Application of Multilayer Perceptron Artificial Neural Network (MLP-ANN) Algorithm for $PM_{2.5}$ Mass Concentration Estimation during Open Biomass Burning Episodes in Thailand

Paluang, P.,[1] Thavorntam, W.[1,2*] and Phairuang, W.[1,3*]
[1]Department of Geography, Faculty of Social Sciences, Chiang Mai University, Muang, Chiang Mai 50200 Thailand, E-mail: Phakphum_p@cmu.ac.th, watinee.thavorntam@cmu.ac.th,* worradorn.ph@cmu.ac.th*
[2]School of Science, Edith Cowan University, Joondalup WA 6027, Australia
E-mail: w.thavorntam@ecu.edu.au
[3]Faculty of Geosciences and Civil Engineering, Institute of Science and Engineering, Kanazawa University Kanazawa, Ishikawa 920-1192, Japan, E-mail: pworradorn@se.kanazawa-u.ac.jp
**Corresponding Author*

DOI: https://doi.org/10.52939/ijg.v20i7.3401

Abstract
Open biomass burning (OBB) is the main cause of air pollution in Northern Thailand, where $PM_{2.5}$ concentrations exceed Thailand's air quality standards annually during the January–April (dry season). The air emissions from databases that detail the pollutants discharged into the atmosphere from specific sources of air pollution are crucial for monitoring air pollution. However, this data has been poorly studied in Thailand. This study estimated ground-level $PM_{2.5}$ concentration in Northern Thailand using the Multilayer Perceptron Artificial Neural Networks (MLP-ANN) model, integrating the in-depth data as input variables. The 10-fold cross-validation approach was applied to validate the model's performance. The meteorological and aerosol optical depth (AOD), which were the satellite data detected by a sensor from the Moderate Resolution Imaging Spectroradiometer (MODIS) Terra and Aqua satellites, were used as input variables. One important input variable was the OBB emission, which consists of forest fire and crop residue burning air emissions. The best modeling was found with the optimal architecture networks, with 8-16-1 indicating the lowest mean absolute error (MAE) values at 0.0187 and root mean square error (RMSE) at 0.0282. The model result was observed as an underestimate between the model result and the actual data from the Pollution Control Department (PCD), for which the limitation of the training datasets was the reason. Moreover, this model was also applied to estimates in Khon Kaen, Nakhon Ratchasima, and Ubon Ratchathani provinces. The model results were still underestimated compared data from the PCD. The primary reasons were the difference in the geographical characteristics and air pollution, including the error model. The highlight finding in this study indicated that integrating the input variables as the meteorological data, the AOD, and the OBB emissions processing with MLP-ANN enhances model performance and can be extended to further estimation in areas without air monitoring stations.

Keywords: Air Pollution, AOD, Artificial Neural Networks, Biomass Burning, Meteorological Data

1. Introduction

Ambient particulate matter (PM) is a major air pollutant comprising a heterogeneous mixture of particle sizes and chemical components [1][2][3] and [4]. Significantly, prolonged exposure to fine PM (PM with a diameter less than 2.5 microns; $PM_{2.5}$) concentrations can harm human health, particularly impacting lung function and potentially leading to more severe consequences that can affect the entire body. This is especially concerning for respiratory-related illnesses such as inflammation in the respiratory tract and the development of the lungs [5] and [6]. Fine particles have directly impacted the Earth's atmospheric radiation force, increasing global warming over the past decade [3]. Furthermore, particles serve as cloud condensation nuclei (CCN), indirectly affecting the albedo and lifespan of clouds [2].

$PM_{2.5}$ can originate from various sources, including biomass burning, transportation, and industrial sectors. In developing countries, agricultural waste residues are a significant source of air pollution [7] and [8].

Various methodologies have been developed to measure the level of air pollutant concentration. One prominent method is remote sensing, which has advanced with the development of satellite technology. This is evident from the spatial resolution of satellite images, such as Landsat-8 OLI (Operational Land Imager), TIRS (Thermal Infrared Sensor), and Landsat-9 OLI/TRI, which have a resolution of 15–30 meters [9][10] and [11].

Additionally, satellites like Japan's Greenhouse Gases Observing Satellite-2 (GOSAT-2), India's HySIS satellite, and China's TanSat satellite, have been developed to monitor air pollution with varying resolutions and data details. There is a variety of information available on the issue of air pollution, which has been a persistent problem for a significant period. One solution that has been increasingly utilized is the implementation of machine learning (ML) techniques, which are particularly adept at processing and analyzing vast quantities of complex data [12] and [13]. Furthermore, several algorithms, such as Random Forest (RF), Support Vector Machine (SVM), and Classification and Regression Trees (CART), can be applied to develop models for air pollution concentration forecasting. Selecting an appropriate algorithm for each specific task is crucial, as it can increase the efficiency of the model and provide insight into patterns in the dataset.

In Northern Thailand, the level of $PM_{2.5}$ concentration exceeds the Thailand Air Quality Index (AQI) annually during the dry season (January to April) due to open biomass burning (OBB), which is the main cause of this problem. Forest fires are the primary source, releasing substantial air emissions and contributing significantly to air pollution in the upper north. This is exacerbated by geographical features such as steep mountains, which trap accumulated air pollutants at ground level. Over the past decade, forest fires have increasingly released significant $PM_{2.5}$ emissions. For instance, the difference between 2019 and 2013 was 32,226.7 tons, with releases in 2019 amounting to 90,424.7 tons/year, up from 58,198 tons/year in 2013 [14]. Additionally, crop waste residues from three key economic crops rice, corn, and sugarcane are significant sources of air pollution in the northern region, with high density in the lower north. Corn waste emissions are the highest in the agricultural sector due to the genetic characteristics of corn that allow it to grow with minimal water, leading to increased cultivation and exacerbating the issue. Unlike other types of OBB, sugarcane waste emissions are released in two phases: first, from burning crop waste in plantation areas, and second from agricultural processing, particularly in sugar factories that use biomass (bagasse) to generate power for manufacturing [15]. Additionally, meteorological characteristics significantly impact air pollution severity, particularly during high temperatures in the dry season [16] and [17].

Agricultural exports have led to a substantial increase in production for farmers due to swift economic growth in the sector. Regrettably, this has resulted in the incineration of agricultural regions, a preferred technique for its cost-effectiveness, which emits a significant quantity of air pollutants [3]. Meteorology uses mathematical models to measure air pollution levels. There are now two primary categories for models: mathematical and non-mathematical, each with distinct characteristics, especially in the data they utilize [18]. Mathematical models consider aspects like geography, air composition, and emissions from sources like transportation and industry. Non-mathematical models use previous statistical data on air quality, such as meteorological indicators and ground-level $PM_{2.5}$ concentration, to create non-linear relationships and forecast future air quality levels through ML techniques [13] and [18].

Therefore, this study aims to improve a model for estimating the concentration of $PM_{2.5}$ by integrating the principles of mathematical and non-mathematical models. Specifically, we included air emissions from OBB as an input variable, representing the air pollution from sources in the area. Other input variables were meteorological data and aerosol optical depth (AOD) from remote sensing data. The non-linear relationship between input and predicted variables is processed and cleaned to enhance data quality, preparing the dataset for model development using an artificial neural network algorithm. The performance of the best model is assessed based on the lowest mean absolute error (MAE) and root mean square error (RMSE). This model for estimating ground-level $PM_{2.5}$ concentration could be applied in areas with a similar context, particularly in regions lacking air monitoring stations.

2. Materials and Methods

2.1 Conceptual Framework

The modeling for estimating this study's ground level of $PM_{2.5}$ concentration will develop with the ML approach. The Northern Meteorological Center contributed data as maximum temperature (MAX), minimum temperature (MIN), relative humidity (RH), air pressure (AP), and wind speed (WS), including the ground-level $PM_{2.5}$ concentration

obtained from 18 air monitor stations that were contributed by the Pollution Control Department (PCD). The remote sensing data for this study is the AOD obtained from measurements of the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite, which will access the data through the Google Earth Engine (GEE) platform. Moreover, the OBB was also one of the input variables, in which the OBB represented the amount of air pollutants from the source of air pollution, and it was a challenge to determine as an input variable because of a lack of previous studies. The data mentioned above were determined as an input variable from generating the training and test datasets for creating models by learning from various situations that occurred in the past. Multilayer Perceptron Artificial Neural Networks (MLP-ANN) are algorithms with a working process that simulates the functioning of the human brain network, which will make them highly effective. The conceptual framework for this study is shown in Figure 1.

2.2 *Study Area*

The study was conducted in 17 provinces in Northern Thailand, classified into the upper north, including 9 provinces, and the lower north, consisting of 8 provinces, as shown in Figure 2. The northern part of Thailand is located from latitude 5.37′N to 20.27′N and longitude 97.22′E to 105.37′E, covering an approximate area of 169,644.3 km^2, with a population of 12,115,915 million people, or 18.24% of the total population in Thailand in 2021. Furthermore, the study areas displayed clear spatial differences, particularly in the upper north, where high mountains cover approximately 61,165.92 km^2, or 63.66% of the area, resulting in a tropical savanna climate. On the other hand, the lower north is mainly comprised of floodplains, with only forests covering approximately 28,736.82 km^2, or 37.85% of the total area. The dry season is plagued by air pollution that directly impacts the region. Specifically, the ground level of $PM_{2.5}$ concentration does not comply with the AQI standard and is a major environmental problem in Thailand. The OBB, especially forest fires and crop waste burning such as rice, corn, and sugarcane, is an important source of air pollutants in this area.

2.3. Data Collection

2.3.1 Meteorological data

Meteorological data, including maximum temperature (MAX, °C), minimum temperature (MIN, °C), air pressure (AP, hPa), wind speed (WS, m/s), and relative humidity (RH, %), were sourced from 17 meteorological stations located in each province, maintained by the Northern Meteorological Center. Concurrently, ground-level $PM_{2.5}$ concentrations ($\mu g/m^3$) were obtained from 18 air monitoring stations overseen by the PCD. The dataset spans three years (2019–2021), focusing on the dry season. Likewise, information for estimation was gathered in Nakhon Ratchasima, northeastern Thailand. The ground-level $PM_{2.5}$ concentration ($\mu g/m^3$) in these regions was obtained from 8 air monitoring ground stations maintained by the PCD. Furthermore, meteorological information for these regions was obtained from the Northeastern Meteorological Center, which has 20 stations, and the Upper North-Eastern Region Meteorological Center.

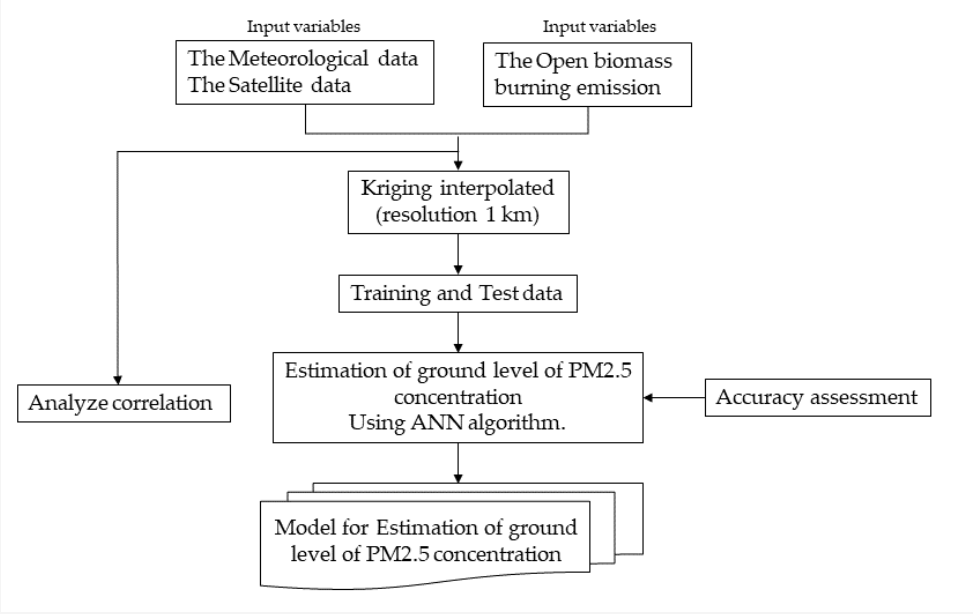

Figure 1: Conceptual framework

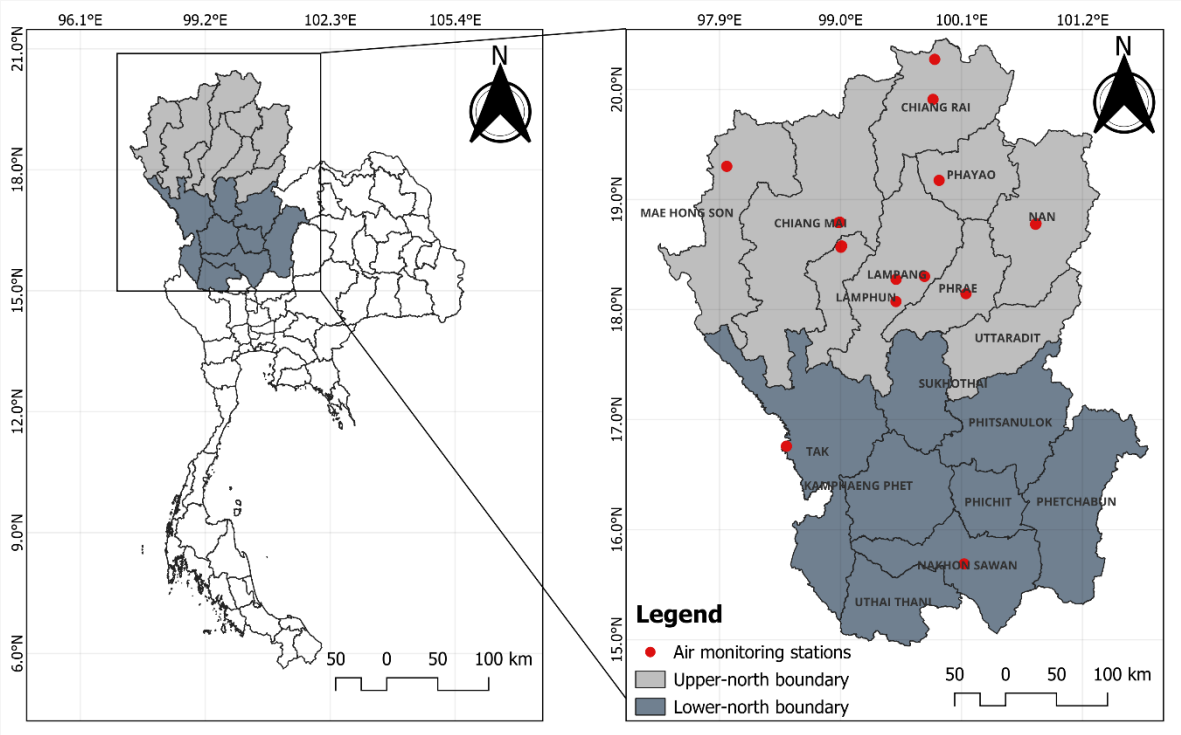

Figure 2: Location of the study area. Red dots illustrate the existing air quality monitoring stations maintained by the PCD

2.3.2 Satellite data

Satellite data utilized in these studies focuses on the AOD, a measure of the transmittance of solar radiation reaching Earth through haze and dust. These particles can absorb and reflect solar radiation, blocking a portion and allowing only a fraction to reach the Earth's surface [19]. The AOD gridded Level 2 product (MCD19A2 V6.1), which was generated daily at a spatial resolution of 1 km^2, was used in our study. This product was obtained from the MODIS Terra and Aqua combined Multi-Angle Implementation of Atmospheric Correction (MAIAC), which this satellite passes over twice daily in Thailand, around 10:30 a.m. and 1:30 p.m. The GEE platform was used to access the MCD19A2 V6.1 data. The data collection focused on the dry season for three years (2019-2021), specifically utilizing the "MODIS/061/MCD19A2_ GRANULES" collection and selecting the data band labeled “Optical_Depth_055”.

2.3.3 The Emission Inventory (EI).

The Emission Inventory (EI) in these studies represents the air emissions emitted from the OBB emission, including forest fire and crop waste residues, especially in rice, corn, and sugarcane. The EI was obtained from a monthly report, which is necessary to convert this monthly report into daily values. For conversion, follow the Atmospheric Brown Cloud (ABC) Emission Inventory Manual [20]. The daily emission can be determined from Equation 1.

$$ED_{n,d} = EM_n \times FD_d$$

Equation 1

Where:

n,d	= Month and day of air emission
$ED_{n,d}$	= Daily emission
EM_n	= Monthly emission
FD_d	= Daily fraction

It was important to note that the factors for daily emissions FD_d can be constructed by relating monthly activity data such as the number of fire hotspots in each land used and land cover.

2.4 Data Analysis and Data Integration

This section contains the data necessary for accurately estimating the ground level of $PM_{2.5}$ concentration in our study area. To achieve this, the daily average meteorological data such as MAX, MIN, RH, AP, WS, and AOD as input variables with the MLP-ANN algorithms. The EI, which focused on the OBB emission, was also used as an input variable, representing the amount of air pollutants emitted from the source of air pollution. This presented an outstanding challenge, as previous studies did not determine air emissions as an input variable. Moreover, all the above data, except AOD, will be

interpolated using the Kriging method, which was processed in a geographic information system program with a pixel resolution of 1 km^2.

2.5 Artificial Neural Networks (ANN) Structure

One of the techniques used in data mining is artificial neural networks (ANN). This mathematical model processes information through connectionist calculations to simulate the functioning of neural networks in the human brain [21] and [22]. The ANN algorithm comprises three fundamental components: the input, hidden, and output layers. Each neuron in one layer is interconnected with all neurons in the adjacent layer, with the connections determined by weights. Each neuron comprises two integral parts: the first involves summarizing all weighted inputs, called weight bias, and the second is the activation function, which translates the cumulative information into the output [23]. Equations 2 and 3 can represent the weight bias and activation function. Additionally, the number of nodes in the hidden layer is determined through the training process. Figure 3 shows this neural network configuration's intricate connections and layers.

$$u_i = \sum_{j=1}^{n} w_{ij} + x_j + b_i$$

Equation 2

Where:

x_j = the input signal
wi_j = the synaptic weights of the neuron k
u_i = transferred using a scalar to the activation function
$f(u_i)$ = the unit's activation

$$y_i = f(u_i)$$

Equation 3

This study employed the Multilayer Perceptron (MLP) neural network, which is just one of several types of ANN algorithms, to create a model that estimates the level of $PM_{2.5}$ concentration. The model used MLP-ANN algorithm processing with WEKA software, an open-source program developed by the University of Waikato, New Zealand.

2.5.1 Multilayer Perceptron (MLP) neural network structure

The MLP is a modern feedforward neural network algorithm that falls under the umbrella of the ANN algorithm. Its complex multi-layer structure enables it to process intricate data patterns that may not be linearly separable [24]. The MLP is classified as a supervised type of ANN, and backpropagation is used in the training process. The architecture of an MLP neural network involves specifying the number of neurons in each layer and the transfer functions utilized in those layers. Standard transfer functions include the log sigmoid function, a popular choice in ANNs, and the tan sigmoid function, as represented by Equations 4 and 5 [25].

$$f\left(v_i^m\right) = \frac{2}{1+\exp(-2v_i^m)} - 1$$

Equation 4

$$f\left(v_i^m\right) = \frac{2}{1+\exp(-\theta v_i^m)}$$

Equation 5

Design considerations, such as the number of neurons in each layer and the specific transfer functions employed, play a crucial role in determining the performance and capabilities of a network. These parameters are adjusted based on the task to optimize the network's ability to learn and generalize patterns from the data [25] and [26]. The MLP-ANN utilized in this study is shown in Figure 4.

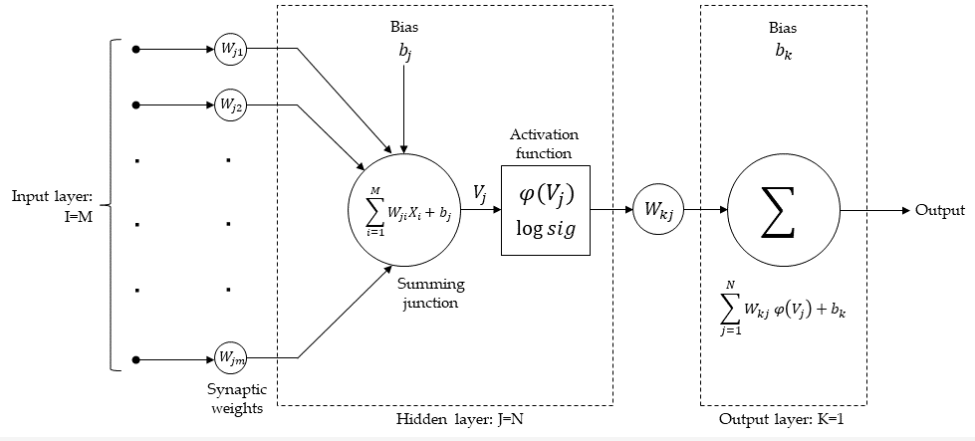

Figure 3: Schematic of the neural MLP computation

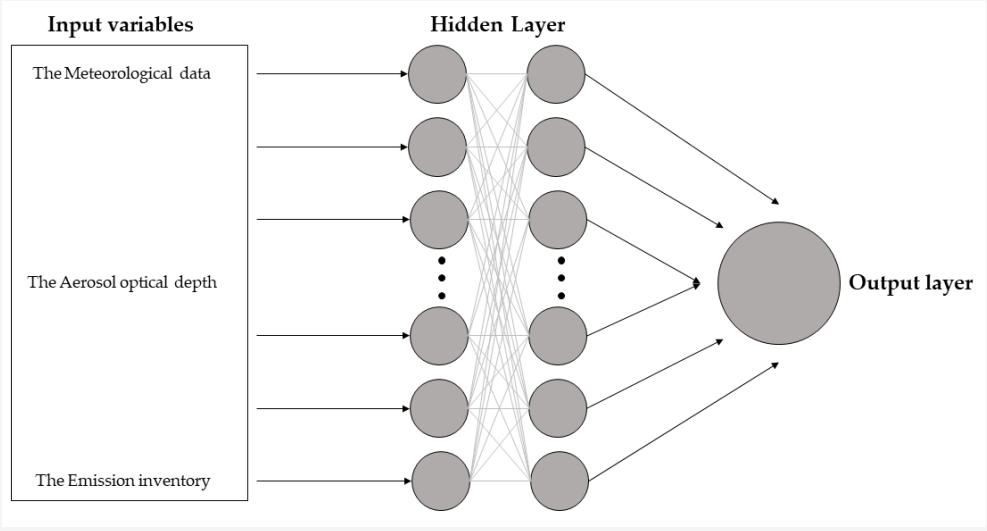

Figure 4: Schematic of the MLP-ANN in this study

2.6 Training and Testing the MLP-ANN

A 10-fold cross-validation method was used to subset the data for training and testing the MLP-ANN algorithm. The data were randomly divided into 10 parts; one part was used for testing, while the remaining parts were used as training data. The processes were repeated 10 times, and the results were average. The backpropagation algorithm was applied to train the feed-forward back propagation neural network.

2.7 Model Assessment

The model performance that indicated the lower statistical parameter; the highest correlation coefficient RMSE and MAE as defined in Equations 6 to 8 were used.

$$r = \frac{\sum\left(O_i - \bar{O}\right)(P_i - \bar{O})}{\sqrt{\sum (O_i - \bar{O})^2 (P_i - \bar{O})^2}}$$

Equation 6

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - P_i)^2}$$

Equation 7

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|O_i - P_i\right|$$

Equation 8

Where:

N = the number of times point
O_i = the observed value
P_i = the predicted value
$\bar{O}$ = the average of observed data

3. Results

3.1 The Characteristics of Input Variables

Table 1 shows descriptive statistics for the input variables obtained during dry seasons over three years (2019–2021), which were used for training and testing with the MLP-ANN algorithm. To ensure accuracy, we used a data cleansing process to remove any outliers or incomplete data from the dataset. After this process, we retained 13,393 data points, with 70% allocated to training data and 30% to test data. From Table 1, the average ground level of $PM_{2.5}$ concentration was 53.95 μg/m^3, ranging from 1.64 μg/m^3 to 310.10 μg/m^3. Meteorological variables revealed an average MAX of 35.0 °C, an average MIN of 20.61 °C, an RH of 62.73%, an AP of 1010.64 hPa, and a WS of 19.09 m/s. Over the study period, the average AOD was 0.48, which is considered a very hazy condition based on the classification range of the AOD from the Global Monitoring Laboratory [27].

The time series of the average ground-level $PM_{2.5}$ concentration and input variables such as AP, RH, AOD, and OBB are presented in Figure 5. Figure 5(d) shows the change in $PM_{2.5}$ concentration levels associated with OBB during the dry season in relation to variations in AOD values, which are detected from the MODIS Terra and Aqua satellites. The primary reason for this relationship is due to the AOD data obtained from sensors specifically developed to track air pollution problems. AOD values measure the transmittance of solar radiation through fog and dust to Earth. Dust, smoke, and pollution particles scattered in the atmosphere can block solar radiation by absorbing and reflecting it, allowing only some radiation to reach the Earth's surface

Table 1: Statistics of measured variables in Northern Thailand in dry seasons during 2019-2021

Variable	Mean	Min	Max	SD	Range
$PM_{2.5}$ concentration (μg/m^3)	53.95	1.64	310.10	29.93	308.46
Maximum temperature (°C)	35.00	13.00	43.18	4.15	30.18
Minimum temperature (°C)	20.61	7.60	31.17	3.70	23.57
Relative humidity (%)	62.73	24.44	95.02	9.20	70.58
Air pressure (hPa)	1010.64	1001.91	1022.00	3.14	20.09
Wind speed (m/s)	19.09	3.47	72.43	5.53	68.96
Open biomass burning emission (ton)	14.77	0.00	312.34	17.29	312.34
Aerosol Optical Depth (Unitless)	0.48	0.00	2.70	0.37	2.70

Note: Mean is the average of the data, Min and Max are the minimum and maximum of the data, SD is the standard deviation, and Range is the difference between the maximum and minimum values

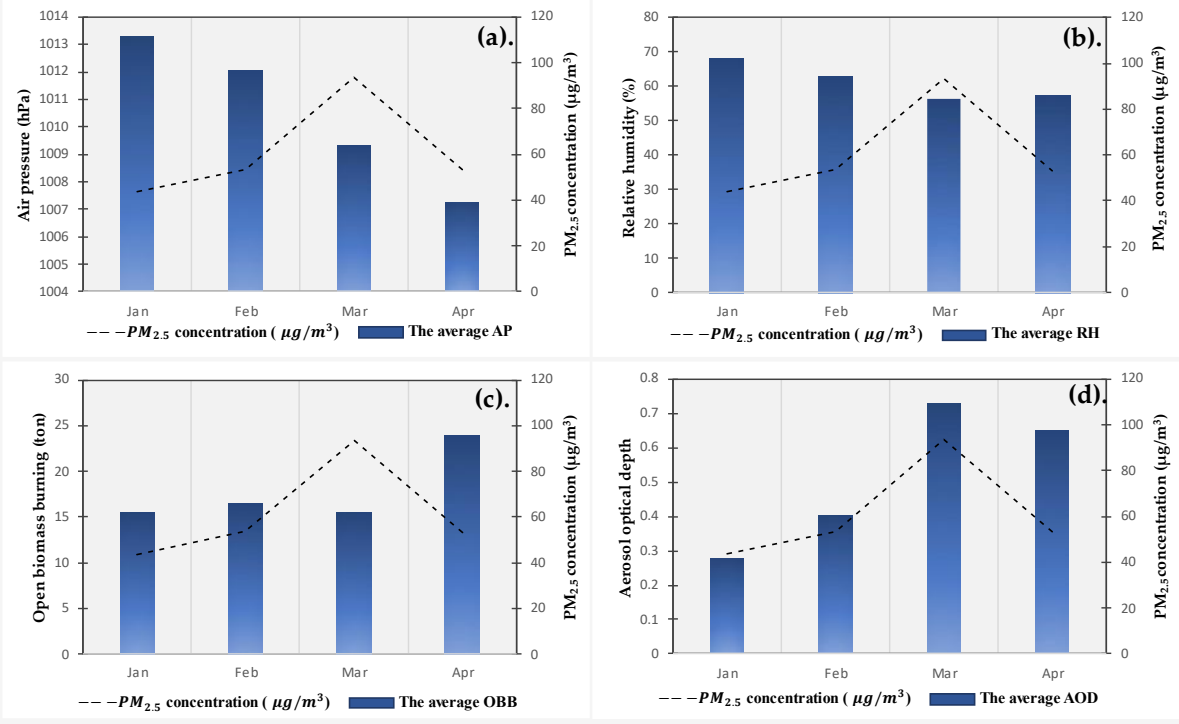

Figure 5: The $PM_{2.5}$ concentration and input variables during the dry season in 2019-2021 (a) AP (b) RH (c) OBB and (d) AOD

Furthermore, the time series of ground-level $PM_{2.5}$ concentration along with AP, RH, and OBB indicated inconsistent changes, as presented in Figure 5(a)-(c). Specifically, the time series of average AP and $PM_{2.5}$ concentration levels showed inconsistencies with the established relationship between AP and air pollution levels. This principle indicates that high AP (>1010.15 hPa) typically leads to increased air pollution, including higher levels of $PM_{2.5}$ concentration. High AP creates stable atmospheric conditions that trap pollutants near the ground, limiting their vertical dispersion. Additionally, high AP can cause temperature inversions, where a layer of warmer air sits above cooler air near the surface, preventing pollutants from rising and dispersing. As a result, pollutants accumulate close to the ground, leading to higher concentrations of air pollution. On the other hand, it was observed that AP values ranging between 1007.25 and 1009.59 hPa, characterized as low air pressure, corresponded to an increase in the level of $PM_{2.5}$ concentration. This difference may be explained by the fact that AP values exceeding 1010.15 hPa were observed annually in January and February, a period that does not coincide with the peak of OBB, the primary source of air pollution in this area.

Additionally, the time series of the average ground-level $PM_{2.5}$ concentration, RH, and the OBB revealed inconsistencies in trend changes, as presented in Figure 5(b)-(c). These inconsistencies were influenced by multiple factors, particularly the La Niña phenomenon and the impact of the COVID-19 pandemic. Among these factors, RH was the most significantly affected. The La Niña phenomenon, which occurred during 2020–2021, had a notable impact on RH. This climate pattern is characterized by cooler-than-average sea surface temperatures in the central and eastern tropical Pacific Ocean, which often leads to increased rainfall in various regions [28] and [29]. Specifically, during the La Niña period, there was above-average rainfall, which in turn increased RH levels. The heightened RH reduced the potential for biomass combustion, as the higher moisture content in the air makes it more difficult for fires to ignite and sustain. This natural suppression of fire activity contributed to variations in $PM_{2.5}$ concentrations observed during this period. Moreover, the COVID-19 pandemic further complicated the OBB trends. The pandemic decreased agricultural product demand due to various factors, including reduced consumer spending and disruptions in the global supply chain. Additionally, epidemic control measures, such as lockdowns and restrictions on movement, significantly disrupted agricultural exports. These disruptions caused fluctuations in the OBB, as agricultural practices and transportation emissions were directly affected by the pandemic's economic and social impacts [30][31] [32] and [33].

3.2 Correlation between the Input Variables and the Ground Level of $PM_{2.5}$ Concentrations

Table 2 shows a correlation matrix that illustrates the relationship between input variables and the ground level of $PM_{2.5}$ concentrations. The analysis reveals significant correlations at the 0.001 level (2-tailed). Positive correlations were found between the ground level of $PM_{2.5}$ concentration and MAX (r = 0.299), AOD (r = 0.530), and the OBB (r = 0.186), indicating that an increase in these factors can directly contribute to the rise in the ground level of $PM_{2.5}$ concentration. Conversely, negative correlations were identified between RH (r = -0.429) and AP (r = -0.317) and $PM_{2.5}$ concentration, indicating that a decrease in RH and AP leads to an increase in the ground level of $PM_{2.5}$ concentration. The significance of the RH variable was in demonstrating the humidity level in the surrounding air. It directly impacts the potential for combustion in the area; when the air is dry due to low humidity, it also affects the moisture of fuel, leading to the release of pollutants. Additionally, the study found that the AP characteristics in the area are opposite to the relationship between the AP and the concentration of air pollutants, which means that a means that a high AP, or AP greater than 1010.15 hPa, is associated with higher pollutant levels because it can cause a temperature inversion, directly affecting air pollution accumulation at ground level [34].

3.3 The Model for Estimation of the Concentration of $PM_{2.5}$

For the model training, the total dataset of 13,393 records was split into training data (9,375 records) and test data (4,018 records), and an MLP-ANN was employed to estimate the ground level of $PM_{2.5}$ concentrations. The best modeling was to identify the optimal architecture network with the lowest value of MAE and RMSE and a high correlation coefficient. The results are displayed in Table 3, which reveals that the optimal architecture is 8-16-1, which was the best modeling.

Table 2: The $PM_{2.5}$ correlation with meteorological factors, AOD, and The OBB variables5

	MIN	**MAX**	**RH**	**AP**	**WS**	**AOD**	**OBB**	**$PM_{2.5}$**
MIN	**1**	0.635***	-0.468***	-0.639***	0.517***	0.369***	0.049***	0.059***
MAX	0.635***	**1**	-0.532***	-0.621***	0.357***	0.367***	0.098***	0.299***
RH	-0.468***	-0.532***	**1**	0.619***	-0.445***	-0.305***	-0.251***	-0.429***
AP	-0.639***	-0.621***	0.619***	**1**	-0.473***	-0.345***	-0.155***	-0.317***
WS	0.517***	0.357***	-0.445***	-0.473***	**1**	0.168***	0.050***	0.010
AOD	0.369***	0.367***	-0.305***	-0.345***	0.168***	**1**	0.070***	0.530***
OBB	0.049***	0.098***	-0.251***	-0.155***	0.050***	0.070***	**1**	0.186***
$PM_{2.5}$	**0.059*** **	**0.299*** **	**-0.429*** **	**-0.317*** **	**0.010**	**0.530*** **	**0.186*** **	**1**

Note:
* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).
*** Correlation is significant at the 0.001 level (2-tailed).

Table 3: The model performance with different hidden layers

Model	r	MAE	RMSE
8-1-a	0.7142	0.0217	0.0323
8-2-1	0.6247	0.0261	0.0365
8-3-1	0.6737	0.0236	0.034
…	…	…	…
8-15-1	0.7364	0.0214	0.0316
8-16-1	**0.7787**	**0.0211**	**0.0313**
8-17-1	0.7355	0.0214	0.0315

Table 4: MAE, RMSE of Learning rate (LR) and momentum

Learn rate: LR	Momentum	MAE	RMSE
0.1	**0.3**	**0.0187**	**0.0282**
0.1	0.5	0.0190	0.0289
0.1	0.7	0.0193	0.0292
0.3	0.3	0.0221	0.0321
0.3	0.5	0.0240	0.0339
0.3	0.7	0.0291	0.0391
0.5	0.3	0.0242	0.0341
0.5	0.5	0.0256	0.0355
0.5	0.7	0.0287	0.0390

This architecture comprises 8 neurons in the input layer, 16 in the hidden layer, and 1 in the output layer. The lowest values of MAE and RMSE were 0.0211 and 0.0313, respectively. Moreover, the correlation coefficient was found to be 0.7787, which indicates a strong connection between the ground level of $PM_{2.5}$ concentration between the model result and the data observed by PCD. To enhance the model performance, forget the lower values of MAE and RMSE, the network architecture's learning rate (LR) and momentum were tested. The best values for LR and momentum were 0.1 and 0.3, respectively, yielding the model's lower MAE and RMSE values (Table 4). The model result was also used as the test data to validate the model performance, as shown in Figure 6. The test data, composed of 4,018 records from March to April 2020 and January to April 2021, indicated a discontinuity in Figure 6 of the data cleaning process. The time series plot in Figure 6 compares the ground level of $PM_{2.5}$ between the model's result and the actual data from the PCD. Figure 6 shows that the model underestimates the ground-level $PM_{2.5}$ concentrations compared to the actual data from the PCD. Specifically, there are two distinct ranges where the model's estimates significantly deviate from the observed values. Despite this tendency to underestimate, the model's overall performance still reflects the general trends and changes in $PM_{2.5}$ concentrations, aligning with the actual data to a considerable extent. The primary reasons for the model's underestimation are related to the limitations of the training dataset. These limitations arose because the training data covered only specific periods, which constrained the model's ability to learn from a broader range of scenarios. External factors, particularly the La Niña phenomenon, further complicated the data. The La Niña phenomenon led to unmeasurable variations in certain data types, resulting in atypical weather patterns, such as increased rainfall and higher relative humidity. These conditions affected the $PM_{2.5}$ concentrations and introduced complexities the model was not adequately trained to handle.

Moreover, this modeling was also used to estimate the ground level of $PM_{2.5}$ concentrations during dry seasons in 2021, which focused on Khon Kaen, Nakhon Ratchasima, and Ubon Ratchathani provinces. These provinces are located in the northeastern part of Thailand, where air pollution problems are exacerbated by agricultural waste burning, such as rice, corn, and sugarcane. The results demonstrated a close alignment between the model's estimates and the data reported by the PCD, as shown in Figure 7. While the model results were generally consistent with the actual data, they tended to be slightly underestimated.

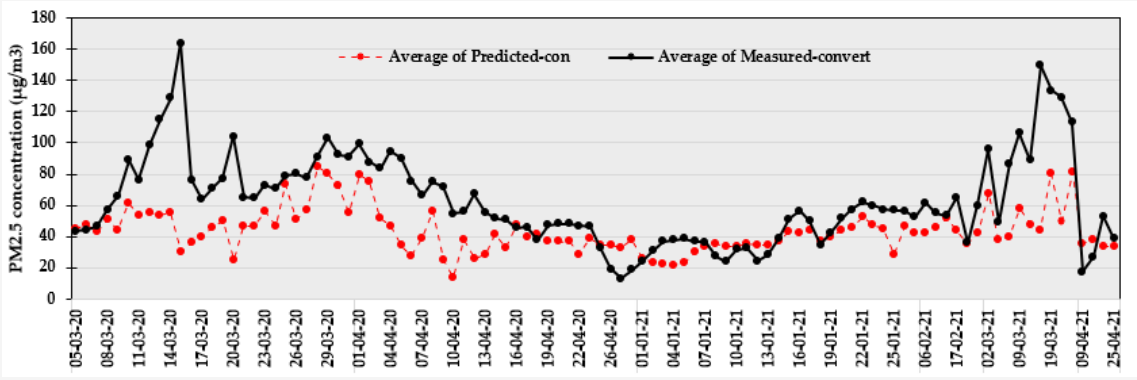

Figure 6: Comparison level of $PM_{2.5}$ concentration between the model's estimates with the actual data reported by the PCD

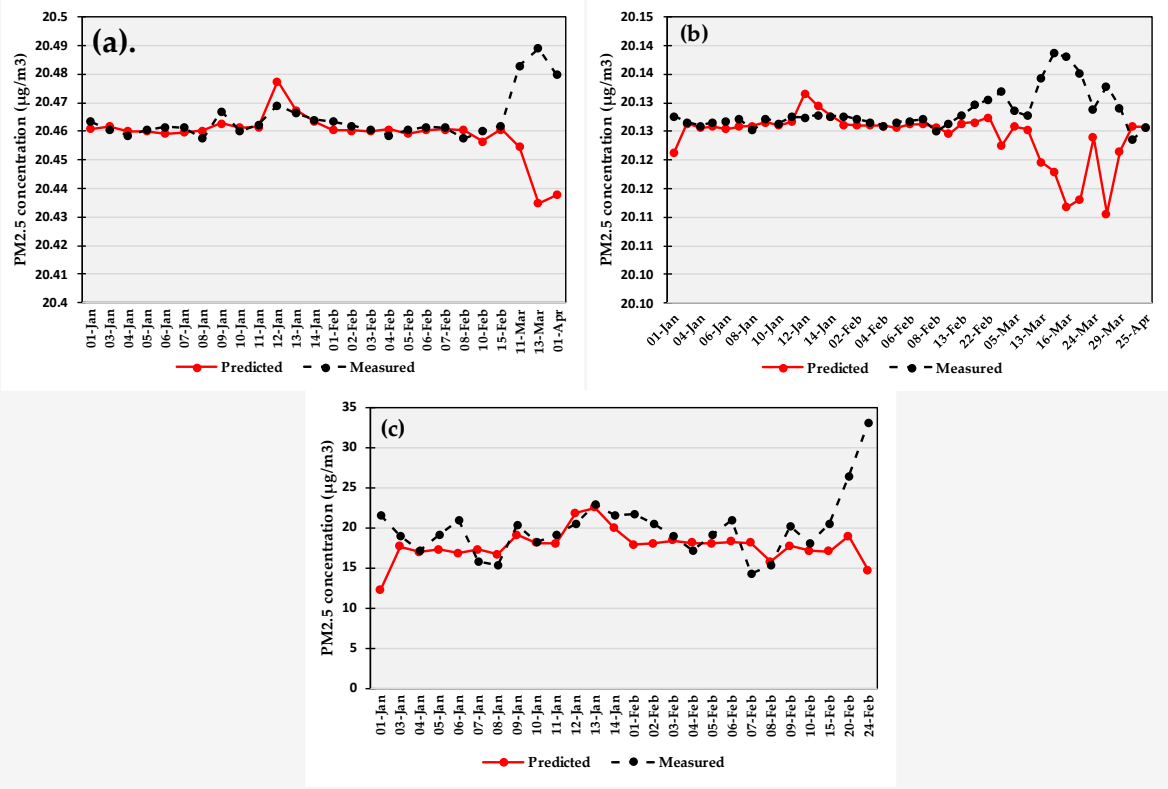

Figure 7: Map of the spatial distribution of average concentration of $PM_{2.5}$ from estimation during January – April 2021 (a) in Khon Kaen (b) Nakhon Ratchasima and (c) Ubon Ratchathani

This underestimation was particularly notable during March and April, where significant gaps between the estimated and actual data were observed, similar to the discrepancies identified in the test dataset discussed previously. Despite the tendency of the model to underestimate, the differences were generally minor. For instance, in Figure 7(a), the estimated $PM_{2.5}$ concentration in Khon Kaen was 20.43 µg/m³, only marginally lower than the measured value of 20.48 µg/m³. Similarly, in Nakhon Ratchasima on March 15, 2018, the model estimated a concentration of 20.12 µg/m³, just 0.02 µg/m³ lower than the measured value of 20.14 µg/m³. However, Figure 7(c) highlights a more significant discrepancy in February, where the model's estimates differed markedly from the actual data. The primary reasons for these differences can be attributed to errors inherent in the model. Additionally, geographical characteristics and sources of air pollution in the regions studied contributed to these discrepancies. The provinces used for estimation, such as Khon Kaen and Nakhon Ratchasima, are characterized by plateau areas where agricultural residue burning is a primary source of pollution.

In contrast, the model was developed using data from the northern region, surrounded by steep mountains where forest fires are the main sources of air pollution. This discrepancy in the types of pollution sources and geographical features between the regions influenced the accuracy of the model's estimates.

4. Discussion

4.1 The Relationship between the Input Variables on the Ground Level of $PM_{2.5}$ Concentrations

In this section, we focused on examining the relationship between various input variables and ground-level $PM_{2.5}$ concentration, particularly noting that our analysis's correlation coefficient (r-value) was not particularly strong. First, we found an inverse relationship between RH and ground-level $PM_{2.5}$ concentrations, with an r-value of -0.429. These findings are consistent with previous studies that focused on meteorological factors influencing PM_{10} concentrations in Saraburi, Thailand, which reported an r-value of -0.581 [35], as well as studies in Nepal that indicated a similar inverse relationship between RH and the level of $PM_{2.5}$ concentrations [36].

Additionally, the MAX demonstrated a positive relationship with the level of $PM_{2.5}$ concentrations in our analysis. This aligns with prior studies consistently showing such positive relationships [37]. When MAX increases, it enhances combustion potential, leading to higher pollutant emissions and an increased level of $PM_{2.5}$ concentrations [38]. Conversely, our analysis revealed a very low r-value for the MIN, indicating a weak relationship. This is similar to findings in a study on meteorological influences on PM_{10} concentrations in Chiang Mai, Thailand, which reported an r-value of 0.059 [16]. Despite the low r-value, previous research has demonstrated that such factors can still significantly impact air pollution levels. For instance, principal component analysis (PCA) has been used in earlier studies to improve the clarity of results when r-values are weak. Even with low r-values, PCA can reveal a clear relationship between meteorological factors and increases in air pollution [16].

Moreover, the findings revealed an intriguing relationship, particularly between the AP and ground-level $PM_{2.5}$ concentration, in which a negative correlation was identified. When the AP decreases, the ground-level $PM_{2.5}$ concentration tends to increase. This finding contradicts the established principle that high AP corresponds to increased ground-level $PM_{2.5}$ concentrations. High AP can inhibit air buoyancy, accumulating air pollution at the ground level as the air cannot circulate freely [34] and [35]. This finding was influenced by periods of highest $PM_{2.5}$ concentrations coinciding with low-pressure conditions (1007.25–1009.59 hPa), consistent with a study on meteorological influences on PM_{10} in Chiang Mai [16], which also reported a negative relationship and indicated the reason was due to a period of peak air pollution problems not consistent with high AP covering the area.

However, our results differ from studies in Bangkok and Saraburi, central Thailand, where a positive correlation between AP and air pollution was found [36] and [37]. These discrepancies underscore the significant impact of local geographic characteristics and sources of air pollution. Additionally, the AOD showed a significant positive correlation with $PM_{2.5}$ concentration ($r = 0.530$) in our analysis. AOD measures the extent to which particulate matter in the atmosphere absorbs or scatters sunlight, with higher $PM_{2.5}$ concentrations leading to higher AOD values. Similar studies in Bangkok, central Thailand, and northern Thailand have integrated AOD as a predictive variable for $PM_{2.5}$ concentrations, consistently finding a positive correlation that enhances predictive model accuracy [38] and [39].

4.2 *The Accuracy of Model Performance*

The next part was focused on the model developed, in which the MLP-ANN was used to estimate the ground level of $PM_{2.5}$ concentration in these studies. The input variables include the OBB, AOD, and meteorological data like MAX, MIN, RH, AP, and WS. The optimal network architecture with 8-16-1 had the lowest MAE and RMSE values, 0.0187 and 0.0282, respectively. These findings are comparable to a study that compared four algorithms for estimating $PM_{2.5}$ concentrations in Zhangdian District, China, with ANN being one of the algorithms [22]. The input variables in that study included meteorological data and industrial waste gas emissions. The results showed that the ANN algorithm demonstrated high model performance, with the lowest MAE and RMSE values being 17.53 and 26.72, respectively. The differences in geographic variations, weather circumstances, and sources of air pollution were the primary reasons for the higher MAE and RMSE values compared with our study. Especially geographically, the Zhangdian district is mostly in a plain with 72.43% of the total area, where the ceramic sector is a significant source of pollution. In contrast, our study has steep mountains, and the OBB, especially forest fire, was the primary source of air pollution. The higher MAE and RMSE values compared to our study were also found in a study that estimated $PM_{2.5}$ concentrations in Liaocheng, China [13]. This study used meteorological data as input variables and processed them using the ANN algorithm, with the lowest MAE

and RMSE values being 4.6 and 6.6, respectively. The primary reason for the difference is the variation in sources of air pollution; in this study, motor vehicles and industrial sectors were the main sources of air pollution. Turning to studies with similar geographic characteristics and sources of air pollution, we found that the MAE and RMSE values were also higher than those in our study. Firstly, a study in Udon Thani, Thailand, compared four algorithms for estimating $PM_{2.5}$ concentrations, of which the ANN algorithm was one [40]. This study reported the lowest RMSE value of 2.945. The primary reasons for the difference are the geographic characteristics and sources of pollution, particularly agricultural waste residues, which are the main sources of air pollution in these areas. Likewise, the studies focused on developing methods for estimating the concentration of $PM_{2.5}$ in Nan, Thailand [41]. The results from this study showed an MAE value of 3.50, which is still higher than in our study because of the difference in input variables, which were determined only by meteorological data as input variables. Besides adding input variables, utilizing various algorithmic approaches is another factor that enhances model performance efficiency.

Based on the previous studies mentioned above, it was found that differences in geographical characteristics, meteorological factors, and sources of air pollution were key variables that led to varying study results. Moreover, most previous studies focused on determining only meteorological data as input variables, significantly impacting model performance. In particular, studies that incorporate in-depth data, especially data that also affects the increase in air pollutants, show increased model performance efficiency. This finding aligns with our study, which aims to develop a model by incorporating in-depth data, namely the OBB emissions. Our results show that incorporating OBB emissions as input variables leads to the lowest values of MAE and RMSE, which are coefficients indicating the effectiveness of model performance.

5. Conclusions

This study aimed to estimate the level of $PM_{2.5}$ concentration using MLP-ANN algorithms. A significant aspect of the model development was the integration of OBB emissions as an input variable, alongside meteorological data and AOD. This approach presented a notable challenge due to the limited attention given to OBB emissions in previous studies. The results indicate that the optimal network architecture, 8-16-1, achieved the lowest MAE and RMSE values of 0.0187 and 0.0282, respectively. The model's accuracy was validated using testing data closely aligned with data from the PCD. However, the model tended to underestimate the level of $PM_{2.5}$ concentrations, primarily due to limitations in the training datasets. Additionally, the model was successfully applied in three provinces: Khon Kaen, Nakhon Ratchasima, and Ubon Ratchathani. Although the estimates generally slightly underestimated actual PCD data, the differences were minimal. The main reasons for underestimation include discrepancies in model development and variations in geographic characteristics and sources of air pollution. Nevertheless, our findings underscore the importance of integrating OBB emissions, meteorological data, and AOD to enhance model performance. Moreover, future efforts should prioritize in-depth analyses of air pollution, particularly focusing on ultrafine particles, to improve our understanding of air pollution dynamics. These insights are critical for guiding strategies to mitigate air pollution and safeguard public health and the environment.

Acknowledgements

The author would like to thank the Northern Meteorological Center (CMMET) and the Pollution Control Department (PCD) in Thailand for providing the data essential for achieving the purpose of this study. Additionally, we would like to extend our gratitude to the Land Development Department (LDD) in Thailand for supplying crucial information on land use and land cover in each area, which was instrumental in understanding the spatial relationships. Furthermore, the authors wish to thank Mr. Tammanoon Iemsupap for his general support in sampling and spatial data analysis.

References

[1] Inerb, M., Phairuang, W., Paluang, P., Hata, M., Furuuchi, M. and Wangpakapattanawong, P., (2022). Carbon and Trace Element Compositions of Total Suspended Particles (TSP) and Nanoparticles ($PM_{0.1}$) in Ambient Air of Southern Thailand and Characterization of their Sources. *Atmosphere*, Vol. 13(4). https://doi.org/10.3390/atmos13040626.

[2] Ahmad, M., Chen, J., Panyametheekul, S., Yu, Q., Nawab, A., Khan, M. T., Zhang, Y., Ali, S. W. and Phairuang, W., (2024). Fine Particulate Matter from Brick Kilns Site and Roadside in Lahore, Pakistan: Insight into Chemical Composition, Oxidative Potential, and Health Risk Assessment. *Heliyon*, Vol. 10(4). https://doi.org/10.1016/j.heliyon.2024.e25884.

[3] Phairuang, W., Amin, M., Hata, M. and Furuuchi, M., (2022). Airborne Nanoparticles ($PM_{0.1}$) in Southeast Asian Cities: A Review. *Sustainability*, Vol. 14(16). https://doi.org/10.3390/su141610074.

[4] Tial, M. K. S., Suriyawong, P., Chetiyanukornkul, T., Paluang P., Amin M., Putri, R. M., Hata, M., Furuuchi, M. and Phairuang, W., (2024). Characterization of $PM_{0.1}$ Mass Concentrations and Elemental and Organic Carbon in Upper Southeast Asia. *Atmospheric Pollution Research*, Vol. 15(8). https://doi.org/10.1016/j.apr.2024.102157.

[5] Wang, T. H., Huang, K. Y., Chen, C. C., Chang, Y. H., Chen, H. Y., Hsueh, C., Liu, Y. T., Yang, S. C., Yang, P. C. and Chen, C. Y., (2023). $PM_{2.5}$ Promotes Lung Cancer Progression through Activation of the AhR-TMPRSS2-IL18 Pathway. *EMBO Molecular Medicine*, Vo. 15(6). https://doi.org/10.15252/emmm.202217014.

[6] Xiong, R., Jiang, W., Li, N., Liu, B., He, R., Wang, B. and Geng, Q., (2021). $PM_{2.5}$-Induced Lung Injury is Attenuated in Macrophage-Specific NLRP3 Deficient Mice. *Ecotoxicology and Environmental Safety*, Vol. 221. https://doi.org/10.1016/j.ecoenv.2021.112433.

[7] Chang, D. and Song, Y., (2010). Estimates of Biomass Burning Emissions in Tropical Asia Based on Satellite-Derived Data. *Atmospheric Chemistry and Physics*, Vol. 10(5), 2335–2351. https://doi.org/10.5194/acp-10-2335-2010.

[8] Shi, Y., Matsunaga, T. and Yamaguchi, Y., (2015). High-Resolution Mapping of Biomass Burning Emissions in Three Tropical Regions. *Environmental Science & Technology*, Vol. 49(18), 10806-10814. https://doi.org/10.1021/acs.est.5b01598.

[9] Claverie, M., Masek, J. G., Ju, J. C. and Dungan, J. L., (20147). *Harmonized Landsat-8 Sentinel-2 (HLS) Product User's Guide*. National Aeronautics and Space Administration (NASA), Washington, DC, USA, 1-17. https://hls.gsfc.nasa.gov/wp-content/uploads/2017/08/HLS.v1.3.UserGuidev2.pdf.

[10] Wulder, M. A., Markham, B., McCorkel, J., Crawford, C. J., James Storey, J. and Del T. Jenstrom, D. T., (2020). Landsat 9: Empowering Open Science and Applications through Continuity. *Remote Sensing of Environment*, Vol. 248. https://doi.org/10.1016/j.rse.2020.111968.

[11] Pagano, T. S. and Durham, R. M., (1993). Moderate Resolution Imaging Spectroradio meter (MODIS). Sensor Systems for the Early Earth Observing System Platforms, Proc. SPIE 1939, Bellingham, WA, USA, https://doi.org/10.1117/12.152835.

[12] Mao, X., Shen, T. and Feng, X., (2017). Prediction of Hourly Ground-Level $PM_{2.5}$ Concentrations 3 Days in Advance using Neural Networks with Satellite Data in Eastern China. *Atmospheric Pollution Research*, Vol. 8(6), 1005-1015. https://doi.org/10.1016/j.apr.2017.04.002.

[13] He, Z., Guo, Q., Wang, Z. and Li, X.., (2022). Prediction of Monthly PM2.5 Concentration in Liaocheng in China Employing Artificial Neural Network. *Atmosphere*, Vol. 13(8). https://doi.org/10.3390/atmos13081221.

[14] Paluang, P., Thavorntam, W. and Phairuang, W., (2024). The Spatial–Temporal Emission of Air Pollutants from Biomass Burning during Haze Episodes in Northern Thailand. *Fire*, Vol. 7(4). https://doi.org/10.3390/fire7040122.

[15] Phairuang, W., (2021). *Biomass Burning and their Impacts on Air Quality in Thailand*. Biomass Burning in South and Southeast Asia: Impacts on the Biosphere. CRC Press. 1-21.

[16] Chotamonsak, C. and Lapyai, D., (2018). Meteorological Factors Related to Air Pollution in Chiang Mai Province. *Journal of Research Unit on Science, Technology and Environment for Learning*, Vol. 9, 237-249.

[17] Kliengchuay, W., Worakhunpiset, S., Limpanont, Y., Meeyai, A. C. and Tantrakarnapa, K., (2021). Influence of the Meteorological Conditions and some Pollutants on PM10 Concentrations in Lamphun, Thailand. *Journal of Environmental Health Science and Engineering*, Vol. 19(1), 237-249. https://doi.org/10.1007/s40201-020-00598-2.

[18] Guo, Q., He, Z. and Wang, Z., (2023). Predicting of Daily PM(2.5) Concentration Employing Wavelet Artificial Neural Networks Based on Meteorological Elements in Shanghai, China. *Toxics*, Vol. 11(1). https://doi.org/10.3390/toxics11010051.

[19] Putham, K., Wetchayon, P. and Choosakul, N., (2020). A Comparison of Methods for Estimating Fine Particulate Matter Concentrations from Himawari-8 Satellite Over Northern Thailand. *Research on Modern Science and Utilizing Technological Innovation Journal*, Vol. 14, 55-67.

[20] Shrestha, R. M., Kim Oanh, N. T., Shrestha, R., Rupakheti, M., Rajbhandari, S., Agustian, D., Kanabkaew, T. and Iyngararasan, M., (2013), *Atmospheric Brown Clouds: Emission Inventory Manual*. Nairobi: United Nations Environment Programme.

[21] Auwatanamongkoll, S., (2014). *Data Mining*. Classification. 2014, Bangkok, Thailand. National Institute of Development Administration. 31-64.

[22] Li, R., Zhou, R. and Zhang, J., (2018). Function of PM2.5 in the Pathogenesis of Lung Cancer and Chronic Airway Inflammatory Diseases. *Oncol Lett*, Vol. 15(5), 7506-7514. https://doi.org/10.3892%2Fol.2018.8355.

[23] Hambli, R., (2011). Multiscale Prediction of Crack Density and Crack Length Accumulation in Trabecular Bone Based on Neural Networks and Finite Element Simulation. *International Journal for Numerical Methods in Biomedical Engineering*, Vol. 27(4), 461-475. https://doi.org/10.1002/cnm.1413.

[24] Asadollahfardi, G., Zangooei, H. and Aria, S. H., (2016). Neural Networks and Markov Chain, a Case Study Karaj City. *Asian Journal of Atmospheric Environment*, Vol. 10(2), 67-79. https://doi.org/10.5572/ajae.2016.10.2.067.

[25] Rogers, T. T. and McClelland, J. L., (2014). Parallel Distributed Processing at 25: Further Explorations in the Microstructure of Cognition. *Cognitive Science*, Vol. 38(6), 1024-1077. https://doi.org/10.1111/cogs.12148.

[26] Goudarzi, G., Hopke, P. K. and Yazdani, M., (2021). Forecasting PM(2.5) Concentration Using Artificial Neural Network and its Health Effects in Ahvaz, Iran. *Chemosphere*, Vol. 283. https://doi.org/10.1016/j.chemosphere.2021.131285.

[27] Global Radiation and Aerosols, (2023). *Laboratory Global Monitoring*. Available: https://gml.noaa.gov/grad/surfrad/aod. [Accessed Feb.23, 2023].

[28] Duc, H. N., Bang, H. Q., Quan, N. H., Quang, N. X., (2019). Impact of Biomass Burnings in Southeast Asia on Air Quality and Pollutant Transport during the End of the 2019 Dry Season. *Environmental Monitoring and Assessment*, Vol. 193(9). https://doi.org/10.1007/s10661-021-09259-9.

[29] Kraisitnitikul, P., Thepnuan, D., Chansuebsri, S., Yabueng, N., Wiriya, W., Saksakulkrai, S., Shi, Z. and Chantara, S., (2024). Contrasting Compositions of PM(2.5) in Northern Thailand during La Nina (2017) and El Nino (2019) Years. *J Environ Sci (China)*, Vol. 135, 585-599. https://doi.org/10.1016/j.jes.2022.09.026.

[30] Sapbamrer, R., Chittrakul, J., Sirikul, W., Kitro, A., Chaiut, W., Panya, P., Amput, P., Chaipin, E., Sutalangka, C., Sidthilaw, S., Promrak, P., Pailinrak Kamolsan, P. and Hongsibsong, S., (2022). Impact of COVID-19 Pandemic on Daily Lives, Agricultural Working Lives, and Mental Health of Farmers in Northern Thailand. Sustainability, Vol. 14(3). https://doi.org/10.3390/su14031189.

[31] Sinha, S. and Swain, M., (2022). Response and Resilience of Agricultural Value Chain to COVID-19 Pandemic in India and Thailand. *Pandemic Risk, Response, and Resilience*, 363-381. https://doi.org/10.1016/B978-0-323-99277-0.00002-4.

[32] Tansuchat, R., Suriyankietkaew, S., Petison, P., Punjaisri, K. and Nimsai, S., (2022). Impacts of COVID-19 on Sustainable Agriculture Value Chain Development in Thailand and ASEAN. *Sustainability*, Vol. 14(20). https://doi.org/10.3390/su142012985.

[33] Thammachote, P. and Trochim, J. I., (2021). *The Impact of the COVID-19 Pandemic on Thailand's Agricultural Export Flows*. Michigan State University, Department of Agricultural, Food and Resource Economics, Food Security Group. PRCI Research Paper.

[34] Whiteman, C. D., Pospichal, B., Eisenbach, S., Weihs, P., Clements, C. B., Steinacker, R., Mursch-Radlgruber, E. and Dorninger, M., (2004). Inversion Breakup in Small Rocky Mountain and Alpine Basins. *Journal of Applied Meteorology*, Vol. 43(8), 1069-1082. https://doi.org/10.1175/1520-0450(2004)043<1069:IBISRM>2.0.CO;2.

[35] Kwanma, P., Pukngam, S. and Arunpraparu, W., (2019). Meteorological Factors Affecting Concentration of Pm10 at Na Phra Lan Sub - District, Chaloem Phra Kiat District, Saraburi Province. *PSRU Journal of Science and Technology*, Vol. 4(2), 85-94. https://ph01.tci-thaijo.org/index.php/Scipsru/article/view/180950.

[36] Aryal, R., Lee, B., Karki, R., Gurung, A., Kandasamy, J., Pathak, B. K., Sharma, S. and Giri, N., (2008). Seasonal PM10 Dynamics in Kathmandu Valley. *Atmospheric Environment*, Vol. 42(37), 8623-8633. https://doi.org/10.1016/j.atmosenv.2008.08.016.

[37] Galindo, N., Varea, M., Gil-Molto, J., Yubero, E. and Nicolás, J. F., (2011). The Influence of Meteorology on Particulate Matter Concentrations at an Urban Mediterranean Location. *Water, Air, & Soil Pollution*, Vol. 215, 365-372. https://doi.org/10.1007/s11270-010-0484-z.

[38] Schoennagel, T., Veblen, T. T. and Romme, W. H., (2004). The Interaction of Fire, Fuels, and Climate across Rocky Mountain Forests. *BioScience*, Vol. 54(7), 661-676. https://doi.org/10.1641/0006-3568(2004)054[0661:TIOFFA]2.0.CO;2.

[39] Aguado, E. and Burt, J., (2015). *Energy and Mass, in Understanding Weather and Climate*. England: Pearson Education Limited. 28-145.

[40] Cheewinsiriwat, P., Duangyiwa, C., Sukitpaneenit, M. and Stettler, M. E. J., (2022). Influence of Land Use and Meteorological Factors on PM2.5 and PM10 Concentrations in Bangkok, Thailand. *Sustainability*, Vol. 14(9). https://doi.org/10.3390/su14095367.

[41] Peng-In, B., Sanitluea, P., Monjatturat, P., Boonkerd, P. and Phosri, A., (2022). Estimating Ground-Level PM(2.5) Over Bangkok Metropolitan Region in Thailand using Aerosol Optical Depth Retrieved by MODIS. *Air Qual Atmos Health*, Vol. 15(11), 2091-2102. https://doi.org/10.1007/s11869-022-01238-4.