River Delta Region of China

using improved linear mixed effects model. *Atmospheric Environment, 133*, 156-164. doi:10.1016/j.atmosenv.2016.03.040

Maiersperger, T. K., Scaramuzza, P. L., Leigh, L., Shrestha, S., Gallo, K. P., Jenkerson, C. B., and Dwyer, J. L. (2013). Characterizing LEDAPS surface reflectance products by comparisons with AERONET, field spectrometer, and MODIS data. *Remote Sensing of Environment, 136*, 1-13. doi:10.1016/j.rse.2013.04.007

Mano, Y., Hashimoto, T., and Okuyama, A. (2009). Verification of satellite-derived aerosol optical thickness over land with AERONET data. *Papers in Meteorology and Geophysics, 60*, 7-16. doi:10.2467/mripapers.60.7

Mansha, M., and Ghauri, B. (2011). Assessment of fine particulate matter (PM2.5) in metropolitan Karachi through satellite and ground-based measurements. *Journal of Applied Remote Sensing, 5*(1). doi:10.1117/1.3625615

Mao, L., Qiu, Y., Kusano, C., and Xu, X. (2012). Predicting regional space–time variation of PM2. 5 with land-use regression model and MODIS data. *Environmental Science and Pollution Research, 19*(1), 128-138.

Mao, X., Shen, T., and Feng, X. (2017). Prediction of hourly ground-level PM2. 5 concentrations 3 days in advance using neural networks with satellite data in eastern China. *Atmospheric Pollution Research, 8*(6), 1005-1015.

Martin, R. V. (2008). Satellite remote sensing of surface air quality. *Atmospheric Environment, 42*(34), 7823-7843. doi:10.1016/j.atmosenv.2008.07.018

Masuda, K., Sasaki, M., Takashima, T., and Ishida, H. (1999). Use of polarimetric measurements of the sky over the ocean for spectral optical thickness retrievals. *Journal of Atmospheric and Oceanic Technology, 16*(7), 846-859.

Mathur, R. (2008). Estimating the impact of the 2004 Alaskan forest fires on episodic particulate matter pollution over the eastern United States through assimilation of satellite-derived aerosol optical depths in a regional air quality model. *Journal of Geophysical Research-Atmospheres, 113*(D17), 13. doi:10.1029/2007jd009767

McClellan, R. O. (2002). Setting ambient air quality standards for particulate matter. *Toxicology, 181*, 329-347.

Mei, L., Xue, Y., Kokhanovsky, A., von Hoyningen-Huene, W., de Leeuw, G., and Burrows, J. (2014). Retrieval of aerosol optical depth over land surfaces from AVHRR data. *Atmospheric Measurement Techniques, 7*(8), 2411-2420.

Mhawish, A., Banerjee, T., Broday, D. M., Misra, A., and Tripathi, S. N. (2017). Evaluation of MODIS Collection 6 aerosol retrieval algorithms over Indo-Gangetic Plain: Implications of aerosols types and mass loading. *Remote Sensing of Environment, 201*, 297-313. doi:10.1016/j.rse.2017.09.016

Misra, A., Jayaraman, A., and Ganguly, D. (2008). Validation of MODIS derived aerosol optical depth over Western India. *Journal of Geophysical Research: Atmospheres, 113*(D4).

Mukherjee, A., and Agrawal, M. (2017). A global perspective of fine particulate matter pollution and its health effects. In *Reviews of environmental contamination and toxicology* (pp. 5-51): Springer.

Myhre, G., Stordal, F., Johnsrud, M., Diner, D. J., Geogdzhayev, I. V., Haywood, J. M., Holben, B. N., Holzer-Popp, T., Ignatov, A., Kahn, R. A., Kaufman, Y. J., Loeb, N., Martonchik, J. V., Mishchenko, M. I., Nalli, N. R., Remer, L. A., Schroedter-Homscheidt, M., Tanre, D., Torres, O., and Wang, M. (2005). Intercomparison of satellite retrieved aerosol optical depth over ocean during the period September 1997 to December 2000. *Atmospheric Chemistry and Physics, 5*, 1697-1719.

Nakhle, M. M., Farah, W., Ziade, N., Abboud, M., Salameh, D., and Annesi-Maesano, I. (2015). Short-term relationships between emergency hospital admissions for respiratory and cardiovascular diseases and fine particulate air pollution in Beirut, Lebanon. *Environ Monit Assess, 187*(4), 196. doi:10.1007/s10661-015-4409-6

Natunen, A., Arola, A., Mielonen, T., Huttunen, J., Komppula, M., and Lehtinen, K. E. J. (2010). A multi-year comparison of PM2.5 and AOD for the helsinki region. *Boreal Environment Research, 15*(6), 544-552. doi:10.1029/2007JD009573

North, P., Brockmann, C., Fischer, J., Gomez-Chova, L., Grey, W., Heckel, A., Moreno, J., Preusker, R., and Regner, P. (2008). *MERIS/AATSR synergy algorithms for cloud screening, aerosol retrieval and atmospheric correction.* Paper presented at the Proc. 2nd MERIS/AATSR User Workshop, ESRIN, Frascati.

Organization, W. H. (2006). *Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide*: World Health Organization.

Othman, J., Sahani, M., Mahmud, M., and Ahmad, M. K. (2014). Transboundary smoke haze pollution in Malaysia: inpatient health impacts and economic valuation. *Environ Pollut, 189*, 194-201. doi:10.1016/j.envpol.2014.03.010

Ou, Y., Chen, F., Zhao, W., Yan, X., and Zhang, Q. (2017). Landsat 8-based inversion methods for aerosol optical depths in the Beijing area. *Atmospheric Pollution Research, 8*(2), 267-274.

Paciorek, C. J., and Liu, Y. (2009). Limitations of remotely sensed aerosol as a spatial proxy for fine particulate matter. *Environmental Health Perspectives, 117*(6), 904-909. doi:10.1289/ehp.0800360

Paciorek, C. J., Liu, Y., Moreno-Macias, H., and Kondragunta, S. (2008). Spatiotemporal associations between GOES aerosol optical depth retrievals and ground-level PM2.5. *Environmental Science and Technology, 42*(15), 5800-5806. doi:10.1021/es703181j

Pan, L., Che, H., Geng, F., Xia, X., Wang, Y., Zhu, C., Chen, M., Gao, W., and Guo, J. (2010). Aerosol optical properties based on ground measurements over the Chinese Yangtze Delta Region. *Atmospheric Environment, 44*(21), 2587-2596.

Pascal, M., Falq, G., Wagner, V., Chatignoux, E., Corso, M., Blanchard, M., Host, S., Pascal, L., and Larrieu, S. (2014). Short-term impacts of particulate matter (PM10, PM10-2.5, PM2.5) on mortality in nine French cities. *Atmospheric Environment, 95*, 175-184. doi:10.1016/j.atmosenv.2014.06.030

Philip, S., Martin, R. V., van Donkelaar, A., Lo, J. W.-H., Wang, Y., Chen, D., Zhang, L., Kasibhatla, P. S., Wang, S., and Zhang, Q. (2014). Global chemical composition of ambient fine particulate matter for exposure assessment. *Environmental Science & Technology, 48*(22), 13060-13068.

Pinto, J. P., and Grant, L. D. (1999). Approaches to monitoring of air pollutants and evaluation of health impacts produced by biomass burning. *Health Guidelines for Vegetation Fire Events: Background Papers*, 147-185.

Polichetti, G., Cocco, S., Spinali, A., Trimarco, V., and Nunziata, A. (2009). Effects of particulate matter (PM(10), PM(2.5) and PM(1)) on the cardiovascular system. *Toxicology, 261*(1-2), 1-8. doi:10.1016/j.tox.2009.04.035

Pope, R. J., Savage, N. H., Chipperfield, M. P., Arnold, S. R., and Osborn, T. J. (2014). The influence of synoptic weather regimes on UK air quality: Analysis of satellite column NO2. *Atmospheric Science Letters, 15*(3), 211-217. doi:10.1002/asl2.492

Pour□Biazar, A., Khan, M., Wang, L., Park, Y. H., Newchurch, M., McNider, R. T., Liu, X., Byun, D. W., and Cameron, R. (2011). Utilization of satellite observation of ozone and aerosols in providing initial and boundary condition for regional air quality studies. *Journal of Geophysical Research: Atmospheres, 116*(D18).

Prather, K. A., Hatch, C. D., and Grassian, V. H. (2008). Analysis of atmospheric aerosols. *Annu. Rev. Anal. Chem., 1*, 485-514.

Qiao, T., Zhao, M., Xiu, G., and Yu, J. (2016). Simultaneous monitoring and compositions analysis of PM 1 and PM 2.5 in Shanghai: Implications for characterization of haze pollution and source apportionment. *Science of the Total Environment, 557*, 386-394.

Radojevic, M. (2003). Chemistry of forest fires and regional haze with emphasis on Southeast Asia. *Pure and Applied Geophysics, 160*(1-2), 157-187.

Rajab, J. M., Lim, H. S., and MatJafri, M. Z. (2013). Monthly distribution of diurnal total column ozone based on the 2011 satellite data in Peninsular Malaysia. *The Egyptian Journal of Remote Sensing and Space Science, 16*(1), 103-109. doi:10.1016/j.ejrs.2013.04.003

Reid, J. S., Hyer, E. J., Johnson, R. S., Holben, B. N., Yokelson, R. J., Zhang, J., Campbell, J. R., Christopher, S. A., Di Girolamo, L., and Giglio, L. (2013). Observing and understanding the Southeast Asian aerosol system by remote sensing: An initial review and analysis for the Seven Southeast Asian Studies (7SEAS) program. *Atmospheric Research, 122*, 403-468.

Remer, L. A., Kaufman, Y., Tanré, D., Mattoo, S., Chu, D., Martins, J. V., Li, R.-R., Ichoku, C., Levy, R., and Kleidman, R. (2005). The MODIS aerosol algorithm, products, and validation. *Journal of the Atmospheric Sciences, 62*(4), 947-973.

Remer, L. A., Mattoo, S., Levy, R. C., and Munchak, L. A. (2013). MODIS 3 km aerosol product: algorithm and global perspective. *Atmospheric Measurement Techniques, 6*(7), 1829-1844. doi:10.5194/amt-6-1829-2013

Remer, L. A., Tanre, D., Kaufman, Y. J., Ichoku, C., Mattoo, S., Levy, R., Chu, D. A., Holben, B., Dubovik, O., and Smirnov, A. (2002). Validation of MODIS aerosol retrieval over ocean. *Geophysical Research Letters, 29*(12).

Rhee, J., and Im, J. (2014). Estimating high spatial resolution air temperature for regions with limited in situ data using MODIS products. *Remote Sensing, 6*(8), 7360-7378.

Román, R., Bilbao, J., and de Miguel, A. (2014). Uncertainty and variability in satellite-based water vapor column, aerosol optical depth and Angström exponent, and its effect on radiative transfer simulations in the Iberian Peninsula. *Atmospheric Environment, 89*, 556-569. doi:10.1016/j.atmosenv.2014.02.027

Sadavarte, P., Venkataraman, C., Cherian, R., Patil, N., Madhavan, B. L., Gupta, T., Kulkarni, S., Carmichael, G. R., and Adhikary, B. (2016). Seasonal differences in aerosol abundance and radiative forcing in months of contrasting emissions and rainfall over northern South Asia. *Atmospheric Environment, 125*, 512-523. doi:10.1016/j.atmosenv.2015.10.092

Sahani, M., Zainon, N. A., Wan Mahiyuddin, W. R., Latif, M. T., Hod, R., Khan, M. F., Tahir, N. M., and Chan, C.-C. (2014). A case-crossover analysis of forest fire haze events and mortality in Malaysia. *Atmospheric Environment, 96*, 257-265. doi:10.1016/j.atmosenv.2014.07.043

Salinas, S. V., Chew, B. N., Miettinen, J., Campbell, J. R., Welton, E. J., Reid, J. S., Liya, E. Y., and Liew, S. C. (2013a). Physical and optical characteristics of the October 2010 haze event over Singapore: A photometric and lidar analysis. *Atmospheric Research, 122*, 555-570.

Salinas, S. V., Chew, B. N., Mohamad, M., Mahmud, M., and Liew, S. C. (2013b). First measurements of aerosol optical depth and Angstrom exponent number from AERONET's Kuching site. *Atmospheric Environment, 78*, 231-241.

San José, R., Pérez, J., Morant, J., and González, R. (2008). Elevated PM10 and PM2. 5 concentrations in Europe: a model experiment with MM5-CMAQ and WRF-CHEM. *WIT Transactions on Ecology and the Environment, 116*, 3-12.

Saunders, R. O., Kahl, J. D. W., and Ghorai, J. K. (2014). Improved estimation of PM2.5 using Lagrangian satellite-measured aerosol optical depth. *Atmospheric Environment, 91*, 146-153. doi:10.1016/j.atmosenv.2014.03.060

Savtchenko, A., Ouzounov, D., Ahmad, S., Acker, J., Leptoukh, G., Koziana, J., and Nickless, D. (2004). Terra and Aqua MODIS products available from NASA GES DAAC. *Advances in Space Research, 34*(4), 710-714. doi:10.1016/j.asr.2004.03.012

Sayer, A., Hsu, N., Bettenhausen, C., Jeong, M., Holben, B., and Zhang, J. (2012). Global and regional evaluation of over-land spectral aerosol optical depth retrievals from SeaWiFS.

Schaap, M., Apituley, A., Timmermans, R. M. A., Koelemeijer, R. B. A., and De Leeuw, G. (2009). Exploring the relation between aerosol optical depth and PM2.5 at Cabauw, the Netherlands. *Atmospheric Chemistry and Physics, 9*(3), 909-925. doi:10.5194/acp-9-909-2009, 2009

Schaeffer, B. A., Hagy, J. D., Conmy, R. N., Lehrter, J. C., and Stumpf, R. P. (2012). An approach to developing numeric water quality criteria for coastal waters using the SeaWiFS satellite data record. *Environmental Science & Technology, 46*(2), 916-922.

Schwartz, C. S., Liu, Z., Lin, H. C., and McKeen, S. A. (2012). Simultaneous three-dimensional variational assimilation of surface fine particulate matter and MODIS aerosol optical depth. *Journal of Geophysical Research Atmospheres, 117*(13). doi:10.1029/2011JD017383

Semire, F. A., Mohd-Mokhtar, R., Ismail, W., Mohamad, N., and Mandeep, J. S. (2012). Ground validation of space-borne satellite rainfall products in Malaysia. *Advances in Space Research, 50*(9), 1241-1249. doi:10.1016/j.asr.2012.06.031

Shi, L., Zanobetti, A., Kloog, I., Coull, B. A., Koutrakis, P., Melly, S. J., and Schwartz, J. D. (2015). Low-concentration PM2. 5 and mortality: estimating acute and chronic effects in a population-based study. *Environmental Health Perspectives, 124*(1), 46-52.

Shon, Z. H. (2015). Long-term variations in PM2.5 emission from open biomass burning in Northeast Asia derived from satellite-derived data for 2000-2013. *Atmospheric Environment, 107*, 342-350. doi:10.1016/j.atmosenv.2015.02.038

Song, W., Jia, H., Huang, J., and Zhang, Y. (2014). A satellite-based geographically weighted regression model for regional PM2. 5 estimation over the Pearl River Delta region in China. *Remote Sensing of Environment, 154*, 1-7.

Song, Y. Z., Yang, H. L., Peng, J. H., Song, Y. R., Sun, Q., and Li, Y. (2015). Estimating PM2.5 concentrations in Xi'an City using a generalized additive model with multi-source monitoring data. *Plos One, 10*(11). doi:10.1371/journal.pone.0142149

Sorek-Hamer, M., Strawa, A. W., Chatfield, R. B., Esswein, R., Cohen, A., and Broday, D. M. (2013). Improved retrieval of PM2.5 from satellite data products using non-linear methods. *Environmental Pollution, 182*, 417-423. doi:10.1016/j.envpol.2013.08.002

Sreekanth, V., Mahesh, B., and Niranjan, K. (2017). Satellite remote sensing of fine particulate air pollutants over Indian mega cities. *Advances in Space Research, 60*(10), 2268-2276.

Stieb, D. M., Chen, L., Beckerman, B. S., Jerrett, M., Crouse, D. L., Omariba, D. W. R., Peters, P. A., van Donkelaar, A., Martin, R. V., and Burnett, R. T. (2015). Associations of pregnancy outcomes and PM2. 5 in a national Canadian study. *Environmental Health Perspectives, 124*(2), 243-249.

Strandgren, J., Mei, L., Vountas, M., Burrows, J., Lyapustin, A., and Wang, Y. (2014). Study of satellite retrieved aerosol optical depth spatial resolution effect on particulate matter concentration prediction. *Atmospheric Chemistry and Physics Discussions*(18), 25869-25899.

Suhaila, J., Deni, S. M., Zin, W. Z. W., and Jemain, A. A. (2010). Trends in Peninsular Malaysia rainfall data during the southwest monsoon and northeast monsoons seasons: 1975-2004. *Sains Malaysiana, 39*(4), 533-542.

Sun, J., Fu, J. S., Huang, K., and Gao, Y. (2015). Estimation of future PM2. 5-and ozone-related mortality over the continental United States in a changing climate: an application of high-resolution dynamical downscaling technique. *Journal of the Air & Waste Management Association, 65*(5), 611-623.

Tahir, N. M., Suratman, S., Fong, F. T., Hamzah, M. S., and Latif, M. T. (2013). Temporal distribution and chemical characterization of atmospheric particulate matter in the eastern coast of Peninsular Malaysia. *Aerosol and Air Quality Research, 13*(2), 584-595.

Tan, F., Lim, H., Abdullah, K., Yoon, T., and Holben, B. (2014). Variations in optical properties of aerosols on monsoon seasonal change and estimation of aerosol optical depth using ground-based meteorological and air quality data. *Atmospheric Chemistry & Physics Discussions, 14*(13).

Tan, F., San Lim, H., Abdullah, K., Yoon, T. L., and Holben, B. (2015). AERONET data–based determination of aerosol types. *Atmospheric Pollution Research, 6*(4), 682-695.

Tan, K. C., San Lim, H., and Jafri, M. Z. M. (2017). Study on solar ultraviolet erythemal dose distribution over Peninsular Malaysia using Ozone Monitoring Instrument. *The Egyptian Journal of Remote Sensing and Space Science*.

Tao, J., Zhang, M., Chen, L., Wang, Z., Su, L., Ge, C., Han, X., and Zou, M. (2013). A method to estimate concentrations of surface-level particulate matter using satellite-based aerosol optical thickness. *Science China Earth Sciences, 56*(8), 1422-1433.

Themistocleous, K., Hadjimitsis, D. G., Retalis, A., and Chrysoulakis, N. (2012). *The development of air quality indices through image-retrieved AOT and PM 10 measurements in Limassol Cyprus.* Paper presented at the Remote Sensing of Clouds and the Atmosphere XVII; and Lidar Technologies, Techniques, and Measurements for Atmospheric Remote Sensing VIII.

Tian, J., and Chen, D. (2010a). A semi-empirical model for predicting hourly ground-level fine particulate matter (PM2.5) concentration in southern Ontario from satellite remote sensing and ground-based meteorological measurements. *Remote Sensing of Environment, 114*(2), 221-229. doi:10.1016/j.rse.2009.09.011

Tian, J., and Chen, D. (2010b). Spectral, spatial, and temporal sensitivity of correlating MODIS aerosol optical depth with ground-based fine particulate matter (PM2.5) across southern Ontario. *Canadian Journal of Remote Sensing, 36*(2), 119-128. doi:10.5589/m10-033

Tian, J., and Chen, D. M. (2007). Evaluating satellite-based measurements for mapping air quality in Ontario, Canada. *Journal of Environmental Informatics, 10*(1), 30-36. doi:10.3808/jei.200700097

Torres, O., Bhartia, P., Herman, J., Sinyuk, A., Ginoux, P., and Holben, B. (2002). A long-term record of aerosol optical depth from TOMS observations and comparison to AERONET measurements. *Journal of the Atmospheric Sciences, 59*(3), 398-413.

Torres, O., Tanskanen, A., Veihelmann, B., Ahn, C., Braak, R., Bhartia, P. K., Veefkind, P., and Levelt, P. (2007). Aerosols and surface UV products from Ozone Monitoring Instrument observations: An overview. *Journal of Geophysical Research: Atmospheres, 112*(D24).

Toth, T. D., Zhang, J., Campbell, J. R., Hyer, E. J., Reid, J. S., Shi, Y., and Westphal, D. L. (2014). Impact of data quality and surface-to-column representativeness on the PM 2.5/satellite AOD relationship for the contiguous United States. *Atmospheric Chemistry and Physics, 14*(12), 6049-6062.

Tsai, T. C., Jeng, Y. J., Chu, D. A., Chen, J. P., and Chang, S. C. (2011). Analysis of the relationship between MODIS aerosol optical depth and particulate matter from 2006 to 2008. *Atmospheric Environment, 45*(27), 4777-4788. doi:10.1016/j.atmosenv.2009.10.006

Vadrevu, K. P., Lasko, K., Giglio, L., and Justice, C. (2014). Analysis of Southeast Asian pollution episode during June 2013 using satellite remote sensing datasets. *Environmental Pollution, 195*, 245-256.

van Donkelaar, A., Martin, R. V., Brauer, M., and Boys, B. L. (2015a). Use of satellite observations for long-term exposure assessment of global concentrations of fine particulate matter. *Environ Health Perspect, 123*(2), 135-143. doi:10.1289/ehp.1408646

van Donkelaar, A., Martin, R. V., Brauer, M., Hsu, N. C., Kahn, R. A., Levy, R. C., Lyapustin, A., Sayer, A. M., and Winker, D. M. (2016). Global Estimates of Fine Particulate Matter using a Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors. *Environ Sci Technol, 50*(7), 3762-3772. doi:10.1021/acs.est.5b05833

van Donkelaar, A., Martin, R. V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., and Villeneuve, P. J. (2010). Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: Development and application. *Environmental Health Perspectives, 118*(6), 847-855. doi:10.1289/ehp.0901623

van Donkelaar, A., Martin, R. V., and Park, R. J. (2006). Estimating ground-level PM2.5 using aerosol optical depth determined from satellite remote sensing. *Journal of Geophysical Research-Atmospheres, 111*(D21), 10. doi:10.1029/2005jd006996

van Donkelaar, A., Martin, R. V., Spurr, R. J. D., and Burnett, R. T. (2015b). High-Resolution Satellite-Derived PM2.5 from Optimal Estimation and Geographically Weighted Regression over North America. *Environmental Science & Technology, 49*(17), 10482-10491. doi:10.1021/acs.est.5b02076

Vijayaraghavan, K., Snell, H. E., and Seigneur, C. (2008). Practical aspects of using satellite data in air quality modeling. *Environmental Science and Technology, 42*(22), 8187-8192. doi:10.1021/es7031339

von Hoyningen-Huene, W., Freitag, M., and Burrows, J. B. (2003). Retrieval of aerosol optical thickness over land surfaces from top-of-atmosphere radiance. *Journal of Geophysical Research: Atmospheres, 108*(D9), n/a-n/a. doi:10.1029/2001jd002018

Wallace, J., and Kanaroglou, P. (2007). *An investigation of air pollution in southern Ontario, Canada, with MODIS and MISR aerosol data.* Paper presented at the 2007 IEEE International Geoscience and Remote Sensing Symposium.

Wang, B. Z., and Chen, Z. (2016). High-resolution satellite-based analysis of ground-level PM2.5 for the city of Montreal. *Science of the Total Environment, 541*, 1059-1069. doi:10.1016/j.scitotenv.2015.10.024

Wang, C., Liu, Q., Ying, N., Wang, X., and Ma, J. (2013). Air quality evaluation on an urban scale based on MODIS satellite images. *Atmospheric Research, 132-133*, 22-34. doi:10.1016/j.atmosres.2013.04.011

Wang, J. (2003). Intercomparison between satellite-derived aerosol optical thickness and PM2.5mass: Implications for air quality studies. *Geophysical Research Letters, 30*(21), ASC 4-1 - ASC 4-4. doi:10.1029/2003gl018174

Wang, J., and Martin, S. T. (2007). Satellite characterization of urban aerosols: Importance of including hygroscopicity and mixing state in the retrieval algorithms. *Journal of Geophysical Research: Atmospheres, 112*(D17).

Wang, K., Liu, J., Zhou, X., Sparrow, M., Ma, M., Sun, Z., and Jiang, W. (2004). Validation of MODIS global land surface albedo product using ground measurements in a semidesert region on the Tibetan Plateau. *Journal of Geophysical Research D: Atmospheres, 109*(5), D05107 05101-05109. doi:10.1029/2003JD004229

Wang, L. L., Wang, Y. S., Xin, J. Y., Li, Z. Q., and Wang, X. Y. (2010). Assessment and comparison of three years of Terra and Aqua MODIS Aerosol Optical Depth Retrieval (C005) in Chinese terrestrial regions. *Atmospheric Research, 97*(1-2), 229-240. doi:10.1016/j.atmosres.2010.04.004

Wang, M., Knobelspiesse, K. D., and McClain, C. R. (2005). Study of the Sea□ Viewing Wide Field□of□ View Sensor (SeaWiFS) aerosol optical property data over ocean in combination with the ocean color products. *Journal of Geophysical Research: Atmospheres, 110*(D10).

Wang, Q., Ni, J., and Tenhunen, J. (2005). Application of a geographically□weighted regression analysis to estimate net primary production of Chinese forest ecosystems. *Global ecology and biogeography, 14*(4), 379-393.

Wang, S., Fang, L., Gu, X., Yu, T., and Gao, J. (2011). Comparison of aerosol optical properties from Beijing and Kanpur. *Atmospheric Environment, 45*(39), 7406-7414.

Wang, Y., Chen, L., Li, S., Wang, X., Yu, C., Si, Y., and Zhang, Z. (2017). Interference of Heavy Aerosol Loading on the VIIRS Aerosol Optical Depth (AOD) Retrieval Algorithm. *Remote Sensing, 9*(4), 397.

Wang, Y., Kloog, I., Coull, B. A., Kosheleva, A., Zanobetti, A., and Schwartz, J. D. (2016). Estimating causal effects of long-term PM2. 5 exposure on mortality in New Jersey. *Environmental Health Perspectives, 124*(8), 1182-1188.

Wang, Z., Chen, L., Tao, J., Zhang, Y., and Su, L. (2010). Satellite-based estimation of regional particulate matter (PM) in Beijing using vertical-and-RH correcting method. *Remote Sensing of Environment, 114*(1), 50-63.

Watson, J. G. (2002). Visibility: Science and regulation. *Journal of the Air & Waste Management Association, 52*(6), 628-713.

Webster, R., and Beckett, P. (1968). Quality and usefulness of soil maps. *Nature, 219*(5155), 680.

Wilson, W. E., and Suh, H. H. (1997). Fine particles and coarse particles: concentration relationships relevant to epidemiologic studies. *Journal of the Air & Waste Management Association, 47*(12), 1238-1249.

Winker, D. M., Hunt, W. H., and McGill, M. J. (2007). Initial performance assessment of CALIOP. *Geophysical Research Letters, 34*(19).

Xie, Y. Y., Wang, Y. X., Zhang, K., Dong, W. H., Lv, B. L., and Bai, Y. Q. (2015). Daily Estimation of Ground-Level PM2.5 Concentrations over Beijing Using 3 km Resolution MODIS AOD. *Environmental Science & Technology, 49*(20), 12280-12288. doi:10.1021/acs.est.5b01413

Xin, J., Zhang, Q., Wang, L., Gong, C., Wang, Y., Liu, Z., and Gao, W. (2014). The empirical relationship between the PM2. 5 concentration and aerosol optical depth over the background of North China from 2009 to 2011. *Atmospheric Research, 138*, 179-188.

Xiong, X. Z., Chen, L. F., Liu, Y., Cortesi, U., and Gupta, P. (2015). Satellite Observation of Atmospheric Compositions for Air Quality and Climate Study. *Advances in Meteorology*, 2. doi:10.1155/2015/932012

Yaacob, W. F. W., Noor, N. S. M., Bakar, N. I. C. A., Zin, N. A. M. a., and Taib, F. (2016). The impact of haze on the adolescent's acute respiratory disease: A single institution study. *Journal of Acute Disease, 5*(3), 227-231.

Yadav, S., Praveen, O. D., and Satsangi, P. G. (2015). The effect of climate and meteorological changes on particulate matter in Pune, India. *Environ Monit Assess, 187*(7), 402. doi:10.1007/s10661-015-4634-z

Yadav, S., and Satsangi, P. G. (2013). Characterization of particulate matter and its related metal toxicity in an urban location in South West India. *Environ Monit Assess, 185*(9), 7365-7379. doi:10.1007/s10661-013-3106-6

Yan, X., Shi, W., Li, Z., Li, Z., Luo, N., Zhao, W., Wang, H., and Yu, X. (2017). Satellite-based PM 2.5 estimation using fine-mode aerosol optical thickness over China. *Atmospheric Environment, 170*, 290-302. doi:10.1016/j.atmosenv.2017.09.023

Yang, Y., Liao, H., and Lou, S. (2015). Decadal trend and interannual variation of outflow of aerosols from East Asia: Roles of variations in meteorological parameters and emissions. *Atmospheric Environment, 100*, 141-153. doi:10.1016/j.atmosenv.2014.11.004

Yao, F., Si, M., Li, W., and Wu, J. (2018). A multidimensional comparison between MODIS and VIIRS AOD in estimating ground-level PM2.5 concentrations over a heavily polluted region in China. *Sci Total Environ, 618*, 819-828. doi:10.1016/j.scitotenv.2017.08.209

Yao, L., Lu, N., and Jiang, S. (2012). *Artificial neural network (ANN) for multi-source PM2.5 estimation using surface, MODIS, and meteorological data.* Paper presented at the 2012 International Conference on Biomedical Engineering and Biotechnology, iCBEB 2012, Macau.

You, W., Zang, Z., Zhang, L., Li, Y., Pan, X., and Wang, W. (2016a). National-scale estimates of ground-level PM2.5 concentration in China using geographically weighted regression based on 3 km resolution MODIS AOD. *Remote Sensing, 8*(3). doi:10.3390/rs8030184

You, W., Zang, Z., Zhang, L., Li, Y., and Wang, W. (2016). Estimating national-scale ground-level PM25 concentration in China using geographically weighted regression based on MODIS and MISR AOD. *Environmental Science and Pollution Research, 23*(9), 8327-8338.

You, W., Zang, Z. L., Pan, X. B., Zhang, L. F., and Chen, D. (2015). Estimating PM2.5 in Xi'an, China using aerosol optical depth: A comparison between the MODIS and MISR retrieval models. *Science of the Total Environment, 505*, 1156-1165. doi:10.1016/j.scitotenv.2014.11.024

You, W., Zang, Z. L., Zhang, L. F., Li, Y., Pan, X. B., and Wang, W. Q. (2016b). National-Scale Estimates of Ground-Level PM2.5 Concentration in China Using Geographically Weighted Regression Based on 3 km Resolution MODIS AOD. *Remote Sensing, 8*(3), 13. doi:10.3390/rs8030184

Yu, C., Di Girolamo, L., Chen, L. F., Zhang, X. Y., and Liu, Y. (2015). Statistical evaluation of the feasibility of satellite-retrieved cloud parameters as indicators of PM2.5 levels. *Journal of Exposure Science and Environmental Epidemiology, 25*(5), 457-466. doi:10.1038/jes.2014.49

Yu, S., Mathur, R., Schere, K., Kang, D., Pleim, J., Young, J., Tong, D., Pouliot, G., McKeen, S. A., and Rao, S. (2008). Evaluation of real□time PM2. 5 forecasts and process analysis for PM2. 5 formation over the eastern United States using the Eta□CMAQ forecast model during the 2004 ICARTT study. *Journal of Geophysical Research: Atmospheres, 113*(D6).

Zhang, F., Xu, L., Chen, J., Chen, X., Niu, Z., Lei, T., Li, C., and Zhao, J. (2013). Chemical characteristics of PM 2.5 during haze episodes in the urban of Fuzhou, China. *Particuology, 11*(3), 264-272.

Zhang, H., Hoff, R. M., and Engel-Cox, J. A. (2009). The relation between moderate resolution imaging spectroradiometer (MODIS) aerosol optical depth and PM2.5 over the United States: A geographical comparison by U.S. Environmental Protection Agency regions. *Journal of the Air and Waste Management Association, 59*(11), 1358-1369. doi:10.3155/1047-3289.59.11.1358

Zhang, J., Reid, J. S., and Holben, B. N. (2005). An analysis of potential cloud artifacts in MODIS over ocean aerosol optical thickness products. *Geophysical Research Letters, 32*(15).

Zhang, X., Chu, Y., Wang, Y., and Zhang, K. (2018). Predicting daily PM 2.5 concentrations in Texas using high-resolution satellite aerosol optical depth. *Science of the Total Environment, 631-632*, 904-911. doi:10.1016/j.scitotenv.2018.02.255

Zhang, Y., and Li, Z. (2015). Remote sensing of atmospheric fine particulate matter (PM2.5) mass concentration near the ground from satellite observation. *Remote Sensing of Environment, 160*, 252-262. doi:10.1016/j.rse.2015.02.005

Zhao, N., Yang, Y., and Zhou, X. (2010). Application of geographically weighted regression in estimating the effect of climate and site conditions on vegetation distribution in Haihe Catchment, China. *Plant Ecology, 209*(2), 349-359.

Zheng, C., Zhao, C., Zhu, Y., Wang, Y., Shi, X., Wu, X., Chen, T., Wu, F., and Qiu, Y. (2017). Analysis of Influential Factors for the Relationship between PM2.5 and AOD in Beijing. *Atmos. Chem. Phys. Discuss., 2017*, 1-57. doi:10.5194/acp-2016-1170

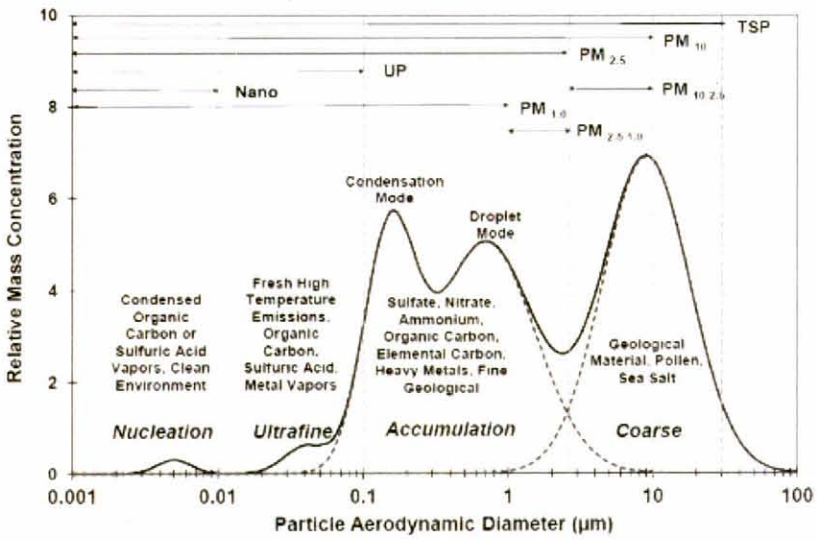
Zheng, Y. X., Zhang, Q., Liu, Y., Geng, G. N., and He, K. B. (2016). Estimating ground-level PM2.5 concentrations over three megalopolises in China using satellite-derived aerosol optical depth measurements. *Atmospheric Environment, 124*, 232-242. doi:10.1016/j.atmosenv.2015.06.046

Zoogman, P., Jacob, D. J., Chance, K., Zhang, L., Le Sager, P., Fiore, A. M., Eldering, A., Liu, X., Natraj, V., and Kulawik, S. S. (2011). Ozone air quality measurement requirements for a geostationary satellite mission. *Atmospheric Environment, 45*(39), 7143-7150. doi:10.1016/j.atmosenv.2011.05.058

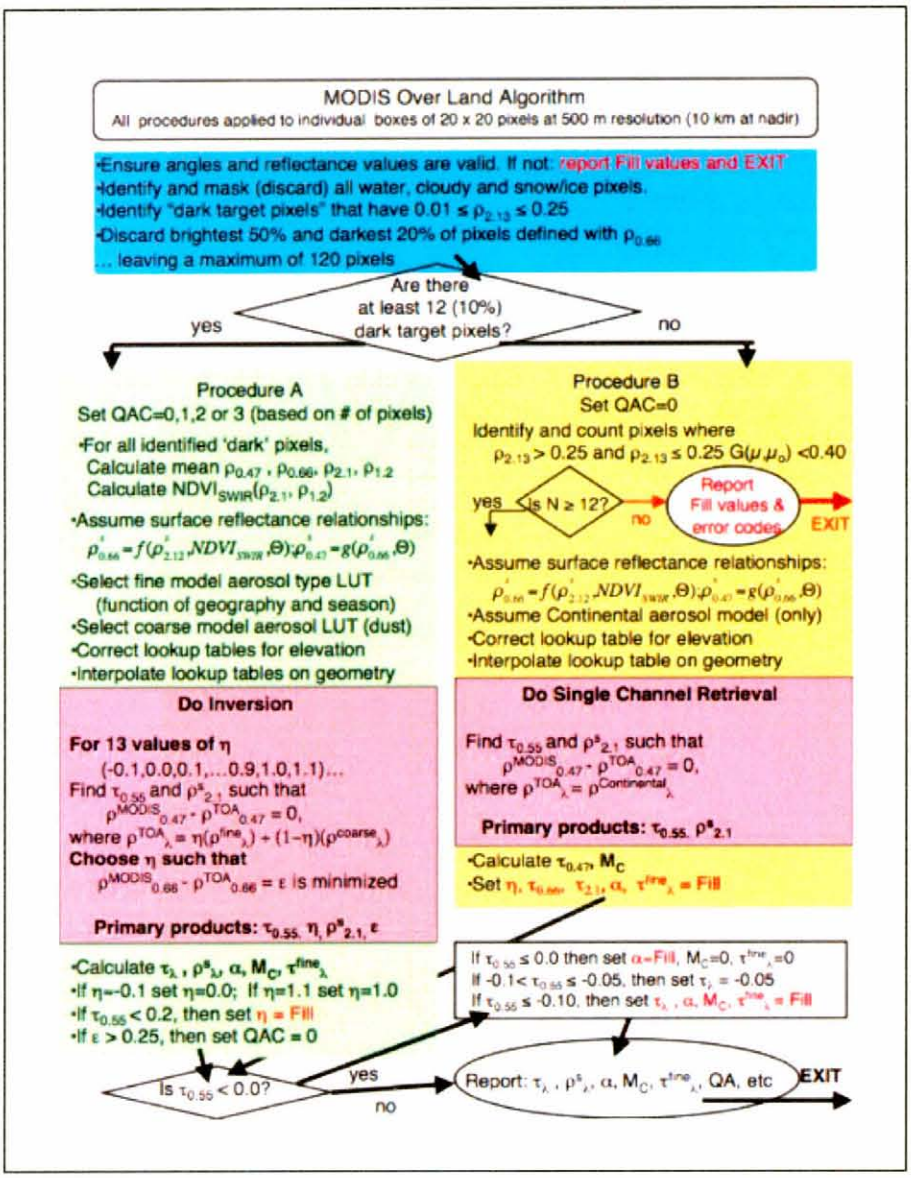
Zou, B., Pu, Q., Bilal, M., Weng, Q., Zhai, L., and Nichol, J. E. (2016). High-Resolution Satellite Mapping of Fine Particulates Based on Geographically Weighted Regression. *Ieee Geoscience and Remote Sensing Letters, 13*(4), 495-499.

APPENDICES

Appendix 1 : Ambient particles' size distribution
(Cao, Chow, Lee and Watson, 2013)

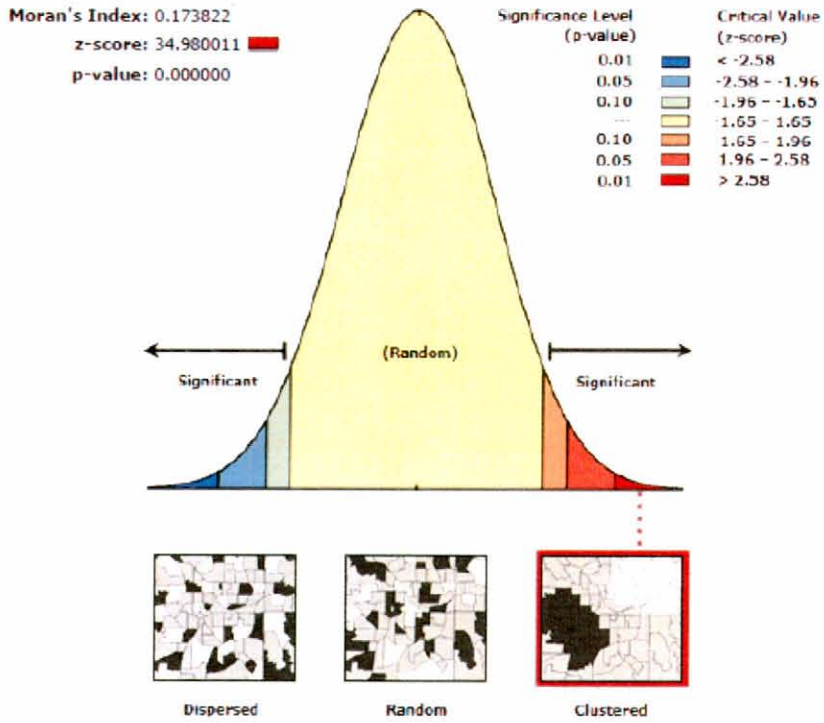


Appendix 2 : Flowchart illustrating the derivation of aerosol over land for V5.2. (Levy et al., 2006)



Appendix 3 : Morans'I Dailly Spatial Autocorrelation Report

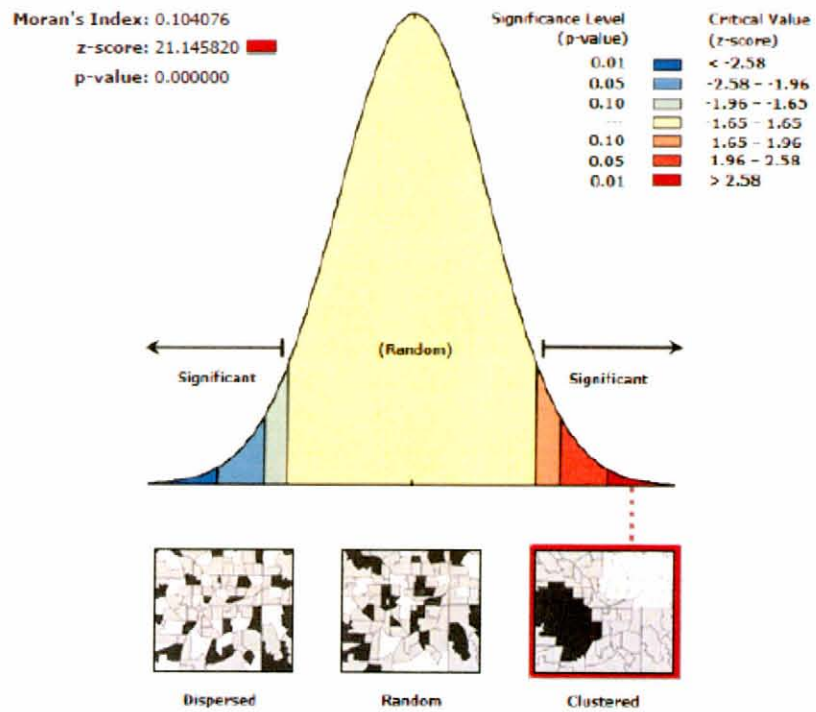
Spatial Autocorrelation Report



Given the z-score of 34.9800112256, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

Appendix 4 : Morans'I Hourly Spatial Autocorrelation Report

Spatial Autocorrelation Report



Given the z-score of 21.1458199594, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

Appendix 5 : Geographically Weighted Regression Report

```
*************************************************************
*          Semiparametric Geographically Weighted Regression
*
*                  Release 1.0.90 (GWR 4.0.90)
*
*                       12 May 2015
*            (Originally coded by T. Nakaya: 1 Nov 2009)                *
*          Tomoki Nakaya(1), Martin Charlton(2), Chris Brunsdon (2)
*
*          Paul Lewis (2), Jing Yao (3), A Stewart Fotheringham (4)
*
*                (c) GWR4 development team
*
* (1) Ritsumeikan University, (2) National University of Ireland, Maynooth,  *
*       (3) University of Glasgow, (4) Arizona State University               *
*************************************************************
******

Program began at 5/20/2019 2:31:02 PM

*************************************************************
*********
Session: All Station
Session control file: C:\Users\User\Desktop\Geographically Weighted Regression
Mode\GWR\All Station\All Daily 2 Parameters.ctl
*************************************************************
*********
Data filename: C:\Users\User\Desktop\Geographically Weighted Regression
Mode\GWR\Stations data 5x5 pixel.csv
Number of areas/points: 685

Model settings--------------------------------
Model type: Gaussian
Geographic kernel: adaptive Gaussian
Method for optimal bandwidth search: Golden section search
Criterion for optimal bandwidth: AIC
Number of varying coefficients: 2
Number of fixed coefficients:  0

Modelling options-----------------------------------
Standardisation of independent variables: On
Testing geographical variability of local coefficients: On
Local to Global Variable selection: OFF
Global to Local Variable selection: OFF
Prediction at non-regression points: OFF

Variable settings-----------------------------------
```

Area key: field1: Site_Id
Easting (x-coord): field3 : Logitude
Northing (y-coord): field2: Latitude
Cartesian coordinates: Euclidean distance
Dependent variable: field6: Daily_PM2.5
Offset variable is not specified
Intercept: varying (Local) intercept
Independent variable with varying (Local) coefficient: field15:
MODIS_AOT_(5x5_Pixel)

****

****

Global regression result

****

< Diagnostic information >

Residual sum of squares:	106116.013526
Number of parameters:	2

(Note: this num does not include an error variance term for a Gaussian model)

ML based global sigma estimate:	12.446441
Unbiased global sigma estimate:	12.464651
-2 log-likelihood:	5398.311331
Classic AIC:	5404.311331
AICc:	5404.346573
BIC/MDL:	5417.899588
CV:	156.907493
R square:	0.484591
Adjusted R square:	0.483079

Variable	Estimate	Standard Error	t(Est/SE)
Intercept	28.917283	0.476250	60.718704
MODIS_AOT_(5x5_Pixel)	12.065308	0.476120	25.340896

****

GWR (Geographically weighted regression) bandwidth selection

****

Bandwidth search <golden section search>
Limits: 44, 685
Golden section search begins...
Initial values

pL	Bandwidth:	144.265	Criterion:	5304.626
p1	Bandwidth:	155.775	Criterion:	5304.626
p2	Bandwidth:	162.888	Criterion:	5304.984
pU	Bandwidth:	174.399	Criterion:	5319.476

```
iter   1 (p1) Bandwidth:   155.775 Criterion:   5304.626 Diff:     7.114
Best bandwidth size 155.000
Minimum AIC     5304.626

***********************************************************************
*********
GWR (Geographically weighted regression) result
***********************************************************************
*********
Bandwidth and geographic ranges
Bandwidth size:             155.774739
Coordinate          Min           Max         Range
--------------- --------------- --------------- ---------------
X-coord          100.297733     102.175900       1.878167
Y-coord            2.265400       5.811183       3.545783
Diagnostic information
Residual sum of squares:       90617.023499
Effective number of parameters (model: trace(S)):                 6.235148
Effective number of parameters (variance: trace(S'S)):            4.695150
Degree of freedom (model: n - trace(S)):                        678.764852
Degree of freedom (residual: n - 2trace(S) + trace(S'S)):       677.224854
ML based sigma estimate:          11.501636
Unbiased sigma estimate:          11.567472
-2 log-likelihood:              5290.156085
Classic AIC:                    5304.626381
AICc:                           5304.802461
BIC/MDL:                        5337.397396
CV:                              137.762710
R square:                          0.559870
Adjusted R square:                 0.554809

***********************************************************
<< Geographically varying (Local) coefficients >>
***********************************************************
Estimates of varying coefficients have been saved in the following file.
Listwise output file: C:\Users\User\Desktop\Geographically Weighted Regression
Mode\GWR\All Station\All Daily 2 Parameters_listwise.csv

Summary statistics for varying (Local) coefficients
Variable                Mean          STD
------------------- --------------- ---------------
Intercept             28.856546       3.586368
MODIS_AOT_(5x5_Pixel)     11.585797       0.478084

Variable                Min           Max         Range
------------------- --------------- --------------- ---------------
Intercept             24.302907       34.518240       10.215333
MODIS_AOT_(5x5_Pixel)     10.831048       12.262744       1.431696
```

Variable	Lwr Quartile	Median	Upr Quartile
Intercept	26.214887	27.165567	31.246405
MODIS_AOT_(5x5_Pixel)	11.276570	11.538636	12.005011

Variable	Interquartile R	Robust STD
Intercept	5.031518	3.729813
MODIS_AOT_(5x5_Pixel)	0.728441	0.539986

(Note: Robust STD is given by (interquartile range / 1.349))

****

GWR ANOVA Table

****

Source	SS	DF	MS	F
Global Residuals	106116.014	683.000		
GWR Improvement	15498.990	5.775	2683.740	
GWR Residuals	90617.023	677.225	133.806	20.056887

****

Geographical variability tests of local coefficients

****

Variable	F	DOF for F test		DIFF of Criterion
Intercept	49.762230	2.270	678.765	-100.921798
MODIS_AOT_(5x5_Pixel)	1.716251	1.984	678.765	0.540133

Note: positive value of diff-Criterion (AICc, AIC, BIC/MDL or CV) suggests no spatial variability in terms of model selection criteria.
F test: in case of no spatial variability, the F statistics follows the F distribution of DOF for F test.

****

Program terminated at 5/20/2019 2:31:20 PM

```
*****************************************************************
******
*          Semiparametric Geographically Weighted Regression          *
*                  Release 1.0.90 (GWR 4.0.90)
*
*                    12 May 2015
*
*              (Originally coded by T. Nakaya: 1 Nov 2009)
*            Tomoki Nakaya(1), Martin Charlton(2), Chris Brunsdon (2)
*      *           Paul Lewis (2), Jing Yao (3), A Stewart Fotheringham (4)
*
*                  (c) GWR4 development team
*
* (1) Ritsumeikan University, (2) National University of Ireland, Maynooth,
*
*       (3) University of Glasgow, (4) Arizona State University
*
*****************************************************************
******

Program began at 5/20/2019 2:37:09 PM

*****************************************************************
********
Session: All Station
Session control file: C:\Users\User\Desktop\Geographically Weighted Regression
Mode\GWR\All Station\All Daily 6 Parameters.ctl
*****************************************************************
********
Data filename: C:\Users\User\Desktop\Geographically Weighted Regression
Mode\GWR\Stations data 5x5 pixel.csv
Number of areas/points: 685

Model settings--------------------------------
Model type: Gaussian
Geographic kernel: adaptive Gaussian
Method for optimal bandwidth search: Golden section search
Criterion for optimal bandwidth: AIC
Number of varying coefficients: 6
Number of fixed coefficients:  0

Modelling options-----------------------------------
Standardisation of independent variables: On
Testing geographical variability of local coefficients: On
Local to Global Variable selection: OFF
Global to Local Variable selection: OFF
Prediction at non-regression points: OFF

Variable settings-----------------------------------
```

Area key: field1: Site_Id
Easting (x-coord): field3 : Logitude
Northing (y-coord): field2: Latitude
Cartesian coordinates: Euclidean distance
Dependent variable: field6: Daily_PM2.5
Offset variable is not specified
Intercept: varying (Local) intercept
Independent variable with varying (Local) coefficient: field11: Daily_WS
Independent variable with varying (Local) coefficient: field12: Daily_WD
Independent variable with varying (Local) coefficient: field13: Daily_T
Independent variable with varying (Local) coefficient: field14: Daily_RH
Independent variable with varying (Local) coefficient: field15:
MODIS_AOT_(5x5_Pixel)
****
****

****
****
Global regression result
****
****
< Diagnostic information >
Residual sum of squares: 97252.351113
Number of parameters: 6
(Note: this num does not include an error variance term for a Gaussian model)
ML based global sigma estimate: 11.915294
Unbiased global sigma estimate: 11.967823
-2 log-likelihood: 5338.563025
Classic AIC: 5352.563025
AICc: 5352.728461
BIC/MDL: 5384.268957
CV: 146.111554
R square: 0.527642
Adjusted R square: 0.523462

Variable	Estimate	Standard Error	t(Est/SE)
Intercept	28.917463	0.457267	63.239718
Daily_WS	-2.242307	0.578554	-3.875710
Daily_WD	2.583835	0.514125	5.025698
Daily_T	-0.908688	0.487174	-1.865223
Daily_RH	-2.536161	0.555976	-4.561635
MODIS_AOT_(5x5_Pixel)	11.554462	0.466753	24.754959

****
****
GWR (Geographically weighted regression) bandwidth selection
****
****

```
Bandwidth search <golden section search>
 Limits: 52, 685
Golden section search begins...
Initial values
 pL       Bandwidth:  135.150 Criterion:  5268.240
 p1       Bandwidth:  146.854 Criterion:  5268.240
 p2       Bandwidth:  154.088 Criterion:  5268.240
 pU       Bandwidth:  165.792 Criterion:  5268.628
Best bandwidth size 0.000
Minimum AIC     5238.801

*****************************************************************
********
 GWR (Geographically weighted regression) result
*****************************************************************
********
 Bandwidth and geographic ranges
Bandwidth size:              0.000000
Coordinate          Min          Max          Range
--------------- ------------ ------------ ------------
X-coord          100.297733   102.175900     1.878167
Y-coord            2.265400     5.811183     3.545783

 Diagnostic information
Residual sum of squares:       74152.478881
Effective number of parameters (model: trace(S)):              42.000000
Effective number of parameters (variance: trace(S'S)):         42.000000
 (Warning: trace(S) is smaller than trace(S'S). It means the variance of the
predictions is inadequately inflated.)
 (Note: n - trace(S) is used for computing the error variance as the degree of
freedom.)
Degree of freedom (model: n - trace(S)):                      643.000000
Degree of freedom (residual: n - trace(S)):                   643.000000
ML based sigma estimate:         10.404412
Unbiased sigma estimate:         10.738839
-2 log-likelihood:             5152.800848
Classic AIC:                   5238.800848
AICc:                        5244.704124
BIC/MDL:                       5433.565858
CV:                          138.554511
R square:                        0.639839
Adjusted R square:               0.616277

*********************************************************
 << Geographically varying (Local) coefficients >>
*********************************************************
Estimates of varying coefficients have been saved in the following file.
   Listwise output file: C:\Users\User\Desktop\Geographically Weighted Regression
Mode\GWR\All Station\All Daily 6 Parameters_listwise.csv
```

Summary statistics for varying (Local) coefficients

Variable	Mean	STD
Intercept	28.343203	5.039667
Daily_WS	-1.475687	2.083447
Daily_WD	1.962618	1.996984
Daily_T	0.328762	2.267491
Daily_RH	-1.321866	1.471140
MODIS_AOT_(5x5_Pixel)	11.242135	1.825591

Variable	Min	Max	Range
Intercept	21.315254	35.217797	13.902543
Daily_WS	-4.888977	1.913928	6.802905
Daily_WD	-1.078159	5.483853	6.562012
Daily_T	-3.507472	3.797215	7.304686
Daily_RH	-4.120065	2.142922	6.262987
MODIS_AOT_(5x5_Pixel)	9.399044	15.730563	6.331518

Variable	Lwr Quartile	Median	Upr Quartile
Intercept	23.546143	29.448338	31.544202
Daily_WS	-2.848110	-1.627484	-0.523024
Daily_WD	0.269870	2.076994	2.274606
Daily_T	-1.031583	-0.121538	1.092832
Daily_RH	-2.049331	-1.559651	-0.768798
MODIS_AOT_(5x5_Pixel)	10.108894	10.628079	11.541896

Variable	Interquartile R	Robust STD
Intercept	7.998059	5.928880
Daily_WS	2.325086	1.723563
Daily_WD	2.004737	1.486091
Daily_T	2.124415	1.574807
Daily_RH	1.280534	0.949247
MODIS_AOT_(5x5_Pixel)	1.433002	1.062270

(Note: Robust STD is given by (interquartile range / 1.349))

****

GWR ANOVA Table

****

Source	SS	DF	MS	F
Global Residuals	97252.351	679.000		
GWR Improvement	23099.872	36.000	641.663	
GWR Residuals	74152.479	643.000	115.323	5.564067

```
*************************************************************************
*****
 Geographical variability tests of local coefficients
*************************************************************************
*****
 Variable          F              DOF for F test  DIFF of Criterion
------------------- ----------------- --------------- ----------------
Intercept              8.578128   6.000  643.000      -40.746687
Daily_WS               1.258461   6.000  643.000        4.002893
Daily_WD               2.279199   6.000  643.000       -2.415686
Daily_T                2.209287   6.000  643.000       -1.977979
Daily_RH               0.760676   6.000  643.000        7.155001
MODIS_AOT_(5x5_Pixel)       2.564635   6.000  643.000       -4.199845
------------------- ----------------- --------------- ----------------
Note: positive value of diff-Criterion (AICc, AIC, BIC/MDL or CV) suggests no
spatial variability in terms of model selection criteria.
 F test: in case of no spatial variability, the F statistics follows the F distribution of
DOF for F test.

*************************************************************************
*********
Program terminated at 5/20/2019 2:38:40 PM


*************************************************************************
******
*          Semiparametric Geographically Weighted Regression
        *
*                  Release 1.0.90 (GWR 4.0.90)
*
*                      12 May 2015
*
*            (Originally coded by T. Nakaya: 1 Nov 2009)
*
*
*
*          Tomoki Nakaya(1), Martin Charlton(2), Chris Brunsdon (2)
*
*          Paul Lewis (2), Jing Yao (3), A Stewart Fotheringham (4)
*
*                (c) GWR4 development team
*
* (1) Ritsumeikan University, (2) National University of Ireland, Maynooth,
*
*       (3) University of Glasgow, (4) Arizona State University
*
*************************************************************************
*********
```

Program began at 5/20/2019 2:33:59 PM

****
Session: All Station
Session control file: C:\Users\User\Desktop\Geographically Weighted Regression Mode\GWR\All Station\All Hourlay 6 Parameters.ctl
****
Data filename: C:\Users\User\Desktop\Geographically Weighted Regression Mode\GWR\Stations data 5x5 pixel.csv
Number of areas/points: 685

Model settings----------
Model type: Gaussian
Geographic kernel: adaptive Gaussian
Method for optimal bandwidth search: Golden section search
Criterion for optimal bandwidth: AIC
Number of varying coefficients: 6
Number of fixed coefficients: 0

Modelling options----------
Standardisation of independent variables: On
Testing geographical variability of local coefficients: On
Local to Global Variable selection: OFF
Global to Local Variable selection: OFF
Prediction at non-regression points: OFF

Variable settings----------
Area key: field1: Site_Id
Easting (x-coord): field3 : Logitude
Northing (y-coord): field2: Latitude
Cartesian coordinates: Euclidean distance
Dependent variable: field5: Hourly_PM2.5
Offset variable is not specified
Intercept: varying (Local) intercept
Independent variable with varying (Local) coefficient: field7: Hourly_WS
Independent variable with varying (Local) coefficient: field8: Hourly_WD
Independent variable with varying (Local) coefficient: field9: Hourly_T
Independent variable with varying (Local) coefficient: field10: Hourly_RH
Independent variable with varying (Local) coefficient: field15: MODIS_AOT_(5x5_Pixel)
****
****
Global regression result
****
****

```
 < Diagnostic information >
Residual sum of squares:            35159.779506
Number of parameters:                   6
 (Note: this num does not include an error variance term for a Gaussian model)
ML based global sigma estimate:         7.164366
Unbiased global sigma estimate:         7.195951
-2 log-likelihood:                   4641.639671
Classic AIC:                         4655.639671
AICc:                                4655.805106
BIC/MDL:                             4687.345603
CV:                                  54.193457
R square:                             0.657815
Adjusted R square:                        0.654787

Variable            Estimate    Standard Error     t(Est/SE)
------------------- ----------- -------------- ---------------
Intercept           19.238141       0.274944      69.971099
Hourly_WS             -0.476982       0.278113      -1.715065
Hourly_WD              0.405452       0.276192       1.468008
Hourly_T               0.475486       0.283414       1.677709
Hourly_RH              0.342160       0.285245       1.199530
MODIS_AOT_(5x5_Pixel)      9.806556       0.278228      35.246502

*************************************************************************
*********
 GWR (Geographically weighted regression) bandwidth selection
*************************************************************************
*********

Bandwidth search <golden section search>
 Limits: 52, 685
 Golden section search begins...
 Initial values
  pL        Bandwidth:  135.150 Criterion:  4557.132
  p1        Bandwidth:  146.854 Criterion:  4557.132
  p2        Bandwidth:  154.088 Criterion:  4557.132
  pU        Bandwidth:  165.792 Criterion:  4557.442
Best bandwidth size  0.000
Minimum AIC     4508.783

*************************************************************************
*********
 GWR (Geographically weighted regression) result
*************************************************************************
*********
 Bandwidth and geographic ranges
Bandwidth size:              0.000000
Coordinate          Min           Max         Range
--------------- --------------- --------------- ---------------
```

X-coord	100.297733	102.175900	1.878167
Y-coord	2.265400	5.811183	3.545783

Diagnostic information

Residual sum of squares:	25544.034217
Effective number of parameters (model: trace(S)):	42.000000
Effective number of parameters (variance: trace(S'S)):	42.000000
Degree of freedom (model: n - trace(S)):	643.000000
Degree of freedom (residual: n - 2trace(S) + trace(S'S)):	643.000000
ML based sigma estimate:	6.106600
Unbiased sigma estimate:	6.302883
-2 log-likelihood:	4422.782851
Classic AIC:	4508.782851
AICc:	4514.686127
BIC/MDL:	4703.547861
CV:	50.523331
R square:	0.751398
Adjusted R square:	0.735134

****
<< Geographically varying (Local) coefficients >>
****
Estimates of varying coefficients have been saved in the following file.
Listwise output file: C:\Users\User\Desktop\Geographically Weighted Regression Mode\GWR\All Station\All Hourlay 6 Parameters_listwise.csv

Summary statistics for varying (Local) coefficients

Variable	Mean	STD
Intercept	20.580748	4.172102
Hourly_WS	0.304683	1.430924
Hourly_WD	0.249319	0.885814
Hourly_T	0.419597	2.120635
Hourly_RH	0.456380	1.773911
MODIS_AOT_(5x5_Pixel)	9.844295	2.983254

Variable	Min	Max	Range
Intercept	15.519216	29.268407	13.749191
Hourly_WS	-1.892487	3.189074	5.081561
Hourly_WD	-1.033707	2.186593	3.220300
Hourly_T	-2.970145	5.300924	8.271069
Hourly_RH	-1.559930	4.593108	6.153039
MODIS_AOT_(5x5_Pixel)	5.224508	14.707590	9.483081

Variable	Lwr Quartile	Median	Upr Quartile
Intercept	17.537217	19.399168	23.256973
Hourly_WS	-0.715777	0.203533	1.017438

```
Hourly_WD              -0.954795      0.499931      0.513612
Hourly_T               -0.430882      0.899645      1.073200
Hourly_RH              -0.766302     -0.394674      1.820659
MODIS_AOT_(5x5_Pixel)   7.883977     11.153229     12.824435

Variable          Interquartile R    Robust STD
----------------- ---------------- --------------
Intercept              5.719755       4.239997
Hourly_WS              1.733216       1.284815
Hourly_WD              1.468406       1.088515
Hourly_T               1.504082       1.114961
Hourly_RH              2.586961       1.917688
MODIS_AOT_(5x5_Pixel)  4.940458       3.662312
 (Note: Robust STD is given by (interquartile range / 1.349) )

*************************************************************************
*********
 GWR ANOVA Table
*************************************************************************
*********
Source                 SS        DF          MS          F
---------------- ------------------ ---------- ------------- ----------
Global Residuals         35159.780    679.000
GWR Improvement           9615.745     36.000       267.104
GWR Residuals            25544.034    643.000        39.726   6.723601

*************************************************************************
*****
 Geographical variability tests of local coefficients
*************************************************************************
*****
 Variable          F              DOF for F test  DIFF of Criterion
----------------- ---------------- --------------- ----------------
Intercept              8.386448    6.000  643.000      -39.611350
Hourly_WS              2.634458    6.000  643.000       -4.635580
Hourly_WD              2.044177    6.000  643.000       -0.943144
Hourly_T               3.434752    6.000  643.000       -9.610149
Hourly_RH              1.894759    6.000  643.000       -0.005315
MODIS_AOT_(5x5_Pixel)     19.500652    6.000  643.000     -102.517517
----------------- ---------------- --------------- ----------------
Note: positive value of diff-Criterion (AICc, AIC, BIC/MDL or CV) suggests no
spatial variability in terms of model selection criteria.
 F test: in case of no spatial variability, the F statistics follows the F distribution of
DOF for F test.

*************************************************************************
*********
Program terminated at 5/20/2019 2:35:29 PM
```

```
*********************************************************************
*
*          Semiparametric Geographically Weighted Regression          *
*                  Release 1.0.90 (GWR 4.0.90)                        *
*                       12 May 2015                                   *
*             (Originally coded by T. Nakaya: 1 Nov 2009)             *
*                                                                     *
*          Tomoki Nakaya(1), Martin Charlton(2), Chris Brunsdon (2)   *
*          Paul Lewis (2), Jing Yao (3), A Stewart Fotheringham (4)   *
*                  (c) GWR4 development team                          *
* (1) Ritsumeikan University, (2) National University of Ireland, Maynooth, *
*       (3) University of Glasgow, (4) Arizona State University       *
*********************************************************************
*********

Program began at 5/20/2019 2:28:07 PM

*********************************************************************
*********
Session: All Station
Session control file: C:\Users\User\Desktop\Geographically Weighted Regression
Mode\GWR\All Station\All Hourlay 2 Parameters.ctl
*********************************************************************
*********
Data filename: C:\Users\User\Desktop\Geographically Weighted Regression
Mode\GWR\Stations data 5x5 pixel.csv
Number of areas/points: 685

Model settings--------------------------------
Model type: Gaussian
Geographic kernel: adaptive Gaussian
Method for optimal bandwidth search: Golden section search
Criterion for optimal bandwidth: AIC
Number of varying coefficients: 2
Number of fixed coefficients:  0

Modelling options-----------------------------
Standardisation of independent variables: On
Testing geographical variability of local coefficients: On
Local to Global Variable selection: OFF
Global to Local Variable selection: OFF
Prediction at non-regression points: OFF

Variable settings-----------------------------
Area key: field1: Site_Id
Easting (x-coord): field3 : Logitude
Northing (y-coord): field2: Latitude
Cartesian coordinates: Euclidean distance
Dependent variable: field5: Hourly_PM2.5
```

```
Offset variable is not specified
Intercept: varying (Local) intercept
Independent variable with varying (Local) coefficient: field15:
MODIS_AOT_(5x5_Pixel)
*************************************************************************
*********

*************************************************************************
*********
 Global regression result
*************************************************************************
*********
 < Diagnostic information >
Residual sum of squares:            35633.196596
Number of parameters:                   2
(Note: this num does not include an error variance term for a Gaussian model)
ML based global sigma estimate:         7.212438
Unbiased global sigma estimate:         7.222990
-2 log-likelihood:                   4650.801471
Classic AIC:                         4656.801471
AICc:                               4656.836713
BIC/MDL:                             4670.389727
CV:                                 54.232001
R square:                            0.653207
Adjusted R square:                       0.652190

Variable              Estimate   Standard Error     t(Est/SE)
-------------------- ------------ -------------- ---------------
Intercept             19.237186      0.275976      69.705912
MODIS_AOT_(5x5_Pixel)      9.895878      0.275901      35.867487

*************************************************************************
*********
 GWR (Geographically weighted regression) bandwidth selection
*************************************************************************
*********

Bandwidth search <golden section search>
 Limits: 44,  685
 Golden section search begins...
 Initial values
 pL        Bandwidth:  144.265 Criterion:  4559.547
 p1        Bandwidth:  155.775 Criterion:  4559.547
 p2        Bandwidth:  162.888 Criterion:  4559.755
 pU        Bandwidth:  174.399 Criterion:  4560.246
iter   1 (p1) Bandwidth:  155.775 Criterion:  4559.547 Diff:     7.114
Best bandwidth size 155.000
Minimum AIC     4559.547
```

```
*********************************************************************
*********
 GWR (Geographically weighted regression) result
*********************************************************************
*********
 Bandwidth and geographic ranges
Bandwidth size:          155.774739
Coordinate          Min          Max          Range
---------- ---------- ---------- ----------
X-coord      100.297733   102.175900     1.878167
Y-coord        2.265400     5.811183     3.545783

 Diagnostic information
Residual sum of squares:      30536.864359
Effective number of parameters (model: trace(S)):          6.235148
Effective number of parameters (variance: trace(S'S)):     4.695150
Degree of freedom (model: n - trace(S)):                 678.764852
Degree of freedom (residual: n - 2trace(S) + trace(S'S)):  677.224854
ML based sigma estimate:       6.676778
Unbiased sigma estimate:       6.714996
-2 log-likelihood:          4545.076467
Classic AIC:                4559.546763
AICc:                       4559.722843
BIC/MDL:                    4592.317778
CV:                           50.887843
R square:                      0.702806
Adjusted R square:             0.699389

*********************************************************
 << Geographically varying (Local) coefficients >>
*********************************************************
Estimates of varying coefficients have been saved in the following file.
   Listwise output file: C:\Users\User\Desktop\Geographically Weighted Regression
Mode\GWR\All Station\All Hourlay 2 Parameters_listwise.csv

Summary statistics for varying (Local) coefficients
Variable              Mean          STD
-------------------- ---------- ----------
Intercept           19.225508     1.264990
MODIS_AOT_(5x5_Pixel)     9.634842     1.353813

Variable              Min          Max          Range
-------------------- ---------- ---------- ----------
Intercept           17.247450    20.875613     3.628163
MODIS_AOT_(5x5_Pixel)     7.320746    11.344249     4.023503

Variable          Lwr Quartile      Median    Upr Quartile
-------------------- ---------- ---------- ----------
Intercept           18.406485    18.894830    20.155986
```

MODIS_AOT_(5x5_Pixel) 9.169508 9.476977 10.796979

Variable	Interquartile R	Robust STD
Intercept	1.749501	1.296888
MODIS_AOT_(5x5_Pixel)	1.627471	1.206428

(Note: Robust STD is given by (interquartile range / 1.349))

****

GWR ANOVA Table

****

Source	SS	DF	MS	F
Global Residuals	35633.197	683.000		
GWR Improvement	5096.332	5.775	882.459	
GWR Residuals	30536.864	677.225	45.091	19.570558

****

Geographical variability tests of local coefficients

****

Variable	F	DOF for F test		DIFF of Criterion
Intercept	22.993093	2.270	678.765	-46.213059
MODIS_AOT_(5x5_Pixel)	33.754838	1.984	678.765	-60.472678

Note: positive value of diff-Criterion (AICc, AIC, BIC/MDL or CV) suggests no spatial variability in terms of model selection criteria.

F test: in case of no spatial variability, the F statistics follows the F distribution of DOF for F test.

****

Program terminated at 5/20/2019 2:28:24 PM

Appendix 6 : Descriptive statistics (Quantitative data) Kuching AERONET station from 2012 to 2015

Statistic	AERONET 550 nm	AERONET (550 nm) ± 30 minutes	MODIS 550 nm
Nbr. of observations	137	137	137
Minimum	0.041	0.057	0.030
Maximum	2.734	2.136	4.841
Freq. of minimum	1	1	1
Freq. of maximum	1	1	1
1st Quartile	0.120	0.119	0.160
Median	0.188	0.165	0.247
3rd Quartile	0.347	0.276	0.402
Mean	0.344	0.291	0.484
Variance (n)	0.204	0.131	0.609
Standard deviation (n)	0.452	0.362	0.781
Standard deviation (n-1)	0.453	0.363	0.783
Variation coefficient	1.314	1.244	1.613
Skewness (Pearson)	3.302	3.335	3.890
Standard error of the mean	0.039	0.031	0.067
Standard error of the variance	0.025	0.016	0.074
Standard error(Skewness (Fisher))	0.207	0.207	0.207
Mean absolute deviation	0.266	0.209	0.409
Harmonic mean	0.160	0.159	0.218

Appendix 7 : Descriptive statistics (Quantitative data) Pontiank AERONET station from 2012 to 2015

Statistic	AERONET 550 nm	AERONET (550 nm) ± 30 minutes	MODIS 550 nm
Nbr. of observations	192	192	192
Minimum	0.032	0.013	0.073
Maximum	4.785	4.285	4.823
Freq. of minimum	1	1	1
Freq. of maximum	1	1	1
1st Quartile	0.125	0.108	0.171
Median	0.212	0.193	0.271
3rd Quartile	0.494	0.382	0.505
Mean	0.503	0.402	0.564
Variance (n)	0.584	0.391	0.639
Standard deviation (n)	0.764	0.625	0.799
Standard deviation (n-1)	0.766	0.627	0.801
Variation coefficient	1.519	1.554	1.417
Skewness (Pearson)	3.182	3.501	3.109
Standard error of the mean	0.055	0.045	0.058
Standard error of the variance	0.060	0.040	0.066
Standard error(Skewness (Fisher))	0.175	0.175	0.175
Mean absolute deviation	0.477	0.369	0.488
Harmonic mean	0.165	0.121	0.241

Appendix 8 : Descriptive statistics (Quantitative data) Singapore AERONET station from 2012 to 2015

Statistic	AERONET 550 nm	AERONET (550 nm) ± 30 minutes	MODIS 550 nm
Nbr. of observations	160	160	160
Minimum	0.104	0.057	0.123
Maximum	1.876	1.105	1.638
Freq. of minimum	1	1	1
Freq. of maximum	1	1	1
1st Quartile	0.215	0.140	0.262
Median	0.303	0.227	0.349
3rd Quartile	0.497	0.380	0.495
Mean	0.394	0.294	0.454
Variance (n)	0.072	0.043	0.092
Standard deviation (n)	0.268	0.207	0.303
Standard deviation (n-1)	0.269	0.208	0.304
Variation coefficient	0.681	0.705	0.668
Skewness (Pearson)	2.011	1.577	1.866
Standard error of the mean	0.021	0.016	0.024
Standard error of the variance	0.008	0.005	0.010
Standard error(Skewness (Fisher))	0.192	0.192	0.192
Mean absolute deviation	0.198	0.156	0.218
Harmonic mean	0.283	0.196	0.336

Appendix 9 : Descriptive statistics (Quantitative data) Songkhla_met_Sta AERONET station from 2012 to 2015

Statistic	AERONET 550 nm	AERONET (550 nm) ± 30 minutes	MODIS 550 nm
Nbr. of observations	187	187	187
Minimum	0.048	0.035	0.069
Maximum	5.025	5.562	4.436
Freq. of minimum	1	1	1
Freq. of maximum	1	1	1
1st Quartile	0.120	0.100	0.183
Median	0.199	0.167	0.268
3rd Quartile	0.324	0.265	0.371
Mean	0.291	0.279	0.357
Variance (n)	0.182	0.345	0.190
Standard deviation (n)	0.427	0.587	0.436
Standard deviation (n-1)	0.428	0.589	0.437
Variation coefficient	1.468	2.100	1.220
Skewness (Pearson)	8.025	7.767	6.198
Standard error of the mean	0.031	0.043	0.032
Standard error of the variance	0.019	0.036	0.020
Standard error(Skewness (Fisher))	0.178	0.178	0.178
Mean absolute deviation	0.184	0.207	0.201
Harmonic mean	0.168	0.134	0.231

Appendix 10 : Descriptive statistics (Quantitative data) USM Penang AERONET station from 2012 to 2015

Statistic	AERONET 550 nm	AERONET (550 nm) ± 30 minutes	MODIS 550 nm
Nbr. of observations	303	303	303
Minimum	0.064	0.035	0.059
Maximum	3.506	3.706	4.055
Freq. of minimum	1	1	1
Freq. of maximum	1	1	1
1st Quartile	0.180	0.125	0.204
Median	0.288	0.206	0.307
3rd Quartile	0.467	0.349	0.454
Mean	0.389	0.305	0.421
Variance (n)	0.140	0.134	0.166
Standard deviation (n)	0.374	0.366	0.407
Standard deviation (n-1)	0.375	0.367	0.408
Variation coefficient	0.961	1.201	0.966
Skewness (Pearson)	3.937	4.994	3.945
Standard error of the mean	0.022	0.021	0.023
Standard error of the variance	0.011	0.011	0.014
Standard error(Skewness (Fisher))	0.140	0.140	0.140
Mean absolute deviation	0.226	0.201	0.243
Harmonic mean	0.244	0.163	0.268

Appendix 11 : Descriptive statistics (Quantitative data) All AERONET station from 2012 to 2015

Statistic	AERONET 550 nm	AERONET (550 nm) ± 30 minutes	MODIS 550 nm
Nbr. of observations	979	979	979
Minimum	0.032	0.013	0.030
Maximum	5.025	5.562	4.841
Freq. of minimum	1	1	1
Freq. of maximum	1	1	1
1st Quartile	0.156	0.118	0.194
Median	0.242	0.196	0.295
3rd Quartile	0.413	0.323	0.456
Mean	0.387	0.315	0.451
Variance (n)	0.238	0.211	0.318
Standard deviation (n)	0.487	0.460	0.564
Standard deviation (n-1)	0.488	0.460	0.564
Variation coefficient	1.259	1.457	1.249
Skewness (Pearson)	4.635	5.954	4.440
Standard error of the mean	0.016	0.015	0.018
Standard error of the variance	0.011	0.010	0.014
Standard error(Skewness (Fisher))	0.078	0.078	0.078
Mean absolute deviation	0.271	0.230	0.303
Harmonic mean	0.198	0.150	0.255

Appendix 12 : Daily descriptive statistics (Quantitative data) in Banting station during 2013

Statistic	Daily $PM_{2.5}$ (μg/m3)	MODIS AOT(550 nm)	Daily WS m/s	Daily WD °	Daily TC°	Daily RH%
Nbr. of observations	96	96	96	96	96	96
Minimum	17.429	0.124	2.301	88.417	22.375	50.100
Maximum	115.000	1.620	8.054	299.292	28.633	98.611
Freq. of minimum	1	1	1	1	1	4
Freq. of maximum	1	1	1	1	1	1
1st Quartile	27.886	0.307	3.800	130.958	25.196	74.510
Median	34.830	0.381	4.310	148.462	26.596	82.979
3rd Quartile	45.792	0.520	4.932	173.479	27.415	87.417
Mean	39.481	0.451	4.438	158.385	26.399	80.869
Variance (n)	351.427	0.058	0.800	1949.780	1.765	97.303
Standard deviation (n)	18.746	0.241	0.895	44.156	1.329	9.864
Standard deviation (n-1)	18.845	0.243	0.899	44.388	1.336	9.916
Variation coefficient	0.475	0.535	0.202	0.279	0.050	0.122
Skewness (Pearson)	1.887	1.846	0.831	1.044	-0.358	-1.183
Standard error of the mean	1.923	0.025	0.092	4.530	0.136	1.012
Standard error of the variance	51.527	0.009	0.117	285.882	0.259	14.267
Standard error(Skewness (Fisher))	0.246	0.246	0.246	0.246	0.246	0.246
Mean absolute deviation	13.000	0.175	0.699	33.271	1.121	7.637
Harmonic mean	33.481	0.358	4.268	147.736	26.330	79.392

Appendix 13 : Hourly descriptive statistics (Quantitative data) in Banting station during 2013

Statistic	Hourly $PM_{2.5}$ (µg/m3)	MODIS AOT(550 nm)	Hourly WS m/s	Hourly WD °	Hourly TC°	Hourly RH%
Nbr. of observations	96	96	96	96	96	96
Minimum	5.000	0.124	2.800	1.000	22.375	50.000
Maximum	69.000	1.620	15.700	358.000	32.400	98.611
Freq. of minimum	2	1	1	1	1	2
Freq. of maximum	1	1	1	1	1	1
1st Quartile	14.000	0.307	4.600	58.000	26.800	65.750
Median	19.500	0.381	6.150	122.500	28.850	74.029
3rd Quartile	25.000	0.520	8.250	265.750	30.000	85.854
Mean	21.313	0.451	6.652	155.250	28.422	74.566
Variance (n)	119.715	0.058	6.592	12931.625	4.421	148.653
Standard deviation (n)	10.941	0.241	2.568	113.717	2.103	12.192
Standard deviation (n-1)	10.999	0.243	2.581	114.314	2.114	12.256
Variation coefficient	0.513	0.535	0.386	0.732	0.074	0.164
Skewness (Pearson)	1.496	1.846	1.057	0.541	-0.484	-0.145
Standard error of the mean	1.123	0.025	0.263	11.667	0.216	1.251
Standard error of the variance	17.553	0.009	0.967	1896.068	0.648	21.796
Standard error(Skewness (Fisher))	0.246	0.246	0.246	0.246	0.246	0.246
Mean absolute deviation	7.996	0.175	2.061	97.693	1.721	10.253
Harmonic mean	16.483	0.358	5.829	37.304	28.259	72.456

Appendix 14 : Daily descriptive statistics (Quantitative data) in Bukit Rambai station during 2013

Statistic	Daily $PM_{2.5}$ (µg/m3)	MODIS AOT(550 nm)	Daily WS m/s	Daily WD °	Daily TC°	Daily RH%
Nbr. of observations	60	60	60	60	60	60
Minimum	13.750	0.125	3.267	33.333	26.542	55.333
Maximum	139.333	1.564	7.700	281.583	30.071	75.083
Freq. of minimum	1	1	5	1	2	4
Freq. of maximum	1	1	4	1	1	2
1st Quartile	24.647	0.224	3.575	77.344	27.406	58.000
Median	30.611	0.327	4.656	109.646	28.167	65.396
3rd Quartile	39.128	0.426	5.600	167.188	28.379	69.604
Mean	35.610	0.387	4.900	124.758	27.953	64.390
Variance (n)	439.140	0.059	1.914	4360.350	0.484	37.393
Standard deviation (n)	20.956	0.244	1.384	66.033	0.696	6.115
Standard deviation (n-1)	21.133	0.246	1.395	66.590	0.702	6.167
Variation coefficient	0.588	0.629	0.282	0.529	0.025	0.095
Skewness (Pearson)	3.273	2.336	0.747	0.767	0.335	0.084
Standard error of the mean	2.728	0.032	0.180	8.597	0.091	0.796
Standard error of the variance	82.223	0.011	0.358	816.412	0.091	7.001
Standard error(Skewness (Fisher))	0.309	0.309	0.309	0.309	0.309	0.309
Mean absolute deviation	12.364	0.166	1.133	53.295	0.591	5.562
Harmonic mean	29.664	0.296	4.559	92.532	27.936	63.810

Appendix 15 : Hourly descriptive statistics (Quantitative data) in Bukit Rambai station during 2013

Statistic	Hourly $PM_{2.5}$ (µg/m3)	MODIS AOT(550 nm)	Hourly WS m/s	Hourly WD °	Hourly TC°	Hourly RH%
Nbr. of observations	60	60	60	60	60	60
Minimum	5.000	0.125	2.800	1.000	27.165	45.000
Maximum	101.000	1.564	7.700	354.000	31.800	73.043
Freq. of minimum	1	1	1	2	1	1
Freq. of maximum	1	1	4	1	1	1
1st Quartile	12.000	0.224	3.575	41.500	28.167	54.000
Median	16.000	0.327	4.621	56.500	28.650	56.250
3rd Quartile	24.250	0.426	5.600	233.500	30.500	60.052
Mean	19.900	0.387	4.990	119.383	29.211	57.207
Variance (n)	194.023	0.059	2.135	11391.970	2.101	40.381
Standard deviation (n)	13.929	0.244	1.461	106.733	1.450	6.355
Standard deviation (n-1)	14.047	0.246	1.474	107.634	1.462	6.408
Variation coefficient	0.700	0.629	0.293	0.894	0.050	0.111
Skewness (Pearson)	3.511	2.336	0.590	0.732	0.289	0.468
Standard error of the mean	1.813	0.032	0.190	13.895	0.189	0.827
Standard error of the variance	36.328	0.011	0.400	2132.981	0.393	7.561
Standard error(Skewness (Fisher))	0.309	0.309	0.309	0.309	0.309	0.309
Mean absolute deviation	8.733	0.166	1.221	97.319	1.307	4.902
Harmonic mean	15.021	0.296	4.604	16.980	29.140	56.522

Appendix 16 : Daily descriptive statistics (Quantitative data) in Cheras station during 2013

Statistic	Daily $PM_{2.5}$ (µg/m3)	MODIS AOT(550 nm)	Daily WS m/s	Daily WD °	Daily TC°	Daily RH%
Nbr. of observations	105	105	105	105	105	105
Minimum	12.238	0.117	1.080	88.000	23.840	73.000
Maximum	151.375	3.387	5.267	343.000	30.371	100.000
Freq. of minimum	1	1	1	1	1	1
Freq. of maximum	1	1	1	1	1	1
1st Quartile	22.381	0.287	2.925	191.583	27.350	88.208
Median	28.917	0.415	3.321	212.542	28.288	92.750
3rd Quartile	37.833	0.599	3.833	228.458	29.075	95.250
Mean	33.161	0.509	3.334	212.428	28.167	91.354
Variance (n)	353.052	0.167	0.475	1608.483	1.385	27.177
Standard deviation (n)	18.790	0.408	0.689	40.106	1.177	5.213
Standard deviation (n-1)	18.880	0.410	0.692	40.298	1.182	5.238
Variation coefficient	0.567	0.802	0.207	0.189	0.042	0.057
Skewness (Pearson)	3.059	3.930	-0.078	0.158	-0.475	-0.932
Standard error of the mean	1.842	0.040	0.068	3.933	0.115	0.511
Standard error of the variance	49.430	0.023	0.066	225.201	0.194	3.805
Standard error(Skewness (Fisher))	0.236	0.236	0.236	0.236	0.236	0.236
Mean absolute deviation	12.192	0.247	0.555	28.139	0.965	4.213
Harmonic mean	27.148	0.363	3.163	203.862	28.117	91.037

Appendix 17 : Hourly descriptive statistics (Quantitative data) in Cheras station during 2013

Statistic	Hourly $PM_{2.5}$ (µg/m3)	MODIS AOT(550 nm)	Hourly WS m/s	Hourly WD °	Hourly TC°	Hourly RH%
Nbr. of observations	105	105	105	105	105	105
Minimum	7.000	0.117	0.900	2.000	23.400	59.000
Maximum	92.000	3.387	10.800	357.000	37.100	100.000
Freq. of minimum	1	1	1	1	1	1
Freq. of maximum	1	1	1	1	1	4
1st Quartile	14.000	0.287	4.000	39.000	32.800	75.000
Median	21.000	0.415	4.900	94.000	34.200	83.000
3rd Quartile	28.000	0.599	5.900	209.000	34.900	87.000
Mean	24.476	0.509	5.130	124.860	33.654	81.419
Variance (n)	269.926	0.167	2.784	11104.516	4.799	83.615
Standard deviation (n)	16.429	0.408	1.668	105.378	2.191	9.144
Standard deviation (n-1)	16.508	0.410	1.676	105.883	2.201	9.188
Variation coefficient	0.671	0.802	0.325	0.844	0.065	0.112
Skewness (Pearson)	2.200	3.930	0.637	0.772	-1.991	-0.162
Standard error of the mean	1.611	0.040	0.164	10.333	0.215	0.897
Standard error of the variance	37.792	0.023	0.390	1554.726	0.672	11.707
Standard error(Skewness (Fisher))	0.236	0.236	0.236	0.236	0.236	0.236
Mean absolute deviation	10.989	0.247	1.271	88.936	1.547	7.434
Harmonic mean	18.034	0.363	4.493	33.364	33.487	80.349

Appendix 18 : Daily descriptive statistics (Quantitative data) in Ipoh station during 2013

Statistic	Daily $PM_{2.5}$ (µg/m3)	MODIS AOT(550 nm)	Daily WS m/s	Daily WD °	Daily TC°	Daily RH%
Nbr. of observations	125	125	125	125	125	125
Minimum	7.550	0.073	3.281	34.917	22.450	40.806
Maximum	77.200	1.151	9.450	180.125	31.067	84.304
Freq. of minimum	1	1	1	1	1	1
Freq. of maximum	1	1	1	1	1	1
1st Quartile	14.708	0.212	6.104	90.333	27.646	63.458
Median	19.500	0.317	6.863	108.708	28.563	68.500
3rd Quartile	25.167	0.454	7.521	126.375	29.296	74.417
Mean	21.692	0.375	6.871	107.413	28.476	68.744
Variance (n)	108.208	0.044	0.938	794.671	1.293	53.874
Standard deviation (n)	10.402	0.210	0.968	28.190	1.137	7.340
Standard deviation (n-1)	10.444	0.211	0.972	28.303	1.142	7.369
Variation coefficient	0.480	0.560	0.141	0.262	0.040	0.107
Skewness (Pearson)	2.291	1.270	-0.205	-0.285	-1.064	-0.399
Standard error of the mean	0.934	0.019	0.087	2.532	0.102	0.659
Standard error of the variance	13.853	0.006	0.120	101.737	0.166	6.897
Standard error(Skewness (Fisher))	0.217	0.217	0.217	0.217	0.217	0.217
Mean absolute deviation	7.195	0.162	0.765	22.352	0.898	6.033
Harmonic mean	18.384	0.283	6.720	97.827	28.428	67.896

Appendix 19 : Hourly descriptive statistics (Quantitative data) in Ipoh station during 2013

Statistic	Hourly $PM_{2.5}$ (µg/m3)	MODIS AOT(550 nm)	Hourly WS m/s	Hourly WD °	Hourly TC°	Hourly RH%
Nbr. of observations	125	125	125	125	125	125
Minimum	4.000	0.073	3.900	7.000	26.200	43.000
Maximum	66.000	1.151	15.800	359.000	33.900	86.000
Freq. of minimum	1	1	1	1	1	1
Freq. of maximum	1	1	1	1	1	1
1st Quartile	11.000	0.212	5.700	27.000	29.900	54.000
Median	16.000	0.317	6.600	42.000	31.300	57.000
3rd Quartile	23.000	0.454	8.100	203.000	32.100	64.000
Mean	18.520	0.375	7.229	115.744	30.924	59.272
Variance (n)	127.274	0.044	4.805	12763.934	2.259	93.078
Standard deviation (n)	11.282	0.210	2.192	112.978	1.503	9.648
Standard deviation (n-1)	11.327	0.211	2.201	113.432	1.509	9.687
Variation coefficient	0.609	0.560	0.303	0.976	0.049	0.163
Skewness (Pearson)	1.701	1.270	1.407	0.835	-0.503	0.943
Standard error of the mean	1.013	0.019	0.197	10.146	0.135	0.866
Standard error of the variance	16.294	0.006	0.615	1634.094	0.289	11.916
Standard error(Skewness (Fisher))	0.217	0.217	0.217	0.217	0.217	0.217
Mean absolute deviation	8.289	0.162	1.658	100.661	1.244	7.516
Harmonic mean	13.559	0.283	6.705	36.260	30.849	57.866

Appendix 20 : Daily descriptive statistics (Quantitative data) in Kelantan Tanah Merah station during 2013

Statistic	Daily $PM_{2.5}$ (μg/m3)	MODIS AOT(550 nm)	Daily WS m/s	Daily WD °	Daily TC°	Daily RH%
Nbr. of observations	97	97	97	97	97	97
Minimum	8.000	0.039	3.417	20.000	24.900	68.958
Maximum	83.917	0.987	8.479	260.000	29.604	88.083
Freq. of minimum	1	1	1	1	1	1
Freq. of maximum	1	1	1	1	1	1
1st Quartile	19.100	0.189	5.179	148.875	26.817	74.292
Median	24.286	0.295	5.592	174.417	27.329	76.750
3rd Quartile	30.773	0.455	6.083	201.625	28.325	80.125
Mean	26.172	0.327	5.629	166.306	27.445	77.350
Variance (n)	122.007	0.029	0.641	2384.717	0.991	16.900
Standard deviation (n)	11.046	0.171	0.800	48.834	0.995	4.111
Standard deviation (n-1)	11.103	0.172	0.805	49.087	1.001	4.132
Variation coefficient	0.422	0.523	0.142	0.294	0.036	0.053
Skewness (Pearson)	1.977	0.995	0.134	-1.007	0.041	0.357
Standard error of the mean	1.127	0.017	0.082	4.984	0.102	0.420
Standard error of the variance	17.794	0.004	0.093	347.790	0.145	2.465
Standard error(Skewness (Fisher))	0.245	0.245	0.245	0.245	0.245	0.245
Mean absolute deviation	7.739	0.136	0.608	37.178	0.809	3.413
Harmonic mean	22.404	0.236	5.511	135.579	27.409	77.135

Appendix 21 : Hourly descriptive statistics (Quantitative data) in Kelantan Tanah Merah station during 2013

Statistic	Hourly $PM_{2.5}$ (µg/m3)	MODIS AOT(550 nm)	Hourly WS m/s	Hourly WD °	Hourly TC°
Nbr. of observations	97	97	97	97	97
Minimum	2.000	0.039	3.500	4.000	26.100
Maximum	39.000	0.987	12.300	354.000	36.300
Freq. of minimum	1	1	1	1	1
Freq. of maximum	1	1	2	1	1
1st Quartile	10.000	0.189	5.300	116.000	29.100
Median	13.000	0.295	6.300	184.917	29.900
3rd Quartile	18.000	0.455	7.900	207.000	30.900
Mean	14.505	0.327	6.659	166.992	29.902
Variance (n)	53.508	0.029	3.237	3970.398	2.540
Standard deviation (n)	7.315	0.171	1.799	63.011	1.594
Standard deviation (n-1)	7.353	0.172	1.808	63.338	1.602
Variation coefficient	0.504	0.523	0.270	0.377	0.053
Skewness (Pearson)	1.095	0.995	0.899	-0.403	0.375
Standard error of the mean	0.747	0.017	0.184	6.431	0.163
Standard error of the variance	7.804	0.004	0.472	579.047	0.370
Standard error(Skewness (Fisher))	0.245	0.245	0.245	0.245	0.245
Mean absolute deviation	5.614	0.136	1.422	50.433	1.192
Harmonic mean	10.925	0.236	6.228	98.289	29.818

Appendix 22 : Daily descriptive statistics (Quantitative data) in Putrajaya station during 2013

Statistic	Daily $PM_{2.5}$ (µg/m3)	MODIS AOT(550 nm)	Daily WS m/s	Daily WD °	Daily TC°	Daily RH%
Nbr. of observations	69	69	69	69	69	69
Minimum	11.682	0.151	2.983	55.696	22.375	50.100
Maximum	163.792	1.708	5.712	273.542	26.963	98.611
Freq. of minimum	1	1	1	1	2	2
Freq. of maximum	1	1	1	1	2	1
1st Quartile	20.913	0.329	3.629	111.083	24.950	73.042
Median	25.292	0.477	3.974	126.042	25.200	76.125
3rd Quartile	37.304	0.674	4.267	169.125	25.921	85.833
Mean	31.768	0.545	3.967	140.752	25.268	78.095
Variance (n)	466.005	0.084	0.235	2275.853	1.048	83.116
Standard deviation (n)	21.587	0.291	0.485	47.706	1.024	9.117
Standard deviation (n-1)	21.745	0.293	0.489	48.055	1.031	9.184
Variation coefficient	0.680	0.533	0.122	0.339	0.041	0.117
Skewness (Pearson)	3.847	1.460	0.594	0.797	-0.824	-0.277
Standard error of the mean	2.618	0.035	0.059	5.785	0.124	1.106
Standard error of the variance	81.094	0.015	0.041	396.045	0.182	14.464
Standard error(Skewness (Fisher))	0.289	0.289	0.289	0.289	0.289	0.289
Mean absolute deviation	12.488	0.219	0.379	38.183	0.748	7.222
Harmonic mean	25.426	0.425	3.910	126.288	25.225	76.938

Appendix 23 : Hourly descriptive statistics (Quantitative data) in Putrajaya station during 2013

Statistic	Hourly $PM_{2.5}$ (µg/m3)	MODIS AOT(550 nm)	Hourly WS m/s	Hourly WD °	Hourly TC°	Hourly RH%
Nbr. of observations	69	69	69	69	69	69
Minimum	7.000	0.151	3.300	14.000	22.375	50.100
Maximum	92.000	1.708	8.100	350.000	32.200	98.611
Freq. of minimum	1	1	1	1	1	2
Freq. of maximum	1	1	1	1	1	1
1st Quartile	15.000	0.329	4.300	136.000	25.664	60.000
Median	22.000	0.477	5.000	175.000	28.100	66.750
3rd Quartile	27.000	0.674	5.900	220.000	29.600	85.833
Mean	24.203	0.545	5.211	183.797	27.714	71.016
Variance (n)	214.539	0.084	1.488	6859.814	5.421	177.361
Standard deviation (n)	14.647	0.291	1.220	82.824	2.328	13.318
Standard deviation (n-1)	14.754	0.293	1.229	83.431	2.345	13.415
Variation coefficient	0.605	0.533	0.234	0.451	0.084	0.188
Skewness (Pearson)	2.309	1.460	0.689	0.459	-0.169	0.427
Standard error of the mean	1.776	0.035	0.148	10.044	0.282	1.615
Standard error of the variance	37.334	0.015	0.259	1193.749	0.943	30.865
Standard error(Skewness (Fisher))	0.289	0.289	0.289	0.289	0.289	0.289
Mean absolute deviation	9.511	0.219	0.983	63.318	1.938	11.780
Harmonic mean	18.528	0.425	4.953	129.289	27.514	68.653

Appendix 24 : Daily descriptive statistics (Quantitative data) in USM station during 2013

Statistic	Daily $PM_{2.5}$ (µg/m3)	MODIS AOT(550 nm)	Daily WS m/s	Daily WD °	Daily TC°	Daily RH%
Nbr. of observations	133	133	133	133	133	133
Minimum	7.000	0.081	2.896	39.375	24.013	56.583
Maximum	100.000	2.820	7.838	244.292	30.125	99.292
Freq. of minimum	1	1	1	1	2	1
Freq. of maximum	1	1	1	1	1	2
1st Quartile	15.000	0.232	4.533	88.353	27.004	72.458
Median	18.000	0.297	5.042	121.000	27.867	79.542
3rd Quartile	27.000	0.485	5.763	214.417	28.513	84.542
Mean	22.180	0.393	5.178	144.552	27.772	78.957
Variance (n)	165.937	0.088	0.847	3949.159	1.315	80.995
Standard deviation (n)	12.882	0.297	0.920	62.842	1.147	9.000
Standard deviation (n-1)	12.930	0.298	0.924	63.080	1.151	9.034
Variation coefficient	0.581	0.755	0.178	0.435	0.041	0.114
Skewness (Pearson)	2.756	4.613	0.529	0.212	-0.430	-0.126
Standard error of the mean	1.121	0.026	0.080	5.470	0.100	0.783
Standard error of the variance	20.580	0.011	0.105	489.790	0.163	10.045
Standard error(Skewness (Fisher))	0.210	0.210	0.210	0.210	0.210	0.210
Mean absolute deviation	8.850	0.183	0.727	58.226	0.911	7.144
Harmonic mean	17.839	0.293	5.020	116.991	27.724	77.892

Appendix 25: Hourly descriptive statistics (Quantitative data) in USM station during 2013

Statistic	Hourly $PM_{2.5}$ (µg/m3)	MODIS AOT(550 nm)	Hourly WS m/s	Hourly WD °	Hourly TC°	Hourly RH%
Nbr. of observations	133	133	133	133	133	133
Minimum	6.263	0.081	1.800	8.000	25.200	41.000
Maximum	63.682	2.820	12.700	348.000	32.500	100.000
Freq. of minimum	1	1	1	1	1	1
Freq. of maximum	1	1	1	2	1	7
1st Quartile	11.095	0.232	6.400	105.000	28.500	60.000
Median	13.500	0.297	7.400	169.000	29.500	67.000
3rd Quartile	16.773	0.485	8.900	211.167	30.200	79.000
Mean	14.808	0.393	7.647	160.784	29.234	70.414
Variance (n)	41.376	0.088	4.292	5058.383	1.954	187.611
Standard deviation (n)	6.432	0.297	2.072	71.122	1.398	13.697
Standard deviation (n-1)	6.457	0.298	2.079	71.391	1.403	13.749
Variation coefficient	0.434	0.755	0.271	0.442	0.048	0.195
Skewness (Pearson)	3.916	4.613	0.123	-0.054	-0.691	0.663
Standard error of the mean	0.560	0.026	0.180	6.190	0.122	1.192
Standard error of the variance	5.132	0.011	0.532	627.361	0.242	23.268
Standard error(Skewness (Fisher))	0.210	0.210	0.210	0.210	0.210	0.210
Mean absolute deviation	4.028	0.183	1.620	56.940	1.077	11.138
Harmonic mean	13.245	0.293	6.950	101.409	29.164	67.952

Appendix 26 : Daily descriptive statistics (Quantitative data) Northeast season during 2013

Statistic	Daily $PM_{2.5}$ (µg/m3)	MODIS AOT(550 nm)	Daily WS m/s	Daily WD °	Daily TC°	Daily RH%
Nbr. of observations	276	276	276	276	276	276
Minimum	7.000	0.039	2.125	33.333	22.400	55.333
Maximum	60.111	1.359	8.129	343.000	30.438	99.292
Freq. of minimum	1	1	1	1	1	1
Freq. of maximum	1	1	1	1	1	2
1st Quartile	18.391	0.220	3.697	98.554	26.966	69.781
Median	24.402	0.313	4.935	135.833	27.546	75.500
3rd Quartile	30.784	0.472	5.932	201.014	28.379	84.438
Mean	25.277	0.369	4.919	146.332	27.550	77.307
Variance (n)	84.705	0.040	2.016	3747.137	1.413	107.367
Standard deviation (n)	9.204	0.200	1.420	61.214	1.189	10.362
Standard deviation (n-1)	9.220	0.200	1.422	61.325	1.191	10.381
Variation coefficient	0.364	0.541	0.289	0.418	0.043	0.134
Skewness (Pearson)	0.655	1.411	0.158	0.455	-0.687	0.219
Standard error of the mean	0.555	0.012	0.086	3.691	0.072	0.625
Standard error of the variance	7.250	0.003	0.173	320.719	0.121	9.190
Standard error(Skewness (Fisher))	0.147	0.147	0.147	0.147	0.147	0.147
Mean absolute deviation	7.280	0.154	1.187	51.508	0.898	8.498
Harmonic mean	21.930	0.283	4.485	119.057	27.496	75.926

Appendix 27 : Hourly descriptive statistics (Quantitative data) Northeast season during 2013

Statistic	Hourly $PM_{2.5}$ (µg/m3)	MODIS AOT(550 nm)	Hourly WS m/s	Hourly WD °	Hourly TC°	Hourly RH%
Nbr. of observations	276	276	276	276	276	276
Minimum	3.000	0.039	1.800	1.000	22.400	37.000
Maximum	63.000	1.359	13.200	359.000	36.300	99.000
Freq. of minimum	1	1	1	2	1	1
Freq. of maximum	1	1	1	1	1	1
1st Quartile	11.939	0.220	4.500	44.750	28.900	56.000
Median	15.000	0.313	5.600	96.500	30.000	63.000
3rd Quartile	20.000	0.472	7.200	198.750	31.600	73.000
Mean	16.830	0.369	5.997	125.565	30.335	65.761
Variance (n)	64.534	0.040	3.978	9116.728	5.717	139.620
Standard deviation (n)	8.033	0.200	1.994	95.482	2.391	11.816
Standard deviation (n-1)	8.048	0.200	1.998	95.655	2.395	11.838
Variation coefficient	0.477	0.541	0.333	0.760	0.079	0.180
Skewness (Pearson)	1.735	1.411	0.965	0.732	0.316	0.547
Standard error of the mean	0.484	0.012	0.120	5.758	0.144	0.713
Standard error of the variance	5.524	0.003	0.340	780.304	0.489	11.950
Standard error(Skewness (Fisher))	0.147	0.147	0.147	0.147	0.147	0.147
Mean absolute deviation	5.952	0.154	1.571	81.884	1.818	9.744
Harmonic mean	13.753	0.283	5.405	34.363	30.149	63.753

Appendix 28 : Daily descriptive statistics (Quantitative data) Southwest season during 2013

Statistic	Daily $PM_{2.5}$ ($\mu g/m^3$)	MODIS AOT (550 nm)	Daily WS m/s	Daily WD °	Daily TC°	Daily RH%
Nbr. of observations	222	222	222	222	222	222
Minimum	7.550	0.067	1.080	38.750	22.375	50.100
Maximum	163.792	3.387	9.450	299.292	30.208	100.000
Freq. of minimum	1	1	1	1	1	2
Freq. of maximum	1	1	1	1	1	1
1st Quartile	19.705	0.281	3.960	114.406	26.299	68.583
Median	31.392	0.423	4.921	156.833	27.533	76.604
3rd Quartile	45.887	0.622	6.300	206.833	28.578	85.240
Mean	37.676	0.522	5.163	159.736	27.404	76.132
Variance (n)	636.335	0.160	2.335	3139.668	2.568	128.243
Standard deviation (n)	25.226	0.400	1.528	56.033	1.602	11.324
Standard deviation (n-1)	25.283	0.401	1.532	56.159	1.606	11.350
Variation coefficient	0.670	0.767	0.296	0.351	0.058	0.149
Skewness (Pearson)	2.005	3.161	0.446	0.139	-0.490	-0.137
Standard error of the mean	1.697	0.027	0.103	3.769	0.108	0.762
Standard error of the variance	60.809	0.015	0.223	300.029	0.245	12.255
Standard error(Skewness (Fisher))	0.163	0.163	0.163	0.163	0.163	0.163
Mean absolute deviation	17.728	0.261	1.265	47.815	1.308	9.266
Harmonic mean	26.434	0.344	4.686	137.252	27.307	74.359

Appendix 29 : Hourly descriptive statistics (Quantitative data) Southwest season during 2013

Statistic	Hourly PM2.5 (µg/m3)	MODIS AOT(550 nm)	Hourly WS m/s	Hourly WD °	Hourly TC°	Hourly RH%
Nbr. of observations	222	222	222	222	222	222
Minimum	2.000	0.067	0.900	4.000	23.400	44.000
Maximum	101.000	3.387	15.800	358.000	36.300	100.000
Freq. of minimum	1	1	1	1	1	1
Freq. of maximum	1	1	1	1	1	6
1st Quartile	13.000	0.281	5.100	93.000	28.384	61.000
Median	19.000	0.423	6.500	171.000	29.550	67.000
3rd Quartile	29.000	0.622	8.075	215.750	31.000	77.000
Mean	24.328	0.522	6.783	162.566	29.769	70.025
Variance (n)	301.331	0.160	5.785	8943.341	4.365	155.022
Standard deviation (n)	17.359	0.400	2.405	94.569	2.089	12.451
Standard deviation (n-1)	17.398	0.401	2.411	94.783	2.094	12.479
Variation coefficient	0.714	0.767	0.355	0.582	0.070	0.178
Skewness (Pearson)	1.900	3.161	0.893	0.143	0.421	0.588
Standard error of the mean	1.168	0.027	0.162	6.361	0.141	0.838
Standard error of the variance	28.795	0.015	0.553	854.632	0.417	14.814
Standard error(Skewness (Fisher))	0.163	0.163	0.163	0.163	0.163	0.163
Mean absolute deviation	12.313	0.261	1.831	76.028	1.620	10.235
Harmonic mean	15.954	0.344	5.867	58.146	29.625	67.944

BIODATA OF STUDENT

Khaled Ali Ahmed Ben Youssef was born on 5th June 1973 in Libya. He received his Bachelor degree in Environmental science from Faculty of Engineering and Technology, Sebha University, Sebha (Libya) in 1997. He obtained his MSc in Environmental science from Faculty of Engineering and Technology, Sebha University in 2008. Apart from his past working experience in Libya, He worked as a University Teaching Assistant in Environmental Science department from 1999 to 2008. He worked as assistance lecture in Environmental Science department since 2008 Faculty of Engineering and Technology, Sebha University. Currently, he is a PhD candidate in the field of Environmental Pollution Control Technology in, Faculty of Environmental Studies, Universiti Putra Malaysia (joined in first semester 2013/2014)

LIST OF PUBLICATIONS

Khaled Ali Ahmed Ben Youssef, A. M. A., Helmi Zulhaidi Mohd Shafri, Zulfa Hanan Ashaari, Danladi Yusuf Gumel, Sara Yasina Yusof (2016). Estimation of aerosols dispersion & urban air quality evaluation over Malaysia using MODIS satellite. International Journal of Advanced Scientific and Technical Research, 3(6).

Khaled Ali Ahmed Ben Youssef, Ahmad Makmom Abdullah, Helmi Zulhaidi Mohd Shafri, Zulfa Hanan Ashaari, Danladi Yusuf Gumel, Fadel M. M. Ali Binyehmed (2016). Aerosol optical depth investigated with satellite remote sensing observations over Peninsular Malaysia. 1st International Research Conference on Engineering, Science and Humanities 2016 (IRCESH 2016). W-206.

Ben Youssef, Khaled Ali Ahmed, Ahmad Makmom Abdullah, Shafri, Helmi Z. M. Shafri, Zulfa Hanan Ash'aari (2018). Valuation of aerosol optical thickness over malaysia based on multi-source ground and satellite data. Pertanika Journal of Science & Technology (JST). (Accepted)

Ben Youssef, Khaled Ali Ahmed, Abdullah Ahmad Makmom, Shafri, Helmi Zulhaidi Mohd, Ashaari, Zulfa Hanan (2018). The development of air quality indices through image-retrieved AOD during haze events in Peninsular Malaysia. EnvironmentalAsia. (Submitted)

Ben Youssef, Khaled Ali Ahmed, Ahmad Makmom Abdullah, Shafri, Helmi Z. M. Shafri, Zulfa Hanan Ash'aari (2019). The Development of Air Quality Indices through Image-Retrieved AOD during Haze Events in Peninsular Malaysia. 2019 Theory & Technique International Aerosol Conference and Malaysia Air Quality Annual Symposium (2019 T&T IAC and MAQAS). (Accepted)

Khaled Ali Ahmed Ben Youssef, Ahmad Makmom Abdullah, Helmi Zulhaidi Mohd Shafri (2019). Satellite remote sensing of fine particulate matter (PM2.5) air quality over Peninsular Malaysia using MODIS. Environmental Pollution (Under revision)

Co-authored publicatons and have not been included as part of this Thesis

Mande, K. H., Abdullah, Ahmad Makmom, Nuruddin, Ahmad Ainuddin, Bose, Mahmud Mohammad, Binyehmed, Fadel Mohamed, Ben Youssef, Khaleed Ali Ahmed, Gabriel Emmanuel (2014a). Forest recovering and soil respiration rate. IOSR Journal of Environmental Science, Toxicology and Food Technology, 8(12), 44-49.

Mande, K. H. A., Ahmad Nuruddin, Abdullah Ahmad Makmom, Aris, Ahmad Zaharin, Binyemed, Fadel Mohamed, Ben Youssef, Khaleed Ali Ahmed (2014b). Forest logging and its impact on soil carbon dioxide efflux in the

tropical forest, Peninsular Malaysia. IOSR Journal of Environmental Science, Toxicology and Food Technology, 8(12), 69-77.

Jamalani, M. A., Abdullah, A. M., Azid, A., Ramli, M. F., Baharudin, M. R., Bose, M. M., Elhadi, R. E., Youssef, K., Gnadimzadeh, A., and Gumel, D. Y. (2016). Monthly analysis of PM10 in ambient air of Klang Valley, Malaysia. Malaysian Journal of Analytical Sciences, 20(5), 1159-1170.

Proposed Articles for Submission

Ben Youssef, Khaled Ali Ahmed, Abdullah Ahmad Makmom., Shafri, Helmi Zulhaidi Mohd., Ashaari, Zulfa Hanan. Exploring the relation between satellite aerosol optical depth and PM_{10} at Peninsular Malaysia

Ben Youssef, Khaled Ali Ahmed, Abdullah Ahmad Makmom, Shafri, Helmi Zulhaidi Mohd, Ashaari, Zulfa Hanan. Collection 6 MODIS aerosol products validation based on AERONET distributed data around the Malaysia region.


UPM
UNIVERSITI PUTRA MALAYSIA
PRPi16
SOARING UPWARDS
BRONZE MEDAL
This certificate of Award is presented to
AHMAD MAKMOM ABDULLAH, KHALED ALI AHMAD BEN YOUSSEF, HELMI Z. M. SHAFRI, ZULFA HANAN ASHA'ARI, DANLADI YUSUF GUMEL, SARA Y. YUSOF
for the invention/innovation of
ESTIMATING GROUND-LEVEL PM2.5 IN PENINSULAR MALAYSIA USING SATELLITE REMOTE SENSING
at
PAMERAN REKACIPTA, PENYELIDIKAN DAN INOVASI 2016
on
15 - 16 November 2016
Dewan Besar, PKKSSAAS, UPM Serdang
PROF. DATIN PADUKA DR. AINI IDERIS, FASc,
DSIS, PSK, KMN
Vice Chancellor
Universiti Putra Malaysia

UPM
UNIVERSITI PUTRA MALAYSIA

FACULTY OF ENVIRONMENTAL STUDIES

Certificate of Participation

This is to certify that

Khaled Ali Ahmed Ben Youssef

had participated in the

Three Minute Thesis Competition 2018
(Faculty Level)

on

22 January 2018

held at

Faculty of Environmental Studies
Universiti Putra Malaysia

Associate Professor Dr. Mohammad Firuz Ramli
Deputy Dean
(Postgraduate Studies, Research and Development)

UNIVERSITI PUTRA MALAYSIA

STATUS CONFIRMATION FOR THESIS / PROJECT REPORT AND COPYRIGHT

ACADEMIC SESSION : First Semester 2019/2020

TITLE OF THESIS / PROJECT REPORT :

ESTIMATION OF GROUND LEVEL $PM_{2.5}$ CONCENTRATION USING AEROSOL OPTICAL THICKNESS FROM MODIS IMAGES IN PENINSULAR MALAYSIA

NAME OF STUDENT: KHALED ALI AHMED BEN YOUSSEF

I acknowledge that the copyright and other intellectual property in the thesis/project report belonged to Universiti Putra Malaysia and I agree to allow this thesis/project report to be placed at the library under the following terms:

1. This thesis/project report is the property of Universiti Putra Malaysia.
2. The library of Universiti Putra Malaysia has the right to make copies for educational purposes only.
3. The library of Universiti Putra Malaysia is allowed to make copies of this thesis for academic exchange.

I declare that this thesis is classified as :

*Please tick (√)

☐	**CONFIDENTIAL**	(Contain confidential information under Official Secret Act 1972).
☐	**RESTRICTED**	(Contains restricted information as specified by the organization/institution where research was done).
☑	**OPEN ACCESS**	I agree that my thesis/project report to be published as hard copy or online open access.

This thesis is submitted for :

☐ **PATENT** Embargo from _____ until _____
(date) (date)

(Signature of Student)
New IC No/ Passport No.: R2KK1116F

Date : 16/01/2020

Approved by:
PROFESSOR DR. AHMAD MAKMOM ABDULLAH
Deputy Dean (Academic and Student Affairs)
Faculty of Environmental Studies
Universiti Putra Malaysia
43400 UPM Serdang

(Signature of Chairman of Supervisory Committee)
Name:

Date : 16/01/2020

[Note : If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization/institution with period and reasons for confidentially or restricted.]

Fine particulate matter is particulate matter lower in diameter than 2.5 μm ($PM_{2.5}$). It affects the public health, economic development, and the regional climate. Governments worldwide have been concerned about the levels of the $PM_{2.5}$ in the atmosphere and long ago began monitoring their levels continuously using air quality monitoring stations. The Malaysian government invested much in building ground monitoring stations. Most of these stations have been located in urban areas. In this context, use of remote sensing (RS) techniques and the geographic information system (GIS) in estimating the levels of the ambient $PM_{2.5}$ become more widespread. The Moderate Resolution Imaging Spectroradiometer (MODIS) Aerosol Optical thickness (AOT) product of the Terra Satellite can be used to estimate the $PM_{2.5}$ levels with an accuracy that depends on the statistical relationship between $PM_{2.5}$ and AOT. This study aimed at estimating the $PM_{2.5}$ mass concentration in Peninsular Malaysia by linking the MODIS sensor data with the measured $PM_{2.5}$ concentrations and meteorological parameters in the year 2013. The results showed that the the map of distribution of AOT indicated that the AOT concentrated in the western coast of Peninsular Malaysia. Mostly, the spatial and temporal AOT values in the southwest monsoon were higher than in the northeast monsoon throughout the study period. The MLRA and GWRM both gave almost identical estimates of the $PM_{2.5}$ concentrations. The analysis outcomes revealed that the R^2 value for the hourly $PM_{2.5}$ regression model (0.66) was higher than that for the daily $PM_{2.5}$ (R^2 = 0.53). Comparison with the literature uncovers that results of estimation of $PM_{2.5}$ using AOT from MODIS for Peninsular Malaysia are similar to the results of other studies in other parts of the world. Furthermore, assessment of the accuracies of the hourly and daily estimates of $PM_{2.5}$ disclosed that the 5 x 5-pixel size model had the lowest values of the mean-squared error, root mean-squared error, and relative root mean-squared error. The relatively low error values associated with the 5 x 5-pixel size model indicate the accuracy of this model in Peninsular Malaysia.

School of Graduate Studies
Universiti Putra Malaysia
43400 UPM Serdang
Selangor, Malaysia