VOL 5, NO.2 (2025) 325-355
https://jagst.utm.my

JAGST
Journal of Advanced Geospatial Science & Technology

VALIDATION OF SATELLITE-DERIVED AEROSOL OPTICAL DEPTH USING LONG-TERM SUNPHOTOMETER DATA IN MALAYSIA

Karthikraajah Nadarajah[1] and Mohd Nadzri Md Reba[1,2]
[1]Department of Geoinformation, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia
[2]Geoscience and Digital Earth Centre (INSTeG), Research Institute for Sustainable Environment (RISE), Universiti Teknologi Malaysia, 8310 UTM Johor Bahru, Malaysia

*Corresponding author: nadzri@utm.my

Abstract – Aerosol Optical Depth (AOD) is a dimensionless measure of the extinction of solar radiation by aerosols in the vertical column of the atmosphere and it is widely used as an indicator of air quality. It can be measured from space using satellite sensors such as MODIS, which observes the total atmospheric column from the top of the atmosphere to the surface, or from the ground using AERONET sunphotometers, which measure the attenuation of direct solar radiation from the ground to the atmospheric boundary layer. In Malaysia, satellite-derived AOD retrieval is often challenging due to frequent cloud cover, high humidity, surface heterogeneity, and spectral mismatch between MODIS wavelengths (470 nm and 550 nm) and AERONET wavelengths (440 nm and 500 nm). These differences in measurement time, spatial scale, and wavelength often cause MODIS to overestimate AOD, especially during high aerosol loading. AOD is closely linked to fine particulate matter (PM2.5) making it important for assessing air quality impacts on health and the environment. Globally, accurate AOD information supports climate and haze studies, while in Malaysia, it is essential for monitoring seasonal haze and urban pollution. This study applied spatial, temporal, and spectral correction techniques to minimise discrepancies between MODIS MAIAC AOD and AERONET Level 2.0 data from USM Penang and Kuching, Malaysia improving the reliability of satellite AOD for air quality monitoring in tropical conditions. Long-term MODIS MAIAC data (MCD19A2.061) from 2012 to 2024 for Penang and from 2012 to 2020 for Kuching were used, along with AERONET data from USM and Kuching stations. Spatial correction was applied using regression on a 3×3 MODIS pixel grid. Spectral correction used interpolation to estimate AOD at 470 nm and 550 nm. Temporal correction was performed by filtering AERONET data between 1000h until 1100h to match the MODIS Terra overpass. The corrected MODIS AOD values showed improved agreement with AERONET. In Penang, the 550 nm result showed R = 0.72 and RMSE = 0.08, while in Kuching, R = 0.97 and RMSE = 0.15. Seasonal and frequency analysis revealed that MODIS tends to overestimate AOD, especially during haze events. This study demonstrates that calibrated satellite AOD data using long-term observations can support more reliable air quality monitoring in urban Malaysian settings.

Keywords *– Air quality, MODIS AOD, AERONET, Satellite data validation*

Article History: Received 1 August 2025, Accepted 28 August 2025, Published 31 August 2025

How to cite: Nadarajah, K. R. and Md Reba, M. N. (2025). Validation Of Satellite Derived Aerosol Optical Depth Data Using Continuous Sunphotometer Measurement. Journal of Advanced Geospatial Science & Technology, 6(1), 325–355; DOI: https://doi.org/10.11113/jagst.v5n2.118

1.0 Introduction

Aerosol Optical Depth (AOD) is a key parameter for understanding atmospheric aerosol concentrations and their influence on air quality, climate and human health. It represents the amount of aerosol particles in the atmospheric column that block or scatter incoming solar radiation. These particles may originate from various natural and anthropogenic sources, including dust, smoke, industrial emissions and biomass burning (Kokhanovsky et al., 2007; Antuña-Marrero et al., 2018). A higher AOD value indicates greater aerosol loading, which can significantly impact visibility, atmospheric heating and human exposure to fine particulate matter. Monitoring AOD is particularly important in Malaysia due to frequent haze events caused by a combination of local emissions and transboundary pollution from large-scale forest and peatland fires in Indonesia. These events contribute to persistent episodes of poor air quality, especially during the southwest monsoon season, and have been linked to adverse environmental, health and economic impacts (Salinas, Chew, Miettinen, et al., 2013; Kanniah et al., 2014).

AOD can be measured using both satellite remote sensing and ground-based observation networks. Satellite measurements, such as the Moderate Resolution Imaging Spectroradiometer (MODIS), offer broad spatial coverage at 1x1 km MODIS spatial resolution and daily global sampling, making them suitable for large-scale aerosol monitoring. MODIS onboard NASA's Terra and Aqua satellites has been widely used in Malaysia for AOD retrieval (Kanniah et al., 2014). MODIS products provide valuable data for air quality analysis due to their global coverage and long-term data availability (Tan et al., 2015; Song et al., 2024). However, retrieval accuracy in tropical regions is often affected by factors such as cloud contamination, high surface reflectance and complex land–atmosphere interactions (Hee et al., 2014). These factors can cause MODIS to overestimate aerosol loading, particularly during haze episodes (Kanniah et al., 2016; Li et al., 2005). Ground-based measurements, especially those from the Aerosol Robotic Network (AERONET), are widely recognised for their accuracy and standardisation in AOD retrieval. AERONET sun photometers measure direct solar irradiance and use radiative transfer models to estimate AOD at specific wavelengths. The system operates with a temporal resolution of approximately 15 minutes, providing frequent observations throughout the day. Despite its reliability, AERONET coverage in Malaysia is limited, with only 11 stations established nationwide. Among these, the USM Penang and Kuching Meteorological Station sites have the most continuous and long-term data records, making them the primary ground-based references for this study (Tripathi et al., 2005; Holben et al., 1998). Previous studies in Malaysia showed that MODIS AOD had moderate correlation around 0.55 with AERONET data, but overestimation was observed especially in high aerosol regions like Penang and Kuching. Most of these studies used only one year of data and combined both Terra and Aqua satellite products. The Angstrom exponent method was commonly used to estimate AERONET AOD at 440, 675, 870 nm to match the MODIS bands at 470 nm and 550 nm (Hee et al., 2014; Kanniah et al., 2014). Their results indicated that MODIS tended to overestimate aerosol concentration, particularly during haze periods (June to September). The presence of temporal variation and differences in overpass time also affected the quality of AOD retrieval.

AOD measurement in Malaysia presents several challenges. Overestimation of MODIS AOD has been reported by (Hee et al. 2014) in Penang and Kuching, where high aerosol loading and complex surface reflectance conditions adversely affected retrieval accuracy. Data availability is also constrained by frequent cloud cover, which diminishes the number of valid retrievals, and the presence of invalid values (−9999) that necessitate filtering prior to analysis. In numerous instances, daily matchups between MODIS and AERONET do not satisfy the recommended validation threshold of $N \geq 30$, particularly when a high-quality AERONET station is not situated in close proximity to the study area. Previous validation studies primarily employed parametric statistical methods such as linear regression, correlation coefficient, and root mean square error. While these methods offer a general assessment, they are limited in their ability to capture non-linear relationships and may not fully account for spatial and temporal discrepancies between datasets. Accurate and reliable Aerosol Optical Depth (AOD) data are crucial for effective air quality monitoring and comprehensive haze event assessment in Malaysia. While AERONET offers highly precise AOD measurements, its spatial coverage is severely restricted, with only two operational stations in Malaysia during the study period. The limited number of viable data points hindered comprehensive validation efforts, as other stations experienced short-term operations, frequent downtime, significant data gaps, or failed to satisfy the minimum data quality criteria for matchup analysis. This sparse distribution impedes the ability to monitor aerosols across broader regions, particularly in urban and coastal areas where aerosol concentrations can exhibit substantial variability over short distances (Salinas, Chew, Mohamad, et al., 2013).

Even though MODIS proved the synoptic and long-term data capability, its AOD retrieval in tropical environments are significantly affected by cloud cover, high surface reflectance, and high humidity, often leading to overestimation during haze events (Hee et al., 2014; Kanniah et al., 2014). Further errors arise from spectral mismatch, since MODIS measures AOD at fixed wavelengths such as 470 nm and 550 nm, while AERONET records at different bands such as 440 nm and 500 nm. Comparing the MODIS product with the ground measurement is not a straightforward task. Temporal difference is evident when MODIS Terra and Aqua capture the AOD at the observation point only once per day (around 1030h and 1330h respectively) to compare with the ground AERONET which collects continuous AOD for every 15 minutes. Without spectral interpolation, direct comparisons can introduce wavelength-dependent biases due to differences in the spectral response of MODIS and AERONET instruments (Antuña-Marrero et al., 2018; Tripathi et al., 2005). Temporal mismatch is also present because MODIS Terra captures data only once per day around 1030h, whereas AERONET collects continuous observations. Without selecting AERONET measurements within the satellite overpass window, diurnal variations in aerosol concentration introduce further bias (Kanniah et al., 2016). Spatial mismatch is another issue, as MODIS data represent pixel-averaged values at 1-km resolution, while AERONET measures at a single point location. Without pixel-level regression or spatial adjustment, these scale differences can reduce correlation (Ichoku et al., 2002; Kokhanovsky et al., 2007).

Most past studies in Malaysia have focused on basic validation techniques or seasonal analysis without combining spatial, spectral and temporal corrections in a single framework. Few have investigated how such integrated corrections could improve MODIS AOD accuracy in multiple haze-affected locations, particularly when analysing both Penang and Kuching with dual wavelengths. This lack of comprehensive correction approaches limits the usefulness of satellite-derived AOD for precise local air quality monitoring. The aim of this study is to improve the accuracy of MODIS MAIAC AOD retrievals for Malaysian cities by applying integrated spatial, temporal, and spectral correction techniques using AERONET Level 2.0 data as reference. The objectives of this study are; (1) to calibrate the spatial and radiometric differences between MODIS and AERONET derived AOD by applying spatial interpolation and spectral regression techniques; (2) to compute the accuracy of MODIS AOD through temporal collocation with AERONET observations using different parametric analyses; and (3) to analyse the validation results in terms of accuracy and statistical performance at collocated spatial and radiometric resolutions. This study validated MODIS AOD values at 470 nm and 550 nm against AERONET observations from USM Penang and Kuching by implementing spatial correction through interpolation to reduce pixel-to-point mismatch, temporal correction by selecting AERONET measurements within a one-hour window of the MODIS Terra overpass, and spectral correction using regression-based interpolation to match wavelengths. The analysis was conducted using a long-term dataset covering 12 years for USM Penang and 8 years for Kuching and applied a linear validation model to examine spatial distribution, temporal trends, and radiometric relationships of AOD using cloud-based platforms and GIS tools.

This research improves our understanding and monitoring of aerosol distribution in Malaysia by generating corrected MODIS AOD datasets. These datasets address spatial, spectral, and temporal discrepancies, leading to reduced biases common in tropical environments. The result is more dependable estimates for use in air quality management, haze assessment, and environmental health studies. The methodology developed here is also applicable to other haze-prone areas with limited ground-based observations. Furthermore, this study offers crucial data for drought analysis by examining the impact of aerosols on the hydrological cycle, which is a factor that either inhibits or promotes precipitation. Improved Aerosol Optical Depth (AOD) products offer significant benefits for decision-makers and environmental agencies, enabling the creation of early warning systems, informing mitigation strategies, and enhancing public health protection during haze episodes. This research aligns with the United Nations Sustainable Development Goals, specifically Goal 3 (good health and well-being) and Goal 11 (sustainable cities and communities) (Nations, 2024), thereby supporting sustainable development. Additionally, the integration of long-term satellite and ground-based measurements with correction techniques advances remote sensing applications in tropical atmospheric studies and bolsters Malaysia's capacity for data-driven environmental management.

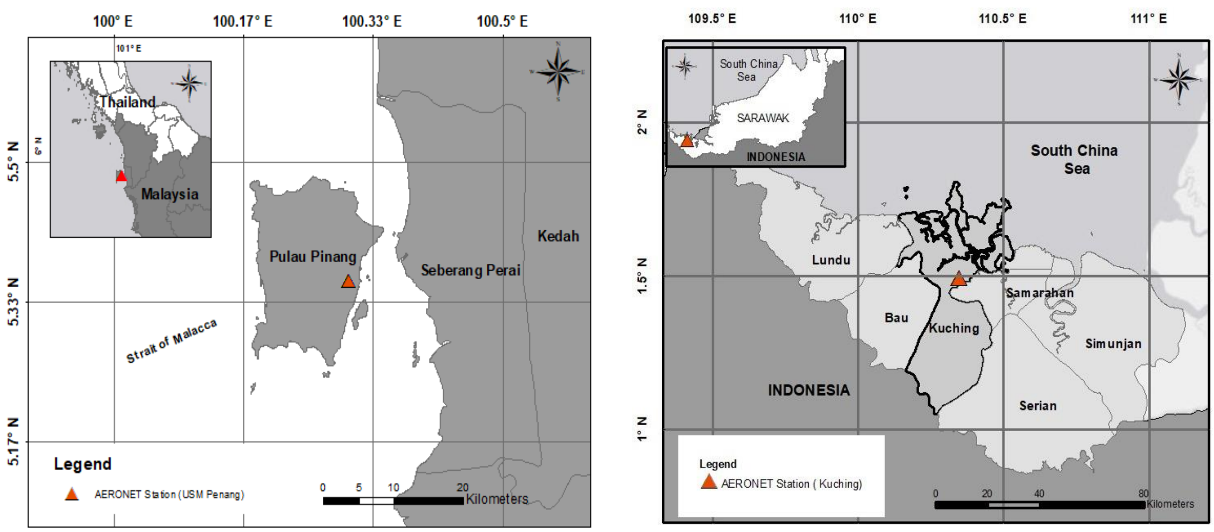
2.0 Materials and Methods

2.1 Study Area

This study was conducted in two locations in Malaysia, namely Universiti Sains Malaysia (USM) in Penang and Kuching in Sarawak. Figures 1(a) and 1(b) show the location of the two study areas respectively. USM Penang is located in the northeastern part of Penang Island, in the northwestern region of Peninsular Malaysia. Geographically, the island is bordered by the Strait of Malacca to the west and north, and the mainland state of Kedah to the east across a narrow channel. The climate is tropical, with temperatures generally ranging between 24 °C and 33 °C throughout the year. During the dry season, particularly from June to September, daytime temperatures can reach 32 °C to 35 °C, with lower rainfall and reduced relative humidity. The wet season occurs from November to March, influenced by the northeast monsoon, with higher rainfall and average humidity above 80 %. The mean annual rainfall exceeds 2,400 mm. Prevailing winds during the dry season are from the southwest, often transporting the transboundary smoke from Sumatra. The topography of Penang Island consists of coastal lowlands and a central hilly region, with elevations ranging from sea level to approximately 833 m at Penang Hill. The land use in the northeastern area, where USM is located, is predominantly urban, with dense residential, commercial, and industrial activities. Penang has experienced haze episodes, when PM2.5 and AOD levels rose sharply due to regional biomass burning (Kanniah et al., 2014).

Kuching, the capital city of Sarawak, is situated in the western part of Borneo Island along the banks of the Sarawak River and faces the South China Sea. The city experiences a tropical rainforest climate with minimal temperature variation throughout the year. Average temperatures range between 23 °C and 32 °C, with peak temperatures of around 34 °C during the dry season from June to September. The wet season occurs between November and March, bringing heavy rainfall exceeding 4,000 mm annually, with humidity levels above 85 %. Winds during the dry season are often light to moderate, but haze episodes are common when winds carry smoke from large-scale forest and peatland fires in Kalimantan, Indonesia. The topography of Kuching is generally low-lying, with elevations ranging from sea level to about 90 m, and surrounding areas consist of flat floodplains with small hill ranges inland. Land use in Kuching is characterised by a mix of urban development, coastal areas, and surrounding green zones with forest cover. Severe haze events have been documented in Kuching, especially in 2015 when PM2.5 concentrations exceeded hazardous levels and visibility dropped significantly due to transboundary smoke (Salinas, Chew, Mohamad, et al., 2013; Kanniah et al., 2014).

Both locations were selected for this study because they are high-aerosol regions with available long-term AERONET Level 2.0 datasets. The dataset for USM Penang covers the period from 2012 to 2024, while the Kuching dataset spans from 2012 to 2020. The contrasting geographical settings, climatic patterns, and pollution sources at these sites provided an ideal basis for assessing MODIS AOD performance and seasonal variability under different tropical environments.



(a) (b)

Figure 1. Map of study area (a) Penang and (b) Kuching, Sarawak

2.2 Methodology Overview

The methodology for this study was structured into five main phases to ensure systematic processing and analysis of aerosol data. The first phase, data acquisition, involved obtaining MODIS MAIAC AOD at 1 km resolution and AERONET Level 2.0 measurements from USM Penang and Kuching. The second phase, data pre-processing, included quality screening, filtering, and temporal–spatial matching between satellite and ground datasets. The third phase, data processing, applied spatial, spectral, and temporal corrections to minimise bias between MODIS and AERONET observations. The fourth phase, data analysis, statistically evaluated the corrected MODIS AOD against AERONET values, with performance assessment by season. The fifth phase, data visualisation, produced maps, plots, and tables for interpretation and reporting. These phases are illustrated in the methodological framework shown in Figure 2.

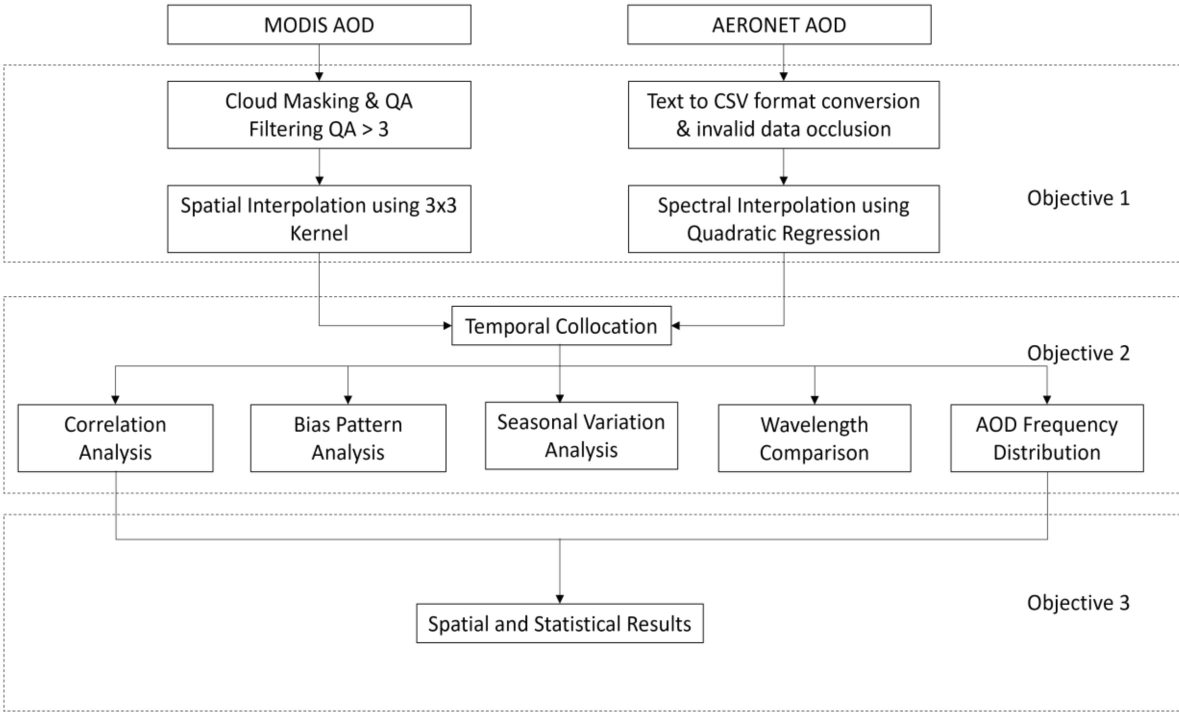


Figure 2. Methodology flowchart explains the process of evaluating the AOD MODIS product with the respective objectives.

2.3 Data Acquisition

This study used both satellite-based and ground-based Aerosol Optical Depth (AOD) datasets to evaluate and improve the accuracy of satellite AOD retrievals over USM Penang and Kuching, Sarawak. The satellite dataset was obtained from the MODIS MAIAC product, while the ground dataset from the AERONET Level 2.0 archive.

(a) MODIS AOD Data

The study utilized data from MODIS (Moderate Resolution Imaging Spectroradiometer), a multispectral sensor on NASA's Terra satellite. Specifically, the MCD19A2.061 Multi-Angle Implementation of Atmospheric Correction (MAIAC) AOD product was used, which offers enhanced aerosol retrieval over both land and water. This dataset provides AOD values at 470 nm and 550 nm wavelengths, derived through an advanced atmospheric correction algorithm. The Level 2 MAIAC AOD data includes information crucial for geolocation, radiometric, and atmospheric corrections. Data preparation involved cloud masking and aerosol quality filtering using QA flags, retaining only pixels with QA values of three or greater. MODIS MAIAC data has a 1 km × 1 km spatial resolution and is collected daily for each location during the Terra satellite's overpass, around 1030h local time. A 3x3 pixel grid, centered on the AERONET station coordinates, was used to estimate the collocated AOD, minimizing spatial mismatch between satellite pixels and ground observations. The processed AOD values were then exported as GeoTIFF for mapping and CSV for statistical analysis. For USM Penang, data was collected from January 2012 to March 2024, while for Kuching, the period covered January 2012 to December 2020. Both datasets encompass significant haze events, particularly those in 2015, which are crucial for regional aerosol studies. MODIS AOD data was accessed and processed via the Google Earth Engine (GEE) platform. The data was clipped to the study area using shapefiles, cloud-affected pixels were removed, and monthly mean composites were generated.

(b) AERONET AOD Data

AERONET (AErosol RObotic NETwork) is a federation of ground-based remote sensing aerosol networks established by NASA and LOA-PHOTONS (CNRS) in collaboration with national agencies, universities and research partners worldwide. It provides a long-term, standardised and publicly accessible database of aerosol optical, microphysical and radiative properties for satellite validation and atmospheric studies. Measurements are performed using Cimel CE-318 automatic sun/sky radiometers, which record direct solar irradiance and sky radiances at multiple wavelengths. The instruments operate autonomously, track the sun, and are regularly calibrated against reference instruments at sites such as Mauna Loa Observatory to ensure consistency and accuracy across the network. AERONET products are distributed in three quality levels, namely Level 1.0 (unscreened), Level 1.5 (cloud-screened), and Level 2.0 (cloud-screened and quality-assured), with Level 2.0 being the highest-quality dataset available (Holben et al., 1998).

For this study used Level 2.0 AOD data, which is cloud-screened and quality-assured. The instruments measure the attenuation of direct solar radiation at multiple wavelengths, allowing the calculation of spectral AOD. For this study, data at 440 nm and 500 nm was used for spectral interpolation to match the MODIS 470 nm and 550 nm bands. AERONET data for USM Penang, covering January 2012 to March 2024, and for Kuching, spanning January 2012 to December 2020, were utilized in this study. Both datasets captured significant haze events, particularly during the dry seasons of 2013, 2015, and 2019, which were linked to transboundary biomass burning in Indonesia. The original data, downloaded in TXT format from the official NASA AERONET website (NASA, 2025), underwent quality control. This process involved the removal of missing values and invalid records (marked as -999). Subsequently, the cleaned datasets were converted to CSV format for further analysis, including spectral interpolation and temporal matching with MODIS overpass times.

2.4 MODIS AOD spatial and temporal collocation and spectral matching

To ensure accurate validation, both MODIS and AERONET AOD datasets underwent pre-processing, including spatial collocation. Since MODIS AOD represented an average over a 1km x 1km pixel and AERONET provided point-based measurements at the exact ground station, a direct comparison could lead to spatial mismatch, particularly in urban coastal areas with mixed land cover. To mitigate this, a 3x3 pixel window surrounding the AERONET station coordinates was extracted from the MODIS MAIAC product. If valid pixels were present within this window, a least squares regression model was applied, following the methodology by Ichoku et al. (2002)., using the equation

$$z = aX + bY + c \qquad (3.1)$$

where z represents the MODIS AOD [unitless], and X and Y are the longitude [degree] and latitude [degree] of each pixel. The coefficients a, b, and c were calculated. The MODIS AOD at the AERONET station was then estimated using

$$\bar{z} = aX_0 + bY_0 + c \qquad (3.2)$$

where X_0 and Y_0 represent the ground station coordinates. This spatial correction is essential for accurately determining local aerosol gradients and minimizing bias when comparing gridded data with single-point observations. Du et al. (2025) recently endorsed this approach, demonstrating that spatial matching enhances satellite-ground agreement and reduces spatial bias in AOD validation studies.

AERONET data which is originally recorded in Coordinated Universal Time (UTC), underwent time normalization to Malaysian local time (GMT+8). Data recorded within a 1-hour window centered around 1030h were chosen to align with the typical MODIS Terra overpass time. This temporal collocation aimed to minimize atmospheric variation and guarantee that both satellite and ground-based observations represented nearly the same time. Given that AERONET data is recorded every 15 minutes (Holben et al., 1998; Ichoku et al., 2002), the median value was used to represent daily Aerosol Optical Depth (AOD) when multiple readings were available within the chosen window. This temporal filtering, implemented using a fixed kernel window, served to reduce diurnal mismatch and retain data only during satellite-representative hours, a crucial step for accurate validation in recent aerosol studies (Du et al., 2025).

Spectral matching was crucial due to the differing wavelengths at which MODIS and AERONET report AOD measurements. MODIS provides AOD at 470 nm and 550 nm, whereas AERONET provides data at 440 nm, 500 nm, and 675 nm. To estimate the true AOD at the interpolated band, spectral interpolation was performed on the AERONET data, utilizing both linear and quadratic methods. Linear interpolation was calculated using

$$AOD(\lambda) = m\lambda + c \tag{3.3}$$

where λ is the desired wavelength, and m and c are the slope and intercept derived from two nearby bands. Quadratic interpolation was also used to improve accuracy, based on the equation:

$$AOD(\lambda) = a\lambda^2 + b\lambda + c \tag{3.4}$$

where a, b, and c were calculated from three AERONET wavelengths. This correction was implemented using the spectral regression approach, a methodology established by Tripathi et al. (2005) and frequently employed in AOD research. For accurate validation of merged or long-term datasets, especially when comparing MODIS–AERONET data, precise spectral alignment is essential. This aligns with recent findings by Gupta et al. (2024). This spectral correction was anticipated to reduce wavelength mismatch bias and enhance consistency between both datasets. Subsequently, the processed AOD values were exported as GeoTIFF for mapping and CSV for statistical analysis.

3.0 Results and Analysis

3.1 Comparison between AOD of MODIS and AERONET

From the conducted analysis, this study has discovered several significant results that show the effectiveness of the applied correction methods. Table 1 shows the statistical metrics of the linear regression between the AOD derived by MODIS and AERONET respectively. Table 1 presents the quadratic interpolation results for USM Penang, which show lower accuracy compared to Kuching. At 550 nm, a moderate correlation was observed, with r = 0.7169, r^2 = 0.5139, and an RMSE of 0.0757. The bias of 0.0247 suggests a slight overestimation of MODIS AOD when compared to AERONET observations. Performance declined at 470 nm, with r = 0.7015, r^2 = 0.4920, and a higher bias of 0.0439. The slope also deviated further from 1, indicating a reduced agreement between MODIS and AERONET. The R^2 value, approximately 50%, defined a moderate positive relationship, suggesting that about half of the variance in MODIS can be explained by AERONET. The MAE was slightly lower than the RMSE, implying that extremely large errors had less influence on the RMSE's skewness. This could be attributed to factors such as increased cloud cover, surface reflectance, or urban influence in Penang. Overall, MODIS AOD showed a slight overestimation with a positive bias, and the 550 nm wavelength AOD data yielded better results than the 470 nm data, even in Penang.

Table 1. Statistical Comparison Between MODIS and AERONET (470nm & 550nm) – USM, Penang and Sarawak

Location	Wavelength	n	R	R^2	RMSE	NRME	MAE	Bias	Unbias
Penang	470 nm	146	0.7015	0.4920	0.1110	0.1637	0.0971	0.0439	0.1020
	550 nm	146	0.7169	0.5139	0.0757	0.1822	0.0637	0.0247	0.0716
Kuching	470 nm	82	0.9746	0.9499	0.1955	0.4458	0.1553	0.1037	0.1658
	550 nm	82	0.9728	0.9463	0.1551	0.4233	0.1092	0.0370	0.1506

For Kuching, quadratic interpolation regression results show that at 550 nm, MODIS AOD demonstrates high accuracy and strong positive correlation (r = 0.9728, r^2 = 0.9463) with a low RMSE of 0.1551. This indicates that approximately 94% of the variance in MODIS AOD is explained by AERONET AOD, highlighting significant agreement between the two datasets. The small bias of 0.0370 suggests a slight overestimation of AOD by MODIS, and the regression slope close to 1 confirms good agreement after correction. At 470 nm, the correlation is slightly stronger (r = 0.9746), but the RMSE is higher at 0.1955, and the bias increases to 0.1037. The larger magnitude of error at this wavelength is potentially due to the generally higher absolute AOD in Kuching. While the correlation remains good, the increased errors suggest that 550 nm is a more stable and reliable wavelength for MODIS AOD validation. The scatter plots in Figure 3 provide a crucial visual context for understanding the relationship between MODIS and AERONET AOD data, allowing for an assessment of data distribution and spread that numerical summaries alone cannot fully capture. In Penang, the Figure 3(a) confirm the moderate correlation where the data points at 470 nm is widely scattered around the 1:1 line. This would suggest that for higher AOD values, the MODIS estimate is generally lower than the AERONET measurements and visually demonstrated by the regression slope (0.65). This suggests that the retrieval algorithm struggles to accurately capture high aerosol loads in a complex urban environment. Figure 3(b) shows better regression for data points in Penang at 550-nm band. The data points are a bit more clustered around the regression line, confirming a better correlation (R=0.72). The regression line's slope is 0.67, which, while still below 1, is closer to the ideal. This improvement at 550 nm is likely due to lower reflectance noise and better atmospheric penetration at this wavelength. Figure 3(c) is AOD comparison for Kuching at 470 nm where the data points are tightly clustered around both the regression line and the 1:1 line, visually confirming the strong correlation (R=0.97). The regression slope of 0.96, which is very close to 1, demonstrates excellent agreement across the entire measurement range. The plot for 550 nm (Figure 3d) reinforces this strong agreement. The data points are similarly tight, and the regression slope of 0.98 is even closer to 1, indicating that the 550 nm channel provides a more accurate and stable retrieval. The intercept being very close to zero further shows a low bias at low AOD values. Overall, Kuching shows a much stronger agreement than Penang, which is likely due to its more homogeneous and darker surface (tropical rainforest), providing a clearer background for the sensor.

3.1 Impact of air-quality episodes and seasonal variation towards monthly AOD measurement

This section presents the monthly trends of MODIS and AERONET AOD values based on the full collocated dataset. The analysis included 146 matched days for Penang and 82 matched days for Kuching, covering multiple years. The trend was analysed for both 470 nm and 550 nm wavelengths. Monthly mean AOD values were calculated and plotted, with error bars indicating the standard deviation. These error bars reflect how much the AOD values varied within each month. Larger bars showed high variability, often caused by haze events or unstable weather. Smaller bars indicated more consistent aerosol levels. Figure 4 (a) and Figure 4 (b) display the AOD trend for Penang at 470 nm and 550 nm respectively. For both wavelengths, AOD values remained relatively low from January to May. A sharp increase was observed from June onwards, with a clear peak in August and September. This pattern matched the southwest monsoon season, which is commonly associated with transboundary haze events due to forest fires in Sumatra and Kalimantan. AOD values started to drop in October and returned to low levels by December.

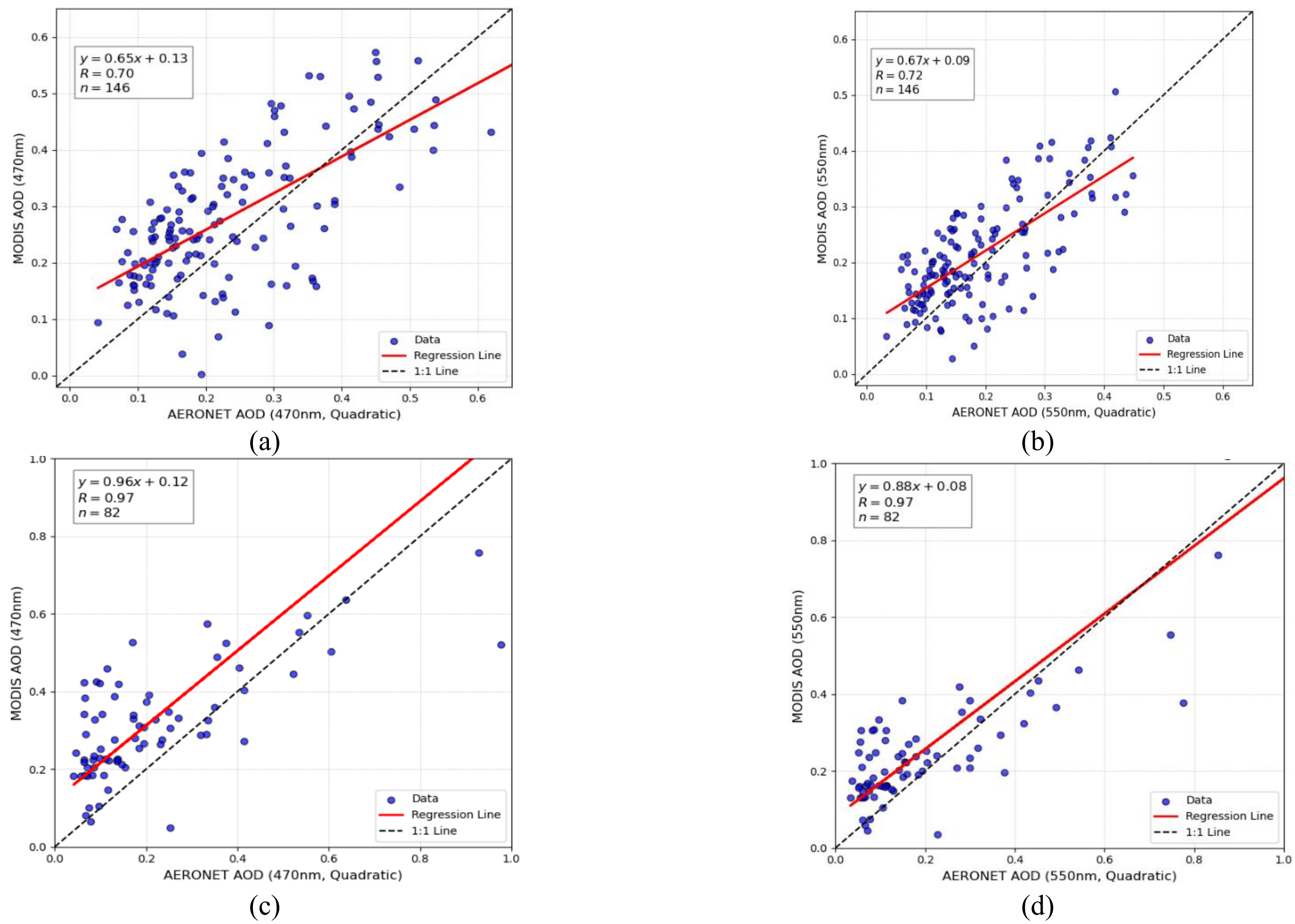


Figure 3. Scatter plot comparing MODIS and AERONET AOD for Penang (a) 470 nm and (b) 550 nm and Kuching at (c) 470 nm and (d) 550 nm, respectively.

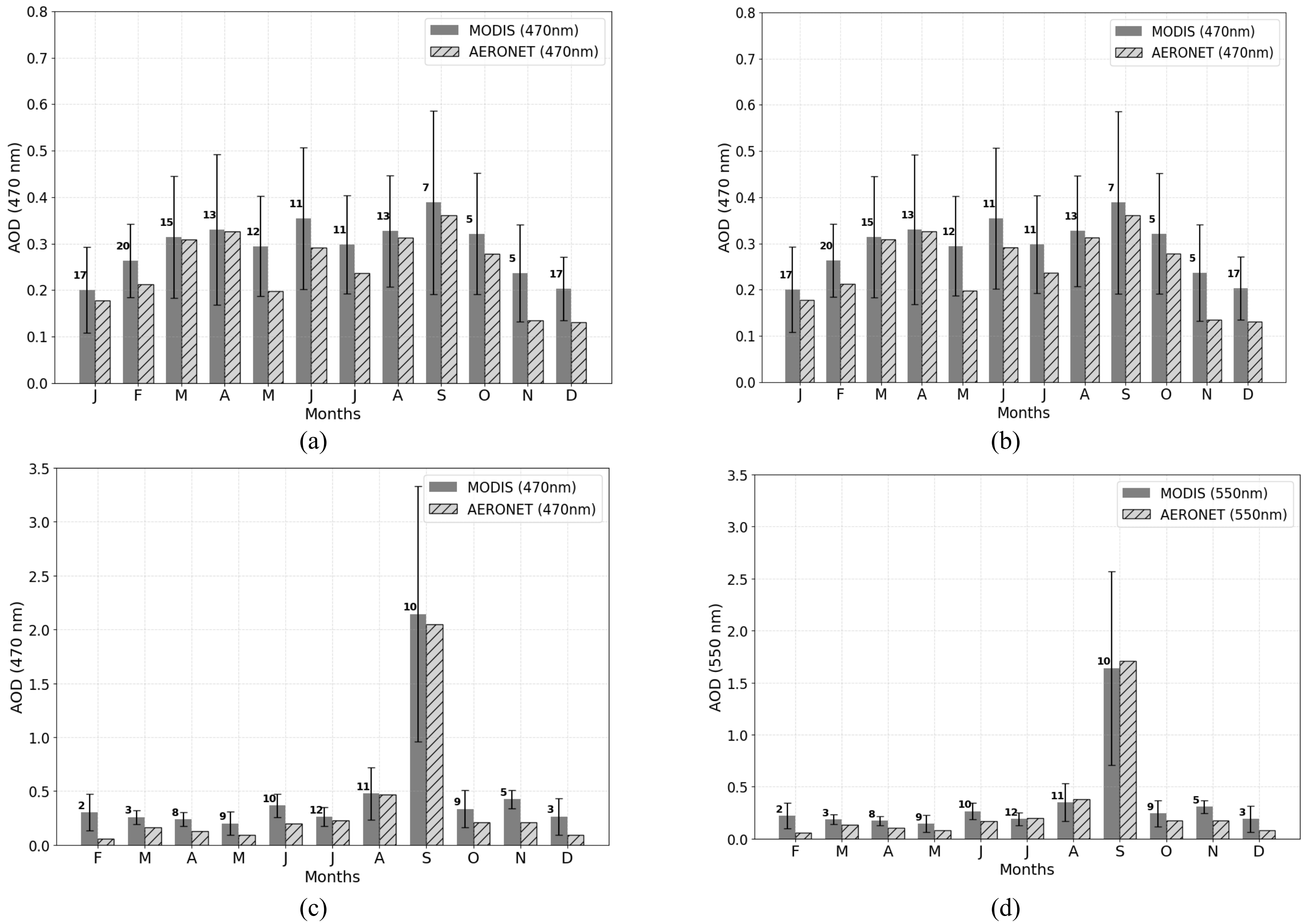


Figure 4. Monthly averaged AOD of MODIS and AERONET at (a) 470nm and (b) 550 nm for Penang and Kuching at (c) 470nm and (b) 550 nm, respectively.

The trend showed that Penang experienced elevated aerosol loading during mid-year due to external pollution sources. The high standard deviation during August and September also confirmed the occurrence of unstable and high-pollution episodes during these months. Figure 4 (a) and Figure 4 (b) show the AOD trend in Kuching for 470 nm and 550 nm respectively. Similar to Penang, Kuching showed low AOD values in the early months and a significant increase from July to September. The highest mean AOD values occurred in August, aligning with major regional fire activities in Borneo and Indonesia. In contrast to Penang, Kuching showed slightly higher AOD levels during the haze period, especially at 470 nm. The error bars during peak months were also larger, indicating high aerosol variation. After September, the AOD values in Kuching gradually declined, suggesting the end of haze season and more stable atmospheric conditions When comparing both sites, Kuching showed higher AOD magnitudes than Penang during the dry season, particularly at 470 nm. This may be due to Kuching's proximity to the major biomass burning sources and local meteorological factors that trapped aerosols over the region. At 550 nm, the difference was slightly less, but Kuching still recorded higher peak values. This suggested that Kuching faced more intense haze loading compared to Penang during 2015.

Overall, the monthly trend analysis confirmed that the highest aerosol concentrations occurred during the southwest monsoon season (June to September) in both locations. This period coincided with regional haze transport and stagnant air masses. The corrected MODIS AOD values aligned well with AERONET trends, proving that the spectral and temporal corrections improved the monthly comparison. The error bars helped visualise uncertainty and confirmed that aerosol loading varied more during the dry season compared to inter-monsoon and wet periods.

The seasonal analysis of AOD for both Penang and Kuching at 470 nm and 550 nm wavelengths is presented in Table 2. The data were categorised into three main seasons based on the Malaysian monsoon cycle: wet season (November to March), inter-monsoon (April to May and October), and dry season (June to September). The aim of this analysis was to identify how aerosol concentrations varied across different seasonal phases and to assess the performance of MODIS AOD compared to AERONET measurements during these periods. Seasonal averages were calculated from the MODIS and collocated AERONET datasets for each station and wavelength. Figures and tables were generated to display the variation in mean AOD for each season. Standard deviation was also included to reflect the variation in AOD values, especially during high-aerosol periods. During the inter-monsoon, the performance was moderate with R = 0.7625, RMSE =

0.0921, and a bias of 0.0670. MODIS still tracks the AERONET trend, but with reduced accuracy, possibly due to unstable weather and mixed atmospheric conditions. This section presents the seasonal regression analysis between MODIS and AERONET AOD values during the wet, dry, and inter-monsoon periods for both Penang and Kuching at 470 nm and 550 nm wavelengths. The statistical results are summarised in Table 2, while the corresponding scatter plots are shown in Figure 5. In Penang at 470 nm, the dry season showed the highest correlation of r equal to 0.7529 with a low bias of 0.0419, indicating better MODIS retrieval during haze periods. As shown in Table 2, this was also supported by lower RMSE and unbiased RMSE values. During the wet season, the correlation dropped to R equal to 0.5619, with increased error likely due to cloud contamination. The inter-monsoon period showed moderate correlation of R equal to 0.6813 with slightly higher error values.

In Penang at 550 nm, MODIS AOD showed fairly stable performance throughout all three seasons. The dry season achieved a strong correlation of R equal to 0.7218 with low RMSE and bias, confirming accuracy under high aerosol load. As presented in Table 2, the wet season showed the lowest RMSE and minimal bias, with R equal to 0.6691, suggesting better atmospheric stability. The inter-monsoon period had slightly better correlation of R equal to 0.7414 with moderate error. For Kuching at 550 nm, the dry season showed the strongest correlation of R equal to 0.9738 with nearly zero bias at negative 0.0043, indicating excellent agreement. The wet season had higher RMSE and bias due to cloud interference, while the inter-monsoon period showed stable correlation of R equal to 0.7625 with the lowest unbiased RMSE of 0.0632.

Overall, the dry season consistently gave stronger MODIS-AERONET agreement across all wavelengths and sites. This was likely due to clearer skies and the presence of elevated aerosol levels from biomass burning, which improved satellite retrieval accuracy. In contrast, the wet season showed the weakest correlation, primarily due to retrieval interference from clouds and water vapour. The performance at 550 nm was generally better than 470 nm, since the longer wavelength was less sensitive to Rayleigh scattering and more robust under tropical atmospheric conditions.

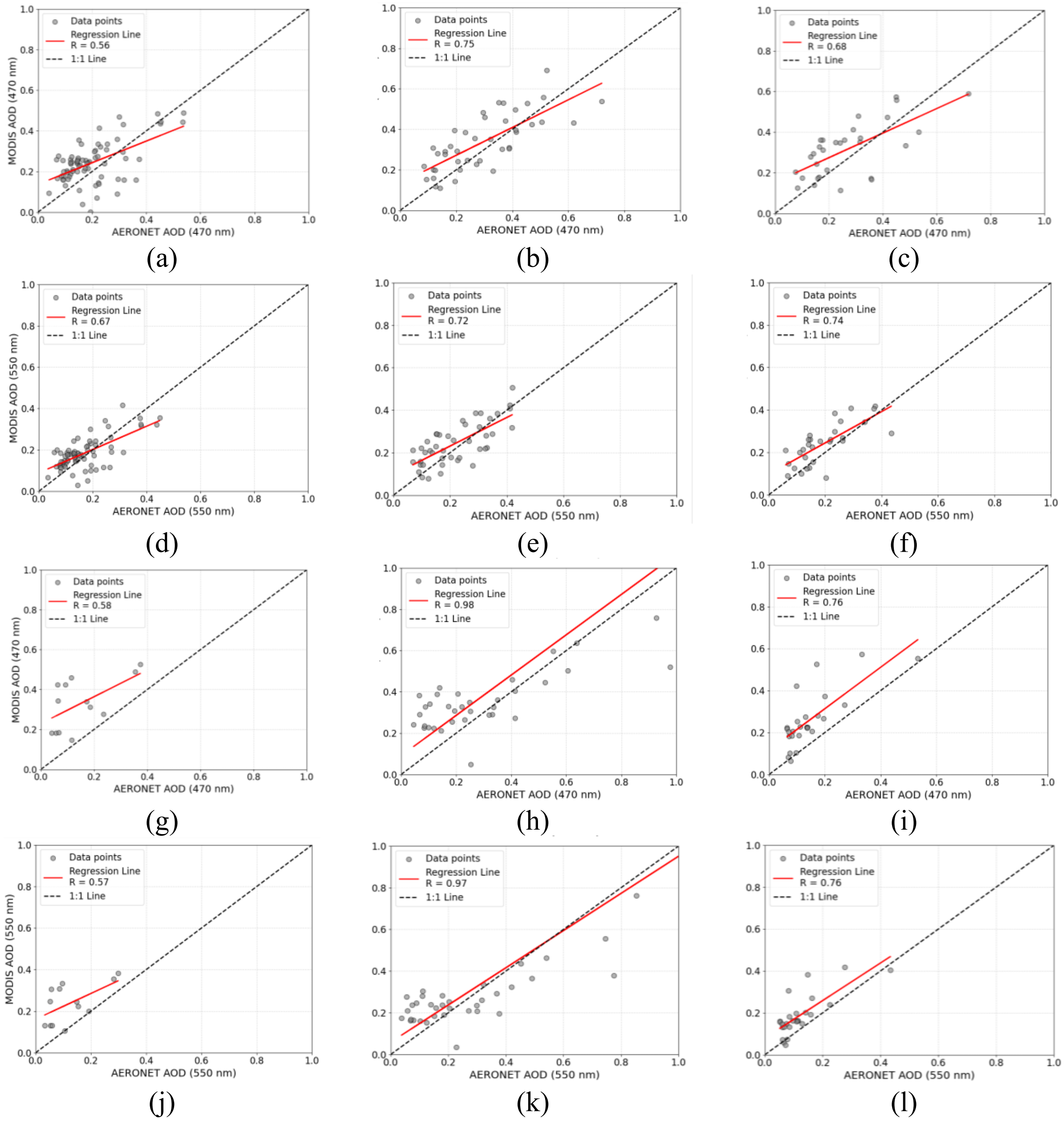


Figure 5. Seasonal regression between MODIS and AERONET AOD for USM Penang and Kuching at 470 nm and 550 nm during Wet, Dry, and Inter-monsoon periods: (a–c) Penang 470 nm – Wet, Dry, and Inter-monsoon; (d–f) Penang 550 nm – Wet, Dry, and Inter-monsoon; (g–i) Kuching 470 nm – Wet, Dry, and Inter-monsoon; (j–l) Kuching 550 nm – Wet, Dry, and Inter-monsoon.

Table 2. Regression statistics of MODIS and AERONET AOD at 470 nm and 550 nm for Penang and Kuching based on monsoon seasons.

Location & AOD wavelength	Season	N	R	RMSE	Bias	Unbiased RMSE
Penang 470 nm	Wet Season	74	0.5619	0.1065	0.0434	0.0973
	Dry Season	42	0.7529	0.1101	0.0419	0.1019
	Inter-monsoon	30	0.6813	0.1226	0.0478	0.1129
Penang 550 nm	Wet Season	72	0.6691	0.0698	0.0154	0.0681
	Dry Season	44	0.7218	0.0798	0.0250	0.0758
	Inter-monsoon	30	0.7414	0.0828	0.0463	0.0686
Kuching 470 nm	Wet Season	13	0.5797	0.2101	0.1808	0.1070
	Dry Season	43	0.9764	0.2171	0.0740	0.2041
	Inter-monsoon	26	0.7560	0.1433	0.1143	0.0865
Kuching 550 nm	Wet Season	13	0.5719	0.1398	0.1139	0.0810
	Dry Season	43	0.9738	0.1867	-0.0043	0.1866
	Inter-monsoon	26	0.7625	0.0921	0.0670	0.0632

3.5 AOD Mapping

The final AOD maps were generated from the spatially corrected MODIS MAIAC data using Google Earth Engine and processed into GeoTIFF format. These maps represented monthly average AOD distributions at 470 nm and 550 nm for Penang and Kuching. The GeoTIFF layers were imported into ArcMap to produce spatial distribution maps using standard cartographic format. Classification was done using defined AOD thresholds to represent low, moderate and high aerosol concentrations. The classification thresholds were determined based on previous local studies and site-specific AERONET records. Earlier work in Malaysia reported that background conditions were associated with low AOD values, while haze periods were characterised by higher AOD values that frequently exceeded 0.8 (Kanniah et al., 2014). These local findings were used as a reference for distinguishing between normal and haze-influenced conditions. In addition, long-term AERONET Level 2.0 data from USM Penang and Kuching were examined to establish site-specific aerosol ranges. These datasets provided statistical evidence of the typical AOD levels observed during both background and haze periods. By combining the results from previous studies with AERONET site climatology, the AOD thresholds were defined in this study to represent low, moderate and high aerosol loading.

Then, Custom colour schemes were applied using the symbology tool in ArcMap to improve visual clarity. In Penang, high AOD values were generally observed in the southwest and central regions, especially during haze periods. The coastal zones showed more variability due to mixed surface reflectance and cloud contamination. In Kuching, elevated AOD levels were more dominant during dry season months, especially in the southern and central parts of the city. These patterns were consistent with transboundary haze episodes from nearby regions such as Kalimantan. This map is not from the corrected AOD dataset and is not part of the study objectives. It is included as visual support to show general AOD patterns in Kuching. The result helps support seasonal trends and is useful for showing how AOD changes over time and across the area.

The spatial patterns clearly showed seasonal variation in aerosol concentrations. Months with higher AOD values corresponded well with known haze periods, while lower values matched with wet season or inter-monsoon periods. Differences between the two wavelengths were also noticeable. At 470 nm, maps tended to show more scattered high values, while at 550 nm, the spatial patterns were smoother due to less sensitivity to fine particles. Cloud contamination was still present in some months, especially in coastal and high-humidity areas. This caused missing pixels or small gaps in the output maps. Despite this limitation, the spatial mapping provided clear insights into AOD distribution across both study areas and supported the statistical and seasonal trend findings. These maps served as a useful reference to understand aerosol behaviour over urban and semi-urban environments under varying atmospheric conditions.

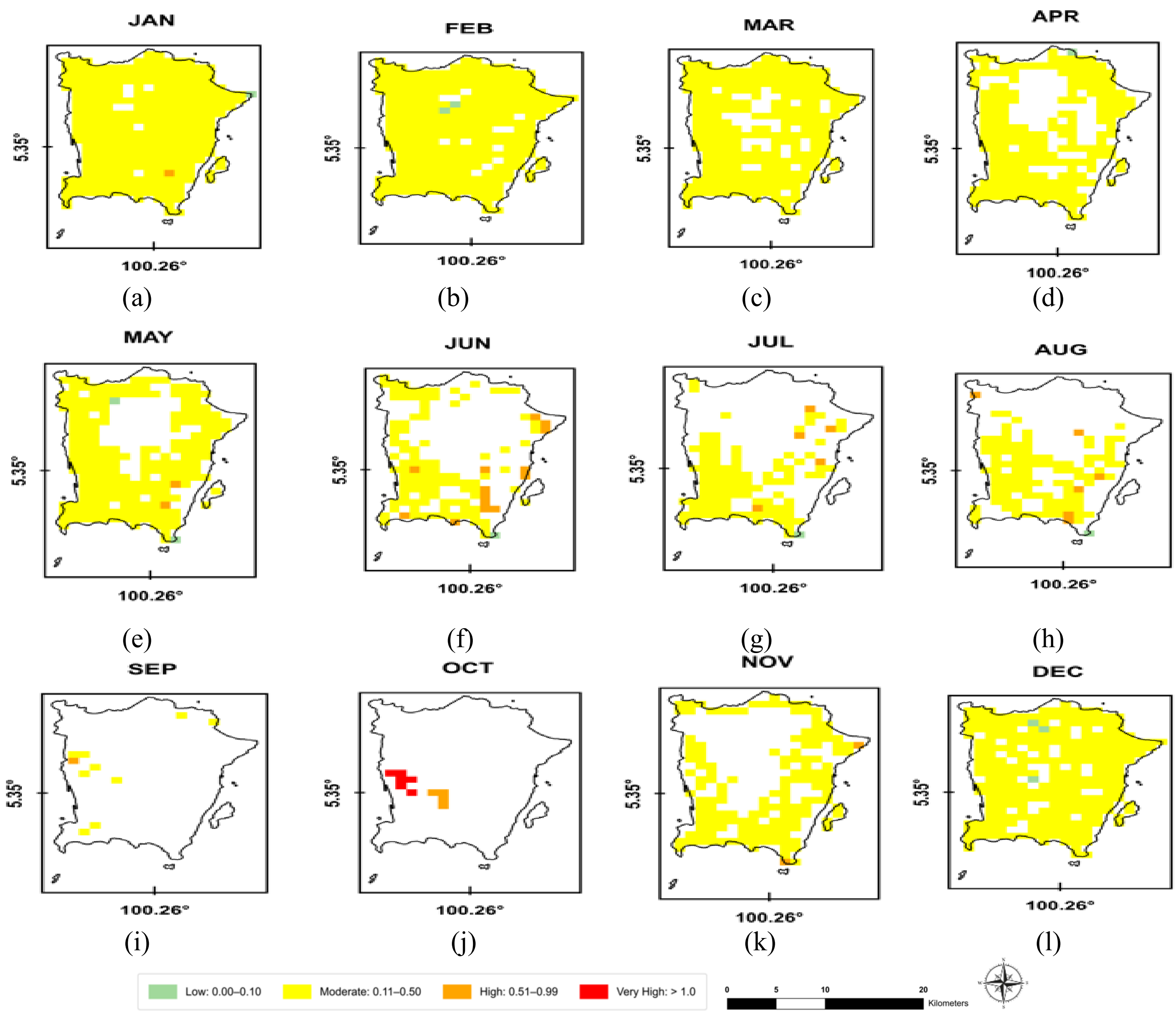


Figure 6. MODIS AOD distribution at 470nm over Penang for (a) January, (b) February, (c) March, (d) April, (e) May, (f) June, (g) July, (h) August, (i) September, (j) October, (k) November, and (l) December in 2015.

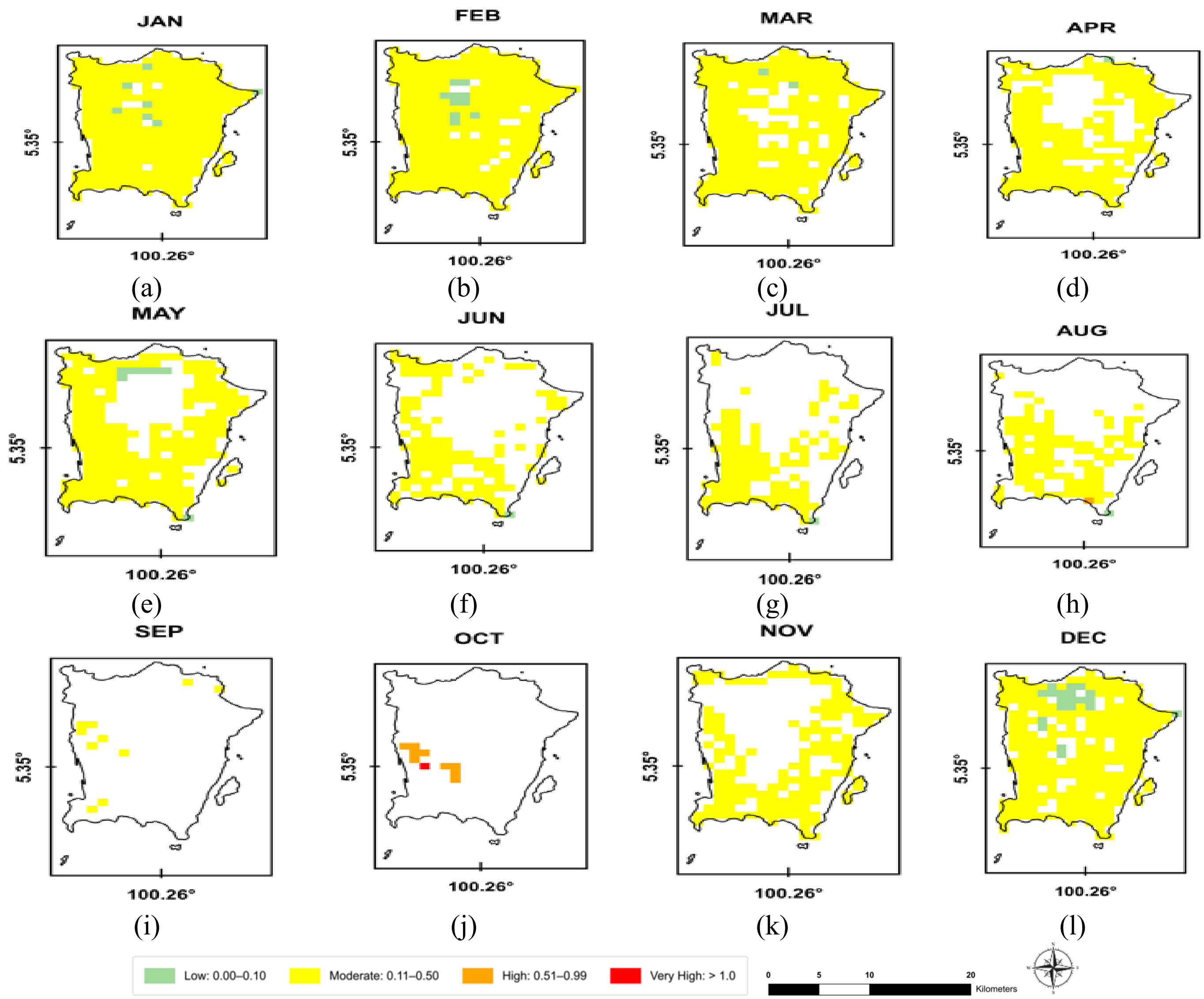


Figure 7. MODIS AOD distribution at 550nm over Penang for (a) January, (b) February, (c) March, (d) April, (e) May, (f) June, (g) July, (h) August, (i) September, (j) October, (k) November, and (l) December in 2015.

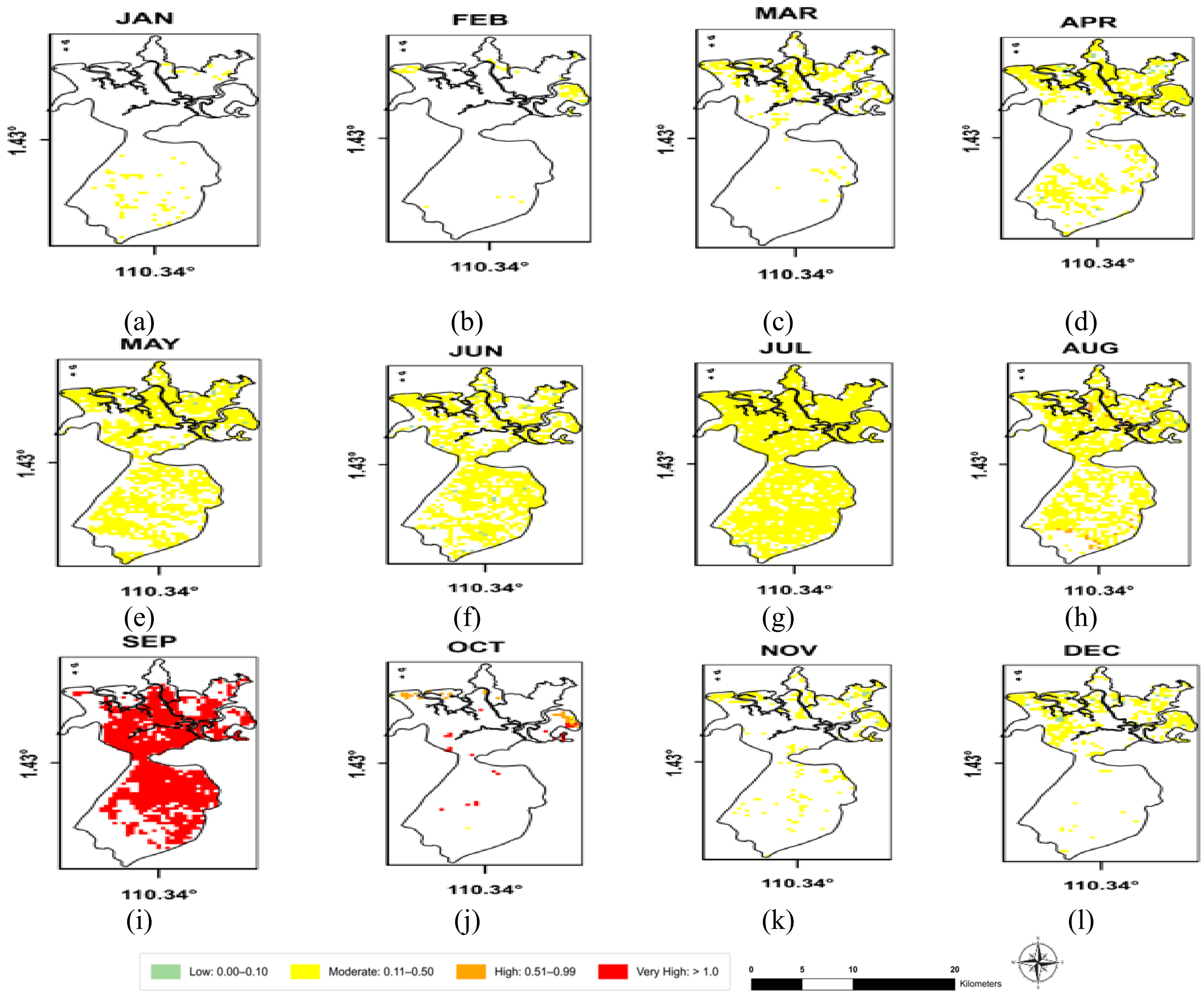


Figure 8. MODIS AOD distribution at 470nm over Kuching for (a) January, (b) February, (c) March, (d) April, (e) May, (f) June, (g) July, (h) August, (i) September, (j) October, (k) November, and (l) December in 2015.

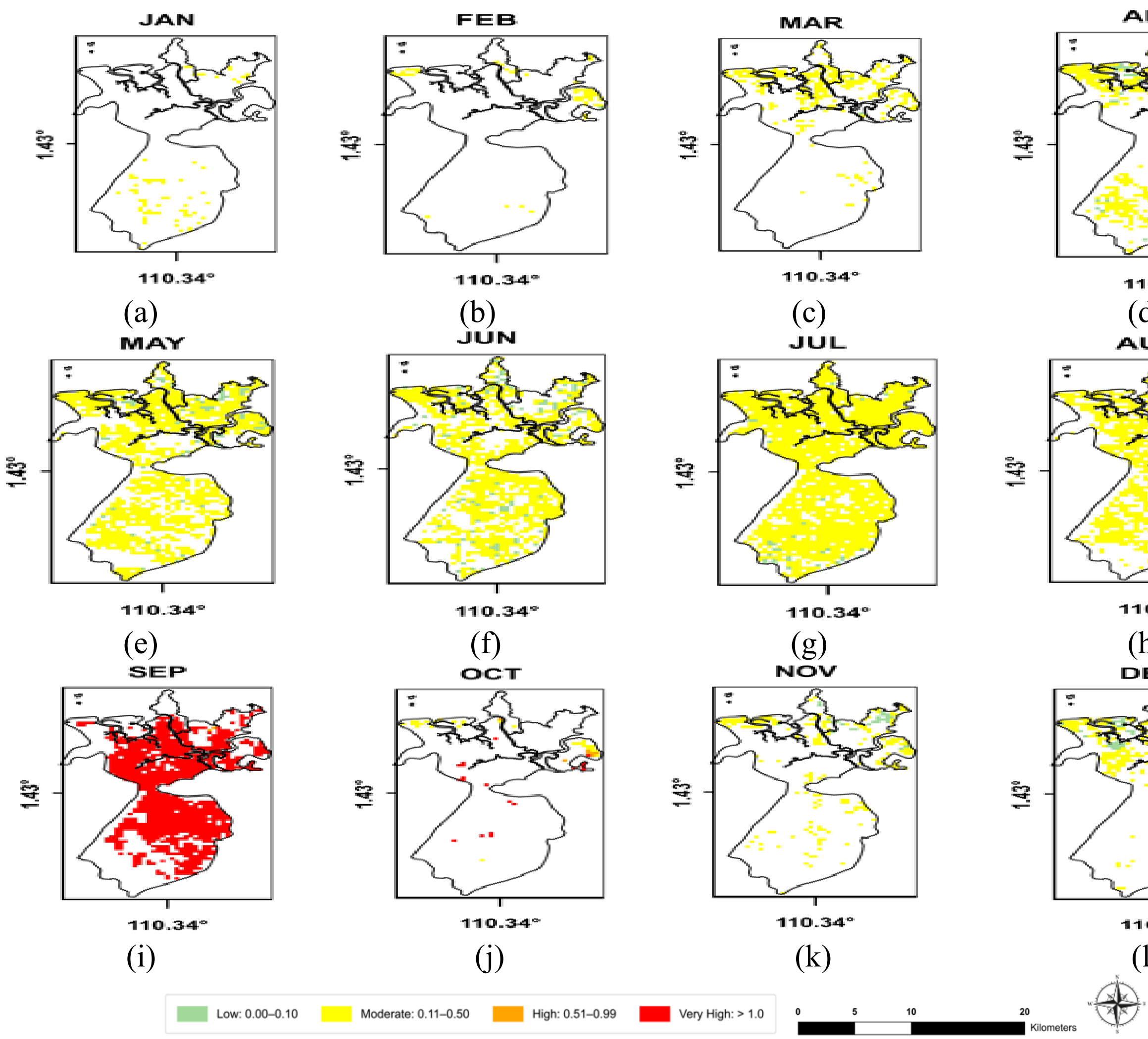


Figure 9. MODIS AOD distribution at 550nm over Kuching for (a) January, (b) February, (c) March, (d) April, (e) May, (f) June, (g) July, (h) August, (i) September, (j) October, (k) November, and (l) December in 2015.

4.0 Discussion

The results show that MODIS AOD overestimates compared to AERONET, but the correction methods applied improved the accuracy significantly. From the regression analysis, 550 nm performed better than 470 nm with lower RMSE and bias values. In Kuching, 550 nm gave a strong correlation, R = 0.9728 and low RMSE of 0.1551, while 470 nm had slightly higher error and bias. This justifies the use of 550 nm is suitable for further analysis. The comparison between Kuching and Penang also showed that Kuching had better validation results. Penang regression values were lower, and MODIS showed more variation. This may be due to surface reflectance issues near the coast, frequent cloud cover, or differences in aerosol type. Kuching, located in land and exposed to transboundary haze, gave more stable results. This supports using Kuching as the main study focus.

The seasonal regression analysis confirms that MODIS AOD is more reliable during the dry season. The dry monsoon (June–September) showed the best performance with R = 0.9738 minimal bias. During this period, the sky is clearer, and MODIS captures aerosol signals better. However, during the wet season (November–March), the accuracy dropped due to high cloud cover and moisture interference, leading to more retrieval errors. The monthly AOD trend showed a clear peak in September, which matches the annual haze episode caused by biomass burning in Sumatra and Kalimantan. This regional smoke was transported across to Sarawak during the dry monsoon, that leading the AOD levels significantly. The spatial AOD map for Kuching in 2015 supports this, showing a visible increase in AOD coverage during the haze months.

Overall, the applied spatial, spectral, and temporal corrections successfully reduced the gap between MODIS and AERONET data. The combination of pixel-based regression, time normalization, and spectral interpolation made MODIS AOD more reliable for air quality monitoring. However, there are still some limitations. MODIS data can be affected by cloud contamination, especially during the wet and inter-monsoon seasons. The number of AERONET stations are also limited, which affects spatial representation. Spectral interpolation may introduces estimation error, especially when using only a few bands. Despite these issues, the corrected MODIS AOD still provides a valuable tool for air quality monitoring with limited ground data.

5.0 Conclusion

This study focused on improving the accuracy of MODIS satellite-derived Aerosol Optical Depth (AOD) by applying correction methods using AERONET Level 2.0 data. The analysis was conducted for two sites in Malaysia, namely USM Penang and Kuching, Sarawak, using MODIS MAIAC AOD at 470 nm and 550 nm. The main objective was to reduce the bias in MODIS AOD data through spatial, spectral, and temporal correction techniques.

The regression results showed that MODIS AOD at 550 nm performed better than 470 nm, with lower RMSE and bias values. Kuching gave more reliable validation results compared to Penang, with a high correlation of $R = 0.9728$ and RMSE = 0.1551 after correction. On the other hand, Penang showed moderate correlation and higher error values, likely due to surface reflectance and frequent cloud cover. These findings confirm that 550 nm wavelength AOD datasets is more suitable for AOD validation and was selected for further analysis.

The seasonal analysis also proved that MODIS AOD is more accurate during the dry monsoon season, where less cloud contamination occurs. The best regression performance was recorded in the dry period from June to September, while the wet season had the lowest accuracy. The monthly AOD trend also showed a clear spike in September, which is aligns with haze events which caused by biomass burning in Sumatra and Kalimantan. The spatial AOD map supported this seasonal pattern, showing higher aerosol loading during haze months.

The correction methods used in this study, including 3×3 spatial regression, spectral interpolation, and time normalization, successfully improved the MODIS AOD data. The corrected data is now more aligned with AERONET and suitable for long-term air quality analysis, especially in areas with limited ground-based stations. Although some limitations still exist, such as cloud interference and limited AERONET coverage, the approach used in this study helps improve the reliability of satellite-based aerosol monitoring in Malaysia.

The findings of this study, after validation, were expected to provide insights into several important aspects related to aerosols and their implications. One important aspect was air quality and its impact on health. Higher AOD values were taken as a proxy for lower air quality because denser aerosol concentrations attenuated solar radiation more strongly. This indicated that fine particles were more abundant in the atmosphere, which gave a direct impact on human respiratory health. During haze episodes, this link became clear as AOD increased significantly and corresponded to greater exposure to pollutants that could aggravate respiratory and cardiovascular diseases. Moreover, aerosols also played a role in climate impact. Seasonal variation strongly influenced aerosol loading in Malaysia. During the dry season, limited rainfall allowed aerosols to remain in the atmosphere for longer periods, leading to higher AOD values and stronger radiative effects. Meanwhile, during the wet season, rainfall washed out most of the aerosols, resulting in much lower AOD values. These patterns showed the strong connection between climate variability and aerosol persistence, which could also be affected by long-term climate change.

In addition, AOD was also influenced by the transport of dust, smoke and other pollutants. Different particle types, whether coarse dust, fine smoke, or urban and industrial pollution, contributed differently to total AOD. The variability in AOD therefore reflected the combined influence of these sources. The analysis provided evidence of long-range transport, particularly during transboundary haze episodes, when smoke and pollutants moved across national boundaries and intensified aerosol loading in Malaysia. Furthermore, understanding AOD variability from MODIS images supported better use of models and policies for decision-making. Satellite-based AOD extended the coverage to regions without AERONET stations and complemented ground-based measurements. This integration allowed for more accurate representation of aerosol distribution and provided reliable datasets for haze management, air quality models and climate simulations. Such information was important for guiding policy and planning at both local and regional scales.

Finally, the findings also addressed the limitation of ground-based measurements and highlighted the advantage of satellite products. AERONET provided accurate and consistent point-based data, but its limited coverage restricted wider applications. MODIS AOD, available at multiple spectral bands, overcame this limitation by offering continuous spatial coverage. These datasets could be integrated into atmospheric radiative transfer models such as MODTRAN and LOWTRAN, improving the representation of aerosol variability and strengthening both research and practical monitoring applications.

Acknowledgement

The research was funded by Jabatan Tanah dan Survei, Sarawak through Universiti Teknologi Malaysia, Geran Kursi Premier Sarawak (R.J130000.7309.1R055).

References

Antuña-Marrero, J. C., Cachorro Revilla, V., García Parrado, F., de Frutos Baraja, Á., Rodríguez Vega, A., Mateos, D., Estevan Arredondo, R., & Toledano, C. (2018). Comparison of aerosol optical depth from satellite (MODIS), sun photometer and broadband pyrheliometer ground-based observations in Cuba. Atmospheric Measurement Techniques, 11(4), 2279-2293.

Du, B., Zhong, B., Cai, H., Wu, S., Qiao, Y., Wang, X., Yang, A., Wu, J., Liu, Q., & Jiang, J. (2025). Improving AOD Algorithm Evaluation: A Spatial Matching Method for Minimizing Quality Control Bias. Remote Sensing, 17(7), 1235.

Gupta, P., Levy, R. C., Mattoo, S., Remer, L., Zhang, Z., Sawyer, V., Wei, J., Zhao, S., Oo, M., & Kiliyanpilakkil, V. P. (2024). Increasing aerosol optical depth spatial and temporal availability by merging datasets from geostationary and sun-synchronous satellites. Atmospheric Measurement Techniques Discussions, 2024, 1-32.

Hee, W., Tan, F., Lim, H., & Matjafri, M. (2014). Haze event monitoring and investigation in Penang Island, Malaysia using a ground-based backscatter Lidar. IOP Conference Series: Earth and Environmental Science.

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

Ichoku, C., Chu, D. A., Mattoo, S., Kaufman, Y. J., Remer, L. A., Tanré, D., Slutsker, I., & Holben, B. N. (2002). A spatio-temporal approach for global validation and analysis of MODIS aerosol products. Geophysical Research Letters, 29(12), MOD1-1-MOD1-4.

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

Kanniah, K. D., Lim, H. Q., Kaskaoutis, D. G., & 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. http://dx.doi.org/10.1016/j.atmosres.2013.11.018

Kokhanovsky, A., Breon, F.-M., Cacciari, A., Carboni, E., Diner, D., Di Nicolantonio, W., Grainger, R., Grey, W., Höller, R., & Lee, K.-H. (2007). Aerosol remote sensing over land: A comparison of satellite retrievals using different algorithms and instruments. Atmospheric Research, 85(3-4), 372-394.

Li, R.-R., Kaufman, Y., Mattoo, S., Remer, L., Chu, D., Martins, J., Kleidman, R., Levy, R., Ichoku, C., & Tanre, D. (2005). The MODIS aerosol algorithm, products and validation.

NASA. (2025). AERONET: Aerosol Robotic Network. Retrieved 10 March 2025 from https://aeronet.gsfc.nasa.gov/

Nations, U. (2024). Sustainable Development Goals. Retrieved 16 december 2024 from https://sdgs.un.org/goals

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

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

Song, S., Kang, Y., & Im, J. (2024). Enhanced continuous aerosol optical depth (AOD) estimation using geostationary satellite data: focusing on nighttime AOD over East Asia. arXiv preprint arXiv:2405.13334. https://doi.org/10.48550/arXiv.2405.13334

Tan, F., Lim, H., Abdullah, K., Yoon, T., & Holben, B. (2015). Monsoonal variations in aerosol optical properties and estimation of aerosol optical depth using ground-based meteorological and air quality data in Peninsular Malaysia. Atmospheric Chemistry and Physics, 15(7), 3755-3771. https://doi.org/:10.5194/acp-15-3755-2015

Tripathi, S., Dey, S., Chandel, A., Srivastava, S., Singh, R. P., & Holben, B. (2005). Comparison of MODIS and AERONET derived aerosol optical depth over the Ganga Basin, India. Annales Geophysicae.