

Which model to choose? Performance comparison of statistical and machine learning models in predicting PM_{2.5} from high-resolution satellite aerosol optical depth

Padmavati Kulkarni ^{a,*}, V. Sreekanth ^a, Adithi R. Upadhyaya ^b, Hrishikesh Chandra Gautam ^a

^a Center for Study of Science, Technology & Policy, Bengaluru, 560 094, India

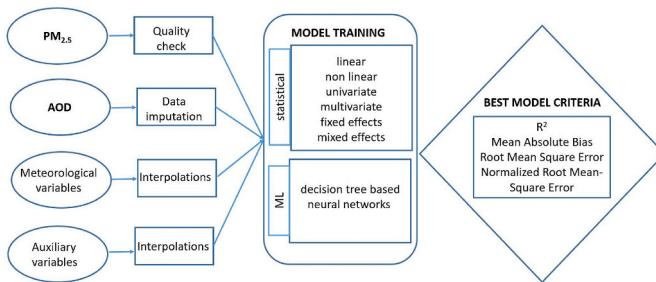
^b ILK Labs, Bengaluru, 560 046, India



HIGHLIGHTS

- Ten models were investigated for their accuracy in predicting PM_{2.5} from AOD.
- Models included linear mixed-effects, Random Forest, Deep Learning, etc.
- Machine learning models performed better than statistical models.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
MODIS-MAIAC
Regression
Machine learning

ABSTRACT

The mathematical solution to estimate surface fine particulate matter (PM_{2.5}) from columnar aerosol optical depth (AOD) includes complex variables and involves a bunch of assumptions. Hence, researchers tend to use training-based models to predict PM_{2.5} from AOD. Here, we integrated regulatory composite PM_{2.5} measurements, high-resolution satellite AOD, reanalysis meteorological parameters, and a few other auxiliary parameters to train ten different regression models. The performance of these (seven statistical and three machine learning) models was evaluated and inter-compared to identify the best performing model. The accuracies of the model predicted PM_{2.5} were quantified based on the coefficient of determination (R^2), mean absolute bias (MAB), normalized root mean square error (NRMSE), and other relevant regression coefficients. The model's performance on unseen data was investigated in terms of 10-fold cross-validation (CV) and Leave-one station-out CV (LOOCV). For this exercise, we considered the case of NCT-Delhi due to: (i) the availability of dense regulatory PM_{2.5} measurements, (ii) the possibility of understanding the model performance over a large range of PM_{2.5} (the daily mean PM_{2.5} values ranged between ~ 4 and 492 $\mu\text{g m}^{-3}$ during the study period), and (iii) the scope of better understanding the influence of extreme meteorological conditions (e.g. the ambient surface temperature varies between ~5 and 40 °C during a calendar year) on the AOD-PM_{2.5} relationship. All the models were trained using data collected for the year 2019 (a non-COVID year). Among models under investigation, Machine

* Corresponding author.

E-mail address: padma.narl@gmail.com (P. Kulkarni).

Learning (ML) models performed better with R^2 , MAB, and NRMSE values for the CV exercises ranging between 0.88 and 0.93, 14.1 and 18.2 $\mu\text{g m}^{-3}$, and 0.18 and 0.23, respectively. The generalizability of the results obtained in this study was discussed.

1. Introduction

Particulate matter (PM) is one of the most commonly tracked and regulated air pollutants globally (WHO, 2006). Fine PM ($\text{PM}_{2.5}$: particulate matter with an aerodynamic diameter less than 2.5 μm) is recognized as an environmental threat to human health (Cohen et al., 2017; Burnett, 2018). Several studies over the last decade continued to show a positive association between $\text{PM}_{2.5}$ exposure and morbidity/mortality (Balluz et al., 2007; West et al., 2016). In India, the air pollution-attributable deaths in the year 2019 were around 1.67 million, among which a major share (0.98 million deaths) was found to be associated with ambient particulate matter pollution (Pandey, 2020). High-resolution $\text{PM}_{2.5}$ spatial maps are highly sought after, for accurate estimation of $\text{PM}_{2.5}$ health impacts. Due to limited regulatory monitoring of $\text{PM}_{2.5}$, especially in low and middle-income countries (Brauer et al., 2019), alternate low-cost methods to produce spatial maps are needed. $\text{PM}_{2.5}$ maps thus generated can also be used for understanding and quantifying $\text{PM}_{2.5}$ spatial gradients and identifying $\text{PM}_{2.5}$ hotspots.

Satellite data-based $\text{PM}_{2.5}$ prediction models are well established (e.g. Christopher and Gupta, 2020) due to the near-global daily coverage of satellite (particularly, polar-orbiting) observations. Aerosol optical depth (AOD), which contains information on aerosols, is the most suitable predictor and is available publicly from a variety of satellite sensors. Earlier studies have shown a good positive correlation between AOD and ground-based $\text{PM}_{2.5}$ (e.g. Wang and Christopher, 2003). Several statistical and machine learning (ML) models predicting $\text{PM}_{2.5}$ from satellite AOD were explored. As there is no gold standard procedure available for choosing the model and predictors, several simple-to-advanced statistical and sophisticated ML models were tried and tested. Statistical models included linear and non-linear regression, linear mixed-effects (LME), generalized additive model (GAM), geographical weighted regression (GWR), etc. (Lee et al., 2011; Sreecanth et al., 2017; Guo et al., 2017; Meng et al., 2018; Maheshwarkar and Raman, 2021). ML models included Neural Networks (NN), Random Forest (RF), Support Vector Machine (SVM), etc. (Liu et al., 2018; Hu et al., 2017; Pak et al., 2019). In addition to AOD, several auxiliary variables (which can influence the AOD- $\text{PM}_{2.5}$ relationship) were also explored as predictors in these models to improve the $\text{PM}_{2.5}$ prediction accuracy (e.g. Mhawish et al., 2020). In the majority of cases, these auxiliary variables included meteorological and land-use parameters.

With the availability of high-resolution satellite AOD (e.g. 1 km \times 1 km AOD product from Moderate Resolution Imaging Spectroradiometer-MODIS), these models can be trained to predict $\text{PM}_{2.5}$ at the same spatial resolution. However, satellite-retrieved AOD has some limitations, including spatial heterogeneity in accuracy and bias in AOD retrieved from different satellite sensors (Farahat, 2019). In addition, missing satellite-AOD due to cloud cover (during monsoon season) can result in less number of data points available for model training and hinders the model from completely understanding the problem scenario. AOD (being an optical quantity) is sensitive to aerosol microphysical properties and to the wavelength at which it is measured, while $\text{PM}_{2.5}$ is not. Also, AOD (being a columnar quantity) accounts for elevated aerosol load too, while $\text{PM}_{2.5}$ accounts for only surface PM load. Polar-orbiting satellite-derived AOD being a snapshot measurement, most of the research studies temporally match the instantaneous AOD data with the hourly mean $\text{PM}_{2.5}$ centered on the satellite overpass time. But for policymakers, predicting daily mean $\text{PM}_{2.5}$ can be more beneficial (national

ambient air pollution standards being set on the daily and annual mean $\text{PM}_{2.5}$). As AOD and $\text{PM}_{2.5}$ observe strong diurnal variations, using instantaneous AOD to predict daily mean $\text{PM}_{2.5}$ can degrade the model performance. To overcome these issues, techniques such as data imputation, daily calibration of AOD- $\text{PM}_{2.5}$ relationship, localized regressions, etc. have been adopted (e.g. Huang et al., 2018), with few underlying assumptions. Still, these predictions are highly helpful by virtue of their cost-effectiveness (compared to monitoring infrastructure) and for producing high-resolution spatial maps.

In this study, we compared the performance of several popular statistical models (including their configuration variants) and machine learning (ML) models in predicting $\text{PM}_{2.5}$. This exercise was carried out using the regulatory $\text{PM}_{2.5}$ measurements, satellite AOD, reanalysis meteorological parameters, and satellite-retrieved land-use proxy, column water vapor, and elevation. The National Capital Territory (NCT, Delhi) of India is considered as the study region and 2019 is the study year. NCT is located in the Indo-Gangetic Plains, which is one of the most highly polluted regions in the world (Guttikunda and Goel, 2013; Beig et al., 2019). The major sectors contributing to the PM pollution load of NCT include transport (exhaust and non-exhaust), domestic, industrial, and transported components (biomass burning and desert dust). The share of each sector to the observed $\text{PM}_{2.5}$ varies seasonally (Guttikunda and Goel, 2013; IITM, 2018). The occurrence of the elevated aerosol layers over NCT during summer and monsoon seasons is also well-understood and reported (Sarkar et al., 2019). The city also experiences severe weather conditions. For the city, $\text{PM}_{2.5}$ data were available from around 38 continuous ambient air quality monitoring stations (CAAQMS). The geographical locations of the monitoring stations from which the $\text{PM}_{2.5}$ data are used for training the statistical and ML models are shown in Fig. 1.

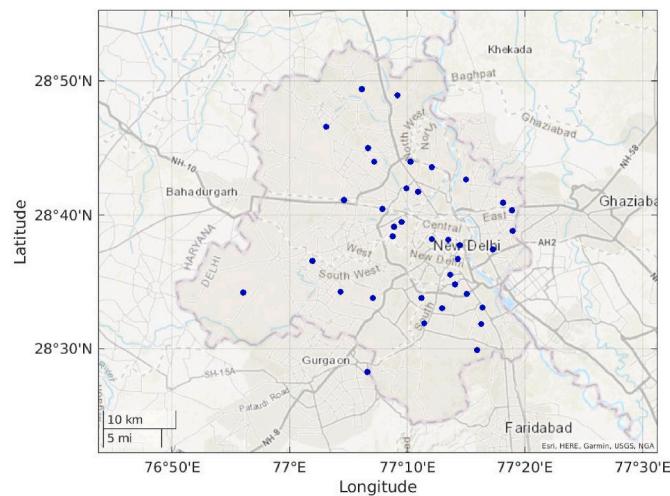


Fig. 1. Geographical locations (blue dots) of the $\text{PM}_{2.5}$ monitoring sites in the National Capital Territory (NCT) region. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

2. Datasets

The datasets used in this study include publicly available pollution control board PM_{2.5}, 1 km MODIS-MAIAC AOD, ERA-Land/ERA-5 meteorological data, 1 km MODIS NDVI and CWV, and ETOPO2 elevation.

2.1. PM_{2.5}

Hourly mean PM_{2.5} data for the study year was acquired from the CAAQMS in NCT, established and maintained by the state and central pollution control boards. At CAAQMS, a USEPA (the United States Environmental Protection Agency)-approved beta attenuation monitor (BAM1020, Met One Instruments, Inc., Grants Pass, USA) is generally used to measure hourly PM_{2.5}. As part of the quality check criteria of the hourly PM_{2.5}, we identified and removed unrealistic (large negative values, fill values, and outliers) PM_{2.5} values. From the hourly values, daily mean PM_{2.5} was calculated only when at least 75% of data (18 hourly data points) was available for that day.

2.2. MODIS-MAIAC AOD

The Multi-Angle Implementation of Atmospheric Correction (MAIAC) algorithm is an advanced and relatively new algorithm for aerosol retrieval from MODIS data over both dark vegetated and bright desert surfaces (Lyapustin et al., 2018). MAIAC aerosol products have several advantages over other operational aerosol products, such as high spatial resolution (1 km), improved retrieval accuracy, and the ability to discriminate between fine and coarse aerosols. In this study, we have used the MAIAC Terra-Aqua combined AOD product (MCD19A2); only AOD data of highest quality (quality flag “clear”) were retained for further analysis. To increase AOD coverage, we imputed AOD based on the predictions from a simple linear univariate (Terra-AOD being the response variable and Aqua-AOD being the predictor) regression model. The model equation was used to predict missing Terra/Aqua AOD when the counterpart was available. To build the model, we used all the Terra-Aqua simultaneous AOD data over the entire study region. We built different models for different days in order to account for any temporal variability in the Terra-AOD and Aqua-AOD relationship. Statistics (correlation and regression coefficients, accuracy metrics) of the models developed are provided in Fig. S1. Models (and thereby, imputation procedures) that resulted in non-significant regression coefficients (which mostly happened when the number of data points available for model training were low) were discarded. For the PM_{2.5} prediction models, we used the Terra-AOD and Aqua-AOD mean value as predictor.

2.3. MODIS NDVI and CWV

Other MODIS-retrieved parameters used in the study include Normalized Difference Vegetation Index (NDVI) and Columnar Water Vapor (CWV). Level-3 gridded MODIS NDVI is produced at 16-day temporal intervals at different horizontal resolutions. NDVI is a normalized transform of the NIR to red reflectance ratio, which varies between -1 and +1. The MODIS vegetation index (VI) algorithm generates each sensor's 16-day NDVI composite. Based on the identical nature of the MODIS sensors aboard Terra and Aqua satellites, combining data from both the sensors can enable higher temporal (8-day) resolution NDVI. Temporal interpolation (spline interpolation technique) was applied on each 1 km grid NDVI to estimate daily NDVI from the 8-day data. MAIAC data also includes column water vapor (CWV, retrieved using MODIS NIR band's data) at 1 km spatial resolution.

2.4. ERA meteorological parameters

Meteorological parameters used in the study include surface (2m) air temperature (ST, K), relative humidity (RH, %), wind speed (WS, ms⁻¹), wind direction (WD,^o), surface (2m) pressure (SP, hPa), and planetary boundary layer height (PBLH, m). Data on these variables were acquired from the ECMWF Reanalysis-5 Land (ERA5-Land) at a native resolution of ~9 km. PBLH was acquired from the ERA-5 product at a native resolution of ~30 km. All these variables were spatially interpolated using the bilinear interpolation technique to match up to the AOD spatial resolution (1 km). More details on ERA-5 meteorological products can be found in Muñoz-Sabater et al. (2021).

2.5. ETOPO elevation

Data on elevation (ELV) was acquired from the global digital elevation model, ETOPO2 of the National Oceanic and Atmospheric Administration (NOAA). ETOPO2 is a 2 arc-minute global DEM, which includes both land topography and ocean bathymetry. The horizontal resolution of ETOPO2 data is ~3.7 km, which was spatially interpolated to a 1 km resolution.

3. Models

3.1. Statistical models

The statistical models trained and validated in the present study include simple linear regression models (univariate and multivariate), linear mixed-effects models, and a generalized additive model. Daily mean PM_{2.5} from ground measurements is considered as the response variable, while daily mean AOD, CWV, ST, RH, SP, WS, WD, PBLH, NDVI, and ELV are considered as predictors for model building. ELV is temporally invariant. Mathematical forms of the different models trained in this study are shown below:

1. Univariate linear regression model (ulr):

$$PM_{2.5\ ij} = a + b_1 \times AOD_{ij} + \epsilon_{ij} \quad (1)$$

2. Multilinear model (mlr):

$$\begin{aligned} PM_{2.5\ ij} = & a + b_1 \times AOD_{ij} + b_2 \times ST_{ij} + b_3 \times RH_{ij} + b_4 \times WD_{ij} + b_5 \\ & \times WS_{ij} + b_6 \times PBLH_{ij} + b_7 \times SP_{ij} + b_8 \times NDVI_{ij} + b_9 \times CWV_{ij} + b_{10} \\ & \times ELV_{ij} + \epsilon_{ij} \end{aligned} \quad (2)$$

3. Linear mixed-effects model (lme1: includes day-specific random slope and random intercept in addition to the fixed slope and intercept in the AOD-PM_{2.5} relationship):

$$PM_{2.5\ ij} = (a + a_{day}) + (b_1 + b_{day}) \times AOD_{ij} + \epsilon_{ij} \quad (3)$$

4. Linear mixed-effects model (lme2: lme1+site specific random intercept):

$$PM_{2.5\ ij} = (a + a_{day}) + (b_1 + b_{day}) \times AOD_{ij} + s_i + \epsilon_{ij} \quad (4)$$

5. Linear mixed-effects model (lme3: lme1+ meteorological and auxiliary parameters as fixed effects):

$$\begin{aligned} PM_{2.5\ ij} = & (a + a_{day}) + (b_1 + b_{day}) \times AOD_{ij} + b_2 \times ST_{ij} + b_3 \times RH_{ij} + b_4 \\ & \times WD_{ij} + b_5 \times WS_{ij} + b_6 \times PBLH_{ij} + b_7 \times SP_{ij} + b_8 \times NDVI_{ij} + b_9 \\ & \times CWV_{ij} + b_{10} \times ELV_{ij} + \epsilon_{ij} \end{aligned} \quad (5)$$

6. Linear mixed-effects model (lme4: lme2+ meteorological and auxiliary parameters as fixed effects):

$$\begin{aligned} PM_{2.5\ ij} = & (a + a_{day}) + (b_1 + b_{day}) \times AOD_{ij} + b_2 \times ST_{ij} + b_3 \times RH_{ij} + b_4 \\ & \times WD_{ij} + b_5 \times WS_{ij} + b_6 \times PBLH_{ij} + b_7 \times SP_{ij} + b_8 \times NDVI_{ij} + b_9 \\ & \times CWV_{ij} + b_{10} \times ELV_{ij} + s_i + \epsilon_{ij} \end{aligned} \quad (6)$$

7. Generalized additive model (gam):

$$\begin{aligned} PM_{2.5\ ij} = & a + f(AOD_{ij}) + f(ST_{ij}) + f(RH_{ij}) + f(WD_{ij}) + f(WS_{ij}) \\ & + f(PBLH_{ij}) + f(SP_{ij}) + f(NDVI_{ij}) + f(CWV_{ij}) + f(ELV_{ij}) + \epsilon_{ij} \end{aligned} \quad (7)$$

In the above equations, i corresponds to the monitoring site location for $PM_{2.5}$, while for the other variables, it represents the nearest (to the corresponding monitoring site) grid. j corresponds to the day. And a is the intercept term and b_1 to b_{10} are the regression coefficients. For lme models, a_{day} and b_{day} correspond to day-specific random intercept and random slope, respectively. ϵ_{ij} is the error term at the site i on day j . s_i is the site-specific random intercept. The f in the gam model represents the smooth function.

3.2. Machine learning (ML) models

Machine Learning has gained great success in predicting $PM_{2.5}$ from satellite aerosol and meteorological parameters in recent years (e.g. Mhawish et al., 2020). In the present study, we trained three established ML models (Random Forest, Deep Learning, eXtreme Gradient Boosting) to understand the $PM_{2.5}$ -AOD relationship and to predict $PM_{2.5}$. We used the open-source machine platform H2O (Luna et al., 2021) to train the ML models. We used the H2O platform to avoid any possible differences due to the choice of programming languages. H2O provides distributed systems and in-memory computing to accelerate ML. Each model is tuned using the respective hyperparameters. The daily mean AOD, CWV, ST, RH, SP, WS, WD, PBLH, NDVI, Latitude, Longitude and ELV are considered as the numeric variables while Julian day is a categorical variable. The workflow is to train a regressor, apply the model, assess the accuracy, and tune the hyperparameters using nested cross-validation. Finally, the model was trained using the whole dataset again using the updated parameters. During the hyperparameter tuning, none of the models failed. The early stopping criteria used for the random grid search is maximum 5400 s, and the stopping metric was Mean Square Error (MSE) with a tolerance of 0.0001. Details on the different ML models trained in this study are given below:

1. Random Forest (RF): RF is a supervised ML algorithm that builds many decision trees, and then, combines them to get a more accurate and stable prediction (Breiman, 2001). In this method, a random subset is considered for splitting a node. Each tree that is built can have additional randomness using random thresholds. Multiple decision trees are modelled using samples drawn from the data set without replacement—also known as the bagging approach. The results from the decision trees are then averaged to give the final result. In this study, the best prediction was achieved when n_{tree} and m_{try} were set to 200 and 7.
2. Deep Learning (DL): DL is a ML method that is based on its ancestor—the artificial neural network. DL is trained with stochastic gradient descent using backpropagation, in other words, mimicking the way a human brain operates to understand underlying relationships and patterns. Due to significant developments in hardware and algorithms, deeper hidden layers with more neurons per layer can be implemented, performing deep neural network modelling. DL is a fairly new technique applied to air quality data. It uses the benefit of multiple layer learning to capture the relationship between input and

output for better prediction. In this study, the activation function of “Tanh” was used and the number of hidden layers was set to 4.

3. eXtreme Gradient Boosting (XGBoost): XGBoost is a successful ML library based on a gradient-boosting algorithm (Chen and Guestrin, 2016) and has better control on overfitting by using a more regularized model strategy (Chen et al., 2015). Another decision tree-based algorithm uses an optimized gradient-boosting method through parallel processing, tree pruning, and regularization to avoid overfitting. Similar to RF, XGBoost is tuned using hyperparameters. The values of the parameters $n_{estimators}$, $subsample$, max_depth , $colsample_bytree$, and min_child_weight used for the final model building are 500, 0.81, 12, 0.9, and 2, respectively.

3.3. Performance evaluation

The accuracy of the predicted daily mean $PM_{2.5}$ by each of the models is quantified using a set of statistical parameters. They include coefficient of determination (R^2), mean absolute bias (MAB), root mean square error (RMSE), and normalized root mean square error (NRMSE). The standard 10-fold cross-validation (CV) exercises were carried out for all the models to investigate the models' performance on unseen data. The equations of the statistical parameters are as shown below.

$$R^2 = \left(\frac{\sum_{i=1}^n P_i O_i - \frac{1}{n} \sum_{i=1}^n P_i \sum_{i=1}^n O_i}{\sqrt{\sum_{i=1}^n P_i^2 - \frac{1}{n} (\sum_{i=1}^n P_i)^2} \sqrt{\sum_{i=1}^n O_i^2 - \frac{1}{n} (\sum_{i=1}^n O_i)^2}} \right)^2 \quad (8)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}} \quad (9)$$

$$MAB = \frac{\sum_{i=1}^n |P_i - O_i|}{n} \quad (10)$$

$$NRMSE = \frac{RMSE}{\frac{1}{n} \sum_{i=1}^n O_i} \quad (11)$$

where P and O are the predictions and observations, respectively, and n is the number of data pairs.

4. Results and discussion

4.1. General statistics of the response variable and predictors

The list of predictors used for the model training, along with details on their spatial and temporal resolutions and their data sources, is summarized in Table 1. The Pearson correlation coefficients among the variables (predictors and response variable) used in the study are shown in Table S1. The daily $PM_{2.5}$ in NCT ranged between $4 \mu g m^{-3}$ and $492 \mu g m^{-3}$. The lowest values were observed during the monsoon season (JJAS; mean \pm standard deviation: $46 \pm 23.2 \mu g m^{-3}$), followed by summer (MAM, $85 \pm 35.3 \mu g m^{-3}$), winter (JF, $165 \pm 85.9 \mu g m^{-3}$), and post-monsoon season (OND, $168 \pm 94.5 \mu g m^{-3}$). Station-wise annual mean $PM_{2.5}$ and the number of data availability days are given in Table S2. Over the study region, the overall mean AOD is 0.67 ± 0.539 , and the seasonal mean values for winter, summer, monsoon, and post-monsoon are 0.63 ± 0.424 , 0.38 ± 0.178 , 0.70 ± 0.476 , and 1.10 ± 0.668 , respectively. Occurrence frequency distribution plots for $PM_{2.5}$ and all the predictors are shown in Fig. S2. After applying the data imputation techniques (for AOD) and quality check criteria, 5577 AOD- $PM_{2.5}$ data pairs were available for model building.

4.2. Performance comparison of the models

Fig. 2 shows the scatter plots (for the 10-fold CV exercise) between the models-predicted daily mean $PM_{2.5}$ and daily mean $PM_{2.5}$ derived

Table 1

The list of predictors used for model training.

Predictor	Data source	Spatial resolution	Temporal resolution	Remarks
Aerosol Optical Depth (AOD)	MODIS satellite sensor (aboard Terra and Aqua)	1 km	Daily	Used linear relationship (between Terra-AOD and Aqua-AOD) to increase the coverage of AOD
Columnar Water Vapor (CWV)	MODIS satellite sensor (aboard Terra and Aqua)	1 km	Daily	
Normalized Difference Vegetation Index (NDVI)	MODIS satellite sensor (aboard Terra and Aqua)	1 km	8-day	Interpolated to daily temporal resolution
2 m temperature (ST)	ERA5-Land	9 km	Hourly	Interpolated to 1 km spatial resolution and matched with the MAIAC AOD grid
Relative Humidity (RH)	Derived from ERA5-Land ST and dew point temperature	9 km	Hourly	
Wind fields (speed (WS) and direction (WD))	Derived from ERA5-Land zonal and meridional wind components	9 km	Hourly	
Planetary Boundary Layer Height (PBLH)	ERA5	30 km	Hourly	
Surface pressure (SP)	ERA5-Land	9 km	Hourly	
Elevation (ELV)	ETOPO2	3.7 km	-	

from hourly measurements. The model performance indicators (R^2 , MAB, RMSE, and NRMSE) are given in [Table 2](#). Normal Quantile-Quantile (QQ) plots of the model residuals are shown in [Fig. 3](#). QQ plots are shown here to examine the distribution of the model residuals. If the points in the plot appear approximately as a straight line, they indicate a normally distributed data. QQ plots also indicate skewness and kurtosis of the dataset. The discussion made further in the manuscript is based on the 10-fold CV results. The ulr model has shown a positive association between AOD and $PM_{2.5}$, but its performance is quite poor (R^2 , RMSE, and NRMSE values being 0.26, $66.6 \mu g m^{-3}$, and 0.56, respectively, [Fig. 2a](#)). But the model regression coefficients ($a = 69.12$; $b_1 = 74.12$) are found statistically significant with p-values < 0.05. The use of the ulr model (AOD being the lone predictor) to predict $PM_{2.5}$ has been also demonstrated by several earlier studies (e.g. [Chu et al., 2003](#); [Wang and Christopher, 2003](#); [Al-Saadi et al., 2005](#)). The ulr model residuals deviated from a normal distribution ([Fig. 3](#)), and exhibited a skewed distribution. This also shows that the model lacks potential predictors. The regression coefficients obtained (and their p-values) for all the statistical models are listed in [Table 3](#). The inclusion of meteorological and land-use proxy to ulr has considerably increased the model performance (mlr, [Fig. 2b](#)). This signifies the influence of meteorology and geographical data on the AOD- $PM_{2.5}$ relationship (see [Xu et al., 2021](#) for more explanation). A similar result has been observed by [Sreekanth et al. \(2017\)](#) over the Indian region, and by [Gupta and Christopher \(2009\)](#) over the contiguous United States. The R^2 value for mlr doubled (0.54) when compared to the R^2 of ulr. The MAB, RMSE, and NRMSE values obtained for the mlr predicted $PM_{2.5}$ are $37.7 \mu g m^{-3}$, $52.2 \mu g m^{-3}$, and 0.44, respectively. All the regression coefficients obtained for mlr are found statistically significant except for ELV (p = 0.71) and the intercept (p = 0.393, see [Table 3](#)). Negative regression coefficients are observed for almost all the predictors (except AOD and SP), indicating that an increase in these variables will result in a decrease in $PM_{2.5}$. The effect of particle microphysical properties on the AOD- $PM_{2.5}$ relationship has not yet been compensated in the ulr and mlr models. As the variables corresponding to the aerosol microphysical properties are difficult and expensive to measure or monitor, the effect can be compensated statistically by introducing a random effect to the mlr. [Lee et al. \(2011\)](#) proposed an lme model with day-specific and

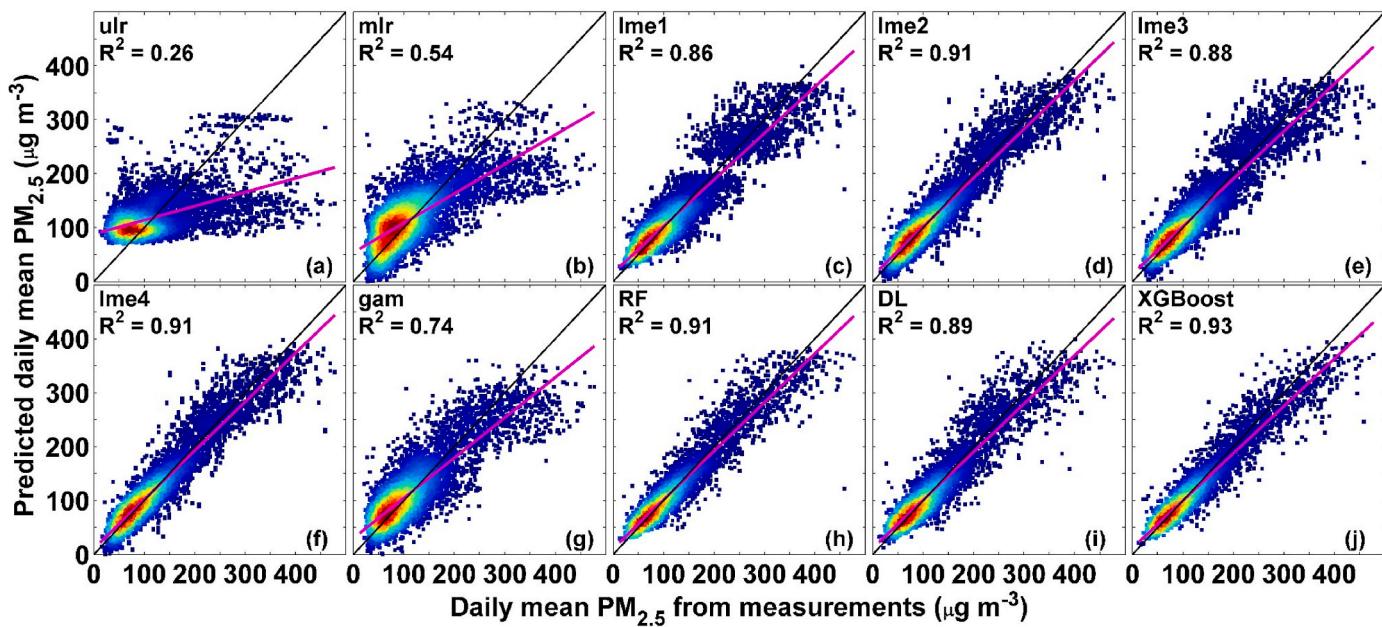
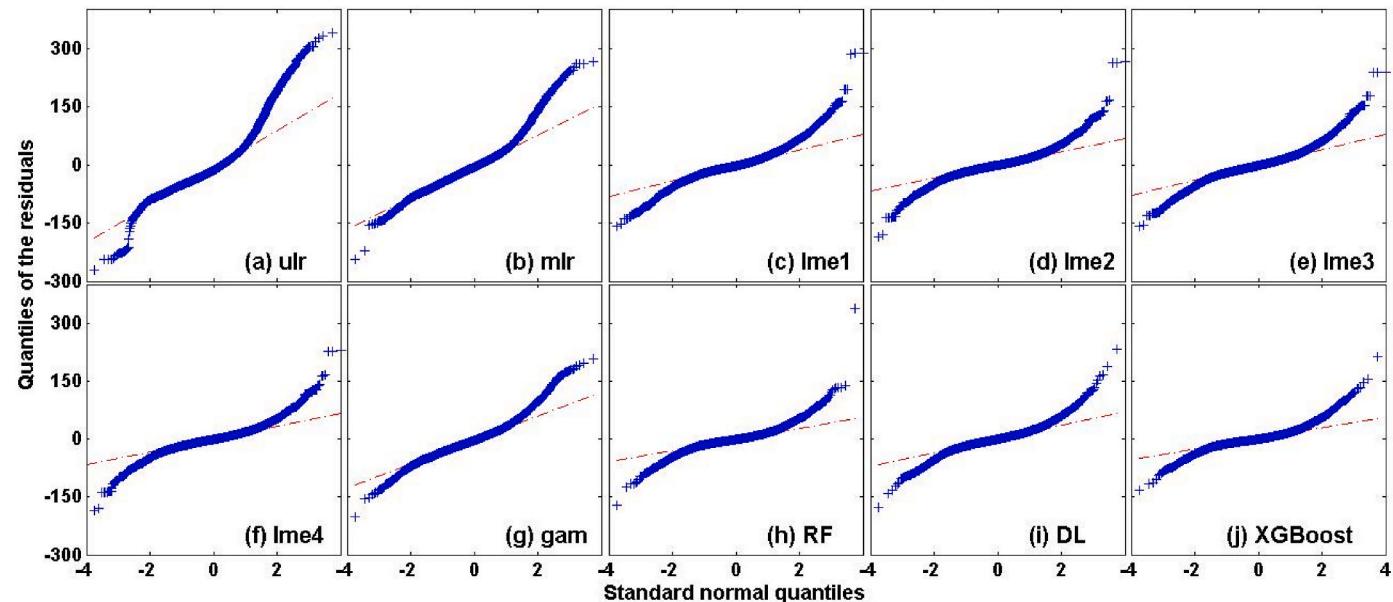


Fig. 2. Density scatter plots of predicted daily mean $PM_{2.5}$ versus daily mean $PM_{2.5}$ from measurements for the different models' 10-fold cross-validation (CV) exercise. The pink and black lines represent the linear regression and 1:1 line, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 2

Model performance indicators for 10-fold cross-validation (CV) exercise.

Models	ulr	mlr	lme1	lme2	lme3	lme4	gam	RF	DL	XGBoost
R ² (linear fit)	0.26	0.54	0.86	0.91	0.88	0.91	0.74	0.91	0.89	0.93
MAB ($\mu\text{g m}^{-3}$)	47.5	37.7	20.0	16.3	18.7	16.3	28.7	15.0	17.6	14.1
RMSE ($\mu\text{g m}^{-3}$)	66.6	52.2	28.8	23.8	26.6	23.7	39.4	22.7	25.8	21.7
NRMSE	0.56	0.44	0.24	0.20	0.22	0.20	0.33	0.19	0.22	0.18

**Fig. 3.** Normal Quantile-Quantile (QQ) plots of the model residuals.**Table 3**

Mean regression coefficients of different statistical models (based on 10-fold CV).

	ulr	mlr	lme1	lme2	lme3	lme4	gam
model-R ²	0.26	0.54	0.88	0.92	0.90	0.92	0.74
Intercept (p-value)	69.12 (0)	-354.99 (0.393)	100.99(0)	107.04 (0)	-11488 (0)	-5326 (0.001)	118.95
AOD coefficient (p-value)	74.12 (0)	77.96 (0)	25.91 (0)	20.07 (0)	18.05 (0)	21.48 (0)	NA (<2e-16)
ST coefficient (p-value)	NA	-3.65 (0)	NA	NA	5.20 (0)	0.18 (0.92)	NA (<2e-16)
RH coefficient (p-value)	NA	-1.98 (0)	NA	NA	0.64 (0.08)	0.03 (0.79)	NA (<2e-16)
WD coefficient (p-value)	NA	-0.03 (0)	NA	NA	0.004 (0.67)	0.009 (0.45)	NA (<2e-16)
WS coefficient (p-value)	NA	-13.22 (0)	NA	NA	-4.85 (0.24)	-12.66 (0.002)	NA (<2e-16)
PBLH coefficient (p-value)	NA	-0.04 (0)	NA	NA	-0.038 (0.03)	0.004 (0.73)	NA (<2e-16)
SP coefficient (p-value)	NA	1.77 (0)	NA	NA	10.16 (0)	5.57 (0)	NA (<2e-16)
NDVI coefficient (p-value)	NA	-137.62 (0)	NA	NA	-90.16 (0)	-17.50 (0.06)	NA (<2e-16)
CWV coefficient (p-value)	NA	-15.60 (0)	NA	NA	-4.11 (0.08)	-6.04 (0.004)	NA (<2e-16)
ELV coefficient (p-value)	NA	-0.02 (0.71)	NA	NA	0.31 (0)	-0.08 (0.62)	NA (<2e-16)

site-specific random effects (in addition to fixed effect, which is the average effect of AOD on $\text{PM}_{2.5}$) in the AOD- $\text{PM}_{2.5}$ linear relation to compensate for the spatio-temporal variations in AOD- $\text{PM}_{2.5}$ relationship. Fig. 2c shows the performance of an lme model (lme1) trained with only AOD and configured with day-specific random effect (in terms of slope and intercept) in the AOD- $\text{PM}_{2.5}$ relationship. Conceptually, this model can compensate for the day-to-day variability in the AOD- $\text{PM}_{2.5}$ relationship, whether it is induced by variations in meteorological variables and/or in aerosol microphysical properties. The lme1 model performance is superior compared to ulr and mlr models. Statistically significant fixed effects coefficients are observed ($a = 100.99$, $b_1 = 25.91$). The random coefficients varied considerably by day, with standard deviations (SD) in a_{day} and b_{day} being 62.48 and 31.09, respectively. When compared to ground truth, the lme1 $\text{PM}_{2.5}$ predictions have an R^2 of 0.86, MAB of $20 \mu\text{g m}^{-3}$, RMSE of $28.8 \mu\text{g m}^{-3}$, and NRMSE of 0.24. lme1 assumes spatially invariant temporal variability in the

AOD- $\text{PM}_{2.5}$ relationship, and yet, its spatial variability is not addressed in any of the above models. To address it, a site-specific random effect (in terms of intercept, also called site bias) is added to lme1 (lme2, Equation (4)), as demonstrated in Lee et al. (2011) and Xie et al. (2015). The lme2 model exhibited further improvement in predicting $\text{PM}_{2.5}$, with an R^2 of 0.91 and NRMSE of 0.2, along with statistically significant fixed coefficients (Table 3). The site bias across the monitoring locations varied between $-30.91 \mu\text{g m}^{-3}$ and $29.62 \mu\text{g m}^{-3}$ ($SD = 15.87 \mu\text{g m}^{-3}$). Even though the spatial variation in the AOD- $\text{PM}_{2.5}$ relationship is adjusted in lme2 for the training data locations, the incorporation of site-specific random effects hinders prediction of $\text{PM}_{2.5}$ over grids with no monitoring sites (unless the site biases are interpolated over the study region). Mhawish et al. (2020) proposed an extended lme model by adding meteorological variables as fixed effects to a day-specific random effects (in AOD- $\text{PM}_{2.5}$ relationship) model. A similar model (lme3) is trained and its prediction performance is shown in Fig. 2e. Its

performance is better than lme1 (but not better than lme2). Finally, a site-bias term is added to lme3 to investigate any further improvement. Models with site-bias terms performed better than other models. The performance of lme4 is very much similar to that of lme2. The addition of meteorological and other auxiliary parameters as fixed effects to lme2 did not improve the model performance. Also, the lme4 model regression coefficients of ST, RH, PBLH, WD, and ELV are found to be statistically non-significant. See Tables 2 and 3 for the model performance indicators, regression coefficients, and their statistical significance. To address any possibility of a non-linear relation between the response variable ($\text{PM}_{2.5}$) and predictors, a gam (Equation (7)) model is trained and investigated for its performance (Fig. 2g). The splines fitted to the individual predictors are shown in Fig. S3. The gam's performance ($R^2 = 0.74$; NRMSE = 0.33) is intermediate between mle and lme models. The residuals of both lme and gam models exhibited fat tails (extreme values).

The performance of the ML models (RF, DL, and XGBoost) is shown in Fig. 2h-j. Overall, the ML model's performance is superior to statistical models. In a similar study over the Middle East, Chau et al. (2021) also found ML as a powerful tool for predicting $\text{PM}_{2.5}$ from AOD and meteorology. Among the three ML models (and among all the ten models investigated in the study), XGBoost-predicted $\text{PM}_{2.5}$ has the highest precision and accuracy ($R^2 = 0.93$ and NRMSE = 0.18), followed by RF and DL. The variable importance sequence as indicated by the RF model is shown in Table 4. Higher the value of the scaled importance (varies between 0 and 1), the higher predictive power of the variable. In other terms, the variable importance of a predictor refers to how much the RF model uses that predictor to make accurate predictions. Surprisingly, AOD predictive power is intermediate. Mhawish et al. (2020) also observed superior performance of the RF model compared to an lme model (their study area being the Indo-Gangetic Plains) in predicting $\text{PM}_{2.5}$ from MAIAC-AOD. In terms of residuals also, the ML models performed better (residual distribution is closer to normal distribution). In a study over China, Bai et al. (2021) observed a better performance by the RF model compared to other ML models (Gradient Boosting, XGBoost, Support Vector Regression) in estimating $\text{PM}_{2.5}$ from Himawari-8 satellite observations.

To investigate the accuracies of the spatial predictions by these models, in addition to the 10-fold CV, we also performed leave-one station-out CV (LOOCV) (except for lme2 and lme4, in which a random spatial effect (site bias) is added). The LOOCV results are shown in the supplementary information (Fig. S4, Table S3). The performance indicators of the models are comparable for the 10-fold CV and LOOCV exercises. ML models are found to perform better than the statistical models (except lme3, whose performance is comparable to XGBoost) for this exercise too. The LOOCV RMSE and MAB values for the ML models varied by less than $2 \mu\text{g m}^{-3}$.

The annual mean gridded spatial $\text{PM}_{2.5}$ maps based on the individual model's spatial predictions are shown in Fig. S5. To derive the annual

mean $\text{PM}_{2.5}$ map, we used only 227 individual day spatial predictions for which the day-specific random coefficients were available. Note that all the lme models were configured with a day-specific random slope and intercept in the AOD- $\text{PM}_{2.5}$ relationship. Also, we spatially interpolated the site bias for the study region to make the spatial predictions by lme2 and lme4 models. From Fig. S5, spatial biases among model predictions can be observed. AOD-only models (ulr and lme1) failed to capture the fine spatial variations in the $\text{PM}_{2.5}$. But most of the models captured the lower values in the southern part (except gam) and higher values in the central and north-western parts of the study region. Further investigations are needed to comprehensively address the observed spatial biases in the model's predictions.

In addition to the choice of the model, the $\text{PM}_{2.5}$ prediction accuracies can also depend on several other factors. They include: (i) spatio-temporal matchup of the response variable ($\text{PM}_{2.5}$) and the predictors, (ii) accuracy of the ground-truth data (Xie et al., 2015), (iii) methods followed for data imputation (mostly in the case of AOD), (iv) identification of multicollinearity in the predictors (relevant in the case of statistical models), (v) the range of variables used for training the model, (vi) choice of auxiliary variables, (vii) generalizability of model coefficients, and (viii) data cleaning methods (extreme values can influence model fitting) etc. The concept of predicting $\text{PM}_{2.5}$ from AOD benefits the most when satellite-AOD is used, due to its spatial coverage. Satellite-AOD, being gridded, represents an average value over a small region. As air pollution can vary rapidly over small scales in space (at least in urban areas and in developing countries; e.g. Apte et al. (2017)), training a model with an area average AOD and a point $\text{PM}_{2.5}$ measurement can introduce some uncertainty in the model predictions. In most cases, for training the models, network regulatory $\text{PM}_{2.5}$ measurements are being used (e.g. Dey et al., 2012). In such cases, researchers have little information on the instrument periodic calibrations, and it needs to be assumed that the instruments are properly maintained. Also, in India, the operational ground-based $\text{PM}_{2.5}$ continuous monitoring stations were limited to urban areas (Brauer et al., 2019), and their density (number of monitors per unit area) is uneven across cities. Due to this, the amount of ground truth $\text{PM}_{2.5}$ data available for model training and validation becomes spatially uneven and can impact the model performance. As the AOD- $\text{PM}_{2.5}$ relationship is sensitive to aerosol microphysical properties, the location of these monitors can play a crucial role in capturing diverse particulate types. In the majority of cases, meteorological parameters are being used as auxiliary variables in model training (e.g. Mhawish et al., 2020). A strong correlation among the meteorological variables can be expected and the inclusion of collinear variables in statistical models can degrade model performance. Due to inherent bias in the satellite-AOD across different sensors, the application of the regression coefficients derived from a model trained using a particular satellite sensor AOD on other sensors' AOD is not often recommended. Most of the models (except gam) trained in this study (and used in general) are linear models. This assumes a linear variation of the response variable with all the predictors. Also, the models need to be trained with a large range of response variables and predictors. If the range of variables used for training and prediction do not overlap, the accuracy of the predictions needs to be considered with caution. The list of challenges and problems in predicting $\text{PM}_{2.5}$ from satellite AOD is also given by Xu et al. (2021).

5. Conclusions

Satellite remote sensing-based estimates/predictions of surface fine particulate matter pollution is fast emerging as a promising technique. Satellite AOD-based mapping of $\text{PM}_{2.5}$ has also been recommended as part of the hybrid monitoring approach for India (Brauer et al., 2019). Several statistical and ML approaches are available in the literature to estimate $\text{PM}_{2.5}$ from AOD. In addition to AOD, these models include various auxiliary/ancillary variables (meteorological, land use, etc.) due to their possible influence on AOD- $\text{PM}_{2.5}$ relationships. Each of these

Table 4
Variable importance sequence (RF model).

Variable	Scaled importance
Julian day	1.000
PBLH	0.157
AOD	0.065
Season	0.054
Latitude	0.026
CWV	0.021
SP	0.021
ST	0.011
ELV	0.011
Longitude	0.011
RH	0.010
WS	0.007
NDVI	0.007
WD	0.006

models has its advantages and limitations. India is diverse in its geography, population and its density, meteorology, and pollution sources, and therefore, large spatio-temporal heterogeneity in the AOD-PM_{2.5} relationship can be expected. Also, the frequent occurrence of dust stratification in higher altitudes during summers in northern India makes the AOD-PM_{2.5} relationship complex. In view of these, we need to account for the day-to-day and spatial variabilities in the AOD-PM_{2.5} relationship to improve the accuracy of the AOD-based PM_{2.5} estimates.

In the present study, we compared and quantified the accuracies of ten (statistical and ML) PM_{2.5} prediction models (AOD, meteorological variables, land use proxy, etc. being the predictors). The results obtained here are specific to the NCT and for the year 2019, and might vary for other regions and time periods. Prior to adopting any particular training-based model for PM_{2.5} predictions from satellite-AOD, we recommend researchers perform a model performance comparison exercise to identify the best suitable model for their geography and study period. The ML models are computationally intense compared to statistical models and can perform better when a large number of predictors are available. The results obtained in the current study are summarized below:

1. A statistically significant positive AOD coefficient is observed for all the statistical models, indicating that an increase in PM_{2.5} can result in enhanced AOD.
2. The addition of meteorological parameters as predictors improved the model performance.
3. Among statistical models, the lme model configured with day- and site-specific random effects performed better (10-fold CV: R² = 0.91, RMSE = 23.7 µg m⁻³).
4. Among all the models under investigation, the XGBoost model exhibited the best performance (10-fold CV: R² = 0.93, RMSE = 21.7 µg m⁻³).

CRediT authorship contribution statement

Padmavati Kulkarni: Conceptualization, Data curation, Formal analysis, Writing – original draft. **V. Sreekanth:** Conceptualization, Methodology, Writing – review & editing. **Adithi R. Upadhyा:** Software. **Hrishikesh Chandra Gautam:** Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors acknowledge the Central Pollution Control Board (CPCB) and Delhi Pollution Control Committee (DPCC) for installing and maintaining the CAAQMS and making the data open access. MODIS satellite data have been obtained from the National Aeronautics and Space Administration (NASA) Ladsweb data archive (<https://ladsweb.modaps.eosdis.nasa.gov/search/order>) and we acknowledge the MODIS team for developing the aerosol products. ERA-5 data have been obtained from the Climate Data Store (CDS) of the Copernicus Climate Change Service (C3S) (<https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=overview>). We also acknowledge the European Centre for Medium-Range Weather Forecasts (ECMWF), CDS, and C3S teams for the data. We thank Dr. Erin LeDell and Abhiram R for their guidance on the Machine Learning packages. We acknowledge the two anonymous reviewers for their valuable suggestions and comments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.atmosenv.2022.119164>.

References

- Al-Saadi, J., Szykman, J., Pierce, R.B., Kittaka, C., 2005. Improving national air quality forecasts with satellite aerosol observations. Bull. Am. Meteor. Soc. 86 (9), 1249. <https://doi.org/10.1175/BAMS-86-9-1249>.
- Apte, J.S., Messier, K.P., Gani, S., Brauer, M., Kirchstetter, T.W., Lunden, M.M., Marshall, J.D., Portier, C.J., Vermeulen, R.C., Hamburg, S.P., 2017. High-resolution air pollution mapping with Google street view cars: exploiting big data. Environ. Sci. Technol. 51 (12), 6999–7008. <https://doi.org/10.1021/acs.est.7b00891>.
- Balluz, L., Wen, X.-J., Town, M., Shire, J.D., Qualter, J., Mokdad, A., 2007. Ischemic heart disease and ambient AirPollution of particulate matter 2.5 in 51 counties in the U.S. Publ. Health Rep. 122, 626–633. <https://doi.org/10.1177/003335490712200510>.
- Bai, H., Zheng, Z., Zhang, Y., Huang, H., Wang, L., 2021. Comparison of satellite-based PM_{2.5} estimation from aerosol optical depth and top-of-atmosphere reflectance. Aerosol Air Qual. Res. 21, 20257. <https://doi.org/10.4209/aaqr.2020.05.0257>.
- Beig, G., et al., 2019. Anatomy of the winter 2017 air quality emergency in Delhi. Sci. Total Environ. 681, 305–311. <https://doi.org/10.1016/j.scitotenv.2019.04.347>.
- Brauer, M., Guttikunda, S.K., Nishad, K.A., Dey, S., Tripathi, S.N., Weagle, C., Martin, R.V., 2019. Examination of monitoring approaches for ambient air pollution: a case study for India. Atmos. Environ. 216, 116940. <https://doi.org/10.1016/j.atmosenv.2019.116940>.
- Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Burnett, R., 2018. Global Estimates of mortality associated with long term exposure to outdoor fine particulate matter. Proc. Natl. Acad. Sci. U.S.A. 115, 9592–9597. <https://doi.org/10.1073/pnas.1803222115>.
- Chau, K., Franklin, M., Lee, H., Garay, M., Kalashnikova, O., 2021. Temporal and spatial autocorrelation as determinants of regional AOD-PM_{2.5} model performance in the Middle East. Rem. Sens. 13 (18), 3790. <https://doi.org/10.3390/rs13183790>.
- Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., 2015. Xgboost: extreme gradient boosting. R Package version 0.4–2, 1–4.
- Chen, T., Guestrin, C., 2016. XGBoost. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794. <https://doi.org/10.1145/2939672.2939785>.
- Christopher, S., Gupta, P., 2020. Global distribution of column satellite aerosol optical depth to surface PM_{2.5} relationships. Rem. Sens. 12 (12), 1985. <https://doi.org/10.3390/rs12121985>.
- Chu, D.A., Kaufman, Y.J., Zibordi, G., Chern, J.D., Mao, J., Li, C., Holben, B.N., 2003. Global monitoring of air pollution over land from the earth observing system-terra moderate resolution imaging spectroradiometer (MODIS). J. Geophys. Res. 108, D21. <https://doi.org/10.1029/2002JD003179>.
- Cohen, A.J., et al., 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. Lancet 389, 1907. [https://doi.org/10.1016/S0140-6736\(17\)30505-6](https://doi.org/10.1016/S0140-6736(17)30505-6).
- Dey, S., Di Girolamo, L., van Donkelaar, A., Tripathi, S.N., Gupta, T., Mohan, M., 2012. Variability of outdoor fine particulate (PM_{2.5}) concentration in the Indian Subcontinent: a remote sensing approach. Rem. Sens. Environ. 127, 153–161. <https://doi.org/10.1016/j.rse.2012.08.021>.
- Farahat, A., 2019. Comparative analysis of MODIS, MISR, and AERONET climatology over the Middle East and north Africa. Ann. Geophys. 37, 49–64. <https://doi.org/10.5194/angeo-37-49-2019>.
- Guo, Y., Tang, Q., Gong, D.Y., Zhang, Z., 2017. Estimating ground-level PM_{2.5} concentrations in Beijing using a satellite-based geographically and temporally weighted regression model. Remote Sens. Environ. 198, 140–149. <https://doi.org/10.1016/j.rse.2017.06.001>.
- Gupta, P., Christopher, S.A., 2009. Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: multiple regression approach. J. Geophys. Res. 114, D14205. <https://doi.org/10.1029/2008JD011496>.
- Guttikunda, S.K., Goel, R., 2013. Health impacts of particulate pollution in a megacity—Delhi, India. Environ. Dev. 6, 8–20. <https://doi.org/10.1016/j.envdev.2012.12.002>.
- Hu, X., Belle, J., Meng, X., Wildani, A., Waller, L.A., Strickland, M.J., Liu, Y., Kelly, J.T., Jang, C.J., Timin, B., et al., 2017. Estimating PM_{2.5} concentrations in the conterminous United States using the random forest approach. Environ. Sci. Technol. 51, 6936–6944. <https://doi.org/10.1021/acs.est.7b01210>.
- Huang, K., Xiao, Q., Meng, X., Geng, G., Wang, Y., Lyapustin, A., Liu, Y., 2018. Predicting monthly high-resolution PM_{2.5} concentrations with random forest model in the North China Plain. Environ. Pollut. 242, 675–683. <https://doi.org/10.1016/j.envpol.2018.07.016>.
- Indian Institute of Tropical Meteorology, Ministry of Earth Sciences, Government of India, 2018. SAFAR-high Resolution Emission Inventory of Delhi Mega City—2018 (SAFAR-Delhi-2018-A). Pune.
- Lee, H.J., Liu, Y., Coull, B.A., Schwartz, J., Koutrakis, P., 2011. A novel calibration approach of MODIS AOD data to predict PM_{2.5} concentrations. Atmos. Chem. Phys. 11 (15) <https://doi.org/10.5194/acp-11-7991-2011>.

- Liu, Y., Cao, G., Zhao, N., Mulligan, K., ye, Xinyue, 2018. Improve ground-level PM_{2.5} concentration mapping using a random forests-based geostatistical approach. Environ. Pollut. 235 <https://doi.org/10.1016/j.envpol.2017.12.070>.
- Luna, A., Bernanke, J., Kim, K., Aw, N., Dworkin, J.D., Cha, J., Posner, J., 2021. Maturity of gray matter structures and white matter connectomes, and their relationship with psychiatric symptoms in youth. Hum. Brain Mapp. 42 (14), 4568–4579. <https://doi.org/10.1002/hbm.25565>.
- Lyapustin, A., Wang, Y., Korkin, S., Huang, D., 2018. MODIS Collection 6 MAIAC algorithm. Atmos. Meas. Tech. 11 (10) <https://doi.org/10.5194/amt-11-5741-2018>.
- Maheshwarkar, P., Raman, R., 2021. Population exposure across central India to PM_{2.5} derived using remotely sensed products in a three-stage statistical model. Sci. Rep. 11 <https://doi.org/10.1038/s41598-020-79229-7>.
- Meng, X., Garay, M.J., Diner, D.J., Kalashnikova, O.V., Xu, J., Liu, Y., 2018. Estimating PM_{2.5} speciation concentrations using prototype 4.4 km-resolution MISR aerosol properties over Southern California. Atmos. Environ. 181, 70–81. <https://doi.org/10.1016/j.atmosenv.2018.03.019>.
- Mhawish, A., Banerjee, T., Sorek-Hamer, M., Bilal, M., Lyapustin, A.I., Chatfield, R., Broday, D.M., 2020. Estimation of high-resolution PM_{2.5} over the Indo-Gangetic Plain by fusion of satellite data, meteorology, and land use variables. Environ. Sci. Technol. 54 (13), 7891–7900. <https://doi.org/10.1021/acs.est.0c01769>.
- Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D., Piles, M., Rodríguez-Fernández, N., Zsoter, E., Buontempo, C., Thepaut, J., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data Discuss. 1–50.
- Pak, U., Ma, J., Ryu, U., Ryom, K., Juhyok, U., Pak, K., Pak, C., 2019. Deep learning-based PM_{2.5} prediction considering the spatiotemporal correlations: a case study of Beijing, China. Sci. Total Environ. 699, 133561. <https://doi.org/10.1016/j.scitotenv.2019.07.367>.
- Pandey, A., et al., 2020. Health and economic impact of air pollution in the states of India: the Global Burden of Disease Study 2019. Lancet Planet. Health. [https://doi.org/10.1016/S2542-5196\(20\)30298-9](https://doi.org/10.1016/S2542-5196(20)30298-9).
- Sarkar, S., Chauhan, A., Kumar, R., Singh, R., 2019. Impact of deadly dust storms (may 2018) on air quality, meteorological, and atmospheric parameters over the northern parts of India. GeoHealth 3. <https://doi.org/10.1029/2018GH000170>.
- Sreekanth, V., Mahesh, B., Niranjan, K., 2017. Satellite remote sensing of fine particulate air pollutants over Indian mega cities. Adv. Space Res. 60 (10), 2268–2276. <https://doi.org/10.1016/j.asr.2017.08.008>.
- Wang, J., Christopher, S.A., 2003. Intercomparison between satellite-derived aerosol optical thickness and PM_{2.5} mass: implications for air quality studies. Geophys. Res. Lett. 30 (21) <https://doi.org/10.1029/2003GL018174>.
- West, J.J., Cohen, A., Dentener, F., Brunekeef, B., Zhu, T., Armstrong, B., Bell, M., Brauer, M., Carmichael, G.R., Costa, D.L., et al., 2016. What we breathe impacts our health: improving understanding of the link between air pollution and health. Environ. Sci. Technol. 50, 4895–4904. <https://doi.org/10.1021/acs.est.5b03827>.
- WHO, 2006. WHO Air Quality Guidelines for Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide: Summary of Risk Assessment. World Health Organization.
- Xie, Y., Wang, Y., Zhang, K., Dong, W., Lv, B., Bai, Y., 2015. Daily estimation of ground-level PM_{2.5} concentrations over Beijing using 3 km resolution MODIS AOD. Environ. Sci. Technol. 49 (20), 12280–12288. <https://doi.org/10.1021/acs.est.5b01413>.
- Xu, X., Zhang, C., Liang, Y., 2021. Review of satellite-driven statistical models PM_{2.5} concentration estimation with comprehensive information. Atmos. Environ. 256, 118302. <https://doi.org/10.1016/j.atmosenv.2021.118302>.