Title	多時期衛星画像の時空間データ分析によるタイ王国のPM2.5濃度の時系列変化推定技術の開発と心肺系死亡への影響評価
Author(s)	SUHAIMEE, BUYA
Citation	
Issue Date	2025-09
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/20063
Rights	
Description	Supervisor: 郷右近 英臣, 先端科学技術研究科, 博士

Doctoral Dissertation

Development of a Spatiotemporal Analysis Method Using Multi-temporal Satellite Imagery to Estimate Temporal Variations in PM2.5 Concentrations and Evaluate Their Impact on Cardiorespiratory Mortality in Thailand

Suhaimee Buya

Supervisor: Hideomi Gokon

Graduate School of Advanced Science and Technology
Japan Advanced Institute of Science and Technology
(Knowledge Science)

September 2025

Abstract

Air pollution, specifically fine particulate matter (PM2.5), is a critical global health concern, contributing to an estimated 4.2 million premature deaths annually. PM2.5 exposure is strongly linked to elevated risks of cardiorespiratory diseases, including cardiovascular conditions, chronic respiratory illnesses, and lung cancer. Thailand's urban and industrialized areas, in particular, face deteriorating air quality, with PM2.5 concentrations often exceeding national and international safety standards. Despite extensive documentation of air pollution's health effects, a significant research gap persists in understanding the long-term mortality impacts of PM2.5 exposure across diverse Thai regions.

This study investigates the relationship between PM2.5 pollution and cardiorespiratory mortality in Thailand from January 2015 to December 2019. Utilizing high-resolution satellite-based PM2.5 data, mortality records, and population statistics, the research employs advanced statistical methodologies, including Poisson regression modeling and Moran's I spatial analysis, to assess regional variations in air pollution and its association with mortality rates.

Findings indicate a significant correlation between PM2.5 exposure and increased cardiorespiratory mortality, with the highest risks observed in the central and northern regions, which experience the most severe pollution levels. Seasonal analysis reveals peak mortality rates during the dry season (November to April), coinciding with heightened air pollution from biomass burning, industrial emissions, and meteorological conditions that worsen pollutant accumulation. Specifically, monthly PM2.5 concentrations above $30 \, \mu g/m^3$ are associated with a 1%-7% increase in mortality risk, while levels below $20 \, \mu g/m^3$ correlate with a 3%-6% reduction.

These results underscore the urgent need for comprehensive air quality management strategies, stricter emission controls, and enhanced monitoring systems to mitigate air pollution's health impacts. Furthermore, spatial analysis identifies high-risk areas, emphasizing the necessity of targeted policy interventions and region-specific mitigation measures. This study contributes to the growing evidence supporting the implementation of more stringent air pollution regulations and public health initiatives to safeguard vulnerable populations in Thailand.

Acknowledgments

Contents

List of Figures

List of Tables

List of Symbols/Abbreviations

PM2.5	Particulate Matter 2.5 microns
WHO	World Health Organization
PCD	Pollution Control Department
BAQ	Bangkok's Air Quality and Noise Management Division
NAAQS	National Ambient Air Quality Standards
AOD	Aerosol Optical Depth
LST	Land Surface Temperature
NDVI	Normalized Difference Vegetation Index
EV	Elevation
WOY	Week of the Year
MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
EOSDIS	Earth Observing System Data and Information System
DAAC	Distributed Active Archive Center
MLR	Multiple Linear Regression
RF	Random Forest
XGBoost	eXtreme Gradient Boosting
SVM	Support Vector Machines
R^2	The coefficient of determination
RMSE	Root Mean Square Error
CI	Confidence Interval
NHSO	National Health Security Office
ICD-10	International Classification of Diseases-10
r	The correlation coefficient
CM	Cardiorespiratory mortality

Chapter 1
Introduction

1.1 Background and Rationale

Air pollution remains a critical global issue, impacting both urban and rural communities and contributing to an estimated 4.2 million premature deaths each year, according to the World Health Organization (WHO) [1]. A major cause of these deaths is exposure to PM2.5, or particulate matter with a diameter of 2.5 microns or less, which is known to lead to severe cardiorespiratory diseases and cancer [2] (Figure 1.1). Alarmingly, around 91% of these premature deaths occur in low- and middle-income countries, with Southeast Asia and the Western Pacific regions bearing the heaviest burden. Thailand, in particular, faces an exceedingly high mortality rate from air pollution-related diseases, underscoring the necessity for localized studies to better understand and tackle this urgent public health challenge.

Figure 1.1: Discover PM2.5 perspectives.

Extensive epidemiological research has consistently demonstrated a strong association between chronic PM2.5 exposure and an increased risk of cardiovascular and non-accidental mortality [3], [4], [5]. Elevated PM2.5 levels have also been linked to higher rates of morbidity and mortality related to cardiorespiratory diseases [6], [7], [8]. While numerous studies have documented the health and economic consequences of air pollution [9], [10], [11], PM2.5 is recognized as particularly harmful compared to larger particulate matter particles [12], [13], [14] (Figure 1.2). However, research in this area remains limited due to inconsistencies in monitoring systems and the underrepresentation of typical populations in health studies, especially in countries like Thailand.
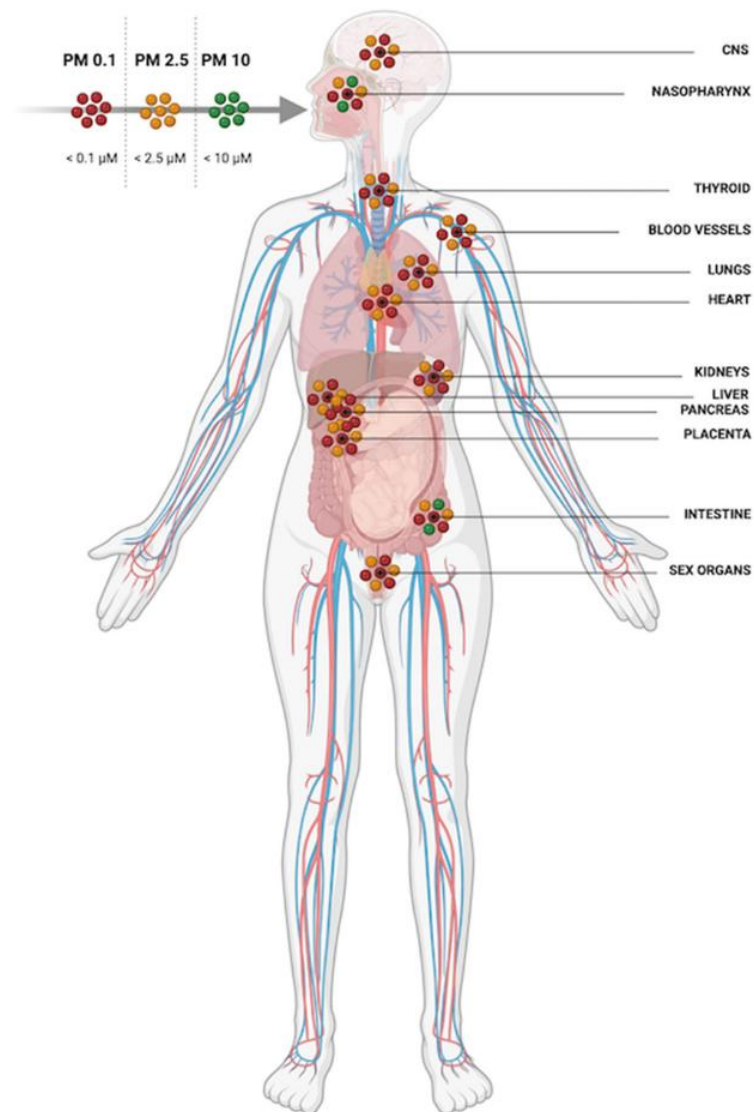
Figure 1.2: The impact of PM2.5 on the cardiorespiratory system [15].

Despite substantial evidence regarding the acute effects of PM2.5, including hospitalizations and symptom exacerbation, there is less emphasis on the long-term impacts of chronic exposure on mortality. Understanding these long-term effects is vital for grasping the cumulative burden of air pollution on public health and the necessity for preventive policies. In Thailand, inconsistent air quality monitoring and limited geographic coverage further impede accurate assessments (Figure 1.3). Therefore, this study aims to address these gaps by examining long-term PM2.5 exposure and its connection to mortality using advanced methods and broader datasets.
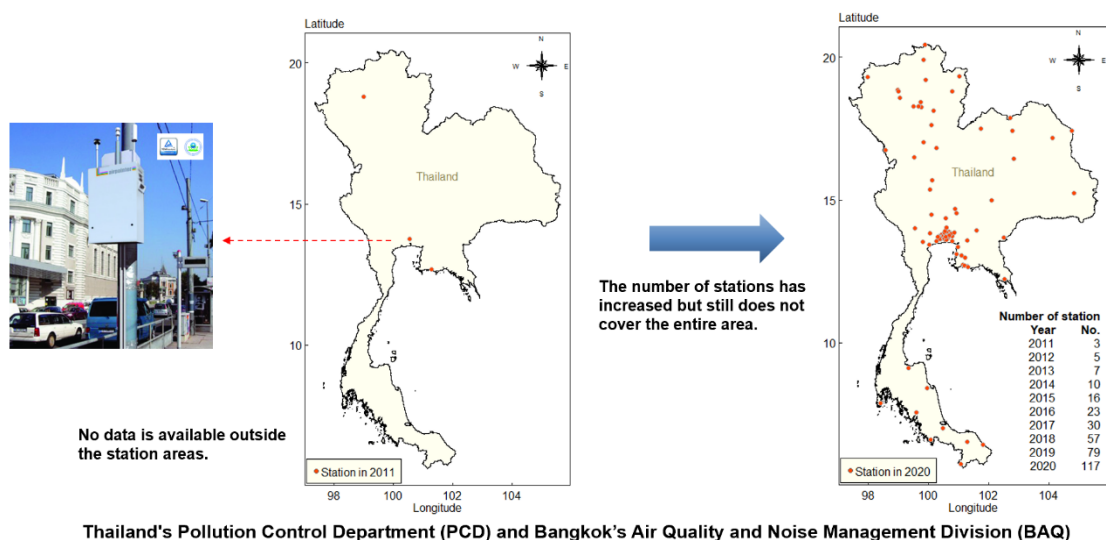
Figure 1.3: Coverage of monitoring stations in Thailand (2011–2020).

1.2 Objectives and Novelty

1.2.1 Objectives of this study

This study aims to estimate PM2.5 concentrations using high-resolution satellite data, analyze their spatiotemporal distribution and long-term impact on cardiorespiratory mortality across Thailand, and provide evidence-based recommendations for targeted public health interventions.

1.2.2 Novelty of this study

This research is among the first comprehensive attempts in Thailand to integrate **satellite-derived PM2.5 estimations with cardiorespiratory mortality data** at a spatiotemporal scale. The novelty lies in:

- Applying **satellite data with advanced modeling** to overcome limitations of sparse ground monitoring stations.

- Incorporating **spatial and temporal epidemiological methods** to capture regional disparities in air pollution impacts.
- Providing **evidence-based insights specific to Thailand**, where limited studies have linked PM2.5 exposure to health outcomes with this level of detail.

By combining remote sensing, epidemiology, and statistical modeling, this study contributes a new framework for **air pollution–health linkage assessment in regions with limited monitoring infrastructure.**

1.3 Literature Review
1.3.1 Estimation of PM2.5 Using Satellite Data

Previous studies on the estimation of PM2.5 using satellite data have utilized a variety of models, but most have selected only one [16]. This study summarizes the relevant work in Table 1.1 to compare model performance. The studies encompass different areas, periods, and factor variables from various data sources. Meteorological factors can influence the dispersion and transportation of small particulate matter; land use and cover can indicate potential PM2.5 sources and areas of concern, and time variables can elucidate seasonal variations and long-term trends in PM2.5 levels. Many studies have employed aerosol optical depth (AOD) as a foundational factor and incorporated additional variables to enhance model performance in predicting PM2.5 using satellite data [17], [18]. However, the significance of factors has varied among studies. Certain analyses, such as the study by Joharestani et al. (2019) in Iran, discovered that satellite-derived AODs could not enhance model performance [19].

Table 1.1 also presents the performance of different models, with the random forest (RF) model displaying a high coefficient of determination (R^2) in three experiments, and the extreme gradient boosting (XGBoost) model showing a high R^2 in two studies. However, it is important to note that the RF model performed similarly to the XGBoost model. Among other machine learning (ML) models, multiple linear regression (MLR) exhibited the lowest accuracy. Despite this, MLR is still widely employed for its simplicity and practicality. Estimating PM2.5 concentrations is a challenge due to the numerous variables that can influence them. As noted by Wei et al. (2019), ML has gained popularity as a method for addressing intricate problems

because it can identify and utilize multiple independent factors that can impact the predicted variable [20].

Table 1.1: Related works for estimation of PM2.5 and compared models' performance.

Description	Factor variables	Models	R² (RMSE (μg/m³))		
			Training	Validation	Testing
Xiao et al. (2018), Area: China, Study period: 2013-2017 [21]	AOD (10 km), EV, PD, RH, PRE, PBL, AAI, and NO2	XGBoost	-	0.78 (21)	-
		RF	-	0.77 (22)	-
		GAM	-	0.65 (28)	-
		LEM+GAM	-	0.64 (27)	-
Xu et al. (2018), Area: Canada, Study period: 2003-2010 [22]	AOD (3 km), Vapor, LST, NDVI, Albedo, PBL, WS, EV, DO, and Month	MLR	-	0.22 (3.24)	-
		BRNN	-	0.31 (3.04)	-
		SVM	-	0.30 (3.13)	-
		LASSO	-	0.24 (3.20)	-
		MARS	-	0.31 (3.05)	-
		RF	-	0.49 (2.67)	-
		XGBoost	-	0.46 (2.71)	-
		Cubist	-	0.48 (2.64)	-
Wei et al. (2019), Area: China, Study period: 2015-2016 [20]	AOD (1 km), NDVI, LUC, NTL, TEM, PRE, ET, RH, SP, PBL, WS, and WD	MLR	0.41 (20.04)	0.41 (20.04)	0.38 (21.97)
		GWR	0.60 (22.83)	0.53 (23.28)	0.44 (26.47)
		RF	0.98 (6.40)	0.81 (17.91)	0.53 (28.09)
Joharestani et al. (2019), Area: Iran, Study period: 2015-2018 [19]	AOD (3 km), TEM, RH, RN, Visibility, WS, SWS, SP, and DP	RF	-	0.66 (15.30)	-
		XGBoost	-	0.67 (15.15)	-
		DNN	-	0.63 (15.89)	-
Danesh Yazdi et al. (2020), Area: UK, Study period: 2005-2013 [23]	AOD (1 km), PD, Cloudiness, BP, WD, DP, TEM, LUC, DW, DH, PBL, NDVI, traffic counts, SDY, CDY, DOY, NOD, APM, Year, NTL, EV, DNR, LM, NB, DNB, ABH, and NBC.	RF	-	0.830 (4.28)	-
		GBM	-	0.826 (4.33)	-
		NN	-	0.793 (4.73)	-
		KNN	-	0.791 (4.72)	-

Factor variable PD: Population density; PRE: Precipitation; PBL: Planetary boundary layer height; AAI: Absorbing aerosol index data in visible light; WS: Wind speed; DO: Distance to ocean; LUC: Land use cover; NTL: Nighttime lights; TEM: Temperature; ET: Evaporation; RH: Relative humidity; SP: Surface pressure; WD: Wind direction; RN: Rainfall; SWS: Sustained wind speed; DP: Dew point; BP: Barometric pressure; DW: Distance to water; DH: Distance to Heathrow airport; SDY: Sine of day of the year; CDY: Cosine of day of the year; DOY: Day of week; NOD: Number of days from time of origin; APM: Average daily PM2.5; DNR: Distance to nearest major road; LM: Length of major road in grid cell; NB: Number of bus stops in grid cell; DNB: Distance to nearest bus stop; ABH: Average building height; NBC: Number of buildings in the grid cell.

Models RMSE: Root Mean Square Error; GAM: Generalized Additive Model; LEM: Linear Mixed Effects Model; BRNN: Bayesian Regularized Neural Networks; SVM: Support Vector Machines; LASSO: Least Absolute Shrinkage and Selection Operator; MARS: Multivariate Adaptive Regression Splines; GWR: Geographically Weighted Regression; DNN: Deep Neural Network; GBM: Gradient Boosting Machine; NN: Neural Network; KNN: K-Nearest Neighbor.

Limited research has been conducted on predicting PM2.5 levels in Thailand using satellite data due to insufficient data from ground stations and satellites. In the Chiang Mai and central regions of Thailand, PM2.5 was estimated through MLR models with AOD at a 10 km resolution in investigations carried out by Kanabkaew (2013) and Phuengsamran & Lalitaporn (2021) [24]. These studies yielded R^2 values of 0.77 and 0.49 when incorporating meteorological parameters from monitoring stations, and 0.22 and 0.11 when not considering them (refer to Table 1.2). A study by Kumharn et al. (2022) in Bangkok, Thailand, using a linear mixed-effects model (LMEM) with only AOD (10 km), reported an R^2 of 0.64, while including other factors from monitoring stations increased the R^2 to 0.87 [25]. Another study in urban areas of Northern Thailand using MLR with AOD (3 km) and other factors reported R^2 values between 0.02 and 0.60 [26].

A review conducted by Chu et al. (2016) on articles predicting ground PM2.5 concentrations using satellite AOD revealed that MLR exhibited the lowest R^2 accuracy (less than 0.5) compared to other models [27]. The low R^2 values depicted above imply that including covariates such as meteorological factors, land use, cover, seasonal variables, etc., in MLR models necessitates further scrutiny [28]. Typically, R^2 values exceeding 0.6 are deemed acceptable for model accuracy. Hence, PM2.5 estimates in Thailand without incorporating meteorological factors are considered unreliable.

Table 1.2: Related works for estimation of PM2.5 using satellite data in Thailand.

Author (Year)	Study area	Satellite	Factor	Model
Kanabkaew (2013) [24]	Chiangmai, Thailand January-April 2007, and 2012	AOD (10 km)	TEM, RH	MLR 0.77[+S], 0.22[-S]
Phuengsamran & Lalitaporn (2021) [29]	Central, Thailand 2008-2018	AOD (10 km)	RH, TEM, WS, PBL	MLR 0.49 [+S], 0.11[-S]
Kumharn et al. (2022) [25]	Bangkok, Thailand 2020	AOD (10 km), NDVI	VRA, T, RH, WS, and HPBL	LMEM 0.87 [+S], 0.64[-S]
Amnuaylojaroen (2022) [26]	Urban Area of Northern 2020	AOD (3 km)	RH, TEM, WS, NO2, CO, SO2, and O3	MLR 0.02-0.60[+S]

±S: Factor from stations; VRA: ventilation rate.

1.3.2 Association of PM2.5 with Cardiorespiratory Disease

Recent studies in Thailand emphasize the varied sources and health impacts of air pollution. For instance, emissions from maritime activities near coastal industrial zones significantly contribute to PM10-bound organic and ionic species, with sulfate and aliphatic carbons being predominant [30]. Similarly, PM2.5-bound pollutants present health risks across different regions.

Table 1.3 presents a comparative summary of key studies, both international and Thai, focusing on PM2.5-related health impacts. Key findings from the literature include:

- A consistent association between PM2.5 exposure and increased mortality from cardiovascular, respiratory, and chronic diseases.
- Elevated health burdens in highly polluted regions, especially northern Thailand.
- Major contributing factors include biomass burning, urban emissions, and seasonal variations.
- Vulnerable populations include the elderly and those with pre-existing health conditions.
- Green spaces and urban greening are identified as potential mitigation strategies.

These insights highlight the need for localized policy actions and in-depth spatiotemporal analysis.

Table 1.3: Summary of PM2.5-Related Articles.

Title	Authors	Location	Key Findings	Health Impact
Spatio-Temporal PM2.5 & PM10 in Iran	Abadi et al. (2025) [31]	Tehran, Isfahan, Khuzestan (Iran)	High seasonal PM2.5/PM10 due to dust; health risks exceed WHO thresholds; ELCR values significant in Tehran	Cancer risk, respiratory issues, HQ > 1
Burden of Disease from Air Pollution in Thailand	Pinichka et al. (2017) [32]	Thailand (nationwide)	Estimated 650–38,410 fatalities; highest PAFs for NO2 and PM2.5	All-cause, cardiovascular, and lung cancer mortality
Air Pollution & Cardiorespiratory Morbidities in Southern Thailand	Buya et al. (2024) [33]	Southern Thailand	PM10/PM2.5 linked to asthma, COPD, CVD with time-lag effects	Increased inpatient admissions for multiple cardiorespiratory diseases
Short-Term Exposure to PM10 in Northern Thailand	Ngamsang et al. (2023) [34]	Northern Thailand	High March PM10 linked to 11–30% respiratory mortality reduction with mitigation	Respiratory mortality, public health risk
Effect of PM2.5 on NCD Mortality in Northern Thailand	Parasin & Amnuaylojaroen (2024) [35]	Northern Thailand	Strong correlations between PM2.5 and NCDs (e.g., hypertension, diabetes, stroke)	Elevated NCD mortality linked to biomass burning PM2.5
PM2.5 & Mortality in Bangkok	Fold et al. (2020) [36]	Bangkok, Thailand	Estimated 4,240 non-accidental deaths/year from PM2.5	Cardiopulmonary, lung cancer, non-accidental mortality
PM2.5 Premature Deaths & Greening (China)	Tang et al. (2025) [37]	Nanjing, China	Greening reduces PM2.5-related premature deaths; strong spatial trends	Respiratory and cardiovascular disease mortality
PM2.5 & CKD Mortality in Thailand	Leonetti et al. (2024) [38]	Thailand (77 provinces)	PM2.5 and components significantly associated with CKD mortality	CKD mortality, especially in Northeast Thailand

HQ: Hazard Quotient, ELCR: The Excess Lifetime Cancer Risk, PAF: Population-attributable fraction, COPD: Chronic Obstructive Pulmonary Disease, CVD: Cardiovascular Disease, NCD: Non-Communicable Diseases, CKD: Chronic Kidney Disease

1.4 Related Studies in Thailand

Research conducted in Thailand has identified several sources of PM2.5, including industrial emissions, vehicle exhaust, agricultural burning, and meteorological conditions [39] [40] [41] . For instance, elevated total suspended particulate matter (TSP) levels in industrial zones like Map Ta Phut highlight the urgent need for control measures. Furthermore, emissions from rubber plantations in northeastern Thailand significantly contribute to the formation of ozone and secondary aerosols.

Most PM2.5-mortality research in Thailand has depended on localized urban monitoring data since 2011. Studies such as Fold et al. (2020) and Pothirat et al. (2019) have primarily focused on Bangkok and Chiang Mai, respectively, while neglecting rural and suburban areas with distinct exposure patterns [36], [42]. Traditional spatial modeling techniques, such as ordinary kriging, have been employed to broaden monitoring coverage; however, they often lack the resolution necessary for detailed analysis [43].

High-resolution satellite data provides an alternative for evaluating long-term regional air pollution trends [44]. Despite its benefits, this method has seen limited use in Thailand for population health studies. This research seeks to address this gap by integrating high-resolution satellite PM2.5 data with cardiorespiratory mortality statistics to examine nationwide patterns and inform public health policy.

1.5 Proposed Research

Spatiotemporal epidemiology analyzes disease incidence and mortality over space and time to identify high-risk areas and associated environmental factors [45], [46], [47]. This study applies such methods to explore the relationship between elevated PM2.5 exposure and cardiorespiratory mortality in Thailand. It also investigates how mortality rates vary across regions and over time, providing insights that could guide health policies and interventions.

1.6 Thesis Structure

This thesis consists of five interconnected chapters that collectively examine the relationship between PM2.5 concentrations and cardiorespiratory mortality in Thailand:

- **Chapter 1 Introduction:** Presents the background, rationale, research objectives, and the importance of understanding the health impacts of PM2.5 in Thailand. This chapter also reviews previous studies related to PM2.5 estimation using satellite data and its association with cardiorespiratory diseases, with a particular focus on research conducted in Thailand.

- **Chapter 2 Materials and Methods:** Describes the study area, data sources—including PM2.5 satellite data, cardiorespiratory mortality, and population data—and the analytical methods used. These methods include descriptive statistics, Spearman's correlation, Poisson regression models, and Moran's I statistics to explore spatial and temporal relationships.

- **Chapter 3 PM2.5 with Satellite Data:** Details the development of the satellite-based PM2.5 estimation model at a high spatial resolution. It explains the selection of predictor variables, model construction, model evaluation, and spatial-temporal distribution of PM2.5 across Thailand.

- **Chapter 4 PM2.5 with Cardiorespiratory Mortality:** Presents the analysis of the relationship between estimated PM2.5 concentrations and cardiorespiratory mortality rates. This includes the application of Poisson regression models and spatial statistics to determine the association patterns across regions in Thailand.

- **Chapter 5 Summary:** Summarizes the key findings of the study, which discusses the estimation of PM2.5 using satellite data. It revealed a significant spatiotemporal association between PM2.5 concentrations and cardiorespiratory mortality in Thailand. The chapter concludes with implications for public health policy and recommendations for future research.

Chapter 2
Materials and Methods

This chapter delineated the study's geographical scope, elucidated the data sources utilized (which encompassed PM2.5 measurements, cardiorespiratory mortality statistics, and population demographics), and expounded upon the methods employed for data analysis, including Spearman's Correlation, the Poisson Regression Model, and Moran's I.

2.1 Study Area

Thailand is traditionally divided into four regions: central, north, northeast, and south. This administrative system, developed by the Ministry of Interior, is widely used for statistical and academic purposes [48]. These regions are further subdivided into 77 provinces. Thailand, located in Southeast Asia, is bordered by the Andaman Sea and the Gulf of Thailand. The country has a population of approximately 70 million and covers a total land area of 513,120 square kilometers (Figure 2.1).

Figure 2.1: The map of Thailand region.

2.2 Data Sources

This study utilized three datasets: PM2.5 data, cardiorespiratory mortality data, and population data. The PM2.5 data represents average monthly concentrations for each province, while the cardiorespiratory mortality and population data consist of monthly counts per province. These datasets were merged based on month and province codes, covering the period from January 1, 2015, to December 31, 2019. Each dataset is described in detail below.

2.2.1 Particulate Matter 2.5 Microns (PM2.5) Data

PM2.5 concentrations were estimated using a developed satellite-based model described by Buya et al. (2023), achieving a spatial resolution of 1-kilometer (km). The model incorporated multiple factors, including AOD, Land Surface Temperature (LST), Normalized Difference Vegetation Index (NDVI), Elevation (EV), Week of the Year (WOY), and year, to enhance accuracy. AOD served as the primary variable, while LST, NDVI, and EV accounted for meteorological conditions and land use, and WOY and year captured temporal variations.

1) Satellite data used to predict PM2.5

This investigation employed AOD, LST, NDVI, and EV data based on images retrieved from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite products. All data were retrieved from the National Aeronautics and Space Administration (NASA) Earth Observing System Data and Information System (EOSDIS) offered by the Distributed Active Archive Center (DAAC).

- Aerosol Optical Depth (AOD) data

The AOD data from the Terra and Aqua satellites in the MCD19A2 data product was processed, which includes the variable "Aerosol Optical Depth at 045 Microns" [49]. The daily AOD data has a spatial resolution of 1 km per pixel and is collected at 10:30 a.m. and 1:30 p.m. local standard time. The average AOD from 2011 to 2020 is presented in Figure 2.2. On average, the AOD yearly is higher in Thailand's central and northern parts. Also, it showed a high value in 2019 (0.31), while other years were between 0.17-0.24.
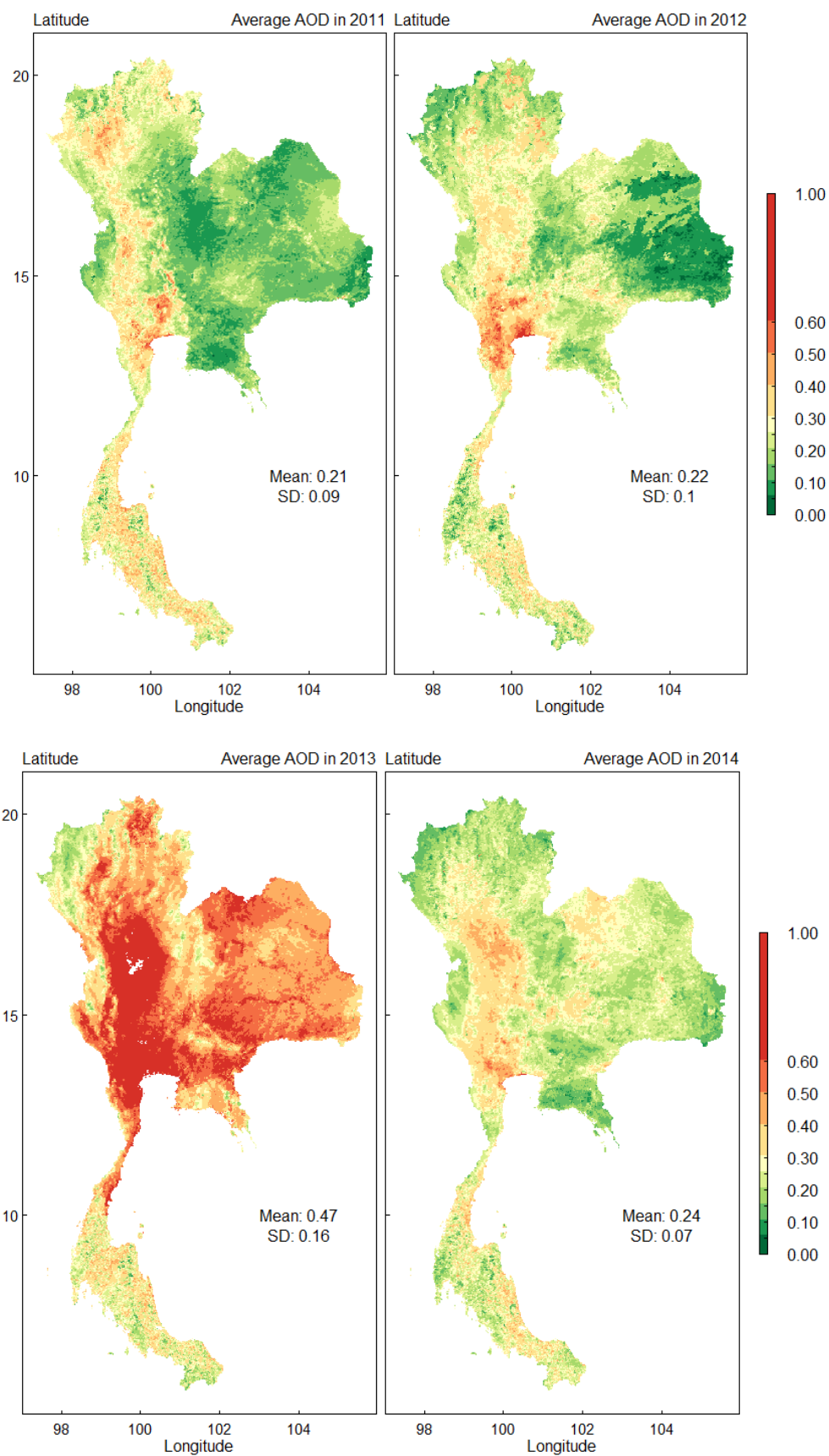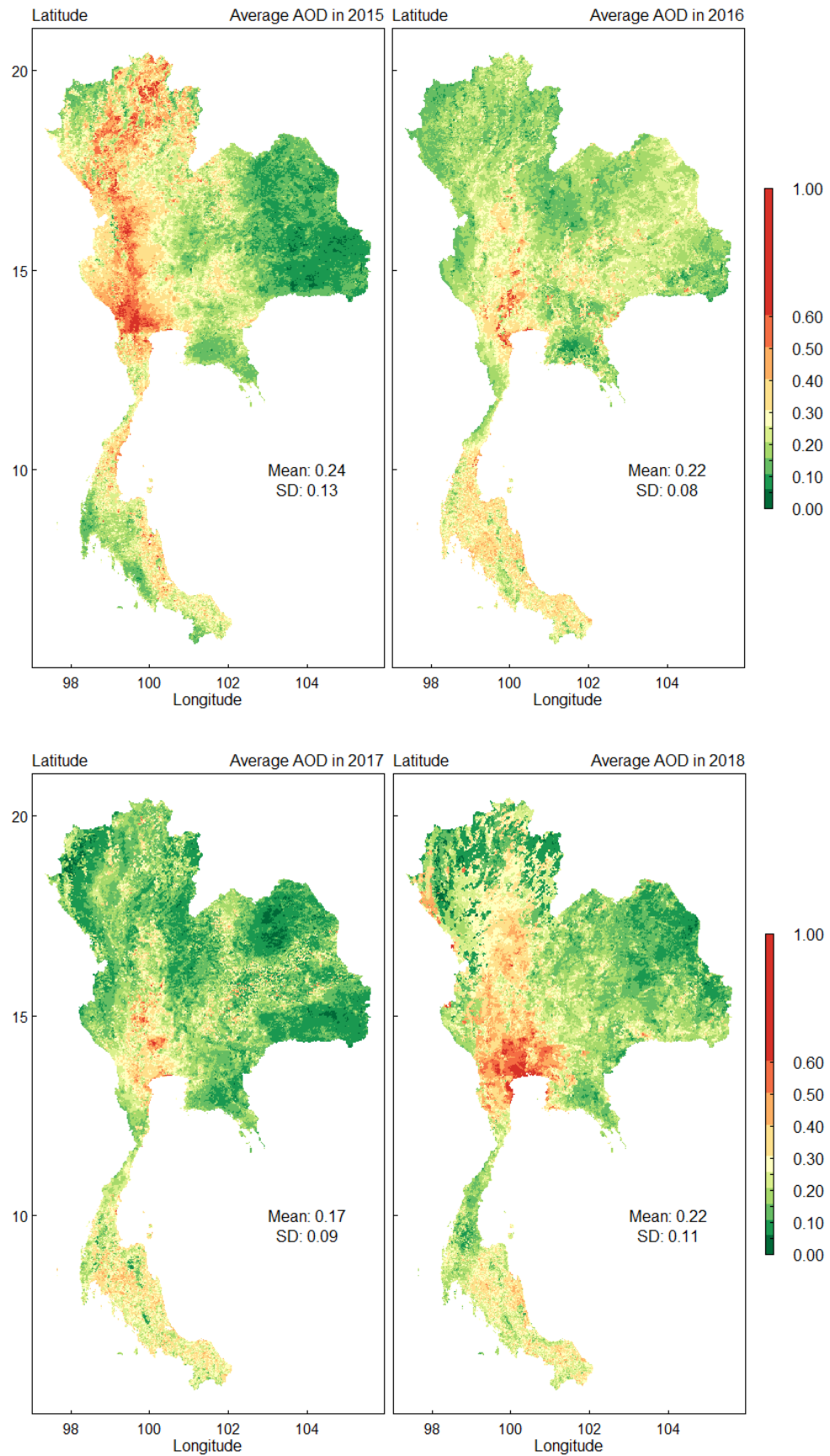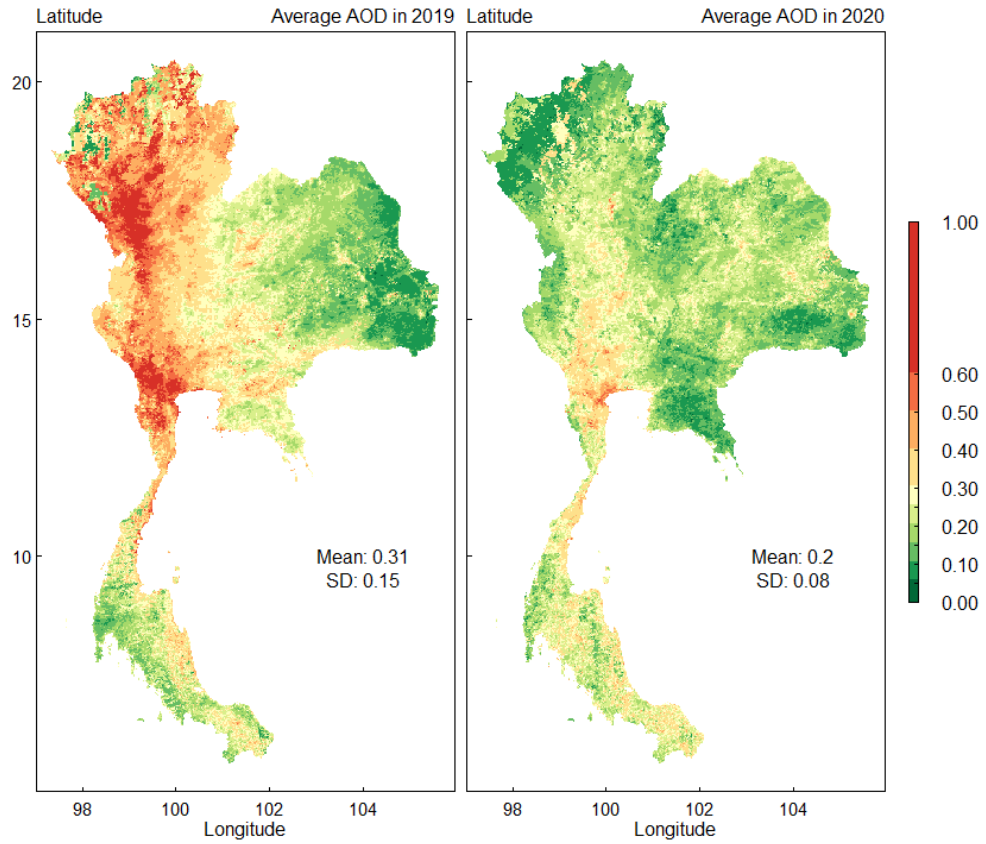
Figure 2.2: The average of AOD in 2011-2020.

- Land Surface Temperature (LST) data

    We utilized the LST data from the Terra and Aqua satellites. The LST data from Terra is a MOD11A1 product [50], while the LST data from Aqua is an MYD11A1 product [51]. We combined the LST measurements from both satellites to enhance the sample size. The daily average LST values were determined by calculating the arithmetic mean of the two satellite measurements. If data from only one satellite was available on a specific day, that data was used as the daily average LST value. Figure 2.3 presents the average LST from 2011 to 2020. On average, the LST has lowest in the northern part and highest in the central, northeast, and east parts of Thailand. The mean LST from 2015-2019 is between 26.56-29.15 °C.
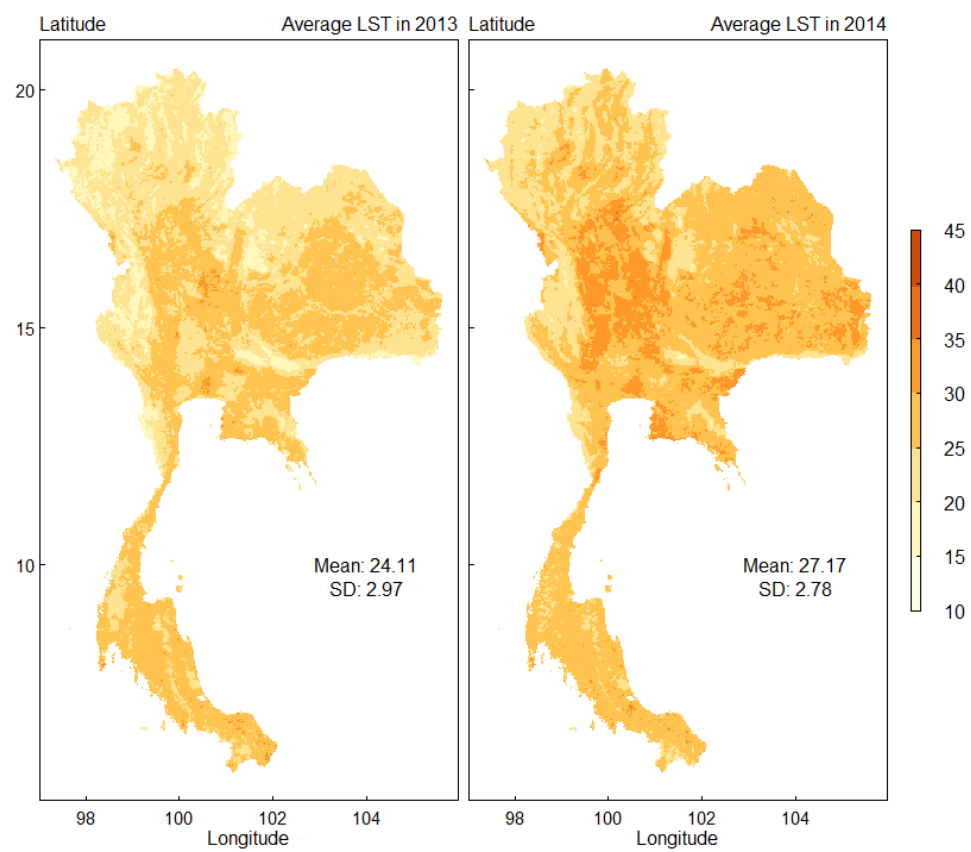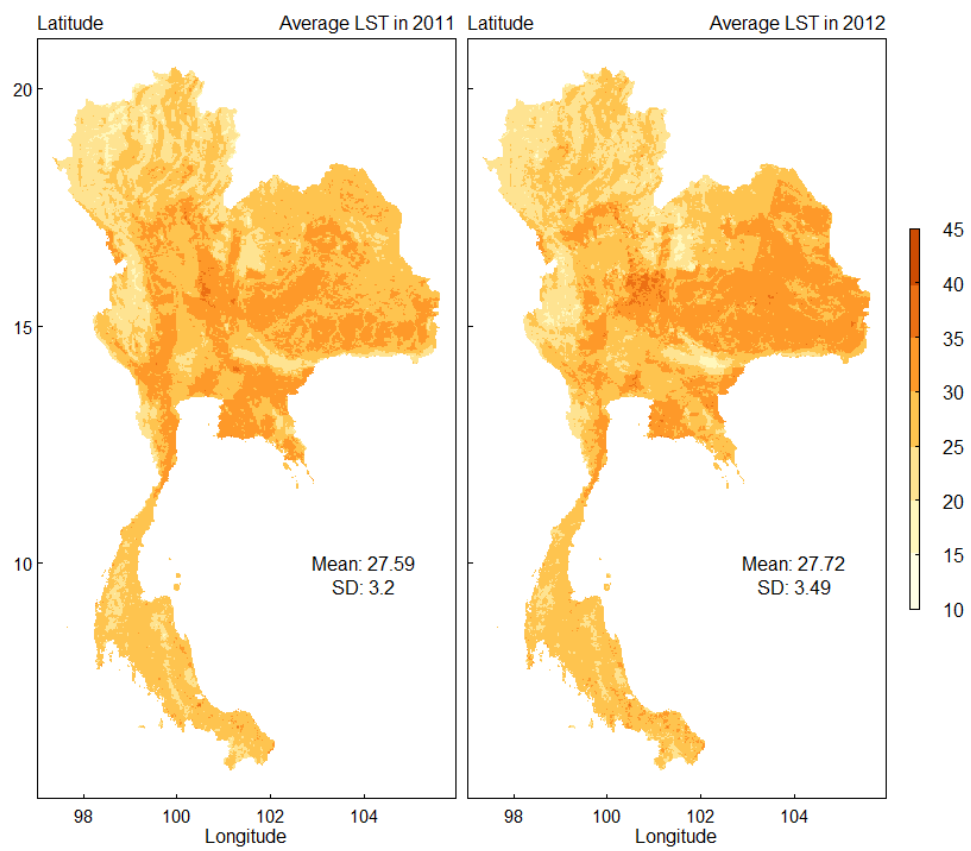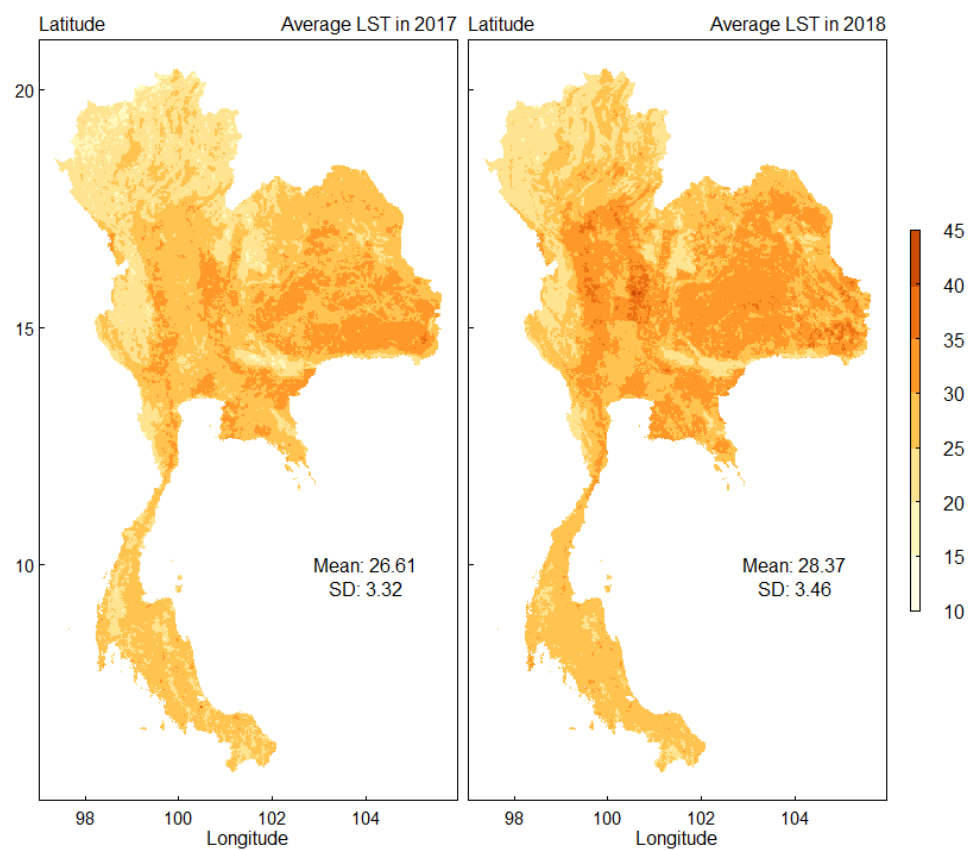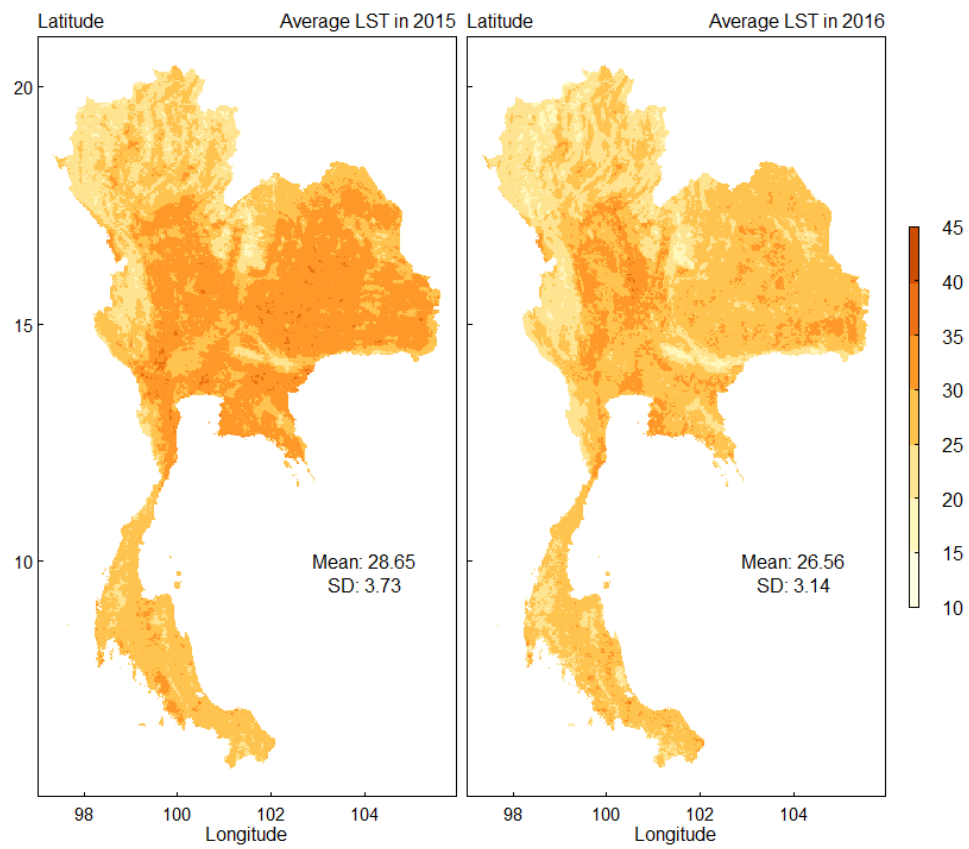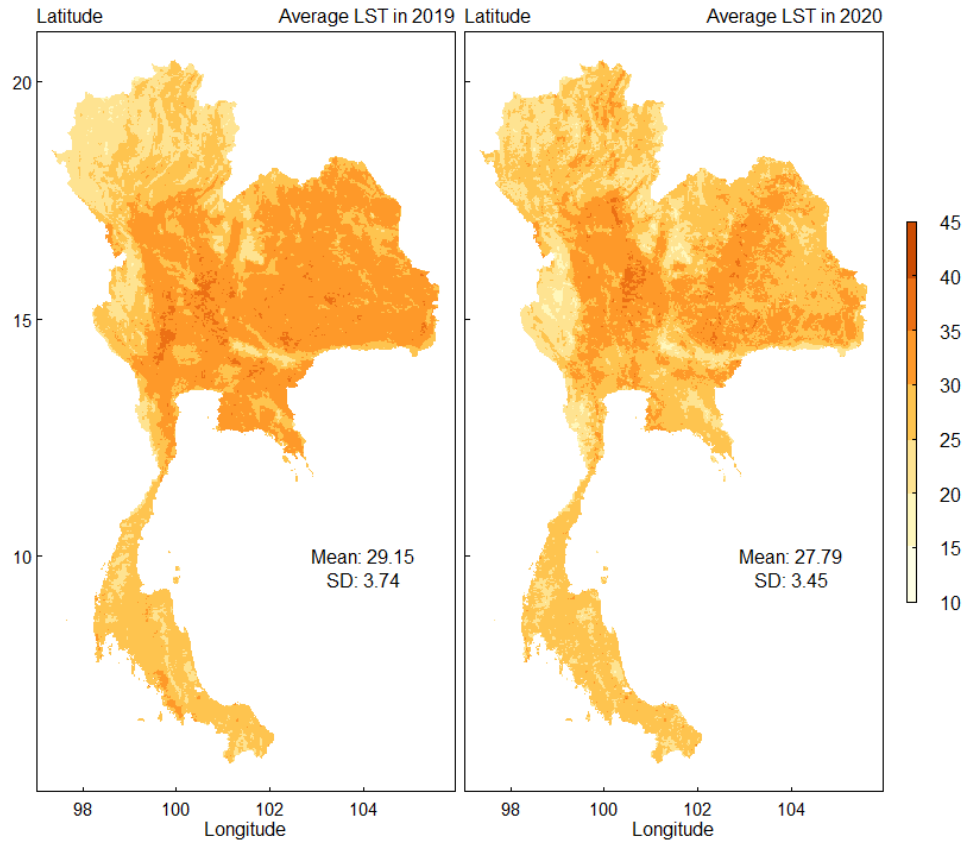
Figure 2.3: The average of LST in 2011-2020.

- Normalized Difference Vegetation Index (NDVI) data

We utilized the NDVI product from MOD13A1, which has a temporal resolution of 16 days and a spatial resolution of 500 meters. NDVI is a widely-used vegetation index that is beneficial in depicting land cover and changes and monitoring vegetation conditions globally. The NDVI data can provide valuable insights for modeling global biogeochemical and hydrologic processes and global and regional climates. NDVI data can also characterize various biophysical features and processes on the ground surface, such as primary production and land cover conversion [52]. Figure 2.4 presents the average NDVI from 2011 to 2020. The NDVI each year in Thailand does not differ much, with a mean value between 0.36-0.42, defined as a sparse vegetation area. It is a higher density of green in an area of the northern and southern parts of Thailand.
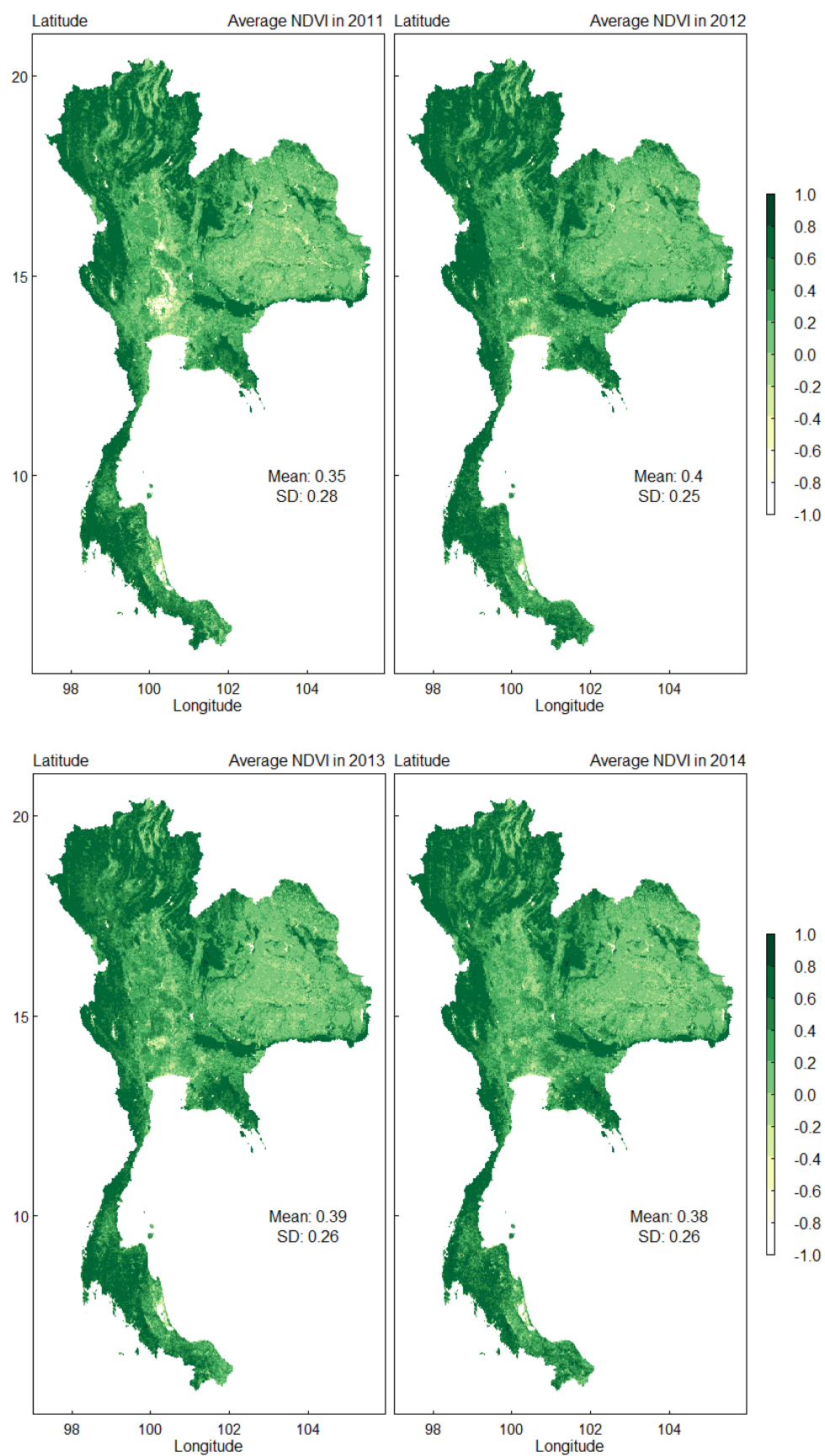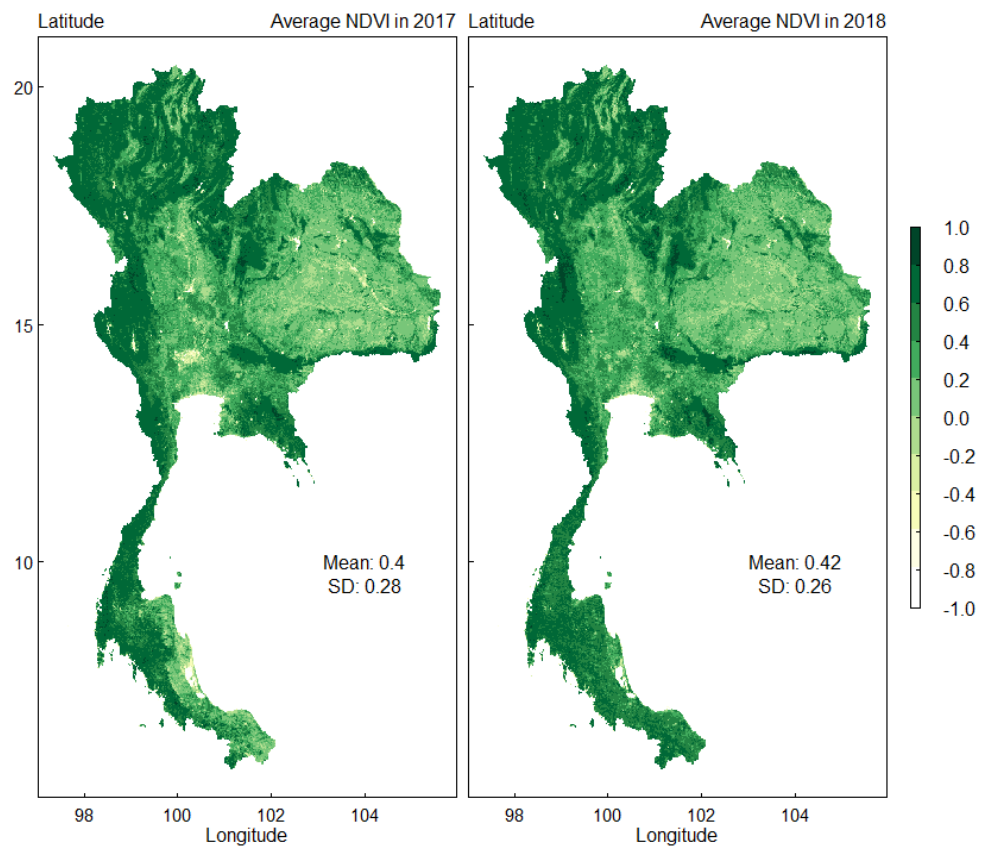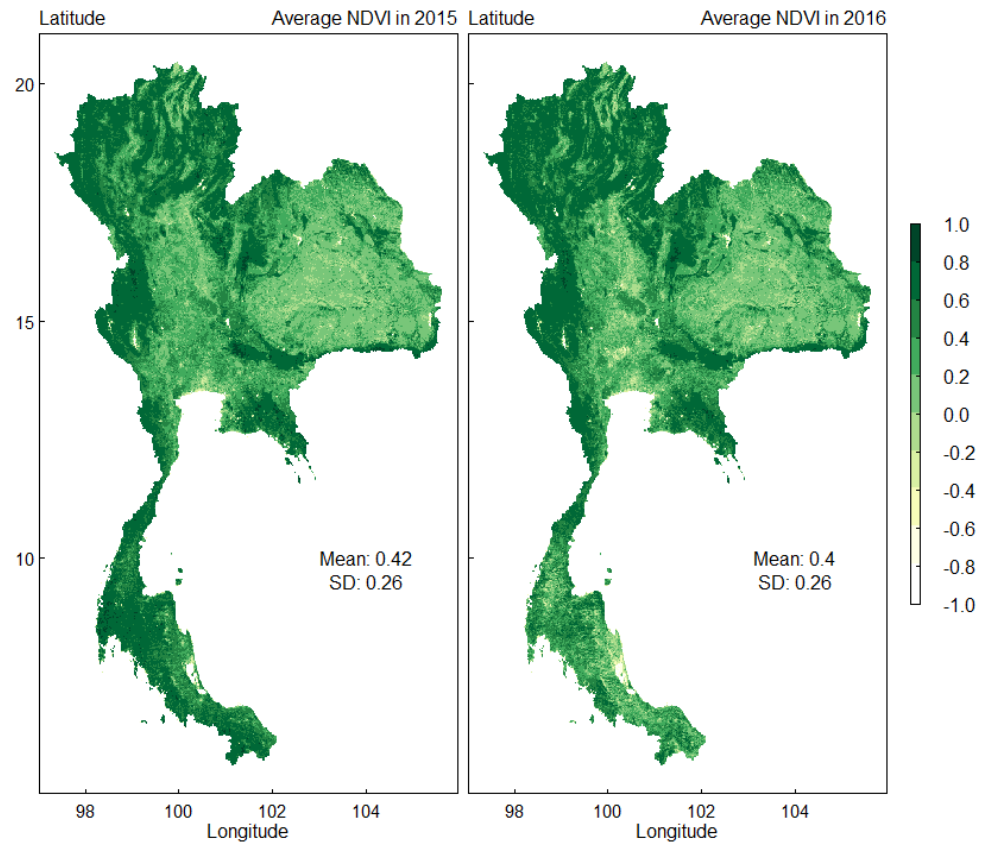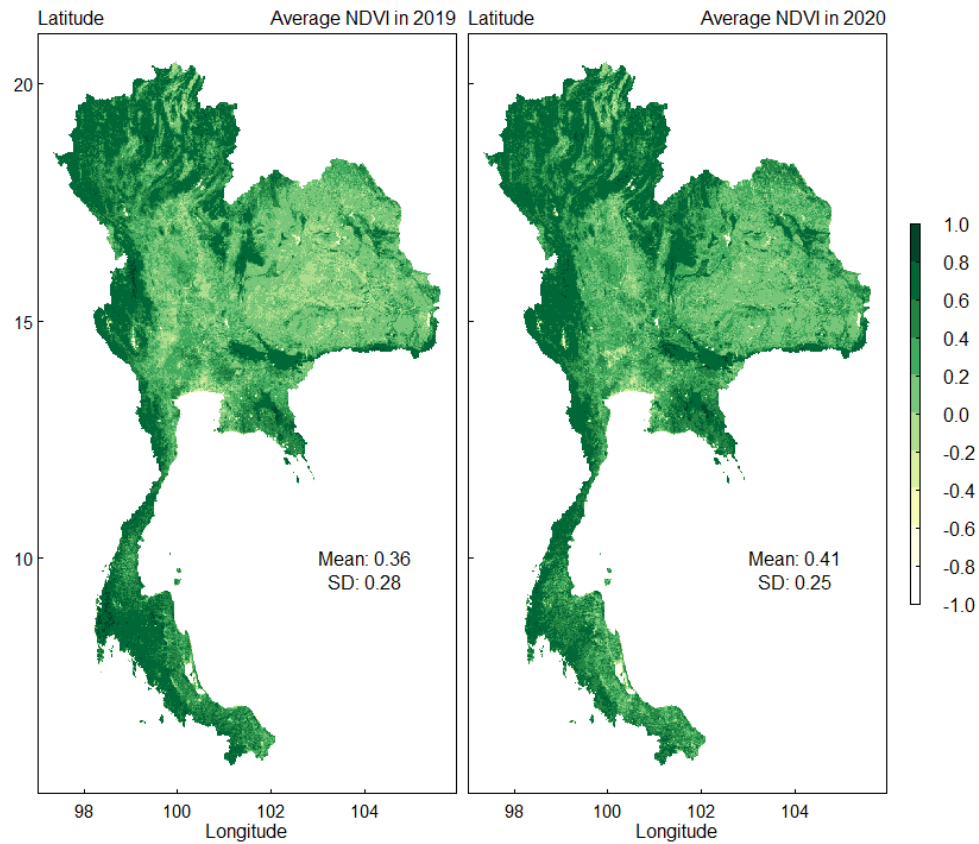
Figure 2.4: The average of NDVI in 2011-2020.

- Elevation (EV) data

EV data from "Land Digital Elevation Model (MODDEM1KM)—Land/sea mask and digital elevation model" were employed. EV data refers to a spatial resolution of 1 km. Figure 2.5 presents the map of EVs in Thailand. The mean of EVs in Thailand is 285.45 meters, this means most of the area in Thailand is a low-lying area.
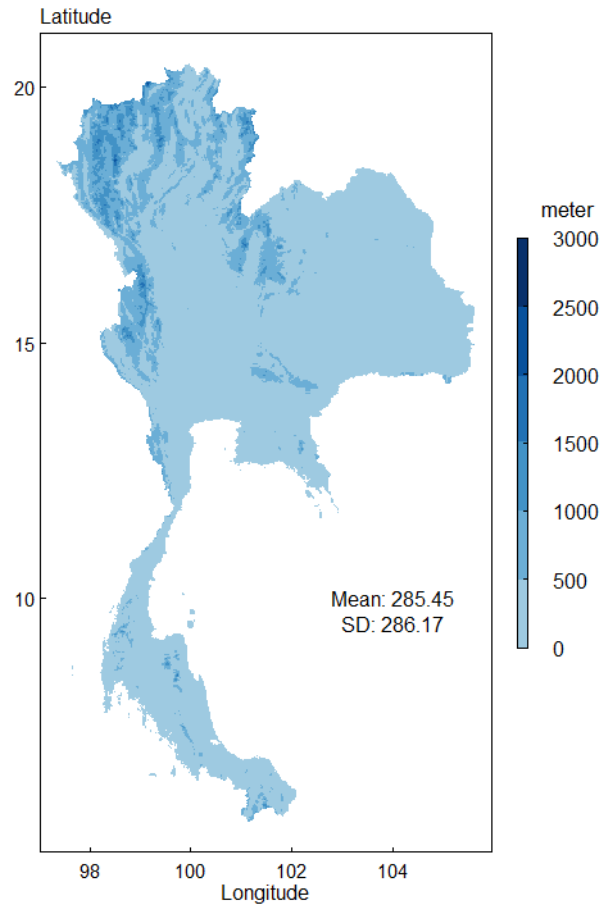
Figure 2.5: The EV in Thailand.

2) Data Distribution

Using the latitude and longitude of the stations, we integrated the daily PM2.5 concentrations for each station from 2011 to 2020. The dataset consists of 100,210 rows of PM2.5 data from 2011 to 2020. The AOD and LST data at 1 km resolution had more than 50% missing values. However, by taking the average over a 5 km radius, the missing values decreased to less than 50%. We will analyze the data at 5 km resolution to determine the association between satellite data and PM2.5 because it has more data and coverage areas. The final data set has 29.1% and 42.1% values for 1 km and 5 km resolution, respectively, as presented in Figure 2.6. When extending this model estimation to other geographical areas, including regions, provinces, districts, and sub-districts, we can utilize the average satellite data within the boundaries of each specific area. Furthermore, data imputation techniques, such as nearest date and pixel, can be employed.

Variables	n (%)	
	1 km	5 km
AOD	38,418 (38.3)	68,382 (68.2)
LST	35,897 (35.8)	57,620 (57.5)
NDVI	100,210 (100)	100,210 (100)
EV	100,210 (100)	100,210 (100)
Matching data	**29,146 (29.1)**	**42,172 (42.1)**
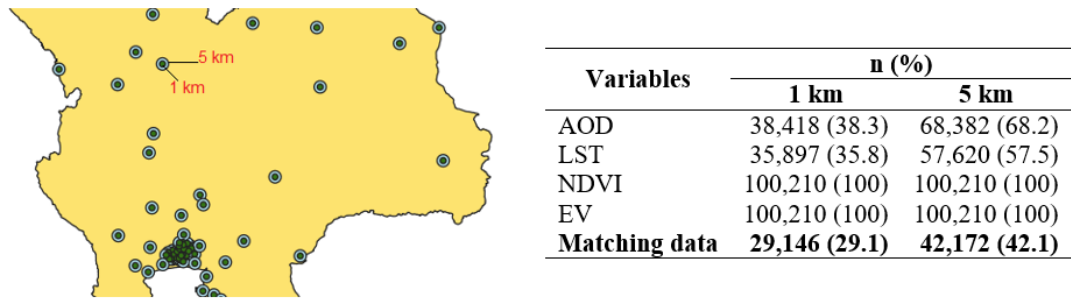
Figure 2.6: Integration of satellite data with PM2.5 data.

Figure 2.7 presents an example of data integration of AOD at stations of PM2.5. The extracted values of AOD were gathered at PM2.5 stations, similar to other satellite data such as LST, NDVI, and EV.
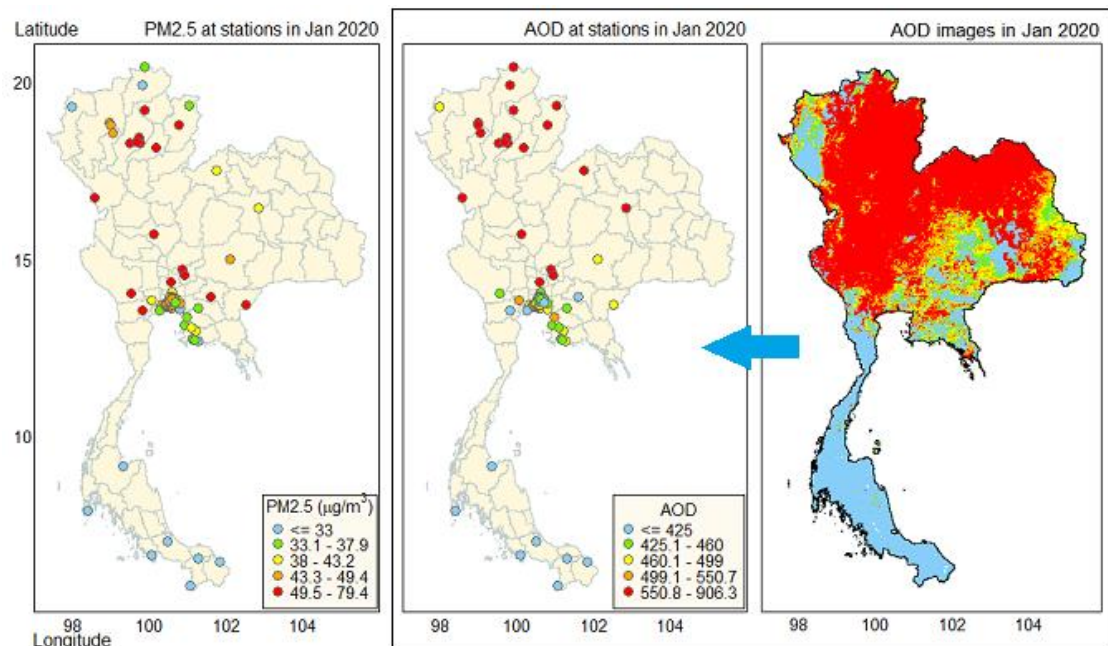
Figure 2.7: An example of data integration of AOD at PM2.5 stations.

The rows of the PCD dataset were randomly shuffled and divided into a training dataset (80%) and a validation dataset (20%) to ensure that model performance comparisons could be made. A consistent random state was used for this purpose. Table 2.1 presents the structure of the PCD and Bangkok's Air Quality and Noise Management Division (BAQ) data. The distribution of the training and validation datasets were similar; however, the testing dataset was different as it only included BAQ data collected in Bangkok provinces.

Table 2.1: The data structure of datasets.

Variables	Types	PCD (n = 34,748)		BAQ (n = 7,339)
		Training (n =27,798)	Validation (n = 6,950)	Testing
Stations	Nominal	68 stations	68 stations	49 stations
Date	Date	2,778 days	1,865 days	734 days
Month	Nominal	12 months	12 months	12 months
Year	Discrete	10 years	10 years	6 years
WOY	Nominal	53 weeks	53 weeks	53 weeks
PM2.5 ($\mu g/m^3$)	Continuous	μ: 32.2, SD: 23.7, IQR: 26	μ: 32.4, SD: 23.8, IQR: 26	μ: 30.1, SD: 16.2, IQR: 21
AOD	Continuous	μ: 0.5, SD: 0.3, IQR: 0.4	μ: 0.5, SD: 0.3, IQR: 0.4	μ: 0.5, SD: 0.3, IQR: 0.4
LST (°C)	Continuous	μ: 33.3, SD: 4.5, IQR: 6	μ: 33.4, SD: 4.5, IQR: 6	μ: 36.1, SD: 3.8, IQR: 4.3
NDVI	Continuous	μ: 0.1, SD: 0.2, IQR: 0.3	μ: 0.1, SD: 0.2, IQR: 0.3	μ: −0.1, SD: 0.1, IQR: 0.2
EV (m)	Continuous	μ: 144.6, SD: 198.9, IQR: 265.3	μ: 142.4, SD: 197.3, IQR: 265.3	μ: 6.8, SD: 1.6, IQR: 2.9

n: Rows; μ: Mean; SD: Standard deviation; IQR: Interquartile range; m: Meter.

Figure 2.8 presents that none of the variables follow a normal distribution. PM2.5 and AOD exhibit similar right-skewed distributions. In our dataset, the mean daily concentration of PM2.5 is 31.8 $\mu g/m^3$ with an SD of 22.6 $\mu g/m^3$, surpassing the WHO guidelines (15 $\mu g/m^3$) but below Thailand National Ambient Air Quality Standards (NAAQS) (50 $\mu g/m^3$). LST demonstrates a left-skewed distribution, while NDVI and EV exhibit right-skewed distributions, indicating that most stations are located in low-lying urban areas. Notably, data is sparser from May to October, likely due to limitations in satellite imagery capturing data during periods of lower PM2.5 levels in Thailand. Given the similarity in distributions between WOY and month, this study employed WOY as a representative variable. Additionally, the number of data rows increased annually due to the growing number of stations included in the investigation.
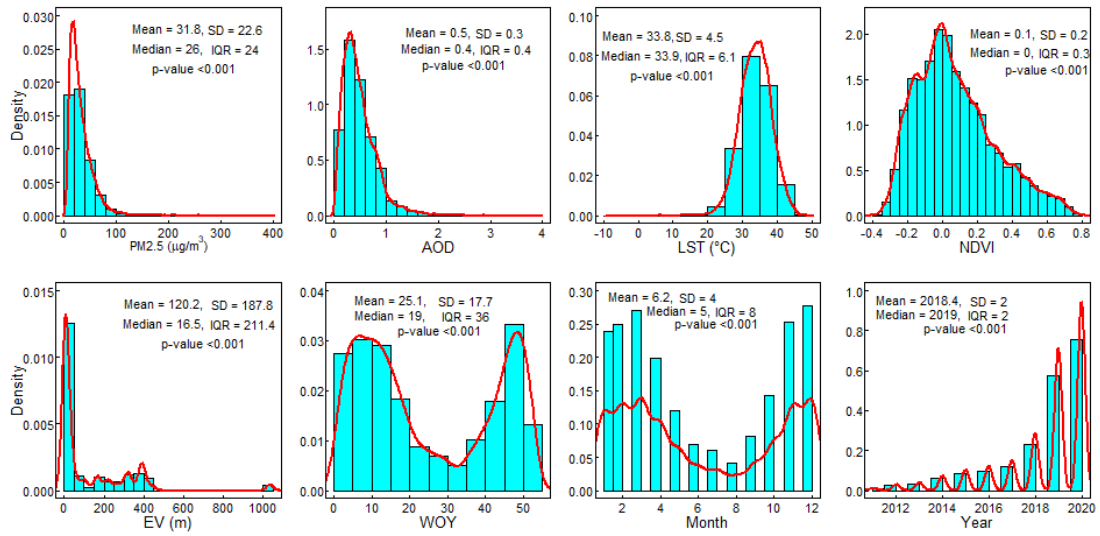
Figure 2.8: The distribution of data.

Figure 2.9 presents significant positive (p < 0.001) correlations between PM2.5 and AOD (r = 0.34), LST (r = 0.17), and EV (r = 0.033), and negative correlations with WOY (r = -0.24), NDVI (r = -0.17), and year (r = -0.069). AOD has the highest positive association, while lower PM2.5 levels are observed during WOY 20-40 in Thailand's rainy season. Dry seasons with increased LST show higher PM2.5 levels, while higher NDVI levels decrease PM2.5. Finally, EV and Year have lower correlation values with PM2.5.
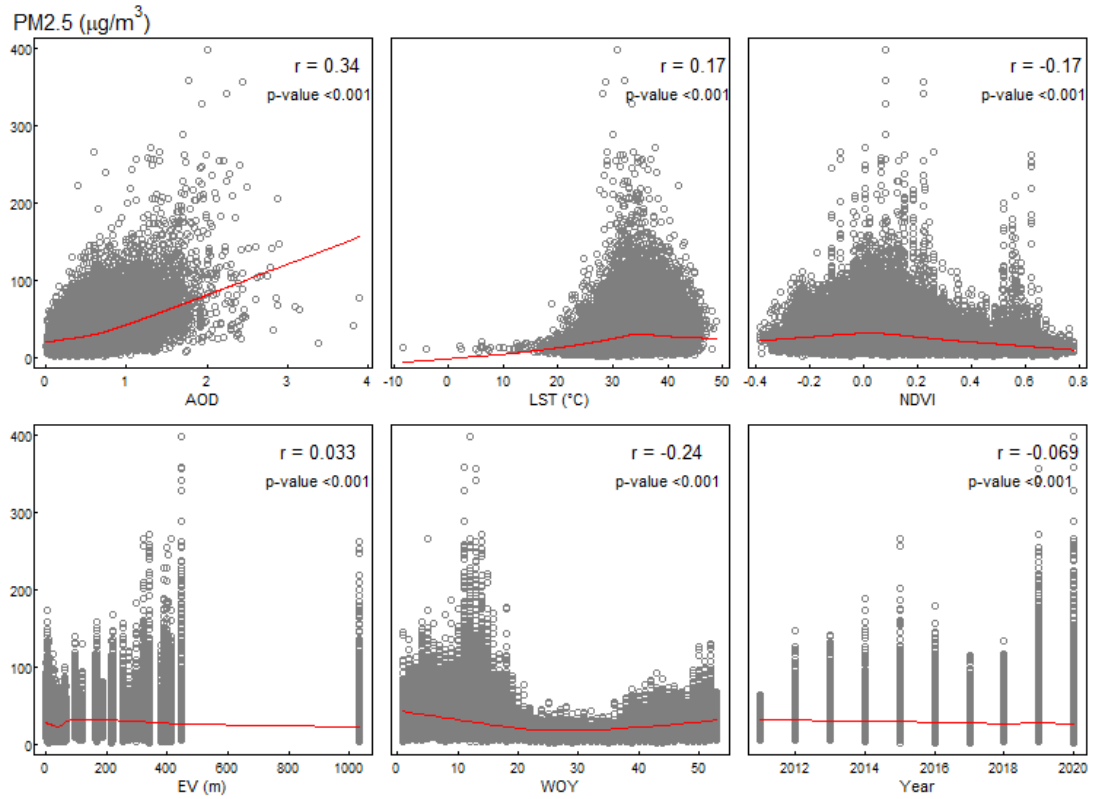
Figure 2.9: The correlation of satellite data with PM2.5.

From selecting possible predictor variables, we created a log-linear regression model by progressively adding and eliminating predictors based on p-values until there were no more variables to add or delete. According to Table 2.2, the final model leaves no variables out with an R^2 of 49.4 since its increases the R^2 values and significantly p-values < 0.001. The factor variable of WOY (+28) and AOD (+15.7) has the highest increase in the accuracy of the model.

Table 2.2: Variable selection in a stepwise regression model.

Response: log(PM2.5)		
Factor variable	**R-squared**	**P-value***
AOD	15.7 (+15.7)	< 0.001
AOD+LST	17.6 (+1.9)	< 0.001
AOD+LST+NDVI	18.1 (+0.5)	< 0.001
AOD+LST+NDVI+EV	20.5 (+2.4)	< 0.001
AOD+LST+NDVI+EV+WOY	48.5 (+28)	< 0.001
AOD+LST+NDVI+EV+WOY+Year	49.4 (+0.9)	< 0.001

*All factor variable has p-values < 0.001.

2) Modeling for Predicting PM2.5

We develop four models to estimate daily PM2.5 levels: MLR, RF, XGBoost, and Support Vector Machines (SVM). Fifth, we will evaluate the model's accuracy using the R^2, the root mean square error (RMSE), and K-fold cross-validation.

The results presented in Table 2.3 suggest that the RF model is performing well in predicting PM2.5, as it consistently produces high R² values, especially in the training dataset (0.95). However, the difference between the R² values for the training (0.95), validation (0.78), and testing (0.71) datasets, as well as the RMSE values for these datasets (5.6 μg/m³, 11.2 μg/m³, and 8.8 μg/m³), raises concerns about potential overfitting. Overfitting occurs when a model learns patterns specific to the training data, making it less generalizable to unseen data, as evidenced by the decline in performance from training to testing.

The significant drop in R² from 0.95 in training to 0.78 in validation (0.71 in testing and 0.74 in cross-validation) suggests that the RF model may be capturing noise or overly complex patterns in the training data that do not generalize well to new, unseen data. The RMSE values further emphasize this, with the training dataset showing a much lower error than the validation and testing datasets, indicating a potential overfitting issue. Cross-validation also shows a moderate performance with an RMSE of 12.2 μg/m³, supporting that while the model is strong in training, its real-world applicability may be more limited. However, we selected models based on their highest accuracy in the validation and testing datasets. As of our review, over 70% accuracy in estimating PM2.5 using satellite data is acceptable [27].

In comparison, XGBoost and SVM show more consistent performance across datasets, with less dramatic discrepancies between training and testing results. This suggests that these models might be less prone to overfitting. The MLR model, on the other hand, performed the worst, likely due to its simpler nature and inability to capture the complex relationships between the predictors and PM2.5 levels.

Table 2.3: The performance of models for estimation of PM2.5.

Models	R^2 (RMSE ($\mu g/m^3$))			
	Training	Validation	Testing	10-fold
MLR				
AOD	0.18 (21.5)	0.19 (21.3)	0.04 (16.8)	0.26 (20.4)
AOD + LST	0.21 (21.2)	0.22 (21.0)	0.01 (17.1)	0.27 (20.2)
AOD + LST + NDVI	0.22 (21.3)	0.22 (21.2)	0.01 (17.3)	0.27 (20.2)
AOD + LST + NDVI + EV	0.25 (20.5)	0.25 (20.38)	0.01 (17.35)	0.31 (19.7)
AOD + LST + NDVI + EV + WOY	0.51 (18.4)	0.51 (17.9)	0.35 (14.1)	0.45 (17.6)
AOD + LST + NDVI + EV + WOY + Year	0.51 (18.3)	0.52 (17.8)	0.35 (13.8)	0.46 (17.5)
RF				
AOD	0.79 (11.4)	0.16 (23.1)	0.02 (20.5)	0.15 (23.3)
AOD + LST	0.86 (10.1)	0.25 (20.9)	0.04 (18.6)	0.26 (20.5)
AOD + LST + NDVI	0.90 (8.8)	0.44 (17.9)	0.10 (16.0)	0.42 (18.0)
AOD + LST + NDVI + EV	0.89 (8.8)	0.60 (15.2)	0.15 (15.0)	0.59 (15.1)
AOD + LST + NDVI + EV + WOY	0.92 (7.2)	0.74 (12.3)	0.60 (10.5)	0.70 (13.0)
AOD + LST + NDVI + EV + WOY +Year	0.95 (5.6)	0.78 (11.2)	0.71 (8.8)	0.74 (12.2)
XGBoost				
AOD	0.31 (19.8)	0.27 (20.3)	0.04 (17.4)	0.28 (20.2)
AOD + LST	0.34 (19.3)	0.30 (19.8)	0.05 (17.6)	0.30 (19.8)
AOD + LST + NDVI	0.40 (18.4)	0.38 (18.7)	0.08 (15.9)	0.34 (19.0)
AOD + LST + NDVI + EV	0.49 (17.0)	0.47 (17.3)	0.12 (15.2)	0.43 (17.8)
AOD + LST + NDVI + EV + WOY	0.61 (14.9)	0.60 (15.1)	0.43 (12.4)	0.53 (16.3)
AOD + LST + NDVI + EV + WOY + Year	0.62 (14.7)	0.60 (15.0)	0.45 (12.1)	0.53 (16.2)
SVM				
AOD	0.28 (20.6)	0.28 (20.7)	0.04 (17.1)	0.27 (20.7)
AOD + LST	0.31 (20.1)	0.31 (20.2)	0.05 (16.9)	0.29 (20.3)
AOD + LST + NDVI	0.39 (18.8)	0.38 (18.9)	0.09 (15.7)	0.38 (18.8)
AOD + LST + NDVI + EV	0.47 (17.6)	0.46 (17.8)	0.14 (15.6)	0.46 (17.6)
AOD + LST + NDVI + EV + WOY	0.59 (15.6)	0.60 (15.4)	0.51 (11.5)	0.60 (15.2)
AOD + LST + NDVI + EV + WOY + Year	0.61 (15.3)	0.62 (15.2)	0.52 (11.6)	0.63 (14.7)

Although the final RF model has a higher R^2 accuracy in the validation dataset than the testing dataset, the testing dataset has a lower RMSE than the validation dataset (Figure 2.10). This means the RF model can estimate PM2.5 in the validation dataset more accurately than in the testing dataset. However, the difference between the actual and estimated PM2.5 in the testing dataset is closer than in the validation dataset due to the lower RMSE. This discrepancy could be attributed to the fact that the testing dataset only covers Bangkok provinces and thus has more data from these areas. In contrast, the validation dataset covers all areas of Thailand.
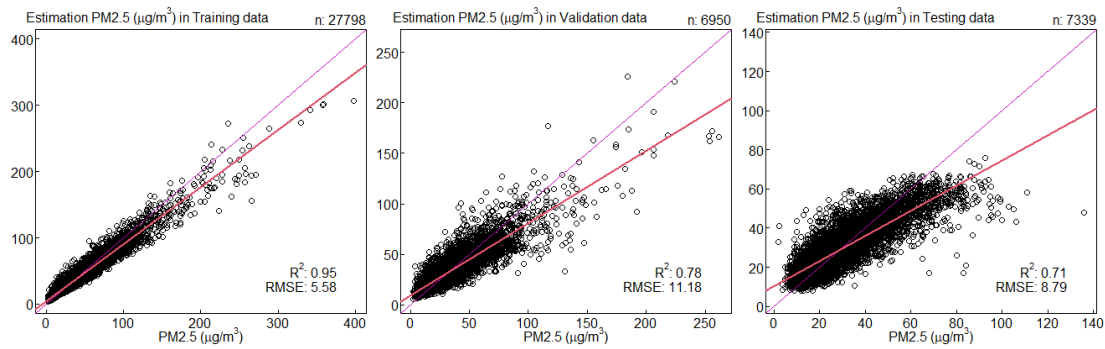
Figure 2.10: The accuracy of the RF model.

RF approaches were used to estimate daily PM2.5 concentrations in Thailand, and it was found that the model that included AOD, LST, NDVI, EV, WOY, and year had the best performance. The RF results also show two alternative measurements of each predictor variable's relative contribution in Figure 2.11. The %IncMSE is a percentage increase in mean square error, equivalent to accuracy-based importance. The IncNodePurity, calculated similarly to Gini-based importance, is based on reducing the sum of squared errors whenever a variable is split. Without WOY, AOD, EV, year, LST, and NDVI as predictors, the %IncMSE was 72.4%, 59.3%, 50.7%, 43.2%, 32.4%, and 31.5%, respectively. The important variables for IncNodePurity were WOY, AOD, EV, NDVI, LST, and year, respectively. These two measurements were calculated using different methods due to their strong association with ground-level PM2.5. Additionally, all the factors were needed to estimate PM2.5 levels in Thailand, where WOY and AOD were the three most essential variables in the two measurements.
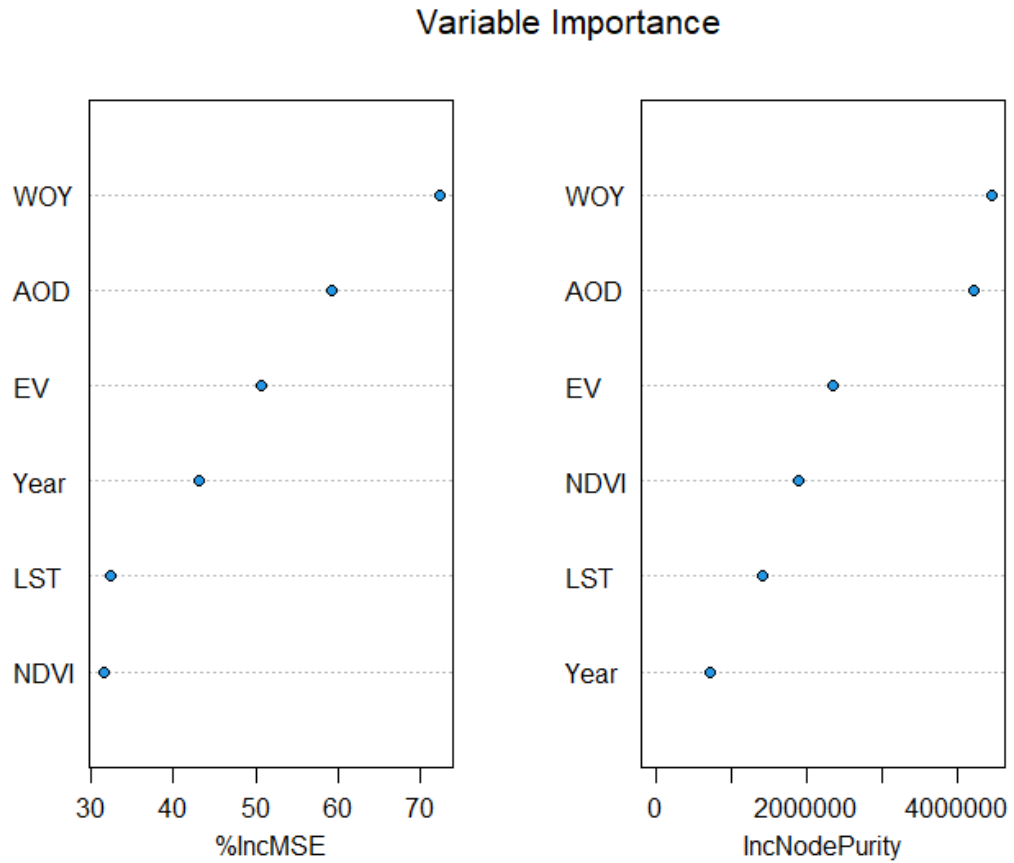
Figure 2.11: The importance variables for estimation of PM2.5.

3) Estimate PM2.5 in 1 km Resolution

Figure 2.12 presents the PM2.5 time series plot and estimation for the training, validation, and testing data. The three plots exhibit a consistent pattern in the observed and estimated PM2.5 concentrations, with the highest concentrations observed during weeks 45 to 53 (November to December) and 1 to 10 (January to March). The difference between the measured and estimated PM2.5 concentrations in the testing dataset was slight in 2015 and 2016 but remained consistent in 2017 and 2020.
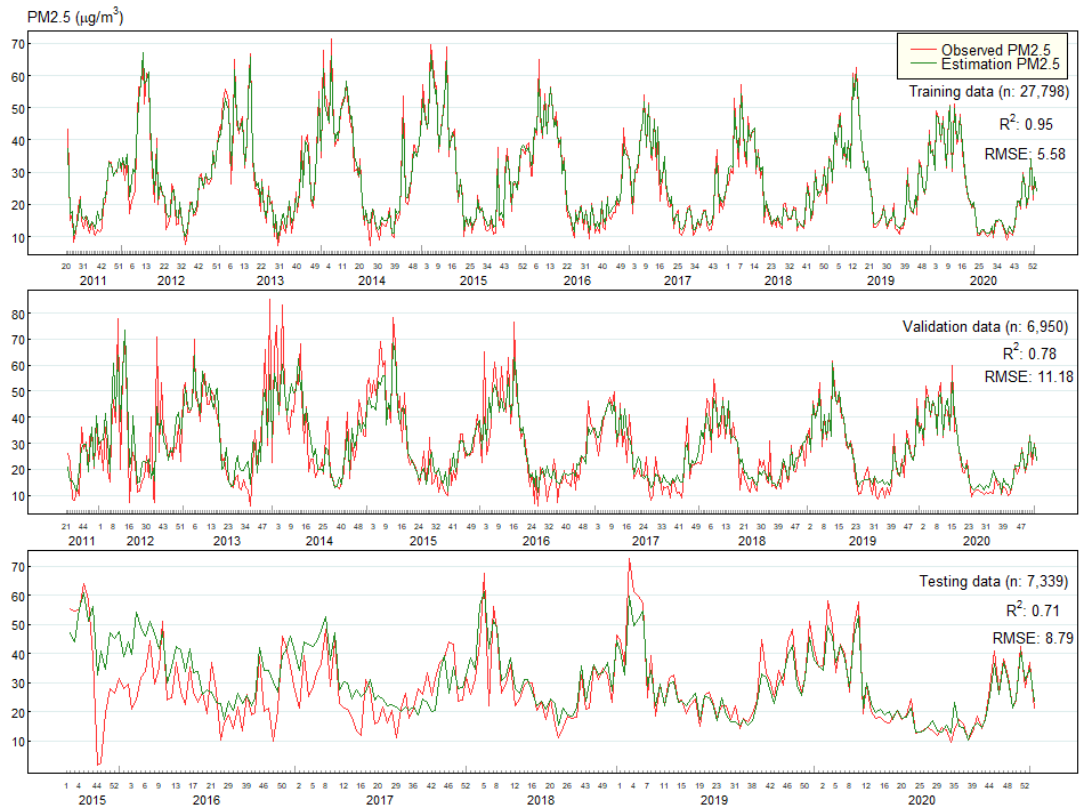
Figure 2.12 Time series (WOY) of PM2.5 observed and estimation of PM2.5 for training, validation, and testing data.

Figure 2.13 presents the estimates of PM2.5 concentrations from 2011 to 2020 at a 1 km resolution using the RF model. The values of PM2.5 at stations and the estimated PM2.5 are comparable. Northern Thailand exhibited the highest PM2.5 concentrations, while Southern Thailand showed the lowest levels. Monthly PM2.5 averages for each province were calculated based on the satellite data at 1 km resolution.

Latitude · PM2.5 at stations in 2011 · Estimation PM2.5 in 2011

PM2.5 (µg/m³)
- < 15
- 15 - 24
- 25 - 34
- 35 - 49
- 50+

Longitude
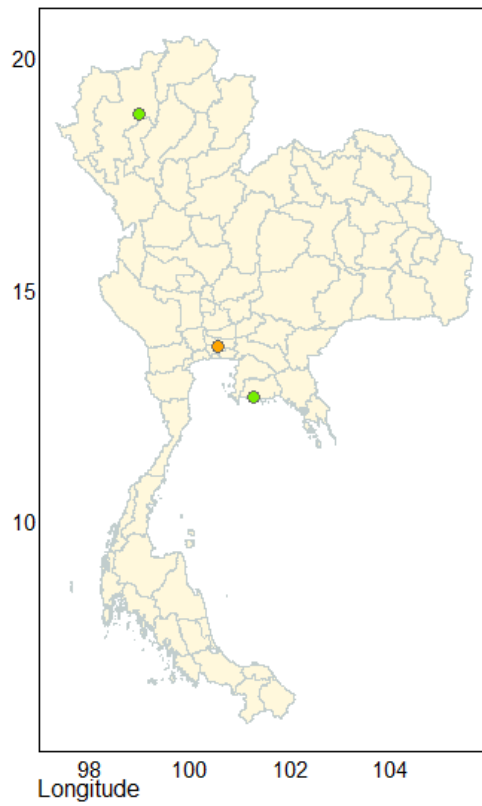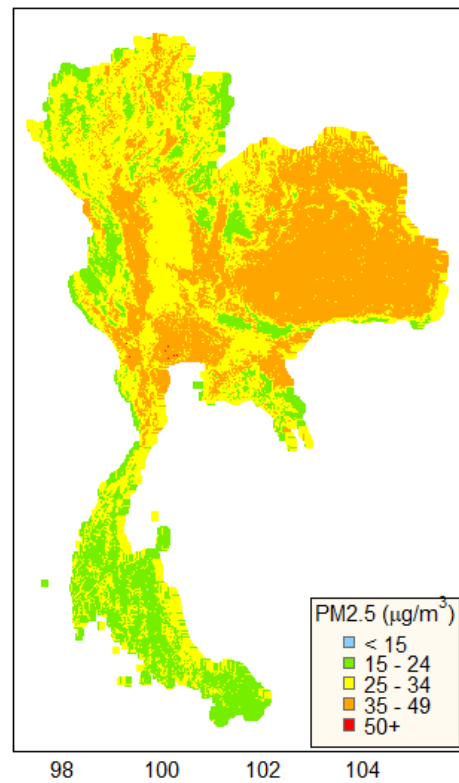
Latitude · PM2.5 at stations in 2012 · Estimation PM2.5 in 2012
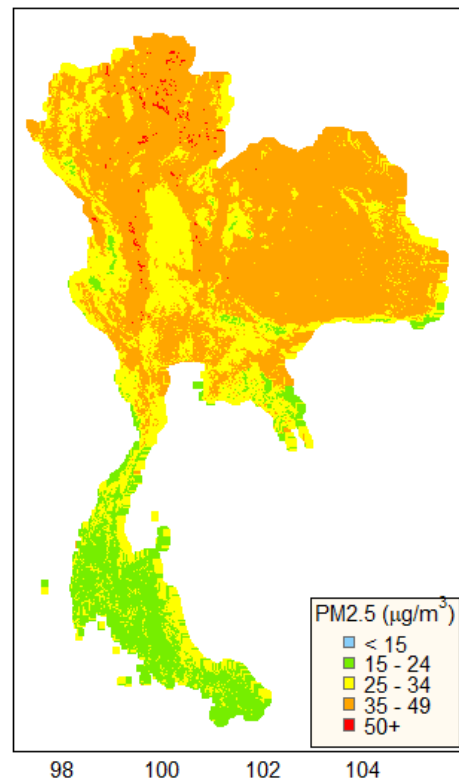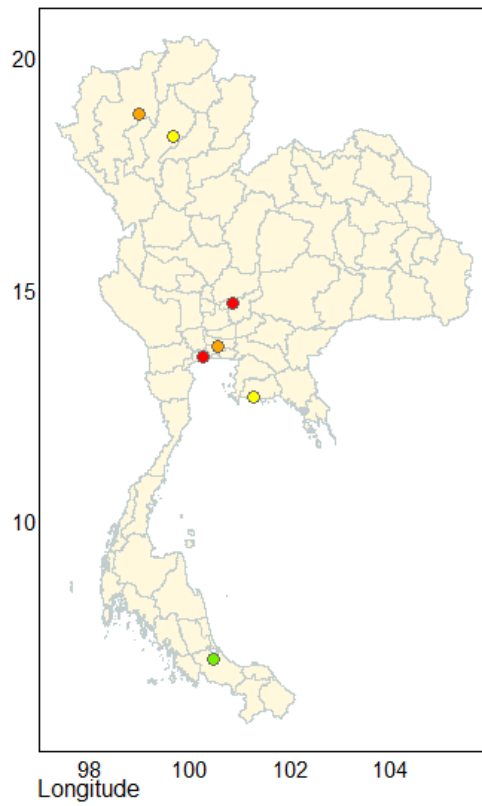
PM2.5 (µg/m³)
- < 15
- 15 - 24
- 25 - 34
- 35 - 49
- 50+

Longitude
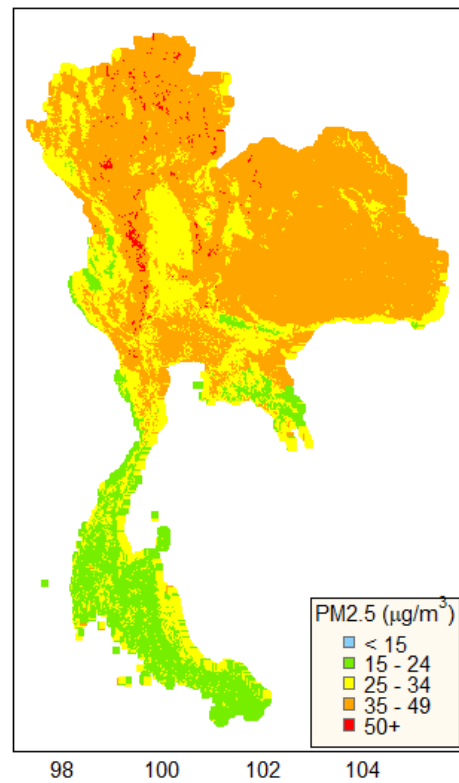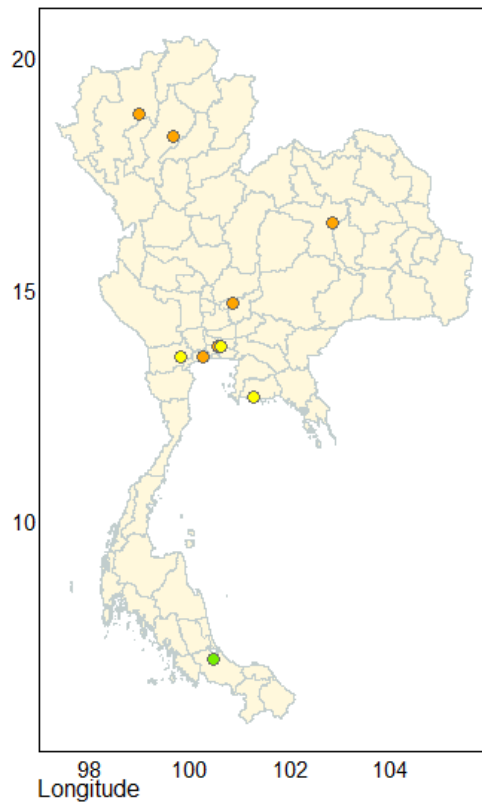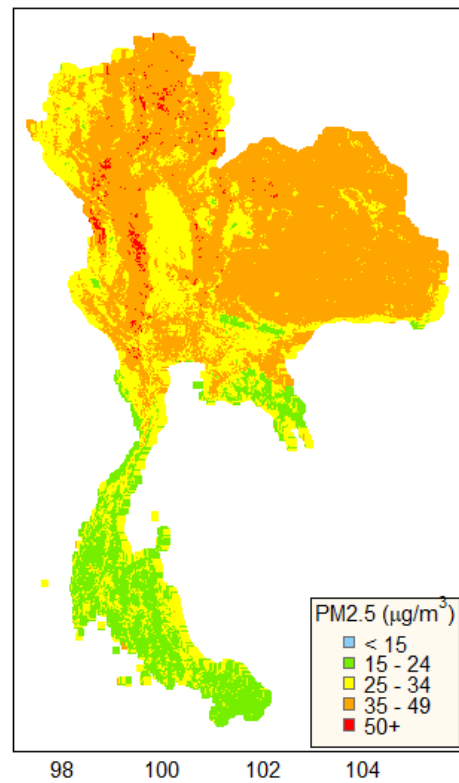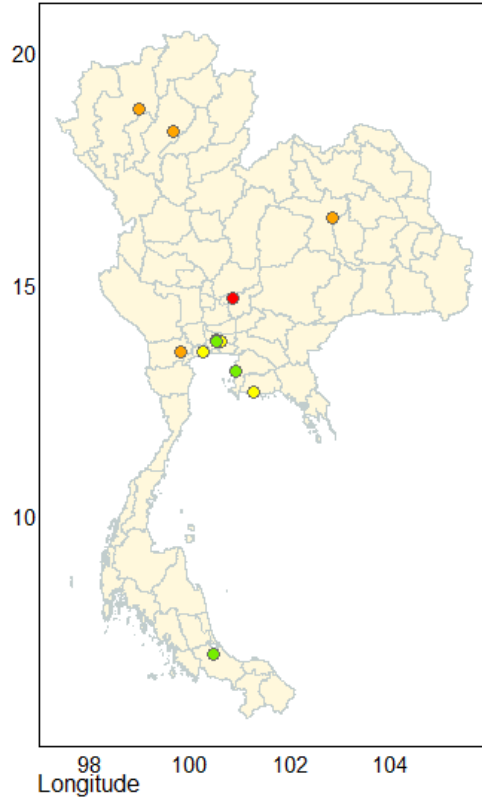
Latitude | PM2.5 at stations in 2015 | Estimation PM2.5 in 2015
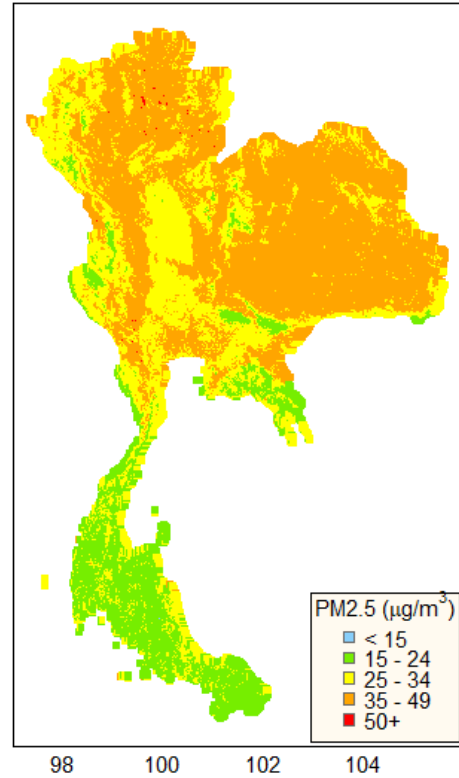
Latitude | PM2.5 at stations in 2016 | Estimation PM2.5 in 2016
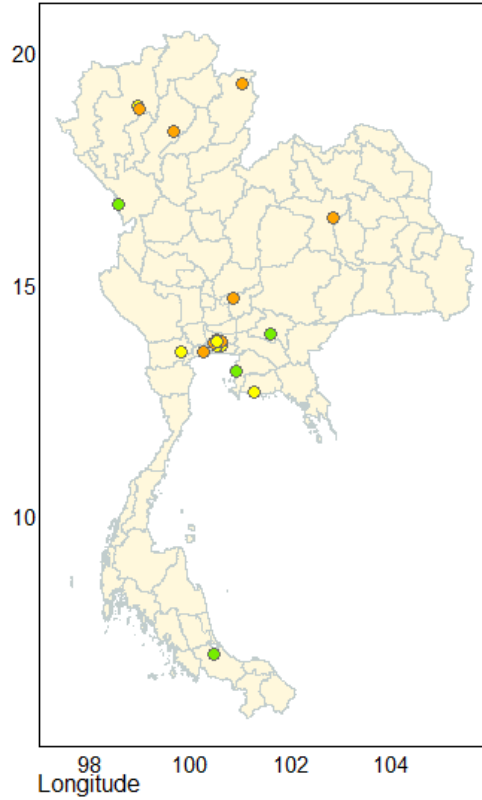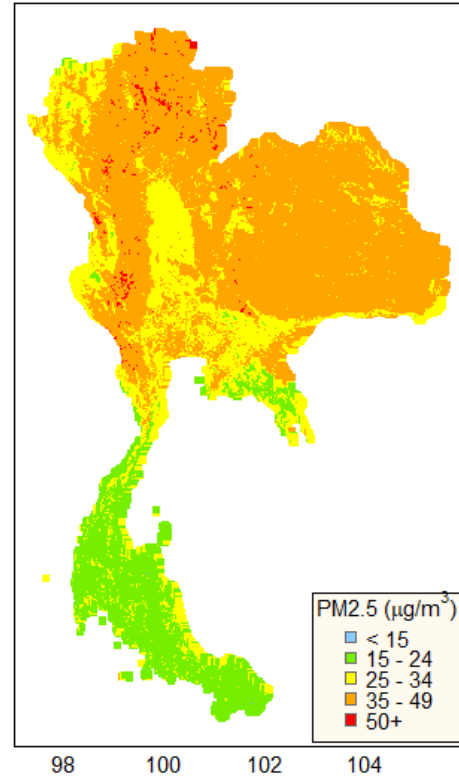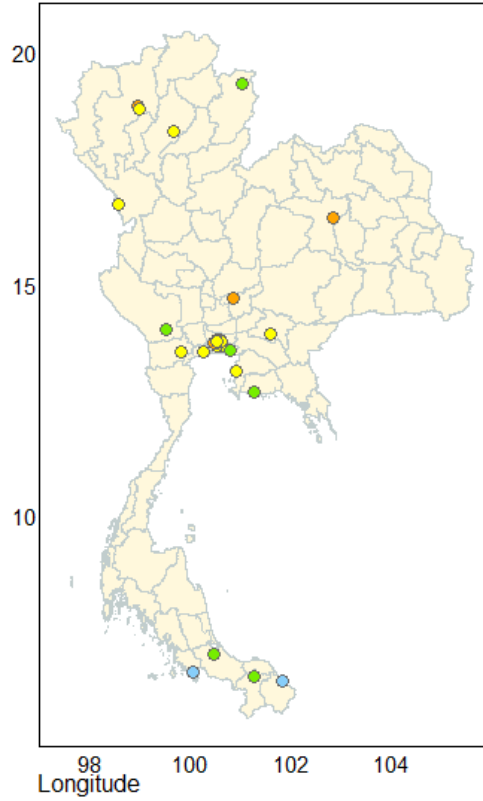
PM2.5 (µg/m³)
< 15
15 - 24
25 - 34
35 - 49
50+
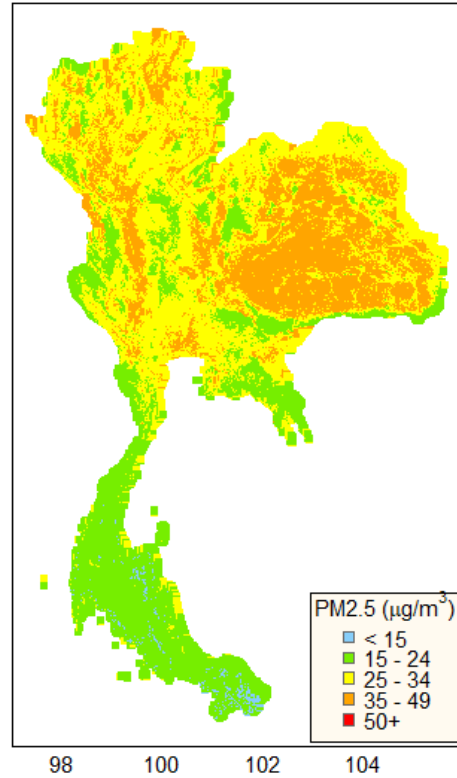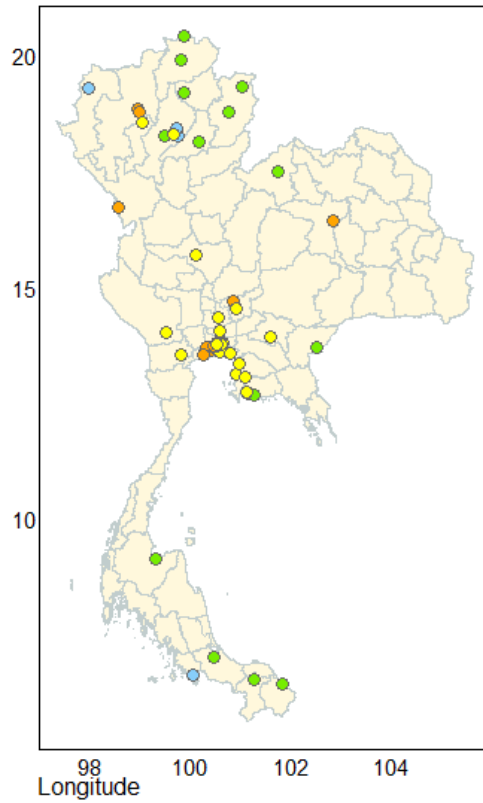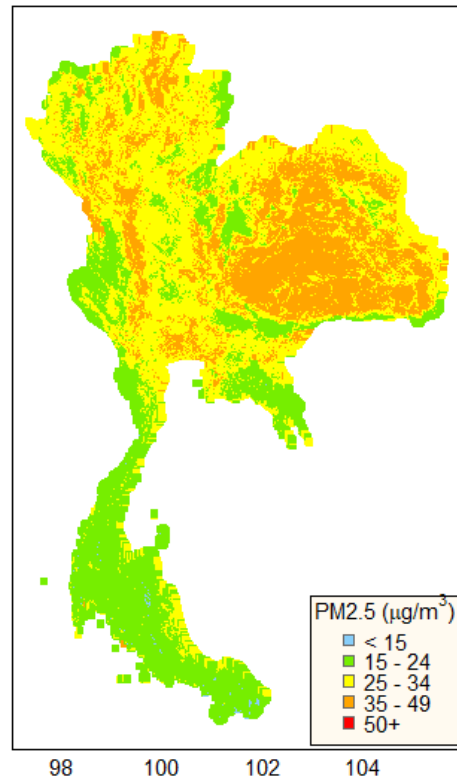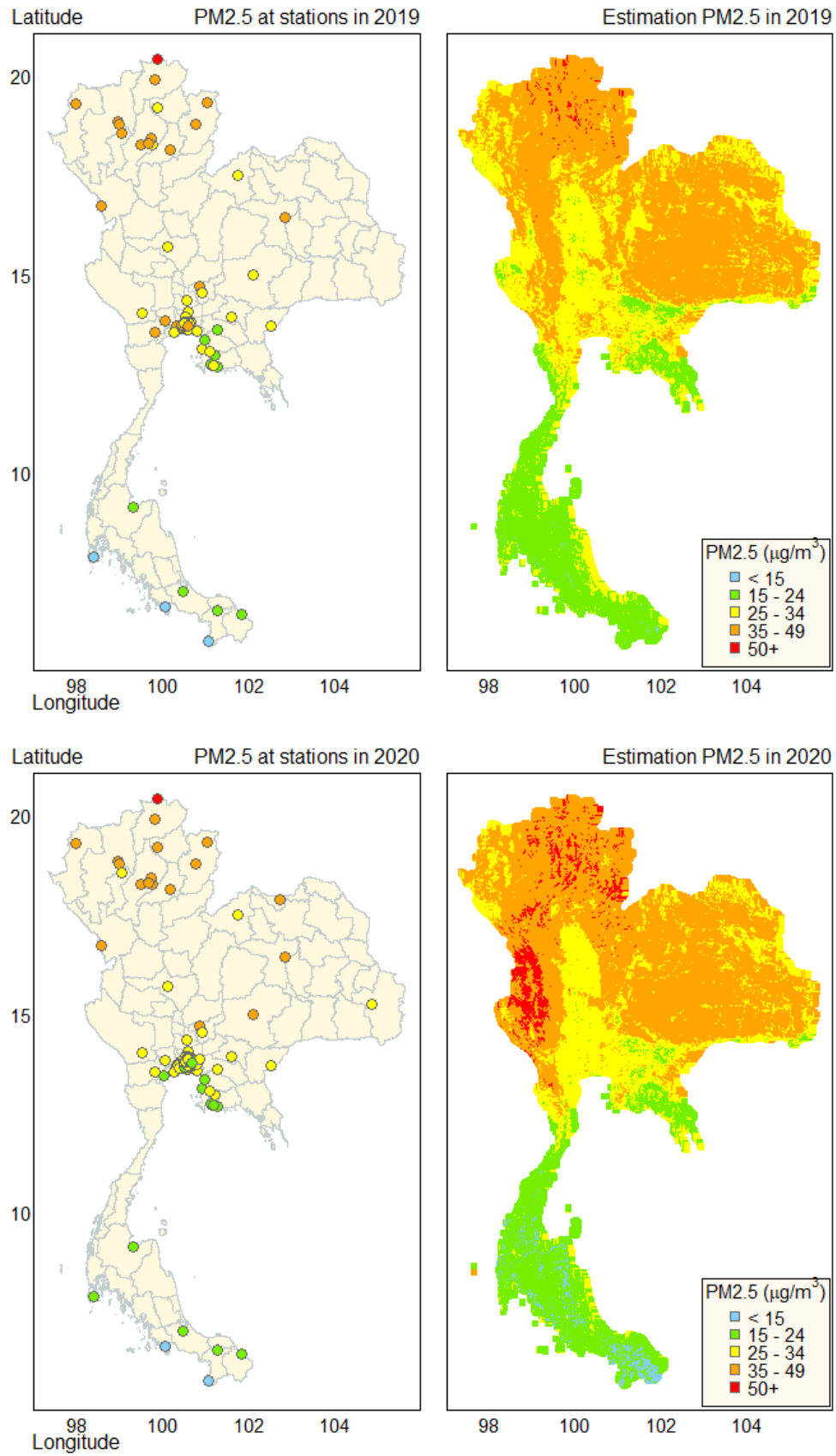
Figure 2.13: Estimation of PM2.5 data using satellite data (2011-2020).

As presented in Table 2.4, R² values for the training dataset were consistently high across regions (89%–96%). However, variations were noted in the validation dataset, with the northern region exhibiting the highest R² values due to the dominance of agricultural burning as a PM2.5 source, which is effectively captured by satellite data. In contrast, the southern region, with its relatively low PM2.5 levels, showed lower R² values, possibly due to the limited ability of remote sensing to capture subtle PM2.5 variations. Despite these differences, the random forest model demonstrated acceptable accuracy, aligning with findings from other studies using satellite data for PM2.5 estimation, where R² values ranged from 49% to 83% and RMSE values from 2.67 to 22 µg/m³ [22], [23].

Table 2.4: Test the performance of daily PM2.5 estimation by region.

Region	Training			Validation		
	Station	N	R^2 (RMSE)	Station	N	R^2 (RMSE)
Central	39	14682	94% (5)	39	3668	69% (10.2)
North	16	8922	96% (6.9)	16	2182	82% (13.5)
Northeast	5	1772	93% (5.6)	5	433	69% (12.3)
South	8	2422	89% (3.1)	8	667	45% (5.9)
Total	68	27798	95% (5.6)	68	6950	78% (11.2)

N: Number of data

The validation dataset reveals that the northern region exhibits a higher R^2 value than other regions. Consequently, we delved into the data for each region, discovering that the trend of PM2.5 levels between 2011 and 2020 in the north remained relatively stable, unlike some regions where a slight decrease was observed (Figure 2.14). These variations may potentially influence the accuracy of the R^2 values.
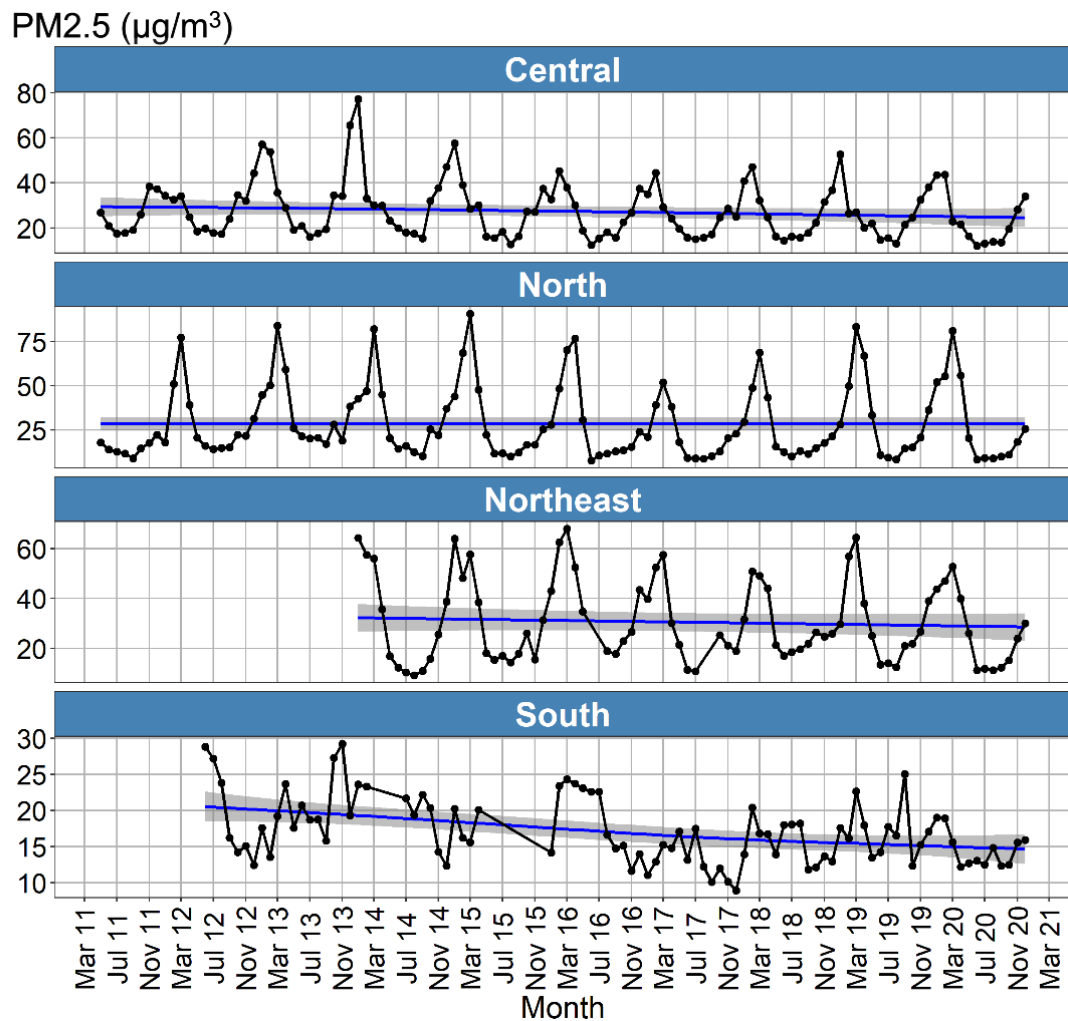
Figure 2.14: The monthly trends of PM2.5 by region.

2.2.2 Cardiorespiratory Mortality Data

The National Health Security Office (NHSO) provided computerized claims data spanning from January 1, 2015, to December 31, 2019. This period was selected for its completeness and relevance in reflecting recent trends in healthcare utilization and outcomes. In compliance with Thailand's data protection laws, no personally identifiable information was collected. Cardiorespiratory mortality was classified using International Classification of Diseases version 10 (ICD-10) diagnosis codes, as presented in Table 2.5. The focus was on deaths attributed to cardiovascular and respiratory causes.

Table 2.5: The ICD-10 of cardiorespiratory mortality.

Disease causes	ICD-10
Cardiovascular disease	
Angina pectoris	I20.0, I20.1, I20.9
Acute Myocardial infarction	I21.0, I21.1, I21.2, I21.3, I21.4, I21.9, I22.0, I22.1, I22.8, I22.9, I24.8, I24.9
Chronic ischemic heart disease	I25.2, I25.8, I25.9
Heart rhythm disturbance	I44.0, I44.1, I44.2, I44.7, I45.5, I45.6, I45.9
Cardiac arrest	I46.0, I46.1, I46.9
Tachycardia	I47.1, I47.2, I47.9, R00.0
Atrial fibrillation and flutter	I48, I48.0, I48.1, I48.2, I48.9
Other cardiac arrhythmias	I49, I49.0, I49.1, I49.3, I49.5, I49.8, I49.9
Heart failure	I50.1, I50.9
Complications and ill-defined descriptions of heart disease	I51.0, I51.1, I51.2, I51.3, I51.4, I51.5, I51.6, I51. 7, I51. 8, I51. 9
Cerebrovascular event	I60.4, I60.6, I60.7, I60.8, I60.9, I61.0, I61.1, I61.2, I61.3, I61.4, I61.5, I61.6, I61.8, I61.9, I62.0, I62.1, I62.9, I63.0, I63.1, I63.2, I63.3, I63.4, I63.5, I63.6, I63.8, I63.9, I64, I65.0, I65.1, I65.2, I65.3, I65.9, I66.0, I66.2, I66.3, I66.8, I66.9, I67.0, I67.4, I67.5, I67.6, I67.9, G45.0, G45.9
Vascular syndromes of brain in cerebrovascular diseases	G46.3
Bradycardia	R00.1
Heart beat abnormality	R00.8
Respiratory system	
Suppurative otitis media	H66, H66.0, H66.3, H66.4, H66.9
Common cold	J00
Acute sinusitis	J01.3, J01.9
Acute pharyngitis	J02.9
Acute tonsillitis	J03.9
Laryngitis	J04.2, J05.0, J05.1, J37.0, J38.4
Laryngopharyngitis	J06.8, J06.9
Pneumonia	J10.0, J10.1, J10.8, J11.0, J11.1, J11.8, J12.0, J12.1, J12.8 , J12.9, J13, J14, J15.0, J15.1, J15.2, J15.3, J15.4, J15.5, J15.6, J15.7, J15.8, J15.9, J16.8, J17.0, J17.1, J17.2, J17.3, J17.8, J18.1, J18.2, J18.8, J18.9
Acute bronchitis	J20, J20.8, J20.9
Acute bronchiolitis	J21.9
Unspecified acute lower respiratory infection	J22

Disease causes	ICD-10
Allergic rhinitis	J30.4, J31.0, J31.1, J31.2
Chronic sinusitis	J32.0, J32.2, J32.4, J32.8, J32.9
Nasal polyp	J33.9
Chronic diseases of tonsils and adenoids	J35.0
Chronic Bronchitis	J40, J68.0, J68.9
Emphysema	J43.1, J43.9, J98.2
Chronic Obstructive Pulmonary Disease (COPD)	J44.0, J44.1, J44.8, J44.9
Asthma	J45.0, J45.1, J45.9, J46
Bronchiectasis	J47

2.2.3 Population Data

Population data were sourced from the National Statistical Office of Thailand. The dataset provided monthly population counts for each province, allowing for the calculation of monthly mortality rates per 10,000 population. The average monthly cardiorespiratory mortality per 10,000 population of each province in Thailand, 2015-2019, is presented in Figure 2.15.
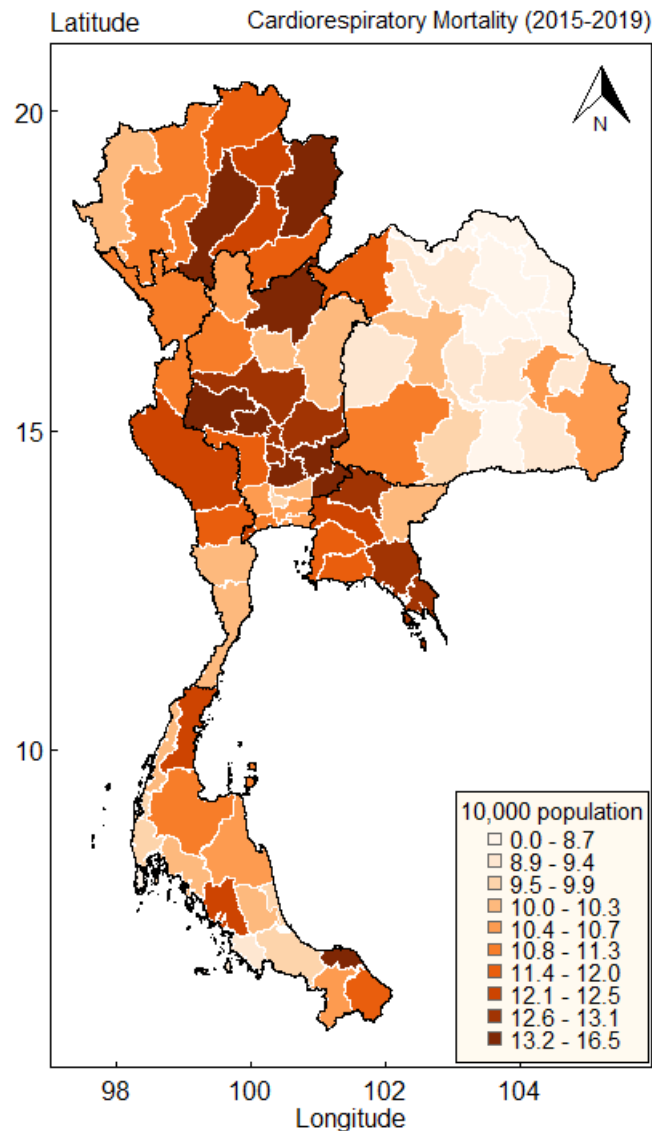
Figure 2.15: Cardiorespiratory mortality per 10,000 population in Thailand.

2.3 Data Analysis

This study investigates the relationship between monthly cardiorespiratory mortality rates and monthly PM2.5 levels. The outcome measure, monthly cardiorespiratory mortality rate per 10,000 population, was calculated by dividing the number of deaths in each province (identified using ICD-10 codes) by the monthly population of the respective province. Two determinant models were analyzed: monthly PM2.5 groups and Region-monthly PM2.5 groups. The latter combined four geographic regions with five categories of monthly PM2.5 levels (Table 2.6).

The analysis was conducted in three steps:

1) Descriptive Statistics, Time Series Plot, and Spearman's Correlation

These analyses explore temporal trends and the correlation between monthly PM2.5 levels and cardiorespiratory mortality rates.

2) Poisson Regression Model with Sum Contrast Confidence Intervals (CI)

This model identifies the statistical association between determinants and outcomes.

3) Moran's I Statistics

This test evaluates the spatial distribution of monthly PM2.5 levels and cardiorespiratory mortality rates.

All analyses were performed using R version 4.1.2 [53].

Table 2.6: The description of data used.

Variables	Types	Description
PM2.5 groups	Ordinal	Group of continuous values: • 0-20 μg/m^3 • 20.1-25 μg/m^3 • 25.1-30 μg/m^3 • 30.1-37.5 μg/m^3 • > 37.5 μg/m^3
Region	Nominal	Group of the province: • Central • North • Northeast • South
Region-PM2.5 groups	Nominal	A combination of four regions and five groups of PM2.5 are 20 groups.
Cardiorespiratory mortality	Discrete	Count data and convert to rate by 10,000 population

2.3.1 Spearman's Correlation

Spearman's correlation is a nonparametric alternative to Pearson's correlation, suitable for analyzing curved, monotonic relationships or ordinal data. The correlation coefficient (r) ranges from -1 to +1. Positive values indicate a monotonic positive relationship, while negative values indicate a monotonic negative relationship. A coefficient close to -1 or 1 represents a strong correlation, whereas a value near 0 indicates no correlation [54].

2.3.2 Poisson Regression Models

A Poisson regression model was employed, suitable for count data where the variance equals the mean [55]. This generalized linear model assumes a Poisson distribution. Sum contrasts, which simplify interpretation by comparing each group to the overall mean, were used in model fitting [56].

Model 1: Monthly PM2.5 Groups

The model estimating cardiorespiratory mortality rates (λ_j) is given by:

$$\log\left(\frac{\lambda_j}{P_j}\right) = \mu + \beta_j \tag{2.1}$$

where:

λ_j: Mean of the Poisson distribution, representing the estimated cardiorespiratory mortality rate per 10,000 population (P_j).

μ: Constant capturing the overall incidence.

β_j: Coefficients for the monthly PM2.5 groups ($j = 1,2,\ldots,5$).

Model 2: Region-Monthly PM2.5 Groups

The model estimating cardiorespiratory mortality rates (λ_j) by region and PM2.5 group is given by:

$$\log\left(\frac{\lambda_j}{P_j}\right) = \mu + \gamma_j \tag{2.2}$$

where:

λ_j: Mean of the Poisson distribution, representing the estimated cardiorespiratory mortality rate per 10,000 population (P_j).

μ: Constant capturing the overall incidence.

γ_j: Coefficients for the Region-monthly PM2.5 groups ($j = 1,2,\ldots,20$).

The fitted models were focused on the use of subjective sum contrasts. This comparison method was used to evaluate the proportion of the result in each group relative to the average proportion and illustrate the 95% CI. In unbalanced designs, weighted sum contrasts, as explained by Tongkumchum & McNeil (2009), were used to generate and display the 95% confidence range for comparing population averages [56]. These CIs of the Poisson regression model provide an easy criterion for classifying

subgroups of factors into three sets: those whose associated CIs exceed, cross, or fall below the overall mean.

We employed Poisson regression with Generalized Additive Models (GAMs) and Distributed Lag Non-Linear Models (DLNMs) to examine the associations between PM2.5 and cardiorespiratory mortality. By combining Poisson regression with GAMs and DLNMs, we create a well-suited modeling framework for capturing the nuanced and dynamic relationships between PM2.5 and mortality rates, considering both non-linear exposure-response patterns and delayed effects over time [57]. The model was fitted using the log-link function and the cubic smoothing spline technique, with the assumption that the error structures followed the Poisson distribution. The cumulative lag effects of PM2.5 concentrations were investigated with a 15-day lag to address PM2.5 and its lag effect on mortality rates. The rate of increase in mortality rates equal to a 10 % rise in local PM2.5 levels was used to determine the health effects of PM2.5. The data are shown as relative risk (RR) or enhanced risk (RR – 1) × 100 %, along with the corresponding 95 % CI.

2.3.3 Moran's I Statistics

Moran's I, developed by Moran (1950), is a statistic used to measure spatial autocorrelation, which is the similarity of values for a variable at nearby locations [58]. It is a weighted correlation coefficient that can detect departures from spatial randomness and determine whether neighboring areas have similar values than expected under the null hypothesis of no spatial autocorrelation. Moran's I is:

$$I = \frac{N \sum_{i=1}^{N} \sum_{j=1}^{N,j \neq i} w_{ij} z_i z_j}{S_0 \sum_{i=1}^{N} z_i^2} \tag{2.3}$$

where N equals the number of provinces, w_{ij} is a weight denoting the strength of the connection between areas i and j, z_i is the rate in province i centered about the mean rate (usings $z_i = x_i - \bar{x}$; x_i is the rate in province i), i is the province number ($i = 1,2,3,…,77$), j is the number around the province of i, and S_0 is the sum of the weights

$$S_0 = \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij,i \neq j} \tag{2.4}$$

Values of Moran's *I* typically range from -1 to +1, with a value close to -1 indicating perfect clustering of dissimilar values, a value close to 0 indicating no autocorrelation, and a value close to 1 indicating perfect clustering of similar values.

Using only one statistic from Moran's *I* to represent an entire research region would be inaccurate, as it is likely that the statistics will vary throughout the study area. To overcome this limitation, the Local Indicator of Spatial Association (LISA) is used, considering that Moran's *I* is a sum of individual cross-products, to assess the clustering in specific units. This is achieved by computing Local Moran's *I* for each geographical unit and determining the statistical significance for each I_i [59]. Using this method, it is possible to determine the presence of spatial autocorrelation at a local level rather than just relying on a single statistic for the entire study area. Local Moran's *I* equation:

$$I_i = \frac{z_i}{m_2} \sum_{j=1}^{N} w_{ij} z_j \qquad (2.5)$$

where, $\qquad m_2 = \frac{\sum_{i=1}^{N} z_i^2}{N}$

then, $\qquad I = \sum_{i=1}^{N} \frac{I_i}{N}$

I is the Moran's *I* measuring global autocorrelation, I_i is local, and *N* is the number of analysis units on the map.

The functions moran.test and localmoran from the spdep package in R were used to compute Moran's I and Local Moran's I, respectively.

2.4 Research Methodology Diagram

Figure 2.16 presents the overall study framework, beginning with **data collection**, which incorporates multiple sources such as satellite-derived PM2.5 concentration estimates, cardiorespiratory mortality records from the NHSO, and population statistics from the NSO. At this stage, factor and model selection are conducted to identify relevant variables and appropriate analytical approaches. Satellite data processing, machine learning–based PM2.5 estimation, and preliminary analyses are also performed to produce initial summaries (**section 1**).

The next step is **data integration**, where these diverse datasets are merged into a unified analytical framework to ensure consistency and comparability across variables.

In the **statistical analysis** phase, descriptive statistics (mean, SD, and range) are computed to provide an overview of the data. Correlations between variables are assessed using Spearman's correlation. Advanced models are then applied: Poisson regression to analyze mortality counts, Moran's I statistic to assess spatial autocorrelation, and more flexible approaches such as GAMs and DLNMs to capture complex exposure–response relationships. Subjective sum contrasts are also applied for clearer interpretation of model results.

The **visualization** stage ensures effective communication of findings through time-series plots (to depict temporal trends), CI plots (to illustrate estimate precision), and spatial maps (to highlight geographic variations).

Finally, the **results** are synthesized, focusing on the impact of PM2.5 exposure on cardiorespiratory mortality. These findings provide scientific evidence that can inform public health strategies and policy interventions aimed at reducing the health burden of air pollution (**section 2**).
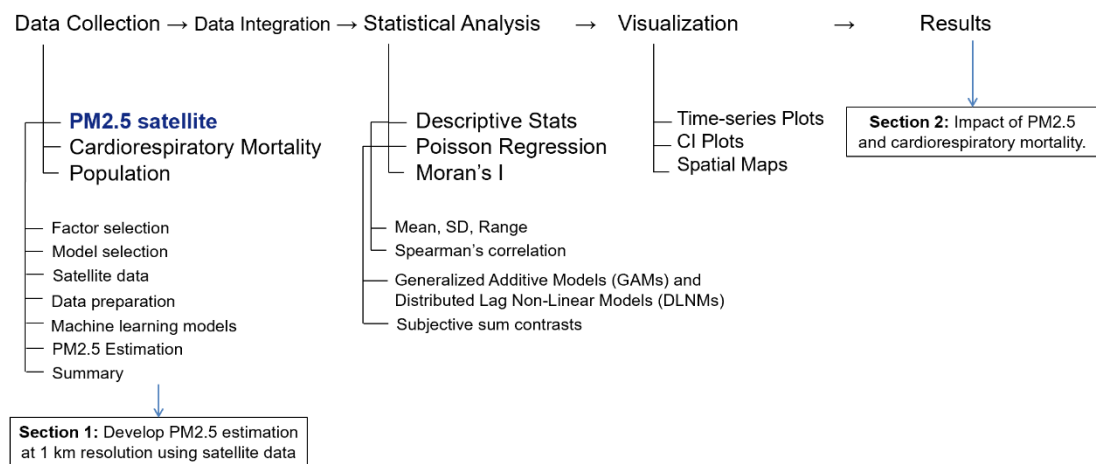
Figure 2.16: The research methodology diagram.

Chapter 3
PM2.5 with Satellite Data

This chapter uses satellite data to reveal Thailand's estimated ground-level hourly PM2.5 concentrations. The approach employs a weighted sum contrast log-linear regression model that incorporates satellite data, enabling the examination of small-scale hourly variations in PM2.5 concentrations. The proposed model has proven valuable for interpretation and practical application, providing comparable estimated hourly PM2.5 concentrations at a 1 km resolution alongside monitoring stations.

3.1 Study Background and Purpose

Suspended particulate matter from human and natural sources significantly impacts the climate and environment [60], [61]. Epidemiological studies indicate that atmospheric particles with an aerodynamic diameter of less than 2.5 μm (PM2.5) can cause severe harm to human health, resulting in increased deaths due to cardiovascular and respiratory diseases and lung cancer [62], [63], [64]. Due to the uneven distribution of monitoring sites, analyzing the spatial and temporal distribution of PM2.5 particles through traditional ground-based monitoring is restricted. Hence, it is crucial to acquire precise and dependable estimates of PM2.5 concentrations with high coverage and resolution to evaluate the impact of air quality on public health.

Various studies have revealed a substantial correlation between surface PM2.5 concentrations and the AOD obtained from satellite data. However, to enhance the accuracy of the modeling, other variable factors are frequently incorporated [27], [65]. With advancements in satellite technology, it is now possible to estimate ground-level PM2.5 concentrations using satellite data products. This approach is expected to provide accurate and reliable estimates of PM2.5 concentrations with high coverage and resolution, overcoming the limitations of conventional ground-based monitoring systems. While the previously mentioned studies primarily used AOD data from the MODIS, other satellite sensors, including the Visible Infrared Imaging Radiometer Suite (VIIRS), Multiangle Imaging Spectro Radiometer (MISR), and Korean Geostationary Ocean Colour Imager (KGOCI), also offer AOD products [66], [67],

[68], [69], [70]. However, these sensors, situated on polar orbit satellites, provide only one observation per day, leading to a limited representation of aerosol distribution during the day [71], [72].

Studies have established a strong correlation between AOD and PM2.5, with statistical models developed to relate the two [73], [74], [75]. By combining satellite remote sensing AOD with ground-based PM2.5 observations and employing statistical methods, it is possible to generate PM2.5 estimates with extensive spatial coverage [76]. In a review on predicting ground PM2.5 concentrations using satellite AOD, MLR (25 articles), Mixed-Effect Model (MEM) (23 articles), Chemical Transport Model (CTM) (16 articles), and Geographically Weighted Regression (GWR) (10 articles) were widely employed [27]. However, no clear "best" model was identified, as each method has its strengths and limitations.

The MLR model has been extensively employed since 2005 to predict PM2.5 levels using satellite AOD data. In this model, AOD is the independent variable, while ground-level PM2.5 is the dependent variable. Previous research utilized the MLR model to forecast PM2.5 concentrations in various regions, including cities, suburbs, and the countryside in the eastern United States during 2001 [77]. They observed significant variations in coefficients among regions, resulting in low R^2 values of 0.420, 0.490, 0.590, and 0.430 in cities, suburban areas, the countryside, and the entire region, respectively. Recent studies have explored covariate factors within the MLR model to enhance the model's performance across different circumstances [78], [79], [80], [81], [82].

To estimate PM2.5 using satellite data analysis with MLR, it is necessary to utilize treatment contrasts when the independent variables are categorical. Treatment contrasts in linear regression involve a linear combination of predictors whose coefficients add up to zero, facilitating the comparison of various treatments [83]. Initially, this method was developed to compare one or more treatment groups with a control group by setting the control group's parameter to zero and allowing the treatment parameters to reflect their respective treatment effects. Nonetheless, the propose an alternative approach for constructing 95% CI to compare means without selecting a reference group [56]. This approach provides informative 95% CIs for comparing each mean with the overall mean and employs different contrasts referred

to as "sum" contrasts when a control group is absent. Sum contrasts restrict the parameters associated with each factor level, indicating the difference between that level and the overall mean outcome. It is worth noting that this technique has not yet been applied to estimate PM2.5 using satellite data.

Weighted sum contrasts have been commonly utilized in previous studies employing linear and logistic regression models. For instance, utilized weighted sum contrasts in linear regression to evaluate the influence of land-cover transformation and EV on decadal changes in LST by comparing the adjusted mean of all factors [84]. Conducted a comparison of blood lead levels among children in the Pattani River region of Thailand [85], while the examined HIV mortality by age group and gender in Thailand between 2014 and 2015 [86]. Additionally, employed weighted sum contrasts logistic regression to investigate land-use change in Thailand [87], [88], [89]. Furthermore, utilized the same technique to explore the increase in LST in Bali, Indonesia, from 2001 to 2020 [90].

PM2.5 concentration exhibits temporal and spatial variations that require high-resolution monitoring, typically not met by polar-orbiting satellites [91]. To address this issue, geostationary meteorological satellites have been increasingly used to estimate PM2.5 [92]. The PM2.5 data collected at ground station sites are typically recorded hourly, necessitating the development of accurate methods to estimate hourly levels to cover all areas of study. No study has been done for the estimation of ground-level hourly PM2.5 in Thailand. Some have been done for the estimation of daily PM2.5 [93], [94]. Our previous research also mentioned the satellite data that can be used to estimate PM2.5 in Thailand and to model daily PM2.5 in Thailand [95], [96]. It begins with AOD as a base factor and then adds other variables to improve accuracy in estimating PM2.5 levels in Thailand. Specifically, we have selected LST, NDVI, and EV data to represent land use and cover, as well as time and WOY as time and seasoning factors. Therefore, this study uses satellite data from our previous research to estimate ground-level hourly PM2.5. We aim to develop an easily interpretable method for estimating hourly PM2.5 levels using a weighted sum contrasts linear regression approach. Our proposed method will be valuable for researchers and policymakers in understanding the hourly fluctuations of PM2.5 and their impact on human health and the environment.

3.2 PM2.5 Data and Area of Study

Thailand, a dynamic nation in Southeast Asia, boasts a diverse geography that encompasses the Andaman Sea and the Gulf of Thailand. With a population of approximately 70 million, Thailand is traditionally divided into four regions: central, north, northeast, and south. The four-region system is the administrative classification developed by the Ministry of Interior and used for statistical or academic purposes. It holds significant regional importance, consisting of 77 provinces and covering an extensive area of 513,120 square kilometers. The Pollution Control Department (PCD) is a legally recognized government agency in Thailand that collects data on air pollution parameters throughout the country. We utilized hourly PM2.5 data from 67 stations in 2020, as presented in Figure 3.1.
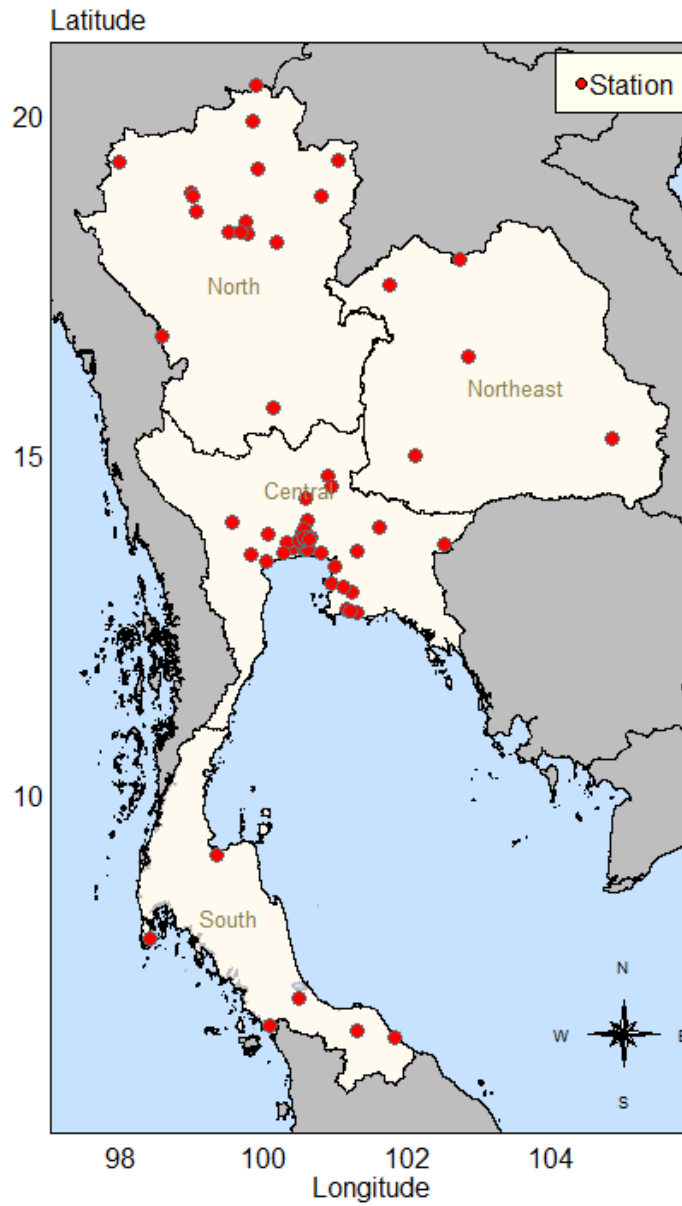
Figure 3.1: PM2.5 station in 2020.

3.3 Satellite Data

In this study, we utilized AOD, LST, NDVI, and EV data obtained from the satellite products of MODIS. All the data used in this research were retrieved from the DAAC, made available through the NASA EOSDIS.

We processed data obtained from the MCD19A2 product, including AOD (AOD at 045 Microns) retrieved from Terra and Aqua satellites. The AOD data is collected twice daily at 10:30 a.m. and 1:30 p.m. local standard time, with a spatial

resolution of 1 km per pixel. The LST data from Terra (MOD11A1 product) and Aqua (MYD11A1 product) satellites were combined to calculate the daily average LST values, taking the arithmetic mean when data from both satellites were available or using the data from a single satellite if only one was operational on a specific day.

Additionally, the MOD13A1 NDVI product, with a temporal resolution of 16 days and a spatial resolution of 500 meters, was employed to depict land cover changes and monitor global vegetation conditions. This dataset provides valuable insights for modeling biogeochemical and hydrologic processes, understanding climates at global and regional scales, and characterizing various biophysical features and processes on the ground surface. Lastly, the EV data from the "Land Digital Elevation Model (MODDEM1KM) - Land/sea mask and digital elevation model" with a spatial resolution of 1 km were utilized.

- Downloading satellite data

Satellite data can be obtained from the Find Data - LAADS DAAC (nasa.gov) portal (https://ladsweb.modaps.eosdis.nasa.gov/search/). Users are required to register for an EARTHDATA LOGIN, as presented in Figure 3.2. There are a total of five steps involved in downloading the data.

Figure 3.2: Register for EARTHDATA LOGIN.

First, select the PRODUCT that needs to be downloaded. Figure 3.3 presents an example of the Terra and Aqua satellites' AOD data from the MCD19A2 product.

Figure 3.3: Select PRODUCT.

Second, select the TIME of the period of study. Figure 3.4 presents an example of date selection from 1-15 January 2020.
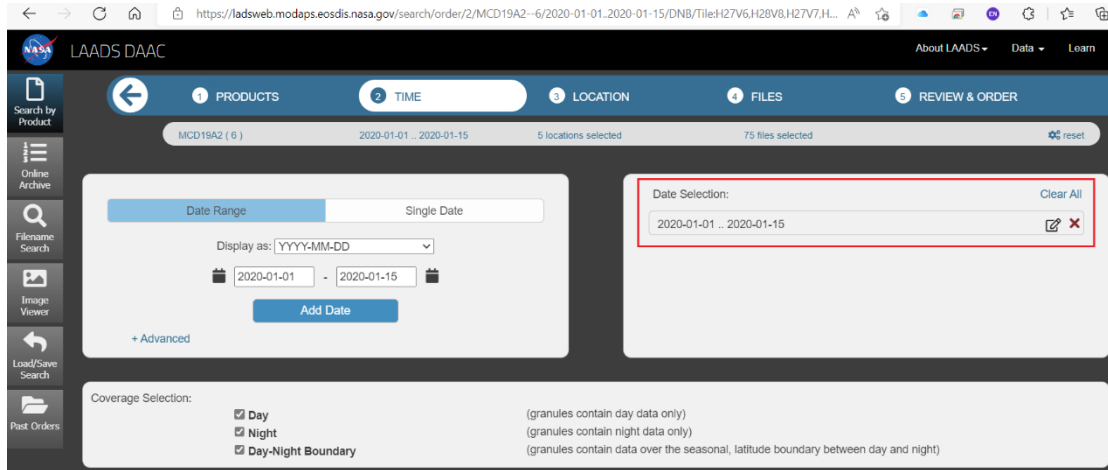
Figure 3.4: Select TIME.

Third, select the LOCATION of the study. Figure 3.5 presents the location of Thailand. The satellite image retrieval consists of five layers, including *h27v06, h27v07, h27v08, h28v07,* and *h28v08*. The horizontal tile represents the *h*, while the vertical tile is the *v*.
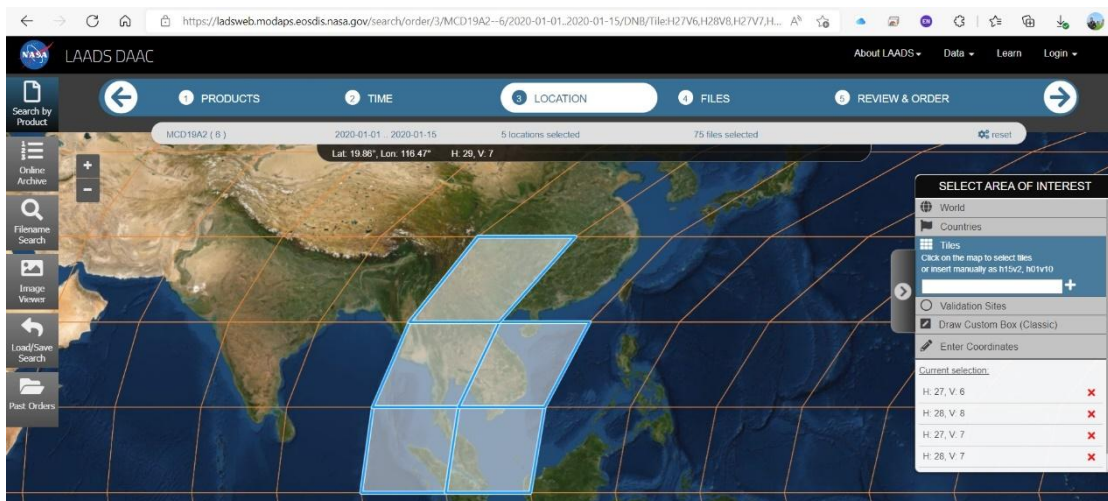
Figure 3.5: Select LOCATION.

Fourth, select the FILES that need to be downloaded. Figure 3.6 presents the file for 15 days, which includes 75 files. We can select all to download all files.
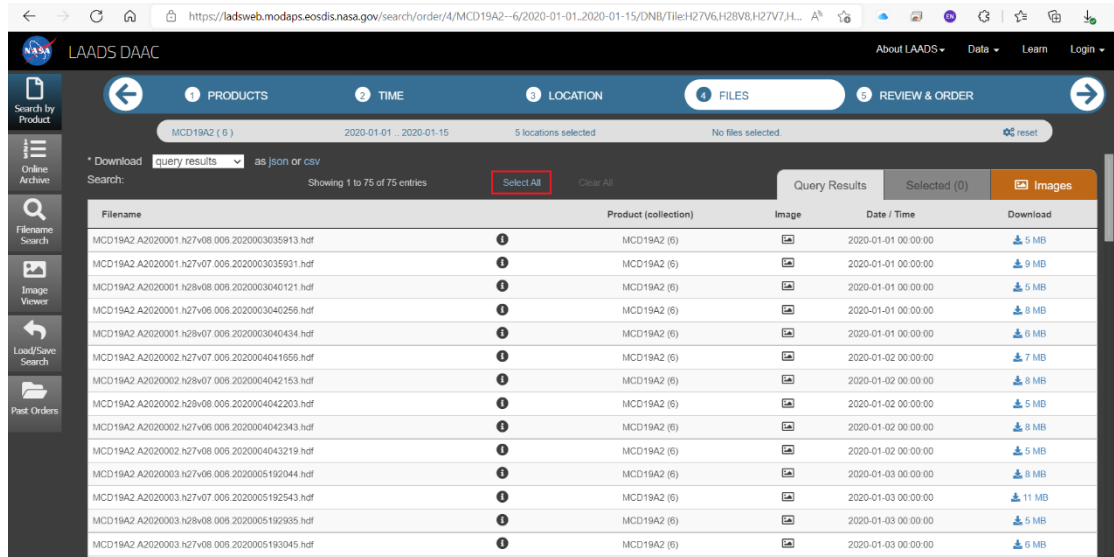
Figure 3.6: Select FILES.

Fifth, select REVIEW & ORDER. Figure 3.7 presents the tab of three steps. (1) Sds and select "Optical_Depth_047", (2) Reformat select "Convert products GeoTIFF format" and (3) Submit Order.
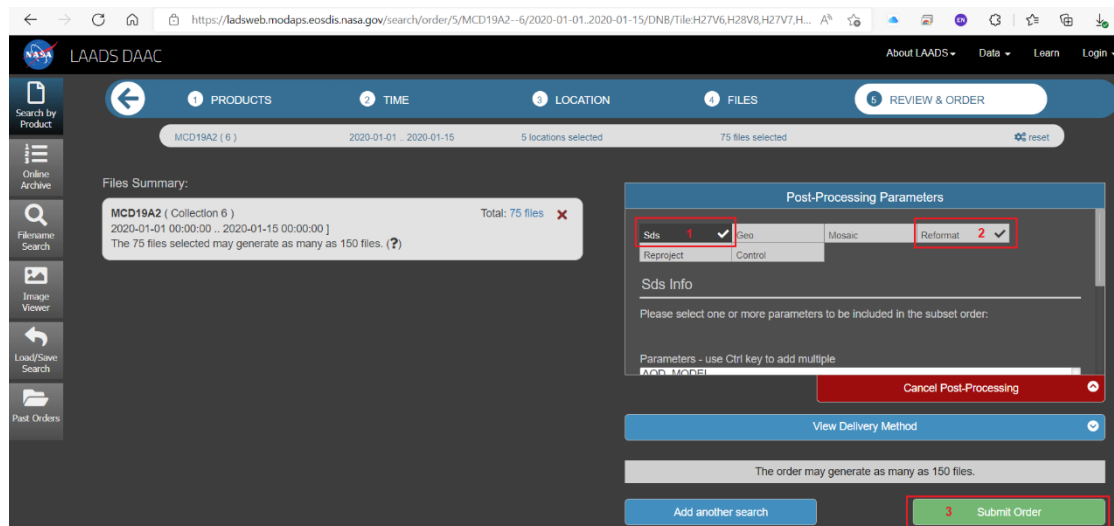
Figure 3.7: Select REVIEW & ORDER.

After submitting the order data, the order notification will be sent via email. We can click the link to access the download data (Figure 3.8).
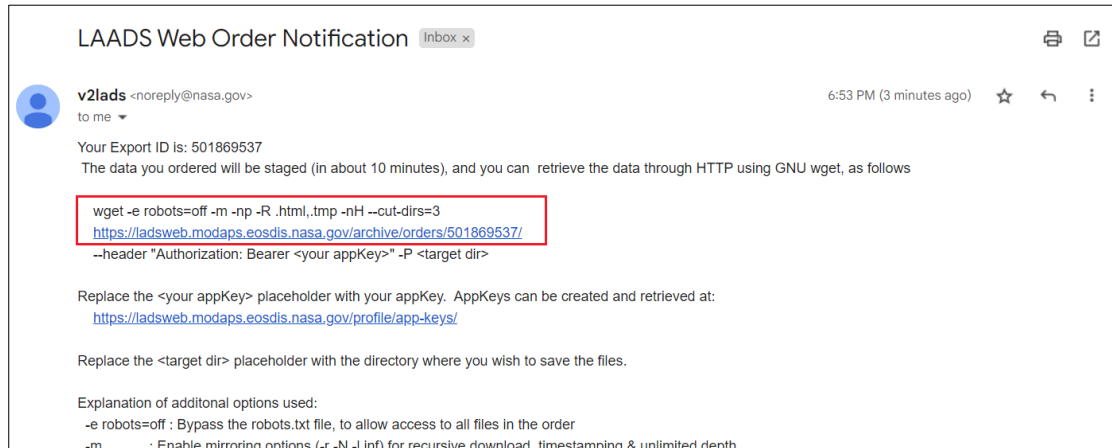
Figure 3.8: Order Notification.

As each location's TIFF file is typically divided into 1-5 files, it is necessary to merge them. An example of this can be seen in Figure 3.9, which shows the merging of three files for the *h27v07* location on January 1, 2020.
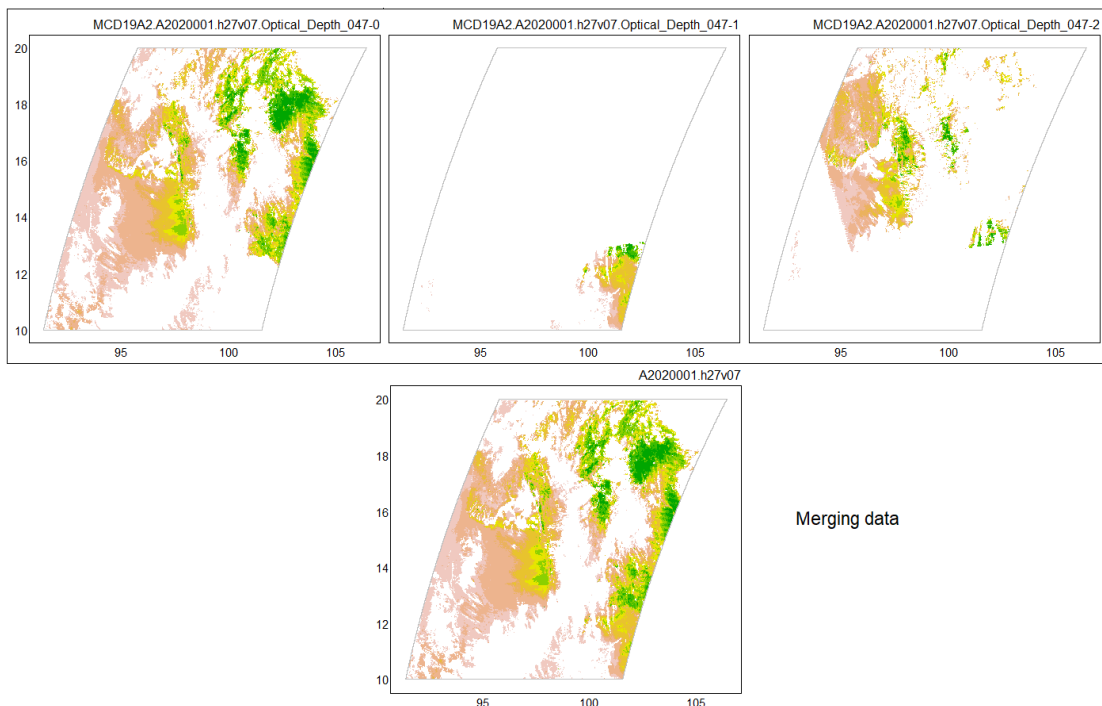
Figure 3.9: Merging data.

Following the merging process, it is necessary to combine the data. For example, Figure 3.10 presents the combination of data for the locations *h27v06, h27v07, h27v08, h28v07,* and *h28v08*.
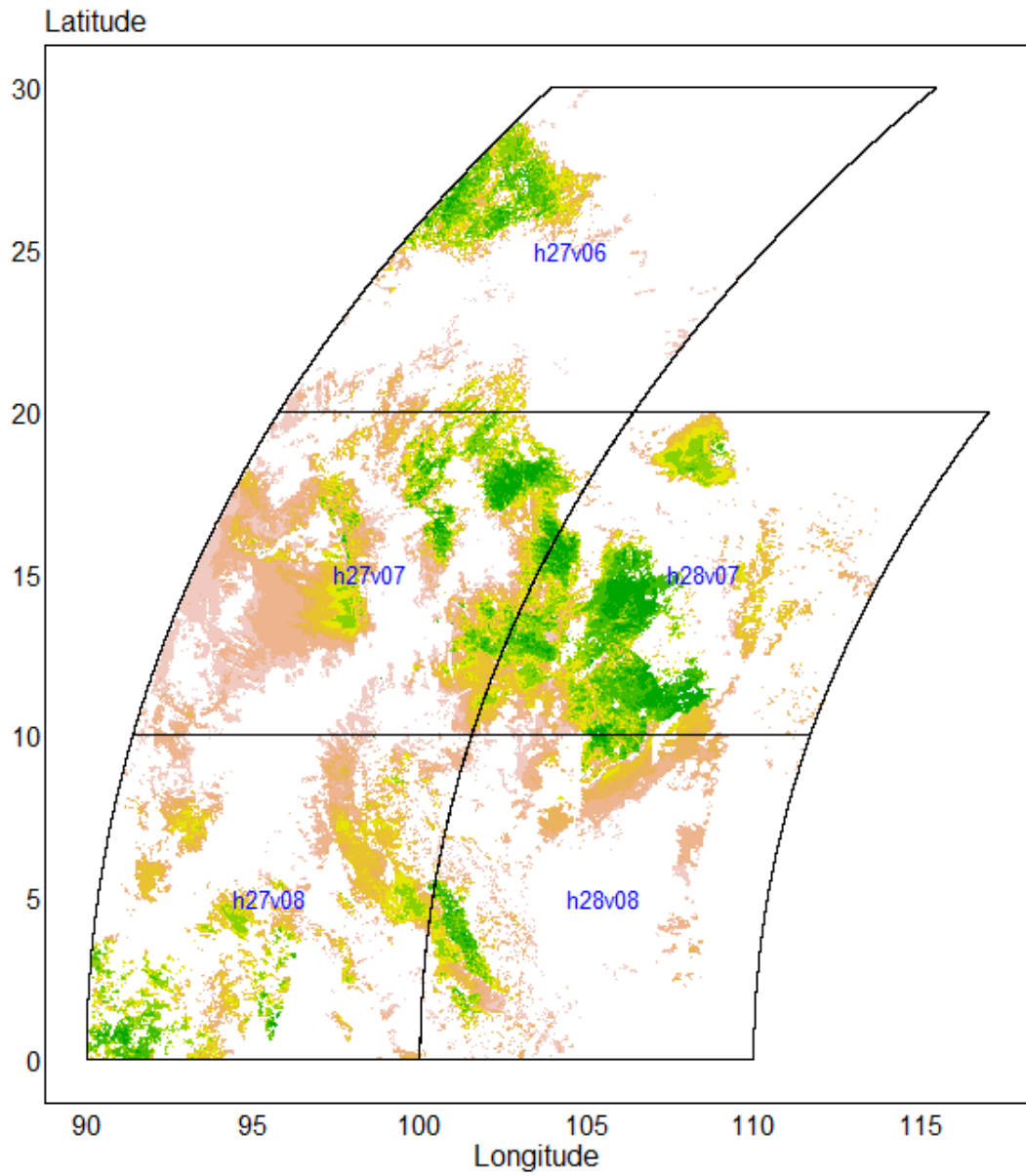
Figure 3.10: Combining data.

Cropping is performed to select a specific region of interest to refine the data further. An example of this can be seen in Figure 3.11, which shows the cropping of data boundaries for Thailand.
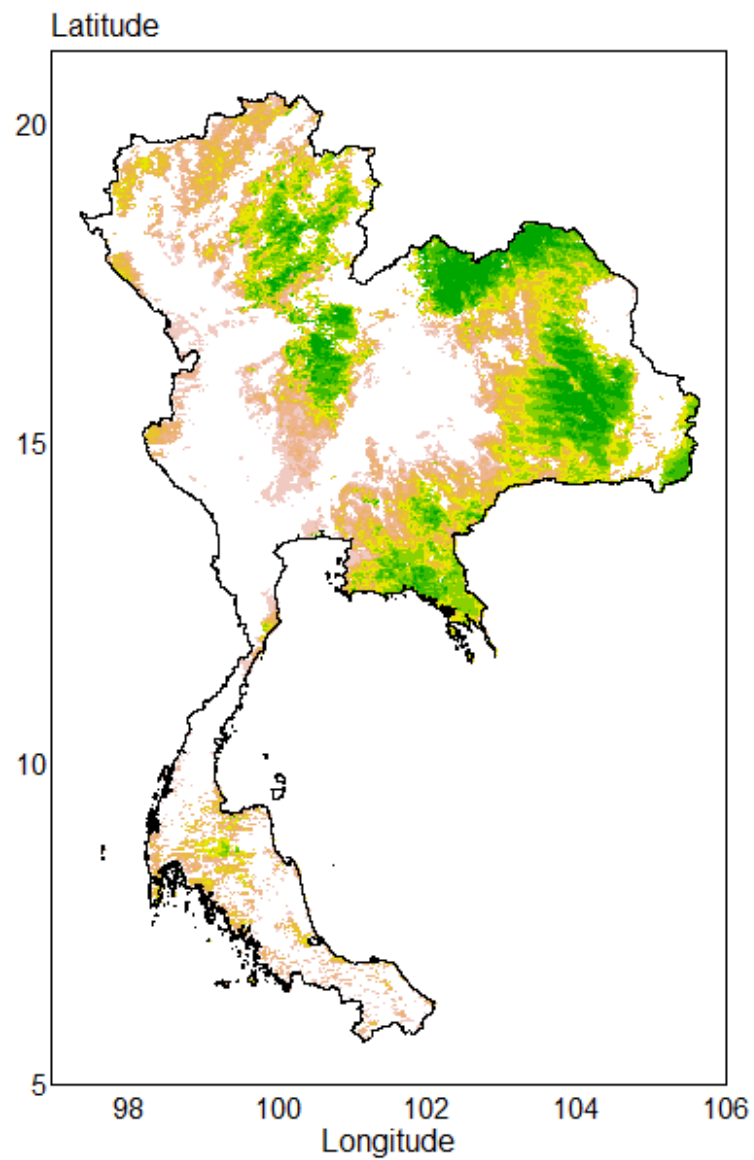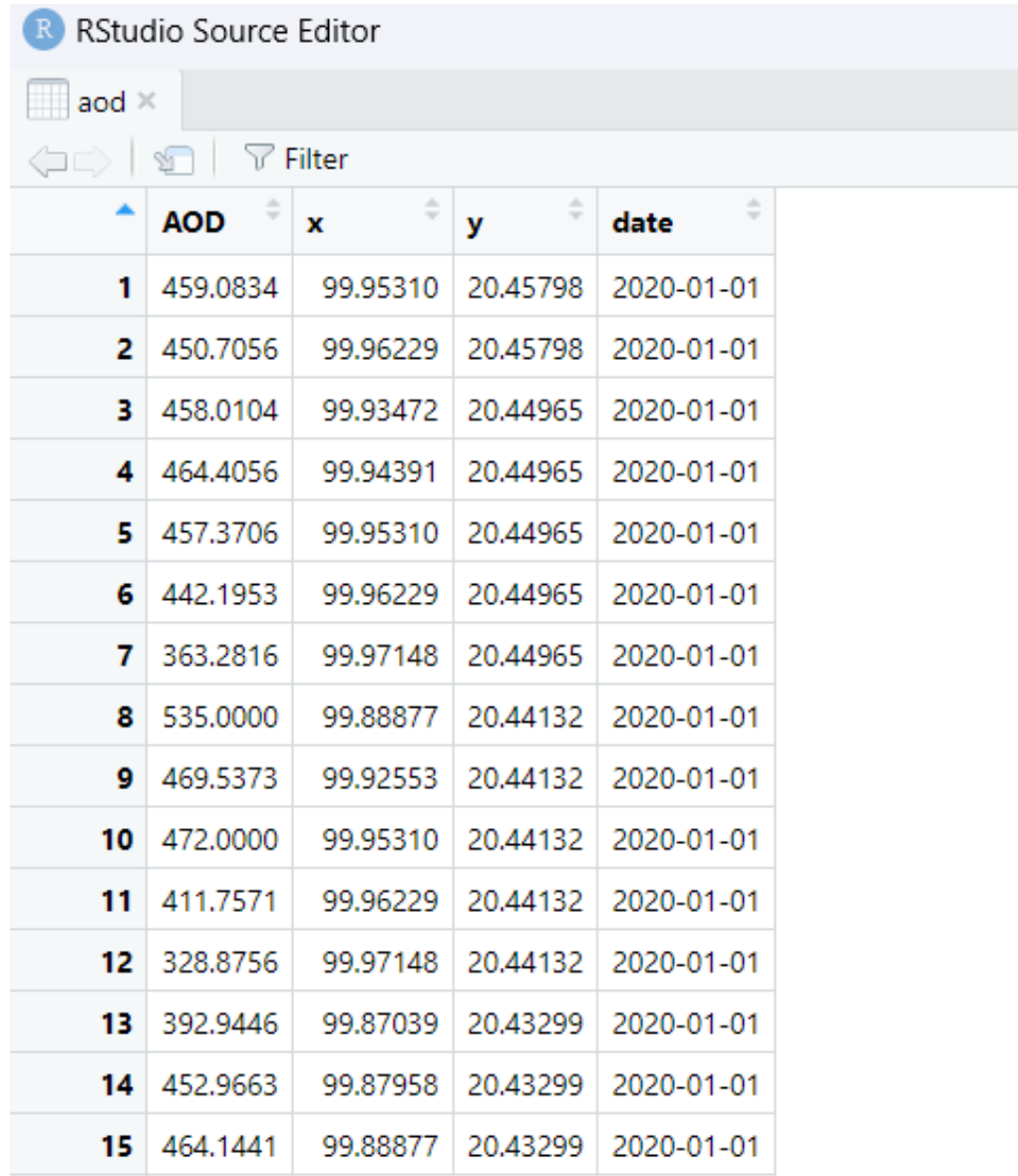
Figure 3.11: Cropping data.

Extracting data is to extract satellite data for estimating PM2.5 in 1 km resolution. An example of AOD data can be seen in Figure 3.12, which shows the data table that includes AOD values, x (Longitude), y (Latitude), and date.

	AOD	x	y	date
1	459.0834	99.95310	20.45798	2020-01-01
2	450.7056	99.96229	20.45798	2020-01-01
3	458.0104	99.93472	20.44965	2020-01-01
4	464.4056	99.94391	20.44965	2020-01-01
5	457.3706	99.95310	20.44965	2020-01-01
6	442.1953	99.96229	20.44965	2020-01-01
7	363.2816	99.97148	20.44965	2020-01-01
8	535.0000	99.88877	20.44132	2020-01-01
9	469.5373	99.92553	20.44132	2020-01-01
10	472.0000	99.95310	20.44132	2020-01-01
11	411.7571	99.96229	20.44132	2020-01-01
12	328.8756	99.97148	20.44132	2020-01-01
13	392.9446	99.87039	20.43299	2020-01-01
14	452.9663	99.87958	20.43299	2020-01-01
15	464.1441	99.88877	20.43299	2020-01-01

Figure 3.12: Extracting data.

3.4 Statistical Analysis

    1) Study Variables

We examine the association between hourly PM2.5 concentrations and satellite data in 2020. To achieve this, we matched the PM2.5 concentrations for each station with the average satellite data within a 5 km radius. Additionally, we employed data imputation techniques, specifically using the nearest date and pixel. This matching process was accomplished using station latitude-longitude and date variables. We then analyzed the relationship between hourly PM2.5 concentrations and various satellite variables, including AOD, LST, NDVI, and EV, as well as the time and WOY variables. The satellite data were grouped into ten levels using the ten quantiles to facilitate our analysis. It is important to note that the time variable was measured over 24 hours, while the WOY variable was measured over 53 weeks. The summary of the data used in this study, including variables, units, temporal resolution, spatial resolution, and source, is presented in Table 3.1.

Table 3.1: Summary of the data used in this study.

Variable	Unit	Temporal resolution	Spatial resolution	Source
PM2.5	$\mu g/m^3$	1 hour	Site	PCD
AOD	Unitless	1 day	1 km	MODIS
LST	°C	1 day	1 km	MODIS
NDVI	Unitless	16 days	500 m	MODIS
EV	Meter (m)	-	1 km	MODIS
Time	hour	1 hour	-	-
WOY	week	1 week	-	-

    2) Linear Regression Based Weighted Sum Contrasts

This study focused on hourly PM2.5 values as the outcome of interest, along with a group of AOD, LST, NDVI, and EV, as well as categories of WOY and time variables as determinants. When a categorical variable is used as a determinant in a regression model, it is referred to as a factor. The model formula features a set of *k-1* parameters, where *k* refers to the number of distinct categories. The linear regression model formulated as $y = a + factor(x_1) + factor(x_2) + factor(x_3) + factor(x_4) + factor(x_5) + factor(x_6)$, where y is the hourly PM2.5, *a* is the constant term, x_1 is AOD group, x_2 is LST group, x_3 is NDVI group, x_4 is EV group, x_5 is time, and x_6 is WOY. Since the factors only have a *k-1* parameter, then the model was formulated as follows:

$$y = a + \sum_{i=2}^{k=10} b_i x_{1i} + \sum_{i=2}^{k=10} c_i x_{2i} + \sum_{i=2}^{k=10} d_i x_{3i}$$

$$+ \sum_{i=2}^{k=10} e_i x_{4i} + \sum_{i=2}^{k=24} f_i x_{5i} + \sum_{i=2}^{k=53} g_i x_{6i} \tag{3.1}$$

where b_i, c_i, d_i, e_i, f_i, and g_i were the coefficients of x_1, x_2, x_3, x_4, x_5, and x_6 at identity i, respectively, and b_1, c_1, d_1, e_1, f_1, and g_1 were equal to 0 in the case that no contrast option was selected when specifying the model. Usually, without a specified contrast option, the parameters are alphabetically assigned to each factor level, with the first parameter being set to 0. Termed "treatment" contrasts, these contrasts were initially employed in experiments that compared one or more treatment groups to a control group. By setting the parameter corresponding to the control group to 0, these contrasts ensure that the parameters associated with the treatments accurately capture the real treatment effects. However, given that our research did not incorporate a control group, we excluded six categories from our analysis of determinants and covariates.

Our study aimed to compare the mean of all relevant factors to evaluate their impact on hourly PM2.5 levels. In order to achieve our objective, we implemented the weighted sum contrasts [56]. This approach recommends constructing 95% CI to compare means without the need for selecting a reference group, thereby offering informative intervals for comparing each mean with the overall mean. When there is no control group available, we employed "sum" contrasts, which restrict the parameters linked to each factor level, allowing us to measure the difference between that level and the overall mean of the outcome. The formulation bears a resemblance to that of treatment contrasts but includes extra terms (b_1, c_1, d_1, e_1, f_1, and g_1), that is,

$$y = a + b_1 x_{11} + \sum_{i=2}^{k=10} b_i x_{1i} + c_1 x_{21} + \sum_{i=2}^{k=10} c_i x_{2i}$$

$$+ d_1 x_{31} + \sum_{i=2}^{k=10} d_i x_{3i} + e_1 x_{41} + \sum_{i=2}^{k=10} e_i x_{4i}$$

$$+ f_1 x_{51} + \sum_{i=2}^{k=24} f_i x_{5i} + g_1 x_{61} + \sum_{i=2}^{k=53} g_i x_{6i} \tag{3.2}$$

During the linear regression using the weighted sum contrasts method, several statistical model parameters were calculated, including the overall mean (the average hourly PM2.5), overall adjusted R^2 (derived from the regression model without contrast option), crude mean (the PM2.5 mean in each factor category), and adjusted R^2 and p-value for each factor. These parameters were compared and investigated. Furthermore, 95% CI were computed for each factor category using the "democratic" approach to evaluate the mean variation. To compare means effectively, 95% CI for the difference between means were employed, representing the difference between each mean and the overall mean. These 95% CI align with the p-value and do not necessitate the selection of a control group for comparison, rendering them "democratic" in nature.

3.5 Analysis Results

3.5.1 Normal Distribution Test

The left-hand plot of Figure 3.13 presents the quantile-quantile (Q-Q) plot of PM2.5, revealing a highly skewed distribution. In contrast, the right-hand plot depicts the distribution after log-transforming PM2.5, which appears to follow a normal distribution. Consequently, we will utilize the log-transformed PM2.5 outcome in this study, as it aligns with the assumption of the linear regression model.
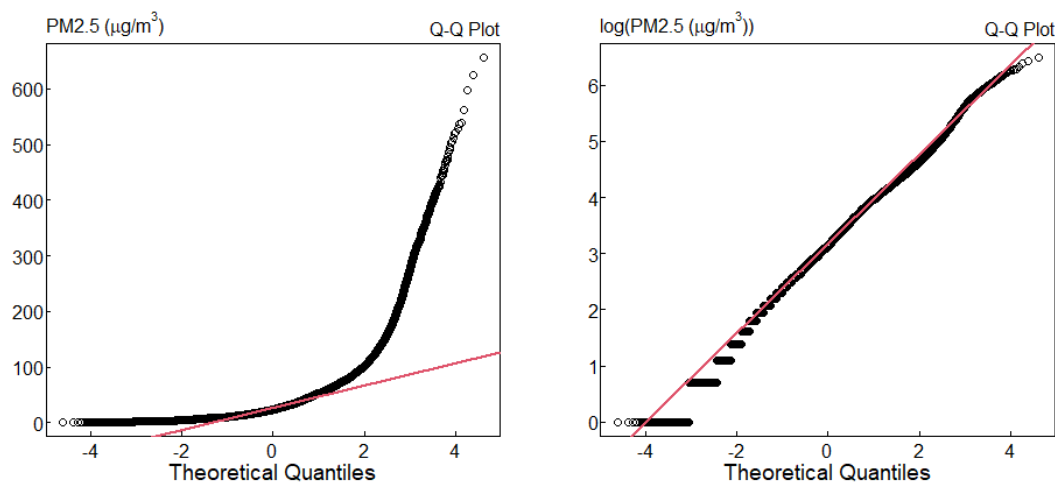
Figure 3.13: Normal Q-Q plots of PM2.5 and log-transformed PM2.5.

3.5.2 Select Variables

We generated a log-linear regression model by iteratively adding and removing predictor variables based on their p-values, starting from a set of potential predictor variables. The final model, presented in Table 3.2, incorporates all the variables, resulting in an R^2 value of 53.8%, supported by significantly low p-values (<0.001). Notably, the factor variables WOY (+22.4) and AOD (+20) contributed the most to the model's accuracy improvement.

Table 3.2: Variable selection in a stepwise regression model.

Response: log(PM2.5) Factor variable	R^2	RMSE	P-value*
AOD+Time	20 (+20)	26.3	< 0.001
AOD+LST+Time	21.7 (+1.7)	26.2 (-0.1)	< 0.001
AOD+LST+NDVI+Time	24.8 (+3.1)	25.8 (-0.4)	< 0.001
AOD+LST+NDVI+EV+Time	31.4 (+6.6)	24.3 (-1.5)	< 0.001
AOD+LST+NDVI+EV+Time+WOY	53.8 (+22.4)	22.2 (-2.1)	< 0.001

*All factor variable has p-value < 0.001.

3.5.3 A Log-Linear Model with Sum Contrast Analysis Results

Figure 3.14 presents the estimated mean of hourly PM2.5 and the comparative 95% CI (plus sign) after adjusting for various factors, including AOD, LST, NDVI, EV, time, and WOY, relative to the overall mean for each factor. The horizontal red lines represent the simple average hourly PM2.5 from all stations, while the blue dots indicate the crude means for each factor group. The "r-sq:" label indicates the adjusted R^2 value obtained from regression fitting with separated determinants. The "Overall r-sq:" label denotes the adjusted R^2 value obtained from the regression model adjusting for all factors.

The overall mean of hourly PM2.5 in 2020 was 23.1 $\mu g/m^3$, falling below the NAAQS of 25 $\mu g/m^3$ but still exceeding the WHO guidelines of 5 $\mu g/m^3$. The overall R^2 value was 53.8%, and all factors were significant, with p-values < 0.001. The individual R^2 values were ranked in descending order, with WOY accounting for 33.8%, followed by AOD (18.7%), NDVI (9.8%), EV (6.4%), LST (4.8%), and time (1.2%).

The analysis revealed that the WOY factor had higher than the overall mean and NAAQS hourly PM2.5 levels from January to March (1-9) and November to December (46-53). The AOD factor was positively correlated with PM2.5, particularly when it exceeded 0.52, resulting in PM2.5 levels higher than the overall mean. Similarly, LST was positively associated with PM2.5, particularly when it exceeded 33.9°C, leading to PM2.5 levels higher than the overall mean. For NDVI, PM2.5 levels were higher than the overall mean within the range of -0.08 to 0.18, indicating areas likely with no green leaves and possibly urbanized. Furthermore, the EV factor showed higher PM2.5 levels than the overall mean in areas above 67.9 m. Lastly, the time variable indicated that PM2.5 levels were higher than the overall mean during 8-11 a.m. and 20-24 p.m.

AOD		LST (°C)		NDVI		EV (m)	
1: 0.02-0.14	6: 0.43-0.51	1: 2.3-27.9	6: 33.9-35	1: -1 - -0.16	6: 0.08-0.13	1: < 5.2	6: 31.7-67.9
2: 0.15-0.21	7: 0.52-0.61	2: 28-29.9	7: 35.1-36.2	2: -0.15 - -0.09	7: 0.14-0.18	2: 5.3-6.1	7: 68-186.8
3: 0.22-0.28	8: 0.62-0.76	3: 30-31.2	8: 36.3-37.6	3: -0.08 - -0.03	8: 0.19-0.27	3: 6.2-8	8: 186.9-299.5
4: 0.29-0.35	9: 0.77-0.96	4: 31.3-32.5	9: 37.7-39.7	4: -0.02-0.01	9: 0.28-0.4	4: 8.1-17.7	9: 299.6-390.2
5: 0.36-0.42	10: 0.97-1	5: 32.6-33.8	10: > 39.7	5: 0.02-0.07	10: 0.41-1	5: 17.8-31.6	10: > 390.2
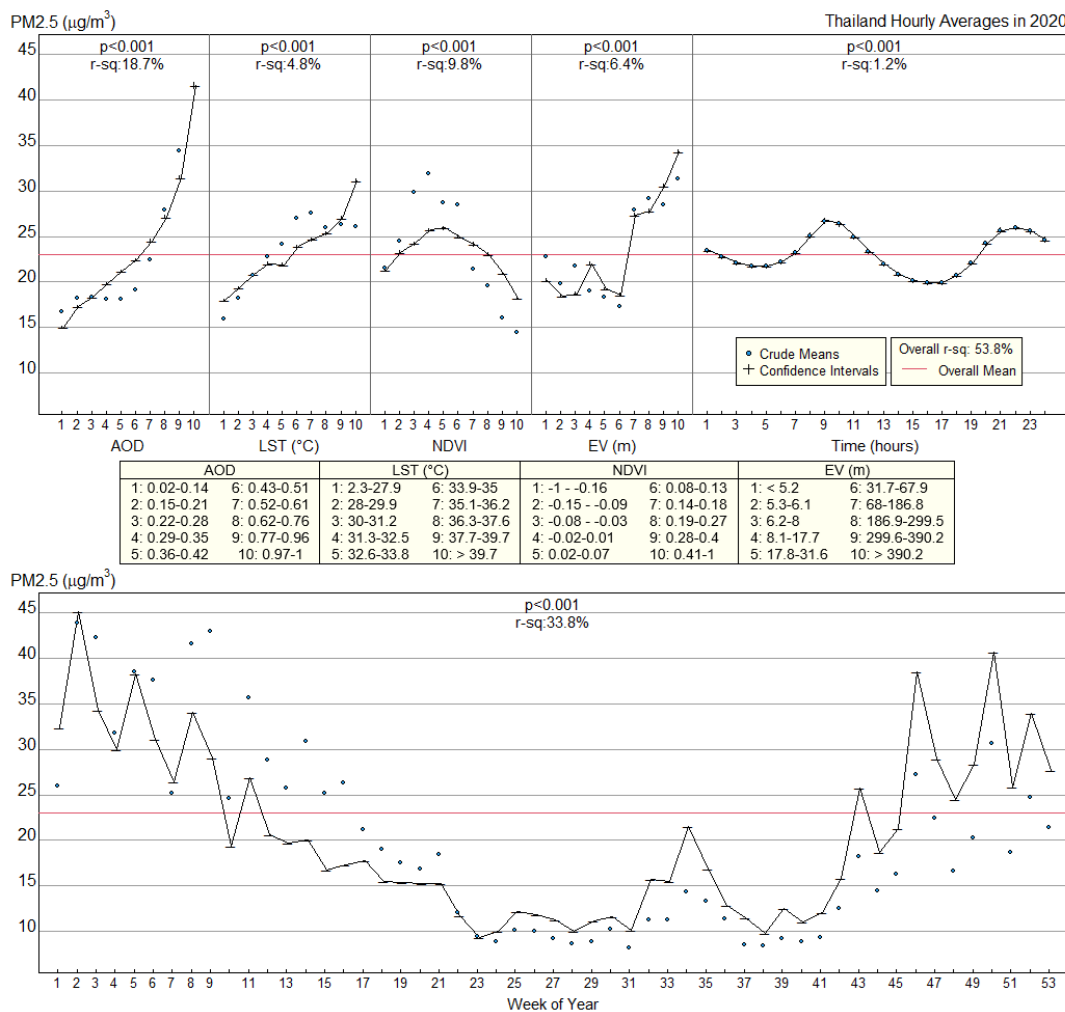
Figure 3.14: A log-linear regression model fitting using weighted sum contrast.

The 95% CI plot demonstrates that it is relatively easy to estimate the hourly concentration of PM2.5 based on each factor. For instance, if we are aware of the values of AOD (0.55), LST (35.5°C), NDVI (0.15), EV (300 m), time (12 am), and WOY (14), we can examine the graph for each factor to determine the corresponding PM2.5 values (24.5, 24.7, 24.2, 30.5, 23.3, and 20, respectively (see actual values in Table 3.3). We can estimate hourly PM2.5 as follows:

where the constant term is -15.69,

$y = log(24.5) + log(24.7) + log(24.2) + log(30.5) + log(23.3) + log(20) – 15.69$

$= 3.465$

$Hourly\ PM2.5 = exp(3.465) = 32\ \mu g/m^3$

Consequently, we can infer that the hourly PM2.5 concentration in this area is around 32 μg/m³. This straightforward approach lets us quickly estimate the hourly PM2.5 concentration by hand and easily interpret the results.

Table 3.3: Weighed sum contrasts log-linear regression model results.

Variables	Crude Means ($\mu g/m^3$)	Adjusted Means ($\mu g/m^3$)	95% CI ($\mu g/m^3$)
AOD			
01: 0.02 – 0.14	16.8	15.0	14.9 – 15.1
02: 0.15 – 0.21	18.3	17.2	17.1 – 17.4
03: 0.22 – 0.28	18.4	18.3	18.2 – 18.4
04: 0.29 – 0.35	18.2	19.8	19.7 – 19.9
05: 0.36 – 0.42	18.1	21.2	21.1 – 21.3
06: 0.43 – 0.51	19.2	22.5	22.3 – 22.6
07: 0.52 – 0.61	22.4	24.5	24.4 – 24.7
08: 0.62 – 0.76	27.9	27.2	27.0 – 27.3
09: 0.77 – 0.96	34.4	31.5	31.3 – 31.7
10: 0.97 – 1	51.8	41.6	41.3 – 41.9
LST (°C)			
01: 2.3 – 27.9	15.9	18.0	17.8 – 18.1
02: 28 – 29.9	18.3	19.4	19.2 – 19.5
03: 30 – 31.2	20.7	20.8	20.7 – 21.0
04: 31.3 – 32.5	22.8	22.0	21.9 – 22.2
05: 32.6 – 33.8	24.1	21.9	21.8 – 22.1
06: 33.9 – 35	27.0	23.9	23.7 – 24.0
07: 35.1 – 36.2	27.6	24.7	24.5 – 24.9
08: 36.3 – 37.6	25.9	25.4	25.2 – 25.6
09: 37.7 – 39.7	26.3	27.0	26.8 – 27.2
10: > 39.7	26.0	31.1	30.8 – 31.3
NDVI			

Variables	Crude Means ($\mu g/m^3$)	Adjusted Means ($\mu g/m^3$)	95% CI ($\mu g/m^3$)
01: -0.33 – -0.16	21.6	21.3	21.2 – 21.5
02: -0.15 – -0.09	24.5	23.3	23.1 – 23.5
03: -0.08 – -0.03	29.9	24.2	24.1 – 24.4
04: -0.02 – 0.01	31.9	25.7	25.5 – 25.9
05: 0.02 – 0.07	28.7	25.9	25.8 – 26.1
06: 0.08 – 0.13	28.5	25.0	24.8 – 25.1
07: 0.14 – 0.18	21.4	24.2	24.0 – 24.3
08: 0.19 – 0.27	19.6	23.1	22.9 – 23.2
09: 0.28 – 0.4	16.1	21.0	20.8 – 21.1
10: 0.42 – 1	14.5	18.2	18.1 – 18.4
EV (m)			
01: 3 – 5.2	22.7	20.2	20.1 – 20.3
02: 5.3 – 6.1	19.9	18.4	18.3 – 18.5
03: 6.2 – 8	21.7	18.7	18.6 – 18.9
04: 8.1 – 17.7	19.0	22.0	21.9 – 22.2
05: 17.8 – 31.6	18.4	19.3	19.1 – 19.4
06: 31.7 – 67.9	17.3	18.6	18.5 – 18.7
07: 68 – 186.8	28.0	27.4	27.2 – 27.5
08: 186.9 – 299.5	29.1	27.8	27.6 – 28.0
09: 299.6 – 390.2	28.5	30.5	30.3 – 30.7
10: > 390.2	31.3	34.3	34.0 – 34.5
Time			
1	23.5	23.5	23.3 – 23.7
2	22.8	22.8	22.5 – 23.0
3	22.1	22.1	21.8 – 22.3
4	21.7	21.7	21.5 – 21.9
5	21.8	21.8	21.6 – 22.0
6	22.2	22.2	22.0 – 22.5
7	23.3	23.3	23.0 – 23.5
8	25.1	25.1	24.9 – 25.4
9	26.7	26.7	26.5 – 27.0
10	26.4	26.4	26.1 – 26.7
11	24.9	24.9	24.7 – 25.2
12	23.3	23.3	23.1 – 23.6
13	22.0	22.0	21.8 – 22.2
14	20.8	20.8	20.6 – 21.0
15	20.2	20.2	20.0 – 20.4
16	19.9	19.9	19.7 – 20.1
17	19.9	19.9	19.7 – 20.1
18	20.7	20.7	20.5 – 20.9
19	22.1	22.1	21.9 – 22.4
20	24.2	24.2	24.0 – 24.5
21	25.7	25.7	25.4 – 25.9
22	26.0	26.0	25.7 – 26.2

Variables	Crude Means ($\mu g/m^3$)	Adjusted Means ($\mu g/m^3$)	95% CI ($\mu g/m^3$)
23	25.6	25.6	25.3 – 25.9
24	24.7	24.7	24.4 – 24.9
WOY			
1	26.0	32.3	31.9 – 32.8
2	43.9	45.1	44.6 – 45.6
3	42.3	34.3	33.9 – 34.7
4	31.8	29.9	29.6 – 30.3
5	38.5	38.3	37.9 – 38.8
6	37.6	31.1	30.7 – 31.4
7	25.1	26.5	26.2 – 26.8
8	41.6	34.1	33.7 – 34.4
9	43.0	29.1	28.7 – 29.5
10	24.6	19.3	19.1 – 19.6
11	35.7	26.9	26.6 – 27.2
12	28.8	20.6	20.3 – 20.9
13	25.8	19.7	19.5 – 20.0
14	30.9	20.0	19.7 – 20.2
15	25.1	16.8	16.5 – 17.0
16	26.3	17.3	17.0 – 17.6
17	21.2	17.7	17.5 – 18.0
18	19.0	15.5	15.2 – 15.8
19	17.6	15.3	15.1 – 15.5
20	16.9	15.3	15.1 – 15.5
21	18.4	15.3	14.8 – 15.8
22	12.0	11.7	11.5 – 11.9
23	9.5	9.4	9.1 – 9.6
24	8.8	10.0	9.8 – 10.2
25	10.1	12.2	11.9 – 12.5
26	10.0	11.9	11.6 – 12.1
27	9.2	11.2	10.9 – 11.6
28	8.6	10.0	9.8 – 10.2
29	8.8	11.1	10.9 – 11.3
30	10.2	11.6	11.4 – 11.9
31	8.2	10.1	9.8 – 10.4
32	11.3	15.7	15.0 – 16.4
33	11.3	15.5	14.7 – 16.4
34	14.3	21.5	20.8 – 22.4
35	13.3	16.9	16.6 – 17.2
36	11.3	12.8	12.5 – 13.1
37	8.5	11.5	11.2 – 11.8
38	8.3	9.8	9.5 – 10.2
39	9.1	12.6	12.3 – 12.8
40	8.8	11.1	10.8 – 11.3
41	9.3	12.0	11.6 – 12.4

Variables	Crude Means ($\mu g/m^3$)	Adjusted Means ($\mu g/m^3$)	95% CI ($\mu g/m^3$)
42	12.5	15.8	15.4 – 16.3
43	18.2	25.7	25.2 – 26.2
44	14.5	18.7	18.3 – 19.1
45	16.3	21.3	20.9 – 21.7
46	27.3	38.5	38.0 – 39.0
47	22.5	29.0	28.6 – 29.3
48	16.6	24.4	24.1 – 24.8
49	20.2	28.4	28.0 – 28.8
50	30.7	40.7	40.3 – 41.2
51	18.6	25.8	25.4 – 26.3
52	24.7	34.0	33.4 – 34.5
53	21.4	27.7	27.2 – 28.2

3.5.4 Estimation Hourly PM2.5 at 1 km Resolution

Figure 3.15 presents the hourly PM2.5 levels recorded by monitoring stations on the left and the estimated levels at a 1 km resolution on the right. The estimation used a log-linear regression model with a weighted sum contrast in WOY 2 (6-12 January) of 2020. Both the hourly PM2.5 levels at monitoring stations and the estimated levels show comparable values. This method enables us to obtain PM2.5 data for every small area since 2000, when MODIS satellite data became available for download.
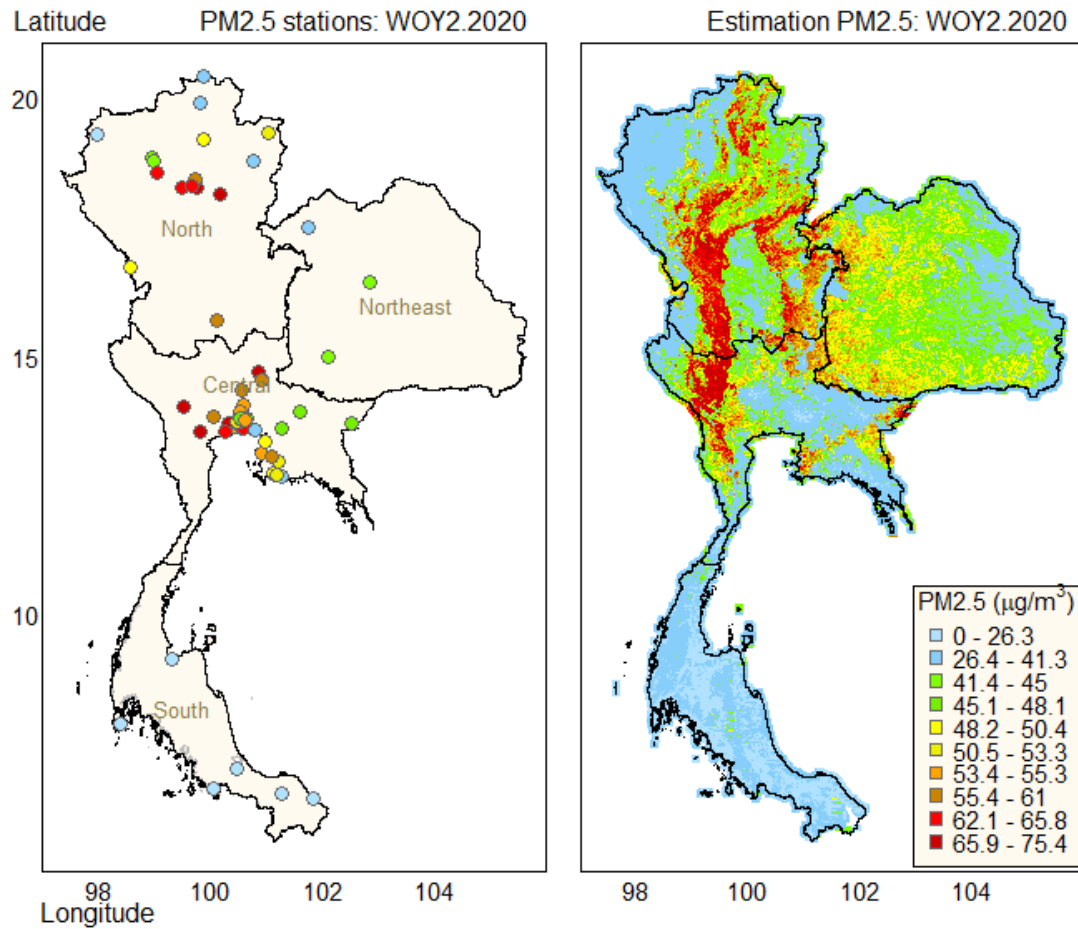
Figure 3.15: PM2.5 at stations and estimated PM2.5 in 1 km resolution.

Figure 3.16 presents the estimation of hourly PM2.5 levels, categorized by region. The R^2 values in the northeast (46.3%), central (44.1%), and north (36.2%) regions are higher than those in the south (11.1%). In the north, the average hourly PM2.5 is 42.7 μg/m^3, a level that may impact health. On the other hand, the northeast and central regions have average hourly PM2.5 levels of 33.3 μg/m^3 and 27.5 μg/m^3, respectively, which represent median range values. Only the south region exhibits an average PM2.5 lower than 15 μg/m^3, indicating good air quality in this area. The standard levels for 24 hours of NAAQS are 37.5 μg/m^3, and WHO recommends 15 μg/m^3.
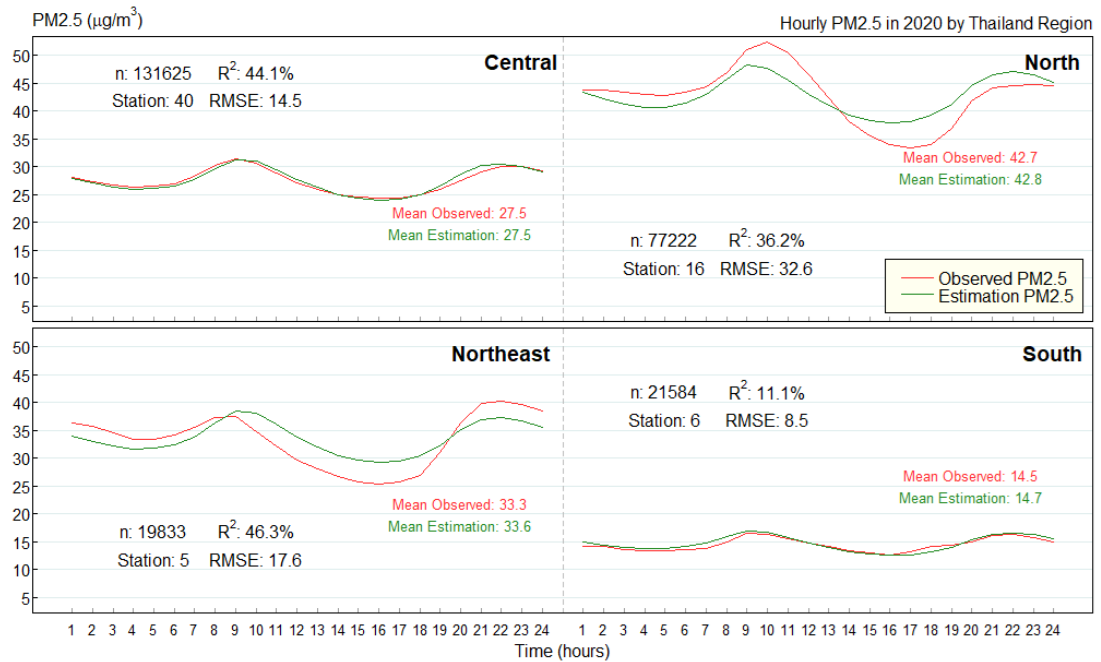
Figure 3.16: Estimation of hourly PM2.5 levels separated by region.

3.6 Interpretation of Results

The log-linear regression model we propose, incorporating weighted sum contrasts, enhances the understanding of the relationship between satellite data and PM2.5, enabling the estimation of hourly PM2.5 concentrations in Thailand with a high spatial resolution of 1 km. This model facilitates the examination of spatiotemporal variations in hourly PM2.5 concentrations at fine scales, offering valuable insights for epidemiological research and empowering individuals to make informed decisions regarding air pollution.

Although we used a linear regression model, assuming that the response variables follow a normal distribution, the Q-Q plot applied to the test revealed that log-transformation was more appropriate for the hourly PM2.5 outcome, given its skewed distribution. Previous studies have also shown that log-transformation can reduce skewness in the distribution of PM2.5, increasing the accuracy of estimation [97], [98], [99], [100]. This finding is consistent with others who similarly found log-transformation to be appropriate for their study on the hospital cost of Chronic-Disease Patients in Southern Thailand, as demonstrated by the Q-Q plot [101]. In statistics, a Q-Q plot is a graphical representation of the differences between observed and expected

values based on the assumption of a normal distribution. A horizontal band close to zero without any discernible pattern indicates that the observed scores are normally distributed [102].

Using the log-linear regression model, we performed an iterative process of adding and removing predictor variables based on their p-values. The variables included satellite variables (AOD, LST, NDVI, and EV), time, and WOY. This process resulted in a final R^2 value of 53.8%. Notably, the WOY and AOD factor variables contributed the most to the model's accuracy improvement. This finding is consistent with previous studies in China, highlighting the significant contribution of AOD to PM2.5 modeling [20], [22]. Moreover, time variables, such as year, month, WOY, and time, explain the seasonal variation in PM2.5 levels and the trend over time [103].

This study has identified various factors associated with elevated levels of PM2.5. Specifically, the analysis has shown that PM2.5 levels were higher than the overall mean and NAAQS hourly PM2.5 levels during particular months, namely January to March and November to December, as indicated by the WOY factor. Additionally, the AOD factor was found to be positively correlated with PM2.5 levels, particularly when exceeding 0.52, while LST was positively associated with PM2.5 levels exceeding 33.9°C. According to research in northern Thailand, the dry season occurring from November to April experiences high levels of ground PM_{10} and PM2.5 concentrations due to AODs [29]. These high levels result from extensive agricultural field burning and open-air biomass burning in the region and neighboring countries. The study also found that NDVI within the range of -0.08 to 0.18 and EV in areas above 67.9 m were associated with higher PM2.5 levels than the overall mean, indicating a relationship with the urbanization area in the central part and higher areas in the northern part of Thailand.

The time variable revealed higher PM2.5 levels during specific hours, specifically 8-11 a.m. and 20-24 p.m., which may be related to working hours in the Bangkok Metropolitan Region (BMR), the capital of central Thailand. In the BMR, automobiles, road dust, biomass burning, and meat cooking are significant sources of PM2.5, contributing to the elevated levels observed in the region [104], [105]. The rapid urbanization of Southeast Asia in recent years has increased the likelihood of air pollution from vehicles, industries, and construction activities [106]. Also, the study in

China described that PM2.5 had greater effects on mortality in urban cities than rural areas [107].

The weighted sum contrast linear regression model helps estimate hourly PM2.5 levels and understand the relationship between determinants and PM2.5. Although some studies indicate that ML models may offer more precise estimates than linear regression [20], [22], the latter is still valuable due to its simplicity and practicality. In contrast, ML models may need more interpretability, require expensive computing resources, and rely on high-resolution satellite images, which can be challenging for researchers with limited budgetary support for analyzing satellite data. Thus, the weighted sum contrast linear regression model remains viable to overcome these limitations. Statistics make population inferences from a sample, while ML identifies generalizable predictive patterns [108].

This method produces hourly PM2.5 estimates at a resolution of 1 km that are similar to those obtained from monitoring stations. By leveraging MODIS satellite data, which has been accessible for download since 2000, this approach allows for the estimation of PM2.5 data in small areas, facilitating the application of the weighted sum contrast linear regression model for hourly PM2.5 estimation. This enables policymakers to acquire data on short-term and long-term PM2.5 fluctuations and their effects on human health and the environment.

3.7 Overall Assessment

This study utilized a weighted sum contrast log-linear regression model to examine the impact of various factors on hourly PM2.5 concentrations in Thailand. The model included six predictor variables: AOD, LST, NDVI, EV, time, and WOY, resulting in an R^2 of 53.8%. The WOY and AOD factors contributed the most to the model's accuracy improvement. The analysis also revealed specific periods and thresholds for each factor correlated with higher hourly PM2.5 levels. The 95% CI plot demonstrated that it is easy to estimate hourly PM2.5 concentrations based on each factor. The map revealed that the spatial distribution of PM2.5 levels was comparable to those observed at the monitoring stations. Overall, the study provides valuable insights into the factors influencing hourly PM2.5 concentrations in Thailand and the potential for utilizing satellite data to estimate hourly PM2.5 levels in small areas.

Chapter 4

PM2.5 with Cardiorespiratory Mortality

This chapter details the preliminary analysis and statistical modeling results of this study. The findings are organized to provide a comprehensive understanding of the spatial distribution of PM2.5 levels, cardiorespiratory mortality rates, and their associations. This includes descriptive statistics, correlation analysis, spatiotemporal distribution patterns, and modeling results, all supporting the study's hypotheses.

4.1 Data Descriptive Analysis

Table 4.1 presents a comprehensive overview of the monthly average PM2.5 levels and cardiorespiratory mortality rates from 2015 to 2019. The analysis reveals that the mean PM2.5 concentration across all years was 27.9 μg/m³ (SD = 13.2), with values ranging between 11.3 and 112 μg/m³. Although this mean falls below the NAAQS threshold of 37.5 μg/m³, it surpasses the more stringent WHO guideline of 15 μg/m³.

The monthly cardiorespiratory mortality rate exhibited a mean of 10.9 per 10,000 population, fluctuating from 1.9 to 25.3 per 10,000 over the study period. The variations in mortality rates suggest possible seasonal trends and regional differences.

Table 4.1: PM2.5 and cardiorespiratory mortality rate by year.

Year	PM2.5 (μg/m³)		N	CM/10^4 population	
	Mean (SD)	Range		Mean (SD)	Range
2015	30.3 (13.7)	14.1-96.7	78,064	10.4 (2.9)	2.1-25.3
2016	29.5 (15.1)	12.8-103.8	87,415	11.6 (3.2)	1.9-24.5
2017	25 (9.7)	11.3-68.2	84,416	10.7 (2.6)	3.9-20.9
2018	26.2 (11.5)	11.4-81.7	83,797	10.4 (2.6)	3.7-22.6
2019	28.4 (14.5)	13.4-112	94,330	11.4 (2.8)	3.3-23.6
Total	27.9 (13.2)	11.3-112	428,022	10.9 (2.9)	1.9-25.3

CM: Cardiorespiratory mortality

4.2 PM2.5 Levels and Cardiorespiratory Mortality Across Regions

Regional differences in PM2.5 concentrations and associated health impacts are crucial for understanding the spatial distribution of pollution-related mortality risks. Table 4.2 presents the comparative PM2.5 levels and cardiorespiratory mortality rates across Thailand's four major regions: North, Northeast, Central, and South.

The central, north, northeast, and south regions exhibited the highest average PM2.5 concentrations, with mean values of 27.2 μg/m³, 29.5 μg/m³, 29.7 μg/m³, and 24.4 μg/m³, respectively. While the monthly PM2.5 levels in all regions were below the 24-hour NAAQS threshold, they exceeded the annual guidelines. These regions of central (12 per 10,000 population) and north (11.7 per 10,000 population) also recorded higher cardiorespiratory mortality rates compared to northeast (8.9 per 10,000 population) and south (10.7 per 10,000 population) regions, likely due to seasonal biomass burning and industrial emissions.

The northern region, particularly provinces near the Myanmar border, experiences significant seasonal air pollution due to agricultural burning and forest fires. This phenomenon coincides with a spike in respiratory illnesses and cardiovascular complications during the dry season (November to April). The northeastern region, while similarly affected by agricultural activities, also faces pollution from cross-border emissions and urban-industrial sources.

The central region, including the Bangkok Metropolitan Area, experiences elevated pollution due to vehicular emissions and industrial activities. However, regulatory measures and improved air quality management have somewhat mitigated its impact on mortality rates. The southern region, characterized by its coastal geography and lower population density, benefits from better air circulation, leading to consistently lower PM2.5 levels and mortality rates.

Table 4.2: PM2.5 and cardiorespiratory mortality by region.

Variables	PM2.5 (µg/m³)		CM/10⁴ population	
	Mean (SD)	Range	Mean (SD)	Range
Central	27.2 (9.9)	12.4-64.9	12 (2.8)	5.4-25.3
North	29.5 (17.8)	12.1-112	11.7 (2.5)	4.2-21.5
Northeast	29.7 (14.9)	13.2-92.4	8.9 (2.2)	1.9-16.6
South	24.4 (7.3)	11.3-44.7	10.7 (2.17)	3.7-22.6

* p-value < 0.05; CM: Cardiorespiratory mortality

4.3 Correlation Analysis of PM2.5 with Cardiorespiratory Mortality

Figure 4.1 presents the relationship between monthly PM2.5 levels and monthly cardiorespiratory mortality per 10,000 population through a month-by-month analysis. The graph highlights trends in monthly PM2.5 levels and monthly cardiorespiratory mortality rates, showing that peak values for both metrics consistently occur from November to April. During this period, average monthly PM2.5 levels rose from 19.3 µg/m³ (observed between May and October) to 36.5 µg/m³, representing an 89% increase. Similarly, monthly cardiorespiratory mortality rates increased by 6% from November to April compared to May to October.
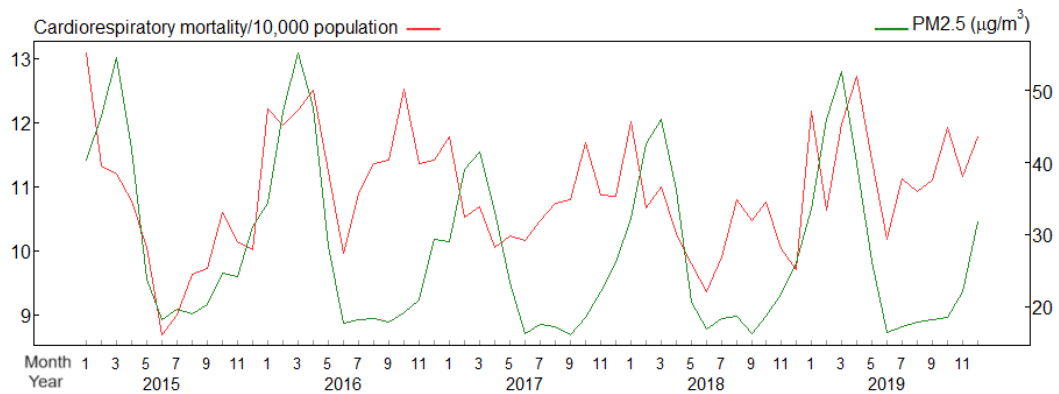
Figure 4.1: Time series plot of PM2.5 with cardiorespiratory mortality.

Figure 4.2 presents Spearman's correlation analysis, which demonstrated a statistically significant positive relationship between monthly PM2.5 concentrations and monthly cardiorespiratory mortality rates, with a correlation coefficient (r) of 0.13 ($p < 0.05$). This indicates that higher monthly PM2.5 levels are associated with

76

increased monthly cardiorespiratory mortality rates. The findings were consistent across all regions included in the analysis, underscoring the potential public health impact of elevated monthly PM2.5 levels during the peak pollution season. The northern (r = 0.35), central (r = 0.13), and northeastern (r = 0.09) regions exhibited a significant (P-value < 0.05) positive correlation between monthly PM2.5 levels and monthly cardiorespiratory mortality. However, no significant association was observed between monthly PM2.5 levels and monthly cardiorespiratory mortality in the southern region. The correlation in the North might provide a basis for further investigation. However, despite statistical significance, the weaker correlations in the other areas suggest limited practical relevance. Additional analysis, such as controlling for confounding variables or using multivariate regression, would be needed to draw more definitive conclusions.
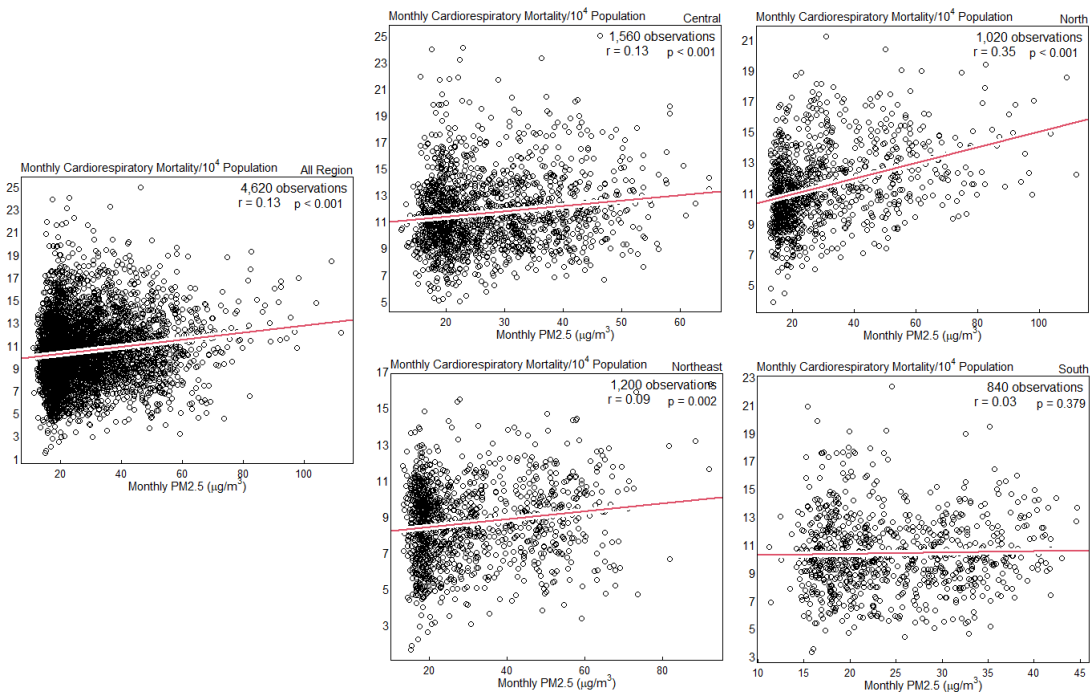
Figure 4.2: Correlation plot of PM2.5 and cardiorespiratory mortality rate.

4.4 Modeling Results

Figure 4.3 presents deviance residual plots against normal quantiles for Poisson regression models fitted with additive effects for monthly PM2.5 groups and Region-monthly PM2.5 groups. The models for monthly cardiorespiratory mortality exhibit a good fit, evidenced by the clustering of standardized residuals along a straight line in

the deviance residual plots. The ratio of deviance to degrees of freedom (Deviance/df < 1) and the alignment of residuals with a straight line further confirm the model's adequacy.
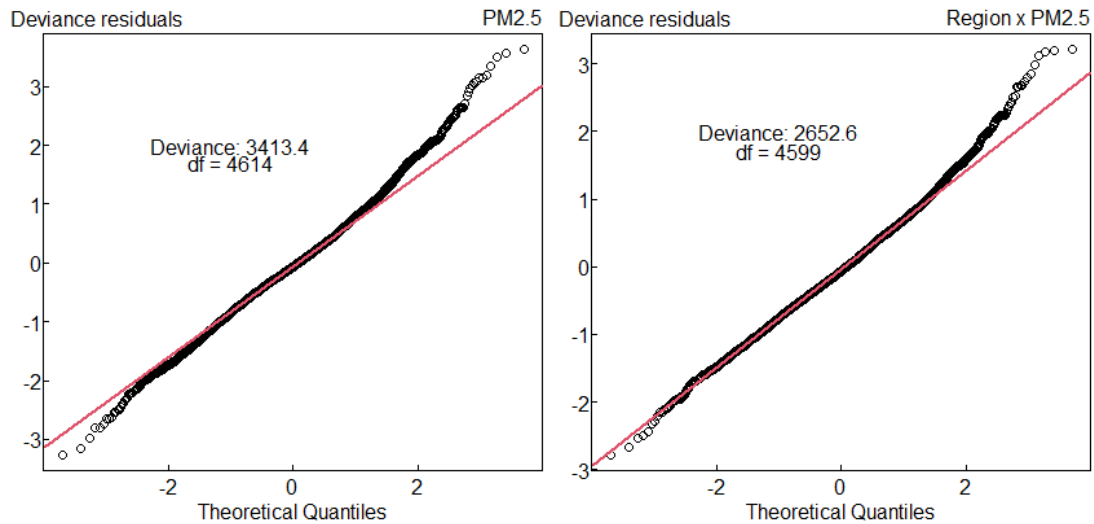
Figure 4.3: Residuals plot of Poisson regression model.

Figure 4.4 presents CIs for monthly cardiorespiratory mortality rates per 10,000 individuals from 2015 to 2019, stratified by monthly PM2.5 concentration groups and regions. A horizontal red line marks the overall average mortality rate of 10.9. In the first panel, blue CIs indicate that regions within the 0–20 µg/m³ PM2.5 group have the lowest monthly mortality rates, while groups with 30.1–37.5 µg/m³ and >37.5 µg/m³ PM2.5 levels exceed the overall mean. This suggests that monthly PM2.5 levels above 30 µg/m³ significantly increase the risk of cardiorespiratory mortality, whereas levels below 20 µg/m³ correspond to a lower risk.

Specifically, mortality rates increase by 3% (95% CI: 1%–5%) in areas with 30.1–37.5 µg/m³ PM2.5 levels and by 5% (95% CI: 3%–7%) in areas exceeding 37.5 µg/m³. Conversely, a 4% (95% CI: 3%–6%) decrease is observed in areas with PM2.5 levels below 20 µg/m³.

The subsequent panels highlight regional variations. The central region consistently shows elevated mortality rates across all PM2.5 level categories. In the northern region, rates increase significantly when PM2.5 levels exceed 25 µg/m³. The northeast region is an exception, with certain PM2.5 groups reporting mortality rates

below the overall mean. In the southern region, rates exceed the overall mean only when PM2.5 levels surpass 37.5 µg/m³.
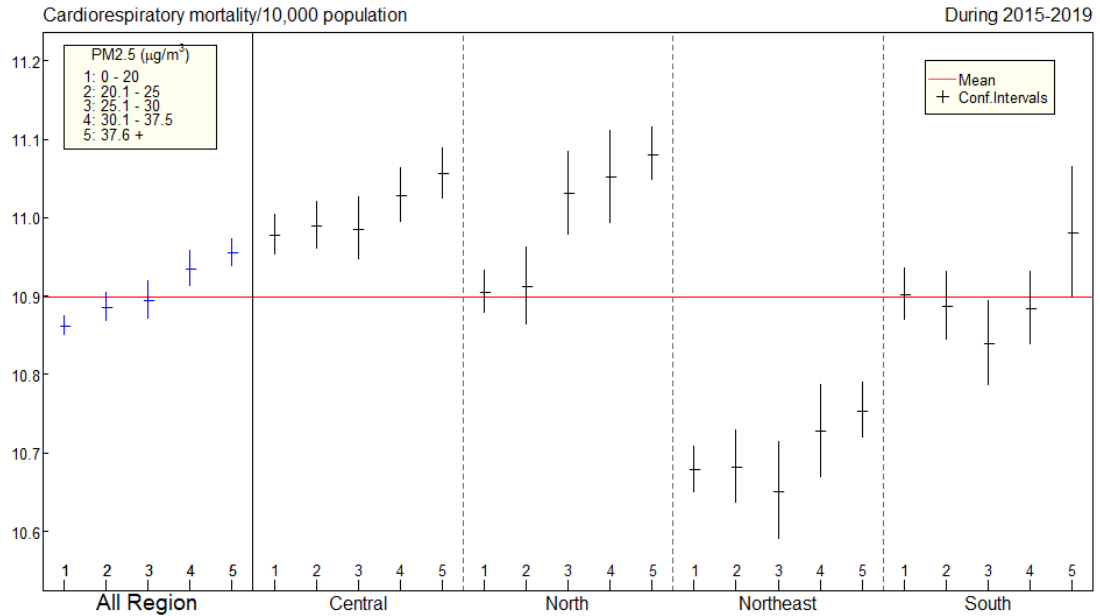
Figure 4.4: Confidence intervals for estimated cardiorespiratory mortality.

4.5 Spatial Distributions

Figure 4.5 presents an analysis of the geographic distribution of monthly PM2.5 levels and cardiorespiratory mortality rates across provinces in Thailand. The central and northern regions consistently show the highest PM2.5 levels and mortality rates, while the southern and northeastern regions report the lowest. Statistical analysis identifies a significant clustering pattern, as indicated by a Moran's I value of 0.334 (p < 0.05), reflecting notable spatial autocorrelation in PM2.5 levels and mortality rates across the country.
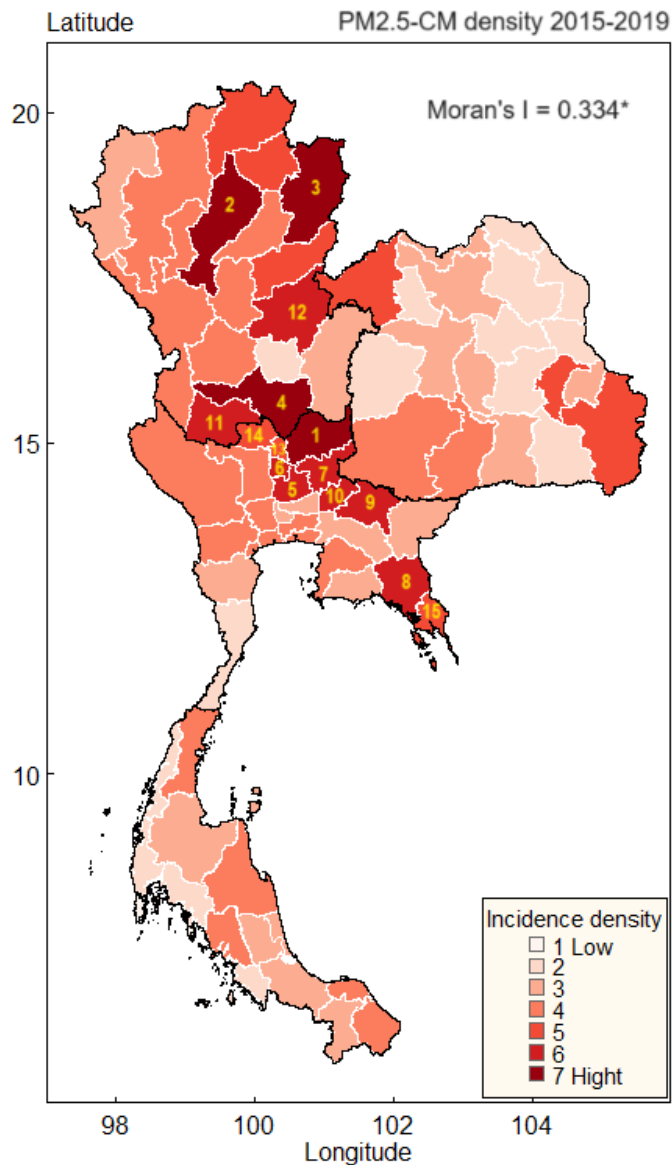
Figure 4.5: Spatial distributions of PM2.5 and cardiorespiratory mortality.

The local Moran's I analysis further corroborates these findings, revealing that the top 15 provinces with the highest rates form distinct clusters, predominantly in the central and northern regions. Table 4.3 presents the local Moran's I values for each province, providing more granular insights. These findings emphasize the spatial consistency of this public health concern and underscore the importance of targeted interventions in high-risk areas.

Table 4.3: Spatial distributions of PM2.5 and cardiorespiratory mortality.

No.	Province	Region	PM2.5	CM	I_i
1	Lop Buri	Central	32.5	13.1	1.604*
2	Lampang	North	30.5	13.1	0.709
3	Nan	North	32.2	14.1	1.288
4	Nakhon Sawan	North	32.8	13.0	1.035
5	Phra Nakhon Si Ayutthaya	Central	26.8	13.7	0.932
6	Ang Thong	Central	27.3	12.5	1.75*
7	Saraburi	Central	29.7	16.2	1.438*
8	Chanthaburi	Central	28.1	12.8	-0.227
9	Prachin Buri	Central	29.0	12.7	0.189
10	Nakhon Nayok	Central	28.8	16.5	0.605
11	Uthai Thani	North	29.9	13.8	1.022
12	Phitsanulok	North	27.6	13.1	0.016
13	Sing Buri	Central	24.2	12.8	1.117*
14	Chai Nat	Central	24.5	13.5	0.948*
15	Trat	Central	24.5	12.8	1.23
16	Ubon Ratchathani	Northeast	30.1	10.4	-0.273
17	Yasothon	Northeast	30.9	10.4	-0.574
18	Loei	Northeast	30.5	11.6	-0.273
19	Uttaradit	North	30.6	11.5	0.807*
20	Phayao	North	33.0	12.1	1.089*
21	Chiang Rai	North	31.3	11.5	0.854
22	Bangkok Metropolis	Central	28.9	10.6	-0.028
23	Samut Prakan	Central	28.2	10.6	-0.028
24	Chon Buri	Central	26.7	11.6	0.016
25	Nakhon Ratchasima	Northeast	28.5	11.1	0.07
26	Buri Ram	Northeast	30.4	9.4	-0.054
27	Chiang Mai	North	29.1	11.2	0.068
28	Lamphun	North	29.2	10.9	0.102
29	Phrae	North	28.3	12.0	0.155*
30	Kamphaeng Phet	North	29.3	10.8	0.068
31	Tak	North	28.5	10.8	0.083*
32	Sukhothai	North	29.9	10.5	0.102*
33	Ratchaburi	Central	27.7	11.9	-0.002
34	Kanchanaburi	Central	26.8	12.2	0.05
35	Suphan Buri	Central	27.0	12.0	0.115*
36	Nakhon Pathom	Central	26.3	10.7	0.028
37	Samut Sakhon	Central	28.3	10.8	0.016
38	Samut Songkhram	Central	26.2	12.4	-0.013
39	Nakhon Si Thammarat	South	25.2	10.4	-0.071
40	Chumphon	South	26.1	12.3	-0.129
41	Trang	South	26.9	12.3	-0.093
42	Pattani	South	19.7	14.2	-0.042

No.	Province	Region	PM2.5	CM	I_i
43	Narathiwat	South	25.1	12.0	-0.028
44	Nonthaburi	Central	29.3	9.4	-0.168
45	Pathum Thani	Central	27.9	10.1	-0.331
46	Rayong	Central	24.6	11.3	-0.461
47	Chachoengsao	Central	24.2	12.3	-0.266
48	Sa Kaeo	Central	29.1	10.1	-0.305
49	Surin	Northeast	30.1	8.2	0.417
50	Amnat Charoen	Northeast	31.1	9.4	-0.071
51	Khon Kaen	Northeast	29.0	10.1	0.276
52	Udon Thani	Northeast	30.6	8.8	0.384
53	Nong Khai	Northeast	30.1	5.9	0.319
54	Maha Sarakham	Northeast	30.2	9.2	0.397
55	Mae Hong Son	North	27.6	10.3	-0.071
56	Phetchabun	North	28.2	9.9	-0.294
57	Phetchaburi	Central	26.0	10.2	0.189
58	Surat Thani	South	24.9	11.0	0.397
59	Songkhla	South	26.8	9.8	0.241
60	Phatthalung	South	26.5	10.2	0.222
61	Yala	South	23.1	10.7	0.059
62	Si Sa Ket	Northeast	29.3	9.1	0.059
63	Chaiyaphum	Northeast	29.2	9.1	-0.374
64	Bueng Kan	Northeast	29.4	7.0	1.288
65	Nong Bua Lam Phu	Northeast	30.0	8.9	0.131
66	Roi Et	Northeast	29.2	9.4	0.854
67	Kalasin	Northeast	29.7	8.6	1.143*
68	Sakon Nakhon	Northeast	28.8	7.0	1.288*
69	Nakhon Phanom	Northeast	29.3	6.6	1.577*
70	Mukdahan	Northeast	28.2	7.5	0.998*
71	Phichit	North	24.1	10.3	-1.025
72	Prachuap Khiri Khan	Central	24.7	9.9	0.276
73	Krabi	South	22.6	9.9	0.493
74	Phangnga	South	24.0	9.6	1.288
75	Ranong	South	23.2	9.9	0.709
76	Satun	South	26.0	9.3	0.42
77	Phuket	South	22.0	8.6	0

*p-value < 0.05; CM: Cardiorespiratory mortality

4.6 Temporal Distributions

Figure 4.6 presents the findings of DLNM combined with GAM for determining the lag–response relationship between PM2.5 and cardiorespiratory mortality, as well as the RR. It depicts the various patterns of mortality as a consequence of a 10 % rise in PM2.5 concentration.

The number of mortality rates on the 2-month to 5-month lags of cardiorespiratory mortality was significantly associated with a 10 % increase in PM2.5 (the mean of 27.9 $\mu g/m^3$). It exhibited the highest increase in cardiorespiratory mortality of 6.4 % (RR = 1.064, 95 % CI = 1.049–1.080).
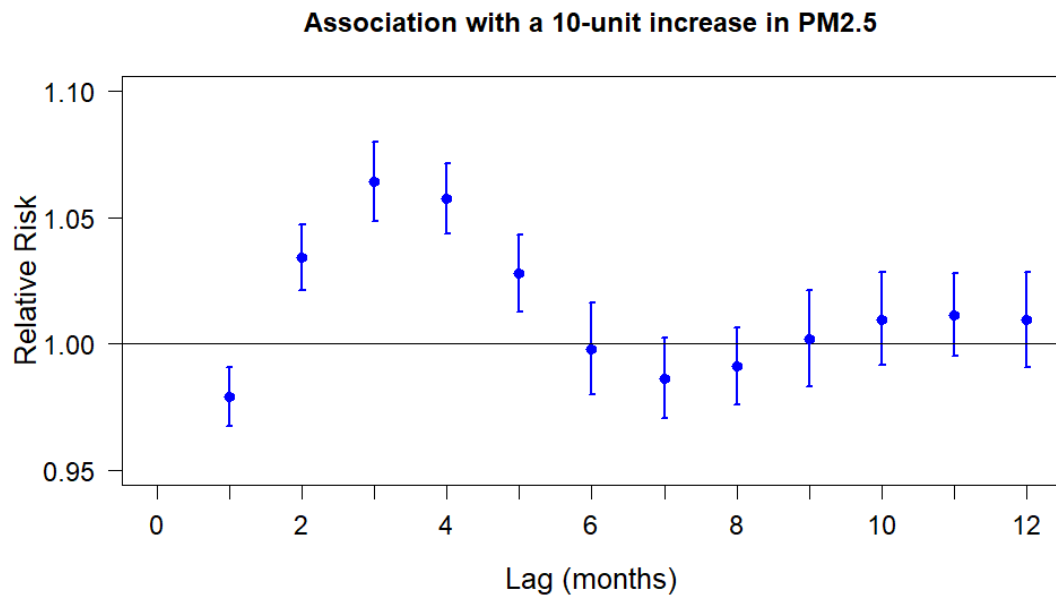
Figure 4.6: Lag–response relationship between PM2.5 and the relative risk for cardiorespiratory mortality.

Chapter 5
Summary

This study identified a significant spatiotemporal association between monthly PM2.5 concentrations and cardiorespiratory mortality rates across Thailand from January 1, 2015, to December 31, 2019. High-resolution satellite-derived PM2.5 concentration estimates at a 1 km spatial resolution were employed to capture the distribution of air pollution exposure. Cardiorespiratory mortality rates, categorized by ICD-10 codes and standardized per 10,000 population, were obtained from the NHSO. The results highlight the necessity for focused strategies to mitigate health risks related to air pollution, especially in areas with persistently elevated PM2.5 levels. This comprehensive approach underlines the value of integrating satellite data with health statistics to reveal nationwide pollution patterns and their health consequences.

5.1 Estimation PM2.5 used Satellite Data

- Variables

The analysis of changes in situ PM2.5 and AOD revealed similar distribution patterns. Most of the data is available during the dry season from November to March. This is likely due to the high LST during the dry season, which causes uneven heating, increased chemical reactions, and increased turbulence among particles, leading to the formation of secondary aerosols in the atmosphere [109]. The NDVI and EV values were found to be the lowest, likely because most monitoring stations were located in low-lying urban areas.

The results of our PM2.5 assessment indicate that Northern Thailand experiences higher levels of PM2.5 than other regions, particularly during the dry seasons of WOY 1–10 (January–March) and WOY 45–53 (November–December). This is attributed to extensive agricultural field burning and open-air biomass burning in Northern Thailand and neighboring countries, as per Phuengsamran & Lalitaporn (2021) [29]. These activities contribute to elevated PM2.5 levels and have a significant impact on climate change. The PM2.5 levels in most areas of Thailand exceed the annual NAAQS limit of 25 $\mu g/m^3$ and the WHO annual standard of 10 $\mu g/m^3$, except

for the Southern region. High PM2.5 levels can negatively impact the health of the population, including respiratory and cardiovascular diseases. Therefore, our developing models of PM2.5 data could be used to identify links between PM2.5 levels and specific geographic areas, such as provinces, districts, and sub-districts.

The PM2.5 at sites was fitted using satellite data within a 5 km radius. We matched the samples for every day in Thailand from 2011 to 2020, and any missing values were excluded. The training and validation models were built using 80% (27,798 rows) and 20% (6,950 rows) of the PM2.5 data from the PCD, respectively. The testing model used 7,339 rows of PM2.5 data from the BAQ. Daily PM2.5 concentrations often exhibit a favorable skewed distribution similar to AOD. Similar to the research conducted by Wei et al. (2019) in China, the bivariate correlation analysis revealed that independent variables such as AOD have a strong association with PM2.5 [20].

Per Spearman's correlation test results, AOD, LST, and EV were positively associated with PM2.5, whereas WOY, NDVI, and year have an inverse relationship with PM2.5. The stepwise regression model with p-values < 0.001 and an R^2 accuracy of 49.4% showed that all possible predictor variables should be included in the model.

- Modeling

In our trials, RF outperformed MLR, XGBoost, and SVM models. Our findings align with previous PM2.5 estimating studies from other countries, with an R^2 of 0.95 (RMSE of 5.58 μg/m^3) for training data, 0.78 (RMSE of 11.18 μg/m^3) for validation data, and 0.71 (RMSE of 8.79 μg/m^3) for testing data. For example, the predicted PM2.5 in Greater London using RF, Gradient Boosting Machine (GBM), and K-Nearest Neighbor (KNN), with RF providing the best estimation with an R^2 of 0.83 and RMSE of 4.28 μg/m^3 [110]. In another study, using remote sensing data and AOD, eight approaches were used to anticipate monthly PM2.5 in British Columbia, and RF was found to be the most reliable ML method, with an R^2 of 0.49 (RMSE of 2.67 μg/m^3) [22]. The predicted daily PM2.5 at a 1 km grid for 2013–2015 in Italy using RF with an R^2 of 0.80 (RMSE = 7.05 μg/m^3) [111]. The computed 1 km-resolution PM2.5 concentrations in China using RF, with an R^2 of 0.98 (RMSE = 6.40 μg/m^3) for model fitting and an R^2 of 0.81 (RMSE = 17.91 μg/m^3) for model validation [112]. Another Chinese study used RF to predict daily PM2.5 from 2005 to 2016, with an R^2 of 0.77

(RMSE of 22 μg/m^3) [113]. These studies demonstrate that estimating PM2.5 from satellite data using the RF model with an R^2 of 0.49–0.83 (RMSE = 2.67–22 μg/m^3) in the validation data is acceptable. On the other hand, the MLR model performed poorly in this study. This may be due to the positively skewed and non-normally distributed nature of PM2.5 data, which may not be well suited for MLR models [114], [115], [116].

The study found that the RF model, utilizing AOD, LST, NDVI, EV, WOY, and year as predictors, produced the best results for estimating daily PM2.5 concentrations in Thailand. The strength of the RF model lies in its ability to avoid overfitting data by utilizing the strength of individual trees in the forest and their correlation. However, the results of our study differ from those of other studies, where other models, such as XGBoost, have been found to outperform RF [113]. This may be due to how these decision tree-based models take in and process training data. Our findings suggest that decision tree-based models are recommended for estimating PM2.5 using satellite data. The results indicate that WOY, AOD, and EV are significant factors in determining PM2.5 concentrations, as shown by the two measurements of the RF model. This is consistent with previous studies, which found AOD and EV to contribute significantly to PM2.5 modeling [22].

Our findings indicate a strong alignment between estimated PM2.5 concentrations and observed values at monitoring stations, consistent patterns evident in the time-series plots for both observed and estimated PM2.5. However, a notable discrepancy emerged between observed and estimated PM2.5 concentrations in the testing dataset during 2015–2017. This inconsistency may be attributed to the relatively homogeneous geographical distribution of pollutants in PM2.5 samples collected before 2018, as proposed by research conducted in the United Kingdom [111].

Upon regional analysis, the estimation of PM2.5 revealed higher R^2 values in the northern region and the lowest values in the southern region. This observation may be linked to satellite data, as remote sensing techniques tend to yield clearer data in areas with elevated PM2.5 levels compared to regions with lower concentrations. Notably, the northern region exhibits higher PM2.5 levels than other regions, particularly during dry seasons. This aligns with a study conducted in Taiwan, which found higher PM2.5 levels during winter associated with elevated R^2 values [118].

Moreover, the northern region demonstrates stable PM2.5 trends between 2011 and 2020, while other regions exhibit slight decreases. This stability in the northern region may contribute significantly to the model's overall accuracy.

5.2 Satellite-Based PM2.5 and Health Analysis

Satellite-based PM2.5 monitoring has become a pivotal method in environmental health research, offering spatially comprehensive and high-resolution data for exposure assessment. This approach overcomes the limitations of sparse ground monitoring networks, especially in developing countries like Thailand. A growing body of international studies demonstrates the integration of satellite-derived PM2.5 data with health outcome analyses to better understand pollution-related morbidity and mortality.

In the United States, for example, Di et al. (2017) utilized satellite-based estimates of daily PM2.5 at a 1 km resolution to examine the link between long-term exposure and all-cause mortality among Medicare beneficiaries. Their findings revealed a significant increase in mortality risk associated with PM2.5 exposure levels even below national standards [119].

In China, Liao et al. (2020) evaluated county-level PM2.5 concentrations using satellite remote sensing combined with land use regression and ML models. Their study associated high PM2.5 levels with increased lung cancer mortality and substantial years of life lost, highlighting the urgent need for pollution control in rapidly urbanizing regions [120].

Similarly, a study in India by Mandal et al. (2023) used satellite-derived PM2.5 to assess its association with the incidence of type 2 diabetes. The analysis revealed that long-term exposure significantly elevated the risk of developing diabetes, particularly in densely populated urban areas [121].

In Iran, Faridi et al. (2022) applied a satellite-based model to estimate PM2.5 concentrations and examined their association with all-cause and cause-specific mortality. The study found notable spatial disparities, with industrial and urban centers experiencing the highest health burdens [122].

In the context of Thailand, while numerous studies have explored PM2.5 exposure and its health impacts using ground-based data, research integrating satellite-

derived PM2.5 estimates remains limited. Existing studies have demonstrated significant associations between PM2.5 and increased risks of cardiovascular, respiratory, and chronic kidney diseases [32], [33], [34], [35], [36], [38]. The current study aims to build on this foundation by using high-resolution satellite PM2.5 data in conjunction with national health statistics to assess cardiorespiratory mortality patterns. This approach will provide a more comprehensive understanding of spatial exposure variations and their implications for public health policy.

5.3 Spatiotemporal Association of PM2.5 and Cardiorespiratory Mortality

This study observed a distinct seasonal trend, with the highest PM2.5 levels and mortality rates recorded from November to April, consistent with Thailand's dry season. During this period, elevated PM10, PM2.5, and AOD were prominent, particularly in Northern Thailand, where pollutant levels often exceeded safety thresholds [29]. Pothirat et al. (2016) associated these seasonal increases with higher emergency visits for asthma and COPD [123], while Buya et al. (2024) reported significant impacts on hospital admissions in Southern Thailand [33].

Poisson regression analysis confirmed a strong association between elevated PM2.5 levels and increased cardiorespiratory mortality, particularly in central and northern regions, consistent with previous research [124], [125]. Chronic and acute PM2.5 exposure has been linked to higher risks of respiratory, cardiovascular, and all-cause mortality [126], [127]. The central region consistently showed high mortality rates across all PM2.5 levels, with Bangkok's emissions primarily originating from vehicle exhaust, road dust, biomass burning, and cooking, which remain stable year-round [105]. In Northern Thailand, mortality rates increased sharply when PM2.5 levels exceeded 25 μg/m³, driven by illegal forest encroachment and agricultural burning during the dry season [128], [129].

The northeast and southern regions displayed distinct patterns [130]. In the northeast, small-scale agricultural burning and limited industrialization contributed to lower PM2.5 concentrations compared to northern forest fires. In the south, coastal and tropical climates facilitated pollutant dispersion, while fishing and tourism activities minimally contributed to PM2.5 emissions. Regional differences in healthcare access, population density, and demographics further explained these variations.

Findings suggest that PM2.5 levels above 30 μg/m³ are associated with higher cardiorespiratory mortality risks, while levels below 20 μg/m³ pose lower risks. Current PM2.5 guidelines include 24-hour and annual limits, recently revised by the NAAQS (June 1, 2023) to 37.5 μg/m³ and 15 μg/m³, respectively, which remain higher than WHO guidelines (15 μg/m³ and 5 μg/m³) [131]. The study's insights highlight the need for stricter air quality standards to protect public health.

5.4 Strengths and Limitations

One of the key strengths of this study is its use of high-resolution satellite data, which enabled a comprehensive analysis of PM2.5 exposure across different regions in Thailand. Unlike previous studies relying on ground-based monitoring stations with limited spatial coverage, this approach provided a more accurate and extensive dataset, ensuring greater reliability of findings.

Additionally, this study applied robust statistical methods, including Poisson regression modeling and spatial autocorrelation analysis (Moran's I), to assess the relationship between PM2.5 and mortality rates. The incorporation of seasonal trends and regional disparities further strengthened the analysis, allowing for targeted policy recommendations.

However, certain limitations must be acknowledged. The study did not account for other potential confounding variables such as socioeconomic status, healthcare access, pre-existing health conditions, and occupational exposure. Additionally, while satellite data provided high spatial resolution, it may still contain inherent uncertainties related to cloud cover and aerosol retrieval techniques. Future studies should incorporate demographic variables and individual health records to refine risk assessment models.

5.5 Conclusions

This study provides compelling evidence of a direct correlation between monthly PM2.5 exposure and increased mortality from cardiorespiratory diseases in Thailand between 2015 and 2019. The seasonal variations observed, with peak pollution levels and mortality rates occurring from November to April, underscore the urgency of implementing targeted interventions, particularly during the dry season. By

utilizing high-resolution satellite data and Poisson regression modeling, significant regional disparities were identified, with the central and northern regions experiencing the highest mortality burden.

The marked spatial clustering of elevated PM2.5 concentrations and mortality rates highlights air pollution as a critical public health concern. The analysis indicates that monthly PM2.5 levels exceeding 30 μg/m³ are associated with a 1%-7% increase in cardiorespiratory mortality risk compared to the overall mean. In contrast, levels below 20 μg/m³ are linked to a 3%-6% reduction in risk compared to the overall mean. These findings emphasize the importance of ongoing research and the implementation of robust policy measures to mitigate air pollution and safeguard vulnerable populations across Thailand.

Policy recommendations include:

- **Strengthening Air Quality Standards:** Aligning national PM2.5 limits with WHO guidelines to minimize health risks.
- **Enhancing Monitoring Infrastructure:** Expanding ground-based air quality monitoring stations and integrating them with satellite-based observations for improved data accuracy.
- **Implementing Seasonal Interventions:** Restricting biomass burning during peak pollution months and enforcing stricter emissions regulations.
- **Public Awareness Campaigns:** Educating communities on pollution-related health risks and promoting protective measures such as air filtration and indoor air quality improvements.
- **Urban Planning and Green Initiatives:** Encouraging green infrastructure, reducing vehicular emissions, and promoting clean energy alternatives.

References

[1]   WHO, "Ambient (outdoor) air pollution," *World Health Organization Fact Sheet*, pp. 5–8, 2019, Accessed: May 02, 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health

[2]   J. Rentschler and N. Leonova, "Global air pollution exposure and poverty," *Nat Commun*, vol. 14, no. 1, p. 4432, 2023.

[3]   D. W. Dockery *et al.*, "An Association between Air Pollution and Mortality in Six U.S. Cities," *New England Journal of Medicine*, vol. 329, no. 24, pp. 1753–1759, Dec. 1993, doi: 10.1056/NEJM199312093292401.

[4]   G. Hoek *et al.*, "Long-term air pollution exposure and cardio- respiratory mortality: a review," *Environmental Health*, vol. 12, no. 1, p. 43, Dec. 2013, doi: 10.1186/1476-069X-12-43.

[5]   J. Lepeule, F. Laden, D. Dockery, and J. Schwartz, "Chronic Exposure to Fine Particles and Mortality: An Extended Follow-up of the Harvard Six Cities Study from 1974 to 2009," *Environ Health Perspect*, vol. 120, no. 7, pp. 965–970, Jul. 2012, doi: 10.1289/ehp.1104660.

[6]   D. Krewski *et al.*, "Mortality and Long-Term Exposure to Ambient Air Pollution: Ongoing Analyses Based on the American Cancer Society Cohort," *J Toxicol Environ Health A*, vol. 68, no. 13–14, pp. 1093–1109, Jul. 2005, doi: 10.1080/15287390590935941.

[7]   US EPA, "Integrated Science Assessment (ISA) for Particulate Matter (Final Report, 2019).," *U.S. Environmental Protection Agency*, no. December 2019, 2019.

[8]   B. Brunekreef *et al.*, "Effects of long-term exposure to traffic-related air pollution on respiratory and cardiovascular mortality in the Netherlands: the NLCS-AIR study.," *Res Rep Health Eff Inst*, no. 139, 2009.

[9]   L. Chen *et al.*, "Assessment of population exposure to PM2.5 for mortality in China and its public health benefit based on BenMAP," *Environmental Pollution*, vol. 221, pp. 311–317, Feb. 2017, doi: 10.1016/j.envpol.2016.11.080.

[10] V. Limaye, W. Schöpp, and M. Amann, "Applying Integrated Exposure-Response Functions to PM2.5 Pollution in India," *Int J Environ Res Public Health*, vol. 16, no. 1, p. 60, Dec. 2018, doi: 10.3390/ijerph16010060.

[11] Z. Qu, X. Wang, F. Li, Y. Li, X. Chen, and M. Chen, "PM2.5-Related Health Economic Benefits Evaluation Based on Air Improvement Action Plan in Wuhan City, Middle China," *Int J Environ Res Public Health*, vol. 17, no. 2, p. 620, Jan. 2020, doi: 10.3390/ijerph17020620.

[12] L. A. Cifuentes, J. Vega, K. Köpfer, and L. B. Lave, "Effect of the fine fraction of particulate matter versus the coarse mass and other pollutants on daily mortality in Santiago, Chile," *J Air Waste Manage Assoc*, vol. 50, no. 8, pp. 1287–1298, 2000, doi: 10.1080/10473289.2000.10464167.

[13] J. Schwartz, D. W. Dockery, and L. M. Neas, "Is Daily Mortality Associated Specifically with Fine Particles?," *J Air Waste Manage Assoc*, vol. 46, no. 10, 1996, doi: 10.1080/10473289.1996.10467528.

[14] R. T. Burnett *et al.*, "ASSOCIATION BETWEEN PARTICULATE- AND GAS-PHASE COMPONENTS OF URBAN AIR POLLUTION AND DAILY MORTALITY IN EIGHT CANADIAN CITIES," *Inhal Toxicol*, vol. 12, no. sup4, pp. 15–39, Jan. 2000, doi: 10.1080/08958370050164851.

[15] J. T. Pryor, L. O. Cowley, and S. E. Simonds, "The Physiological Effects of Air Pollution: Particulate Matter, Physiology and Disease," *Front Public Health*, vol. 10, Jul. 2022, doi: 10.3389/fpubh.2022.882569.

[16] Y. Chu *et al.*, "A review on predicting ground PM$_{2.5}$ concentration using satellite aerosol optical depth," *Atmosphere (Basel)*, vol. 7, no. 10, 2016, doi: 10.3390/atmos7100129.

[17] J. A. Engel-Cox, C. H. Holloman, B. W. Coutant, and R. M. Hoff, "Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality," *Atmos Environ*, vol. 38, no. 16, 2004, doi: 10.1016/j.atmosenv.2004.01.039.

[18] X. Zhang, Y. Chu, Y. Wang, and K. Zhang, "Predicting daily PM$_{2.5}$ concentrations in Texas using high-resolution satellite aerosol optical depth," *Science of the Total Environment*, vol. 631–632, pp. 904–911, 2018, doi: 10.1016/j.scitotenv.2018.02.255.

[19]   M. Z. Joharestani, C. Cao, X. Ni, B. Bashir, and S. Talebiesfandarani, "PM2.5 prediction based on random forest, XGBoost, and deep learning using multisource remote sensing data," *Atmosphere (Basel)*, vol. 10, no. 7, 2019, doi: 10.3390/atmos10070373.

[20]   J. Wei *et al.*, "Estimating 1-km-resolution PM2.5 concentrations across China using the space-time random forest approach," *Remote Sens Environ*, vol. 231, p. 111221, 2019, doi: https://doi.org/10.1016/j.rse.2019.111221.

[21]   Q. Xiao, H. H. Chang, G. Geng, and Y. Liu, "An Ensemble Machine-Learning Model to Predict Historical PM<inf>2.5</inf> Concentrations in China from Satellite Data," *Environ Sci Technol*, vol. 52, no. 22, pp. 13260–13269, 2018, doi: 10.1021/acs.est.8b02917.

[22]   Y. Xu *et al.*, "Evaluation of machine learning techniques with multiple remote sensing datasets in estimating monthly concentrations of ground-level PM2.5," *Environmental Pollution*, vol. 242, 2018, doi: 10.1016/j.envpol.2018.08.029.

[23]   M. Danesh Yazdi *et al.*, "Predicting fine particulate matter (PM2. 5) in the greater london area: An ensemble approach using machine learning methods," *Remote Sens (Basel)*, vol. 12, no. 6, p. 914, 2020.

[24]   T. Kanabkaew, "Prediction of hourly particulate matter concentrations in Chiangmai, Thailand using MODIS aerosol optical depth and ground-based meteorological data," *EnvironmentAsia*, vol. 6, no. 2, 2013.

[25]   W. Kumharn *et al.*, "Improved Hourly and long-term PM2.5 Prediction Modeling Based on MODIS in Bangkok," *Remote Sens Appl*, vol. 28, Nov. 2022, doi: 10.1016/j.rsase.2022.100864.

[26]   T. Amnuaylojaroen, "Prediction of PM2.5in an Urban Area of Northern Thailand Using Multivariate Linear Regression Model," *Advances in Meteorology*, vol. 2022, 2022, doi: 10.1155/2022/3190484.

[27]   Y. Chu *et al.*, "A review on predicting ground PM<inf>2.5</inf> concentration using satellite aerosol optical depth," *Atmosphere (Basel)*, vol. 7, no. 10, 2016, doi: 10.3390/atmos7100129.

[28]   I. Kloog, P. Koutrakis, B. A. Coull, H. J. Lee, and J. Schwartz, "Assessing temporally and spatially resolved PM<inf>2.5</inf> exposures for epidemiological studies using satellite aerosol optical depth measurements,"

Atmos Environ, vol. 45, no. 35, pp. 6267–6275, 2011, doi: 10.1016/j.atmosenv.2011.08.066.

[29] P. Phuengsamran and P. Lalitaporn, "Estimating Particulate Matter Concentrations in Central Thailand Using Satellite Data," *Thai Environmental Engineering Journal*, vol. 35, no. 3, pp. 1–11, 2021.

[30] S. Pongpiachan *et al.*, "Applying synchrotron radiation-based attenuated total reflection-Fourier transform infrared to evaluate the effects of shipping emissions on fluctuations of PM10-bound organic functional groups and ionic species," *Atmos Pollut Res*, vol. 13, no. 9, Sep. 2022, doi: 10.1016/j.apr.2022.101517.

[31] A. R. S. Abadi, N. H. Hamzeh, D. G. Kaskaoutis, J. F. Vuillaume, K. A. Shukurov, and M. Gharibzadeh, "Spatio-Temporal Distribution of PM2.5 and PM10 Concentrations and Assessment of Public Health Risk in the Three Most Polluted Provinces of Iran," *Sustainability (Switzerland)*, vol. 17, no. 1, Jan. 2025, doi: 10.3390/su17010044.

[32] C. Pinichka, N. Makka, D. Sukkumnoed, S. Chariyalertsak, P. Inchai, and K. Bundhamcharoen, "Burden of disease attributed to ambient air pollution in Thailand: A GIS-based approach," *PLoS One*, vol. 12, no. 12, Dec. 2017, doi: 10.1371/journal.pone.0189909.

[33] S. Buya, A. Lim, R. Saelim, S. Musikasuwan, T. Choosong, and N. Taneepanichskul, "Impact of air pollution on cardiorespiratory morbidities in Southern Thailand," *Clin Epidemiol Glob Health*, vol. 25, p. 101501, Jan. 2024, doi: 10.1016/j.cegh.2023.101501.

[34] P. Ngamsang, T. Amnuaylojaroen, N. Parasin, and S. Pimonsree, "Health Impact Assessment of Short-Term Exposure to Particulate Matter (PM10) in Northern Thailand," *J Environ Public Health*, vol. 2023, pp. 1–9, May 2023, doi: 10.1155/2023/1237768.

[35] N. Parasin and T. Amnuaylojaroen, "Effect of PM2.5 on burden of mortality from non-communicable diseases in northern Thailand," *PeerJ*, vol. 12, no. 9, 2024, doi: 10.7717/peerj.18055.

[36] N. R. Fold *et al.*, "An Assessment of Annual Mortality Attributable to Ambient PM2.5 in Bangkok, Thailand," *Int J Environ Res Public Health*, vol. 17, no. 19, p. 7298, 2020, doi: 10.3390/ijerph17197298.

[37] P. Tang, T. Liu, X. Zheng, and J. Zheng, "Spatiotemporal Dynamics of PM2.5-Related Premature Deaths and the Role of Greening Improvement in Sustainable Urban Health Governance," *Atmosphere (Basel)*, vol. 16, no. 2, Feb. 2025, doi: 10.3390/atmos16020232.

[38] A. Leonetti, U. Peansukwech, J. Charnnarong, U. Cha'on, S. Suttiprapa, and S. Anutrakulchai, "Effects of particulate matter (PM2.5) concentration and components on mortality in chronic kidney disease patients: a nationwide spatial–temporal analysis," *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-67642-1.

[39] S. Pongpiachan *et al.*, "Combined use of principal component analysis/multiple linear regression analysis and artificial neural network to assess the impact of meteorological parameters on fluctuation of selected PM2.5-bound elements," *PLoS One*, vol. 19, no. 3 March, 2024, doi: 10.1371/journal.pone.0287187.

[40] S. Pongpiachan *et al.*, "Using synchrotron based ATR-FTIR, EXAFS, and XRF to characterize the chemical compositions of TSP in industrial estate area," *Heliyon*, vol. 10, no. 20, Oct. 2024, doi: 10.1016/j.heliyon.2024.e39215.

[41] S. Khruengsai *et al.*, "Seasonal and height dynamics of volatile organic compounds in rubber plantation: Impacts on ozone and secondary organic aerosol formation," *Science of the Total Environment*, vol. 945, Oct. 2024, doi: 10.1016/j.scitotenv.2024.173984.

[42] C. Pothirat *et al.*, "Acute effects of air pollutants on daily mortality and hospitalizations due to cardiovascular and respiratory diseases," *J Thorac Dis*, vol. 11, no. 7, 2019, doi: 10.21037/jtd.2019.07.37.

[43] K. Paoin *et al.*, "Long-term air pollution exposure and self-reported morbidity: A longitudinal analysis from the Thai cohort study (TCS)," *Environ Res*, vol. 192, 2021, doi: 10.1016/j.envres.2020.110330.

[44] S. Cole and E. Ellen, "New NASA satellite maps show human fingerprint on global air quality," *National Aeronautical Space Administration NASA: Washington, DC, USA*, 2015.

[45]   C. Lippi *et al.*, "The Social and Spatial Ecology of Dengue Presence and Burden during an Outbreak in Guayaquil, Ecuador, 2012," *Int J Environ Res Public Health*, vol. 15, no. 4, p. 827, Apr. 2018, doi: 10.3390/ijerph15040827.

[46]   A. S. Siraj *et al.*, "Data Descriptor: Spatiotemporal incidence of Zika and associated environmental drivers for the 2015-2016 epidemic in Colombia," 2018. doi: 10.1038/sdata.2018.73.

[47]   N. Wang *et al.*, "Lung cancer and particulate pollution: A critical review of spatial and temporal analysis evidence," 2018. doi: 10.1016/j.envres.2018.03.034.

[48]   Wikipedia contributors, "Regions of Thailand," In Wikipedia, The Free Encyclopedia. Accessed: Jan. 11, 2025. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Regions_of_Thailand&oldid=12474 17101

[49]   A. Lyapustin and Y. Wang, "MCD19A2 MODIS/Terra+ aqua land aerosol optical depth daily L2G global 1km SIN grid V006 [data set]," *NASA EOSDIS Land Processes DAAC*, 2018.

[50]   Z. Wan, S. Hook, and G. Hulley, "MOD11A1 MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V006. 2015, Distributed by NASA EOSDIS Land Processes DAAC," 2015.

[51]   Z. Wan, S. Hook, and G. Hulley, "MYD11A1 MODIS/Aqua land surface temperature/emissivity daily L3 global 1km SIN Grid V006. NASA EOSDIS LP DAAC," 2015.

[52]   K. Didan, "MOD13Q1 MODIS/Terra vegetation indices 16-day L3 global 250m SIN grid V006," *NASA EOSDIS Land Processes DAAC*, vol. 10, no. 1, 2015.

[53]   R Core Team., "R: A Language and Environment for Statistical Computing.," 2021, *R Foundation for Statistical Computing., Vienna, Austria.*: 4.1.2. Accessed: May 20, 2024. [Online]. Available: https://www.r-project.org/

[54]   L. Myers and M. J. Sirois, "Spearman Correlation Coefficients, Differences between," in *Encyclopedia of Statistical Sciences*, Hoboken, NJ, USA: John Wiley & Sons, Inc., 2006. doi: 10.1002/0471667196.ess5050.pub2.

[55] P. C. Consul and F. Famoye, "Generalized poisson regression model," *Commun Stat Theory Methods*, vol. 21, no. 1, pp. 89–109, Jan. 1992, doi: 10.1080/03610929208830766.

[56] P. Tongkumchum and D. McNeil, "Confidence intervals using contrasts for regression model," *Songklanakarin Journal of Science and Technology*, vol. 31, no. 2, 2009.

[57] A. Gasparrini, B. Armstrong, and M. G. Kenward, "Distributed lag non-linear models," *Stat Med*, vol. 29, no. 21, pp. 2224–2234, Sep. 2010, doi: 10.1002/sim.3940.

[58] P. A. P. Moran, "Notes on Continuous Stochastic Phenomena," *Biometrika*, vol. 37, no. 1/2, p. 17, Jun. 1950, doi: 10.2307/2332142.

[59] L. Anselin, "Local Indicators of Spatial Association-LISA," *Geogr Anal*, vol. 27, no. 2, pp. 93–115, Sep. 2010, doi: 10.1111/j.1538-4632.1995.tb00338.x.

[60] F. Mao, M. Duan, Q. Min, W. Gong, Z. Pan, and G. Liu, "Investigating the impact of haze on MODIS cloud detection," *J Geophys Res*, vol. 120, no. 23, pp. 12, 212–237, 247, 2015, doi: 10.1002/2015JD023555.

[61] Z. Pan, F. Mao, W. Wang, B. Zhu, X. Lu, and W. Gong, "Impacts of 3D aerosol, cloud, and water vapor variations on the recent brightening during the South Asian monsoon season," *Remote Sens (Basel)*, vol. 10, no. 4, 2018, doi: 10.3390/rs10040651.

[62] F. Dominici *et al.*, "Fine Particulate Air Pollution and Hospital Admission for Cardiovascular and Respiratory Diseases," *JAMA*, vol. 295, no. 10, pp. 1127–1134, 2006, doi: 10.1001/jama.295.10.1127.

[63] W. J. Gauderman *et al.*, "The Effect of Air Pollution on Lung Development from 10 to 18 Years of Age," *New England Journal of Medicine*, vol. 351, no. 11, pp. 1057–1067, 2004, doi: 10.1056/NEJMoa040610.

[64] J. Cai *et al.*, "Association of developmental coordination disorder with early-life exposure to fine particulate matter in Chinese preschoolers," *The Innovation*, vol. 4, no. 1, p. 100347, 2023, doi: https://doi.org/10.1016/j.xinn.2022.100347.

[65] X. Hu, L. A. Waller, A. Lyapustin, Y. Wang, and Y. Liu, "10-year spatial and temporal trends of $PM_{2.5}$ concentrations in the southeastern US estimated using

high-resolution satellite data," *Atmos. Chem. Phys.*, vol. 14, no. 12, pp. 6301–6314, 2014, doi: 10.5194/acp-14-6301-2014.

[66] X. Meng, M. J. Garay, D. J. Diner, O. V Kalashnikova, J. Xu, and Y. Liu, "Estimating PM2.5 speciation concentrations using prototype 4.4 km-resolution MISR aerosol properties over Southern California," *Atmos Environ*, vol. 181, pp. 70–81, 2018, doi: https://doi.org/10.1016/j.atmosenv.2018.03.019.

[67] W. Wang, F. Mao, L. Du, Z. Pan, W. Gong, and S. Fang, "Deriving Hourly PM2.5 Concentrations from Himawari-8 AODs over Beijing–Tianjin–Hebei in China," 2017. doi: 10.3390/rs9080858.

[68] J. Wu, F. Yao, W. Li, and M. Si, "VIIRS-based remote sensing estimation of ground-level PM2.5 concentrations in Beijing–Tianjin–Hebei: A spatiotemporal statistical model," *Remote Sens Environ*, vol. 184, pp. 316–328, 2016, doi: https://doi.org/10.1016/j.rse.2016.07.015.

[69] J. W. Xu *et al.*, "Estimating ground-level PM$_{2.5}$ in eastern China using aerosol optical depth determined from the GOCI satellite instrument," *Atmos. Chem. Phys.*, vol. 15, no. 22, pp. 13133–13144, 2015, doi: 10.5194/acp-15-13133-2015.

[70] L. Zang, F. Mao, J. Guo, W. Gong, W. Wang, and Z. Pan, "Estimating hourly PM1 concentrations from Himawari-8 aerosol optical depth in China," *Environmental Pollution*, vol. 241, pp. 654–663, 2018, doi: https://doi.org/10.1016/j.envpol.2018.05.100.

[71] J. Guo *et al.*, "Impact of diurnal variability and meteorological factors on the PM2.5 - AOD relationship: Implications for PM2.5 remote sensing," *Environmental Pollution*, vol. 221, pp. 94–104, 2017, doi: https://doi.org/10.1016/j.envpol.2016.11.043.

[72] M. Feiyue *et al.*, "Assimilating moderate resolution imaging spectroradiometer radiance with the weather research and forecasting data assimilation system," *J Appl Remote Sens*, vol. 11, no. 3, p. 36002, 2017, doi: 10.1117/1.JRS.11.036002.

[73] X. Ma, J. Wang, F. Yu, H. Jia, and Y. Hu, "Can MODIS AOD be employed to derive PM2.5 in Beijing-Tianjin-Hebei over China?," *Atmos Res*, vol. 181, pp. 250–256, 2016, doi: https://doi.org/10.1016/j.atmosres.2016.06.018.

[74] Q. Xu, X. Chen, S. Yang, L. Tang, and J. Dong, "Spatiotemporal relationship between Himawari-8 hourly columnar aerosol optical depth (AOD) and ground-

level PM2.5 mass concentration in mainland China," *Science of The Total Environment*, vol. 765, p. 144241, 2021, doi: https://doi.org/10.1016/j.scitotenv.2020.144241.

[75] Q. Yang, Q. Yuan, L. Yue, T. Li, H. Shen, and L. Zhang, "The relationships between PM2. 5 and aerosol optical depth (AOD) in mainland China: About and behind the spatio-temporal variations," *Environmental Pollution*, vol. 248, pp. 526–535, 2019.

[76] Q. Xiao *et al.*, "Full-coverage high-resolution daily PM2.5 estimation using MAIAC AOD in the Yangtze River Delta of China," *Remote Sens Environ*, vol. 199, pp. 437–446, 2017, doi: https://doi.org/10.1016/j.rse.2017.07.023.

[77] Y. Liu, J. A. Sarnat, V. Kilaru, D. J. Jacob, and P. Koutrakis, "Estimating Ground-Level PM2.5 in the Eastern United States Using Satellite Remote Sensing," *Environ Sci Technol*, vol. 39, no. 9, pp. 3269–3278, 2005, doi: 10.1021/es049352m.

[78] M. Green, S. Kondragunta, P. Ciren, and C. Xu, "Comparison of GOES and MODIS Aerosol Optical Depth (AOD) to Aerosol Robotic Network (AERONET) AOD and IMPROVE PM2.5 Mass at Bondville, Illinois," *J Air Waste Manage Assoc*, vol. 59, no. 9, pp. 1082–1091, 2009, doi: 10.3155/1047-3289.59.9.1082.

[79] N. Kumar, D. Liang, A. Comellas, A. D. Chu, and T. Abrams, "Satellite-based PM concentrations and their application to COPD in Cleveland, OH," *J Expo Sci Environ Epidemiol*, vol. 23, no. 6, pp. 637–646, 2013, doi: 10.1038/jes.2013.52.

[80] L. Jean-Francois, L. Cathy, G.-L. Corinne, D. Thierno, and C. Hélène, "Monitoring of ambient fine particulate matter concentrations from space: application to European and African cities," in *Proc.SPIE*, 2010, p. 78262A. doi: 10.1117/12.864954.

[81] M. Schaap, A. Apituley, R. M. A. Timmermans, R. B. A. Koelemeijer, and G. de Leeuw, "Exploring the relation between aerosol optical depth and PM$_{2.5}$ at Cabauw, the Netherlands," *Atmos. Chem. Phys.*, vol. 9, no. 3, pp. 909–925, 2009, doi: 10.5194/acp-9-909-2009.

[82] J. Tao *et al.*, "A method to estimate concentrations of surface-level particulate matter using satellite-based aerosol optical thickness," *Sci China Earth Sci*, vol. 56, no. 8, pp. 1422–1433, 2013, doi: 10.1007/s11430-012-4503-3.

[83] G. Casella, *Statistical Design*. New York, NY: Springer New York, 2008. doi: 10.1007/978-0-387-75965-4.

[84] S. Abdulmana, A. Lim, S. Wongsai, and N. Wongsai, "Effect of land cover change and elevation on decadal trend of land surface temperature: a linear model with sum contrast analysis," *Theor Appl Climatol*, vol. 149, no. 1–2, pp. 425–436, 2022, doi: 10.1007/s00704-022-04038-z.

[85] G. Af, "Blood lead levels among school children living in the Pattani River Basin: Two contamination scenarios?," *J Environ Med*, vol. 2, pp. 11–16, 2000.

[86] H. D. Tulu, A. Lim, A. Ma-A-Lee, K. Bundhamcharoen, and N. Makka, "Prediction of HIV mortality in Thailand using three data sets from the national AIDS program database," *Sains Malays*, vol. 49, no. 1, pp. 155–160, 2020, doi: 10.17576/jsm-2020-4901-19.

[87] S. Buya, P. Tongkumchum, and B. E. Owusu, "Modelling of land-use change in Thailand using binary logistic regression and multinomial logistic regression," *Arabian Journal of Geosciences*, vol. 13, no. 12, 2020, doi: 10.1007/s12517-020-05451-2.

[88] S. Buya, P. Tongkumchum, K. Rittiboon, and S. Chaimontree, "Logistic Regression Model of Built-Up Land Based on Grid-Digitized Data Structure: A Case Study of Krabi, Thailand," *Journal of the Indian Society of Remote Sensing*, vol. 50, no. 5, 2022, doi: 10.1007/s12524-022-01503-0.

[89] P. Chuangchang, O. Thinnukool, and P. Tongkumchum, "Modelling urban growth over time using grid-digitized method with variance inflation factors applied to spatial correlation," *Arabian Journal of Geosciences*, vol. 9, no. 5, p. 342, 2016, doi: 10.1007/s12517-016-2375-0.

[90] S. Buya, P. Chuangchang, and B. A. Owusu, "Analysis of land surface temperature with land use and land cover and elevation from NASA MODIS satellite data: a case study of Bali, Indonesia," *Environ Monit Assess*, vol. 194, no. 8, 2022, doi: 10.1007/s10661-022-10252-z.

[91] L. Zhang, J. P. Wilson, B. MacDonald, W. Zhang, and T. Yu, "The changing PM2.5 dynamics of global megacities based on long-term remotely sensed observations," *Environ Int*, vol. 142, p. 105862, 2020, doi: https://doi.org/10.1016/j.envint.2020.105862.

[92] T. Zhang, L. Zang, Y. Wan, W. Wang, and Y. Zhang, "Ground-level PM2.5 estimation over urban agglomerations in China with high spatiotemporal resolution based on Himawari-8," *Science of The Total Environment*, vol. 676, pp. 535–544, 2019, doi: https://doi.org/10.1016/j.scitotenv.2019.04.299.

[93] B. Peng-in, P. Sanitluea, P. Monjatturat, P. Boonkerd, and A. Phosri, "Estimating ground-level PM2.5 over Bangkok Metropolitan Region in Thailand using aerosol optical depth retrieved by MODIS," *Air Qual Atmos Health*, vol. 15, no. 11, pp. 2091–2102, 2022, doi: 10.1007/s11869-022-01238-4.

[94] P. Wongnakae, P. Chitchum, R. Sripramong, and A. Phosri, "Application of satellite remote sensing data and random forest approach to estimate ground-level PM2.5 concentration in Northern region of Thailand," *Environmental Science and Pollution Research*, 2023, doi: 10.1007/s11356-023-28698-0.

[95] S. Buya, S. Usanavasin, H. Gokon, and J. Karnjana, "An Estimation of Daily PM2.5 Concentration in Thailand Using Satellite Data at 1-Kilometer Resolution," *Sustainability*, vol. 15, no. 13, 2023, doi: 10.3390/su151310024.

[96] S. Buya, S. Usanavasin, G. Hideomi, and J. Karnjana, "The Association of Satellite Data with PM $_{2.5}$ Data from Ground Monitoring Stations in Thailand," in *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, Jul. 2023, pp. 1533–1536. doi: 10.1109/IGARSS52108.2023.10282670.

[97] Z.-Y. Chen, T.-H. Zhang, R. Zhang, Z.-M. Zhu, C.-Q. Ou, and Y. Guo, "Estimating PM2.5 concentrations based on non-linear exposure-lag-response associations with aerosol optical depth and meteorological measures," *Atmos Environ*, vol. 173, pp. 30–37, 2018, doi: https://doi.org/10.1016/j.atmosenv.2017.10.055.

[98] Y. Liu, M. Franklin, R. Kahn, and P. Koutrakis, "Using aerosol optical thickness to predict ground-level PM2.5 concentrations in the St. Louis area: A

comparison between MISR and MODIS," *Remote Sens Environ*, vol. 107, no. 1, pp. 33–44, 2007, doi: https://doi.org/10.1016/j.rse.2006.05.022.

[99]   J. Tian and D. Chen, "A semi-empirical model for predicting hourly ground-level fine particulate matter (PM2.5) concentration in southern Ontario from satellite remote sensing and ground-based meteorological measurements," *Remote Sens Environ*, vol. 114, no. 2, pp. 221–229, 2010, doi: https://doi.org/10.1016/j.rse.2009.09.011.

[100] W. Song, H. Jia, J. Huang, and Y. Zhang, "A satellite-based geographically weighted regression model for regional PM2.5 estimation over the Pearl River Delta region in China," *Remote Sens Environ*, vol. 154, pp. 1–7, 2014, doi: https://doi.org/10.1016/j.rse.2014.08.008.

[101] W. Thongpeth, A. Lim, S. Kraonual, A. Wongpairin, and T. Thongpeth, "Determinants of hospital costs for management of chronic-disease patients in southern thailand," *Journal of Health Science and Medical Research*, vol. 39, no. 4, pp. 313–320, 2021, doi: 10.31584/jhsmr.2021787.

[102] K. Rani Das, "A Brief Review of Tests for Normality," *American Journal of Theoretical and Applied Statistics*, vol. 5, no. 1, p. 5, 2016, doi: 10.11648/j.ajtas.20160501.12.

[103] M. Danesh Yazdi *et al.*, "Predicting Fine Particulate Matter (PM2.5) in the Greater London Area: An Ensemble Approach using Machine Learning Methods," 2020. doi: 10.3390/rs12060914.

[104] N. Chuersuwan, S. Nimrat, S. Lekphet, and T. Kerdkumrai, "Levels and major sources of PM2.5 and PM10 in Bangkok Metropolitan Region," *Environ Int*, vol. 34, no. 5, pp. 671–677, 2008, doi: https://doi.org/10.1016/j.envint.2007.12.018.

[105] C. ChooChuay *et al.*, "Impacts of PM2.5 sources on variations in particulate chemical compounds in ambient air of Bangkok, Thailand," *Atmos Pollut Res*, vol. 11, no. 9, pp. 1657–1667, 2020, doi: 10.1016/j.apr.2020.06.030.

[106] D. Yang, C. Ye, X. Wang, D. Lu, J. Xu, and H. Yang, "Global distribution and evolvement of urbanization and PM2.5 (1998–2015)," *Atmos Environ*, vol. 182, pp. 171–178, 2018, doi: https://doi.org/10.1016/j.atmosenv.2018.03.053.

[107] T. Liu *et al.*, "Urban-rural disparity of the short-term association of PM2.5 with mortality and its attributable burden," *The Innovation*, vol. 2, no. 4, p. 100171, Nov. 2021, doi: 10.1016/j.xinn.2021.100171.

[108] D. Bzdok, N. Altman, and M. Krzywinski, "Statistics versus machine learning," *Nat Methods*, vol. 15, no. 4, 2018, doi: 10.1038/nmeth.4642.

[109] J. Li, X. Ge, Q. He, and A. Abbas, "Aerosol optical depth (AOD): spatial and temporal variations and association with meteorological covariates in Taklimakan desert, China," *PeerJ*, vol. 9, 2021, doi: 10.7717/peerj.10542.

[110] M. Danesh Yazdi *et al.*, "Predicting fine particulate matter (PM2. 5) in the greater london area: An ensemble approach using machine learning methods," *Remote Sens (Basel)*, vol. 12, no. 6, p. 914, 2020.

[111] M. Stafoggia *et al.*, "Estimation of daily PM$_{10}$ and PM$_{2.5}$ concentrations in Italy, 2013–2015, using a spatiotemporal land-use random-forest model," *Environ Int*, vol. 124, pp. 170–179, 2019, doi: 10.1016/j.envint.2019.01.016.

[112] J. Wei *et al.*, "Estimating 1-km-resolution PM$_{2.5}$ concentrations across China using the space-time random forest approach," *Remote Sens Environ*, vol. 231, 2019, doi: 10.1016/j.rse.2019.111221.

[113] Q. Xiao, H. H. Chang, G. Geng, and Y. Liu, "An Ensemble Machine-Learning Model to Predict Historical PM$_{2.5}$ Concentrations in China from Satellite Data," *Environ Sci Technol*, vol. 52, no. 22, pp. 13260–13269, 2018, doi: 10.1021/acs.est.8b02917.

[114] A. L. Boulesteix and M. Schmid, "Machine learning versus statistical modeling," 2014. doi: 10.1002/bimj.201300226.

[115] K. Kourou, T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis, and D. I. Fotiadis, "Machine learning applications in cancer prognosis and prediction," 2015. doi: 10.1016/j.csbj.2014.11.005.

[116] D. Bzdok, N. Altman, and M. Krzywinski, "Statistics versus machine learning," *Nat Methods*, vol. 15, no. 4, 2018, doi: 10.1038/nmeth.4642.

[117] C.-C. Chen, Y.-R. Wang, H.-Y. Yeh, T.-H. Lin, C.-S. Huang, and C.-F. Wu, "Estimating monthly PM2.5 concentrations from satellite remote sensing data, meteorological variables, and land use data using ensemble statistical modeling

and a random forest approach," *Environmental Pollution*, vol. 291, p. 118159, 2021, doi: https://doi.org/10.1016/j.envpol.2021.118159.

[118] Q. Di *et al.*, "Air Pollution and Mortality in the Medicare Population," *New England Journal of Medicine*, vol. 376, no. 26, pp. 2513–2522, Jun. 2017, doi: 10.1056/nejmoa1702747.

[119] W. Bin Liao, K. Ju, Q. Zhou, Y. M. Gao, and J. Pan, "Forecasting PM2.5-induced lung cancer mortality and morbidity at county level in China using satellite-derived PM2.5 data from 1998 to 2016: a modeling study," *Environmental Science and Pollution Research*, vol. 27, no. 18, pp. 22946–22955, Jun. 2020, doi: 10.1007/s11356-020-08843-9.

[120] S. Mandal *et al.*, "PM 2.5 exposure, glycemic markers and incidence of type 2 diabetes in two large Indian cities," *BMJ Open Diabetes Res Care*, vol. 11, no. 5, Oct. 2023, doi: 10.1136/bmjdrc-2023-003333.

[121] S. Faridi *et al.*, "Health burden and economic loss attributable to ambient PM2.5 in Iran based on the ground and satellite data," *Sci Rep*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-18613-x.

[122] C. Pothirat, A. Tosukhowong, W. Chaiwong, C. Liwsrisakun, and J. Inchai, "Effects of seasonal smog on asthma and COPD exacerbations requiring emergency visits in Chiang Mai, Thailand," *Asian Pac J Allergy Immunol*, vol. 34, no. 4, 2016, doi: 10.12932/AP0668.

[123] R. W. Atkinson, I. C. Mills, H. A. Walton, and H. R. Anderson, "Fine particle components and health—a systematic review and meta-analysis of epidemiological time series studies of daily mortality and hospital admissions," *J Expo Sci Environ Epidemiol*, vol. 25, no. 2, pp. 208–214, Mar. 2015, doi: 10.1038/jes.2014.63.

[124] L. Fajersztajn, P. Saldiva, L. A. A. Pereira, V. F. Leite, and A. M. Buehler, "Short-term effects of fine particulate matter pollution on daily health events in Latin America: a systematic review and meta-analysis," *Int J Public Health*, vol. 62, no. 7, pp. 729–738, Sep. 2017, doi: 10.1007/s00038-017-0960-y.

[125] O. Raaschou-Nielsen *et al.*, "Traffic air pollution and mortality from cardiovascular disease and all causes: A Danish cohort study," *Environ Health*, vol. 11, no. 1, 2012, doi: 10.1186/1476-069X-11-60.

[126] E. Samoli *et al.*, "Associations of short-term exposure to traffic-related air pollution with cardiovascular and respiratory hospital admissions in London, UK," *Occup Environ Med*, vol. 73, no. 5, 2016, doi: 10.1136/oemed-2015-103136.

[127] P. Pengchai, S. Chantara, K. Sopajaree, S. Wangkarn, U. Tengcharoenkul, and M. Rayanakorn, "Seasonal variation, risk assessment and source estimation of PM 10 and PM10-bound PAHs in the ambient air of Chiang Mai and Lamphun, Thailand," *Environ Monit Assess*, vol. 154, no. 1–4, pp. 197–218, Jul. 2009, doi: 10.1007/s10661-008-0389-0.

[128] S. Pongpiachan, M. Hattayanone, and J. Cao, "Effect of agricultural waste burning season on PM 2.5 -bound polycyclic aromatic hydrocarbon (PAH) levels in Northern Thailand," *Atmos Pollut Res*, vol. 8, no. 6, pp. 1069–1080, Nov. 2017, doi: 10.1016/j.apr.2017.04.009.

[129] P. Suriyawong *et al.*, "Airborne particulate matter from biomass burning in Thailand: Recent issues, challenges, and options," *Heliyon*, vol. 9, no. 3, p. e14261, Mar. 2023, doi: 10.1016/j.heliyon.2023.e14261.

[130] WHO, *WHO global air quality guidelines: particulate matter (PM2. 5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. World Health Organization, 2021. Accessed: Jan. 24, 2025. [Online]. Available: https://www.who.int/publications/i/item/9789240034228

Appendix

Ethics of cardiorespiratory data

      The cardiorespiratory mortality data used in this research received approval from the Research Ethics Review Committee for Research Involving Human Research Participants, Chulalongkorn University (certificate of analysis no. 093/65). This approval aligns with the Office for Human Research Protections (OHRP Exempt Categories) 45 CFR part 46.101(b).

The Research Ethics Review Committee for Reseach Involving Human Research Participants,
Group I, Chulalongkorn University

Chamchuri 1 Building, 2nd Floor, 254 Phayathai Road, Pathumwan, Bangkok 10330 Thailand

Telephone: 02-218-3202, 02-218-3049 Email: eccu@chula.ac.th

COA No. 093/65

Certificate of Approval

Exemption for Ethics Review

Study Title No. 650042	:	REVISED NATIONAL AMBIENT AIR QUALITY STANDARD FOR PARTICULATE MATTER (PM2.5)
Principal Investigator	:	Assistant Prof. Dr. Nutta Taneepanichskul
Place of Proposed Study/institution	:	College of Public Health Sciences, Chulalongkorn University

      This Research proposal is exempted for ethics review in compliance with the Office for Human Research Protections (OHRP Exempt Categories) 45 CFR part 46.101(b).

Certified under condition: To conduct this research project, the researcher (s) must strictly adhere to research proposal approved by the committee. If there is any amendment, it must be sent to the committee for review before carrying on the project.

Signature *Prida Tasanapradit*

(Associate Prof. Prida Tasanapradit)

Chairman

Signature *Raveenan Mingpaleanu.*

(Assistant Prof. Dr. Raveenan Mingpakanee)

Secretary

Date of Approval : 21 April 2022

Remarks

Final report (AF 01-15) and abstract is required for a one year (or less) research/project and report within 30 days after the completion of the research/project.

Thesis publications

List of International Journals:

1. The first international journal paper, **"An Estimation of Daily PM2.5 Concentration in Thailand Using Satellite Data at 1-Kilometer Resolution,"** presents the PM2.5 dataset used in this study, covering all areas of Thailand.

   • Buya, S., Usanavasin, S., Gokon, H., & Karnjana, J. (2023). An Estimation of Daily PM2.5 Concentration in Thailand Using Satellite Data at 1-Kilometer Resolution. *Sustainability, 15*(13). https://doi.org/10.3390/su151310024. [IF: **3.9**] (Peer-Reviewed International Journal, **Q1**)

2. The second paper, **"Estimating Ground-Level Hourly PM2.5 Concentrations in Thailand Using Satellite Data: A Log-Linear Model With Sum Contrast Analysis,"** explores the characteristics of satellite data in relation to hourly PM2.5 concentrations.

   • Buya, S., Gokon, H., Dam, H. C., Usanavasin, S., & Karnjana, J. (2024). Estimating Ground-Level Hourly PM 2.5 Concentrations in Thailand Using Satellite Data: A Log-Linear Model With Sum Contrast Analysis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17*, 8203–8212. https://doi.org/10.1109/JSTARS.2024.3384964. [IF: **5.5**] (Peer-Reviewed International Journal, **Q1**)

3. The third paper, **"Spatiotemporal Association Between Monthly PM2.5 Levels and Cardiorespiratory Mortality in Thailand (2015–2019),"** examines the spatiotemporal relationship between PM2.5 exposure and cardiorespiratory mortality across Thailand from 2015 to 2019, addressing a critical research gap in geographical coverage.

   • Buya, S., Gokon, H., Huynh, V. N., Dam, H. C., Usanavasin, S., Karnjana, J., & Taneepanichskul, N. (2025). Spatiotemporal association between monthly PM2. 5 levels and cardiorespiratory mortality in Thailand (2015–2019). *International Journal of Environmental Health Research*, 1-12. https://doi.org/10.1080/09603123.2025.2458726. [IF: **2.2**] (Peer-Reviewed International Journal, **Q1**)

List of International Conferences:

4. **"The Association of Satellite Data With PM2.5 Data From Ground Monitoring Stations in Thailand"** discusses the selection of satellite data for PM2.5 estimation.

- Buya, S., Usanavasin, S., Hideomi, G., & Karnjana, J. (2023, July). The Association of Satellite Data with PM 2.5 Data from Ground Monitoring Stations in Thailand. In *IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium* (pp. 1533-1536). IEEE. Pasadena, CA, USA, 16 – 21 July 2023. https://doi.org/10.1109/IGARSS52108.2023.10282670.　　　(Peer-Reviewed International Conference, **Q3**)

5. **"Evaluating MODIS and VIIRS Satellite AOD for Ground-Level PM2.5 Estimation in Thailand**.". The second conference paper compares different satellite products used for PM2.5 estimation.

- Buya, S., Gokon, H., Usanavasin, S., & Karnjana, J. Evaluating MODIS and VIIRS Satellite AOD for Ground-Level PM2.5 Estimation in Thailand. Submission to an *IEEE International Geoscience and Remote Sensing Symposium* (IGARSS) 2025. Brisbane, Australia, on 3-8 August 2025. https://www.2025.ieeeigarss.org/

The papers have been attached to this section.

Bibliography

Name	Suhaimee Buya
Education	2013: Bachelor of Science (Applied Mathematics) Faculty of Science and Technology, Prince of Songkla University, Thailand
	2016: Master of Science (Research Methodology) Faculty of Science and Technology, Prince of Songkla University, Thailand
	2024: Doctor of Philosophy (Engineering and Technology) Sirindhorn International Institute of Technology, Thammasat University, Thailand

Publications

List of Publications in International Journals:

[1] Buya, S., & Gokon, H. (2025). Topographic transformations and coastal implications: assessing the impact of the 2024 Noto earthquake. *Coastal Engineering Journal*, 1–9. https://doi.org/10.1080/21664250.2025.2540718

[2] Buya, S., Gokon, H., Huynh, V. N., Dam, H. C., Usanavasin, S., Karnjana, J., & Taneepanichskul, N. (2025). Spatiotemporal association between monthly PM2. 5 levels and cardiorespiratory mortality in Thailand (2015–2019). *International Journal of Environmental Health Research*, 1-12. http://dx.doi.org/10.1080/09603123.2025.2458726

[3] Buya, S., Gokon, H., Dam, H. C., Usanavasin, S., & Karnjana, J. (2024). Estimating Ground-Level Hourly PM 2.5 Concentrations in Thailand Using Satellite Data: A Log-Linear Model With Sum Contrast Analysis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17*, 8203–8212. https://doi.org/10.1109/JSTARS.2024.3384964

[4] Buya, S., Usanavasin, S., Gokon, H., & Karnjana, J. (2023). An Estimation of Daily PM2.5 Concentration in Thailand Using Satellite Data at 1-Kilometer Resolution. *Sustainability*, *15*(13). https://doi.org/10.3390/su151310024

[5] Buya, S., Lim, A., Saelim, R., Musikasuwan, S., Choosong, T., & Taneepanichskul, N. (2024). Impact of air pollution on cardiorespiratory morbidities in Southern Thailand. *Clinical Epidemiology and Global Health*, *25*, 101501. https://doi.org/10.1016/j.cegh.2023.101501

[6] Buya, S., Chuangchang, P., & Owusu, B. A. (2022). Analysis of land surface temperature with land use and land cover and elevation from NASA MODIS satellite data: a case study of Bali, Indonesia. *Environmental Monitoring and Assessment*, *194*(8), 566. https://doi.org/10.1007/s10661-022-10252-z

[7] Buya, S., Tongkumchum, P., Rittiboon, K., & Chaimontree, S. (2022). Logistic Regression Model of Built-Up Land Based on Grid-Digitized Data Structure: A Case Study of Krabi, Thailand. *Journal of the Indian Society of Remote Sensing*, *50*(5), 909-922. https://doi.org/10.1007/s12524-022-01503-0

[8] Buya, S., Tongkumchum, P., & Owusu, B. E. (2020). Modelling of land-use change in Thailand using binary logistic regression and multinomial logistic regression. *Arabian journal of geosciences*, *13*, 1-12. https://doi.org/10.1007/s12517-020-05451-2

[9] Munawar, M., McNeil, R., Jani, R., Buya, S., & Tarmizi, T. (2025). Variations in land surface temperature increase in South-East Asian Cities. *Environmental Monitoring and Assessment*, *197*(2), 190. http://dx.doi.org/10.1007/s10661-024-13604-z

[10] Lim, A., Owusu, B. A., Thongrod, T., Khurram, H., Pongsiri, N., Ingviya, T., & Buya, S. (2024). Trend and Association Between Particulate Matters and Meteorological Factors: A Prospect for Prediction of PM2.5 in Southern Thailand. *Polish Journal of Environmental Studies.* https://doi.org/10.15244/pjoes/190787

[11] Vichitkunakorn, P., Assanangkornchai, S., Thaikla, K., Buya, S., Rungruang, S., Talib, M., Duangpaen, W., Bunyanukul, W., & Sittisombut, M. (2024). Alcohol outlet density and adolescent drinking behaviors in Thailand, 2007–2017: A spatiotemporal mixed model analysis. *PLoS ONE, 19*(10). https://doi.org/10.1371/journal.pone.0308184

[12] Paoin, K., Pharino, C., Vathesatogkit, P., Phosri, A., Buya, S., Ueda, K., Seposo, X.T., Ingviya, T., Saranburut, K., Thongmung, N., & Yingchoncharoen, T. (2023). Associations between residential greenness and air pollution and the incident metabolic syndrome in a Thai worker cohort. *International Journal of Biometeorology*, *67*(12), 1965-1974. https://doi.org/10.1007/s00484-023-02554-9

[13] Munawar, M., Prasetya, T. A. E., McNeil, R., Jani, R., & Buya, S. (2023). Spatio and Temporal Analysis of Indonesia Land Surface Temperature Variation During 2001–2020. *Journal of the Indian Society of Remote Sensing*. https://doi.org/10.1007/s12524-023-01713-0

[14] Paoin, K., Pharino, C., Vathesatogkit, P., Buya, S., Saranburut, K., Phosri, A., Ueda, K., Seposo, X. T., Ingviya, T., Kitiyakara, T., Thongmung, N., & Sritara, P. (2023). Long-term associations of air pollution exposure with liver enzymes among adult employees of the Electricity Generating Authority of Thailand: A longitudinal cohort study. *Atmospheric Environment*, *299*. https://doi.org/10.1016/j.atmosenv.2023.119648

[15] Paoin, K., Pharino, C., Vathesatogkit, P., Phosri, A., Buya, S., Saranburut, K., Ueda, K., Seposo, X. T., Ingviya, T., Kitiyakara, C., Thongmung, N., & Sritara, P. (2023). Residential greenness and kidney function: A cohort study of Thai employees. *Health and Place*, *80*. https://doi.org/10.1016/j.healthplace.2023.102993

[16] Geater, S. L., Chaniad, P., Trakunram, K., Keeratichananont, W., Buya, S., Thongsuksai, P., & Raungrut, P. (2022). Diagnostic and prognostic value of serum miR-145 and vascular endothelial growth factor in non-small cell lung cancer. *Oncology Letters*, *23*(1), 1-11. https://doi.org/10.3892/ol.2021.13130

[17] Paoin, K., Ueda, K., Vathesatogkit, P., Ingviya, T., Buya, S., Dejchanchaiwong, R., Phosri, A., Seposo, X.T., Kitiyakara, C., Thongmung, N., Honda, A., Takano, H., Sritara, P., & Tekasakul, P. (2022). Long-term air pollution exposure and decreased kidney function: A longitudinal cohort study in Bangkok Metropolitan Region, Thailand from 2002 to

2012. *Chemosphere*, *287*, 132117. https://doi.org/10.1016/j.chemosphere.2021.132117

[18] Paoin, K., Ueda, K., Ingviya, T., Buya, S., Phosri, A., Seposo, X.T., Seubsman, S.A., Kelly, M., Sleigh, A., Honda, A., & Takano, H. (2021). Long-term air pollution exposure and self-reported morbidity: A longitudinal analysis from the Thai cohort study (TCS). *Environmental Research*, *192*, 110330. https://doi.org/10.1016/j.envres.2020.110330

[19] Paoin, K., Ueda, K., Vathesatogkit, P., Ingviya, T., Buya, S., Phosri, A., Seposo, X.T., Thongmung, N., Yingchoncharoen, T., Honda, A., Takano, H., & Sritara, P. (2021). Long-term air pollution exposure and serum lipids and blood sugar: A longitudinal cohort study from the electricity generating authority of Thailand study. *Atmospheric Environment*, *259*, 118515. https://doi.org/10.1016/j.atmosenv.2021.118515

[20] Paoin, K., Ueda, K., Vathesatogkit, P., Ingviya, T., Buya, S., Phosri, A., Seposo, X.T., Thongmung, N., Yingchoncharoen, T., Honda, A., Takano, H., & Sritara, P. (2021). Effects of long-term air pollution exposure on ankle-brachial index and cardio-ankle vascular index: A longitudinal cohort study using data from the Electricity Generating Authority of Thailand study. *International Journal of Hygiene and Environmental Health*, *236*, 113790. https://doi.org/10.1016/j.ijheh.2021.113790

[21] Trakunram, K., Chaniad, P., Geater, S.L., Keeratichananont, W., Chittithavorn, V., Uttayamakul, S., Buya, S., Raungrut, P., & Thongsuksai, P. (2020). Serum miR-339-3p as a potential diagnostic marker for non-small cell lung cancer. *Cancer Biology & Medicine*, *17*(3), 652. https://doi.org/10.20892%2Fj.issn.2095-3941.2020.0063

List of Publications in International Conferences:

[1] Buya, S., Usanavasin, S., Hideomi, G., & Karnjana, J. (2023, July). The Association of Satellite Data with PM 2.5 Data from Ground Monitoring Stations in Thailand. In *IGARSS 2023-2023 IEEE International Geoscience*

and Remote Sensing Symposium (pp. 1533-1536). IEEE. Pasadena, CA, USA, 16 – 21 July 2023. https://doi.org/10.1109/IGARSS52108.2023.10282670

[2] Paoin, K., Pharino, C., Vathesatogkit, P., Phosri, A., Buya, S., Ueda, K., Seposo, X. T., Ingviya, T., Saranburut, K., & Thongmung, N. (2023, March). Long-Term Associations between Greenness and Air Pollution with Risk Factors for Cardiovascular Disease: A Cohort Study. In *the 8th World Congress on Civil, Structural, and Environmental Engineering (CSEE'23) Conference. Paper No.* ICEPTP 146. March 29 – 31, 2023, in Lisbon, Portugal –. http://dx.doi.org/10.11159/iceptp23.146