

Spatio-Temporal Analysis of PM_{2.5} Concentrations through Integration of Aerosol Optical Depth and Sentinel-5P Data Using Google Earth Engine

Worawit Suppawimut*, Phonpat Hemwan, Watinee Thavorntam, and Arisara Charoenpanyanet

Department of Geography, Faculty of Social Sciences, Chiang Mai University, Chiang Mai, Thailand

ARTICLE INFO

Received: 4 Sep 2025
Received in revised: 20 Jan 2026
Accepted: 26 Jan 2026
Published online: 1 May 2026
DOI: 10.32526/ennrj/24/20250246

Keywords:

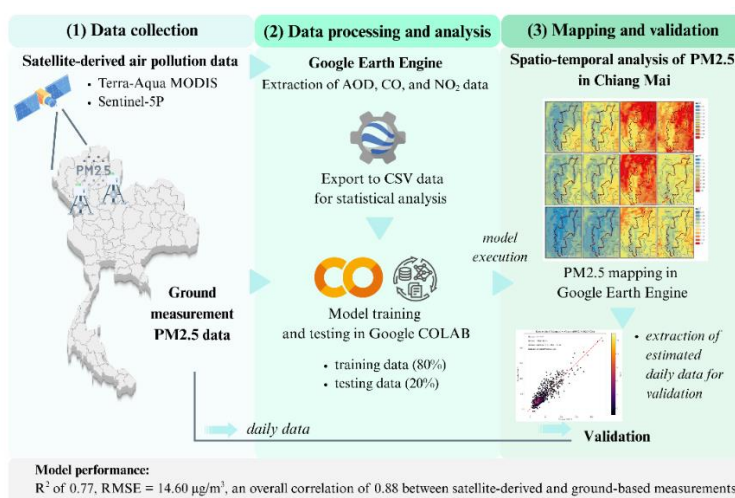
Air pollution/ Remote sensing/
Particulate matter/ Satellite data/
Aerosol optical depth

* Corresponding author:

E-mail: worawit_sup@cmu.ac.th

ABSTRACT

GRAPHICAL ABSTRACT



Air pollution in Northern Thailand is a persistent environmental issue, particularly during the dry season, posing significant health risks. This study used multiple remote-sensing data sources to explore spatio-temporal PM_{2.5} variations in Chiang Mai Province from 2020 to 2024. Ground-based PM_{2.5} measurements were merged with Terra-Aqua aerosol optical depth (AOD) and Sentinel-5P carbon monoxide (CO) and nitrogen dioxide (NO₂) data within the Google Earth Engine platform. Multiple linear regression models were employed to estimate ground-level PM_{2.5} concentrations and generate spatial distribution maps. Central lowland and northern areas of Chiang Mai consistently exhibited higher PM_{2.5} levels, particularly in March and April. The model achieved an R² of 0.77 and a root mean square error (RMSE) of 14.60 µg/m³, with an overall correlation of 0.88 between satellite-derived and ground-based measurements. Seasonal analysis revealed enhanced model performance during the burning period (January-April; R²=0.72, RMSE=17.31 µg/m³), compared with the non-burning period (May-December; R²=0.26, RMSE=8.32 µg/m³). During the 2020-2024 burning periods, average PM_{2.5} concentrations were 51.33, 45.62, 25.18, 49.81, and 41.11 µg/m³, respectively, peaking at 79.18±16.90 µg/m³ in March 2023. Integrating AOD with Sentinel-5P CO and NO₂ data improved estimation accuracy and hotspot identification, highlighting the potential of cloud-based geospatial platforms for comprehensive air quality monitoring.

HIGHLIGHTS

- Multi-source satellite data integration (AOD, CO, NO₂) within the Google Earth Engine successfully estimated PM_{2.5} concentrations, achieving high accuracy R²=0.77, RMSE=14.60 µg/m³, and R=0.88 correlation.
- Model performance showed significant seasonal variation, with higher accuracy during burning periods (R²=0.72) versus non-burning periods (R²=0.26), while combining aerosol and gas species data from TROPOMI substantially improved estimation over AOD-only models.
- This cloud-based remote sensing approach effectively mapped PM_{2.5} hotspots, demonstrating strong potential for air quality monitoring in areas with limited ground-based measurements.

Citation: Suppawimut W, Hemwan P, Thavorntam W, Charoenpanyanet A. Spatio-temporal analysis of PM_{2.5} concentrations through integration of aerosol optical depth and Sentinel-5P data using google earth engine. Environ. Nat. Resour. J. 2026;24(4):382-396. (<https://doi.org/10.32526/ennrj/24/20250246>)

1. INTRODUCTION

Air pollution, specifically fine particulate matter (PM_{2.5}), is a critical environmental issue in Thailand, significantly impacting public health (Chankaew et al., 2022). According to the Pollution Control Department (2023), annual average PM_{2.5} concentrations from 87 stations nationwide ranged from 12 to 49 µg/m³, with many stations exceeding the national average of 26 µg/m³. This issue is particularly severe in Northern Thailand during the dry season. Epidemiological studies have consistently demonstrated associations between elevated PM_{2.5} levels and respiratory and cardiovascular diseases (Buya et al., 2026; Fatima et al., 2023). Notably, Sapbamrer et al. (2024) revealed direct associations between short-term PM_{2.5} exposure and COPD fatality rates in the north of Thailand, while Buya et al. (2026) found significant correlations between monthly PM_{2.5} levels and cardiorespiratory mortality during 2015-2019.

Agricultural residue burning and forest fires are the primary pollution sources in Northern Thailand, particularly during the dry season (Junpen et al., 2018; Paluang et al., 2024). Road transportation is another major contributor, with traffic emissions accounting for 76% (27.2 µg/m³) of PM_{2.5} during the non-burning period in Chiang Mai (Chansuebsri et al., 2024). This proportion decreased to 13% during the burning period, when biomass burning became predominant (51%). Transboundary pollution and limited public transportation development also create significant barriers to effective air quality management (Bran et al., 2022; Chansuebsri et al., 2024).

Ground-based PM_{2.5} measurements are essential for air quality monitoring and health impact assessment (Holloway et al., 2021). However, the insufficient number and poor distribution of monitoring stations limit spatial coverage, resulting in data gaps in rural regions affected by biomass burning (Son et al., 2023; Song et al., 2019). Accurate and spatially extensive PM_{2.5} estimation is crucial for addressing these limitations and enhancing exposure assessments (Chen et al., 2018).

Satellite-based remote sensing is essential for wide spatial coverage. Previous studies have derived surface PM_{2.5} from aerosol optical depth (AOD), employing methodologies including statistical regression (Hands Schuh et al., 2022) and machine learning (Song et al., 2025; Wongnakae et al., 2023). However, processing these multi-temporal satellite datasets requires substantial computational resources.

Google Earth Engine (GEE) has emerged as a prominent cloud-based platform that mitigates this challenge by enabling efficient analysis of petabyte-scale geospatial data without the need for high-performance local infrastructure. Recent studies have demonstrated its effectiveness in spatio-temporal air quality monitoring and pollutant mapping (Ghasempour et al., 2021; Paluang et al., 2024; Rirugchart et al., 2025).

Despite these advancements, localized studies considering region-specific characteristics are still critically needed (Pardthaisong et al., 2018). Chiang Mai is characterized by seasonal biomass burning, basin topography, and specific meteorology, creating a complex context that requires specialized approaches. This study examines the spatio-temporal distribution of PM_{2.5} in Chiang Mai from 2020 to 2024 by integrating ground data with satellite AOD, CO, and NO₂ data using multiple linear regression (MLR) within GEE. It aims to clarify spatial distribution patterns in unmonitored areas, offering a framework applicable to different contexts and providing spatial data for epidemiological research and air pollution monitoring.

2. METHODOLOGY

2.1 Study area

The study area is Chiang Mai Province, located in Northern Thailand between latitudes 17°14' and 20°8'N and longitudes 98°00' and 99°34'E, covering approximately 20,107 km² with a 2024 population of 1.79 million (Chiang Mai Provincial Statistical Office, 2024). The terrain is predominantly mountainous with central lowlands, with an average annual temperature of 25.4°C (range: 20.1-31.8°C). This basin-shaped topography (Figure 1) restricts atmospheric ventilation during biomass burning, exacerbating pollutant accumulation and public health risks (Chansuebsri et al., 2024). The study area and the distribution of air quality monitoring stations are illustrated in Figure 1.

2.2 Data collection

2.2.1 Ground-based PM_{2.5}

Daily average PM_{2.5} concentrations from 2020 to 2024 were obtained from nine air quality monitoring stations, provided by the Pollution Control Department of Thailand (PCD). Due to the limited number of ground monitoring stations in Chiang Mai, the dataset includes measurements from PM_{2.5} stations located in both Chiang Mai (35T and 36T) and neighboring provinces: Lampang (37T and 40T),

Chiang Rai (57T and 73T), Mae Hong Son (58T), Lamphun (68T), and Tak (76T). The MLR model uses this data as its dependent variable (Y). Details and sources of the data are presented in Table 1. Prior to analysis, the ground-based PM_{2.5} data underwent quality control, with missing data, zeros, and invalid values removed from the dataset.

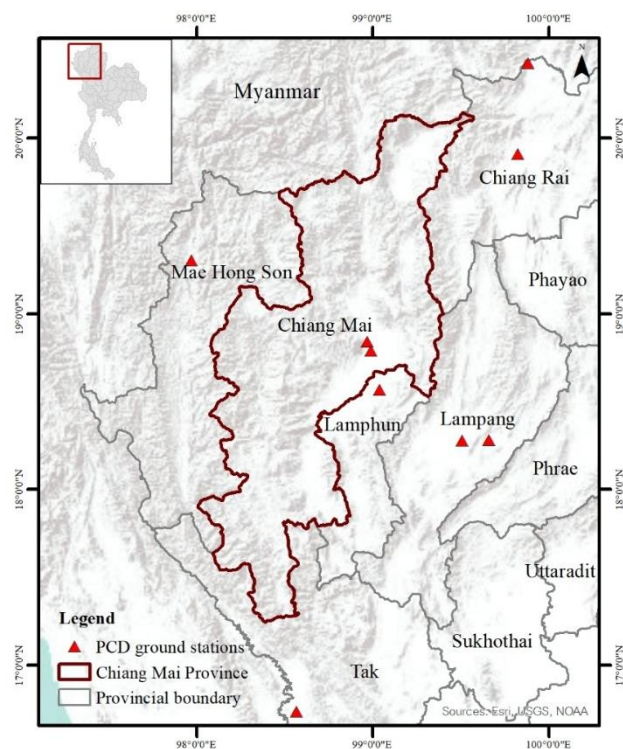


Figure 1. Map of the study area covering Chiang Mai Province and neighboring provinces in Northern Thailand, showing the locations of air quality ground stations.

2.2 Data collection

2.2.1 Ground-based PM_{2.5}

Daily average PM_{2.5} concentrations from 2020 to 2024 were obtained from nine air quality monitoring stations, provided by the Pollution Control Department of Thailand (PCD). Due to the limited number of ground monitoring stations in Chiang Mai,

the dataset includes measurements from PM_{2.5} stations located in both Chiang Mai (35T and 36T) and neighboring provinces: Lampang (37T and 40T), Chiang Rai (57T and 73T), Mae Hong Son (58T), Lamphun (68T), and Tak (76T). The MLR model uses this data as its dependent variable (Y). Details and sources of the data are presented in Table 1. Prior to analysis, the ground-based PM_{2.5} data underwent quality control, with missing data, zeros, and invalid values removed from the dataset.

2.2.2 Aerosol optical depth

AOD is a unitless measure of aerosols distributed within a column of air from Earth's surface to the top of the atmosphere. In this study, the Level 2 AOD product was obtained from the combined Terra and Aqua MODIS sensors using the Multi-angle Implementation of Atmospheric Correction (MAIAC) algorithm (MCD19A2). Crucially, this algorithm offers a fine spatial resolution of 1×1 km², which has been shown to significantly improve the characterization of PM_{2.5} spatial distribution compared with coarser datasets (Holloway et al., 2021; Rirugchart et al., 2025; Song et al., 2025). The 0.55 µm band of this AOD data for ground-based PM_{2.5} monitoring locations, serving as the primary independent variable in the model, was extracted using GEE.

2.2.3 TROPOMI

Sentinel-5P is an air quality monitoring satellite developed by the European Space Agency as part of the Copernicus Programme. It carries the Tropospheric Monitoring Instrument (TROPOMI), which measures atmospheric chemical concentrations at high resolution. TROPOMI daily data of carbon monoxide (CO) and nitrogen dioxide (NO₂) were retrieved from GEE at 1×1 km² grid (Table 1). CO and NO₂ have been effectively utilized in air pollution studies to identify areas with elevated biomass burning and traffic emissions, respectively (Pathakoti et al., 2021; Sakti et al., 2023).

Table 1. List of datasets used in this study

Data	Variables	Units	Temporal resolution	Data source
Ground-based PM _{2.5} (2020-2024)	Daily average PM _{2.5}	µg/m ³	Daily	PCD, Thailand
MODIS AOD (2020-2024)	Optical depth at 0.55 µm	Unitless	Daily	GEE
Sentinel-5P: CO (2020-2024)	CO column number density	mol/m ²	Daily	GEE
Sentinel-5P: NO ₂ (2020-2024)	NO ₂ column number density	mol/m ²	Daily	GEE

2.3 Data processing

A multi-phase methodology was used to estimate PM_{2.5} in Chiang Mai Province and parts of

Northern Thailand, employing remote sensing data and ground-based observations within a cloud-based framework. GEE served as the primary platform for

data extraction and satellite data processing, while Python (Google Colab) was used for regression model development.

The initial phase involved daily extraction of satellite data from 2020 to 2024. In the GEE environment, we retrieved AOD at 0.55 μm from the MCD19A2 product and CO and NO₂ column number densities from Sentinel-5P (TROPOMI). These variables were sampled at nine PCD PM2.5 ground monitoring stations, resulting in a daily paired dataset of satellite-derived atmospheric properties and ground-measured PM2.5 observations across Chiang Mai and neighboring provinces, with days containing missing AOD or TROPOMI data excluded. The data were then exported as CSV files for model training.

The second phase focused on regression model development. An MLR model was developed in Google Colab to quantify the relationship between ground-measured PM2.5, AOD, CO, and NO₂. Predictor variables (AOD, CO, and NO₂) were used in their raw satellite-derived units to maintain physical interpretability and ensure compatibility with operational satellite inputs. Model performance was evaluated using R² and root mean square error (RMSE). The resulting coefficients were applied within the GEE environment to generate PM2.5 maps. Following the statistical framework established in previous satellite-based air quality studies (Hands Schuh et al., 2022; Kaewmesri et al., 2024), the estimation model was expressed as:

$$\text{Estimated PM2.5} = \beta_0 + \beta_1 \cdot \text{AOD} + \beta_2 \cdot \text{CO} + \beta_3 \cdot \text{NO}_2 \quad (1)$$

Where; estimated PM2.5 represents the predicted PM2.5 concentration. β_0 is the intercept, representing the baseline PM2.5 concentration when all satellite-derived variables are zero. AOD, CO, and NO₂ are the AOD, CO, and NO₂ column number densities, respectively, obtained from satellite observations. β_1 , β_2 , and β_3 are the regression coefficients for AOD, CO, and NO₂, respectively.

The final phase involved daily PM2.5 estimation and monthly averaging. Using GEE, the regression model was applied to every pixel across the study area to generate daily estimates from 2020 to 2024. Model performance was evaluated against PCD ground data using R² and RMSE. To mitigate short-term variability and data discontinuities in daily satellite measurements, daily estimates were averaged into monthly maps. The analysis focused on the “burning period” (January-April), corresponding to

peak biomass-burning activity. Seasonal averages were calculated using GEE’s mean reducer function for all pixels within Chiang Mai Province. While this study presents maps covering adjacent areas for regional comparison, statistical values were calculated exclusively within Chiang Mai. Future research will use these data to analyze PM2.5 drivers and assess the associated health impacts in Chiang Mai Province.

3. RESULTS AND DISCUSSION

3.1 Statistical analysis

A total of 5,377 valid daily records combining AOD, CO, and NO₂ satellite data with ground-measured PM2.5 daily data were obtained for analysis, with dates containing any missing values excluded. Correlation analysis demonstrated statistical relationships ($p < 0.001$) between ground PM2.5 and all three predictive variables. CO showed the strongest positive correlation with PM2.5 ($r = 0.82$), followed by AOD ($r = 0.72$) and NO₂ ($r = 0.32$) (Figure 2). This finding is consistent with the results of Amnuaylojaroen (2022), who demonstrated a strong correlation between CO and PM2.5 levels in Chiang Mai Province. To ensure the stability of the regression model, variance inflation factors (VIFs) were calculated to detect potential multicollinearity among the predictor variables. The VIF values for AOD, CO, and NO₂ were 3.36, 3.47, and 1.10, respectively, all well below the critical threshold of 5 (Hair et al., 2019), indicating negligible multicollinearity and reliable model coefficients.

The synchronization between satellite-derived precursors and ground-level PM2.5 is strongest during intense burning episodes (February-April). This high temporal covariance is attributed to the co-emission of aerosols and trace gases from biomass-burning activity, which is the dominant source during this period (Junpen et al., 2018; Paluang et al., 2024). Furthermore, stable atmospheric conditions and low planetary boundary layer height (PBLH) during the dry season simultaneously trap these pollutants, reinforcing their correlation. In contrast, this temporal coupling weakens during the wet season due to precipitation scavenging (washout) and a shift toward traffic-dominated emissions, for which the ratio of PM2.5 to gaseous precursors differs from that associated with biomass burning.

In the multiple regression analysis, the 5,377 observations were divided into two subsets: 80% (4,301 observations) for model training and 20% (1,076 observations) for testing and validation. A

regression model was generated by iteratively adding and removing predictor variables and evaluating performance using R² and RMSE for each dataset

combination. The regression results are presented in Table 2.

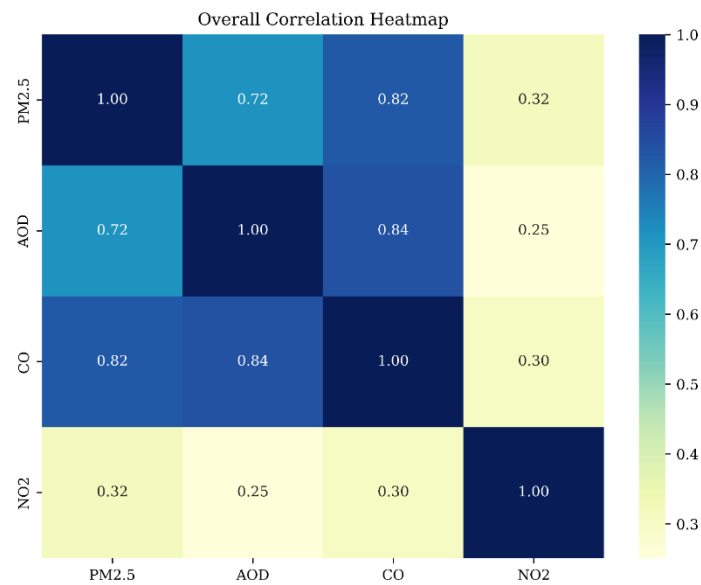


Figure 2. Correlation matrix between PM2.5, AOD, CO, and NO₂

Table 2. Result of regression analysis for different dataset combinations

Variables	R ²	RMSE	p
AOD	0.48	25.30	0.00
AOD+CO	0.65	20.83	0.00
AOD+NO ₂	0.50	24.86	0.00
CO+NO ₂	0.65	20.68	0.00
AOD+CO+NO ₂	0.66	20.66	0.00

The combined dataset of AOD, CO, and NO₂ showed the best performance, with R² and RMSE of 0.66 and 20.66 µg/m³, compared with 0.48 and 25.30 µg/m³ for AOD alone, respectively (Table 2). This result indicates an improvement when incorporating TROPOMI data to estimate ground-based PM2.5. This

finding is consistent with research conducted by Amnuaylojaroen (2022), Han et al. (2022), and Son et al. (2023), who employed Sentinel-5P satellite data for PM2.5 estimation. The estimation model is formally expressed as:

$$\text{Estimated PM2.5} = (12.48 \cdot \text{AOD}) + (1862 \cdot \text{CO}) + (87188.97 \cdot \text{NO}_2) - 54.51 \quad (2)$$

Where; estimated PM2.5 represents the predicted PM2.5 concentration (µg/m³). AOD is the aerosol optical depth (unitless), and CO and NO₂ are the total vertical column densities in moles per square meter (mol/m²). All three predictor variables contribute positively to the estimation of PM2.5 concentrations. The relatively large coefficients for CO and NO₂ reflect their small value ranges rather than greater importance. Therefore, due to differences

in variable scales and units, coefficient magnitudes should not be interpreted as indicators of variable importance. These findings are consistent with previous studies (Lalitaporn and Boonmee, 2019; Tan et al., 2024), which also highlight the relevance of trace gases in satellite-based PM2.5 estimation.

To assess model robustness, a temporal validation was conducted by training the model on 2020-2023 data to estimate PM2.5 concentrations for

2024. The results confirmed the superior performance of multi-variable integration (AOD, CO, NO₂) over the AOD-only approach. The integrated model achieved R² of 0.73 and RMSE of 17.37 µg/m³,

whereas the single-variable model yielded a lower R² of 0.58 and a higher RMSE of 21.687 µg/m³, demonstrating a significant reduction in overestimation (Figure 3).

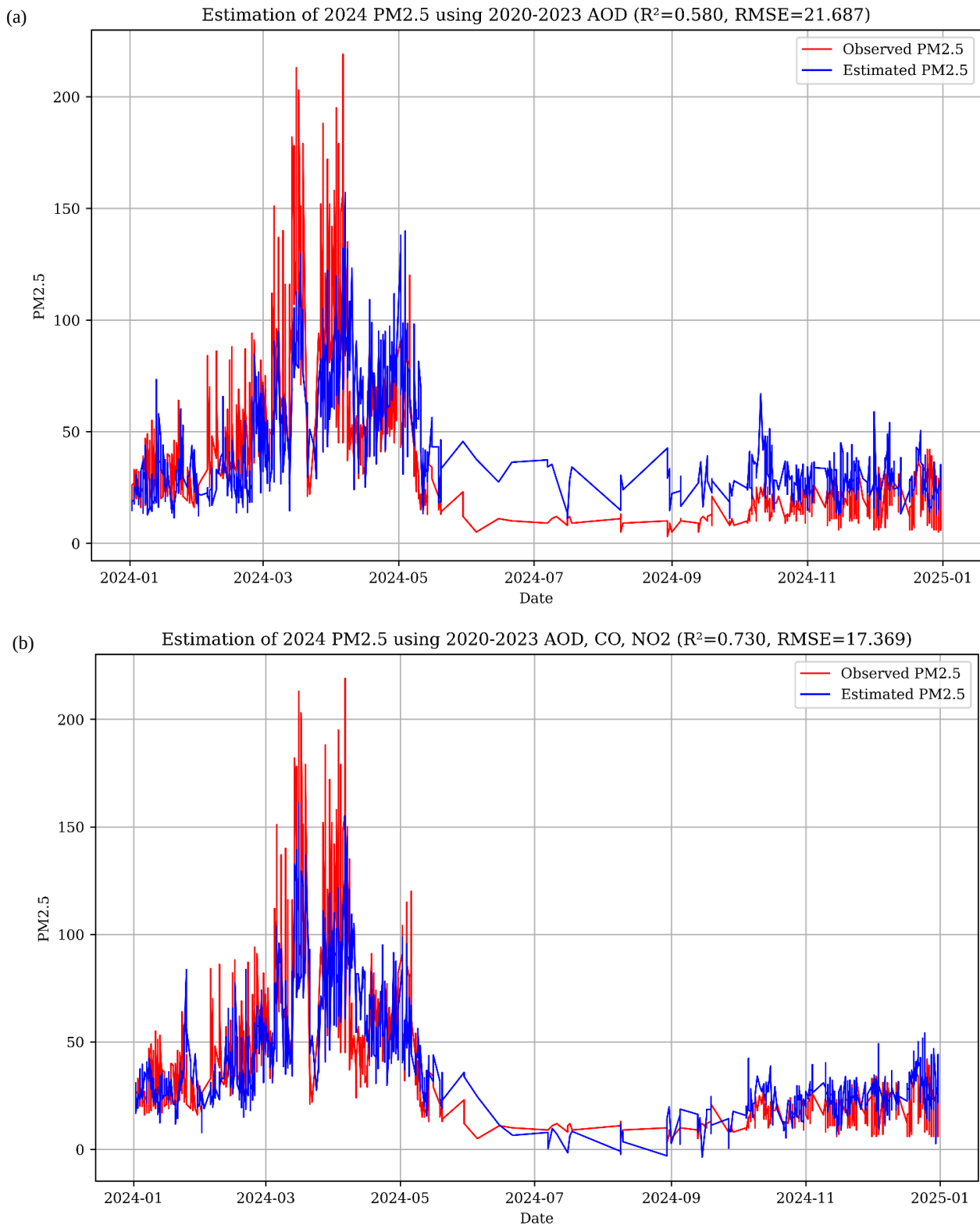


Figure 3. Comparison of ground-observed and estimated PM_{2.5} concentrations (µg/m³) in 2024, based on model trained with 2020-2023 data using (a) AOD only and (b) a combination of AOD, CO, and NO₂

3.2 Estimation of ground PM2.5 in GEE

Within the GEE environment, Equation 2 was applied to estimate daily PM2.5 concentrations (2020-2024). Validation against ground stations (35T and 36T) demonstrated robust overall performance ($r=0.88$, $R^2=0.77$, $RMSE=14.60 \mu g/m^3$). Analysis of the regression parameters ($y=0.91x + 5.55$) revealed range-specific behavior: the positive intercept (+5.55) indicates minor overestimation at low concentrations ($<25 \mu g/m^3$), likely due to surface reflectance interference, while the slope of 0.91 suggests slight underestimation at high concentrations ($>100 \mu g/m^3$), consistent with AOD saturation. Despite these biases, the model maintains high accuracy in the moderate range ($25\text{-}100 \mu g/m^3$), where most observations occur (Figure 4).

To further investigate bias patterns, a residual analysis was performed (Figure S1). The plot reveals a distinct funnel-shaped pattern indicative of heteroscedasticity, with prediction-error variance increasing with PM2.5 concentration. At lower estimated values ($<25 \mu g/m^3$), the residuals are tightly clustered with a slight negative skew, consistent with the model's tendency to overestimate under clean conditions. In contrast, higher concentrations exhibit significantly larger residuals in both directions. While the regression slope (0.91) suggests a general underestimation trend, the residual analysis highlights that prediction uncertainty is markedly higher during peak pollution events.

Compared with results from previous studies in the region, our MLR approach utilizing TROPOMI precursors (CO , NO_2) demonstrates competitive accuracy. While some machine learning-based studies have reported slightly higher correlations ($R^2>0.80$) (Son et al., 2023; Wongnakae et al., 2023), our model offers greater interpretability. For example, Buakhao (2023) used ground-based AERONET AOD in Chiang Mai and reported comparable R^2 values (0.76-0.81) but lower RMSE ($\sim 9.5 \mu g/m^3$). While their ground-based method offers higher precision at specific locations, our satellite-based approach achieves similar predictive performance ($R^2=0.77$) with the added benefit of region-wide spatial coverage, enabling pollution mapping in unmonitored border areas. These results confirm the model's robustness for estimating ground-level PM2.5 in areas with limited ground monitoring.

The implementation of this study within GEE proved to be an effective analytical strategy, enhancing data accessibility and processing efficiency. Its cloud-based computing infrastructure enabled rapid processing of large satellite datasets, establishing it as a powerful platform for environmental monitoring (Ghasempour et al., 2021; Rirugchart et al., 2025). The capacity to retrieve and archive extensive historical air pollution data is another key advantage, enabling historical evaluation in areas without ground-based sensors and supporting comprehensive long-term air quality assessments (Holloway et al., 2021; Wongnakae et al., 2023).

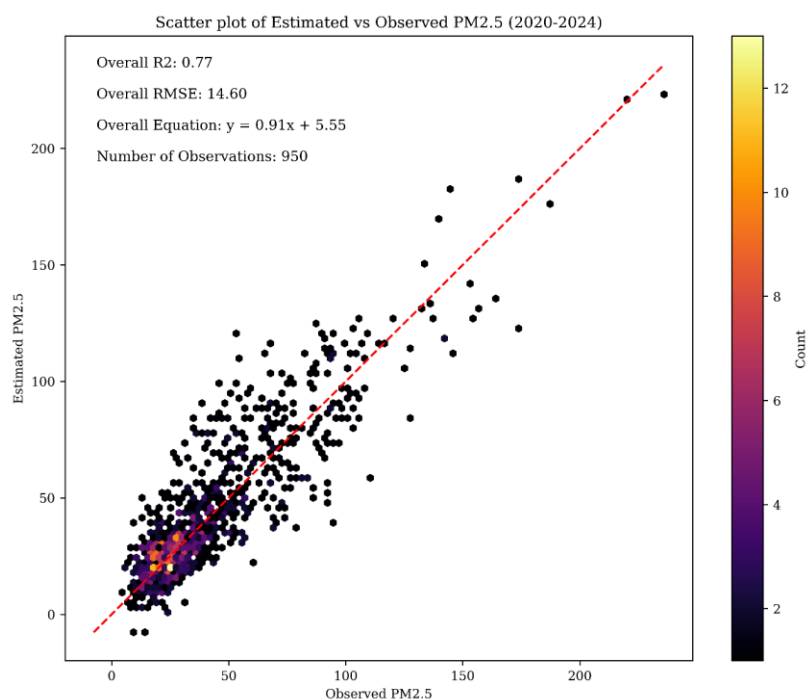


Figure 4. Scatter plot comparing estimated and observed PM2.5

A significant difference in PM_{2.5} estimation model performance was observed between the burning (January-April) and non-burning (May-December) periods. During the burning period, the model demonstrated high accuracy ($R^2=0.72$, RMSE=17.31 $\mu\text{g}/\text{m}^3$; Figure 5(a)), whereas performance declined substantially during the non-burning period ($R^2=0.26$, RMSE=8.32 $\mu\text{g}/\text{m}^3$; Figure 5(b)). This difference highlights the model's stronger predictive capacity

during periods of intensive biomass burning. Consequently, subsequent mapping and analysis focused exclusively on the burning period, when air pollution most strongly affects regional public health, while maps for the non-burning period were excluded due to poor model performance. In addition, extensive missing data caused by cloud cover and rainfall further limited estimation reliability (Son et al., 2023). These issues represent key limitations of the present study.

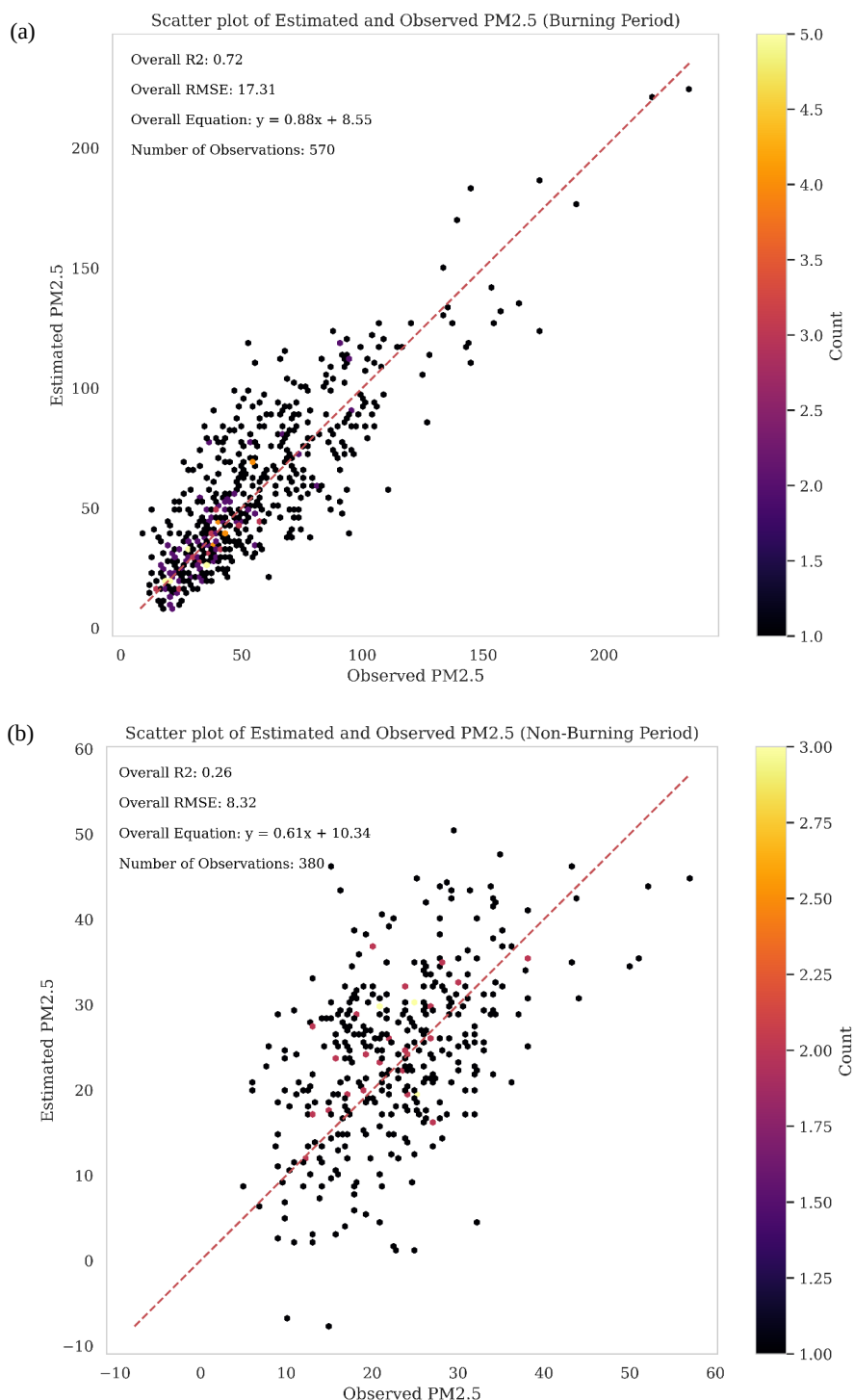


Figure 5. Scatter plot comparing observed and estimated PM_{2.5}: (a) burning period and (b) non-burning period

3.3 Spatio-temporal pattern of PM_{2.5} over Chiang Mai and parts of Northern Thailand

Monthly average PM_{2.5} maps for January–April from 2020 to 2024 are presented in Figures 6 and 7. January consistently exhibited low concentrations across Chiang Mai Province, with mean values of 24.44, 27.11, 11.12, 17.30, and 16.74 $\mu\text{g}/\text{m}^3$, respectively, all within the PCD's standard for good air quality. PM_{2.5} concentrations increased markedly during March and April, particularly in the northern and northeastern regions, where mountainous terrain limits atmospheric ventilation. This temporal surge aligns with Paluang et al. (2024), who identified that forest fire emissions in Upper Northern Thailand reach their peak intensity during these two months. Border districts, such as Chai Prakan and Mae Ai, were particularly affected by both domestic and transboundary biomass burning (Figures 6 and 7). This pattern is consistent with Amnuaylojaroen et al.

(2020) and Chansuebsri et al. (2024), whose WRF-HYSPLIT simulations indicated that transboundary transport from neighboring countries contributes more than domestic sources to PM_{2.5} levels in Northern Thailand.

Figures 6 and 7 show a consistent spatial pattern, with higher values observed in the central and northern parts of Chiang Mai Province. The central part, which includes the provincial center (Mueang District) and several densely populated districts, is the primary residential and urban area. These findings indicate a substantial air pollution burden in urbanized zones, highlighting potential health risks to the local population (Figures 8 and S2). By contrast, the western and southwestern parts of Chiang Mai Province consistently exhibited lower PM_{2.5} concentrations. Specifically, districts such as Mae Chaem, Omkoi, and Galyani Vadhana showed relatively low average PM_{2.5} concentrations throughout the study period.

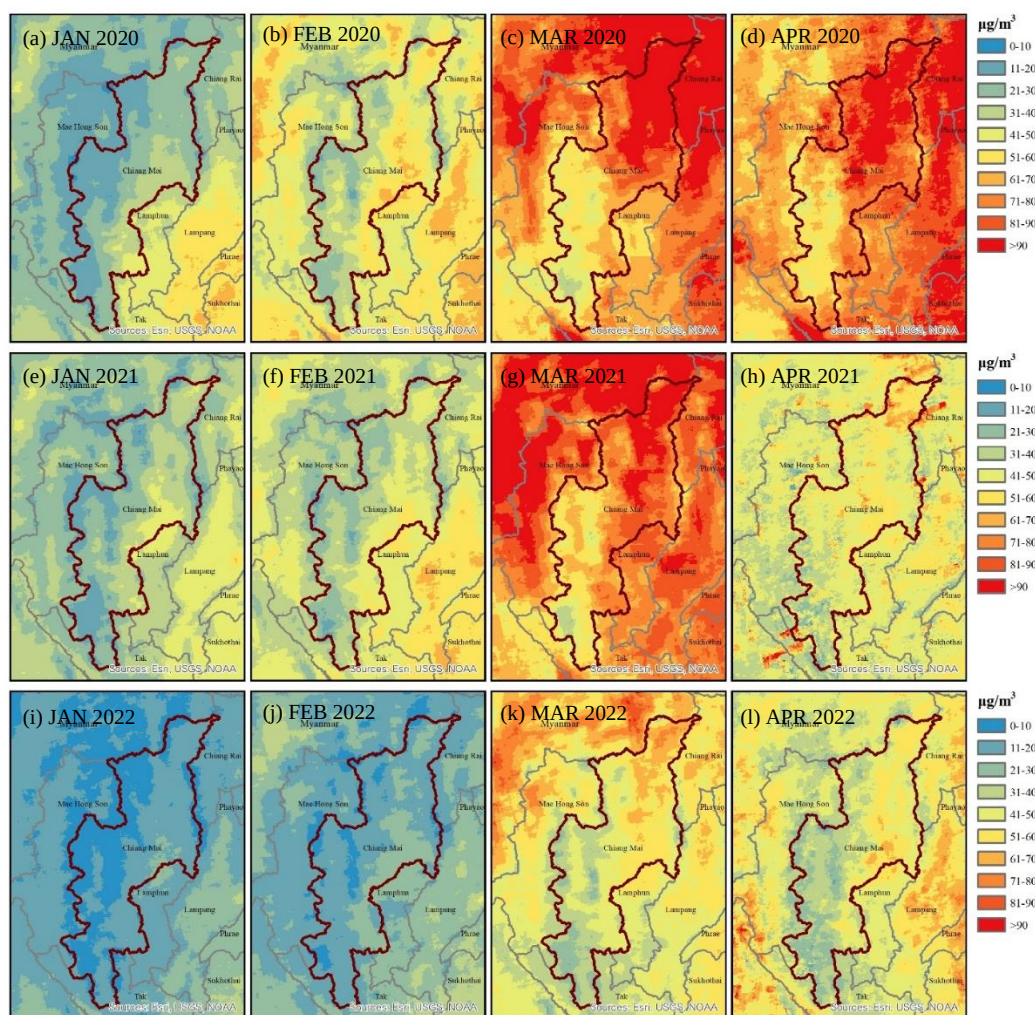


Figure 6. Monthly average satellite-derived PM_{2.5} concentrations for 2020–2022: (a) January 2020, (b) February 2020, (c) March 2020, (d) April 2020, (e) January 2021, (f) February 2021, (g) March 2021, (h) April 2021, (i) January 2022, (j) February 2022, (k) March 2022, and (l) April 2022

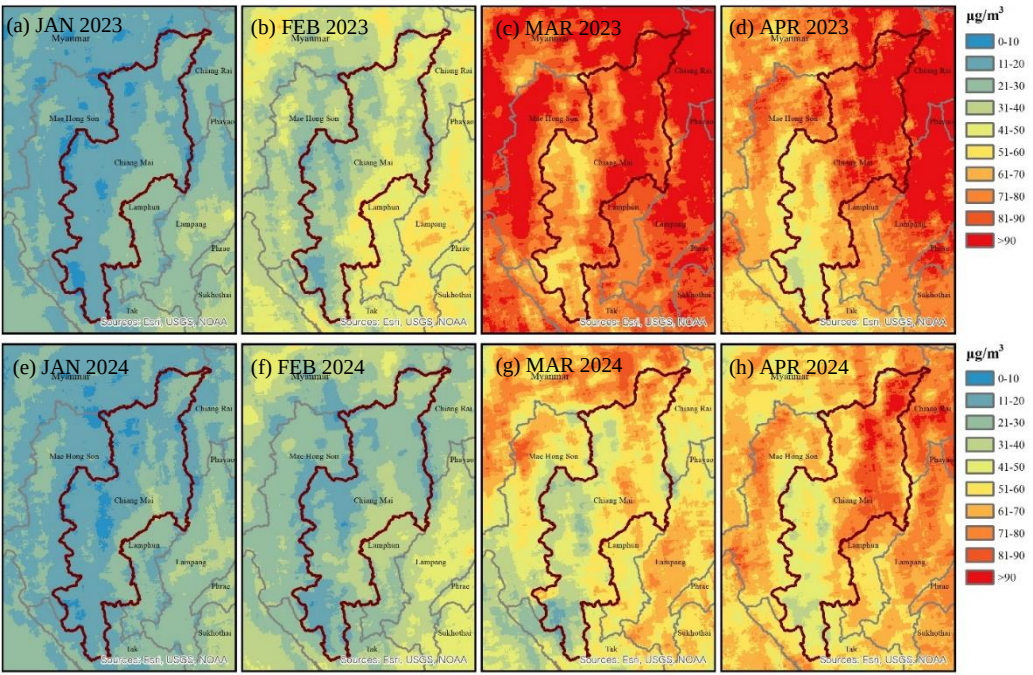


Figure 7. Monthly average satellite-derived PM_{2.5} concentrations for 2023 and 2024: (a) January 2023, (b) February 2023, (c) March 2023, (d) April 2023, (e) January 2024, (f) February 2024, (g) March 2024, and (h) April 2024

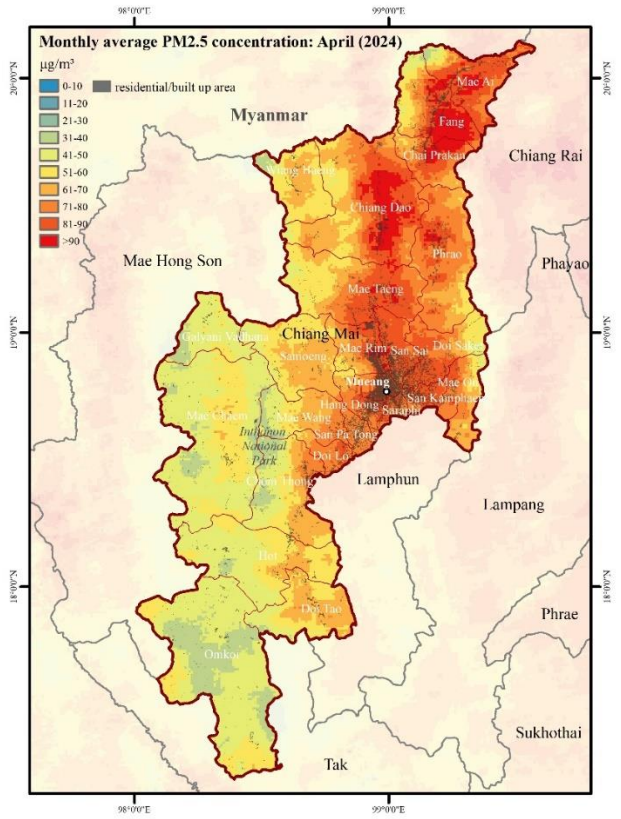


Figure 8. Overlay of average PM_{2.5} concentration (April 2024) on residential areas, illustrating potential population exposure during the burning season

Figure 9 presents monthly minimum, maximum, mean, and standard deviation (SD) values of satellite-derived PM_{2.5} concentrations during the burning period (2022-2024) in Chiang Mai. While January generally shows low air pollution levels, concentrations increase sharply by March,

highlighting the influence of seasonal biomass burning on regional air quality (Lalitaporn and Boonmee, 2019). The highest monthly mean, $79.18 \pm 16.90 \mu\text{g}/\text{m}^3$, was recorded in March 2023, whereas the lowest monthly mean, $11.12 \pm 4.18 \mu\text{g}/\text{m}^3$, was in January 2022. The maximum value, $194.81 \mu\text{g}/\text{m}^3$, was recorded in April 2021.

High SD values in March and April indicate considerable spatial and temporal variability in air pollution levels. This variation may be due to differences in local emission sources and topographical influences. For example, areas in the

southwestern area of Chiang Mai Province generally exhibit lower PM_{2.5} levels, whereas the central and northeastern regions show notably higher concentrations (Figures 6-8). Variations in meteorological conditions and emission sources may result in different adverse health effects across the province (Sapbamrer et al., 2024), requiring further investigation at provincial and regional scales. To illustrate the spatial distribution and provide a long-term overview of air pollution during the burning period, PM_{2.5} maps for 2020-2024 are presented in Figure 10.

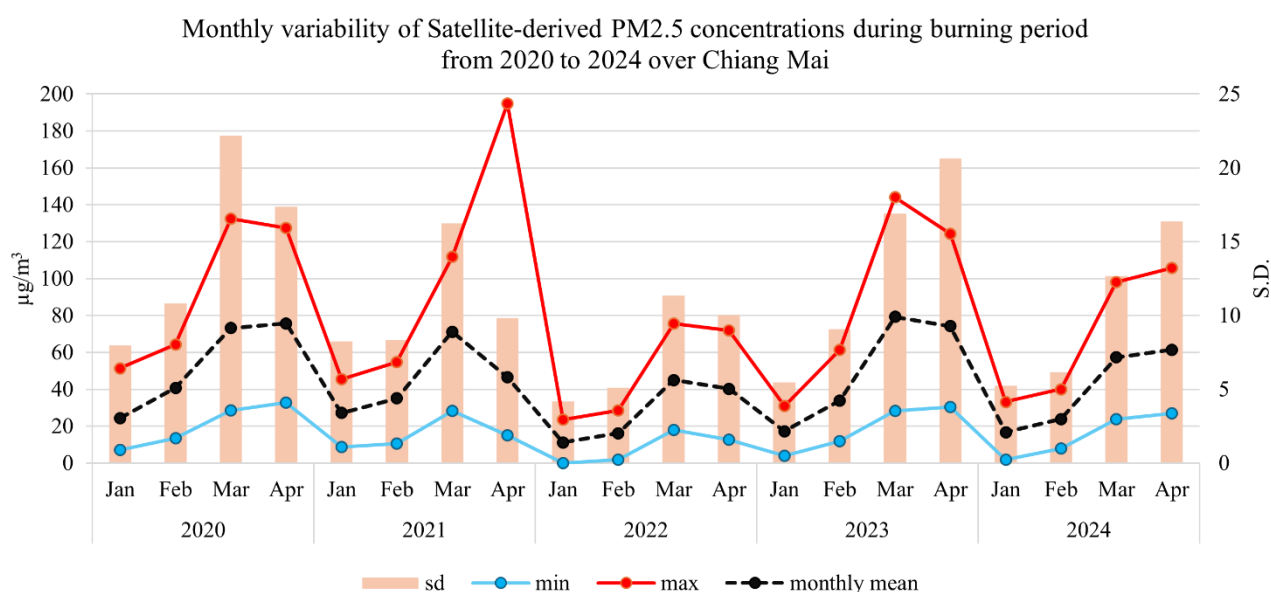


Figure 9. Descriptive statistics (minimum, maximum, mean, and standard deviation) of satellite-derived PM_{2.5} concentrations ($\mu\text{g}/\text{m}^3$) over Chiang Mai (January-April, 2020-2024)

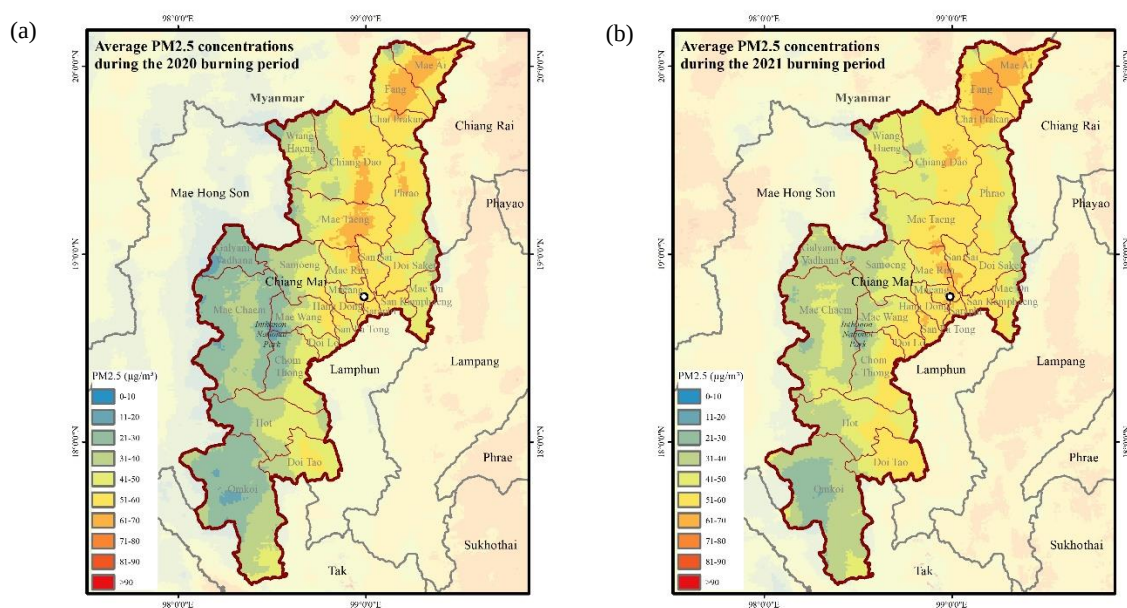


Figure 10. Average PM_{2.5} during the burning period: (a) 2020, (b) 2021, (c) 2022, (d) 2023, and (e) 2024

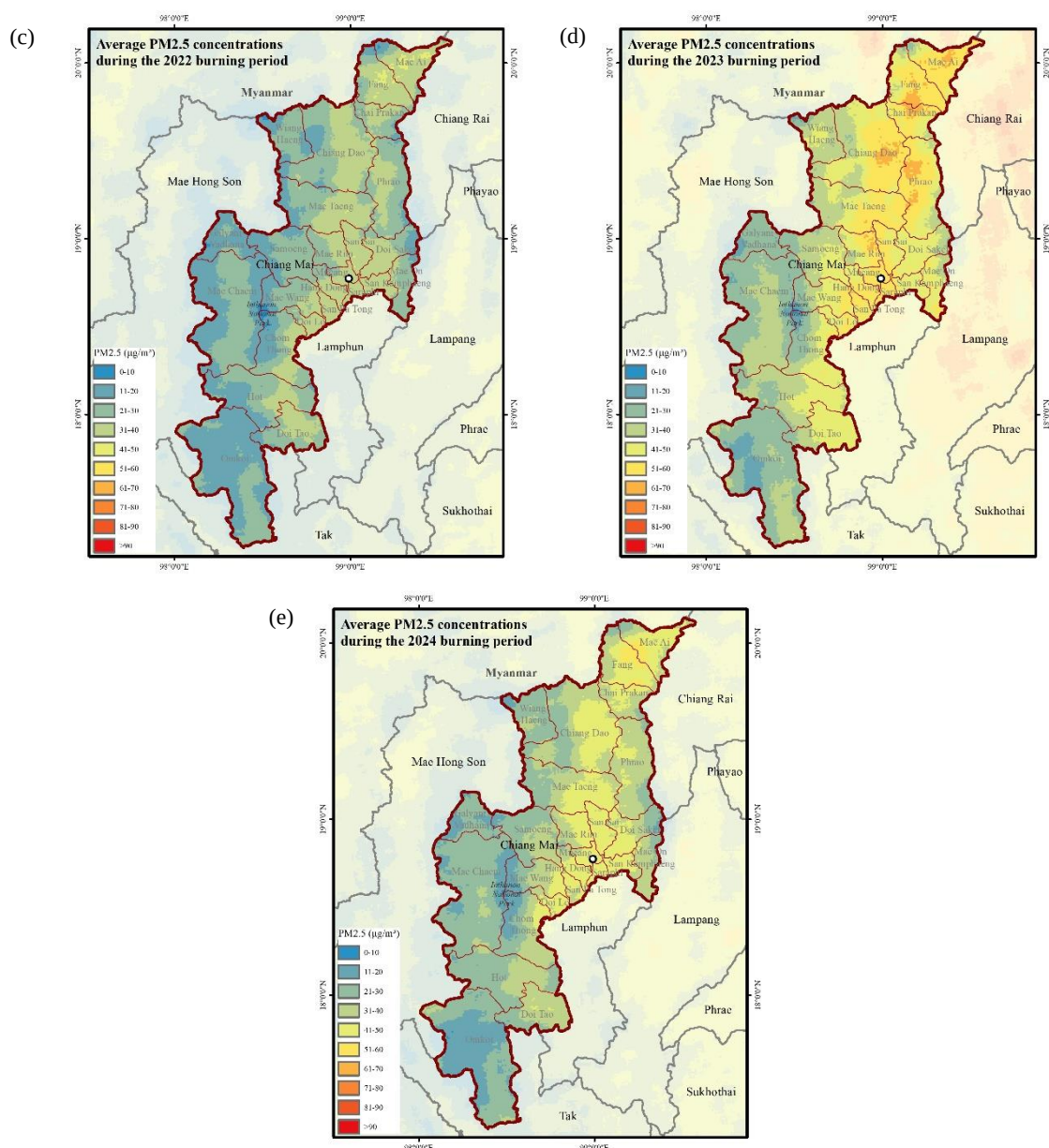


Figure 10. Average PM2.5 during the burning period: (a) 2020, (b) 2021, (c) 2022, (d) 2023, and (e) 2024 (cont.)

3.4 Policy implications and transboundary pollution challenges

The PM2.5 crisis constitutes a complex structural challenge driven by transboundary and socioeconomic factors rather than a singular seasonal event. Intense hotspots in northern border districts, such as Mae Ai, align spatially with active fires in Myanmar, consistent with [Amnuaylojaroen et al. \(2020\)](#)'s finding that westerly winds facilitate the long-range transport of biomass burning aerosols into Northern Thailand. Persistent border hotspots suggest that pollution is due to local burning and transboundary transport within a shared "airshed," where pollutants are trapped regardless of national borders. Effective management requires a dual

strategy: enforcing "haze-free" trade standards for agricultural imports to address supply-chain root causes under the ASEAN Haze Agreement and scientifically delineating the northern ASEAN airshed for joint monitoring.

To mitigate domestic emissions effectively, a holistic strategy encompassing rural agricultural management and urban traffic control is required. Initiatives like the Mae Chaem Model exemplify successful integrated resource management, with [Kamton et al. \(2019\)](#) reporting a continuous decline in forest fires, corroborated by the lower PM2.5 levels detected in this study. However, enforcement-based "zero burning" policies require re-evaluation. [Suriyawong et al. \(2023\)](#) argue that total restrictions

on burning can result in leaf-litter accumulation, leading to severe wildfires; thus, “prescribed burning” for fuel management is recommended alongside socioeconomic solutions such as agricultural carbon credit markets and high-value alternative crops. Chansuebsri et al. (2024) identified traffic emissions as the dominant source in urban contexts, contributing 76% of PM_{2.5} during the non-smoke-haze period. Consequently, policy initiatives must prioritize expanding public transportation and electric vehicles to mitigate vehicular emissions year-round.

Structurally, management frameworks should transition toward a decentralized “northern haze sandbox” model while empowering local authorities with integrated budgets. This requires a professional agency utilizing big data, such as this study’s high-resolution maps, for precise year-round planning. Finally, public health agencies must use these findings to prioritize resource allocation, including N95 masks and clean air rooms, for vulnerable populations in the unmonitored border districts identified in this study.

3.5 Limitations

While this study effectively mapped PM_{2.5} spatial distributions, it has several limitations. First, satellite data continuity is compromised by missing values, particularly during cloud cover or precipitation. Moreover, the coarse temporal resolution of satellite observations (typically one or two measurements per day) may not adequately capture the full diurnal variability of PM_{2.5} concentrations (Kaewmesri et al., 2024). This temporal sparsity can introduce inaccuracies, potentially biasing daily average estimations.

Second, the current model architecture excludes meteorological variables (e.g., wind speed, relative humidity, PBLH) for operational simplicity and to minimize dependency on external datasets. While this omission introduces uncertainty regarding aerosol hygroscopic growth and vertical dispersion, this study assumes that satellite-derived column densities (AOD, CO, NO₂) implicitly integrate these effects; for example, stagnant atmospheric conditions directly result in pollutant accumulation that is captured by the model inputs.

Third, a large performance disparity was observed between seasons. While the integration of AOD and TROPOMI data proved robust during the burning period ($R^2=0.72$), accuracy decreased markedly during the non-burning period ($R^2=0.26$). To quantify this disparity, an independent t-test on

absolute prediction errors confirmed a statistically significant difference ($t=11.19$, $p<0.001$), with a mean absolute error of $13.48 \mu\text{g}/\text{m}^3$ during the burning season, compared with $7.17 \mu\text{g}/\text{m}^3$ in the non-burning season. This reduced predictive capability, verified by statistical divergence, is likely due to weaker physical correlations between ground-level PM_{2.5} and biomass burning proxies (CO, AOD) during the wet season. Specifically, precipitation scavenging and the shift in dominant emission sources reduce the strength of the direct relationship between satellite-derived column densities and surface concentrations. Consequently, the current MLR model is not recommended for year-round operational monitoring. To address this limitation, future studies could develop season-specific models or employ nonlinear machine learning approaches to better capture variability during low-pollution periods.

Finally, validating high-resolution estimates remains a primary challenge. This study used $1 \times 1 \text{ km}^2$ MAIAC AOD to enhance spatial precision, an approach shown by Mei et al. (2019) to improve estimation accuracy. However, the validation process was constrained by the sparse distribution of ground monitoring stations, particularly the limited number of reference-grade stations within Chiang Mai. To mitigate this spatial limitation, the training dataset included stations from neighboring provinces to capture broader regional aerosol characteristics, ensuring the model learned generalized patterns. Additionally, model robustness was reinforced through temporal validation (predicting 2024 data), confirming predictive accuracy despite spatial constraints. Nevertheless, the lack of monitors in remote border areas restricts full validation of the pollution hotspots identified by the satellite model.

For future work, the recently launched GEO-KOMPSAT-2B satellite, equipped with the Geostationary Environment Monitoring Spectrometer, offers higher temporal resolution data, potentially addressing diurnal variability issues and enhancing research performance. Additionally, the authors plan to investigate the associations between PM_{2.5} and meteorological parameters (precipitation, temperature, relative humidity, wind speed, and PBLH) and socioeconomic factors to identify the drivers of PM_{2.5} in Chiang Mai. Finally, future studies could explore substituting AOD data with TROPOMI-derived data when AOD is unavailable, as Son et al. (2023) indicated that TROPOMI data alone can provide accurate estimations.

4. CONCLUSION

This study characterized the spatio-temporal variation of satellite-derived PM_{2.5} in Northern Thailand (2020-2024) by integrating TERRA-MODIS AOD with Sentinel-5P CO and NO₂ data within the GEE platform. Analysis revealed a consistent annual increase in pollution during the burning period (January-April), peaking in March. Seasonal average concentrations ranged from 25.18 µg/m³ (2022) to 51.33 µg/m³ (2020), highlighting notable inter-annual variability.

The model achieved robust performance ($r=0.88$, $R^2=0.77$, $RMSE=14.60$ µg/m³), confirming that combining gas species data (CO, NO₂) significantly enhances estimation accuracy compared with AOD-only methods. Distinct from the complex machine learning approaches increasingly used in Thailand, this study demonstrates that a physically interpretable regression framework enriched with biomass burning tracers offers competitive accuracy with simpler, cloud-based operational implementation.

Future research should incorporate meteorological (e.g., wind speed, temperature) and socioeconomic parameters to further refine model precision. Based on the identification of intense border hotspots, policymakers are recommended to prioritize installing ground monitors and allocating healthcare resources to unmonitored northern districts. Overall, this work highlights the potential of multi-source satellite integration to fill data gaps in areas with limited ground monitoring, providing a crucial baseline for assessing the health impacts of air pollution on local populations.

ACKNOWLEDGEMENTS

This research was conducted as part of a Ph.D. project at the Department of Geography, Faculty of Social Sciences, Chiang Mai University. This study was funded by the Teaching Assistant/Research Assistant Scholarship from Chiang Mai University. The authors express their gratitude to the Pollution Control Department of Thailand for providing ground-based PM_{2.5} data. We also acknowledge the satellite data providers: the NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC) for AOD data, and the European Space Agency (ESA) Copernicus Sentinel Hub for CO and NO₂ data. Finally, we thank the Google Earth Engine development team for providing the high-performance cloud-computing platform that facilitated the data analysis for this research. The datasets generated during the current study are available from the corresponding author on reasonable request.

AUTHOR CONTRIBUTIONS

Conceptualization, W.S. and A.C.; Methodology, W.S., A.C, P.H. and W.T.; Validation, W.S. and A.C.; Formal Analysis, W.S.; Data Curation, W.S.; Writing-Original Draft Preparation, W.S.; Writing-Review and Editing, W.S., P.H., W.T. and A.C.; Visualization, W.S.; Supervision, A.C.; Project Administration, W.S.

DECLARATION OF CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Amnuaylojaroen T. Prediction of PM_{2.5} in an urban area of Northern Thailand using multivariate linear regression model. *Advances in Meteorology* 2022;1:Article No. 3190484.
- Amnuaylojaroen T, Inkom J, Janta R, Surapipith V. Long range transport of Southeast Asian PM_{2.5} pollution to Northern Thailand during high Biomass burning episodes. *Sustainability* 2020;12(23):Article No. 10049.
- Bran SH, Macatangay R, Surapipith V, Chotamonsak C, Chantara S, Han Z, et al. Surface PM_{2.5} mass concentrations during the dry season over Northern Thailand: Sensitivity to model aerosol chemical schemes and the effects on regional meteorology. *Atmospheric Research* 2022;277:Article No. 106303.
- Buakhao L. Model for estimating PM_{2.5} concentration using aerosol optical depth data in the Muang District of Chiang Mai Province. *YRU Journal of Science and Technology* 2023;8(1):50-8 (in Thai).
- Buya S, Gokon H, Huynh VN, Dam HC, Usanavasin S, Karnjana J, et al. Spatiotemporal association between monthly PM_{2.5} levels and cardiorespiratory mortality in Thailand (2015-2019). *International Journal of Environmental Health Research* 2026;36(1):41-52.
- Chankaew K, Sinitkul R, Manuyakorn W, Roekworachai K, Kamalaporn H. Spatial estimation of PM_{2.5} exposure and its association with Asthma Exacerbation: A prospective study in Thai Children. *Annals of Global Health* 2022;88(1):Article No. 15.
- Chansuebsri S, Kolar P, Kraistititkul P, Kantarawilawan N, Yabueng N, Wiriya W, et al. Chemical composition and origins of PM_{2.5} in Chiang Mai (Thailand) by integrated source apportionment and potential source areas. *Atmospheric Environment* 2024;327:Article No. 120517.
- Chen B, Song Y, Kwan MP, Huang B, Xu B. How do people in different places experience different levels of air pollution? Using worldwide Chinese as a lens. *Environmental Pollution* 2018;238:874-83.
- Chiang Mai Provincial Statistical Office. Chiang Mai Provincial Statistical Report: 2024 [Internet]. 2024 [cited 2025 Jun 10]. Available from: <https://chiangmai.nso.go.th/reports-publications/provincial-statistics-report.html>.
- Fatima M, Butt I, Nasar-u-Minallah M, Atta A, Cheng G. Assessment of Air Pollution and its association with population health: Geo-statistical evidence from Pakistan. *Geography, Environment, Sustainability* 2023;16(2):93-101.
- Ghasempour F, Sekertekin A, Kutoglu SH. Google earth engine based spatio-temporal analysis of air pollutants before and during the first wave COVID-19 outbreak over Turkey via

- remote sensing. *Journal of Cleaner Production* 2021;319:Article No. 128599.
- Hair JF, Black WC, Babin BJ, Anderson RE. *Multivariate Data Analysis*. 8th ed. Andover, Hampshire: Cengage Learning EMEA; 2019.
- Han S, Kundhikanjana W, Towashiraporn P, Stratoulis D. Interpolation-based fusion of Sentinel-5P, SRTM, and regulatory-grade ground stations data for producing spatially continuous maps of PM_{2.5} concentrations nationwide over Thailand. *Atmosphere* 2022;13(2):Article No. 161.
- Handschoh J, Erbertseder T, Schaap M, Baier F. Estimating PM_{2.5} surface concentrations from AOD: A combination of SLSTR and MODIS. *Remote Sensing Applications: Society and Environment* 2022;26:Article No. 100716.
- Holloway T, Miller D, Anenberg S, Diao M, Duncan B, Fiore AM, et al. Satellite monitoring for air quality and health. *Annual Review of Biomedical Data Science* 2021;4(1):417-47.
- Junpen A, Pansuk J, Kamnoet O, Cheewaphongphan P, Garivait S. Emission of air pollutants from rice residue open burning in Thailand, 2018. *Atmosphere* 2018;9(11):Article No. 449.
- Kaewmesri P, Uamkasem B, Koedkurang K, Chalermpong P, Sriwilas P, Tupbamroong J, et al. Investigating the relationship between particulate matter (PM_{2.5}) concentration and aerosol optical depth (AOD) over Thailand during the dry period. *IOP Conference Series: Earth and Environmental Science* 2024;1412:Article No. 012004.
- Kamton R, Satienerakul K, Yotapakdee T, Nunthasen K. Haze-relate air pollution and impacts on healthy in Chiang Mai province. *Thai Interdisciplinary and Sustainability Review* 2019;8(1):265-73 (in Thai).
- Lalitaporn P, Boonmee T. Analysis of tropospheric nitrogen dioxide using satellite and ground based data over Northern Thailand. *Engineering Journal* 2019;23(6):19-35.
- Mei L, Strandgren J, Rozanov V, Vountas M, Burrows JP, Wang Y. A study of the impact of spatial resolution on the estimation of particle matter concentration from the aerosol optical depth retrieved from satellite observations. *International Journal of Remote Sensing* 2019;40(18):7084-112.
- Pardthaisong L, Sin-ampol P, Suwanprasit C, Charoenpanyanet A. Haze pollution in Chiang Mai, Thailand: A road to resilience. *Procedia Engineering* 2018;212:85-92.
- Paluang P, Thavorntam W, Phairuang W. The Spatial-temporal emission of air pollutants from biomass burning during haze episodes in Northern Thailand. *Fire* 2024;7(4):Article No. 122.
- Pathakoti M, Muppalla A, Hazra S, Venkata MD, Lakshmi KA, Sagar VK, et al. Measurement report: An assessment of the impact of a nationwide lockdown on air pollution: A remote sensing perspective over India. *Atmospheric Chemistry and Physics* 2021;21(11):9047-64.
- Pollution Control Department. *Thailand Air and Noise Pollution Situation and Management Report, 2023* [Internet]. 2023 [cited 2025 May 15]. Available from: http://air4thai.com/tagoV2/tago_file/books/book_file/eef68f0fe6195accf1e101ff3c204fe0.pdf.
- Rirugchart P, Losiri C, Sitthi A. The Application of google earth engine on PM_{2.5} estimation and its distribution pattern in Saraburi Province, Thailand. *International Journal of Geoinformatics* 2025;21(1):26-42.
- Sapbamrer P, Assavanopakun P, Panumasvivat J. Decadal trends in ambient air pollutants and their association with COPD and lung cancer in Upper Northern Thailand: 2013-2022. *Toxics* 2024;12(5):Article No. 321.
- Sakti AD, Anggraini TS, Ihsan KTN, Misra P, Trang NTQ, Pradhan B, et al. Multi-air pollution risk assessment in Southeast Asia region using integrated remote sensing and socio-economic data products. *Science of the Total Environment* 2023;854:Article No. 158825.
- Son R, Stratoulis D, Kim HC, Yoon JH. Estimation of surface PM_{2.5} concentrations from atmospheric gas species retrieved from TROPOMI using deep learning: Impacts of fire on air pollution over Thailand. *Atmospheric Pollution Research* 2023;14(10):Article No. 101875.
- Song J, Hong X, Yu K, He B, Wu S, Hu K, et al. Spatiotemporal estimation and analysis of PM_{2.5} concentrations in Wuhan utilizing multisource remote sensing data and NO_x as inputs for machine learning models. *IEEE Sensors Journal* 2025;25(4):6812-24.
- Song Y, Huang B, He Q, Chen B, Wei J, Mahmood R. Dynamic assessment of PM_{2.5} exposure and health risk using remote sensing and geo-spatial big data. *Environmental Pollution* 2019;253:288-96.
- Suriyawong P, Chuetor S, Samae H, Piriyaakarnsakul S, Amin M, Furuuchi M, et al. Airborne particulate matter from biomass burning in Thailand: Recent issues, challenges, and options. *Heliyon* 2023;9(3):e14261.
- Tan H, Chen Y, Mao F, Wilson JP, Zhang T, Cui X, et al. PM_{2.5} estimation and its relationship with NO₂ and SO₂ in China from 2016 to 2020. *International Journal of Digital Earth* 2024;17(1):Article No. 2398055.
- Wongnakae P, Chitchum P, Sripramong R, Phosri A. Application of satellite remote sensing data and random forest approach to estimate ground-level PM_{2.5} concentration in Northern region of Thailand. *Environmental Science and Pollution Research* 2023;30(38):88905-17.