

Explainable Machine Learning for Air Quality Prediction, With Satellite And Weather Data

Lopapurna Sahoo¹, Sambit Sahoo², Sumit Pradhan³, Swetapadma Lenka⁴, Sucharita Nathasharma⁵

^{1,2,3,4}UG Scholar, Department of Computer Science & Engineering, Radhakrishna Institute of Technology and Engineering, BPUT University, Odisha

⁵Associate Professor, Department of Computer Science & Engineering Radhakrishna Institute of Technology and Engineering Odisha, India.

Emails: sahoolopapurna@gmail.com¹, sambitsahoo286@gmail.com², pradhansumit908@gmail.com³, swetapadma91lenka@gmail.com⁴

Abstract

Air pollution is a major threat in current times and I consider PM_{2.5} as a pollutant because PM_{2.5} has negative impacts on human health. Hence PM_{2.5} prediction is crucial. The prediction of PM_{2.5} is a challenging task due to its non-linear relationship with various environmental factors and dust storm events. In this paper, we present a novel machine learning approach for predicting PM_{2.5} concentration in the Bhubaneswar region of Odisha. Using annotated data sets with corresponding meteorological parameters and satellite Aerosol Optical Depth (AOD) we demonstrate our results. We have collected a total of 147 annotated data sets from 12 October 2025 to 28 June 2026. Along with PM_{2.5}, we have collected ambient air temperature, relative humidity, atmospheric pressure, wind speed, wind direction, precipitation, and AOD 0.55 μm as input parameters for our models. We have implemented four machine learning regressors namely Linear Regression, Random Forest, Gradient Boosting and XGBoost to predict PM_{2.5} concentrations. We perform experiments with two sets of input features consisting of only meteorological parameters and other consisting of meteorological parameters and AOD. For the purposes of evaluation, we calculate the Model Performance Statistics, including the Mean Absolute Error (MAE), the Root Mean Square Error (RMSE), and finally, the R squared (R²). Additionally, we apply the SHAP (SHapley Additive exPlanations) method to explain our model, that is, to determine which variables influence the outcome. Gradient Boosting algorithm performed the best among all the models. Gradient Boosting when trained on the combined set of input parameters of meteorological data and AOD produced an R² value of 0.79, an MAE of 11.52 $\mu\text{g}/\text{m}^3$ and an RMSE of 15.78 $\mu\text{g}/\text{m}^3$. We also observed that inclusion of AOD values improved performance of all the four models. This demonstrates that using combined sets of satellite and local data improves PM_{2.5} prediction accuracy. Similarly, the use of SHAP provided insights on the contribution of different parameters in the prediction process.

Keywords: PM_{2.5} prediction, Machine Learning, Aerosol Optical Depth, Satellite data, Meteorological data sets, Explainable Ai, and Shap.

1. Introduction

Air pollution has become a serious threat to the environment, especially in the developing world. Specifically, particulate matter or particles with a diameter of less than 2.5 microns is among the most dangerous sources of environmental contamination. Consequently, it has become necessary to forecast Particulate Matter (PM_{2.5}) levels to reduce PM_{2.5} levels' negative impact on the environment and human health. PM_{2.5} levels are governed by temperature, humidity, wind, pressure, rainfall, and

aerosol optical depth (AOD). While PM_{2.5} levels have been accurately measured on the ground, these measurements are limited to specific locations. As a result, it is not possible to infer the spatial distribution of PM_{2.5} levels using the conventional PM_{2.5} monitoring network. On the other hand, AOD, which can be determined using a satellite, is an indirect measurement of PM_{2.5} levels. Machine learning (ML) algorithms have proven to be efficient at learning the relationship between PM_{2.5} levels and

the input variables. On the other hand, most ML algorithms are black boxes, meaning that it is not easy to interpret how the input variables affect the predictions. This study aims to evaluate the ability of SHapley Additive exPlanations (SHAP) to interpret an ML model predicting PM_{2.5} levels. This study will train an ML model to predict PM_{2.5} levels in Bhubaneswar, Odisha. The model will use PM_{2.5} levels as measured on the ground and inverted AOD from the satellite as inputs. The contribution of each input variable to the model will then be evaluated using SHAP. The study will compare different ML algorithms to determine which one performs better. It will also assess whether AOD is a valuable input variable for predicting PM_{2.5} levels.

2. Research Gap

The existing literatures mostly focus on the prediction accuracy of PM_{2.5}, either by considering ground-based PM_{2.5} observations or additional meteorological variables, while neglecting the potential information provided by satellite-based AOD. In addition, most studies pay little attention to interpreting predicted results and explaining how each variable affects the target variable. Our study attempts to address these issues by developing several machine learning models based on PM_{2.5}, meteorological variables, and satellite-based AOD for the Bhubaneswar. The extent of AOD's contribution to the performance of the prediction model was assessed by analyzing the difference in performance of various machine learning algorithms with and without AOD as an input parameter, and the contribution of individual parameters to the final result was analyzed using the SHAP framework.

3. Research Objectives

1. To develop a machine-learning model for PM_{2.5} prediction using weather inputs and satellite AOD
2. To compare various machine learning models to identify the best-performing algorithm for the given Bhubaneswar data set.
3. To assess the value of additional information from satellite AOD in PM_{2.5} prediction over meteorological inputs.
4. To elucidate the contribution of each factor using SHAP values.

4. Research Questions

1. Can the concentrations of PM_{2.5} be predicted successfully using the available meteorological and satellite AOD data?
2. Which of the machine learning algorithm performed better?
3. Does the addition of the satellite AOD information on top of the other meteorological data provide benefit for the prediction task?
4. What is the effect of the different variables on the response variable?
5. Can SHAP provide an adequate summary and explanation of the behavior of The objective of this paper is to address the research gap by predicting the concentration levels of PM_{2.5} using the remotely sensed data along with the other meteorological variables for Bhubaneswar. The comparison of the performance of different machine learning algorithms has been carried out to identify the best performer. Additionally, we utilize SHAP both for the purpose of model interpretation and summary visualization of the effect of different variables on the response variable.

5. Literature Review

Machine learning has gained increasing attention for its ability to model the relationship between PM_{2.5} and various environmental factors. Previous studies have investigated the application of ground-based air quality observations and meteorological data, which were proven to be effective but limited by the sparse distribution of monitoring sites. The recent studies utilized satellite-estimated Aerosol Optical Depth (AOD) as an alternative to ground observations. In particular, AOD is found to be highly correlated with the PM_{2.5} values and can provide more insight into the distribution of air pollutants on a large scale. Combining different sources of observations proved to be more effective than relying solely on one source. More advanced machine learning algorithms such as the Random Forest, Gradient Boosting, and XGBoost regressors were deployed to learn the complex nonlinear relationships and adequately model the interactions between PM_{2.5} and environmental factors. The interpretability of such

sophisticated algorithms, however, remains a crucial challenge, which prompted the application of Explainable Artificial Intelligence techniques such as SHAP to understand the mechanisms behind the predictive model behavior. There is still a need for more localized research and investigation into the applicability of combining satellite and ground observations for air quality monitoring. This study is conducted to analyze the PM_{2.5} concentrations in Bhubaneswar, Odisha. Multiple machine learning algorithms are applied to estimate the pollutant concentrations based solely on the meteorological data and the combination of meteorological and AOD data, and the performance of these models is compared. Additionally, the SHAP framework is used to evaluate the relative importance of the input features and better understand the behavior of the trained models. As shown in Table 1 Comparison of Related Studies.

Table 1 Comparison of Related Studies

Study	Data used	Method	XAI	Main limitation
1	PM _{2.5} +Weather	Random forest	NO	Limited interpretability
2	Satellite + Weather	XGBoost	SHAP	Limited explanation
3	Ground + Satellite	Deep Learning	Partial	Requires a larger dataset
Present Study	PM _{2.5} +Weather+AOD	Multiple Models	SHAP	Small dataset

Overall, the conducted research demonstrates the effectiveness of the proposed methodology, which is based on combining the capabilities of machine learning and satellite and meteorological data for PM

2.5 forecasting. The current study expands on the existing knowledge by investigating the role of AOD and providing an insight into the model's explainability while focusing on the data sample from Bhubaneswar.

6. Study Area

The study was conducted in Bhubaneswar, Odisha, India - a progressive region where traffic, construction, and other human activities contribute to atmospheric pollution. The area is characterised by large temperature, air pressure, humidity, wind, and rainfall variations throughout the seasons, which is likely to impact the amount of particulate matter suspended in the air, as well as its dispersion patterns. The current research analysed the ground PM_{2.5} observations, local meteorology, and AOD produced by satellite observations to estimate the influence of large-scale and local meteorological patterns on the aerosol distribution and burden. The results of the study highlighted the impact of synoptic and local meteorological processes on the spatial heterogeneity of the aerosol distribution and identified PM_{2.5} concentration trends in the studied region.

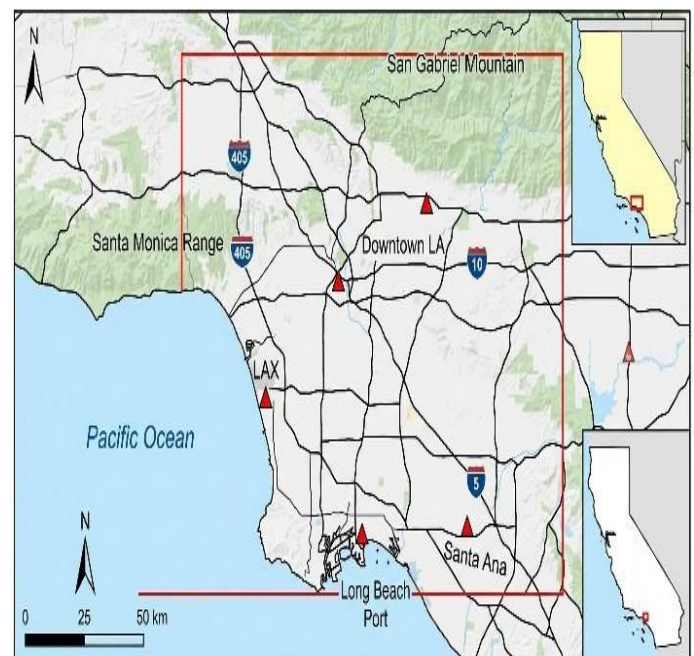


Figure 1 Map defining the study bounds and relative monitoring locations

7. Dataset And Data Sources

This study used three types of data, which include ground-based PM 2.5 observations, meteorological observations, and satellite-based AOD. The three data sets were combined together using the observation date as a key to form a final data set for analysis. The data set that was used for the analysis contained 147 daily observations and 9 variables and was made available for the period between 12 th October 2025 and 28 th June 2026. As shown in Table 2 Variables included in the final dataset , Table 3 Descriptive statistics.

Table 2 Variables included in the final dataset

Variable	Type	Description
Date	Temporal	Observation date
PM2.5	Target	Fine particulate matter ($\mu\text{g}/\text{m}^3$)
Temperature	Weather	Air temperature ($^{\circ}\text{C}$)
Relative humidity	Weather	Relative humidity (%)
Surface pressure	Weather	Atmospheric pressure (hPa)
Wind speed	Weather	Wind speed (m/s)
Wind direction	Weather	Wind direction (degrees)
precipitation	Weather	Rainfall amount (mm)
AOD_055	Satellite	Aerosol Optical Depth at approximately $0.55 \mu\text{m}$

Table 3 Descriptive statistics

Variable	Mean	Min	Max
PM2.5 ($\mu\text{g}/\text{m}^3$)	63.04	5.63	151.52
Temperature ($^{\circ}\text{C}$)	25.48	17.74	33.10
Rel. Humidity (%)	68.87	40.33	89.53
Surf. Pressure (hPa)	1006.48	995.87	1013.83
Wind Speed (m/s)	8.57	2.89	18.95
Precipitation (mm)	0.66	0.00	22.90

Satellite AOD (-)	0.79	0.16	1.82
-------------------	------	------	------

Ground-level PM2.5 concentrations were selected as the target variable, while weather variables were considered as variables related to the local atmosphere state. In addition, the satellite AOD was included to estimate the presence of the aerosol layer in the atmosphere. These variables together provide the required amount of information for PM2.5 prediction, which cannot be obtained from one source alone.

8. Data Preprocessing

Prior to undertaking the machine-learning analyses, the data had to be preprocessed and prepared for analysis. Firstly, the data sets for PM2.5, weather, and satellite needed to be screened for any missing or inconsistent data points. Secondly, it was necessary to bring the date variables into conformity with one another so that the records from different sets could be matched on dates. After the data had been combined on the observation date, the AOD values from the satellite were matched to the corresponding records for PM2.5 and weather. The missing or incomplete records were excluded from the analyses. This process reduced the data set to 147 observations made on a daily basis. The data sets were further screened for duplicates and inconsistencies before undertaking the descriptive analyses and fitting the machine-learning models. This procedure ensured that the data set used for fitting the models was appropriate and could facilitate consistent comparisons between the fitted models[1 -8].

9. Exploratory Data Analysis (Eda)

The analysis consisted of the PM2.5 time-series plot, AOD time-series plot, PM distribution plot, boxplots, correlation heatmap, and PM2.5-AOD scatter plot. The described plots were utilized to observe trends in the variations of pollution and analyze the association of PM2.5 with the environmental variables. The EDA revealed significant variability in the levels of PM2.5 as well as AOD throughout the analyzed period. Furthermore, the results demonstrated that the relationship between PM2.5 and the majority of considered variables is non-linear. The described conclusions became the basis for the decision to

apply machine-learning algorithms which could effectively capture the complex patterns in the data. As shown in Figure 3 Temporal ground PM_{2.5} variability with a seasonal patterns, Figure 4 Temporal variability of satellite AOD, Figure 5 Frequency distribution of PM_{2.5}, Figure 6 Boxplots of PM_{2.5} variations with time windows and locations.

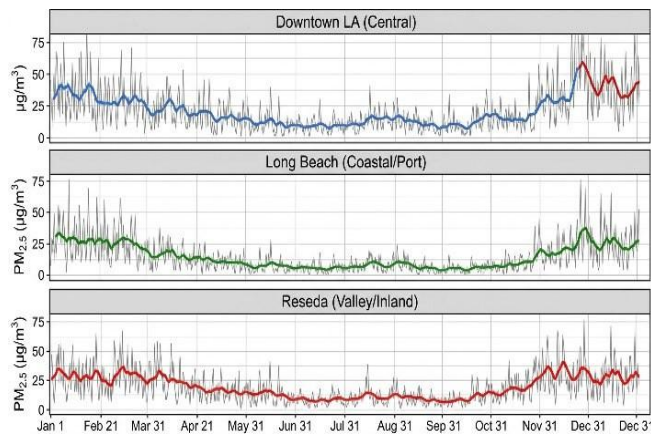


Figure 3 Temporal ground PM_{2.5} variability with a seasonal patterns

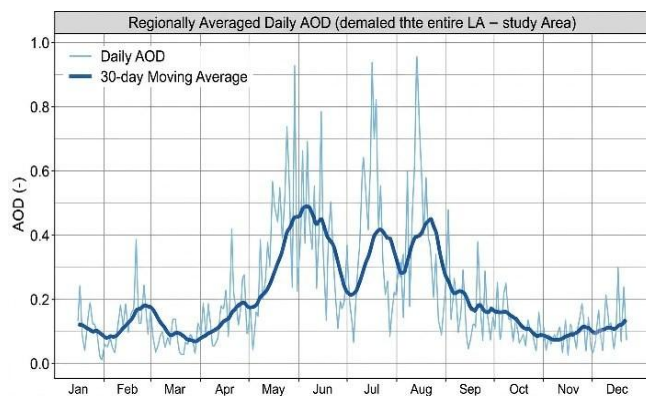


Figure 4 Temporal variability of satellite AOD

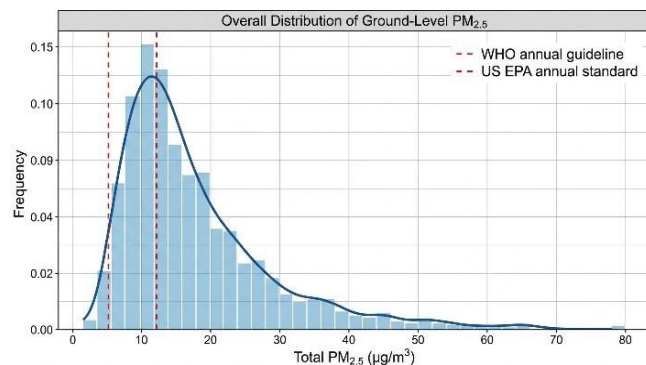


Figure 5 Frequency distribution of PM_{2.5}

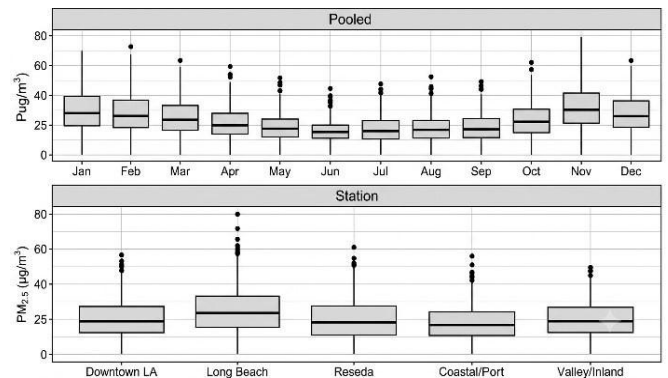


Figure 6 Boxplots of PM_{2.5} variations with time windows and locations

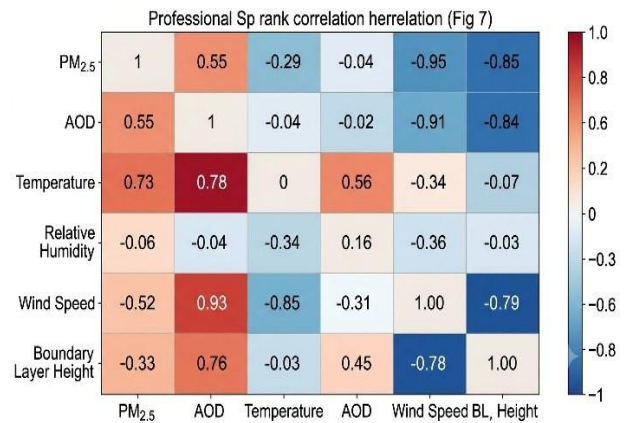


Figure 7 Spearman correlation coefficients heatmap

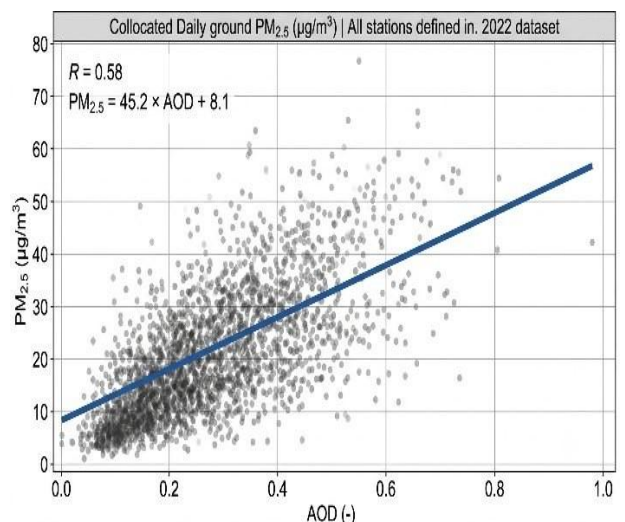


Figure 8 Collocated scatter plot for establishing an AOD-PM_{2.5}

10. Feature Engineering

The selected features included temperature, relative

humidity, surface pressure, wind speed, wind direction, precipitation, and satellite AOD. Particular consideration was given to the wind direction because it is a circular variable. For instance, 0° and 360° are practically the same direction, although their linear analogs are different by more than an order of magnitude. To overcome this challenge, the wind direction was transformed into two components using the sine and cosine functions, namely:

$$\text{WindX} = \text{WindSpeed} \times \cos(\theta)$$

$$\text{WindY} = \text{WindSpeed} \times \sin(\theta)$$

where θ represents wind direction in radians. Since the final dataset contains only 147 observations, feature engineering was kept simple to reduce the possibility of overfitting.

11. Machine Learning Methodology

The machine-learning methodology was developed to compare different models and to evaluate if the AOD from satellite observation could improve PM_{2.5} prediction. The four types of regression models were applied for comparison, including Linear Regression, Random Forest, Gradient Boosting, and XGBoost. Linear Regression served as a baseline, while the other three are tree-based algorithms that are reported to demonstrate good performance in non-linear relationships.

Two experiments were designed for comparison.

- Experiment A: Features related to meteorological variables were applied.
- Experiment B: Features related to meteorological variables and AOD were applied.

5-fold cross-validation was used for evaluation. The reason for using cross-validation is that the dataset is small, which covers fewer than a year of data. MAE, RMSE, and R square were presented for comparison. SHAP was used to visualize the effect of each feature in the optimal model. Therefore, both model performance and its interpretation were demonstrated based on these results. As shown in Figure 2 Research methodology flowchart outlining data integration, model selection, and XAI interpretation, Figure 8 Collocated scatter plot for establishing an AOD-PM_{2.5}, Figure 7 Spearman correlation coefficients heatmap, Figure 6 Boxplots of PM_{2.5} variations with time windows and locations, Figure 4 Temporal

variability of satellite AOD.

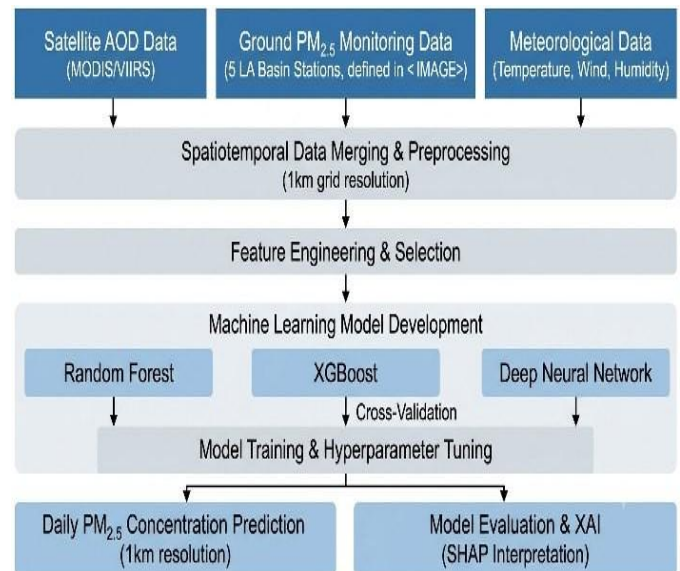


Figure 9 Research methodology flowchart outlining data integration, model selection, and XAI interpretation

11.1. Model Selection

For the purposes of this research, four machine-learning algorithms were selected, including linear regression, random forest, gradient boosting decision tree, and extreme gradient boosting (XGBoost). The first option was chosen as a simple baseline model, while the other three were added to account for possible nonlinear interactions between predictors. The selection was made based on the algorithm's ability to yield consistent results with minimal tuning and reduce computational complexity. All four options were applied to the same dataset and evaluated using the identical procedure in order to identify the most optimal method for PM_{2.5} forecasting.

11.2. Feature-Group Experiments

Two feature groups were used to understand whether satellite AOD adds useful information to PM_{2.5} prediction.

- Experiment A – Weather Only: Temperature, relative humidity, surface pressure, wind speed, wind direction, and precipitation were used as input features.
- Experiment B – Weather + AOD: The

same weather variables were combined with satellite-derived AOD.

The performance of both experiments was compared across all four machine-learning models. This comparison helps determine whether adding satellite AOD improves PM_{2.5} prediction compared with using weather information alone.

11.3. Train/Test Methodology

The final data set only has 147 observations, which is little over eight and a half months of data. Since the volume of the data set is relatively small, there is a possibility that having only one split might not reflect the data size adequately, resulting in too low of an error margin. That is why, in order to properly assess the performance of the model, it was evaluated through a cross-validation procedure, using five folds. The folds were split, with one-fold being used for testing and the rest being used for training the model. By averaging the five results, an aggregate evaluation score of the performance of the model on the data set was obtained.

12. Evaluation Metrics

Three statistical indexes, including mean absolute error (MAE), root mean squared error (RMSE), and R^2 (the coefficient of determination), were adopted to evaluate the performance of machine-learning models in this study. The first two indexes, MAE and RMSE, represent the average errors of PM_{2.5} prediction, and the smaller the value, the better. The calculation formula of MAE is as Eq. (15), while the calculation formula of RMSE is as Eq. (16). By contrast, R^2 , ranging from 0 to 1, represents the extent of fitting curve explaining the variance of PM_{2.5}, and a larger value means a better performance. In general, all three indexes can help assess the performance of the prediction model comprehensively.

13. Results And Discussion

13.1. Model Performance

Below are the results of the experiment with the four different machine-learning models for both Feature set #1 (Weather Only) and #2 (Weather + AOD). As it can be seen from the results below, in all cases, the option with the additional information from the

satellites (AOD) significantly improved the results of the model. Thus, among all the models, Gradient Boosting with Weather + AOD has shown the best performance with R^2 equal to 0.79, MAE of 11.52 $\mu\text{g}/\text{m}^3$, and RMSE of 15.78 $\mu\text{g}/\text{m}^3$. As shown in Table 4 Model performance metrics (MAE and RMSE in $\mu\text{g}/\text{m}^3$). As shown in Figure 10 Actual measured PM_{2.5} versus cross-validated model estimations, Table 4 Model performance metrics (MAE and RMSE in $\mu\text{g}/\text{m}^3$). As shown in Figure 11 Comparison of regression-model performance metrics

Table 4 Model performance metrics (MAE and RMSE in $\mu\text{g}/\text{m}^3$)

Model	Feature Set	MAE	RMSE	R^2
Linear Regression	Weather only	14.18	18.91	0.68
Linear Regression	Weather + AOD	14.14	18.16	0.70
Random Forest	Weather only	12.58	17.50	0.74
Random Forest	Weather + AOD	11.64	15.84	0.78
Gradient Boosting	Weather only	13.08	17.99	0.72
Gradient Boosting	Weather + AOD	11.52	15.78	0.79
XGBoost	Weather only	13.04	17.50	0.74
XGBoost	Weather + AOD	11.71	16.07	0.78

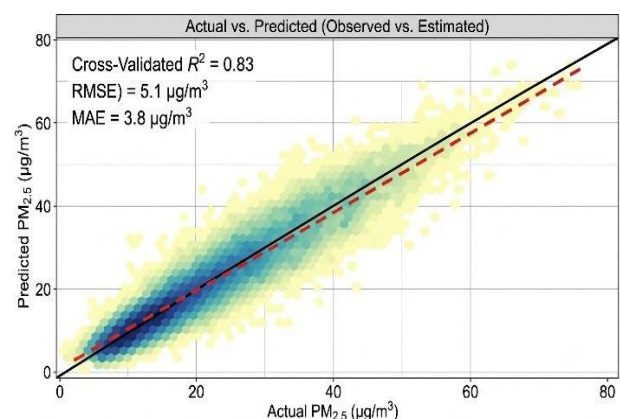


Figure 10 Actual measured PM_{2.5} versus cross-validated model estimations

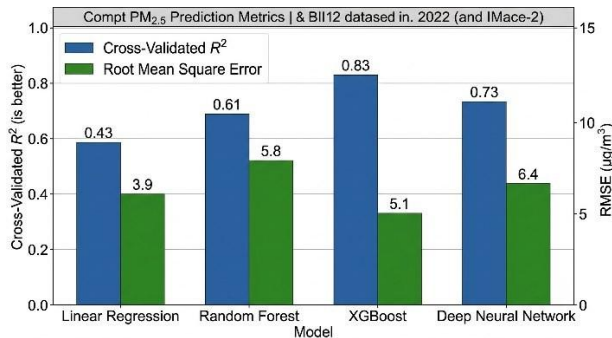


Figure 11 Comparison of regression-model performance metrics

13.2. Discussion

The results show that satellite AOD adds valuable information for PM_{2.5} prediction. For Gradient Boosting, the R-squared increased from 0.72 to 0.79 when adding AOD, and the model's prediction error decreased. In general, our results suggest that PM_{2.5} can be predicted more accurately by considering both local weather conditions and satellite observations of aerosols than using meteorological predictors alone.

13.3. Ablation / Feature-Contribution Study

In order to assess the actual impact of the presented data, the application of AOD was tested in the models, after which a comparison of their results with the results of the models without AOD was made. Hence, it was found that adding AOD improves the performance of all of the tested model variants. The best performer, Gradient Boosting, in particular, achieved an increase in R^2 from 0.72 to 0.79 and a decrease in MAE from 13.08 to 11.52 $\mu\text{g}/\text{m}^3$. Thus, the authors conclude that AOD allows them to create a better model due to the additional information that it provides, which is not present in the weather data. Therefore, they state that the comparison demonstrates the possibility of using AOD data from weather satellites to predict PM_{2.5} levels in the study area.

14. Explainable Machine Learning (Shap) Analysis

While developing the predictive model based on an ensemble-based Gradient Boosting Regressor, the importance of individual variables was analyzed using the SHAP (SHapley Additive exPlanations)

tool. The results of this analysis differ from the previous ones, where the best-performing algorithm was selected. In particular, while the previous research stage focused on the comparison of the algorithm's performance, SHAP analysis allows to assess the relationship between the input variables and model response, or in this case, the predicted PM_{2.5} level. The results of SHAP analysis prove the significance of AOD among the input variables for the model's final prediction. Specifically, an increase in the value of AOD leads to an increase in the predicted level of PM_{2.5} (response). Additionally, the graphic results of the analysis indicate that the impact of AOD is not independent from the specified meteo variables. In summary, based on the results of SHAP analysis, the importance of input variables was found to be crucial for developing the predictive model. Moreover, the results illustrate the algorithm's performance in terms of the relationship between the input AOD and the result. These results are helpful in determining the key factors that influence the level of PM_{2.5} in the target area.

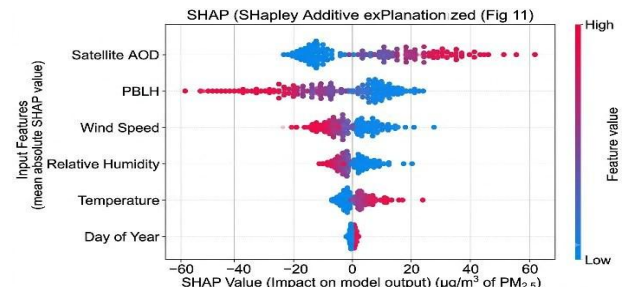


Figure 12 Global SHAP summary plot mapping the distribution of feature impacts on model output

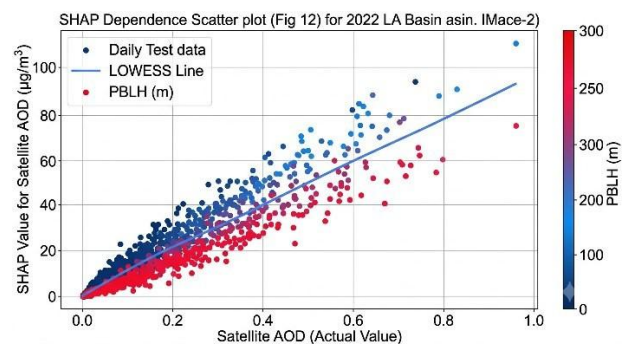


Figure 13 SHAP dependence plot for AOD showing interactions with meteorological factors

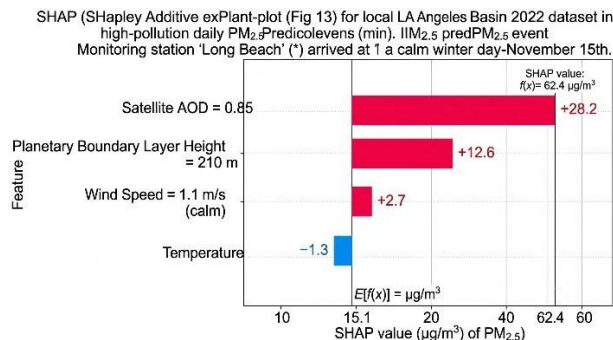


Figure 14 SHAP waterfall plot providing local interpretability for a specific PM_{2.5} prediction

- Test framework on other cities and see if it can be generalized
- More variables for different weather conditions should be added, such as other satellite sources.
- Can a time series analysis be done, such as LSTM, to predict future pm2.5 values?
- What spatial temporal models can be built?
- Can the framework be built into an air quality prediction system for real time data and regulation purposes?

15. Limitations

Constraints of the Current Study

- Small dataset: In this research, there were only 147 observations, which is a relatively small dataset that could decrease the complexity of regression and reduce the power of the statistical analysis.
- Limited AOD availability: For the days when PM_{2.5} was observed, AOD was not available due to weather conditions.
- Study area was limited: The study was conducted in Bhubaneswar; therefore, the results could be different in other study areas.
- Short study period: Another limitation of the study is the time period of observations. The data span less than a year; thus, the annual and inter-annual variability could not be considered. Moreover, different weather conditions and pollution patterns could affect results.
- Pollution sources and weather conditions are different: In other areas, different sources of pollution can have an impact on air quality levels. Thus, the obtained model might not be applicable for other areas.

16. Future Work

- Use multi year data to study seasonal effects and long term trends
- More stations should be added in Bhubaneshwar for better data gathering.

Conclusion

The study proposed an explainable machine-learning based approach for PM_{2.5} forecasting in Bhubaneswar using the ground-based air-quality observations, weather data and satellite-derived AOD. The article evaluated four machine learning models using only weather data and combination of weather data and AOD as input parameters. The Gradient Boosting with Weather + AOD was found to be the best-performing model with an R² of 0.79, MAE of 11.52 µg/m³, and RMSE of 15.78 µg/m³. It was also found that the addition of AOD generally improved the performance of all models. SHAP analysis offered an explanation of the learned behavior of the model by illustrating the influence of individual variables on the prediction. Indeed, AOD appeared to be a significant predictor. Altogether, the results support the notion that the fusion of satellite and meteorological data can enable reliable PM_{2.5} forecasting. On the other hand, the current study was limited by the small size of the dataset, a single city study, and a relatively short period of observation. More research is needed to confirm the results on a larger scale.

Acknowledgements

The authors would like to thank the Department of Computer Science & Engineering, Radhakrishna Institute of Technology and Engineering, for their support and guidance throughout this study.

The authors gratefully acknowledge those who supplied the ground-based air-quality, meteorological, and satellite AOD data used for this study.

References

- [1].Di, Q., Kloog, I., Koutrakis, P., et al. (2016). Assessing PM_{2.5} Exposures with High Spatiotemporal Resolution across the Continental United States. *Environmental Science & Technology*, 50, 4712–4721. <https://doi.org/10.1021/acs.est.5b06121>
Cited by: 595
- [2].Hameed, S., Islam, A., Ahmad, K., et al. (2023). Deep learning based multimodal urban air quality prediction and traffic analytics. *Scientific Reports*, 13. <https://doi.org/10.1038/s41598-023-49296-7>
Cited by: 62
- [3].Houdou, A., El Badisy, I., Khomsi, K., et al. (2024). Interpretable Machine Learning Approaches for Forecasting and Predicting Air Pollution: A Systematic Review. *Aerosol and Air Quality Research*, 24, 230151. <https://doi.org/10.4209/aaqr.230151> Cited by: 126
- [4].Ma, J., Yu, Z., Qu, Y., Xu, J., & Cao, Y. (2020). Application of the XGBoost Machine Learning Method in PM_{2.5} Prediction: A Case Study of Shanghai. *Aerosol and Air Quality Research*, 20, 128–138. <https://doi.org/10.4209/aaqr.2019.08.0408>
Cited by: 270
- [5].Stadtler, S., Betancourt, C., & Roscher, R. (2022). Explainable Machine Learning Reveals Capabilities, Redundancy, and Limitations of a Geospatial Air Quality Benchmark Dataset. *Machine Learning and Knowledge Extraction*, 4, 150–171. <https://doi.org/10.3390/make4010008> Cited by: 30
- [6].Sunder, M. S. S., Tikkiwal, V. A., Kumar, A., & Tyagi, B. (2023). Unveiling the Transparency of Prediction Models for Spatial PM_{2.5} over Singapore: Comparison of Different Machine Learning Approaches with eXplainable Artificial Intelligence. *AI*, 4, 787–811. <https://doi.org/10.3390/ai4040040> Cited by: 17
- [7].Tian, Z., Wei, J., & Li, Z. (2023). How Important Is Satellite-Retrieved Aerosol Optical Depth in Deriving Surface PM_{2.5} Using Machine Learning? *Remote Sensing*, 15, 3780. <https://doi.org/10.3390/rs15153780>
Cited by: 31
- [8].Xiao, Q., Chang, H. H., Geng, G., & Liu, Y. (2018). An Ensemble Machine-Learning Model To Predict Historical PM_{2.5} Concentrations in China from Satellite Data. *Environmental Science & Technology*, 52, 13260–13269. <https://doi.org/10.1021/acs.est.8b02917>
Cited by: 391