Spatiotemporal Patterns of SO_2, NO_2, and CO Using Satellite Data Integrated Machine Learning in EEC Thailand

Pavinee Polsuwan [1*], Zhengfeng Shao [2] and Phattraporn Soytong [3]

[1, 3]Department of Geoinformatics, Faculty of Humanities and Social Sciences, Burapha University, Thailand
[2]State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS); Wuhan University, China
***Corresponding Author E-mail**: polsuwan2@hotmail.com
Received: 13/5/25, **Revised**: 24/5/25, **Accepted**: 26/5/25

Abstract

Air pollution causes a major global environmental challenge especially in industrialized regions such as the Eastern Economic Corridor (EEC) of Thailand. This study assesses the spatiotemporal trends of SO_2, NO_2 and CO concentrations in the EEC (2019–2022) using Sentinel-5P satellite data, ERA5 meteorological variables and machine learning models (Random Forest, XGBoost, LightGBM). Results revealed that pollutant fluctuations were significantly influenced by wind, aerosol index, surface pressure and dew point. XGBoost provided superior accuracy (R^2: SO_2=0.95, NO_2=0.90, CO=0.96). The study demonstrates satellite and machine learning efficacy in air quality monitoring to identify pollution hotspots and critical environmental drivers. This is to support air quality management in Thailand.

Keywords: Air pollution, SO_2, NO_2, CO, Sentinel-5P, Machine learning, Spatiotemporal analysis, EEC

1. Introduction

Thailand's Eastern Economic Corridor (EEC) is a government long-term 20 years strategy aimed at improving infrastructure, increasing investment and enhancing transportation and human development to transform the country into the higher income country. The goal of this initiative is to revitalize the eastern seaboard, specifically the provinces of Chachoengsao, Chonburi and Rayong by promoting innovative and high value industries. However, these industrial activities emit pollutants into the air, producing waste that threatens human health and the environment [1]. Manufacturing industries are especially problematic because they are energy intensive and emit large quantities of pollutants which leads to more severe air pollution compared to other sectors [2]. Furthermore, industrial emissions, construction and urban expansion contribute air pollution. Construction activities release particulate matter (PM), dust and other pollutants exacerbating air quality issues [3]. The rapid urbanization embodied by the development of new housing, commercial spaces, and infrastructure projects. This has led to increasing vehicular emissions and energy consumption which further degrading air quality [4].

Air pollution poses significant risks to health of the public such as sulfur dioxide (SO_2), nitrogen dioxide (NO_2), and carbon monoxide (CO) are particularly concerning due to their adverse health effects. SO_2 can cause respiratory issues and aggravate heart and lung diseases. NO_2 irritates the respiratory system, leading to decreased lung function and higher risks of respiratory infections. CO is a dangerous gas inhibits the blood's ability to transport oxygen resulting in symptoms from headaches to severe cardiovascular issues [5].

2. Research Method

This study integrates remote sensing data from Sentinel 5p atmospheric data, ERA5 meteorological data and NASA SRTM product inquisition by Google earth code engine (GEE). After that, we used ground-based air quality monitoring to estimate air pollution by using three machine learning models: Random Forest Regression, XGBoost, and LightGBM. Thereafter, we implemented and evaluated using metrics. We also calculated the score of feature importance by Python. Finally, we generated models that help estimate air pollution levels and perform spatiotemporal patterns analysis.

2.1 Data collection

The main objective of this study was to monitor trends in SO_2, NO_2 and CO pollutants in EEC Thailand using Sentinel-5 images (TROPOMI) derived from GEE (1 January 2019 to 31 December 2022) as main variates in the inputs. Meanwhile, some other datasets are meteorological data using ERA-5 satellites derived from GEE, Which are adopted as auxiliary variates to enlarge the applicability of the trained models. In addition, the measurements from the Thailand Pollution Control Department are considered the ground truth values.

2.1.1 Ground measurement data

According to the Enhancement and Conservation of National Environmental Quality Act, B.E. 2535 [6], Thailand's Pollution Control Department (PCD) provides official annual reports on ambient air pollutant concentrations across the country to support the promotion and maintenance of environmental quality [7]. In this study, ground-based measurements of SO_2, NO_2 and CO from January 2019 to December 2022 were obtained from the PCD's public database http://air4thai.pcd.go.th.

SO_2 concentrations are measured using two principal methods; The Pararosaniline method and Ultraviolet Fluorescence. The Pararosaniline method involves absorbing SO_2 through potassium tetrachloromercurate solution to form a dichlorosulfitomercurate complex, which reacts with formaldehyde and pararosaniline to

produce methyl sulfonic acid, measurable at 548 nm. The UV Fluorescence method excites SO_2 with ultraviolet light, and fluorescence is emitted and detected at 190–230 nm. These methods offer high sensitivity and are used to derive both 24-hour average and annual mean values [8].

NO_2 is measured primarily using the Chemiluminescence method, where ozone reacts with nitric oxide—converted from NO_2—resulting in light emission detected at wavelengths above 600 nm. Additionally, Cavity Attenuated Phase Shift (CAPS) Spectroscopy is employed, utilizing modulated light in an optical cavity to enable precise, real-time detection. Measurements are typically reported as 1-hour averages and annual means [9].

CO is measured using Non-Dispersive Infrared (NDIR) Detection which identifies CO by its absorption of infrared radiation at specific wavelengths. This technique is known for its high specificity and reliability with data commonly recorded as 8-hour averages. therefore, This method suitable for assessing short-term exposure levels [10].

During the study period, nine air quality monitoring stations were operational in the Eastern Economic Corridor (EEC) including five stations in Rayong, three in Chonburi and one in Chachoengsao. Due to this limited spatial distribution satellite remote sensing is used to complement the ground network, offering extensive spatial coverage. Furthermore, the presence of over 13,000 factories across the EEC, as reported in the 2023–2027 Master Plan for the Eastern Special Development Zone [11], underscores the necessity of satellite-supported monitoring particularly in regions lacking sufficient ground-based measurement infrastructure.

2.1.2 Remote sensing data

Remote sensing data can be used as a valuable alternative tool for monitoring air pollution and can support decision makers as they can be employed in retrieving and mapping of air quality parameters in a synoptic and multi-temporal coverage at regular intervals and dynamic scale [12]. In this study, The TROPOMI instrument was aboard the Sentinel-5P and ERA-5 satellite data integrated with google earth engine to analyze air quality data.

Sentinel-5P operates across wavelengths ranging from UltraViolet (UV) to ShortWave InfraRed (SWIR). This hyperspectral spec-trometer provides daily high-resolution observations of SO_2, NO_2, and CO using passive remote sensing techniques. The standard pixel size for the most spectral bands is 7×3.5 km², except for the UV1 band (7×28 km²) and SWIR bands (7×7 km²) [13]. Evaluations indicate that TROPOMI's atmospheric products meet the mission's accuracy requirements. In this study, annual mean data from January 2019 to December 2022 was collected by google earth code engine (GEE). All data was regraded into 10 km × 10 km grids using bilinear interpolation and main variables include "sulfurdioxide_total_vertical_column"(SO2), "nitrogendioxide_tropospheric_column" (NO2) and "CO_column_density" (CO) with included the "absorbing_aerosol_index"(AAI) as an auxiliary variable.

ERA5 is the fifth generation atmospheric reanalysis dataset that encompasses uncertainty information for all of the variables at reduced spatial and temporal resolutions [14]. The meteorological data were from the fifth generation European center for medium-range weather forecasts atmospheric reanalysis of the global climate (ERA5) produced by the Copernicus climate change service included 2m temperature (TEM), 2m dew point temperature (DEW), 10m u-component of wind (UCOMPO) , 10m v-component of wind (VCOMPO), surface pressure (PRES), total evaporation (EVA). Land surface data the digital elevation model (DEM) and slope from the shuttle radar topography mission (STRM) digital elevation dataset from NASA SRTM V3 product (SRTM Plus).

Table 1 Data Collection

Type Sources	Dataset	Index	Spatial resolution	Periods
Remote sensing Data	Sentinel 5P	SO2	10 km × 10 km	January 2019 to December 2022
		NO2		
		CO		
		AAI		
	ERA5	DEW		
		UCOMPO		
		VCOMPO		
		PRES		
		EVA		
	NASA SRTM	DEM		
		SLOPE		
Ground data	Annual-Pollutants	SO2	Ground monitoring 9 Station	
		NO2		
		CO		

Model Training

Model Training is a critical phase in machine learning that involves building models to recognize patterns and make predictions. This study used three models, Random Forest, LightGBM and XGBoost. They are commonly used due to their high accuracy and robustness in regression and classification tasks. These models rely on ensemble learning, where multiple decision trees or models are combined to enhance performance and reduce overfitting.

The input features included remote sensing variables as tropospheric concentrations of SO_2, NO_2, CO and AAI from the Sentinel-5P satellite. Meteorological auxiliary variables were sourced from the ERA5 reanalysis dataset including 2-meter air temperature, dew point temperature, surface pressure, 10-meter u- and v-component winds and total evaporation. Topographic variables including elevation and slope were derived from NASA's Shuttle Radar Topography Mission (SRTM). All spatial datasets were standardized to a 10 km × 10 km resolution using bilinear interpolation to ensure consistency across variables. Because the coarsest resolution among the input variables was at 10 km. To ensure consistency across datasets and avoid introducing scale mismatches that

could negatively affect model performance all input layers were resampled to this common resolution before model training.

The output variable was the annual average concentration of each pollutant SO_2, NO_2, and CO. These predicted values were then used to generate spatial concentration maps by applying the trained models. The resulting maps visualized the estimated pollutant distribution across the entire study area analysis of air quality from 2019 to 2022.

For model training and validation, ground-based pollution data served as the dependent variable (target label). All input features and target values were aggregated to annual averages to align temporal resolution and minimize the impact of short-term anomalies. The dataset spans from 2019 to 2022. A 10-fold cross-validation approach was used to train the models, ensuring generalizability and reducing the risk of overfitting.

2.2 Implementation and Evaluation Model

In this study measures the predictive results using metrics, to evaluate the different models by classification performance at various levels by Mean Absolute Error (MAE), Root-Mean-Square Error (RMSE) and R-squared (R^2) values.

$$MAE=\frac{1}{n}\sum_{i=1}^{n}\left|y_i-\hat{y}_i\right| \quad (1)$$

MAE quantifies the difference between the predicted and actual values with lower values representing better performance.

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2} \quad (2)$$

RMSE is particularly sensitive to large prediction errors within the training set, making it an important metric for evaluating the accuracy of predictions. A lower RMSE value indicates higher model performance.

$$R^2 = 1 - [\Sigma\,(y_i - \hat{y}_i)^2 / \Sigma\,(y_i - \bar{y})^2] \quad (3)$$

The R^2 score known as the coefficient of determination measures how well the predicted values align with the actual observed values. The score ranges from negative values to a maximum of 1. Higher values are supposed to indicate a better fit. A value closer to 1 signifies stronger alignment between predictions and actual outcomes and reflects higher predictive accuracy. In this formula y_i noted the actual value of the i-th data point. $\hat{y_i}$ is the corresponding predicted value and $\bar{y}$ represents the mean of the observed values in the dataset. Generally, R^2 values range between 0 and 1, where an R^2 of 0 implies that the model cannot predict the target variable. While an R^2 of 1 indicates good prediction. If the R^2 value is negative, it suggests that the model performs worse than simply using the mean of the target variable as the prediction

3. Results and Discussion

3.1 Analysis Spatiotemporal Patterns

The average values shown in Figure 1 are derived from the model outputs which were generated by applying trained machine learning predictions. Annual mean concentrations of three air pollution from 2019 to 2022. Concentrations are reported in parts per billion (ppb). The concentration of sulfur dioxide (SO_2) shows slight fluctuations, starting at 2.41 ppb in 2019, decreasing to 1.83 ppb in 2020, and then rising to 2.29 ppb in 2021 and 2.6 ppb in 2022. Nitrogen dioxide (NO_2) concentrations remain relatively high and stable, starting at 12.22 ppb in 2019, slightly decreasing to 11 ppb in 2020, then increasing to 11.81 ppb in 2021 and slightly decreasing to 11.16 ppb in 2022.Carbon monoxide (CO) shows lower values overall, with a decreasing trend from 0.59 ppb in 2019 to 0.43 ppb in 2022.

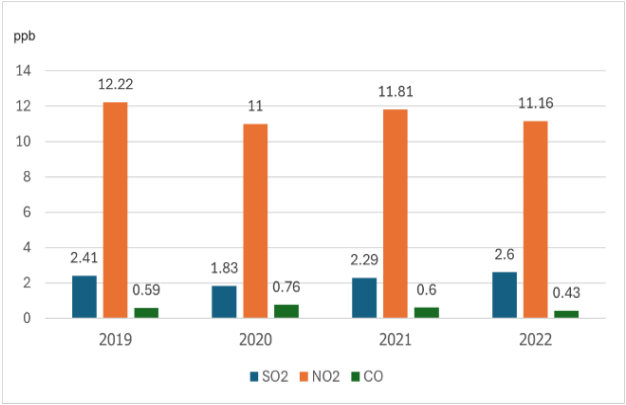


Fig. 1 Average annual concentration of SO_2, NO_2, and CO

3.2 Machine Learning Models Performance

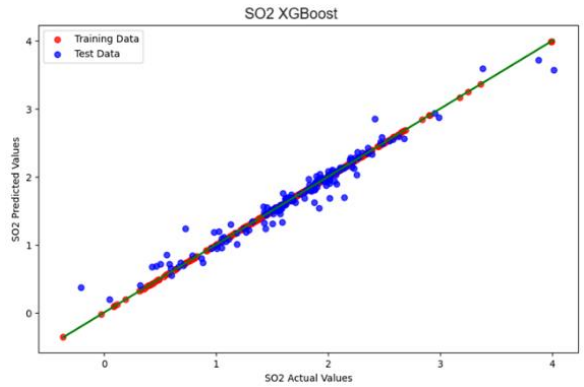
The predictive performance of the three machine learning models XGBoost, LightGBM and Random Forest was assessed using 10-fold cross-validation across SO_2, NO_2, and CO datasets. The evaluation was based on standard regression metrics, including R^2, Mean Absolute Error (MAE), and Root Mean Square Error (RMSE).

For all three pollutants, XGBoost consistently exhibited the highest predictive accuracy with superior alignment between predicted and observed values and the lowest error rates. The model's ability to handle complex nonlinear relationships contributed to its robustness and generalizability.

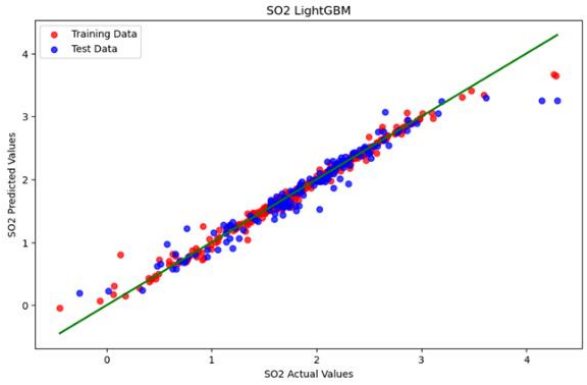
In the case of SO_2, XGBoost achieved the best performance as reflected in its tight clustering of predicted values around the ideal regression line and minimal residuals. LightGBM also performed well although its predictions displayed slightly greater dispersion. Random Forest showed higher variability in predictions and indicating a relatively reduced capacity to generalize in the presence of complex spatial patterns as shown in figure 2.

For NO_2, XGBoost again yielded the most precise results, followed by LightGBM. Random Forest exhibited the widest distribution of errors, suggesting limitations in

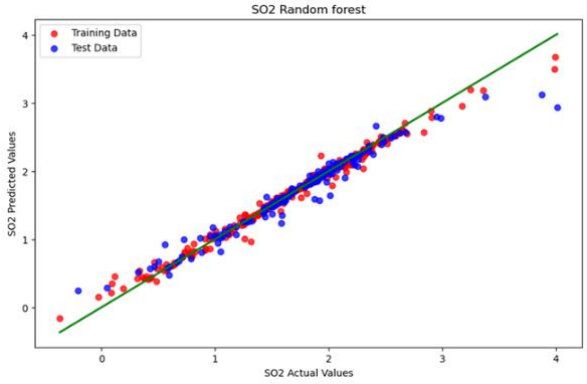
capturing temporal variability and feature interactions as shown in figure 3.



(a)



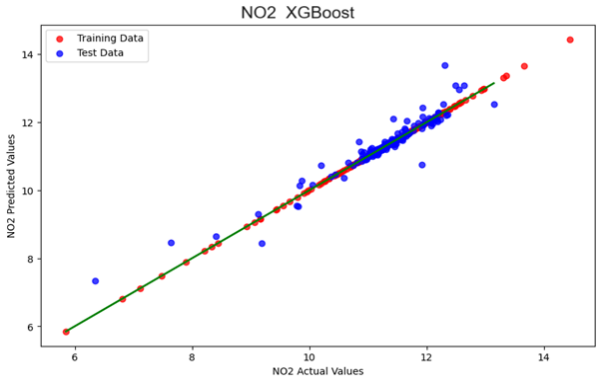
(b)



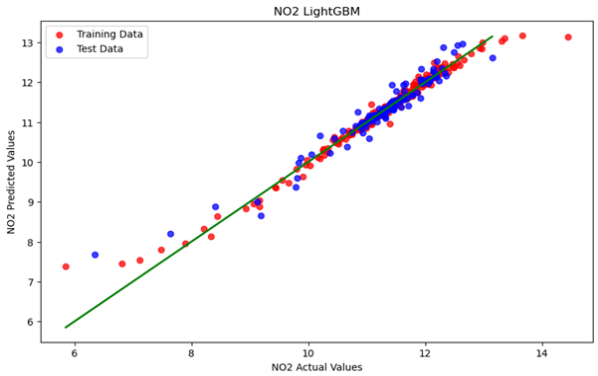
(C)

Fig. 2 (a) XGBoost (b) LightGBM (c) Random Forest SO_2 predicted and actual values

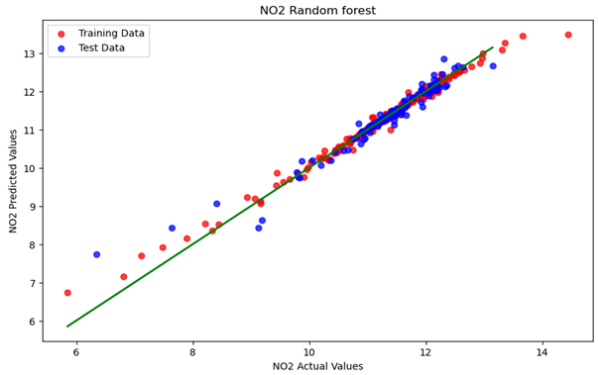
In predicting CO concentrations, XGBoost maintained its leading performance which particularly in capturing both high and low concentration values with minimal error. LightGBM provided similar but slightly less consistent results, whereas Random Forest revealed pronounced deviations, especially in lower concentration ranges as shown in figure 4.



(a)

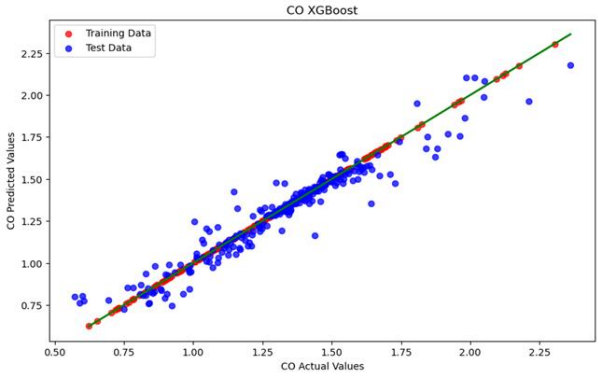


(b)

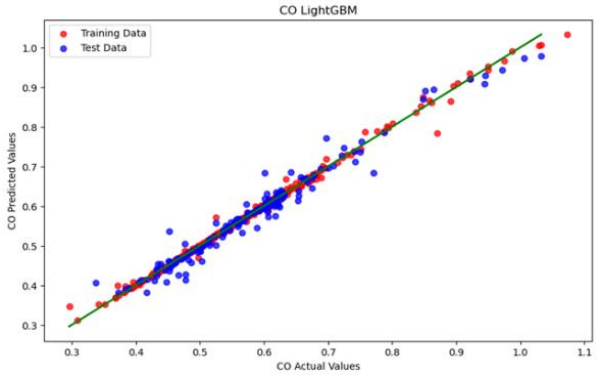


(C)

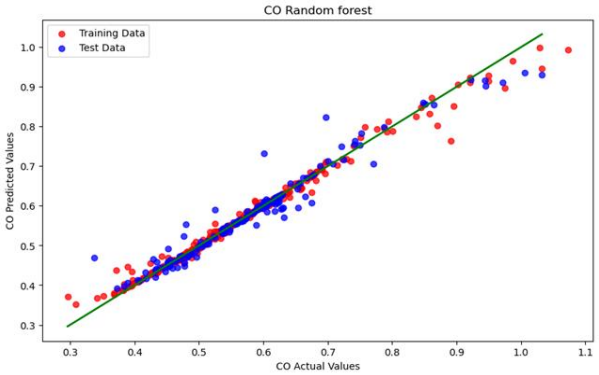
Fig. 3 (a) XGBoost (b) LightGBM (c) Random Forest NO_2 predicted and actual values



(a)



(b)



(C)

Fig. 4 (a) XGBoost (b) LightGBM (c) Random Forest CO predicted and actual values

The comparative results presented in Table 2 and Table 3 further support the conclusion that XGBoost outperforms the other models in terms of both training and testing accuracy. This underscores the suitability of gradient boosting frameworks for spatiotemporal air quality prediction tasks, particularly in settings characterized by complex environmental interactions.

Table 2 Accuracy of Model training data

Air Pollutant	Model	Training R^2	Training MAE	Training RMSE
SO2	XGBoost	0.999997	0.000781	0.001077
SO2	RF	0.983212	0.039601	0.076615
SO2	LightGBM	0.998828	0.009218	0.020243
NO2	XGBoost	0.999999	0.00067	0.000965
NO2	RF	0.984405	0.044299	0.104433
NO2	LightGBM	0.972398	0.060471	0.138935
CO	XGBoost	0.998616	0.000438	0.00058
CO	RF	0.980857	0.000973	0.002158
CO	LightGBM	0.997722	0.000338	0.000745

Table 3 Accuracy of Model testing data

Air Pollutant	Model	Testing R^2	Testing MAE	Testing RMSE
SO2	XGBoost	0.957113	0.080417	0.128924
SO2	RF	0.944955	0.075459	0.146058
SO2	LightGBM	0.94994	0.08287	0.139287
NO2	XGBoost	0.904464	0.125868	0.244183
NO2	RF	0.938802	0.08913	0.195435
NO2	LightGBM	0.937142	0.112321	0.198067
CO	XGBoost	0.966192	0.002073	0.003288
CO	RF	0.959723	0.001824	0.003589
CO	LightGBM	0.917175	0.002419	0.005147

Early studies [15] and [16] used OMI data with random forest spatiotemporal kriging (RF-STK) and Random Forest with Kriging (RF-K) models achieving moderate CV-R^2 values ranging from 0.62 to 0.78. Reference [17] shows RF–STK model using Measurements of Pollution in the Troposphere CO retrievals (MOPITT) data achieving a CV-R^2 of 0.51. This early work was limited by spatial resolution and data quality. Reference [18] shows LightGBM model with Sentinel-5P data to estimate CO concentrations with a CV-R^2 of 0.71 and RMSE of 0.26. Reference [19] employed the Space Time Extra Trees (STET) model with big data to estimate NO_2, SO_2, and CO achieving high predictive accuracy with CV-R^2 values 0.84 and highlighted the potential of combining ground observations with reanalysis data to address gaps in satellite coverage. Nested XGBoost models in study [20] have also shown exceptional performance in predicting NO_2, achieving an CV- R^2 of 0.93 and RMSE of 4.19. Neural Networks PCA hybrid models as demonstrated in [21] enhanced the accuracy of SO_2 achieving CV-R^2 = 0.976 with OMI data. This approach significantly improved data quality for polluted areas by reducing noise. These advancements have improved air quality monitoring. Machine learning models often struggle to generalize across diverse geographies necessitating region specific adaptations. Validation with ground-based measurements remains critical to ensuring model reliability.

3.3 Feature Importance

Among all the pollutant models wind components, surface pressure, and dew point were identified as the main factors affecting air pollution levels. Wind patterns, especially the U component, had the most significant impact on CO and NO_2, highlighting the importance of atmospheric circulation in spreading pollutants. On the other hand, evaporation and the digital elevation model (DEM) play an important role in SO_2 levels, showing that the shape of the land and water are key factors in how SO_2 spreads. These results are illustrated by the average feature importance values shown in Figure 5 , which highlights the key environmental factors affecting pollution in the EEC.

In summary, across all pollutant models, atmospheric circulation, particularly the U component of wind, emerged as the most influential driver of pollution dispersion in the EEC. The effect of surface pressure and dew point was similar to all the models. The role of topography and evaporation was more prominent in SO_2 predictions, indicating that localized geographic and hydrological factors play a more significant role in certain pollutant dynamics. These can inform targeted high pollution control such as focusing on areas with high industrial and vehicular activity during periods of low wind speed or high atmospheric pressure.

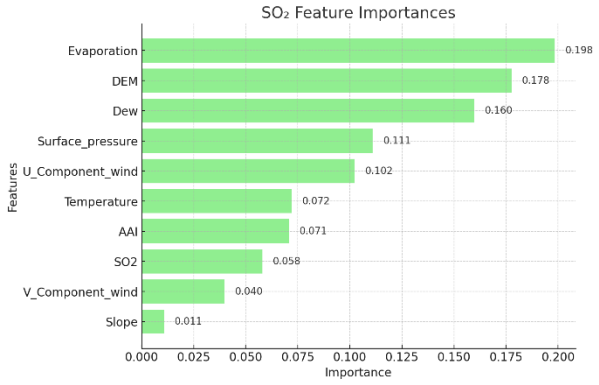
3.4 Spatiotemporal Concentration map

The spatiotemporal distribution of sulfur dioxide (SO_2) concentrations over the Eastern Economic Corridor

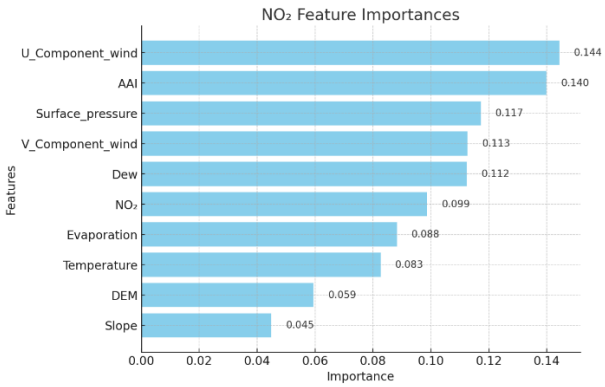
(EEC) of Thailand was analyzed from 2019 to 2022 using Sentinel-5P TROPOMI data integrated with ERA5 and NASASRTM by the most accurate model XGBoost generated into concentration map.

SO_2 is primarily emitted from industrial activities, power plants, and vehicular traffic making it a key indicator of air quality. In this study annual maps were generated to visualize changes in SO_2 concentrations across the study area. The spatial and temporal analysis of SO_2 concentrations from 2019 to 2022 reveals a general decline in pollution levels across the EEC with occasional fluctuations. Chachoengsao consistently exhibited higher concentrations throughout the study period as shown in figure 6.

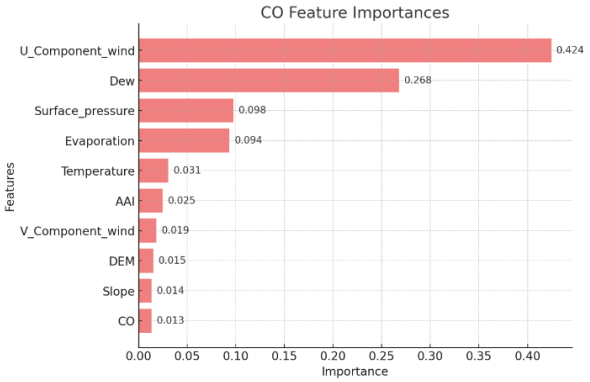
The NO_2 concentrations from 2019 to 2022 Chachoengsao consistently emerges as the most polluted province, with NO_2 concentrations peaking in 2021. The decrease observation in 2022 suggests some improvements in air quality likely due to regulatory actions or changing industrial activity levels. The pollution hotspots in Chachoengsao and Chon Buri linked with the relatively low levels in Rayong show the differences in NO_2 pollution levels across the EEC. As shown in figure 7. The concentration of CO pollutant concentrations from 2019 to 2022 showed a decreasing trend across the study area. The northern areas of Chachoengsao consistently have the highest concentrations of CO, but the overall intensity of pollution has significantly declined since 2019. As shown in figure 8.



(a)



(b)



(c)

Fig. 5 (a) SO_2 (b) NO_2 (c) CO feature importance Performance of pollutants factors

4. Conclusion and Suggestions

This study utilized multisource remote sensing data and machine learning algorithms (XGBoost, Random Forest regression, and LightGBM) to model and predict air pollution trends, explicitly focusing on SO_2, NO_2, and CO concentrations in the Eastern Economic Corridor (EEC) of Thailand from 2019 to 2022 by combining atmospheric data from Sentinel-5P and meteorological data from ERA5 and NASA SRTM products. Also, it can be summarized based on the three main objectives.

1) Feature Importance The key factors influencing air pollution were identified. Across all models, wind components, surface pressure and dew point emerged as the most critical predictors. XGBoost and Random Forest regression showed that the U component of wind, surface pressure and dew point were particularly influential for NO_2 and CO concentrations, highlighting the role of atmospheric circulation in pollutant distribution. For SO_2, features like evaporation and digital elevation model (DEM) were more significant, suggesting that topographical and hydrological factors are important for SO_2 dispersion.

2) Model Performance Using satellite imagery data and environmental data predictive models were developed to estimate SO_2, NO_2, and CO levels. The XGBoost model demonstrated the highest accuracy, outperforming the other models in terms of R^2, RMSE, and MAE metrics. R^2 values of 0.95 for SO_2, 0.90 for NO_2, and 0.96 for CO confirm that it is more accurate than Random Forest regression and LightGBM in this study. Although the model's performance varied depending on the pollutant and the study area, the results are consistent with other studies that found XGBoost to be highly effective for environmental modeling tasks.

3) The analysis of SO_2, NO_2 and CO air pollution concentrations showed that high pollutant concentrations were found in industrial zones especially in Chachoengsao and Rayong provinces. This study emphasized the role of industrial activity in contributing to pollution levels. The temporal trends indicated a decline in CO levels in 2020. This is primarily attributed to reduced industrial activity and transportation during the

COVID-19 pandemic. SO_2 and NO_2 followed similar declining trends in industrial areas.

Future air quality modeling can be significantly enhanced by leveraging higher-resolution satellite data advanced machine learning techniques and expanded environmental datasets. High-resolution satellite imagery offers more granular insights, enabling precise identification of pollution hotspots at the neighborhood or factory level, rather than generalizing pollutant levels across large regions. This finer spatial detail is crucial for urban and industrial pollution management. Advanced machine learning models deep learning could improve the capture of complex spatiotemporal patterns, surpassing the capabilities of XGBoost. Hybrid approaches that integrate physical and machine learning models hold promise for deeper insights into pollutant interactions with environmental factors.

Future research could also bridge air pollution analysis with health data to evaluate the impacts of pollution on public health. By correlating pollutant levels with health outcomes, policymakers can design targeted interventions, enforce emission regulations and raise public awareness, ultimately improving health and well-being in affected regions.

Reference

[1] K. Thongphunchung *et al.*, "Outpatient Department Visits and Mortality with Various Causes Attributable to Ambient Air Pollution in the Eastern Economic Corridor of Thailand," *Int. J. Environ. Res. Public Health*, vol. 19, no. 13, p. 7683, 2022. [Online]. Available: https://www.mdpi.com/1660-4601/19/13/7683

[2] M. Jiang, E. Kim, and Y. Woo, "The Relationship between Economic Growth and Air Pollution—A Regional Comparison between China and South Korea," *Int. J. Environ. Res. Public Health*, vol. 17, no. 8, 2020. [Online]. Available: https://doi.org/10.3390/ijerph17082761

[3] P. R. Shukla, "Construction and its Impact on Air Quality," *Environ. Pollut.*, vol. 242, pp. 616–626, 2018.[Online].Available: https://doi.org/10.1016/j.envpol.2018.07.110

[4] X. Wang, "Urbanization and its Impact on Air Quality," *Urban Climate*, vol. 28, p. 100464, 2019. [Online]. Available: https://doi.org/10.1016/j.uclim.2019.100464

[5] P. M. Mannucci *et al.*, "Effects on health of air pollution: a narrative review," *Intern. Emerg. Med.*, vol. 10, no. 6, pp. 657–662, 2015. [Online]. Available: https://doi.org/10.1007/s11739-015-1276-7

[6] Pollution Control Department, *Thailand National Ambient Air Quality Standards*, Bangkok, Thailand: Ministry of Natural Resources and Environment, 2021. [Online]. Available: https://www.pcd.go.th/

[7] Pollution Control Department, *Manual for Air Quality Measurement Stations*, Bangkok, Thailand: Ministry of Natural Resources and Environment, 2021.

[8] Pollution Control Department, *Standard Methods for Measuring Sulfur Dioxide in Ambient Air*, Bangkok, Thailand, 2021.

[9] Pollution Control Department, *Standard Methods for Measuring Nitrogen Dioxide in Ambient Air*, Bangkok, Thailand, 2021.

[10] Pollution Control Department, *Standard Methods for Measuring Carbon Monoxide in Ambient Air*, Bangkok, Thailand, 2021.

[11] Eastern Economic Corridor Office, *Master Plan for the Development of the Eastern Special Development Zone* (2023–2027), Bangkok, Thailand: EEC Office, 2023.

[12] A. Soleimany, R. Grubliauskas, and V. Šerevičienė, "Application of satellite data and GIS services for studying air pollutants in Lithuania (case study: Kaunas city)," *Air Qual. Atmos. Health*, vol. 14, no. 3, pp. 411–429, 2021. [Online]. Available: https://doi.org/10.1007/s11869-020-00946-z

[13] J. P. Veefkind *et al.*, "TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications," *Remote Sens. Environ.*, vol. 120, pp. 70–83, 2012. [Online].Available: https://doi.org/10.1016/j.rse.2011.09.027

[14] H. Hersbach *et al.*, "The ERA5 global reanalysis," *Q. J. R. Meteorol. Soc.*, vol. 146, no. 730, pp. 1999–2049, 2020. [Online]. Available: https://doi.org/10.100(d)2/qj.3803

[15] Y. Zhan *et al.*, "Spatiotemporal prediction of continuous daily PM2.5 concentrations across China using a spatially explicit machine learning algorithm," *Atmos. Environ.*, vol. 180, pp. 273–282, 2018. [Online]. Available: https://doi.org/10.1016/j.atmosenv.2018.03.031

[16] J. Dou *et al.*, "Estimating ground-level PM2.5 concentrations using random forest and spatiotemporal kriging ensemble models with OMI data," *Remote Sens. Environ.*, vol. 257, p. 112355, 2021.

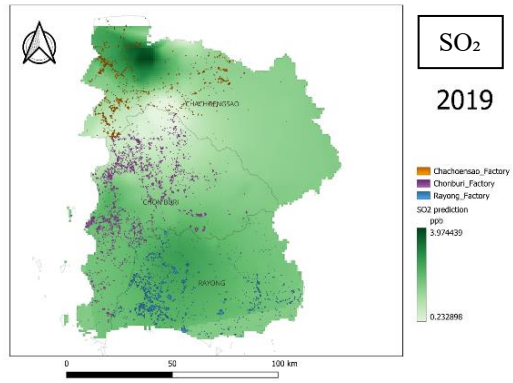
[17] X. Liu *et al.*, "Estimating CO concentrations using MOPITT and Random Forest–STK," *Sci. Total Environ.*, vol. 651, pp. 2390–2400, 2019.

[18] Y. Wang *et al.*, "Estimating urban CO concentrations using Sentinel-5P and LightGBM," *Atmos. Pollut. Res.*, vol. 12, no. 4, p. 101038, 2021.

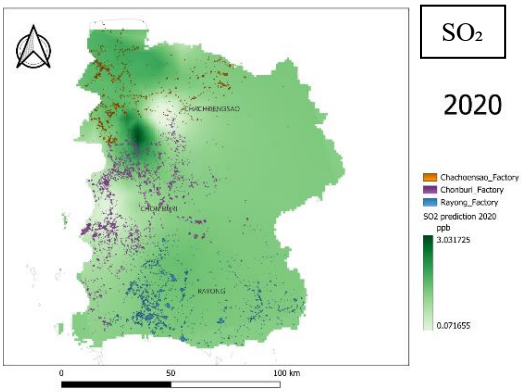
[19] Y. Wei *et al.*, "Estimating air pollutants using Space-Time Extra Trees and big data," *Sci. Total Environ.*, vol. 843, p. 157017, 2023.

[20] S. Ahmad *et al.*, "Predicting NO_2 using nested XGBoost models," *Environ. Monit. Assess.*, vol. 196, no. 2, p. 155, 2024.

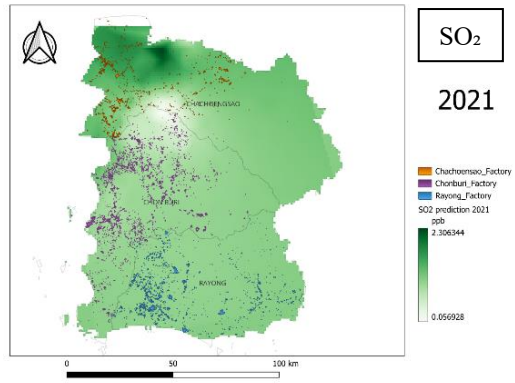
[21] W. Li *et al.*, "Hybrid PCA-Neural Networks model for SO_2 prediction with OMI data," *Atmos. Res.*, vol. 278, p. 106382, 2022.



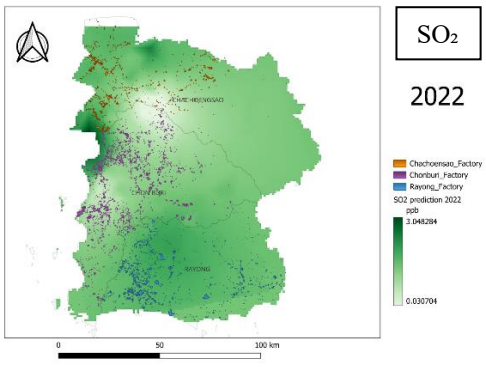
(a)



(b)

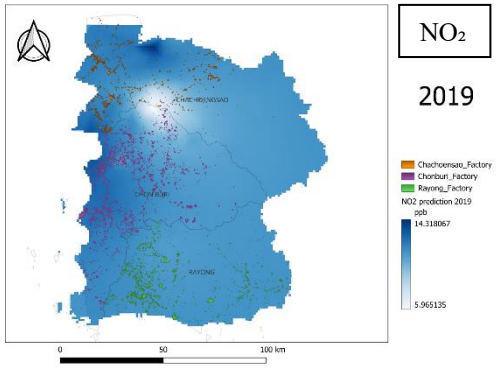


(c)

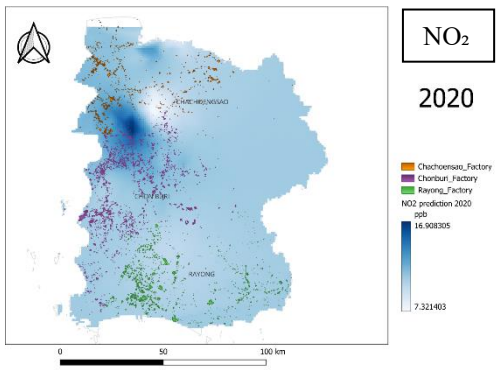


(d)

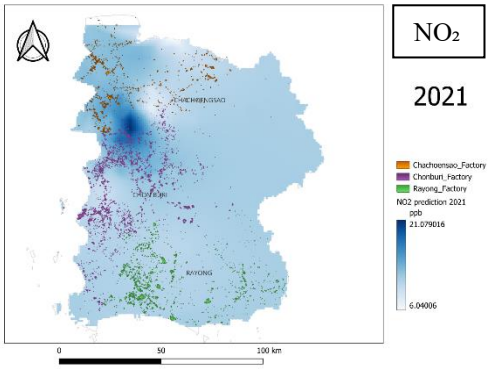
Fig. 6 (a)-(d) SO_2 Concentrations map trends from 2019-2022



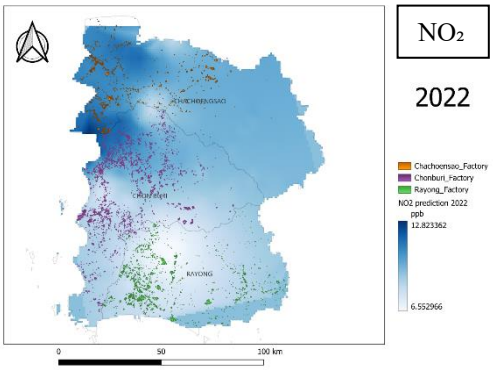
(a)



(b)

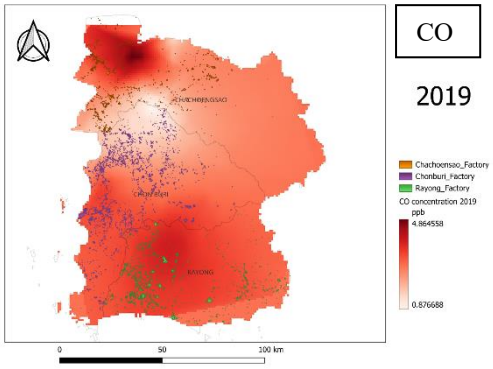


(c)

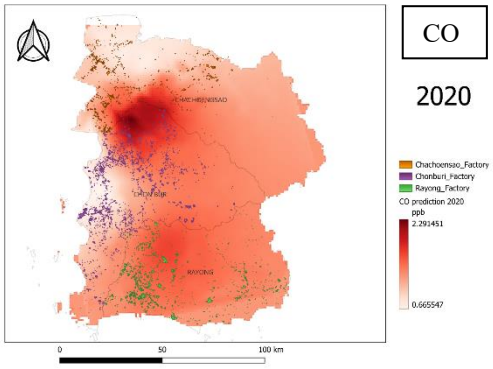


(d)

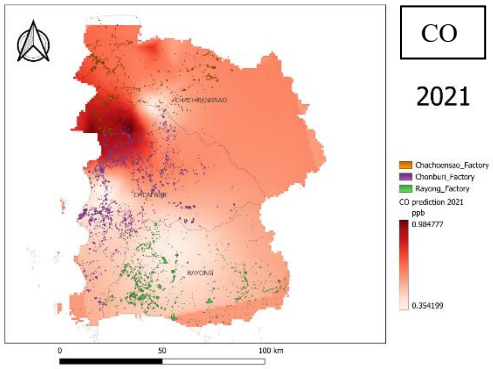
Fig. 7 (a)-(d) NO_2 Concentrations map trends from 2019-2022



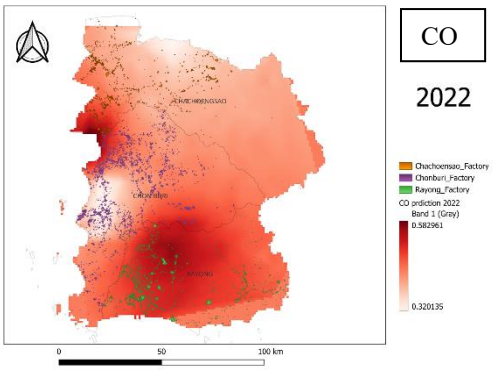
(a)



(b)



(c)



(d)

Fig. 8 (a)-(d) CO Concentrations map trends from 2019-2022