Relationship between Particulate Matter (PM2.5) Concentration and Aerosol Optical Depth (AOD) throughout Peninsular Malaysia

Noorfatekah Talib1*, Nur Amira Mohd Sa’aidi1, Rohayu Haron Narashid1, Nurul Ain Mohd Zaki1, Nurhafiza Md Saad1, Suzanah Abdullah2, Fazly Amri Mohd1

1Environment and Climate Change Research Group, Centre of Studies for Surveying Science & Geomatics, Faculty of Architecture, Planning and Surveying, Universiti Teknologi MARA, Cawangan Perlis, Kampus Arau, 02600 Arau, Perlis, Malaysia

2Centre of Studies for Surveying Science & Geomatics, Faculty of Architecture, Planning and Surveying, Universiti Teknologi MARA, Cawangan Perak, Kampus Seri Iskandar, 32610 Bandar Baru Seri Iskandar, Perak, Malaysia

Email: noorf492@uitm.edu.my

Abstract. A particular matter (PM2.5) is a small particle in the air that reduces visibility and causes air to fade when raised. It is one of the parameters that contribute to air pollution. There are a few causes of air pollution. It can appear in solid and liquid particles and certain gases that are suspended in the air. Consequently, pollution can significantly and negatively impact human health. This study aimed to explore the relationship between concentrations of PM2.5 and aerosol optical depth (AOD) throughout Peninsular Malaysia using the GIS approach. The tabular data of PM2.5 was acquired from the Department of Environment and overlaid with the MODIS satellite image. A correlation test was performed using QGIS software using Inverse Distance Weighting (IDW) technique. Then, by using the Air Pollution Index data and the value of Aerosol Optical Depth, this study can identify the values that influence the pollutants by studying the relationship between PM2.5 and AOD using MODIS satellite images. The linear regression of PM2.5 concentration with the MODIS AOD has shown significant differences. This study aligns with the eleven Sustainable Development Goals which are focused on sustainable cities and communities, to reduce the adverse environmental impact of cities per capita, especially the air quality such as fine particulate matter (PM2.5 and PM10) in cities.

Keyword: PM2.5, pollution, air, Aerosol Optical Depth (AOD)
1. Introduction
Particulate Matter (PM2.5) is one of the main causes of air pollution affected by many factors such as motor vehicles, urbanization, industrial emissions, open combustion, and biomass combustion [1]. PM2.5 means the mass per cubic meter of molecules of interest with an estimated (width) of less than 2.5 micrometers (μm) [2]. Besides, air pollution can also distress human health, such as heart attack, asthma, bronchitis, other respiratory problems, and others. This is because PM2.5 is a small-sized particle in the air [3]. Malaysia also has recorded a higher number of PM2.5 pollutants. [1] have proven the factors coming from motor vehicles, industrial emissions, and Indonesian peat fires during the southwest monsoon.

In this study, the GIS open software is used to identify PM2.5 throughout Peninsular Malaysia. A geographic information system can have great benefits, but some software is quite expensive especially for small companies [4]. This study used open-source software because it can be used freely. QGIS is one of the famous open-source software today. Besides that, this study used QGIS to run the PM2.5 data to identify its ability to analyze PM2.5 using the inverse distance weighted (IDW) interpolation method. This software is very useful in the Geographic Information System (GIS), producing solutions to spatial problems [5]. The IDW interpolation is a precise method that enforces the condition for influencing the estimated value more with the nearest point rather than further away. This means that the IDW point near the sample point gives more weight when calculating the mean [8]. This study also used MODIS satellite images to identify the value that affects the contaminants by reviewing the relationship between certain matters (PM2.5) and the aerosol optical depth (AOD). Aerosol optical depth is a measure of the extinction of solar beams by dust and haze [6]. Therefore, the particle in the atmosphere such as smoke, dust, and other contaminations can block sunlight by absorbing it or with a light of duality. The AOD data is derived from satellite images (MODIS) and they have been widely used to assess air quality. From the results of this study, the highest area of PM2.5 can be analyzed.

1.1. Significance of Study
Particular matter (PM2.5) and aerosol are some of the particles that cause air pollution. These particles are important to researchers because they represent the uncertainty of large areas in order to understand the Earth’s climate system. PM2.5 and aerosols have a very important role in determining the earth’s radiation estimates and impacts on weather changes. Likewise, PM2.5 and aerosol are also significant for human health. Therefore, good air quality is important for our health and the environment. Besides that, the sources of pollution are often difficult to monitor. GIS technology can manage statistical and spatial data to provide tools that show the relationship between poor air quality, and human health, and poor environment. Consequently, GIS aids in monitoring pollutant emissions.

1.2. Scope and Limitation of Research
The particulate in this study was a fine particle of PM2.5 and aerosol optical depth (AOD) using MODIS satellite images to interact with the aerosol data. Secondly, the distribution was shown using the IDW interpolation method, but the scope of the study was limited due to the experimental parameter. In this study, the data is collected as average per month in the year 2019. Thus, there were limitations on the variations.

2. Study Area
The location included entire Peninsular Malaysia with coordinate 03° 08’ 00” N, 101° 41’ 00” E to study the particulate matter and aerosol optical depth of the ambient air. Malaysia is a developing country, where it can contribute to pollution cases. The area size of 131,587sq km consisted of 47 monitoring stations for air pollution. The population is about 32.5 million. Figure 1 shows the location of the study area.
2.1. Data Acquisition

There were three types of data needed in this study. The first data were MODIS satellite images. These data were needed to get the parameter of data aerosol from that satellite image to do a relationship between PM2.5 and AOD. The AOD is useful to estimate the spatiotemporal distribution of PM2.5 levels [7]. Secondly, the data were the base map of Peninsular Malaysia. Then digitized using QGIS software. The third data were the tabular data of particular matter 2.5 index level in the year 2019. The data is obtained from the APIMS Malaysia website.

2.1.1. MODIS Satellite Images.

The collection of aerosol products used in this study was available from NASA’s Level 1 Distribution Archive and System (LAADS). Alternatively, this data could be generated using a MODIS Processing Package for affordable ground stations that received live broadcasts from Terra and Aqua satellites, which could be useful for monitoring air pollution in almost real-time. For this study, Deep Blue Aerosol Optical Depth ground measuring 135 × 203 16-bit integer datatype was abstracted from the MODIS image. The actual ground footprint for this AOD satellite data ranged from 10 to 10 km. The picture is freely downloaded.

2.1.2. Location and Coordinate of Air Quality Monitoring Station.

There were 47 stations available in Peninsular Malaysia and operated by the Department of Environment Malaysia (DOE) [8]. The air quality level and status were continuously monitored by DOE for 24 hours each day and controlled by the automated air quality control remote station. The data were collected starting from August 2018. The calculation of pollutants was according to the PM2.5 calculation based on the Air Pollutant Index (API) until 2019 and taken from DOE or APIMS website. Regarding DOE, selected pollutants were monitored because of their presence as major air pollutants, and they followed the criteria of pollutants from other countries and the World Health Organization. The main air pollutants were sulfur dioxide (SO2), nitrogen dioxide (NO2), ozone levels (O3), carbon monoxide (CO), particle material with a diameter of fewer than 10 microns (PM10), and PM2.5. For this study, PM2.5 was chosen because it was suitable for most industrialized countries, and this would further enhance the Malaysian Government’s efforts towards environmental protection.
3. Method
3.1. Pre-processing
There were several stages of pre-processing that needed to be done, such as subsetting, reprojecting, layer stacking, radiometric correction, and geometric correction. Layer stacking is the process of joining multiple images into one image. The picture must have the same level (number of rows and number of columns) to retrieve another strip that has a spatial resolution that is different from the target resolution. Then, radiometric and geometric corrections were needed to be done to get clearer images, reduce errors, and improve the accuracy of satellite images. Geometric corrections were performed to set the image as the exact location for the feature data while radiometric correction was to increase the consistency correction of the magnitude of the brightness value involved in the relative spatial consistency of the image. A subset is a process of retrieving just the parts of large files which are of interest for a specific purpose. The satellite image was reprojected to local datum projection from WGS84/UTM to MRSO. Reprojection is very important to make sure the position of the image is in the right projection.

3.1.1. Extraction of MODIS Data. After completing the pre-processing for MODIS images, the next process was to extract AOD data using the Extract Sub data tools. Deep Blue Aerosol Optical Depth ground data was used by measuring 135 × 203 16-bit integer data type.

3.2. Post-processing
Georeferenced data, also known as spatial, geographical, or geospatial data, are the basic pieces of information needed to identify the geographic location of phenomena across the Earth’s surface. For this study, this process was done to determine the position of PM2.5 contaminations accurately. Then, geographical geocoding was done to obtain PM2.5 tabular data from a data source, analyzing it to geographic layers, such as shapefile and finally showing address match in QGIS.

3.2.1. Inverse Distance Weighting (IDW) Interpolation. This process involved the process of creating spatial interpolation using the IDW interpolation method. Spatial analysis is the method of monitoring spatial data to extract unused data and the importance of unique information. IDW is a specific interpolator, hence statistical values are well-considered. IDW is the best interpolation method in Indian air quality, and it has smaller faults than the kriging method on all pollutants [9]. GIS offers a practical and important work environment for integrating, analyzing, and visualizing this data along with new spatial data sources. Other than that, according to [9], IDW is the best interpolation method to determine air pollution conditions from Ordinary Kriging (OK) and Universal Kriging (UK) methods. The spatial interpolation method is a process of using points with known values to estimate values at other points. In GIS applications, spatial interpolation is usually applied to the raster with a budget created for all cells. Therefore, spatial interpolation is a method of making surface data from sample points. This study ran the IDW method using QGIS to analyze the spatial pattern of PM2.5 air pollution and PM2.5 variations in Peninsular Malaysia.

3.2.2. Correlation Coefficients Process. The correlation analysis was done to find the relationship between two parameters which were PM2.5 pollutant and AOD particles. The range of correlation coefficients is from -1 to 1. Besides that, the relationship of the correlation coefficient depends on its value. For example, if its value is close to ± 1, then it is deemed to be a perfect correlation while when one variable increases, the other variable tends to also increase (if positive) or decrease (if negative). If the value of the coefficient lies between ± 0.50 and ± 1, then it is said that the correlation is strong between the two parameters. The determination coefficient is the measure of how well the regression line represents the data [10]. The result of the correlation coefficient (r) was extracted from the graph and was recorded within the table. The regression line reflected the value of the correlation coefficient.
3.2.3. **Linear Regression.** In statistical modeling, regression analysis is a set of statistical processes to estimate the link between two variables, including many performances for modeling and analyzing some variables. However, the main concentration is on the relationship between the dependent and independent variables. For more detailed purposes, regression analysis helps to understand how representative values change when any independent variable changes while other independent variables have improved. The analysis shows how many independent variables are connected to the dependent variables and at the same time. Many methods can be used in conducting regression modeling and analysis as the most common method is linear regression and the least common is the squared regression method. In that regression, a function is used to calculate the number of unknown parameters from the data. For the usual smallest square method, it selects linear function parameters from a set of explanations with the least-squares principle. This minimizes the number of different boxes in the data set between the dependent variables observed and those predicted by the linear function.

4. **Results and Discussion**

4.1. **Concentration Map of PM2.5 over Peninsular Malaysia**

   The IDW interpolation method was performed periodically to show that the spread of PM2.5 and the contribution to the highest and lowest index map (Figure 2). The result has shown that the highest concentration at the Shah Alam station with a value of 97.17670. The lowest concentration of the PM2.5 index was at Kota Tinggi stations where the value was 39.00464. The state of Negeri Sembilan also indicated an unhealthy index which is above 60 levels. This is probably because the highest concentration region indicates the urban area where they contribute to economic development. Previous study by [11] stated that the urban area in Selangor has increasing rapidly from year 2016 to year 2019.

![Figure 2. Concentration Map of PM2.5 over Peninsular Malaysia](image)

4.2. **Extraction of Aerosol Optical Depth Map**

   The AOD map in Figure 3 was based on the data from the MODIS satellite images. The green areas show where the aerosol was regularly in larger particles while the red areas show that the aerosol regularly contained small particles. Meanwhile, the yellow areas have shown both of the particles. The white colour represented where the sensors did not collect or capture the data.
4.3. Interpolation map of Aerosol Optical Depth Land
By using QGIS Software, each pixel was converted into a point, and each pixel was associated with calculated AOD. The white area of the image contained no data, but the interpolation method was used to estimate the value from the no-data region. The IDW interpolation tools were applied where the unknown data were interpolated using the weighted average of neighbouring samples. The results displayed the distribution of AOD value in Peninsular Malaysia. Then, the blue colour has shown where the aerosol presented in large particles. The red colour area presented where aerosol regularly contained with small particles. From this map, it could be concluded that the very large particle size was located in Pasir Gudang station with the value of particles 486.00338 (Figure 4).

Figure 3. Result extract Aerosol Optical Depth Map

Figure 4. Interpolation map of Aerosol Optical Depth Land
4.4. Correlation Analysis

4.4.1. Relationship between PM2.5 and AOD.

The scatterplots graph in Figure 5 shows the result of the correlation between PM2.5 pollutant index and AOD particles. Table 1 shows the details of R squared and adjusted R squared. The value of 0.5 and above indicates a positive strong relationship while below 0.4 shows a weak positive relationship [12]. The relationship table shows the PM2.5 pollutants index has a weak positive relationship with the AOD with the R² value of 0.1297. The reason for the weak correlation because of less diurnal variation of the particulate matter over the region. The cloud cover also can be one of the possibilities. The spatial averaging of MODIS AOD to compensate for cloud cover is also expected can reduce the representativeness [12].

Figure 5. Relationship between PM2.5 and AOD

| R Squared (R²) | Adjusted R Squared (R) |
|----------------|------------------------|
| 0.129651064    | 0.360138862             |

5. Conclusion

In this study, MODIS AOD and PM2.5 data were used to perform the spatial analysis and identify the relationship between PM2.5 and AOD using correlation analysis. From the results, it indicates that Selangor and Negeri Sembilan have the highest value of PM2.5 pollution distribution index compared to other states. Based on this, it is proven that Selangor and Negeri Sembilan are the fastest developing areas with high industrial activities, increasing urbanization, and has the highest population density. All these factors are potential in contributing the PM2.5 pollution distribution problems compared to other areas that have shown less development and urbanization. Along with urbanization, industrialization, and rapid economic development, fine particle pollution also has degraded the atmospheric visibility in this region. This is in line with the aspiration of the Ministry of Environment and Water in implementing the Environmental Quality (Clean Air) Regulations 2014 and Environmental Quality Act (1974). The value of PM2.5 can increase over time if the local authorities ignore this problem. This will worsen with the existence of transboundary haze, which might increase the susceptibility of the public to respiratory problems. Lastly, MODIS AOD is not well correlated with PM2.5 concentrations. The weaker correlations are possibly due to the corresponding aerosol observation over the region. In addition, greater satellite coverage will allow a better estimation of surface aerosol concentration [11]. Even though the correlation between PM2.5 and AOD is weak, however, the regression functions showed significant differences among the seasonal variations [13]. Therefore, the study can be continued with different seasons.

Acknowledgement

The authors would like to express their gratitude for Department of Environment, Malaysia for providing the data.
References

[1] Fujii, Y., Mahmud, M., Tohno, S., Okuda, T., & Mizohata, A. (2016b). A case study of PM2.5 characterization in Bangi, Selangor, Malaysia during the southwest monsoon season. A Case Study of PM2.5 Characterization in Bangi, Selangor, Malaysia During the Southwest Monsoon Season., 16 (11), 2685–2691. https://doi.org/10.4209/aaqr.2015.04.0277

[2] Liu, Y., Park, D. J., Li, Q., Kilaru, V., & Sarnat, J. A. (2004). Mapping annual mean ground-level PM 2.5 concentrations using Multiscale Imaging Spectroradiometer aerosol optical thickness over the contiguous United States. 109, 1–10. https://doi.org/10.1029/2004JD005025

[3] Martins and Carrhilo da Graça, (2017) Effects of Airborne Fine Particle Pollution on the Usability of Natural Ventilation in Office Buildings in Three Megacities in Asia, Renewable Energy 117, DOI: 10.1016/j.renene.2017.10.089

[4] Khan, S., & Mohiuddin, K. (2019). Evaluating the parameters of ArcGIS and QGIS for GIS Applications. January 2018.

[5] Benduch, P. (2017). Comparison of the results of Spacial Analyses performed in ArcGIS and QGIS software. Comparison of the Results of Spatial Analyses Performed in Arcgis and QGIS Software, IV(4), 1887–1902. https://doi.org/10.14597/infraeco.2017.4.3.143

[6] Ali, K., Ben, A., Ser, C., Environ, E., Ali, K., Ben, A., Abdullah, A. M., & Zulhaidi, H. (2019). The development of Air Quality Indices using AOD-Retrieved Images during haze events in Peninsular Malaysia the development of Air Quality Indices using AOD-Retrieved Images during haze events in Peninsular Malaysia. https://doi.org/10.1088/1755-1315/373/1/012027

[7] Depth, A. O. (2019). Aerosol Optical Depth. March 2000, 1–2.

[8] Department of Statistics Malaysia Official Portal. (2021). Retrieved from https://www.dosm.gov.my/v1/index.php?r=column/cone&menu_id=d3EvQThuRnFsSGlrQXRtY2E5R0wvZz09

[9] Jumaah, H. J., Ameen, M. H., Kalantar, B., Rizeei, H. M., Jumaah, S. J., Jamal, H., Ameen, M. H., & Kalantar, B. (2019). Air quality index prediction using IDW geostatistical technique and OLS-based GIS technique in Kuala Lumpur, Malaysia technique, and OLS-based GIS technique in Kuala. Natural Hazards, and Risk, 10(1), 2185–2199. https://doi.org/10.1080/19475705.2019.1683084

[10] Asuero, A. G., Sayago, A., & González, A. G. (2006). The correlation coefficient: An overview. Critical Reviews in Analytical Chemistry, 36(1), 41–59. https://doi.org/10.1080/10408340500526766

[11] Talib N., Sallehuddin N.S.M., Sa’aid N.A.M., Saad N.M. (2020) PM2.5 Pollutant Distributions in Years 2016 and 2019 Using GIS Spatial Analyst. In: Alias N., Yusof R. (eds) Charting the Sustainable Future of ASEAN in Science and Technology. Springer, Singapore. https://doi.org/10.1007/978-981-15-3434-8_15

[12] Boon Ning Chew, James R. Campbell, Edward J. Hyer, Santo V. Salinas, Jeffrey S. Reid, Ellsworth J. Welton, Brent N. Holben, Soo Chin Liew (2016). Relationship between Aerosol Optical Depth and Particulate Matter over Singapore: Effects of Aerosol Vertical Distributions. Aerosol and Air Quality Research, 16: 2818–2830, 2016. https://doi.org/10.4209/aaqr.2015.07.0457

[13] Jinyuan Xin, Qing Zhang, Lili Wang, Chongshui Gong, Yuesi Wang, Zirui Liu and Wenkang Gao (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. https://doi.org/10.1016/j.atmosres.2013.11.001