science, engineering and health studies

Research Article

Science, Engineering and Health Studies
https://li01.tci-thaijo.org/index.php/sehs
ISSN (Online): 2630-0087

Predictive models of $PM_{2.5}$ concentration with aerosol optical depth and meteorological data in Bangkok area using machine learning techniques

Ikwan Bensalam, Salang Musikasuwan, and Rattikan Saelim*

Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Pattani Campus, Pattani 94000, Thailand

***Corresponding author**:*
Salang Musikasuwan
salang.m@psu.ac.th

***Received**: 19 April 2024*
***Revised**: 12 July 2024*
***Accepted**: 25 July 2024*
***Published**: 30 May 2025*

***Citation**:*
Bensalam, I., Musikasuwan, S., & Saelim, R. (2025). Predictive models of $PM_{2.5}$ concentration with aerosol optical depth and meteorological data in Bangkok area using machine learning techniques. Science, Engineering and Health Studies, 19, 25020001.

ABSTRACT

Air pollution, particularly fine particulate matter ($PM_{2.5}$), is a significant global concern due to its adverse effects on human health and the environment. In response to this challenge, this study aimed to develop and compare machine learning models for predicting $PM_{2.5}$ concentrations, focusing on two monitoring stations in Bangkok. A comprehensive dataset integrating meteorological data and aerosol optical depth (AOD) information was utilized. The models employed in this research included multiple linear regression (MLR), random forest (RF), and support vector machine (SVM). Notably, the SVM model demonstrated superior predictive performance for ambient stations. The findings underscore the importance of tailoring the machine learning method to the specific monitoring station type. Furthermore, the inclusion of influential gas variables such as NO_2, SO_2, CO, and O_3 significantly enhanced the models' predictive capabilities. Fine-tuning hyperparameters further improved model performance. In conclusion, this research highlights the effectiveness of machine learning models in predicting $PM_{2.5}$ concentrations, with important implications for air quality management in urban environments.

Keywords: $PM_{2.5}$; machine learning; aerosol optical depth; multiple linear regression; random forest; support vector machine

1. INTRODUCTION

Air pollution is a major global issue that adversely impacts both economic and social development. It reduces workforce productivity by causing health problems that lead to increased sick days and lower overall efficiency. In addition, it raises healthcare costs and damages crops and infrastructure (Lohmann et al., 2024), complicating urban planning. Socially, air pollution exacerbates health issues and diminishes quality of life, while environmentally, it harms ecosystems and contributes to climate change. Educationally, it impairs children's cognitive development and leads to higher rate of school absenteeism due to health problems. Air pollution ranks as the fourth most significant risk factor for global fatalities, following high blood pressure, poor diet, and smoking (Brauer et al., 2015). A particular concern in the context of air pollution is fine particulate matter, which poses a growing challenge for many countries, especially in developing countries, due to economic expansion, communication advancements, and industrial activities. Fine particulate matter, such as $PM_{2.5}$, has been linked to severe health problems, including cardiovascular and respiratory diseases (Health Effects Institute, 2004; Mehta et al., 2021; Pope et al., 2009). Exposure to fine particulate matter, particularly $PM_{2.5}$, can penetrate the circulatory system and even reach the brain (Croft et al., 2019; Zhang et al., 2018). According to a study by Shaddick et al. (2020), over 99% of the population in Central Asia, South Asia, East Asia, and Southeast Asia is exposed to $PM_{2.5}$ levels that exceed WHO guidelines. These guidelines specify that annual average concentrations of $PM_{2.5}$ should not exceed 5 μg/m^3, and 24-h average exposures should not exceed 15 μg/m^3 more than 3–4 days per year (World Health Organization, 2021).

In recent years, $PM_{2.5}$ has become one of the most serious environmental and public health issues in Thailand, particularly in the Bangkok metropolitan region (BMR), which consists of Bangkok, Nakhon Pathom, Nonthaburi, Pathum Thani, Samut Prakan, and Samut Sakhon. This region is the most densely populated region in Thailand. Furthermore, the BMR's rapid expanding economy, driven by various commercial activities, significantly contributes to the emission of fine particulate matter. This has resulted in elevated ambient $PM_{2.5}$ concentrations, with daily averages frequently exceeding Thailand's 24-h standard (Thongphunchung et al., 2021). Bangkok confronts persistent challenges from high $PM_{2.5}$ levels, primarily attributed to vehicle emissions, biomass burning, and industrial activities (Peng-in et al., 2022). Thailand air quality standard stipulates that $PM_{2.5}$ concentrations should not exceed 37.5 μg/m^3 for daily averages and 15 μg/m^3 for annual averages (Notification of the National Environment Board, 2022). However, data from the Pollution Control Department (PCD) in 2020 revealed that $PM_{2.5}$ concentrations in many areas, particularly in the central region of Thailand, exceeded these limits. Long-term exposure to fine particulate matter can severely impact lung function, increasing the risk of lung cancer, chronic bronchitis, and heart disease, especially in patients with pre-existing respiratory conditions (Pope et al., 2009). Vulnerable groups, such as children and the elderly, are particularly sensitive to $PM_{2.5}$ and more likely to experience adverse health effects (Health Effects Institute, 2004). The rising concentrations of $PM_{2.5}$ are increasingly affecting human health and quality of life, making this an urgent issue that requires to be resolved in Thailand, especially in the BMR. Therefore, developing effective method for predicting $PM_{2.5}$ concentrations is essential to reduce the risk of exposure and prevent further public health impact.

Numerous research studies related to $PM_{2.5}$ in terms of prediction, forecasting, or estimation utilized machine learning algorithms combined with meteorological data. According to a comparative study of machine learning models, this study found that models from machine learning algorithms can forecast $PM_{2.5}$ with good performance, and the extra tree regression model has strong predictability (Minh et al., 2021). In addition to meteorological and $PM_{2.5}$ emission data, aerosol optical depth (AOD) data can also predict $PM_{2.5}$ using the machine learning method (Zamani Joharestani et al., 2019).

The application of machine learning methods is widespread, including detection, estimation, clustering, and prediction. However, in Thailand, particularly in Bangkok, there is a lack of studies focusing on the prediction of $PM_{2.5}$ concentrations. While some research has explored machine learning methods for $PM_{2.5}$ prediction, this study, the researchers aim to develop machine learning models and compare their efficiency in predicting $PM_{2.5}$ concentrations in Bangkok area.

2. MATERIALS AND METHODS

To analyze $PM_{2.5}$ concentrations in Bangkok, Thailand, several datasets were utilized. These included meteorological data from monitoring stations managed by the PCD in Bangkok, and AOD data from the Himawari satellite, operated by the Japan Meteorological Agency for monitoring dust events in East Asia. Initially, multiple linear regression was applied to examine the relationship between $PM_{2.5}$, meteorological factors, and AOD. Subsequently, advanced machine learning techniques, including random forest and support vector machine (SVM), were employed to develop predictive models for $PM_{2.5}$ concentrations in Bangkok area.

2.1 Study area

This research focused on the urban area of Bangkok, the capital and economic hub of Thailand. Historically, the Bangkok region has been equipped with twelve air quality monitoring stations. For the purposes of this study, we narrowed our focus to two specific stations: the Meteorological Department Bangna Station, representing an ambient monitoring facility, and the Din Daeng community housing station, designed as a roadside monitoring point, as detailed in Table 1. Focusing on these two stations offers several advantages. Firstly, they represent diverse environmental settings. The Bangna Station, located in a suburban area, provides data on air quality and meteorological conditions in regions relatively unaffected by heavy urban pollution sources. In contrast, the Din Daeng Station, situated in a densely populated urban community, captures data reflective of air quality impacts from high population density, traffic emissions, and industrial activities. Secondly, this focus enables a comparative analysis of suburban and urban air quality and meteorological conditions. Thirdly, selecting these two stations ensure comprehensive coverage of the city's air quality, incorporating data from both low- and high-pollution areas. Finally, the findings can inform targeted interventions and policy decisions, with measures tailored to different characteristics of suburban and urban areas. By focusing on these specific stations, the study provides a balanced and detailed understanding of air quality and meteorological patterns in Bangkok. Figure 1 visually depicts the locations of the selected monitoring stations for this research.

Table 1. Description of observation location monitored by the PCD in Bangkok, Thailand

No.	Station name	Code	Latitude	Longitude	Type of station
1	Meteorological Department Bangna	05t	13.67	100.61	Ambient, 8 m above sea level
2	Din Daeng community housing	54t	13.77	100.55	Roadside, 14 m above sea level

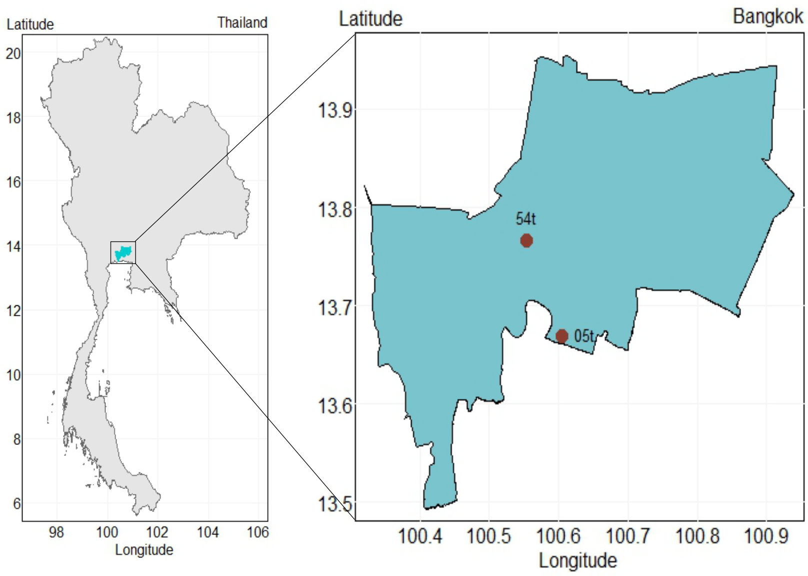

Figure 1. Study locations in Bangkok area

2.2 Data collection

Meteorological data were obtained from air quality monitoring stations managed by the PCD under the Ministry of Natural Resources and Environment. The original dataset consisted of hourly data collected at air quality monitoring stations across the Bangkok area during various periods. This dataset included multiple meteorological variables, such as CO, NO, NO_x, NO_2, SO_2, O_3, wind speed, wind direction, temperature, relative humidity, air pressure, rainfall, PM_{10}, and $PM_{2.5}$. To ensure the consistency across stations, the dataset was standardized to include the following selected variables: CO, NO_2, SO_2, O_3, temperature, wind speed, wind direction, relative humidity, rainfall, air pressure, PM_{10}, and $PM_{2.5}$. The data were then aggregated from hourly to daily averages for the years 2019 to 2022. Missing data were addressed using the imputation method with predictive mean matching.

AOD data, derived from satellite observations were utilized in this study. AOD data were obtained from the Himawari satellite operated by the Japan Meteorological Agency, which monitor dust events in East Asia. AOD is a dimensionless value that quantifies the relationship between the number of particles in the vertical column of the atmosphere and the number of particles measured at ground level.

For this research, AOD data were filtered to include only the desired area, defined by the latitude and longitude coordinates of the monitoring station in Bangkok. Data were collected within a specified range of latitude and longitude, with AOD values aggregated in grid cells, measuring 5x5 km^2. Missing values in the AOD data were handled using the same imputation methods with predictive mean matching applied to the meteorological data.

2.3 Data analysis

Multiple linear regression (MLR) is a statistical tool used to explored the relationship between an outcome variable, such as $PM_{2.5}$ concentration and multiple independent variables, including meteorological factors and AOD. The MLR equation is typically expressed as:

$$Y = \beta_0+\beta_1X_1+\beta_2X_2+...+\beta_nX_n+\varepsilon \qquad (1)$$

This equation quantifies how changes in meteorological variables influence $PM_{2.5}$ levels. MLR is valuable for environmental research as it provides insights into the impact of meteorological conditions on air quality aiding prediction and supporting pollution control and management efforts.

RF regression is a powerful machine learning technique that constructs an ensemble of decision trees during training. Each tree is trained on random subsets of the dataset and input features, ensuring model diversity. Prediction from individual trees are averaged to produce the final output (Breiman, 2001). This method enhances accuracy and mitigates overfitting, making it suitable for both linear and non-linear relationships (Hastie et al., 2009). RF's flexibility, resistance to overfitting, and broad applicability make it a widely used tool in predictive modeling across various domains (Breiman, 2001).

SVM is a machine learning algorithm commonly used to construct hyperplanes for classifying categorical data. When predicting a continuous numeric output, the algorithm is adapted to regression analysis, referred to as support vector regression (SVR). SVR approximate a function g(x) from complex data, represented as follows:

$$G=\{(x_i,y_i)\}\binom{N}{i=1} \qquad (2)$$

The core concept behind SVR involves mapping non-linearly separable data into a higher-dimensional, linearly separable feature space, followed by linear programming for computation (Müller et al., 1997). The regression function, f(x), is expressed as:

$$f(x)= \sum_{i=1}^{D} w_i\Phi_i(x)+b \qquad (3)$$

Here, $\Phi_i(x)$ represents variables that can be estimated from the data. SVR is particularly effective for non-linear data, requiring mapping into a richer feature space for better separability (Liu et al., 2017).

In evaluating the models, the researchers employed three key performance metrics: the coefficient of determination (R^2) to assess the goodness of fit, root mean square error (RMSE) to measure prediction accuracy, and mean absolute error (MAE) to evaluate the magnitude of absolute prediction errors. RMSE is a widely used statistical metric for assessing model performance, particularly in meteorology, air quality, and climate research studies. Similarly, MAE is a common measure for evaluating model accuracy, providing a straightforward interpretation of error magnitude. their extensive use in model evaluation over the years, there is no consensus on which metric is most appropriate for assessing model errors (Hodson, 2022).

3. RESULTS AND DISCUSSION

This study addresses the critical issue of air pollution in Bangkok, focusing specifically on $PM_{2.5}$, a pollutant associated with severe health and environmental consequences. With the city's high population density and rapid economic growth, effective prediction and mitigation of $PM_{2.5}$ are essential. The primary objectives of this research were to develop machine learning models for predicting $PM_{2.5}$ concentrations, using meteorological data and AOD data, and to compare the performance of these models to provide valuable insights into their efficiency in managing $PM_{2.5}$ dynamics in Bangkok.

This section explores the results and interpretation from this study on $PM_{2.5}$ prediction. The analysis incorporates data from two monitoring stations in Bangkok, and the application of three distinct models: MLR, RF, and SVM. The section is structured to present the descriptive analyses, evaluate model performance, and provide comparative insights within the context of air quality management.

3.1 Descriptive analysis

This subsection presents the preliminary analysis of data from two monitoring stations: the Meteorological Department Bangna station (representing an ambient station), and the Din Daeng community housing station, (representing a roadside station). Tables 2 and 3 summarize the meteorological and AOD data results. The analysis revealed that roadside stations exhibited higher mean air pollution levels compared to the ambient station. However, no significant difference were observed in AOD values between the two stations. Notably, the highest recorded $PM_{2.5}$ concentration values exceeded the daily standard (37.5 μg/m^3) set by the PCD of Thailand (Notification of the National Environment Board, 2022). The peak values were 100 μg/m^3 for the ambient stations and 111.96 μg/m^3 for the roadside stations.

Figure 2 illustrates the characteristics of $PM_{2.5}$ concentration data from 2019 to 2022. The data indicate that $PM_{2.5}$ concentrations at both stations were consistency high at the beginning and end of the year, often exceeding the prescribed standards. This seasonal pattern may be attributed to higher air pressure during these periods, which trap warm air and dust within the atmosphere (Amnuaylojaroen, 2022). In contrast, during the middle of the year, $PM_{2.5}$ levels were significantly lower and generally remained within acceptable limits. Despite this seasonal variation, the annual average $PM_{2.5}$ values for each year exceeded both the PCD's annual standards of 15 μg/m^3 and the WHO stricter guideline of 5 μg/m^3 (World Health Organization, 2021), particularly at roadside stations.

3.2 Model performance

Three models (MLR, RF, and SVM) were developed using a meteorological dataset that included $PM_{2.5}$ concentrations and AOD. Initially, MLR models were applied to identify variables significantly influencing $PM_{2.5}$ concentrations at two monitoring stations in Bangkok. For the Meteorological Department Bangna station (ambient station), significant relationships were found between $PM_{2.5}$ and the following variables: PM_{10}, temperature, relative humidity, wind speed, wind direction, air pressure, NO_2, SO_2, and O_3. At the Din Daeng community housing station (a roadside station), $PM_{2.5}$ was significantly related to PM_{10}, interactions between temperature and relative humidity, rainfall, wind speed, air pressure, NO_2, SO_2, and AOD. These station-specific variables were subsequently used to build machine learning models for all three methods. The results are summarized in Table 4. For ambient stations, the SVM model demonstrated superior performance, achieving the highest R^2 and the lowest RMSE and MAE among the three models. The SVM model achieved an R^2, of 89.6%, an RMSE of 3.6448, and an MAE of 2.9701. The MLR model followed closely with an R^2 value of 89.37%, an RMSE of 3.6843, and an MAE of 3.0130. The RF model exhibited the least favorable performance, with R^2, RMSE, and MAE values of 87.37%, 4.0170, and 3.5560, respectively. In contrast, for the roadside stations, the findings indicated that the RF model provided the best predictive performance with an R^2 of 77.74%, RMSE of 5.0251, and MAE of 3.7815. The SVM and MLR models followed with the R^2, RMSE, and MAE values of 74.91% and 71.15%, 5.3346 and 5.7204, 4.0811 and 4.3706, respectively.

These results suggest that the choice of machine learning method should be tailored to the type of monitoring station. The SVM model is particularly effective for ambient stations, while the RF model excels at roadside stations. These findings have significant implications for accurate predicting $PM_{2.5}$ concentration and supporting air quality management efforts in urban environments.

Table 2. Descriptive statistics for ambient station

Variables	Mean	Minimum	Maximum
$PM_{2.5}$	20.92	3.67	100.00
PM_{10}	36.61	9.79	131.21
Temperature	28.63	21.12	34.05
Relative humidity	74.91	39.33	99.00
Rainfall	0.16	0.00	4.52
Wind speed	1.48	0.26	3.34
Wind direction	186.53	53.58	322.17
Air pressure	762.64	756.54	764.46
Carbon monoxide (CO)	0.37	0.03	1.48
Nitrogen dioxide (NO_2)	15.52	0.13	58.22
Sulfur dioxide (SO_2)	0.91	0.00	6.83
Ozone (O_3)	23.49	3.48	71.26
AOD	0.26	0.03	1.46

Table 3. Descriptive statistics for roadside station

Variables	Mean	Minimum	Maximum
$PM_{2.5}$	32.22	13.00	111.96
PM_{10}	63.96	26.00	170.92
Temperature	29.71	21.77	35.24
Relative humidity	66.62	36.21	93.47
Rainfall	0.18	0.00	4.62
Wind speed	0.39	0.02	1.67
Wind direction	186.36	12.17	349.42
Air pressure	756.30	747.50	764.92
Carbon monoxide (CO)	1.41	0.07	5.75
Nitrogen dioxide (NO_2)	29.28	0.52	85.24
Sulfur dioxide (SO_2)	2.51	0.00	7.87
Ozone (O_3)	14.99	0.00	54.91
AOD	0.26	0.03	1.46

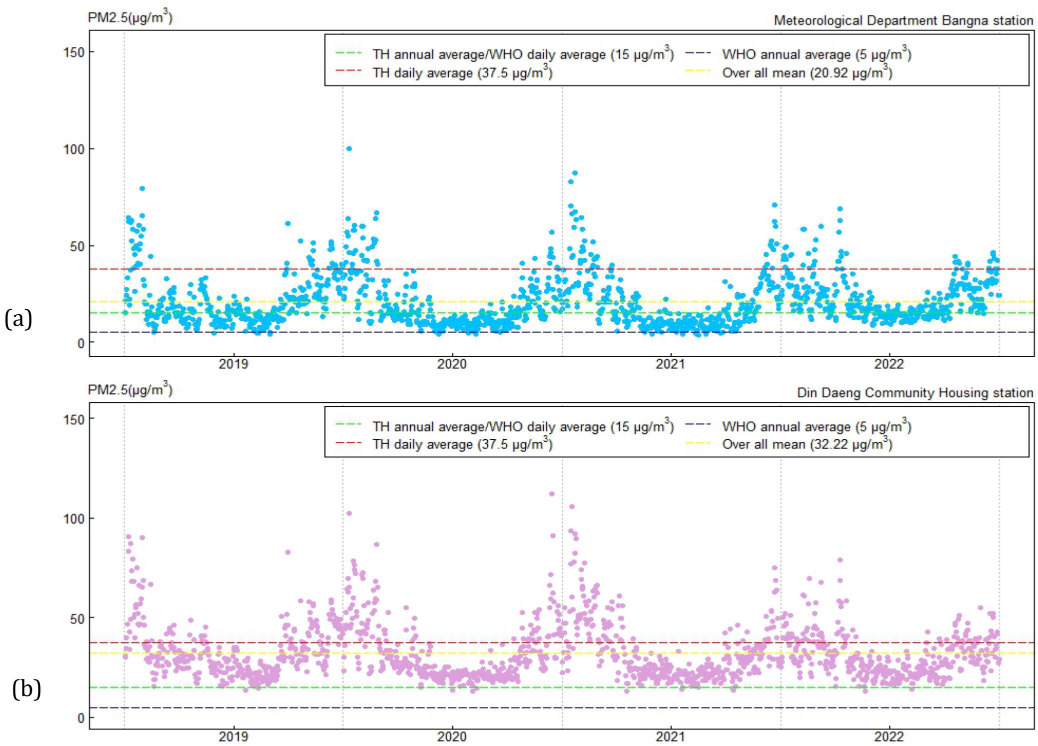

Figure 2. Plots of daily $PM_{2.5}$ concentration values in Bangkok from ambient station (a) and roadside station (b) (2019 to 2022)

Table 4. The evaluation results of the models for predicting $PM_{2.5}$ concentrations

Station	Methods	Train set			Test set		
		R^2	RMSE	MAE	R^2	RMSE	MAE
Ambient	MLR	97.00%	2.3813	1.7622	89.37%	3.6843	3.0130
	RF	96.62%	2.5274	1.7163	87.37%	4.0170	3.5560
	SVM	96.97%	2.3923	1.7483	89.60%	3.6448	2.9701
Roadside	MLR	86.79%	5.4131	4.2021	71.15%	5.7204	4.3706
	RF	89.38%	4.8516	3.4806	77.74%	5.0251	3.7815
	SVM	86.28%	5.5149	4.1770	74.91%	5.3346	4.0811

All three models demonstrated their effectiveness in predicting $PM_{2.5}$ concentration using the dataset obtained from the PCD of Thailand. This effectiveness is evident in the prediction characteristics presented in Figures 3 and 4. Furthermore, as shown in Table 4, all three models perform well in detecting concentration values, with the SVM and RF models superior performance. These finding align with the results of Minh et al. (2021) and Huang et al. (2018), reinforcing the efficacy of machine learning models in predicting and detecting $PM_{2.5}$ concentrations. The selection of variables used in creating these models was guided by Peng-in et al. (2022), which confirmed that the inclusion of gas variables such as NO_2, SO_2, CO, and O_3, significantly enhance model performance. The incorporation of these influential variables has been instrumental in improving efficacy of the models. Furthermore, the fine-tuning of hyperparameters has further improved the performance of these machine learning models, underscoring the importance of parameter adjustment in enhancing predictive capabilities. It is essential to acknowledge that different locations and the nature of variables can yield varying analytical results when predicting $PM_{2.5}$ concentrations. Careful consideration must be given to tailoring these models to the specific characteristics of the data from each location. This is supported by empirical evidence from studies such as Minh et al. (2021) and Huang et al. (2018), which showed that air quality models perform differently across various geographical and climatic conditions. The interaction and impact of pollutants like NO_2, SO_2, CO, and O_3 vary significantly depending on local pollution sources and environmental conditions, as highlighted by Peng-in et al. (2022). Established practices in machine learning and data science emphasize the need for model fine-tuning to adapt to new datasets and environments, ensuring accuracy and robustness. By acknowledging these potential variabilities, models can be applied with tailored adjustments, enhancing their predictive performance and reliability. Furthermore, the researchers' experience with diverse datasets, including the primary dataset from the PCD of Thailand, further underscores the necessity of model calibration to account for local data characteristics. Collectively, these factors support the assertion that geographical and variable-specific differences must be carefully considered to achieve reliable analytical results when predicting $PM_{2.5}$ concentrations.

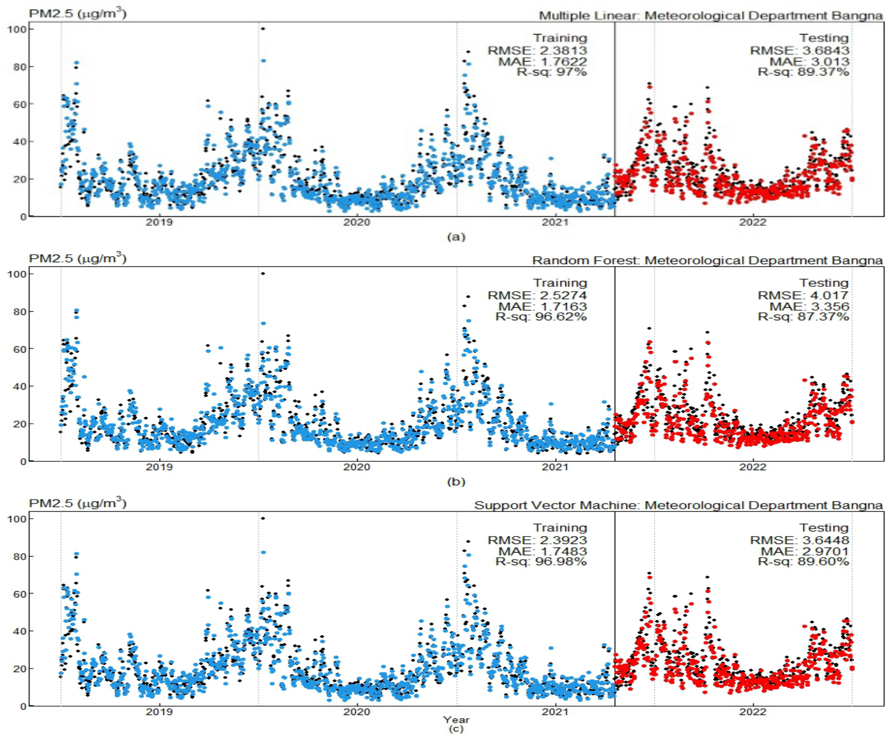

Figure 3. The performance of $PM_{2.5}$ concentration predictive models at ambient stations (a-c)

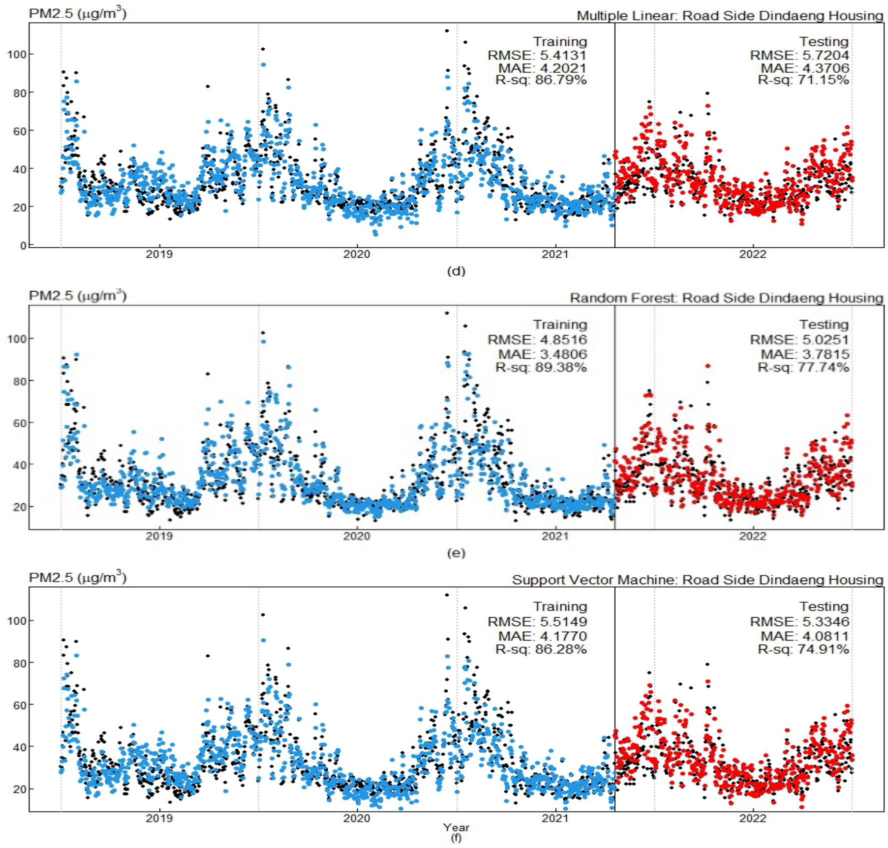

Figure 4. The performance of $PM_{2.5}$ concentration predictive models at roadside stations (d-f)

4. CONCLUSION

This study developed three models (MLR, RF, and SVM) to predict $PM_{2.5}$ concentrations using a comprehensive meteorological dataset that included AOD. The models were designed to specific monitoring station types, with MLR identifying influential variables for both ambient and roadside stations. For ambient stations, the SVM model outperformed the other models, achieving the highest R^2 and the lowest RMSE and MAE, demonstrating its effectiveness in predicting $PM_{2.5}$ concentrations. In contrast, the RF model excelled, at roadside stations, indicating the importance of selecting the appropriate model based on station types and environmental contexts.

The results have significant implications for improving air quality management in urban environments. The inclusion of gas variables such as NO_2, SO_2, CO, and O_3, validated by prior research, significantly enhanced predictive performance. Additionally, fine-tuning hyperparameters further optimized model accuracy. It is essential to acknowledge that the prediction accuracy is influenced by the choice of location and the nature of variables, emphasizing the importance of customizing models to suit specific data characteristics.

In summary, this research supports the effectiveness of machine learning models in predicting and detecting $PM_{2.5}$ concentrations, while underscoring the need for careful consideration of variables and model selection in diverse environmental contexts.

ACKNOWLEDGMENTS

This study was funded by the National Science, Research and Innovation Fund (NSRF) and Prince of Songkla University (Grant No. SAT6601245S-0), Faculty of Science and Technology, Prince of Songkla University, Pattani campus, Thailand and the Graduate School, Prince of Songkla University, Thailand. The authors gratefully acknowledge the helpful comments and suggestions provided by the reviewers, which have significantly improved the quality and clarity of this manuscript.

REFERENCES

Amnuaylojaroen, T. (2022). Prediction of $PM_{2.5}$ in an urban area of Northern Thailand using multivariate linear regression model. *Advances in Meteorology*, *2022*(1), Article 3190484. https://doi.org/10.1155/2022/3190484

Brauer, M., Freedman, G., Frostad, J., van Donkelaar, A., Martin, R. V., Dentener, F., van Dingenen, R., Estep, K., Amini, H., Apte, J. S., Balakrishnan, K., Barregard, L., Broday, D., Feigin, V., Ghosh, S., Hopke, P. K., Knibbs, L. D., Kokubo, Y., Liu, Y., . . . Cohen, A. (2016). Ambient air pollution exposure estimation for the Global Burden of Disease 2013. *Environmental Science & Technology*, *50*(1), 79–88. https://doi.org/10.1021/acs.est.5b03709

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Croft, D. P., Zhang, W., Lin, S., Thurston, S. W., Hopke, P. K., Masiol, M., Squizzato, S., van Wijngaarden, E., Utell, M. J., & Rich, D. Q. (2019). The association between respiratory infection and air pollution in the setting of air quality policy and economic change. *Annals of the American Thoracic Society*, *16*(3), 321–330. https://doi.org/10.1513/AnnalsATS.201810-691OC

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning* (2nd ed.). Springer.

Health Effects Institute. (2004). *Health effects of outdoor air pollution in developing countries of Asia* (Special Report 15 Executive Summary). https://www.healtheffects.org/system/files/SpecialReport15Summ.pdf

Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geoscientific Model Development*, *15*(14), 5481–5487. https://doi.org/10.5194/gmd-15-5481-2022

Huang, K., Xiao, Q., Meng, X., Geng, G., Wang, Y., Lyapustin, A., Gu, D., & Liu, Y. (2018). Predicting monthly high-resolution $PM_{2.5}$ concentrations with random forest model in the North China Plain. *Environmental Pollution*, *242*(Part A), 675–683. https://doi.org/10.1016/j.envpol.2018.07.016

Liu, B.-C., Binaykia, A., Chang, P.-C., Tiwari, M. K., & Tsao, C.-C. (2017). Urban air quality forecasting based on multi-dimensional collaborative support vector regression (SVR): A case study of Beijing-Tianjin-Shijiazhuang. *PLOS ONE*, *12*(7), Article e0179763. https://doi.org/10.1371/journal.pone.0179763

Lohmann, P. M., Probst, B., Gsottbauer, E., & Kontoleon, A. (2024). High levels of air pollution reduce team performance. *Journal of Economic Psychology*, *101*, Article 102705. https://doi.org/10.1016/j.joep.2024.102705

Mehta, U., Dey, S., Chowdhury, S., Ghosh, S., Hart, J. E., & Kurpad, A. (2021). The association between ambient $PM_{2.5}$ exposure and anemia outcomes among children under five years of age in India. *Environmental Epidemiology*, *5*(1), Article e125. https://doi.org/10.1097/EE9.0000000000000125

Minh, V. T. T., Tin, T. T., & Hien, T. T. (2021). $PM_{2.5}$ forecast system by using machine learning and WRF model, a case study: Ho Chi Minh City, Vietnam. *Aerosol and Air Quality Research*, *21*(12), Article 210108. https://doi.org/10.4209/aaqr.210108

Müller, K.-R., Smola, A. J., Rätsch, G., Schölkopf, B., Kohlmorgen, J., & Vapnik, V. (1997). Predicting time series with support vector machines. In W. Gerstner, A. Germond, M. Hasler, & J.-D. Nicoud (Eds.), *Lecture Notes in Computer Science, Volume 1327* (pp. 999–1004). Springer. https://doi.org/10.1007/BFb0020283

Notification of the National Environment Board. (2022, July 8). *Royal Thai Government Gazette*. No. 139 Special Section 163 D. pp. 21–22. https://www.ratchakitcha.soc.go.th/DATA/PDF/2565/E/163/T_0021.PDF [in Thai]

Peng-in, B., Sanitluea, P., Monjatturat, P., Boonkerd, P., & Phosri, A. (2022). Estimating ground-level $PM_{2.5}$ over Bangkok Metropolitan region in Thailand using aerosol optical depth retrieved by MODIS. *Air Quality, Atmosphere & Health*, *15*(11), 2091–2102. https://doi.org/10.1007/s11869-022-01238-4

Pope, C. A., Ezzati, M., & Dockery, D. W. (2009). Fine-particulate air pollution and life expectancy in the United States. *The New England Journal of Medicine*, *360*(4), 376–386. https://doi.org/10.1056/NEJMsa0805646

Shaddick, G., Thomas, M. L., Mudu, P., Ruggeri, G., & Gumy, S. (2020). Half the world's population are exposed to increasing air pollution. *npj Climate and Atmospheric Science*, *3*, Article 23. https://doi.org/10.1038/s41612-020-0124-2

Thongphunchung, K., Phosri, A., Sihabut, T., & Patthanaissaranukool, W. (2021). Short-term effects of particulate matter on outpatient department visits for respiratory diseases among children in the Bangkok Metropolitan Region: A case-crossover study. *Air Quality, Atmosphere & Health*, *14*(11), 1785–1795. https://doi.org/10.1007/s11869-021-01053-3

World Health Organization. (2021). *WHO global air quality guidelines: Particulate matter ($PM_{2.5}$ and PM_{10}), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. https://iris.who.int/handle/10665/345329

Zamani Joharestani, M., Cao, C., Ni, X., Bashir, B., & Talebiesfandarani, S. (2019). $PM_{2.5}$ prediction based on random forest, XGBoost, and deep learning using multisource remote sensing data. *Atmosphere*, *10*(7), Article 373. https://doi.org/10.3390/atmos10070373

Zhang, X., Kang, J., Chen, H., Yao, M., & Wang, J. (2018). $PM_{2.5}$ meets blood: In vivo damages and immune defense. *Aerosol and Air Quality Research*, *18*(2), 456–470. https://doi.org/10.4209/aaqr.2017.05.0167