ESTIMATION AND PREDICTION OF $PM_{2.5}$ AND PM_{10} CONCENTRATION IN KATHMANDU VALLEY USING REMOTE SENSING AND MACHINE LEARNING

Indra Paudel[1], Ashim Poudel[2], Aashish Baral[3], Aayush Chand[4], Gaurav Baral[5]
[12345] Department of Geomatics Engineering, Pashchimanchal Campus, IOE, Tribhuvan University, Nepal
[1]pas077bge022@wrc.edu.np, [2]pas077bge005@wrc.edu.np, [3]pas077bge002@wrc.edu.np, [4]pas077bge003@wrc.edu.np, [5]pas077bge021@wrc.edu.np

KEYWORDS: $PM_{2.5}$, PM_{10}, Remote Sensing, Air Quality, Machine Learning

ABSTRACT:

In recent years, public health and the environment have been seriously threatened by air pollution, especially in low and middle income nations like Nepal, where the challenge is compounded by sparse ground-level monitoring stations that provide limited spatial coverage. To address this data gap, this study uses machine learning algorithms and satellite data to meet the pressing demand for precise $PM_{2.5}$ and PM_{10} prediction. Ground station data from the monitoring stations of the Department of Environment covering the period from January 2022 to December 2023 were combined with satellite-derived parameters including Aerosol Optical Depth (AOD), Land Surface Temperature (LST), Normalized Difference Vegetation Index (NDVI) and soil moisture to train the models. The Gradient Boosting algorithm performed consistently well, with average R^2 scores of 0.82 and 0.84 for $PM_{2.5}$ and PM_{10} respectively, proving efficient in capturing non-linear relationships in air quality data. Likewise, Random Forest also showed reliable accuracy with R^2 values of 0.80 and 0.82 for $PM_{2.5}$ and PM_{10} respectively. Spatial maps created from model predictions highlight pollution hotspots, providing actionable insights for targeted interventions. The study demonstrates the potential of integrating satellite data with machine learning for air quality monitoring in data-scarce regions, supporting evidence-based environmental policy in Nepal and similar contexts. Future research to incorporate high-resolution datasets and additional metrological variables to enhance model reliability and scalability.

1. INTRODUCTION

Air pollution constitutes the presence of toxic substances in the atmosphere, leading to adverse effects on human health and millions of premature deaths each year. Air pollution became the second leading risk factor for death worldwide, with 8.1 million deaths recorded in 2021, including children under five years of age (UNICEF, 2024). Approximately 91% of the burden falls on low and middle income nations in Southeast Asia and the Western Pacific. Major pollutants include sulphur dioxide (SO_2), particulate matter (PM), nitrogen dioxide (NO_2), carbon monoxide (CO) and ozone (O_2).

Particulate matter (PM) with aerodynamic diameters smaller than 2.5 µm ($PM_{2.5}$) and 10 µm (PM_{10}) has received significant attention due to its impact on human health and ecosystems (Bodor et al., 2022; Ghasemi et al., 2019; Yunesian et al., 2019; Zoran et al., 2020). Elevated PM10 levels are linked to higher hospital admissions for lung and heart diseases, while $PM_{2.5}$ poses a greater health risk as it penetrates deeper into the respiratory system (Mamić et al., 2023).

A report by Nepal's Ministry of Health and Population (2020) attributed 42,100 annual deaths to air pollution, reducing average life expectancy by 4.1 years. More recent global estimates (UNICEF, 2024) record 8.1 million air pollution-related deaths worldwide in 2021, underscoring that the burden remains acute and growing. The same report found that more than 90% of these deaths (7.8 million) were linked to exposure to fine particulate matter (PM2.5) from outdoor and household pollution sources. Air quality monitoring in Nepal faces significant challenges due to the sparse network of ground-based sensors, leaving many areas unmonitored (Timilsina et al., 2023). The Air Quality Index (AQI) is widely used to estimate pollution severity, but its ground-based calculation is limited to monitored locations (Lo Re et al., 2014; Bechle et al., 2013).

Satellite remote sensing data, particularly aerosol optical depth (AOD), offers broad spatial coverage and has been widely used to derive $PM_{2.5}$ and PM_{10} concentrations (Martin, 2008; Xue et al., 2019). Geostatistical methods such as Kriging (Jerrett et al., 2005; Mercer et al., 2011) and regression (Eeftens et al., 2012) have been used, though station coverage remains sparse in sub-urban and rural areas. Advanced machine learning approaches integrating satellite data are enhancing exposure assessments in alignment with WHO air quality guidelines.

This study aims to predict $PM_{2.5}$ and PM_{10} concentrations in the Kathmandu Valley by integrating satellite-derived data with ground-based measurements using two ensemble learning models: Random Forest (RF) and Gradient Boosting (GB). Prior to implementing these models, a Linear Regression approach was also evaluated. However, due to the non-linear nature of the relationship between the predictor variables and particulate matter concentrations, the model produced unsatisfactory results. Hence, these two algorithms were selected for their proven ability to handle non-linear relationships, robustness to outliers, and strong performance in prior PM prediction studies (Mamić et al., 2023; Sakti et al., 2023). The findings provide insights for developing targeted intervention strategies and aid air quality monitoring in regions lacking dense ground based station networks.

2. STUDY AREA

The selected study area is the Kathmandu Valley, spanning an area of 933.73 km². The geographical boundaries are defined by latitudes 27.403° N to 27.818° N and longitudes 85.189° E to 85.566° E. Located in the lesser Himalayas of Central Nepal, the valley is bowl-shaped with a central elevation of 1,425 m, surrounded by four mountain ranges: Shivapuri (2,732 m), Phulchowki (2,695 m), Nagarjun (2,095 m), and Chandragiri (2,551 m), with the Bagmati River flowing through it.

The valley comprises three districts ie Kathmandu, Bhaktapur, and Lalitpur, with population densities of 5,169, 3,631, and 1,433 persons per km² respectively. Its subtropical, semi-humid climate has a mean temperature of 18.3°C and mean annual rainfall of 1,439.7 mm. Pollution originates from vehicular emissions, open fires, construction dust and post-earthquake damage, compounded by the valley's topography which limits wind dispersal.

Kathmandu Valley was selected due to its highest concentration of air quality monitoring stations in Nepal: 7 governmental stations operated by the Department of Environment and 44 Purple Air sensors, providing a substantial dataset for model development. Model performance was evaluated using an 80/20 train-test split, with the held-out 20% test subset serving as the basis for accuracy assessment. Once the model performs reliably well, it is intended for extension to the rest of Nepal incorporating high resolution datasets and additional metrological variables.

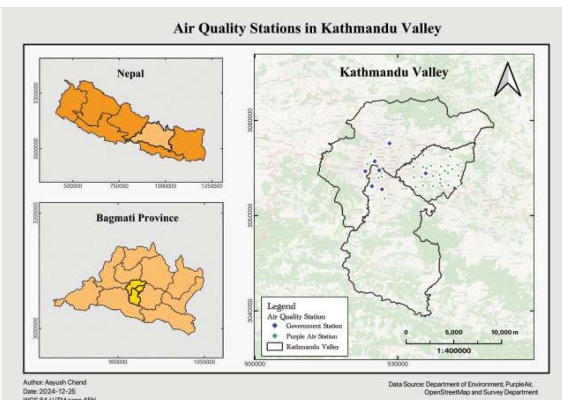

Figure 1. Study Area - Kathmandu Valley

3. MATERIALS AND METHODS

3.1 Description of Datasets

Data from both remote sensing and ground-based observations were utilised, spanning two years from January 1, 2022, to December 31, 2023. Ground level $PM_{2.5}$ and PM_{10} data were sourced from seven monitoring stations of the Department of Environment, Government of Nepal. The two-year period was sufficient to capture seasonal variations, long-term trends and short-term pollution spikes.

Satellite-based AOD data were sourced from the MCD19A2.061 MODIS Terra and Aqua MAIAC product via the Google Earth Engine Catalog, providing daily coverage at 1 km spatial resolution. Additional variables included NDVI from MODIS (1 km, monthly); LST from MODIS MOD11A1 V6.1 (1 km, daily); precipitation from CHIRPS (5.56 km, daily); and soil moisture from SMAP SPL4SMGP/007 (9 km, resampled to 1 km, 3-hour intervals). Precipitation was excluded from model training due to near zero variance (>75% zero values).

Table 1. Description of Datasets (resampled to 1 km)*

Datasets	Source	Type	Spatial Res.	Temporal Res.
Shapefile	Survey Dept.	Vector	-	-
Ground $PM_{2.5}$/PM_{10}	Dept. of Env.	CSV	-	1 Day
AOD	MODIS	Raster	1 KM	1 Day
LST	MODIS	Raster	1 KM	1 Day
NDVI	MODIS	Raster	1 KM	1 Month
Soil Moisture	SMAP	Raster	9 KM*	3 Hour

3.2 Methodology

The study aimed to establish machine learning models to estimate and predict ground-level $PM_{2.5}$ and PM_{10} concentrations in the Kathmandu Valley using heterogeneous satellite and meteorological data.

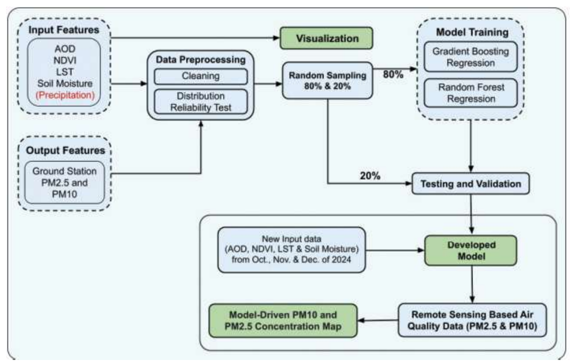

Figure 2. Methodological Flowchart

3.2.1 Data Extraction, Cleaning and Analysis: Ground-based $PM_{2.5}$ and PM_{10} data were collected as daily averages. AOD and LST data shared daily temporal resolution; NDVI values (monthly) were interpolated by assigning the first-day value to all days within ±15 days. Soil moisture values were matched to the closest 3-hour timestamp coinciding with MODIS AOD acquisition. Point values for each remote sensing dataset were extracted for the pixel containing each ground station. Missing values, duplicates and outliers beyond three standard deviations were removed using Python's Pandas library, yielding 987 clean records.

3.2.2 Model Development: Two ensemble algorithms were employed: Random Forest Regression (RFR) and Gradient Boosting Regression (GBR). Random Forest combines the predictions of multiple decision trees, producing a robust model capable of handling complex, non-linear relationships. Gradient Boosting builds models sequentially, with each successive tree correcting the residual errors of the previous one. Hyperparameters for each model were tuned iteratively to avoid underfitting and overfitting. This dataset was split into training and testing subsets. An 80/20 train-test split was applied, a widely adopted ratio in environmental machine learning studies (Mamić et al., 2023; Wu et al., 2024) that balances the need for sufficient training samples against an adequate hold-out set for unbiased performance evaluation. Accordingly, 80% of the records were used for training, while the remaining 20% was reserved for testing purposes, allowing us to validate the models' predictive capabilities with random sampling to ensure representative subsets.

3.2.3 Accuracy Assessment: Model accuracy was assessed using Root Mean Square Error (RMSE), Mean Square Error (MSE), R-squared (R^2), Adjusted R^2, and Mean Absolute Error (MAE). The key metrics are defined as follows:

$$RMSE = \sqrt{[\Sigma(y - \hat{y})^2 / n]} \quad (1)$$

$$MSE = \Sigma(y - \hat{y})^2 / n \quad (2)$$

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

$$Adjusted\ R^2 = 1-(1-R^2)(n-1) / (n-k-1) \quad (4)$$

$$MAE = \Sigma|y - \hat{y}| / n \quad (5)$$

Where,

y = predicted value
$\hat{y}$ = observed value
$\bar{y}$ = mean of observed values
n = number of observations
k = number of predictors
R^2 =Adjusted coefficient of determination

3.2.4 Statistical Distribution of Variables: Prior to model training, the statistical distribution of all predictor and target variables was examined to characterise their behaviour and inter-relationships. Kernel Density Estimation (KDE) plots were produced to visualise the probability distributions of each variable, and boxplots were generated to display central tendency, dispersion, and outliers. Pearson correlation analysis was then applied to quantify linear associations between predictor and target variables, and to assess multicollinearity among predictors. These analyses guided feature selection and informed interpretation of model results.

3.2.5 Visualization of Prediction Results: Trained models were applied to predict $PM_{2.5}$ and PM_{10} concentrations at 935 evenly spaced points (1 km apart) across the valley, using input data collected at 10-day intervals for October, November, and December 2024. The final $PM_{2.5}$ and PM_{10} values were obtained by averaging the outputs from both approaches. The predicted values were then mapped using a color scale ranging from white to red, where white represented lower concentrations and red indicated higher concentrations, effectively capturing spatial pollution variations as shown in figure 14. These visualizations help to highlight pollution hotspots and temporal trends, offering valuable insights into air quality distribution. This approach demonstrates the capability of the model in identifying high risk zones, aiding policymakers in designing targeted air quality management strategies.

4. RESULTS

4.1 Data Visualization

To represent the spatial and temporal variations of air quality and associated environmental parameters, we developed comprehensive maps for each parameter: Aerosol Optical Depth (AOD), Land Surface Temperature (LST), Normalized Difference Vegetation Index (NDVI), and Soil Moisture. These maps were created for four distinct four-month intervals spanning the years 2022 and 2023, providing a seasonal perspective of air quality dynamics in the Kathmandu Valley.

4.1.1 Aerosol Optical Depth (AOD)

The average AOD was extracted from the MCD19A2.061 MODIS Terra and Aqua MAIAC satellite data using the Google Earth Engine Catalog. These datasets, with a 1km spatial resolution, were averaged over the specified intervals to visualize the concentration of aerosols across the study area.

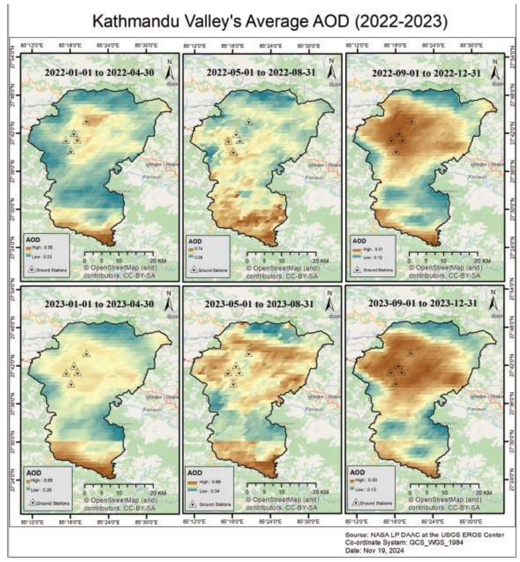

Figure 3. Kathmandu Valley's Average AOD (2022-2023)

4.1.2 Land Surface Temperature (LST)

LST data were derived from the MODIS MOD11A1 V6.1 dataset, which provides daily temperature and emissivity values with a 1km spatial resolution. The data were aggregated to represent average LST for each four-month interval. These maps help correlate temperature variations with aerosol distributions, emphasizing seasonal effects on air quality.

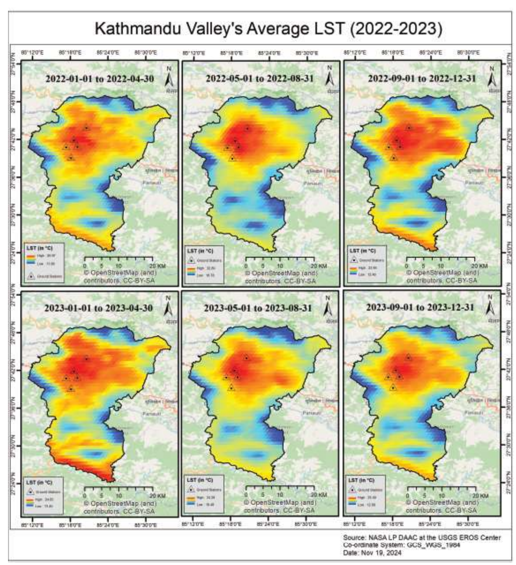

Figure 4. Kathmandu Valley's Average LST (2022-2023)

4.1.3 Normalized Difference Vegetation Index (NDVI)
NDVI data were sourced from MODIS satellite data at a 1km spatial resolution and a monthly temporal resolution. The interval-averaged NDVI maps provide insights into vegetation coverage and health, factors that influence air quality by regulating particulate deposition and secondary aerosol formation.

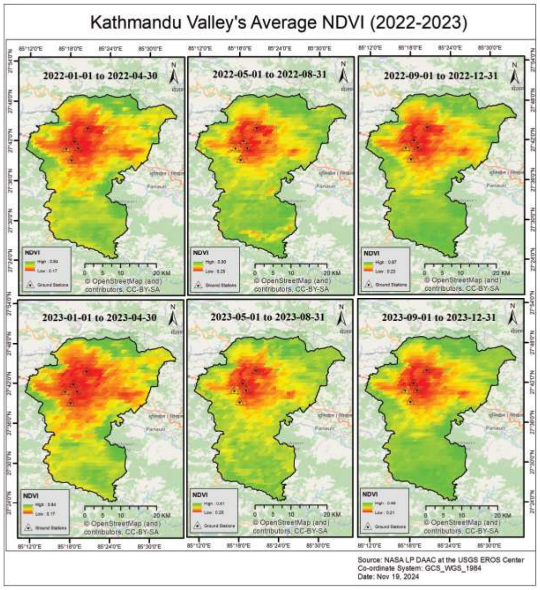

Figure 5. Kathmandu Valley's Average NDVI (2022-2023)

4.1.4 Soil Moisture
Soil moisture data were obtained from the SMAP dataset ("NASA/SMAP/SPL4SMGP/007"), originally available at an 9 km resolution. These data were resampled to a 1km spatial resolution and aggregated for four-month intervals. The resulting maps highlight the influence of soil moisture on surface-level air quality by affecting aerosol resuspension and deposition dynamics.

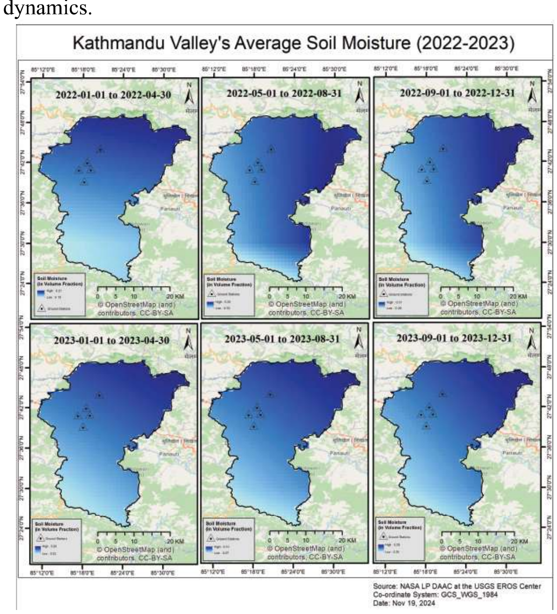

Figure 6. Kathmandu Valley's Average Soil Moisture (2022-2023)

4.2 Statistical Evaluation

As described in Section 3.2.4, statistical evaluation of all predictor and target variables was performed prior to model training. Kernel Density Estimation (KDE) plots revealed that AOD values are positively skewed (0.2–0.7 range), LST follows a near-normal distribution peaking around 27°C, soil moisture also follows a near-normal distribution with most values concentrated between 0.2 and 0.4, and NDVI is moderately skewed (0.2–0.6). PM2.5 values are near-normally distributed (50–150 μg/m³), while PM10 is right-skewed (50–100 μg/m³).

The boxplots illustrate the distribution of environmental variables, highlighting their central tendencies, variability, and outliers. AOD has a median around 0.3 with variability between 0.2 and 0.5, along with several high-value outliers. LST shows a median of approximately 25°C and ranges from 20°C to 30°C, with a few extreme values above 40°C. NDVI values indicate sparse vegetation, with a median near 0.3 and limited higher outliers above 0.6. Precipitation is mostly low, with a median of 0, though occasional heavy rainfall appears as outliers. Soil moisture remains relatively stable around 0.3, with minor variations and some drier outliers. PM10 levels show moderate pollution (median ~75 μg/m³) with occasional severe spikes, while PM2.5 indicates high pollution (median ~100 μg/m³) with some lower-value outliers.

Pearson's correlation analysis showed that AOD has the strongest positive correlation with PM10 ($r = 0.53$) and PM2.5 ($r = 0.49$). Soil moisture exhibited a notable negative correlation with both PM10 ($r = -0.43$) and PM2.5 ($r = -0.43$). NDVI showed a weak negative correlation with PM2.5 ($r = -0.22$). PM2.5 and PM10 were strongly positively correlated ($r = 0.77$). The weak inter-predictor correlations indicated that multicollinearity was not a concern.

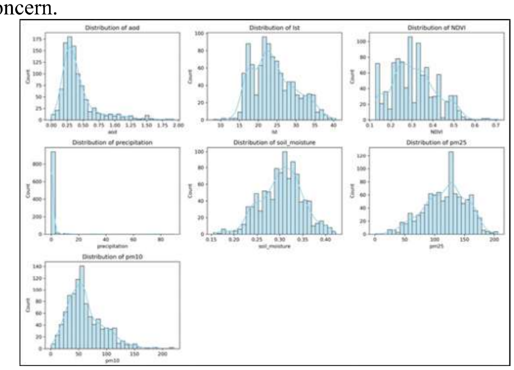

Figure 7. Data Distribution (KDE Plots) of Predictor and Target Variables

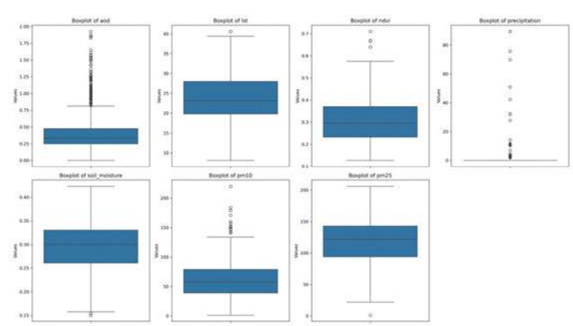

Figure 8. Boxplots of predictor and target variables

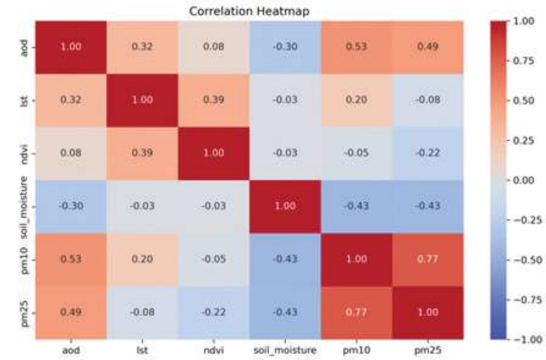

Figure 9. Correlation Heatmap of Predictor and Target Variables

4.3 Random Forest Model Performance Evaluation

The Random Forest model explained 80% of the variance in PM2.5 ($R^2 = 0.80$) with an MAE of 10.98 and RMSE of 15.17. For PM10, the model achieved $R^2 = 0.82$, MAE = 9.32, and RMSE = 15.05. Cross-validation metrics remained stable across all folds, confirming the model's generalization capability.

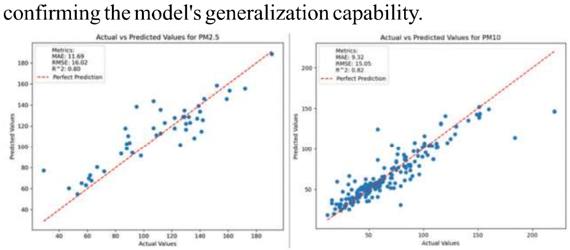

Figure 10. Actual vs. Predicted Values - Random Forest

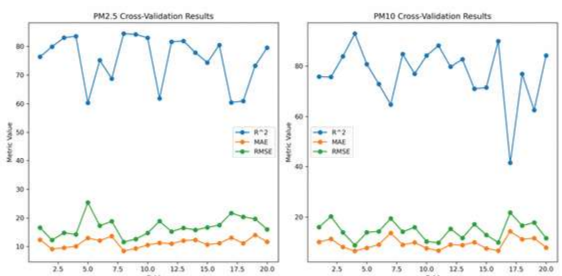

Figure 11. Cross-Validation Results - Random Forest

4.4 Gradient Boosting Model Performance Evaluation

Gradient Boosting achieved $R^2 = 0.82$ (MAE = 10.42, RMSE = 14.13) for PM2.5 and $R^2 = 0.84$ (MAE = 8.16, RMSE = 14.09) for PM10, outperforming Random Forest on all metrics. Cross-validation across 20 folds showed consistently high R^2 values and stable MAE and RMSE, confirming the absence of overfitting. A summary of all accuracy metrics is provided in Table 2.

Table 2. Model Accuracy Metrics (RF = Random Forest; GB = Gradient Boosting)

Metric	RF PM2.5	RF PM10	GB PM2.5	GB PM10
MAE	10.99	9.32	10.42	8.16
MSE	229.97	226.64	199.55	198.40
RMSE	15.17	15.05	14.13	14.09
R^2 Score	0.80	0.82	0.82	0.84
Adj. R^2	0.79	0.82	0.82	0.84

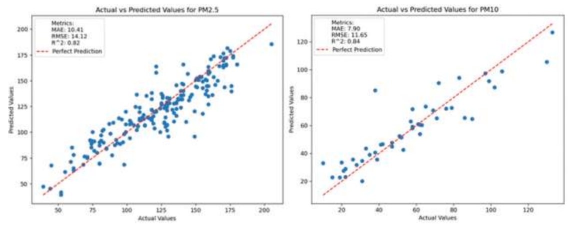

Figure 12. Actual vs. Predicted Values - Gradient Boosting

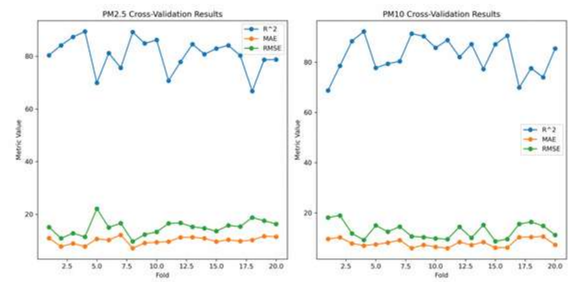

Figure 13. Cross-Validation Results - Gradient Boosting

4.5 PM2.5 and PM10 Concentration Maps

Spatial prediction maps for October to December 2024, at 10 day intervals, reveal PM2.5 concentrations (119.1–158.2 µg/m³) and PM10 concentrations (64.2–108.5 µg/m³) with highest concentration in the central, densely urbanised areas of the valley. These hotspots, concentrated primarily over the Kathmandu metropolitan core (Ratnapark–Sundhara area), Patan in Lalitpur, and the Bhaktapur urban area, are consistent with high vehicular emissions, industrial activity, and urban heating in those districts. Peripheral, less-developed zones (the northern slopes of Shivapuri and the southern foothills near Chandragiri) exhibit markedly lower concentrations.

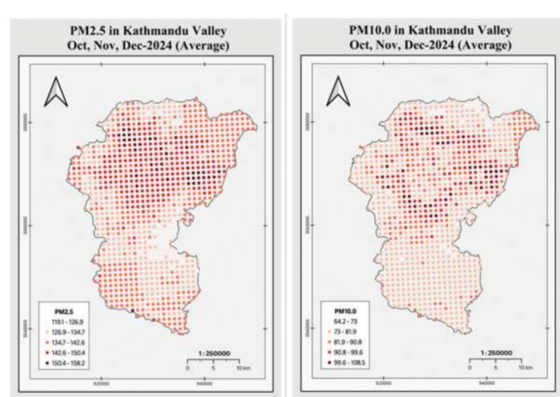

Figure 14. Predicted PM2.5 (left) and PM10 (right) Concentration Maps, Kathmandu Valley

5. DISCUSSION

This study demonstrates that machine learning models integrating satellite-derived parameters and ground-based monitoring stations data can effectively estimate $PM_{2.5}$ and PM_{10} concentrations in a data-scarce urban environment, with both Random Forest and Gradient Boosting models achieving R^2 values above 0.80. These results are consistent with comparable studies in other regions too. Mamić et al. (2023) reported R^2 values of 0.77–0.88 for PM prediction in Europe using similar ensemble methods, while Wu et al. (2024) achieved R^2 of 0.85 for $PM_{2.5}$ estimation in Beijing using satellite data, lending confidence to the methodological approach adopted here.

Gradient Boosting consistently outperformed Random Forest across all metrics for both $PM_{2.5}$ and PM_{10}, consistent with its sequential residual-correction mechanism that allows it to progressively reduce prediction error. This advantage has been noted in other environmental ML studies where the target variable exhibits non-linear, skewed distributions (Sakti et al., 2023). The strong influence of AOD and soil moisture as predictors, compared to LST and NDVI, suggests that aerosol loading and surface moisture conditions are the dominant controls on near-surface PM concentrations in this valley environment. The role of soil moisture as a negative predictor (r = -0.43 with both PM species) likely reflects the suppression of road dust resuspension during and after rainfall periods, a mechanism particularly relevant in the Kathmandu Valley where unpaved roads and construction sites are widespread sources of coarse particulate matter. The relatively weaker signal from LST and NDVI may reflect the coarse monthly resolution of NDVI data and the indirect relationship between surface temperature and aerosol formation at the scale of this study.

Several limitations should be acknowledged. First, the coarse spatial resolution of SMAP soil moisture (9 km, resampled to 1 km) and the monthly temporal resolution of NDVI may introduce interpolation uncertainty at the point-scale used for model training. Second, the 2-year training period (2022–2023), while sufficient to capture seasonal cycles, may not fully represent inter-annual variability driven by exceptional events such as major dust storms or forests fires. Third, key meteorological covariates like wind speed, relative humidity, etc were excluded due to data availability constraints; their inclusion in future iterations could substantially improve performance, particularly for cold-seasons when temperature inversions trap pollutants within the valley. Finally, $PM_{2.5}$ and PM_{10} concentrations at a given location are influenced not only by local emission sources but also by spatial dependence and pollutant transport from surrounding areas. Therefore, restricting the prediction domain strictly to the Kathmandu Valley boundary may introduce boundary effects and underestimate concentrations near the edges of the study area, as neighboring interactions and aerosol transport from outside the valley cannot be fully captured within a hard administrative boundary. Addressing these limitations through higher-resolution inputs, extended training datasets and a spatially buffered prediction domain or spatially continuous prediction grids to better account for cross-boundary aerosol transport represents the primary direction for future work.

6. CONCLUSIONS

This study demonstrates the viability of combining remote sensing data with machine learning for predicting PM2.5 and PM10 concentrations in the Kathmandu Valley, where ground-based monitoring is sparse. Both Random Forest ($R^2 \geq 0.80$) and Gradient Boosting ($R^2 \geq 0.82$) models performed reliably, with Gradient Boosting achieving superior accuracy. Aerosol Optical Depth and soil moisture emerged as the strongest predictors.

The spatial prediction maps identify persistent pollution hotspots in the central valley, providing actionable information for targeted air quality interventions and urban environmental policy. The approach is scalable to the broader national context once additional ground-truth data are available. Future work should incorporate higher-resolution datasets, additional meteorological variables, and longer training periods to improve model reliability and seasonal coverage.

ACKNOWLEDGEMENTS

The authors thank the Department of Environment, Government of Nepal, for providing ground-based PM2.5 and PM10 monitoring data. The authors also acknowledge the support of Asst. Prof. Netra Bahadur Katuwal for constant guidance and supervision throughout the research, and the Department of Geomatics Engineering, Pashchimanchal Campus, IOE, Tribhuvan University, Nepal for providing the platform and resources for this research.

REFERENCES

Arvani, B., Pierce, R.B., Lyapustin, A.I., Wang, Y., Ghermandi, G. and Teggi, S., 2015. High spatial resolution aerosol retrievals for daily PM monitoring over Po Valley, Northern Italy. Atmospheric Chemistry and Physics, 15, pp. 123–155.

Bechle, M.J., Millet, D.B. and Marshall, J.D., 2013. Remote sensing of exposure to NO_2: satellite versus ground-based measurement in a large urban area. Atmospheric Environment, 69, pp. 345–353.

Bodor, K., Szép, R. and Bodor, Z., 2022. The human health risk assessment of particulate air pollution (PM2.5 and PM10) in Romania. Toxicology Reports, 9, pp. 556–562.

Eeftens, M., Beelen, R., de Hoogh, K., Bellander, T., Cesaroni, G., Cirach, M., et al., 2012. Development of land use regression models for PM2.5, PM2.5 absorbance, PM10 and PM(coarse) in 20 European study areas. Environmental Science and Technology, 46(20), pp. 11195–11205.

Filonchyk, M., Peterson, M.P. and Hurynovich, V., 2021. Air pollution in the Gobi Desert region: analysis of dust-storm events. Quarterly Journal of the Royal Meteorological Society, 147, pp. 1097–1111.

Ghasemi, F.F., Dobaradaran, S., Saeedi, R., Nabipour, I., Nazmara, S., Vakil Abadi, D.R., et al., 2019. Levels and ecological and health risk assessment of PM2.5-bound heavy metals in the northern part of the Persian Gulf. Environmental Science and Pollution Research, 27, pp. 5305–5313.

Jerrett, M., Burnett, R.T., Ma, R., Pope, C.A., Krewski, D., Newbold, K.B., et al., 2005. Spatial analysis of air pollution and mortality in Los Angeles. Epidemiology, 16(6), pp. 727–736.

Lo Re, G., Peri, D. and Vassallo, S.D., 2014. Urban Air Quality Monitoring Using Vehicular Sensor Networks. Springer International Publishing, Cham, pp. 311–323.

Mamić, L., Gašparović, M. and Kaplan, G., 2023. Developing PM2.5 and PM10 prediction models on a national and regional scale using open-source remote sensing data. Environmental Monitoring and Assessment, 195.

Martin, R.V., 2008. Satellite remote sensing of surface air quality. Atmospheric Environment, 42(34), pp. 7823–7843.

Mercer, L.D., Szpiro, A.A., Sheppard, L., Lindström, J., Adar, S.D., Allen, R.W., et al., 2011. Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen. Atmospheric Environment, 45(26), pp. 4412–4420.

Mozumder, C., Reddy, K.V. and Pratap, D., 2013. Air pollution modeling from remotely sensed data using regression techniques. Journal of the Indian Society of Remote Sensing, 41, pp. 269–277.

Sakti, A.D., Anggraini, T.S., Ihsan, K.T.N., Misra, P., Trang, N.T.Q., Pradhan, B., et al., 2023. Multi-air pollution risk assessment in Southeast Asia using integrated remote sensing and socio-economic data. Science of the Total Environment, 854, p. 158825.

Timilsina, S., Gautam, P. and Shrestha, K.L., 2023. Relation between MODIS-based aerosol optical depth and particulate matter in Kathmandu using regression model. Journal of Environment Sciences.

UNICEF, 2024. Air pollution accounted for 8.1 million deaths globally in 2021. https://www.unicef.org/rosa/press-releases (accessed June 2024).

Wu, S., Sun, Y., Bai, R., Jiang, X., Jin, C. and Xue, Y., 2024. Estimation of PM2.5 and PM10 mass concentrations in Beijing using GaoFen-1 data at 100 m resolution. Remote Sensing, 16(4).

Xue, T., Zheng, Y., Tong, D., Zheng, B., Li, X., Zhu, T. and Zhang, Q., 2019. Spatiotemporal continuous estimates of PM2.5 concentrations in China, 2000–2016: a machine learning method. Environment International, 123, pp. 345–357.

Yunesian, M., Rostami, R., Zarei, A., Fazlzadeh, M. and Janjani, H., 2019. Exposure to high levels of PM2.5 and PM10 in Tehran and associated health risks during 2016–2017. Microchemical Journal, 150, p. 104174.

Zheng, Z., Yang, Z., Wu, Z. and Marinello, F., 2019. Spatial variation of NO_2 and its impact factors in China: application of Sentinel-5P products. Remote Sensing, 11(16).

Zoran, M.A., Savastru, R.S., Savastru, D.M. and Tautan, M.N., 2020. Assessing the relationship between PM2.5 and PM10 particulate matter impact on COVID-19 in Milan, Italy. Science of the Total Environment, 738, p. 139825.

AUTHOR INFORMATION

Name	:	**Indra Poudel**
Academic Qualification	:	BE Geomatics Engineering
Organization	:	Kent State University
Current Designation	:	Graduate Teaching Assistant Kent State University
Work Experience	:	Geomatics Engineer Intern 2024-2025 May (Ghar Engineering Consultancy)