Available online at
https://journal.uitm.edu.my/ojs/index.php/MJCET

e-ISSN: 2682-8588

Malaysian Journal of Chemical Engineering & Technology 7(2) 2024, 75–89

Malaysian Journal of Chemical Engineering and Technology

Air pollution index evaluation based on haze phenomena in East Malaysia using Giovanni satellite database

Zafirah Muhamad Haitamin[1], Norhusna Mohamad Nor[2*]

[1,2]*Chemical Engineering Studies, Universiti Teknologi MARA, Cawangan Pulau Pinang, Permatang Pauh Campus, 13500 Pulau Pinang, MALAYSIA*
[2]*Waste Management and Resource Recovery (WeResCue) Group, Universiti Teknologi MARA, Cawangan Pulau Pinang, Permatang Pauh Campus, 13500 Pulau Pinang, MALAYSIA*

ARTICLE INFO

Article history:
Received 15 February 2024
Revised 30 September 2024
Accepted 2 October 2024
Online first
Published 31 October 2024

Keywords:
Air Pollution Index
Giovanni
Haze
Linear Regression Analysis
Satellite Database

DOI:
10.24191/mjcet.v7i2.886

ABSTRACT

This work evaluates air pollution related to haze phenomena in Malaysia from 2000 to 2019. The main objective of this work is to evaluate the relationship between the Air Pollution Index (API) during haze occurrences using the Giovanni satellite database. The collected data of the API was denoted as ground-based data, while that of GIOVANNI was denoted as satellite-based data. The air pollutants targeted in the study include $PM_{2.5}$, SO_2, CO, and O_3. Sarawak and Sabah were chosen as the study areas due to the high levels of hazardous haze pollutants observed in these regions. The data analysis utilised a linear regression approach to examine the correlation and relationship between ground-based and satellite-based measurements. Factors contributing to haze occurrence were also investigated by gathering meteorological data from GIOVANNI, including wind speed and surface temperature. The analysis's correlation coefficient (R) values range from weak, moderate and strong, with all p-values below 0.05, indicating statistical significance. Notably, wind speed shows a strong negative correlation with API, with an R-value of −0.8750, demonstrating an inverse relationship between the two variables. Similarly, temperature exhibits a moderate negative correlation with API, reflected in an R-value of −0.7270. The findings indicate a strong inverse relationship between the factors and haze pollution, with correlations from the GIOVANNI database serving as a benchmark for identifying causes of high API during haze.

1. INTRODUCTION

Rapid industrialisation and urbanisation in developing countries have resulted in a significant rise in various environmental and pollution problems, especially air pollution. Air pollution has become a vital and critical component of these issues affecting human health, crops, forest species, and ecosystems (Afroz et al., 2003). Haze is a common air pollution phenomenon that occurred almost yearly in Malaysia over the last

[2*]Corresponding author. *E-mail address*: norhusna8711@uitm.edu.my
https://doi.org/10.24191/mjcet.v7i2.886

few decades. Haze pollutants usually consist of PM_{10}, $PM_{2.5}$, SO_2, NO_2, CO, and O_3, and in Malaysia, the most measured parameters are $PM_{2.5}$ and PM_{10} concentrations (Redzuan et al., 2023). According to the Department of Environment Malaysia (DOE), the particulates or fine particles that cannot be seen by the naked eye and are suspended in the atmosphere at high concentrations contribute to the haze phenomenon (Department of Environment, 2016a). This phenomenon happens when a sufficient concentration of aerosols in the atmosphere scatters the visible light, resulting in a measurable reduction in the visual range (Latif et al., 2018). Haze can also occur because of large concentrations of secondary gas, such as surface ozone, which can combine with volatile organic compounds to form secondary organic aerosols (Latif et al., 2018).

Haze in Malaysia is not a new phenomenon, as it was first recorded in 1983, and the highest API value was recorded in 1997 at Kuching, Sarawak, which was 839. This phenomenon consistently happened almost every year in Malaysia, continuing to 2005, 2006, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2018, and 2019. As reported by BBC News (Vaughn, 2019), one of the primary sources of haze condition is commonly known from the open biomass burning in Indonesia, and this burning usually peaks during Indonesia's dry season when many farmers will take advantage of the weather and conditions to clear vegetation lands for palm oil, pulp and paper plantations using the slash-and-burn method. The influence of wind causes these open biomasses burning from Malaysia's neighbours to cross the borders and contributes to the increased aerosol loadings or smoke in the atmosphere, which releases concentrated particulate matter such as organic matter, graphitic carbon, toxic metals, and acid species that are hazardous to human's health (Othman et al., 2014). Exposure to particle pollution is also considered a public health hazard and causes severe damage to human health as inhalation of these particulate matters can aggravate and cause heart and lung disease as it can travel deep into the lungs, thus also causing cold attacks, chest pains and cause eye-associated illness such as conjunctivitis (Latif & Hamzah, 2016; Othman et al., 2014). Haze can also negatively impact the economy, causing billions of dollars in losses due to healthcare costs and income opportunities lost. It could also affect the tourism industry, as foreigners would be less likely to choose Malaysia as their vacation spot because of the air pollution.

To investigate the amount and concentration of air pollutants caused by the haze phenomenon, satellite databases, commonly known as remote sensing tools, can be used to understand haze conditions' statistics further. Remote Sensing is the science of acquiring information about the Earth's surface without actually being in contact with it, and it is done by sensing and recording reflected or emitted energy and processing, analysing, and applying the information (Al-Alola et al., 2022). The physical characteristics of an area are detected and monitored by measuring its reflected and emitted radiation at a distance from the targeted area (Christopherson et al., 2019). To obtain the information, many researchers used remote sensing tools such as Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY), Geospatial Interactive Online Visualisation and Analysis Infrastructure (GIOVANNI), Moderate Resolution Imaging Spectroradiometer (MODIS), Phased Array L-band Synthetic Aperture Radar (PALSAR), Iterative Self-Organizing Data Analysis Techniques (ISODATA) and more. In Malaysia, the monitoring networks and quality data are generally recorded and managed by ground stations nationwide, monitored by the DOE via a private company known as Alam Sekitar Malaysia Sdn. Bhd. (ASMA) and the Malaysia Meteorological Department (MetMalaysia) (Muzammil & Hanan, 2017). The data collected from remote sensing tools can be analysed using multiple data analyses such as linear regression analysis (LRA), non-linear regression analysis (non-LRA), and multiple regression analysis (MRA). Regression analysis is a proper statistical method that allows you to examine the relationship between two or more variables of interest (Julfikar et al., 2020). These statistical analyses are commonly used to determine the relationship between remotely sensed and ground-based data. It can also test the air prediction model's accuracy by comparing the data provided by government facilities and satellite sensors (Muzammil & Hanan, 2017).

In this research work, the focus will be on data collection of ground-based data from DOE and satellite databases from GIOVANNI. The ground-based data from DOE will be collected through the API readings on the DOE website and updated hourly. API is an index system that is considered a simple, comprehensive

https://doi.org/10.24191/mjcet.v7i2.886

approach to defining air quality status so that the general public can easily understand (Latiffah et al., 2018). Even though the advantage of using API for policies and regulatory actions is that it can reveal air quality status and its effects on human health, it also has its cons, which are only based on the highest sub-index value of the pollutants (Latiffah et al., 2018). By collecting a satellite database from GIOVANNI, the air pollutant concentrations during the haze phenomenon can be specified and compared with the ground-based data. The comparison between the two data is then analysed using statistical data analysis to understand the haze pollutants' trend further.

2. METHODS

Fig.1 shows the overall process flow for this research work.

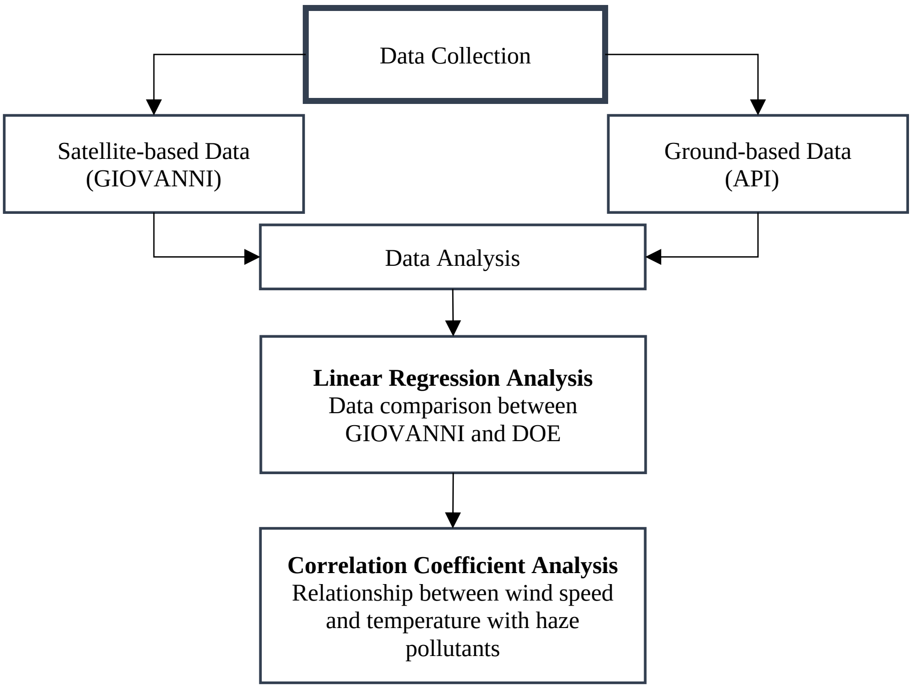

Fig. 1. Process flow of API evaluation based on haze phenomena using GIOVANNI

Source: Authors’ illustration

2.1 Data collection at selected study areas

To collect data on the emission of air pollutants during the haze phenomenon, both ground-based data and satellite databases were collected from Department of Environment (DOE) and GIOVANNI, respectively. The study areas selected for both databases are Sarawak and Sabah. These study areas were selected as some hazardous and high API values spiked and detected at Sarawak and Sabah. The ground-based data were obtained from one of the DOE’s monitoring stations at Miri, Sarawak and Tawau, Sabah. For the satellite-based data, the locations are precisely coordinated at 4.3995° N, 113.9914° E, and 4.2447° N, 117.8912° E, for Miri, Sarawak and Tawau, Sabah, respectively(Fig.2).

https://doi.org/10.24191/mjcet.v7i2.886

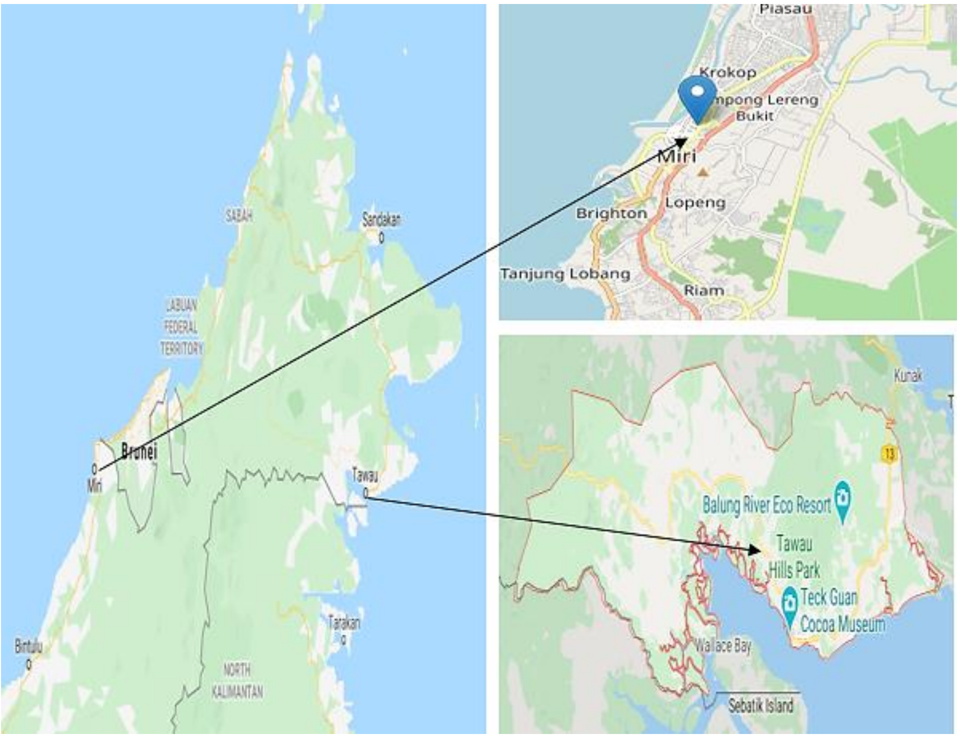

Fig. 2. The selected study areas

Source: Authors' illustration

Ground-based data - DOE

The haze pollutant concentrations were obtained from the DOE from 2000 to 2019, corresponding to hourly data measurement. In addition to the API value readings, DOE provided meteorological data such as wind speed and hourly temperature from 2000 to 2019.

Satellite-based Data - GIOVANNI

As for the satellite database, the data on haze pollutants were collected from the GIOVANNI website from 2000 to 2019. GIOVANNI provides remote sensing data alongside several different basic analytical capabilities, which include spatial maps of data variable values, difference maps, area-averaged time-series, animations, and vertical profiles of atmospheric variables, thus making the data obtained easier to use, has a large number of data sets available, also allows users to customise the analysis, and it is uniquely well suited for air quality event analysis (Prados et al., 2010). The time-series option for Area-averaged was chosen, while the date range was selected from early 2000 to 2019, as shown in Fig. 3. As for the region, the type of shape chosen was the country's shape, Malaysia. The variables for the haze pollutants chosen (SO_2, CO, O_3) were processed, and the data were downloaded in comma-separated values files. Notably, data for $PM_{2.5}$ was selected from 2017 to 2019 due to the inclusion of $PM_{2.5}$ in the API calculation starting from 2017 (Department of Environment Malaysia, 2016b). The selected data were extracted and were averaged according to yearly calculation, according to results presented in this work.

https://doi.org/10.24191/mjcet.v7i2.886

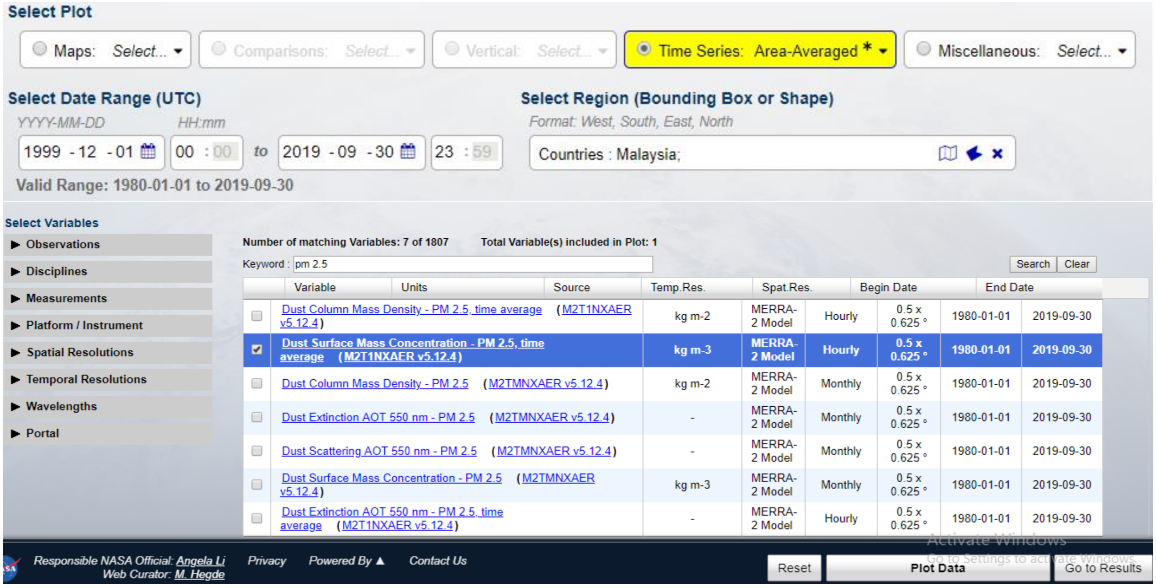

Fig. 3. GIOVANNI's interface for data/variables selection

Source: Authors' own data

2.2 Linear regression analysis

To conduct a regression analysis, two data sets were needed to estimate the relationship between two or more variables. In this case, linear regression was applied to find the relationship between the two parameters: the concentration of haze pollutants from ground-based data (DOE) and satellite-based data (GIOVANNI). The resulting coefficient from the analysis was an indicator of the correlation between the sets of data of ground-based and satellite (Muzammil & Hanan, 2017). Linear regression was analysed for each haze pollutant concentration, which are $PM_{2.5}$, SO_2, O_3, and CO. The application used to model the linear regression analysis was Microsoft Excel using the 'Data Analysis Tool'. Eq. (1) was used for the regression analysis where the y-axis was the ground-based data and the x-axis was the satellite-based data, which are as follows:

$$y=mx+c \tag{1}$$

where the independent variable (x) will be the satellite database obtained from GIOVANNI, while the dependent variable (y) will be the ground-based data from DOE, and m represents the slope between the y and x values.

Once the analysis is done, the R^2 values for each analysis will be shown, ranging from 0 (no correlation) to 1 (perfect correlation). Moreover, the p-value of each model was calculated as it can prove the level of the statistical significance of the analysis. The value of p (probability) must be typically less than 0.05 to show that the analysis is statistically significant and indicates strong evidence against the null hypothesis (McLeod, 2019).

https://doi.org/10.24191/mjcet.v7i2.886

2.3 Correlation coefficient analysis

The relationship between wind speed and atmospheric temperature with haze pollutants was analysed using correlation coefficient analysis to investigate the factors contributing to haze phenomena in Malaysia. Correlation coefficient analysis is a technique to investigate the relation between two variables. Thus, the R values obtained from the analysis represent the correlation coefficient, which measures the strength of the association between the two variables. A value of -1 indicates a 'perfect negative correlation', and a value of +1 indicates a 'perfect positive correlation'. The correlation coefficient analysis was done using Microsoft Excel using the Data Analysis Tool tab.

3. RESULTS AND DISCUSSION

3.1 Air pollution index (API) and remote-sensing data monitoring

The API data of Miri, Sarawak and Tawau, Sabah, from 2000 to 2019, were obtained from the DOE, as illustrated in Fig. 4.

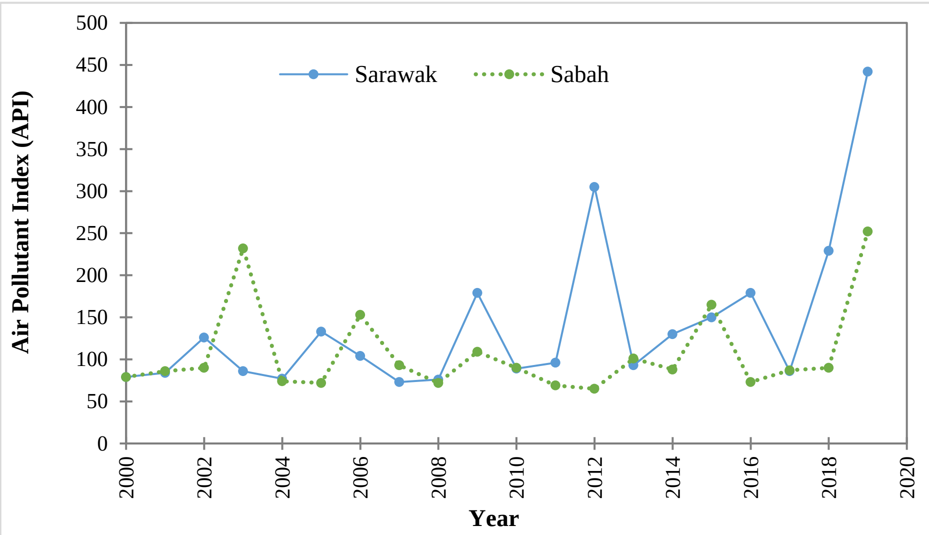

Fig. 4. API monitoring in Sabah and Sarawak from 2000 to 2019

Source: Authors' own data

Fig. 4 shows the trend of API values from 2000 to 2019 at a continuous air monitoring station in Miri, Sarawak and Tawau, Sabah, where the values ranged from moderate, unhealthy, and very unhealthy to hazardous. For Sarawak, the lowest API was 73 in 2007, which is in the moderate range, while the highest was 442 in 2019, which is in the hazardous range. The highest API value of 442 in 2019 was caused by transboundary haze from forest fires during slash-burn activities that affected Malaysia and other countries such as Thailand, Indonesia, Singapore, Brunei, and Vietnam. According to the Asean Specialised Meteorological Centre (ASMC), continuous moderate-to-dense smoke hazes were emanating and blowing from hotspots in South, Central, and West Kalimantan towards western Sarawak (New Straits Times, 2019). This also happened during August, which was during the dry season that usually ran from May to September, resulting in less rainfall (Latiffah et al., 2018). The API values at Sabah were much lower than at Sarawak as they only ranged from moderate, healthy, to very unhealthy, with the lowest being 65 and

https://doi.org/10.24191/mjcet.v7i2.886

the highest being 252. Sabah has two spikes of high API values, 232 and 252, during 2003 and 2019. It was reported that in 2003, the haze phenomenon was more severe on Borneo Island and parts of neighbouring Sabah, while the peninsula of Malaysia was not affected by it (Latif et al., 2018). According to The Borneo Post, the haze in 2019 over Sabah is believed to have been caused by open burning carried out by the locals, which was also further become worse as the hot weather contributed to it (Binti Udin et al., 2021).

Fig. 5 shows the trend for $PM_{2.5}$, SO_2, CO, and O_3 concentrations in Sarawak and Sabah from 2000 to 2019. To be noted, $PM_{2.5}$ data was incorporated into Malaysia's API calculation starting in 2016. The inclusion of $PM_{2.5}$ aimed to provide a more accurate representation of air quality, as $PM_{2.5}$ is a significant component of atmospheric pollution and poses serious health risks (Department of Environment Malaysia, 2016b).

From Fig. 5(a), it can be observed that there are multiple rises in the concentrations of $PM_{2.5}$ for both study areas, but they are generally higher in Sabah than in Sarawak. The highest concentrations for Sarawak and Sabah are during 2019, which are 0.654 and 0.745 µg/m^3, respectively. This correlates with the ground-based data during the haze period of 2019, which has very high API values, too. High concentrations of $PM_{2.5}$ are caused by open burning, thus causing air pollution to be worsened and also contributing to climate change (Akhtar & Palagiano, 2017)The SO_2 concentrations in Sarawak are much higher than in Sabah. Furthermore, the SO_2 concentrations in Sabah are also generally constant, with a slight increase in 2003. For Sarawak, the values fluctuate, with multiple rises in 2009, 2012, and 2016. SO_2 emissions are usually from motor vehicles and fossil fuel combustion at power plants, refineries, and other industrial facilities. This might be why the concentration values do not correlate with haze periods caused by forest fires.

There are also multiple studies conducted in other countries that showed a substantial reduction in SO_2 concentrations over the decade, as shown by the study of (Kan, 2022). Sarawak is shown to have much higher concentrations of CO, which peaked in the early years of 2001, 2004, 2006, and 2007 and eventually decreased to the average of 110 ppb in concentration, while the trend for concentrations of CO in Sabah seemed to have a constant trend with slight rises in 2006, 2009, 2013 and 2015 that ranged from 73.89 ppb to 124.99 ppb. It can be observed that the concentrations of ozone for both study areas fluctuated irregularly throughout the years. Furthermore, it can be observed that this trend is the rapid rise of ozone pollutants during 2019, which have the highest concentrations for both Sarawak and Sabah at 269.01 and 268.11 DU, respectively. According to the National Center for Atmospheric Research, wildfires can cause an increase in ozone concentrations to unhealthy levels even though the location of the fires is from large distances (Binti Udin et al., 2021). This is also supported by previous studies, which stated that a rise in the concentration of ozone was observed to be mainly a result of transboundary pollution and forest fires that also contributed to the deterioration of air quality over the years (Rajab et al., 2013). The primary sources of these pollutants include transportation emissions, industrial activities, domestic open burning, and transboundary haze pollution (Shafii et al., 2017).

https://doi.org/10.24191/mjcet.v7i2.886

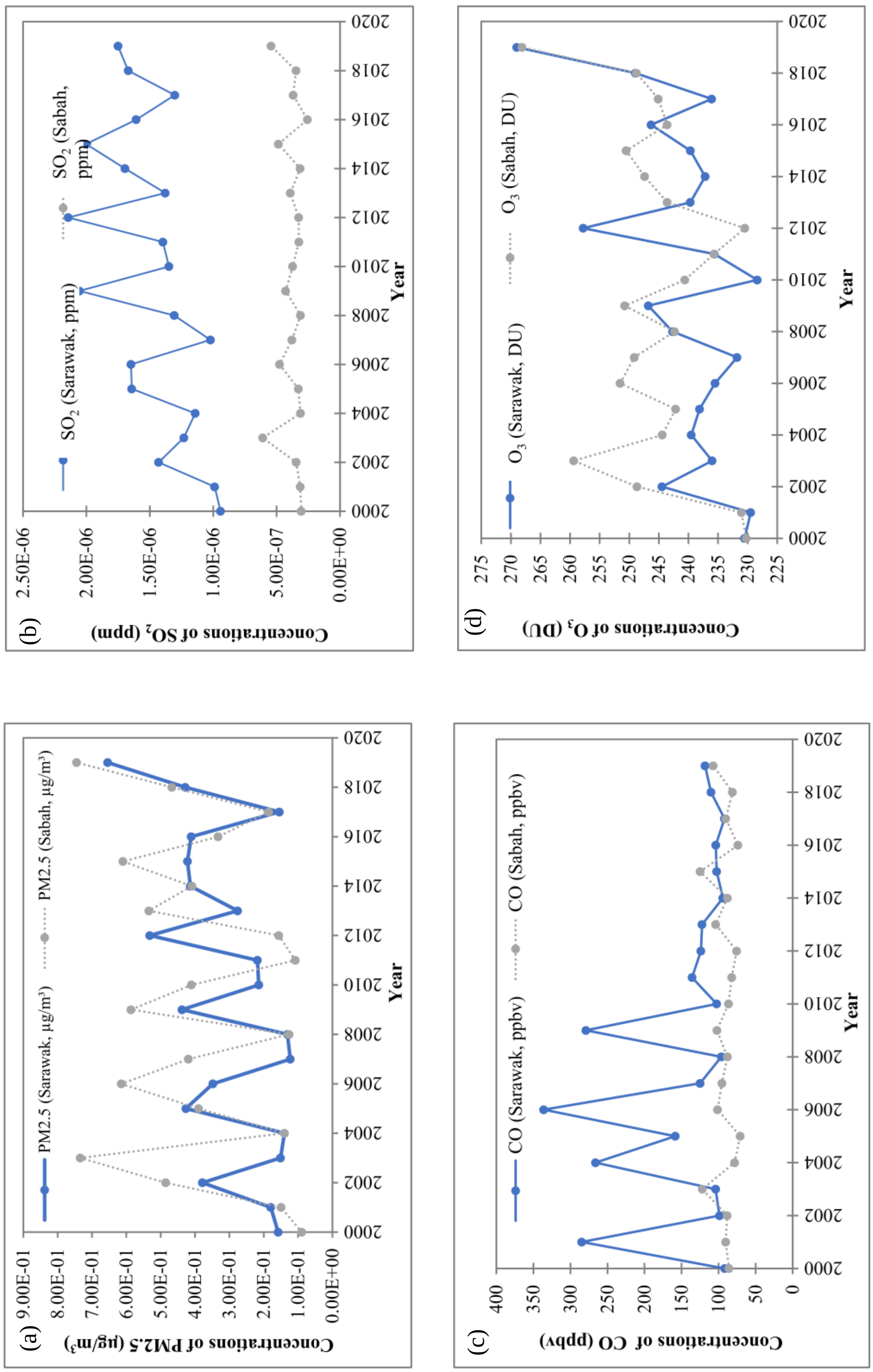

Fig. 5. GIOVANNI's data of a) $PM_{2.5}$, b) SO_2, c) CO, and d) O_3 concentrations in Sabah and Sarawak from 2000 to 2019

Source: Authors' own data

https://doi.org/10.24191/mjcet.v7i2.886

3.2 Linear regression analysis of Air Pollution Index (API) and GIOVANNI Data

Fig. 6 shows the linear regression analysis comparing the data for API (ground-based data, DOE) and $PM_{2.5}$, SO_2, CO, and O_3 (satellite database, GIOVANNI) in Sarawak and Sabah from 2000 to 2019. From Fig. 6(a), the ground-based data (API) and satellite database ($PM_{2.5}$) in Sarawak are analysed using the linear regression analysis method. It can be shown that API positively correlates with the concentration of $PM_{2.5,}$ with the R^2 value being 0.7478, indicating a solid correlation relationship. Thus, based on Table 1, the p-value is $8.27 \times 10^{-7,}$ which indicates that the analysis is statistically significant. For Sabah, it can be shown that the API is positively correlated with the concentration of $PM_{2.5}$, with the R^2 value being 0.6235, indicating a strong correlation. Thus, the p-value is 3.47×10^{-5} which indicates that the analysis is statistically significant.

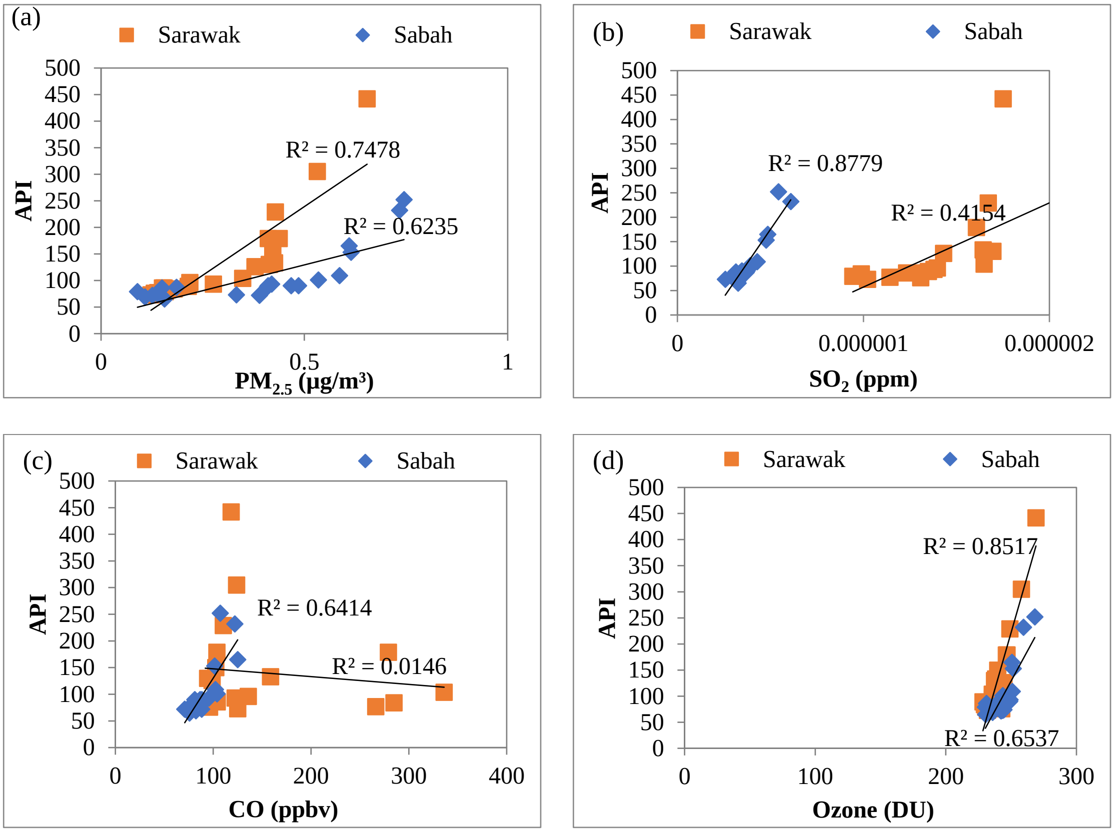

Fig. 6. API (Ground-based data) against GIOVANNI (Satellite-based data) of (a) PM2.5, (b) SO_2, (c) CO, and (d) O_3 concentrations in Sabah and Sarawak

Source: Authors' own data

From Fig. 6(b), the linear regression analysis method is done by plotting the ground-based data (API) against the satellite database (SO_2) in Sarawak. It can be shown that the API value is positively correlated with the SO_2 concentration, with the R^2 value being 0.4154, indicating a moderate correlation relationship. The low R^2 value for Sarawak might be because of the outlier data of the API, which is too high at 442, resulting in an abnormal and massive distance from the other data and values, thus affecting the slope of the linear regression analysis. It can also be shown from Table 1 that the p-value is 0.0022, which indicates

https://doi.org/10.24191/mjcet.v7i2.886

that the analysis is statistically significant. Next, for Sabah, it can be shown that API positively correlates with the SO_2 concentration, with the R^2 value being 0.8779, indicating a solid correlation relationship. Thus, the p-value is 1.19×10^{-9}, which indicates that the analysis is statistically significant.

It can be shown in Fig. 6(c) that the R^2 value of 0.0146 API versus the concentration of CO indicates a weak correlation relationship. Unfortunately, this model resulted in a weak correlation relationship as too much outlier data and values caused the regression analysis to be negative. This is because the CO data obtained from the GIOVANNI database were not aligned with the API trend, resulting in the lowest R^2 (0.0146). Regardless of the lowest R^2 value, the p-value for CO in Sarawak is 0.0027, indicating that the analysis is still within the range and statistically significant. For Sabah, it can be shown that API is positively correlated with the concentration of carbon monoxide. Thus, the R^2 value is 0.6414, indicating a solid correlation relationship. Thus, the p-value is 2.21×10^{-5}, which indicates that the analysis is statistically significant.

The ground-based data (API) and satellite database of O_3 shown in Fig.6(d) in Sarawak show that API positively correlates with the ozone concentration, with the R^2 value being 0.8517, indicating a solid correlation relationship. Thus, from Table 1, the p-value is $2.02\times10^{-8,}$ which indicates that the analysis is statistically significant. For Sabah, it can be observed that the API value has a positive correlation with the O_3 concentration, with the R^2 value being 0.6537, indicating a solid correlation relationship. Thus, the p-value is 1.6×10^{-5}, which indicates that the analysis is statistically significant. A study done by (Deng et al., 2024) demonstrated that during severe haze pollution, aerosols significantly reduced ground-level photolysis rates of NO_2 and ozone by 22% and 29%, respectively. Elevated aerosol levels in the atmosphere can significantly diminish photochemical radiation, affecting ozone and other secondary pollutants in photochemical processes.

Table 1. The R2 and p values for linear regression analysis in Sarawak and Sabah

Component		**Sarawak**	**Sabah**
$PM_{2.5}$	R^2	0.7478	0.6235
	p ($p < 0.05$)	8.27×10^{-7}	3.47×10^{-5}
SO_2	R^2	0.4154	0.8779
	p ($p < 0.05$)	0.0022	1.19×10^{-9}
CO	R^2	0.0146	0.6414
	p ($p < 0.05$)	0.0027	2.21×10^{-5}
Ozone	R^2	0.8517	0.6537
	p ($p < 0.05$)	2.02×10^{-8}	1.6×10^{-5}

Source: Authors' own data

3.3 Correlation coefficient analysis of air pollution index and GIOVANNI data

Wind speed is considered one of the meteorological factors that can affect the haze phenomenon. The increase in wind speed will cause an effect on haze pollutants as it will dilute the concentration of the pollutants. One of the criteria for high air pollution potential forecasts for most vast urban areas is light and low wind speed, as pollutants tend to pile up in calm conditions or weather when wind speeds are not more than about 2.8 m/s (Dai et al., 2020). In previous studies, it is found that the concentration of air pollutants has an inversely proportional relationship to wind speed (Lu & Fang, 2002; Shenfeld, 1970). Temperatures are considered one of the critical determinants of 'mixing height', which is the depth of the layer of air

https://doi.org/10.24191/mjcet.v7i2.886

closest to the ground within which pollutants and aerosols move, where hot air rises. In contrast, cold air stays close to the ground (Li et al., 2023). This section plots the wind speed (m/s) and temperature (℃) against the API to investigate whether the factors have an inverse relationship and fit the theory given, as shown in Fig. 7. Correlation data analysis was also done to determine whether there was a strong correlation. Moreover, correlation data analysis is also done to determine the strength of the relationship between the two variables.

Using the correlation analysis, the R-value obtained was −0.8750, meaning that wind speed negatively correlates with API. Statistically, a perfect negative correlation is represented by the R-value at −1 (Picardo, 2019). Furthermore, according to the rule of thumb in interpreting the strength of a relationship based on the R-value, when R is larger than 0.7, it means that the relationship between the two variables is considered generally strong (Mindrila & Balentyne, 2013). Wind speed impacts air pollution by affecting the dispersion and accumulation of pollutants. High wind speeds disperse pollutants, improving air quality, while low wind speeds lead to more significant accumulation, resulting in higher pollution levels and poor air quality, especially during stable conditions (Kan, 2022). This is also supported by another study conducted by Dickson and Lawrence, which stated that maximum concentrations of pollutants under calm conditions or at very low wind velocities and that increasing wind velocity improved air quality, particularly in the case of low pollution sources (Padmanabhamurty, 2012). This showed that meteorological conditions, such as wind speed, affect the haze phenomenon.

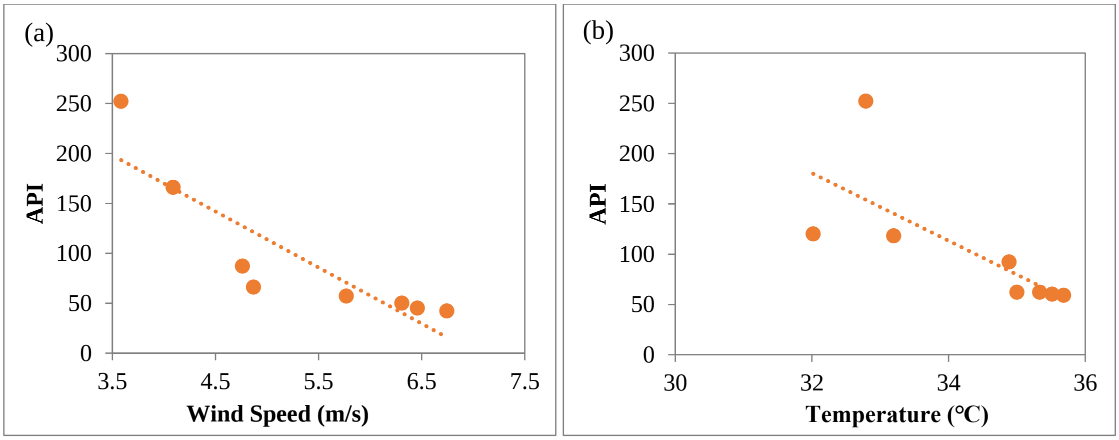

Fig.7. API (Ground-based data) against GIOVANNI (Satellite-based data) of (a) wind speed and (b) temperature in Sabah and Sarawak

Source: Authors' own data

Table 2 shows the correlation coefficient of API versus wind speed and temperature in Sabah and Sarawak. From the correlation data analysis, the relationship between temperature and API value was moderate and negatively correlated, with the R-value being −0.7270. Even though the highest API value is 252, the temperature is 32.79, which is not the lowest value compared to the other temperatures. This data is considered outlier data, which, in this case, when included in the analysis, can reduce a strong relationship to a moderate relationship. Some of the limitations of this analysis are that the temperature range in Malaysia is relatively small as it only ranges from a mild 20 ℃ to 30 ℃ on average throughout the year (MyGoverment, 2016). Because of this condition, comparing the API value with data on drastic temperature differences, such as the difference between the API value during summer and winter, is difficult. Meteorological factors like temperature, humidity, pressure, and wind in the boundary layer significantly

https://doi.org/10.24191/mjcet.v7i2.886

influence pollutant diffusion and accumulation, playing a pivotal role in atmospheric pollution events. This happened because temperature can change the dynamics of the air movement, thus causing warm air to be denser and more buoyant than cooler air. In this case, meteorological conditions such as low temperatures can trap pollutants, increasing the haze pollutants' concentrations (Li et al., 2023).

Table 2. Correlation coefficient of API versus wind speed and temperature in Sabah and Sarawak

Study areas	**Correlation coefficient (R)**
Wind Speed (m/s)	−0.8750
Temperature (°C)	−0.7270

Source: Authors' own data

4. CONCLUSIONS

This study investigated the evaluation of air pollution based on the haze phenomenon by analysing the data collected from ground-based data which is from the Department of Environment (DOE) and from the satellite database which is from GIOVANNI over Malaysia, specifically at Sarawak and Sabah from 2000 to 2019. For the data analysis between the ground-based data (API) and satellite database ($PM_{2.5}$, SO_2, CO, and O_3), linear regression analysis is done where the R^2 values ranged from weak, moderate to strong correlation relationship while all the p values for the analysis are below 0.05 which indicates that the analysis is statistically significant. Based on the correlation analysis, the wind speed has a strong negative correlation with API, with the value of R being −0.8750, indicating that both variables have an inverse relationship. That is, when wind speed increases, the API value decreases. The meteorological condition for temperature also has moderate and negative correlations with API value, resulting in the value of R being −0.7270.

ACKNOWLEDGEMENT

The authors are thankful for the facilities provided by Universiti Teknologi MARA, Cawangan Pulau Pinang, Kampus Permatang Puah, which enabled them to complete this research.

CONFLICT OF INTEREST STATEMENT

The authors agree that this research was conducted without any self-benefits or commercial or financial conflicts and declare the absence of conflicting interests with the funders.

AUTHORS' CONTRIBUTIONS

Zafirah Muhamad Haitamin: Conceptualisation, methodology, formal analysis, investigation, and writing-original draft; **Norhusna Mohamad Nor**: Project administration, conceptualisation, data analysis and validation, supervision, writing- review and editing.

https://doi.org/10.24191/mjcet.v7i2.886

REFERENCES

Afroz, R., Hassan, M. N., & Ibrahim, N. A. (2003). Review of air pollution and health impacts in Malaysia. *Environmental Research, 92*(2), 71–77. https://doi.org/10.1016/S0013-9351(02)00059-2

Akhtar, R., & Palagiano, C. (Eds.). (2018). *Climate change and air pollution: The impact on human health in developed and developing countries*. Springer. https://doi.org/10.1007/978-3-319-61346-8

Al-Alola, S. S., Alkadi, I. I., Alogayell, H. M., Mohamed, S. A., & Ismail, I. Y. (2022). Air quality estimation using remote sensing and GIS-spatial technologies along Al-Shamal train pathway, Al-Qurayyat City in Saudi Arabia. *Environmental and Sustainability Indicators, 15,* Article 100184. https://doi.org/10.1016/j.indic.2022.100184

Australian Government; Department of the Environment and Energy (2019). *Particulate Matter (PM_{10} and $PM_{2.5}$). National Pollutant Inventory*. https://www.dcceew.gov.au/environment/protection/npi/substances/fact-sheets/particulate-matter-pm10-and-pm25

Binti Udin, Q. A., Binti Kasim, A., Binti Rassa, H., Ul, A., Binti, H., & Nasir, A. (2021). Factors that contribute to air pollution in Malaysia. *Malaysian Journal of Business and Economics, 8(*2), 43–58. https://doi.org/10.51200/mjbe.vi.3662

Christopherson, J. B., Ramaseri, C., & Quanbeck, J. Q. (2019). *2019 Joint Agency Commercial Imagery Evaluation - Land Remote Sensing Satellite Compendium*. *U.S. Geological Survey Circular 1455*. https://doi.org/10.3133/cir1455

Dai, Z., Liu, D., Yu, K., Cao, L., & Jiang, Y. (2020). Meteorological variables and synoptic patterns associated with air pollutions in Eastern China during 2013–2018. *International Journal of Environmental Research and Public Health, 17*(7), Article 2528. https://doi.org/10.3390/ijerph17072528

Deng, T., Ouyang, S., He, G., Zhang, X., Leung, J. C. H., Chen, X., Wang, Q., Zhang, Z., Zou, Y., Mai, B., Liu, L., & Deng, X. (2024). Impact of aerosol actinic radiative effect on ozone during haze pollution in the Pearl River Delta region. *Atmospheric Environment, 332*, Article 120610. https://doi.org/10.1016/j.atmosenv.2024.120610

Department of Environment Malaysia. (2016a). *Haze air pollution phenomena. In Department of Environment, Ministry of Natural Resources & Environment (pp. 1–6). Jabatan Alam Sekitar Malaysia, Department of Environment.* https://www.doe.gov.my/en/2021/10/26/haze-air-pollution-phenomenon/

Department of Environment Malaysia. (2016b). *Final Report: Review of Air Pollutant Index (API)*. https://enviro2.doe.gov.my/ekmc/wp-content/uploads/2016/11/API-FINAL-REPORT.pdf

Julfikar, S.K., Ahamed, S., Rehena, Z. (2021). Air quality prediction using regression models. In A. Choudhary, A.P. Agrawal, R. Logeswaran, &B. Unhelkar (Eds.), *Applications of Artificial Intelligence and Machine Learning. Lecture Notes in Electrical Engineering* (vol 778). https://doi.org/10.1007/978-981-16-3067-5_19

Kan, H. (2022). World Health Organization air quality guidelines 2021: Implication for air pollution control and climate goal in China. *Chinese Medical Journal, 13*(5), 513–515. https://doi.org/10.1097/CM9.0000000000002014

Latif, M. T., & Hamzah, W. P. (2016). Air Quality & Haze Episodes in Malaysia (Issue May). Akademi Sains Malaysia.

https://doi.org/10.24191/mjcet.v7i2.886

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., & Jaafar, A. B. (2018). Impact of regional haze towards air quality in Malaysia: A review. *Atmospheric Environment, 177,* 28–44. https://doi.org/10.1016/j.atmosenv.2018.01.002

Latiffah, N., Rani, A., Azid, A., Khalit, S. I., Juahir, H., & Samsudin, M. S. (2018). Air pollution index trend analysis in Malaysia, 2010–15. J*ournal of Environmental Study, 27*(2), 801–807. https://doi.org/10.15244/pjoes/75964

Li, S., Li, X., Deng, Z., Xia, X., Ren, G., An, D., Ayikan, M., & Zhong, Y. (2023). Characteristics of atmospheric boundary layer and its relation with $PM_{2.5}$ during winter in Shihezi, an Oasis city in Northwest China. *Atmospheric Pollution Research, 14*(11), Article 101902. https://doi.org/10.1016/j.apr.2023.101902

Lu, H. C., & Fang, G. C. (2002). Estimating the frequency distributions of PM_{10} and $PM_{2.5}$ by the statistics of wind speed at Sha-Lu, Taiwan. *Science of the Total Environment, 298*(1–3), 119–130. https://doi.org/10.1016/S0048-9697(02)00164-X

MyGoverment (2016). Malaysia Information - Climate https://malaysia.gov.my/portal/content/144

FMcLeod, S. (2019, May). What a p-value tells you about statistical significance. *Simply Psychology*. https://www.simplypsychology.org/p-value.html

Mindrila, D., & Balentyne, P. (2013). *Scatterplots and Correlation. The Basic Practice of Statistics* (6th ed., Vol. 3, Issue 1, pp. 73–90). W. H. Freeman and Company. https://doi.org/10.3109/08958379109145275

Muzammil, M., & Hanan, Z. (2017). Application of remote sensing instruments in air quality monitoring. *Pertanika Journal of Scholarly Research Review, 3*, 93–112.

New Straits Times. (2019, September 22). Haze crisis: Still no breather for much of Malaysia. *New Straits Times*. https://www.nst.com.my/news/nation/2019/09/523413/haze-crisis-still-no-breather-much-malaysia

Othman, J., Sahani, M., Mahmud, M., & Sheikh Ahmad, M. K. (2014). Transboundary smoke haze pollution in Malaysia: Inpatient health impacts and economic valuation. *Environmental Pollution, 189*, 194–201. https://doi.org/10.1016/j.envpol.2014.03.010

Padmanabhamurty, B. (2012). The role of wind in pollution dispersion. J*ournal of the Air Pollution Control Association*, *25*(9), 956–957. https://doi.org/10.1080/00022470.1975.10468120

Picardo, E. (2019). Negative Correlation: How it Works, Examples and FAQ. https://www.investopedia.com/terms/n/negative-correlation.asp

Prados, A. I., Leptoukh, G., Lynnes, C., Johnson, J., Rui, H., Chen, A., & Husar, R. B. (2010). Access, visualization, and interoperability of air quality remote sensing data sets via the Giovanni Online Tool. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 3*(3), 359–370. https://doi.org/10.1109/JSTARS.2010.2047940

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

https://doi.org/10.24191/mjcet.v7i2.886

Redzuan, S. N., Noor, N. M., Rahim, N. A. A. A., Jafri, I. A. M., Baidrulhisham, S. E., Ul-Saufie, A. Z., Sandu, A. V., Vizureanu, P., Zainol, M. R. R. M. A., & Deák, G. (2023). Characteristics of PM_{10} level during haze events in Malaysia based on quantile regression method. *Atmosphere, 14*(2), 407. https://doi.org/10.3390/atmos14020407

Shafii, N. Z., Saudi, A. S. M., Mahmud, M., & Rizman, Z. I. (2017). Spatial assessment on ambient air quality status: a case study in Klang, Selangor. *Journal of Fundamental and Applied Sciences, 9*(4S), 9643–977. https://doi.org/10.4314/jfas.v9i4s.58

Shenfeld, L. (1970). Meteorological aspects of air pollution control. *Atmosphere, 8*(1), 3–13. https://doi.org/10.1080/00046973.1970.9676578

Vaughn, B. (2019, September 16). Indonesia haze: Why do forests keep burning? *BBC News*. https://www.bbc.com/news/world-asia-34265922