AI

MDPI

Article

Unveiling the Transparency of Prediction Models for Spatial $PM_{2.5}$ over Singapore: Comparison of Different Machine Learning Approaches with eXplainable Artificial Intelligence

M. S. Shyam Sunder [1], Vinay Anand Tikkiwal [2], Arun Kumar [3] and Bhishma Tyagi [1,*]

[1] Department of Earth and Atmospheric Sciences, National Institute of Technology Rourkela, Rourkela 769008, India; 521er1012@nitrkl.ac.in
[2] Department of Electronics Engineering, Jaypee Institute of Information Technology, Noida 201309, India; vinay.anand@jiit.ac.in
[3] Department of Computer Science and Engineering, National Institute of Technology Rourkela, Rourkela 769008, India; kumararun@nitrkl.ac.in
* Correspondence: tyagib@nitrkl.ac.in

Abstract: Aerosols play a crucial role in the climate system due to direct and indirect effects, such as scattering and absorbing radiant energy. They also have adverse effects on visibility and human health. Humans are exposed to fine $PM_{2.5}$, which has adverse health impacts related to cardiovascular and respiratory-related diseases. Long-term trends in PM concentrations are influenced by emissions and meteorological variations, while meteorological factors primarily drive short-term variations. Factors such as vegetation cover, relative humidity, temperature, and wind speed impact the divergence in the $PM_{2.5}$ concentrations on the surface. Machine learning proved to be a good predictor of air quality. This study focuses on predicting $PM_{2.5}$ with these parameters as input for spatial and temporal information. The work analyzes the in situ observations for $PM_{2.5}$ over Singapore for seven years (2014–2021) at five locations, and these datasets are used for spatial prediction of $PM_{2.5}$. The study aims to provide a novel framework based on temporal-based prediction using Random Forest (RF), Gradient Boosting (GB) regression, and Tree-based Pipeline Optimization Tool (TP) Auto ML works based on meta-heuristic via genetic algorithm. TP produced reasonable Global Performance Index values; 7.4 was the highest GPI value in August 2016, and the lowest was −0.6 in June 2019. This indicates the positive performance of the TP model; even the negative values are less than other models, denoting less pessimistic predictions. The outcomes are explained with the eXplainable Artificial Intelligence (XAI) techniques which help to investigate the fidelity of feature importance of the machine learning models to extract information regarding the rhythmic shift of the $PM_{2.5}$ pattern.

Keywords: $PM_{2.5}$; machine learning; MODIS; ERA5; XAI; random forest; gradient boosting; tree-based pipeline optimization tool

Citation: Sunder, M.S.S.; Tikkiwal, V.A.; Kumar, A.; Tyagi, B. Unveiling the Transparency of Prediction Models for Spatial $PM_{2.5}$ over Singapore: Comparison of Different Machine Learning Approaches with eXplainable Artificial Intelligence. *AI* **2023**, *4*, 787–811. https://doi.org/10.3390/ai4040040

Academic Editor: Demos T. Tsahalis

Received: 3 August 2023
Revised: 2 September 2023
Accepted: 22 September 2023
Published: 27 September 2023

1. Introduction

Inhalable particulate matter (PM) can cause acute and chronic diseases by infecting human respiratory organs. PM having ≤2.5 µm of particle size in its diameter, known as $PM_{2.5}$, has been reported as hazardous to human health by causing tuberculosis, lung cancer, and damage to the respiratory tract [1,2]. An increase in the short-term $PM_{2.5}$ is exemplary for human health, resulting in affected mortality rates negatively [3–6]. Different deleterious metals present in $PM_{2.5}$ affected the atmosphere, particularly in Asian countries [7]. Lelieveld et al. [8] employed a global atmospheric model to project $PM_{2.5}$ concentrations and utilized exposure–response equations from the work of Burnett et al. [9]. Variability sets the backdrop for evaluating the global premature mortality linked with chronic obstructive pulmonary disease (COPD), cerebrovascular disease (CEVD), ischemic heart disease (IHD), and lung cancer (LC) [10]. $PM_{2.5}$ also exhibited fluctuating patterns of

escalation and reduction across numerous regions globally, particularly in economically advanced areas, during recent decades [11]. For $PM_{2.5}$ predictions, Chen et al. [12] used environmental and meteorological parameters like vegetation cover, relative humidity, temperature, wind speed, and direction as they also impact the divergence in the surface $PM_{2.5}$ concentrations. Ancillary variables (like the Normalized Difference Vegetation Index (NDVI) for vegetation cover, water bodies, forests, urban areas and settlements, barren land, etc.) are closely linked with the emission sources and the mass movement of the air particles.

While considering significant scale prediction, the ground observations data are often deficit to analyze. The satellite data with good spatial and spectral information made it possible to predict $PM_{2.5}$ [13]. Aerosol optical depth (AOD) is a measure of the extinction effect of aerosols on the atmosphere. The particle size, distribution, and composition influence the AOD. The satellite-derived AODs and meteorological parameters are used in predicting $PM_{2.5}$ [14,15]. In addition to AOD, the reanalysis products like Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) and European Center for Medium-Range Weather Forecasting (ECMWF) ERA5 with a spatial resolution of $0.25° \times 0.25°$ [16] widely used source of meteorological data for $PM_{2.5}$ predictions. A global land cover dataset with a 300 m spatial resolution and MODIS Normalized Difference Vegetation Index (NDVI) with a 250 m spatial resolution are also used as supporting parameters in predicting $PM_{2.5}$. Moderate Resolution Imaging Spector-radiometer (MODIS) is a satellite product with Terra/Aqua satellite's; MODIS products like MODIS Dark Target (DT) with a spatial resolution of 10 km, MODIS Deep Blue (DB) with a spatial resolution of 10 km, and MODIS DB Multiangle Implementation of Atmospheric Correction (MAIC) with a 1 km spatial resolution [15,17]. Sekertekin et al. [18] and Xiang et al. [19] have demonstrated the usefulness of parameters such as AOD and land surface temperature (LST) derived from MODIS data in improving $PM_{2.5}$ prediction accuracy. Inverse distance weighted (IDW), a spatial interpolation technique used for interpolating the ground data, was found to have a 24% error difference in the predicted and actual value of $PM_{2.5}$ concentrations for the Delhi monitoring stations [20].

For estimating $PM_{2.5}$ concentrations, Ma et al. [14] used autoregressive integrated moving averages (ARIMA) and multiple linear regression statistical models. Pu and Yoo [15] developed a multi-stage model to attribute unavailable/missing values in spatial data by quantifying the uncertainties and gave meaningful outcomes in $PM_{2.5}$ predictions. The density of data availability is positively correlated with the strength of the outcomes. As a result, specific methods such as wavelet transform were used to boost data abundance [21]. Machine learning (ML) algorithms are assigned via explicit programs to learn and understand structural and practical data-related problems. Especially in climate and extreme weather predictions, with the available labeled climate benchmark datasets, the ML algorithms trained and understood the typical feature-based circumstances [22]. Statistical and ML models have been used to estimate $PM_{2.5}$ concentrations and to identify the specific severity and local impact of emissions on potentially affected areas [23]. ML algorithms, including Bayesian statistics, regression, Random Forest (RF), radial basis function long short-term memory (RBF-LSTM) algorithms, maximum likelihood estimation (MLE), support vector machine (SVM), K-nearest neighbor classifier (KNN), and neural network (NN) models, have also been used to extract key features of $PM_{2.5}$ in time series meteorological data and improve the accuracy of predictions [13,21,24]. Deep learning (DL) techniques can also provide reasonably good predictions by adjusting their critical hyperparameters. In addition, it also investigates how the proposed method can be extended to apply to the other types of datasets in dispersion in the atmospheric chemistry domain.

eXplainable Artificial Intelligence (XAI), a novel technique to explain the transparency and fidelity of models, has garnered interest in clarifying the significance and dependability of features and models [25]. XAI has attracted attention for explicating the importance and trustworthiness of characteristics and models, though some methodologies need more desirable properties and face constraints. Implementing an agnostic method can significantly

help bias control. Many XAI approaches contradict desirable properties (such as "completeness" or "implementation invariance") and often have nontrivial limits for particular issue configurations. In neural networks, the softmax function, for instance, creates values (class weights) with particular probabilistic properties but not actual probabilities. On the other hand, relative frequencies are considered for the possibilities of classes in random forests and random survival forests. They are exact probabilities determined by calculating the percentage of various sample classes at each leaf node [26]. In both cases, it is evident that the degree to which the class probability distributions are accurate depends on the quantity of training data and the ML model employed to forecast the distributions. It is difficult to influence the creation of a post hoc ML model representing an opaque system [27].

Some studies were made on the XAI technique using remote sensing datasets. Kakogeorgiou et al. [28] used saliency methods to qualitatively evaluate the input benchmark datasets BigEarthNet and SEN12MS that are used to fill the gaps via deep learning models. Input × Gradient (InputXGrad), Integrated Gradients (IntGrad), Guided Backpropagation (GuidedBackprop), Grad-CAM, Guided Grad-CAM, Occlusion, DeepLift, Lime, and SmoothGrad (SG) were used, which resulted in achieving lowest max-sensitivity, providing reliable data classification. The Shapley Additive Explanations (SHAP) model was used to understand different datasets (Ecological Spectral Information System; Spectro-radiometer, ASD Field Spec bare fiber) over different scenarios, with the analysis of plant tissues, including information such as contents of nitrogen, leaf area index, and water content, in Israel [29]. Stadtler et al. [30] used a multi-labeled global air quality benchmark dataset over black box models using SHAP, Neural Network Activation, Random Forest Activation, and Explaining Inaccurate Predictions with k-nearest neighbors to explain the prediction of ozone's averaged distribution over different periods. Betancourt et al. [31] and Stirnberg et al. [32] also used the SHAP model to explain the transparency and contribution of datasets in the model. As input, the global dataset AQ-Bench and reanalyzed observations (ReObs) collected at Paris Charles de Gaulle Airport, located northeast of Paris, were used in machine learning models to predict variations in tropospheric ozone and PM_1 concentrations, respectively. The Regression Activation Mapping (RAM) model was used in the Indian wheat belt region to explain the contribution of datasets at each time step in predicting meteorological and satellite-derived vegetation variables at daily temporal resolution.

In this study, the prediction of monthly $PM_{2.5}$ concentrations spatially using satellite products along with targeting sparse observation data measured at different sites for validation in the north, south, central, east, and west directions over Singapore (detailed in Section 3) is undertaken. The environmental and meteorological variables and the AOD data-driven methods are referred for mapping spatial predictions. This study's primary contribution is to explore different ML models for spatial prediction. The secondary contribution focuses on anticipating better ML model performance for predicting $PM_{2.5}$ spatially by incorporating multiple input variables and improving the model performance using a meta-heuristic approach and ensemble models. This study also explains the factors responsible for prediction using novel XAI techniques on different models. The ML models, Random Forest regression, Gradient Boosting regression algorithm, and TP are employed over meteorological variables, including wind, relative humidity, temperature, and vegetation cover, in conjunction with AOD data to forecast surface $PM_{2.5}$. ML techniques have been utilized for one-time-step training and prediction, and sequential information is produced with reasonable accuracy for subsequent time steps. The model is designed to predict each pixel/data point within the study area and is scrutinized using the XAI technique. Although the prediction of $PM_{2.5}$ level has been explored extensively with ML models in recent years, the spatial pixel-based prediction is still to be explored. Even though several popular neural network architectures are used in $PM_{2.5}$ prediction, the neural networks do not focus on peak variations like extreme events [33], and it is tough to interpret the model's transparency, uncertainty, and explainability. They also have more chances of overfitting the exploratory variable when the external factors

are added [33]. ML approaches can learn high-dimensional, complex representative features [34] and are easily interpretable. There is still a research gap in understanding the model complexity, interpretability, temporal and spatial dynamics, feature selection, and engineering and involving multiple/hybrid models that lead to enhanced prediction accuracy in the spatial domain.

The results of this study concentrating on various air pollutants can be beneficial in devising mitigation. This goal seeks to foster sustainable, resilient, and safe cities and human settlements that are inclusive. With remote sensing technology, open source data products, statistics, and ML with XAI models can significantly improve the level of accuracy in $PM_{2.5}$ predictions and inform air quality management for better practices. The novelty of this work includes the identification of the potential of dynamic ML models that incorporate past data and important features in predicting $PM_{2.5}$ with the input of spatial and temporal satellite datasets, including multiple meteorological variables and the identification of input variables that strongly explain the $PM_{2.5}$ predictions over Singapore based on temporal statistical premises, explaining the need of multiple features variables as input and the inter- and intra-variation of different tree-based machine learning models with and without optimization algorithms. The following objectives are taken into consideration when creating and deploying a novel ML framework to forecast spatial $PM_{2.5}$ distributions and analyze temporal changes of delicate particulate matter:

(i) Predicting the spatial $PM_{2.5}$ values over Singapore and validating the outcomes using machine learning models.
(ii) Investigating the fidelity of the model outcomes with XAI.

The scope of this work is to identify a good spatial prediction ML model for $PM_{2.5}$ by comparing the performance of different ML models in the Singapore region. This study is structured into five sections. Section 1 provides an introduction and literature review. Section 2 contains details of the study area and a description of the datasets used in this study. The workflow is described in Section 3. Section 4 presents and discusses the results obtained in detail including exploratory analysis. Finally, Section 5 presents the conclusion of this study.

2. Study Area and Data Description

Located at the southernmost point of the Malay Peninsula, Singapore is a highly urbanized city-state with a population of 5.7 million residents and 3 million daily commuters, situated approximately 137 km north of the Equator [35–37]. With an equatorial climate characterized by year-round rainfall, humidity, and high temperatures, our study focuses on the in situ sites selected to cover the spatiotemporal dynamics of Singapore's air quality, as presented in Figure 1. The National Environmental Agency (NEA) of Singapore regularly monitors ambient air pollutants, including particulate matter ($PM_{2.5}$ and PM_{10}). As already discussed, $PM_{2.5}$ represents finer particulates with a diameter $\leq$2.5 μm, and they are primarily associated with health impacts and large distance transportations [38,39]. While PM_{10} is defined as particles with diameters $\leq$10 μm, it corresponds to larger particulates, respectively [40]. Due to their bigger sizes, their health impacts are not severe, though they are important for the weather and climate processes [41] and are important for studying forest fires [42], industrial transportation [43], and pollution contributions [44]. For the present study, our focus is bounded to $PM_{2.5}$ [6] measured at north, south, east, west, and central sites across Singapore (as shown in Figure 1). The data obtained have been averaged on a monthly analysis for 2014–2019, focusing on June, July, August, and September, which exhibit consecutive rainfall deficits compared to other months.

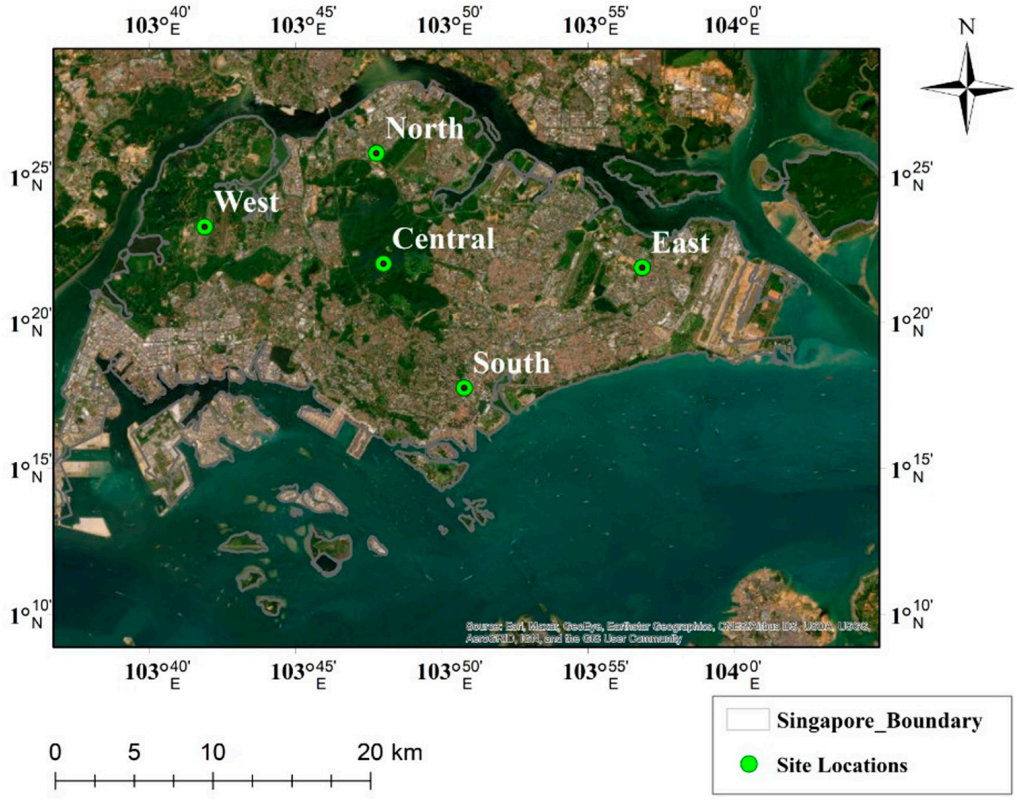

Figure 1. Satellite imagery of study area, Singapore.

The geographical location and visual representation (map) of the study area are depicted in Figure 1, encompassing five monitoring stations strategically placed in the west, north, east, south, and central regions of Singapore to collect observational data.

3. Methodology

The methodological flowchart in Figure 2 depicts the methodology of the work carried out in this study. MODIS and ERA-5 spatial datasets have been utilized. The climate data store (https://cds.climate.copernicus.eu/, accessed on 24 March 2022) makes the ERA5 data available globally in 0.25° × 0.25° grids of latitude–longitude with time scales for the fifth generation of the ECMWF, atmospheric reanalysis of global climate. For this study, the MODISplatform having Terra and Aqua satellites with the Collection of 6.1, level 2 aerosol optical depth product at 550 nm wavelength for both land and ocean [45], with a spatial resolution of 10 km, is used.

Python is mainly used to prepare datasets and run to the ML models, and spatial maps were plotted using ArcGIS 10.3. Preprocessing is performed to transform data into an efficient input format that will be fed to the model. The different preprocessing methods used in this research work include feature variable selection, handling missing values to fill the spatial gaps with observation points, and creating test/training datasets with 30% and 70% for feeding into ML algorithms.

Inverse distance weighted (IDW) is used to interpret observed data points to validate the spatially predicted outputs; an example is shown in Figure 3. The input data are resampled into coarse resolution ~0.25° × 0.25° while pre-processing to feed into the machine learning models. The systematic methodology and the different parameters for different methods used in this work are shown in Figure 2.

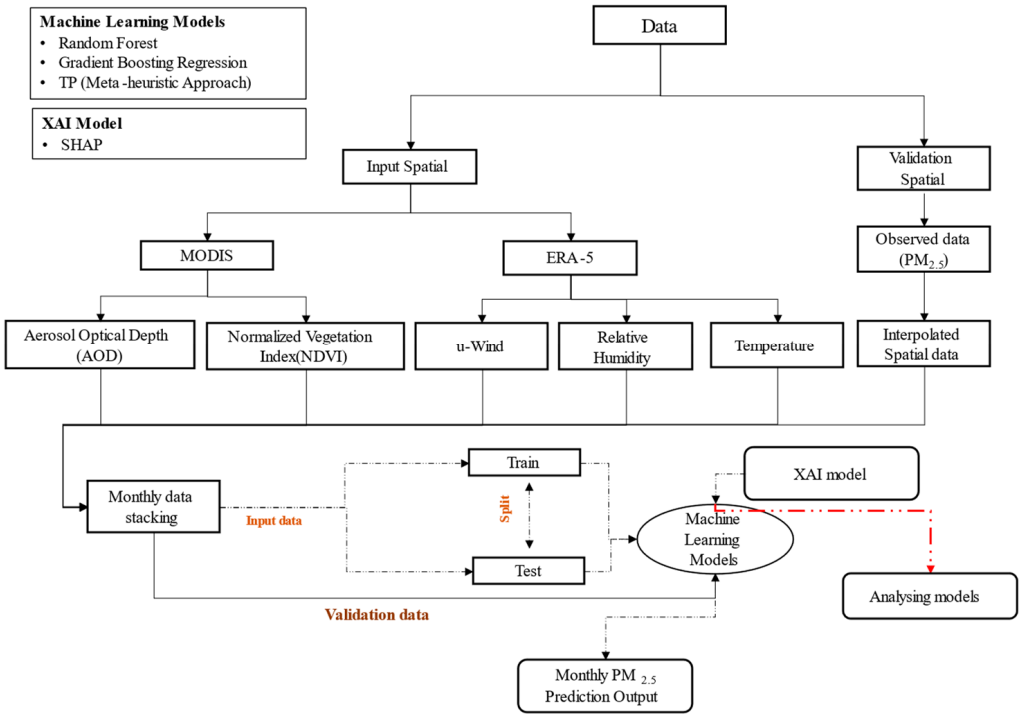

Figure 2. Systematic methodology detailing the data, models, and procedures used.

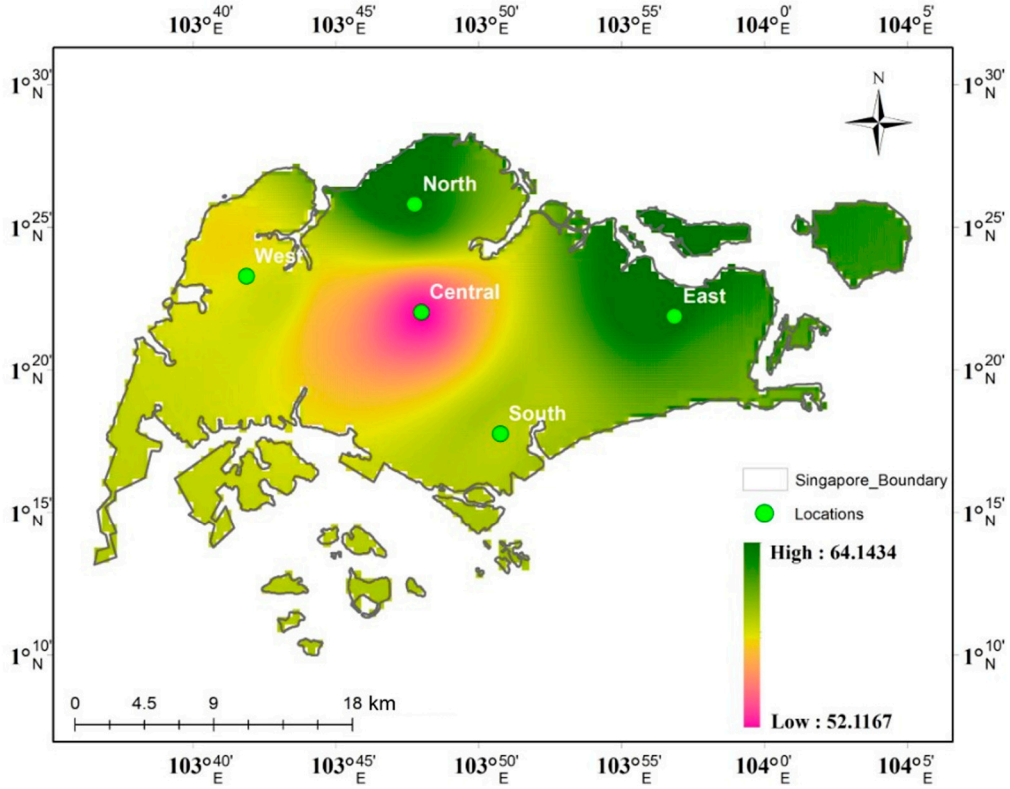

Figure 3. Inverse distance weighted interpolated data using observation data for January 2014 of $PM_{2.5}$ concentration.

3.1. The Machine Learning Models

3.1.1. Random Forest (RF) Regression

RF is a machine learning algorithm for classification and regression problems. It is a bootstrapping tree-based model ensemble with a bagging algorithm that combines multiple decision trees to make accurate predictions. The random forest algorithm builds a forest of decision trees, where each tree is constructed using a different subset of training data and random subset features at each node in the tree. The sample is based on the counts of trees, and the tree growth depends on the best split and the node of the input variables in the dataset [46]. This randomness in the tree construction helps to reduce overfitting and improve the model's generalization ability. The accuracy of the RF algorithm depends on two main parameters: the number of decision trees and the number of features in the random subset at each node.

Adding more decision trees increases the model's accuracy, but the computational cost also increases. The square root of the total number of features in the dataset is used to determine how many features are included in each node's random subset. The subset's feature count can be adjusted to balance the bias and variance. The random forest approach produces the mean of all the individual regression trees' predictions in regression problems. Each regression tree predicts the target variable as a constant value, and the final prediction made by the random forest is the average of all the tree forecasts. This strategy aids in lowering the model's variance and raising the forecast's precision.

3.1.2. Gradient Boosting (GB) Regression

GB, like RF, is also an ensemble learning method, but this model builds them sequentially instead of having multiple decision trees. Integrating two model algorithms with decision trees and a supervised method is used to aggregate the final output prediction. The GB uses the loss function for the converging output to minimize the loss using less complex decision trees [47]. Gradient boosting iteratively adds weak learners to the ensemble, each weak learner attempting to outperform the prior weak learners. The trees are created to remedy the flaws of the preceding tree. At each iteration, the algorithm calculates the negative gradient of a loss function with respect to the predictions of the previous trees. Then it fits a new tree to the negative gradient. One of the advantages of gradient boosting is that it can handle different loss functions, such as mean squared error for regression problems and cross-entropy loss for classification problems. It can also handle missing values and outliers in the data. The performance of gradient boosting depends on several hyperparameters, such as the learning rate, the number of trees, the depth of the trees, and the regularization parameters. The learning rate controls the contribution of each tree to the final prediction, and a lower learning rate generally leads to better performance but slower convergence. The number of trees determines the complexity of the ensemble, and a more significant number of trees can improve the performance and increase the computational cost. The depth of the trees controls the complexity of each weak learner, and a larger depth can lead to overfitting. Finally, the regularization parameters, such as the minimum samples per leaf and the maximum depth, can also help to prevent overfitting.

3.1.3. Extreme Gradient Boosting Regression (XGBoost)

The robust ML algorithm uses a Gradient Boosting algorithm based on a decision tree-based ML algorithm. The model performance is found to outperform the small to medium-sized datasets. Although certain hyperparameters need to be tuned, the parameters are tuned automatically to stop the learning when the best value is reached [48,49].

3.1.4. Tree-Based Pipeline Optimization Tool (TP) Optimization Algorithm

Optimization algorithms are divided into exact algorithms and heuristics mostly. Heuristic algorithms are particular, problem-dependent, and meta-heuristic approaches like the Genetic Algorithm (GA) [50]. Random Search [51], Grid Search, and Evolutionary Algorithm (EA) [52] are common approaches to building AutoML systems for diverse

applications. It is an AutoML tool created to create optimal pipelines through GA effectively, the emerging method to be faced in the irregular research space [53–55]. Many randomly assembled candidate pipelines are evaluated by TP [56], which are used in this study. The complexity of model accuracy is accounted for using the pipeline. TP uses feature selection and feature engineering, model selection, and hyperparameter optimization. Complete pipeline cross-validation is carried out based on their cross-validated score [57], such as balanced accuracy or mean squared error.

3.2. Error Metrics

Error metrics are used to evaluate the models to select or make an efficient ML model. Several metrics are available for different kinds of machine-learning problems. Since the problem at hand is a regression task, Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R^2 score have been used to evaluate our regression model outcomes in predicting $PM_{2.5}$ [58–60].

$$\mathrm{MSE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}(\hat{\mathrm{y}}_{\mathrm{i}} - \mathrm{y}_{\mathrm{i}})^2 \tag{1}$$

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

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{3}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}|\frac{\hat{y}_i - y_i}{y_i}| \times 100\% \tag{4}$$

$$R^2 = \frac{\sum_{i-1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i-1}^{n}(y_i - \bar{y})^2} \tag{5}$$

$$GPI = \sum_{k}^{4}\alpha_k(m_k - n_{ik}) \tag{6}$$

MSE, being a popular metric, efficiently points out the mean squared error of predicted and actual values. To find the deviation between the targeted and the predicted values, the RMSE is used. MAPE is used for calculating the relative absolute error in percentage to compare forecast accuracy between the models [61]. MSE is a commonly used metric to validate ML model performances by measuring the average squared differences between predicted and actual values. The squaring of the differences emphasizes more significant errors in the model predictions. RMSE is another commonly used evaluation metric, which accounts for the scale of the response variables in the dataset, in contrast to MSE, which does not consider the scale. This makes RMSE more sensitive than MSE to differences in the response variable values. R^2 has often been used as a metric to assess the variability in regression model responses. It represents the proportion of variance in the response variables that the model's predictor variables can explain. A higher R^2 value indicates a better fit of the model to the data.

R^2 score is a posted metric calculated using the sum of squared errors. If the sum of the square of error is small, which is near 1, it means the variance of the target variable is wholly captured and vice versa for the high value of the square error of the regression line.

Zhu et al. [62] proposed using the Global Performance Index (GPI) as a method for ranking the performance of machine learning models. The GPI combines multiple metrics into a score to determine which model performs the best. The GPI formula includes a constant αk, set to 1 for metrics like MSE, RMSE, and MAPE and -1 for R^2. The scaled value of each statistical indicator is represented as nik, with mk being the median of the scaled statistical indicator j for all models (where k = 1, 2, 3, and 4). Higher GPI values

indicate better model performance, and the model with the highest GPI value at a given station is considered to have the best predictive capacity.

3.3. XAI Methods

The SHapley Additive exPlanations (SHAP) method is widely used in explainable AI as it explains the importance of features in ML models and can be applied to different types of models since it is model-agnostic. The method is based on game theory, which allows it to determine the optimal contributions of different features in a game [63,64]. SHAP values can be efficiently computed for tree models using Tree SHAP, a tree-based version of the method [64].

Singh et al. [19] demonstrated the effectiveness of SHAP in identifying feature importance in various machine learning models by eliminating individual features and monitoring the changes in their contribution to the overall model. The SHAP method has both global and local explanations. The global explanation provides an overview of the model's feature importance, while the local explanation explains individual predictions.

Local explanation approaches can be model-agnostic and used to explain tree models [65]. However, these approaches may be slow or experience sampling variability when used with models that have many input features. In summary, SHAP is a powerful and widely used XAI method that explains feature importance in machine learning models. It can provide both global and local explanations, which can help interpret the model's output for individual predictions.

4. Results and Discussion

$PM_{2.5}$ predictions use multiple meteorological parameters (NDVI, relative humidity, temperature, and U—wind) with Aerosol Optical Distribution (AOD). Following the proposed systematic procedure shown in the methodology section in Figure 2, the performance of the result is to be analyzed and discussed in subsequent sub-sections. Predictions are made using several ML algorithms, incorporating several machine learning methods. Still, our analysis is focused on regression problems. Furthermore, MSE, RMSE, R^2 score, MAE, and MAPE metrics with XAI (an interpretable ML technique having SHAP, a contemporary algorithm used as a global interpreter) are used to perform inference analysis.

4.1. Performance Comparison with RF and GB Models

The following output is obtained using RF and GB regression models with MSE, RMSE, and R^2 scores. The parameters used are given in Table 1. The performance of the regression models is inferred below.

Table 1. Parameters used to perform Random Forest and Gradient Boosting regression algorithms.

Parameters	Random Forest	Gradient Boosting
n_estimator	10	1200
criterion	mse	friedman_mse
max_depth	None	4
min_sample_split	1	2
min_samples_leaf	1	1
min_density	0.1	None
learning_rate	None	0.01
random_state	None	3
subsample	None	0.5
Out of Bag(OOB)_score	bool	None
n_jobs	1	None

The Random Forest regression algorithm results show the lowest MSE and RMSE values for the test data, 0.01, in June 2019. Table 2 shows that the highest MSE and RMSE values were observed in August 2015, with a value of 0.63 for test data in September. The R^2 values were highest in August 2014, with 0.89 and lowest in June 2017 and July 2019 for test data (0.26). The MAE and MAPE values were also computed in Tables 3 and 4. The lowest MAE values for training data were observed at 0.01 to 0.04, and the highest values at 0.1 to 0.13 over the months of June to September for the years 2014 to 2019. The lowest MAPE values were observed at 0.0003 to 0.0007, while the highest values were observed at 0.0014 to 0.0023 from Table 4.

Table 2. MSE and RMSE value for test and training datasets using RF and GB models for June–September (2014–2019).

RF															
MSE								RMSE							
Training	Test	Training	Test	Training	Test	Training	Test	Training	Test	Training	Test	Training	Test	Training	Test
June		July		August		September		June		July		August		September	
0.04	0.2	0.02	0.07	0.02	0.05	0.03	0.11	0.2	0.49	0.15	0.28	0.16	0.24	0.17	0.33
0.1	0.04	0.08	0.16	0.25	0.05	0.07	0.63	0.2	0.31	0.29	0.4	0.23	0.5	0.27	0.79
0.05	0.07	0.05	0.07	0.09	0.49	0.05	0.03	0.22	0.27	0.23	0.27	0.31	0.7	0.22	0.18
0.04	0.04	0.01	0.02	0.02	0.05	0.06	0.08	0.21	0.22	0.14	0.15	0.16	0.22	0.25	0.29
0.03	0.19	0.04	0.3	0.03	0.14	0.01	0.06	0.18	0.44	0.2	0.54	0.17	0.37	0.1	0.24
0.004	0.01	0.001	0.01	0.003	0.01	0.02	0.12	0.07	0.12	0.04	0.12	0.06	0.13	0.15	0.34
GB															
4.48	0.32	2.77	0.09	0.0002	0.06	5.65	0.18	0.56	0.006	0.005	0.3	0.01	0.25	0.007	0.42
0.21	0.0002	0.002	0.58	0.0006	0.2	0.0005	0.51	0.46	0.01	0.05	0.76	0.02	0.45	0.02	0.71
0.04	4.006	5.02	0.27	0.001	0.8	0.0007	0.08	0.22	0.006	0.52	0.007	0.89	0.04	0.02	0.28
0.11	0.0002	1.57	0.04	0.0002	0.13	0.0002	0.17	0.33	0.01	0.003	0.21	0.01	0.37	0.01	0.41
0.01	4.65	0.003	0.33	0.0001	0.1	0.0001	0.08	0.382	0.006	0.06	0.58	0.01	0.32	0.01	0.29
1.75	0.02	5.43	0.01	9.73	0.01	0.006	0.11	0.004	0.16	0.002	0.1	0.009	0.14	0.07	0.34

Table 2 presents the results of the Gradient Boosting regression algorithm for the test data, the MSE was highest at 4.65 during June 2018, and the lowest value of 0.0002 was observed during June 2017. The RMSE values were highest in August 2016, with 0.89 for the training data, and in July 2015, with 0.76 for the test data, as shown in Table 2. The lowest value of RMSE for the test data was 0.006 in June 2014. The R-squared values were overestimated with training data across all the years and months, with a value of 0.99. The test data showed the highest value of 0.89 in August 2014 and the lowest value of −0.01 in June 2019, as indicated in Table 3.

The minimum values of MAE and MAPE in the training data, as shown in Tables 4 and 5, were 0.02, 0.01, 0.02, and 0.03 for MAE and 0.0005, 0.0002, 0.0003, and 0.0004 for MAPE for June–September for the years 2014 to 2019. The maximum values were 0.07, 0.08, 0.10, and 0.09 for MAE and 0.0014, 0.0017, 0.0018, and 0.0010 for MAPE. The variation in absolute error for the training dataset was at its minimum in July 2019, with an MAE of 0.01 and MAPE of 0.0003 for the RF model and an MAE of 0.01 in July 2019 and a MAPE of 0.0018 in July 2019 for the GB model. The maximum variation error was observed during August 2019, with an MAE of 0.13 and MAPE of 0.0023 for the RF model and an MAE of 0.01 in July 2019 and a MAPE of 0.0018 in July 2019 for the GB model.

Table 3. R2 and MAE values for test and training datasets using RF and GB regression models for June–September (2014–2019).

RF															
R^2								MAE							
Training	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**
June		**July**		**August**		**September**		**June**		**July**		**August**		**September**	
0.95	0.65	0.95	0.82	0.96	0.89	0.95	0.78	0.1	0.24	0.08	0.14	0.08	0.13	0.1	0.21
0.95	0.86	0.93	0.81	0.94	0.62	0.94	0.34	0.07	0.17	0.11	0.22	0.1	0.27	0.12	0.36
0.93	0.8	0.94	0.88	0.93	0.54	0.92	0.87	0.09	0.14	0.08	0.15	0.13	0.33	0.08	0.11
0.86	0.65	0.86	0.54	0.88	0.55	0.87	0.78	0.05	0.1	0.03	0.06	0.04	0.1	0.07	0.15
0.91	0.28	0.92	0.6	0.94	0.74	0.95	0.66	0.07	0.18	0.09	0.26	0.08	0.19	0.04	0.14
0.92	0.35	0.94	0.26	0.96	0.84	0.95	0.74	0.03	0.07	0.01	0.05	0.03	0.06	0.06	0.17
GB															
0.99	0.54	0.99	0.79	0.99	0.89	0.99	0.65	0.07	0.24	0.06	0.17	0.06	0.15	0.08	0.21
0.99	0.7	0.99	0.34	0.99	0.7	0.99	0.46	0.06	0.18	0.08	0.29	0.08	0.27	0.09	0.37
0.99	0.87	0.99	0.54	0.99	0.25	0.99	0.68	0.06	0.16	0.07	0.18	0.1	0.35	0.06	0.14
0.99	0.225	0.99	0.11	0.99	−0.18	0.99	0.57	0.03	0.12	0.02	0.07	0.04	0.12	0.05	0.15
0.99	0.45	0.99	0.55	0.99	0.8	0.99	0.52	0.05	0.2	0.08	0.28	0.05	0.2	0.03	0.13
0.99	−0.01	0.99	0.53	0.99	0.82	0.99	0.75	0.02	0.08	0.01	0.04	0.02	0.07	0.04	0.17

Table 4. MAPE value for test and training datasets using RF and GB regression models for June–September (2014–2019).

RF							
MAPE							
Training	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**
June		**July**		**August**		**September**	
0.002	0.0047	0.0014	0.0026	0.0015	0.002	0.0012	0.0026
0.0016	0.0036	0.0023	0.0048	0.0015	0.004	0.0009	0.0026
0.0017	0.0028	0.0016	0.003	0.0022	0.0057	0.0014	0.002
0.001	0.0023	0.0007	0.0013	0.001	0.0022	0.0013	0.0028
0.0014	0.0036	0.0015	0.0044	0.0014	0.0032	0.0007	0.0022
0.0007	0.0017	0.0003	0.0009	0.0004	0.001	0.0006	0.0017
GB							
0.0014	0.0046	0.0011	0.003	0.0011	0.0029	0.001	0.0027
0.0013	0.0038	0.0017	0.0061	0.0012	0.0039	0.0006	0.0027
0.0013	0.003	0.0013	0.0036	0.0018	0.0061	0.001	0.0025
0.0008	0.0028	0.0005	0.0017	0.0008	0.0027	0.0009	0.0028
0.0011	0.0039	0.0014	0.0047	0.0009	0.0033	0.0005	0.0022
0.0005	0.0019	0.0002	0.0008	0.0003	0.0011	0.0004	0.0017

Table 5. MSE, RMSE, R^2, MAE, and MAPE values for test and training datasets using the TP model for June–September (2014–2019).

Regression Models	**TP**															
Metrics	**MSE**								**RMSE**							
Case	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**
Year/Months	**June**		**July**		**August**		**September**		**June**		**July**		**August**		**September**	
2014	0.27	0.30	0.21	0.30	0.14	0.28	0.15	0.38	0.52	0.55	0.46	0.55	0.38	0.53	0.38	0.62
2015	0.20	0.38	0.57	0.58	0.51	0.32	0.68	0.54	0.45	0.61	0.76	0.76	0.72	0.57	0.82	0.74
2016	0.40	0.21	0.26	0.51	0.99	1.57	0.34	0.27	0.63	0.46	0.51	0.72	1.00	1.25	0.58	0.52
2017	0.17	0.07	0.10	0.04	0.12	0.08	0.35	0.38	0.41	0.27	0.31	0.19	0.35	0.28	0.60	0.61
2018	0.29	0.45	0.32	0.59	0.28	0.28	0.10	0.11	0.54	0.67	0.57	0.77	0.53	0.53	0.31	0.34
2019	0.03	0.03	0.02	0.03	0.05	0.09	0.21	0.24	0.18	0.18	0.14	0.17	0.23	0.30	0.46	0.49
	R^2								MAE							
2014	0.68	0.52	0.59	0.23	0.79	0.52	0.78	0.27	0.27	0.35	0.25	0.37	0.21	0.32	0.24	0.42
2015	0.80	0.45	0.59	0.31	0.54	0.49	0.62	0.41	0.19	0.45	0.38	0.56	0.32	0.38	0.31	0.40
2016	0.56	0.40	0.79	0.03	0.37	−0.59	0.53	−0.16	0.26	0.30	0.19	0.57	0.48	0.88	0.24	0.31
2017	0.63	0.42	0.49	0.18	0.60	0.24	0.14	0.04	0.13	0.15	0.11	0.11	0.13	0.16	0.28	0.41
2018	0.34	−1.05	0.52	0.09	0.52	0.49	0.60	0.18	0.23	0.33	0.26	0.42	0.27	0.36	0.15	0.20
2019	0.55	−0.25	0.30	−0.48	0.53	0.10	0.60	0.42	0.09	0.11	0.05	0.08	0.12	0.20	0.19	0.26
	MAPE															
2014	0.00529	0.0068	0.0045	0.0066	0.0039	0.006	0.003	0.0053								
2015	0.004	0.0095	0.0078	0.0116	0.0046	0.0054	0.0023	0.0029								
2016	0.0047	0.0055	0.0036	0.0113	0.008	0.0148	0.0042	0.0054								
2017	0.0028	0.0032	0.0022	0.0023	0.0029	0.0036	0.0052	0.0074								
2018	0.0045	0.0064	0.0044	0.007	0.0044	0.0059	0.0024	0.0031								
2019	0.0021	0.0027	0.001	0.0014	0.0017	0.003	0.0018	0.0025								

The results of both RF and GB regression models were compared, with RF demonstrating strong performance in predicting $PM_{2.5}$ values. Regarding MAPE, GB produced the lowest error of 0.0002, while RF had the highest error of 0.0007. The maximum percentage error was 0.001 for GB, and the lowest was 0.0023 for RF regression. Although RF did not perform as well as GB in terms of accuracy, it exhibited the lowest error variation, indicating error consistency. On the other hand, GB was found to overestimate, as demonstrated by the R^2 score within the trained dataset. Additionally, the analysis of the variation of absolute error for the training dataset revealed that RF exhibited a minimum error in July 2019, with an MAE of 0.01 and MAPE of 0.0003, while GB exhibited a minimum error in July 2019, with an MAE of 0.01 and MAPE of 0.0002. The maximum error variation was observed in August 2019 for RF, with an MAE of 0.13 and MAPE of 0.0023, and in July 2019 for GB, with an MAE of 0.01 and MAPE of 0.0018.

4.2. Comparative Analysis Using TP (AutoML) Meta-Heuristic Approach Using Genetic Algorithm

Different time periods are used as input, and the regression-based models that TP suggests are unpruned [66]. Based on a pipeline fit with five genetic iterations and a negative mean absolute error, the cross-validation score (5 folds) is calculated. According to the optimum pipeline recommendation of TP, tree-based regression ensemble methods are primarily used in the dataset.

The best pipeline models are generated by TP, each with different regression models and their corresponding hyperparameters. The evaluation metrics over TP revealed that the mean squared error (MSE) values were highest during August 2016, registering at 0.99 and 1.57 for training and test data, respectively. In contrast, the lowest MSE values were observed during July 2019, with a score of 0.0291 for test data, as shown in Table 5. The root mean squared error (RMSE) scores were highest for both training and test data during August 2016, measuring at 1.2535 and 0.9952, respectively, while the lowest RMSE score for training data was 0.18 in June 2019 and for test data, it was 0.17 in July 2019.

The R^2 values for the training data were higher, with a maximum of 0.52 in June 2014 and a minimum of 0.144 in September 2017. In contrast, the R^2 value for test data was highest at 0.80 in June 2014 and lowest at -1.04981 in June 2018. The mean absolute error (MAE) values for both training and test data were the highest in August 2016, with scores of 0.48 and 0.88, respectively. However, the MAPE values were lower in June and July of 2019, registering at 0.11 for test data and 0.05 for training data. The highest MAPE values were recorded in August 2016, with scores of 0.008 and 0.014 for training and test data, respectively, while the lowest values were observed in July 2019 and July 2017 at 0.0023 for training and test data, respectively. MSE, RMSE, MAE, and MAPE metrics were used to infer the outcome for the TP model. Using the XGB Regressor with specific hyperparameters, such as the alpha value of 0.014, max depth of 3, minimum child weight of 3, and 100 estimators, as shown in Table S1, provided in the supplementary material, the August 2016 $PM_{2.5}$ concentration values differ from those of another period.

Although evaluation metrics such as MSE, RMSE, R^2, MAE, and MAPE indicate different accuracy levels with specific models and periods, the GPI values serve as a helpful tool to understand these performance fluctuations and provide insights into the impact of different meteorological variables and patterns on the predictive accuracy of the models.

Meteorological variables and seasonal patterns influence $PM_{2.5}$ predictions on local scales according to changes in time. This leads to variations in the performance of models on a monthly and yearly basis apart from data uncertainties.

As shown in Figure 4, with regard to GPI values for the RF algorithm, August 2016 has the highest predicted value, while August 2019 has the lowest. For the GB algorithm, August 2016 has the highest predicted value, while September 2018 has the lowest. The range of predicted values varies significantly for different algorithms and month combinations. Some have a large range (e.g., August 2016 for GB), while others have a much smaller range (e.g., June 2019 for RF). The TP model's performance indicates its predictive capacity

during different time periods, and the range of values varies significantly across years and months. The GPI values are used to evaluate the performance of the models. TP's highest GPI value was 7.4 in August 2016, and the lowest was −0.6 in June 2019. The TP model exhibits positive performance, with the negative GPI values being less pessimistic than other models.

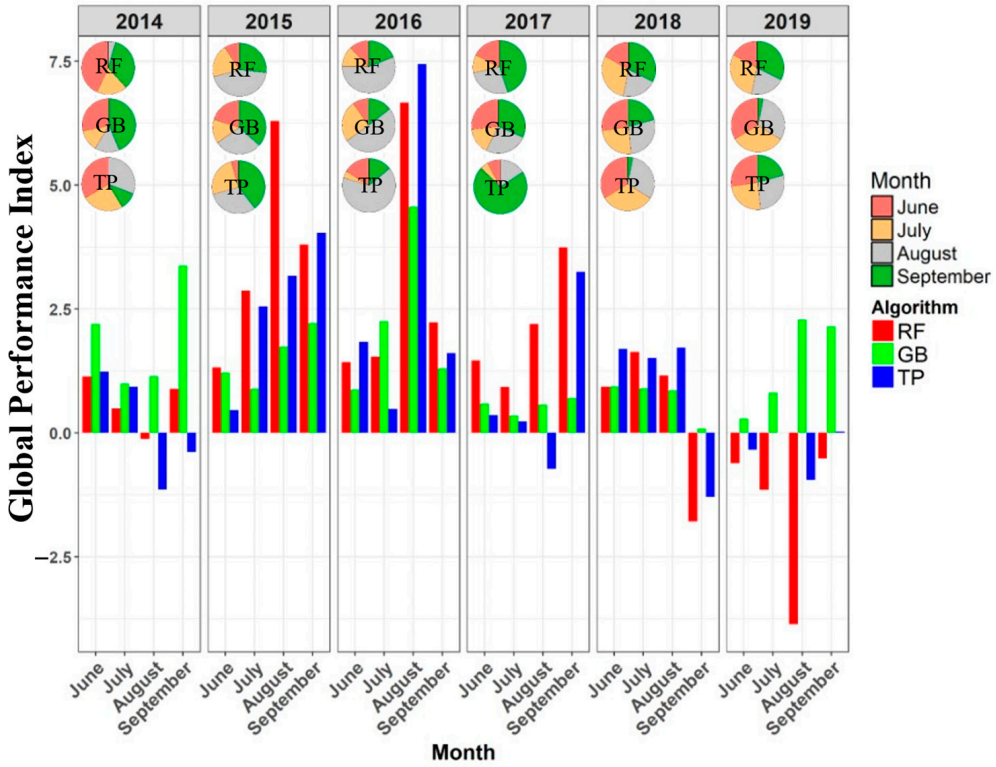

Figure 4. Yearwise monthly variation in Global Performance Index (GPI).

4.3. Global Interpretability and Local Interpretability Using SHAP Model

Global explanation using SHAP values can explain predictors/features' contribution to the output features. The regression models are used as the base model with a tree explainer for calculating SHAP values. Upon analyzing the results of the RF and GB regression models for July 2019 and August 2019, with mean accuracy scores of 0.35 and 0.68 for RF and GB test datasets, a brief study was conducted to investigate the relationship between features and $PM_{2.5}$ values. The results indicated that RH for July 2019 and UWIND for August 2019 were the major contributors in predicting $PM_{2.5}$ values using both RF and GB regression methods, as shown in Figure S1. Specifically, over the southern region, both regression methods showed high positive SHAP values above 0.6 for UWIND in the August 2019 dataset, accounting for approximately 66.30% of the total variation, as depicted in Figures S2 and S3. Furthermore, as AOD strongly correlates with $PM_{2.5}$, the relationship between AOD and influencing features was investigated. Figures S4 and S5 indicated that temperatures above 300 K and RH between ~76 and 78, with a UWIND speed of ~−1.0 m/s, were the most influential features for both RF and GB during July 2019 and August 2019. The results that were consistent with Figures S5–S7 and also with local comparison on a directional basis for both models were shown in Figures S8–S10, which showed that UWIND and RH were the most influential features in predicting $PM_{2.5}$ values.

Fine particulate matter ($PM_{2.5}$) is a significant air pollutant; by studying variability in $PM_{2.5}$ it is possible to understand the factors that contribute to its distribution and can be informed to policymakers to reduce its impact. The SHAP method was used to explain the feature importance of TP's best pipeline results in predicting $PM_{2.5}$ levels. The SHAP values are used to measure the impact of each feature on the model output, and Figure 5

shows the distribution of SHAP values and feature importance for different features. The bee swarm plots are used to visualize the continuous distribution of variables for different categories along with the SHAP values. They help to visualize how data points are spread out to reveal patterns or outliers within categories. NDVI, which measures live green vegetation based on satellite data, had lower importance in SHAP values for all periods due to the scarcity of data availability. This suggests that the model may not be as sensitive to changes in vegetation cover as it is to other factors such as weather patterns or human activity.

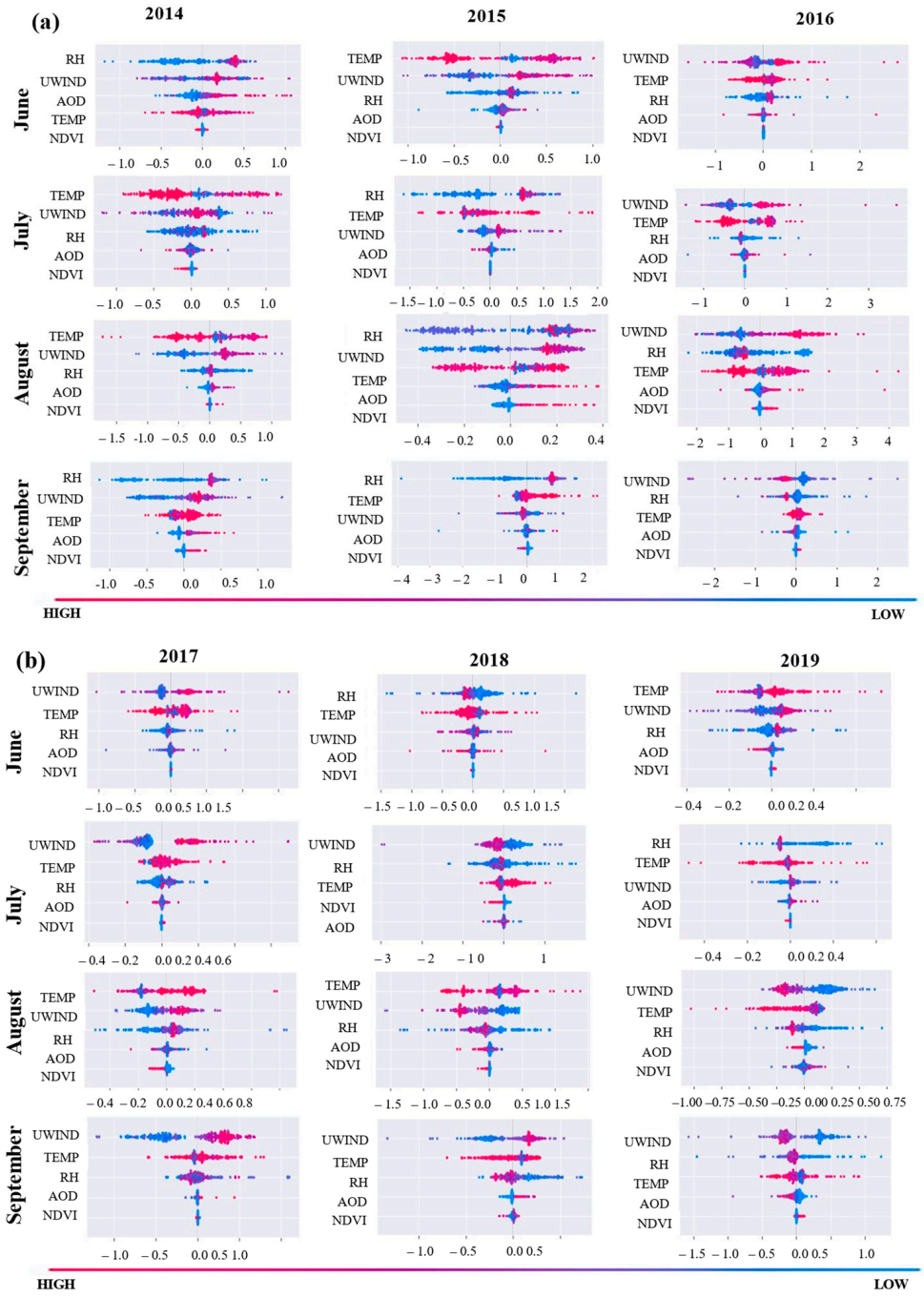

Figure 5. Distribution of SHAP values for RF Tree-based explainer using Bee swarm plot at feature importance levels: (**a**) June–September for the years 2014–2016 and (**b**) June–September for the years 2017–2019. The transition color ramp from red to blue indicates model output values from high to low.

UWIND, which measures the east–west wind speed component, had the highest mean SHAP value in 2016, indicating that it was the most important feature in predicting $PM_{2.5}$ levels for that year. This could be due to specific weather patterns or other more prevalent factors in 2016. RH which measures water vapor content in the air relative to its maximum capacity was an important feature in most months of 2015 except for June. This could be due to varying weather patterns or other factors that affected humidity levels during that year.

The mean SHAP value plot in Figure 6 provides a way to aggregate the SHAP values across all observations and calculate the average impact of each feature on the model output. The histogram shows the distribution of categorical feature importance based on SHAP values in high to low order. The bar plot shows that UWIND had the highest mean SHAP value among all the features, indicating that it was the most important feature overall in predicting $PM_{2.5}$ levels. This could be due to UWIND being strongly correlated with other features that are important in predicting the model output, or because it captures important information about the health or air condition that the model aims to predict. Overall, these findings provide insights into the factors that contribute to $PM_{2.5}$ variability and highlight the importance of considering multiple features in predicting $PM_{2.5}$ levels.

The heat map plot shown in Figure 7a,b presents the global interpretation. The heat map is used to represent the linearized density and continuity in the distribution of data, with values shown as colored lines. In August 2016, UWIND showed the highest feature importance with the highest MAE and MAPE values using the XGB-regression model. In contrast, RH was the dominating feature in Figure 5, with the lowest MAPE value with the XGB regressor in July 2019. Figure 7 shows that the RH variable in July 2019 improved the predictions for more than 150 observations, resulting in a contribution greater than 37% with SHAP values of 0.2. The contribution of temperature ranked second (5–6%), with SHAP values ranging between 20 and 25. Insignificant contribution, with SHAP values close to 0, was observed for the rest features. The SHAP values analysis and global explanation heatmaps provided insights into the model's behavior and the importance of different features in predicting $PM_{2.5}$ concentrations. The study found that UWIND was the most important feature in August 2016, while RH was the most important feature in July 2019.

The study also evaluated different explanation metrics applied to RF, GB, and XGBoost models using July and August 2019 datasets. The Random explainer had larger explanation errors in all the regression tree models, with good performance when excluding positive model output values (Figures S11–S16). The Partition, Permutation Part, Tree, Exact, and Permutation explainers resulted in fewer errors when explaining the RF model. The Tree, Partition, and Permutation Part explainers had fewer explanation errors with the July and August 2019 datasets.

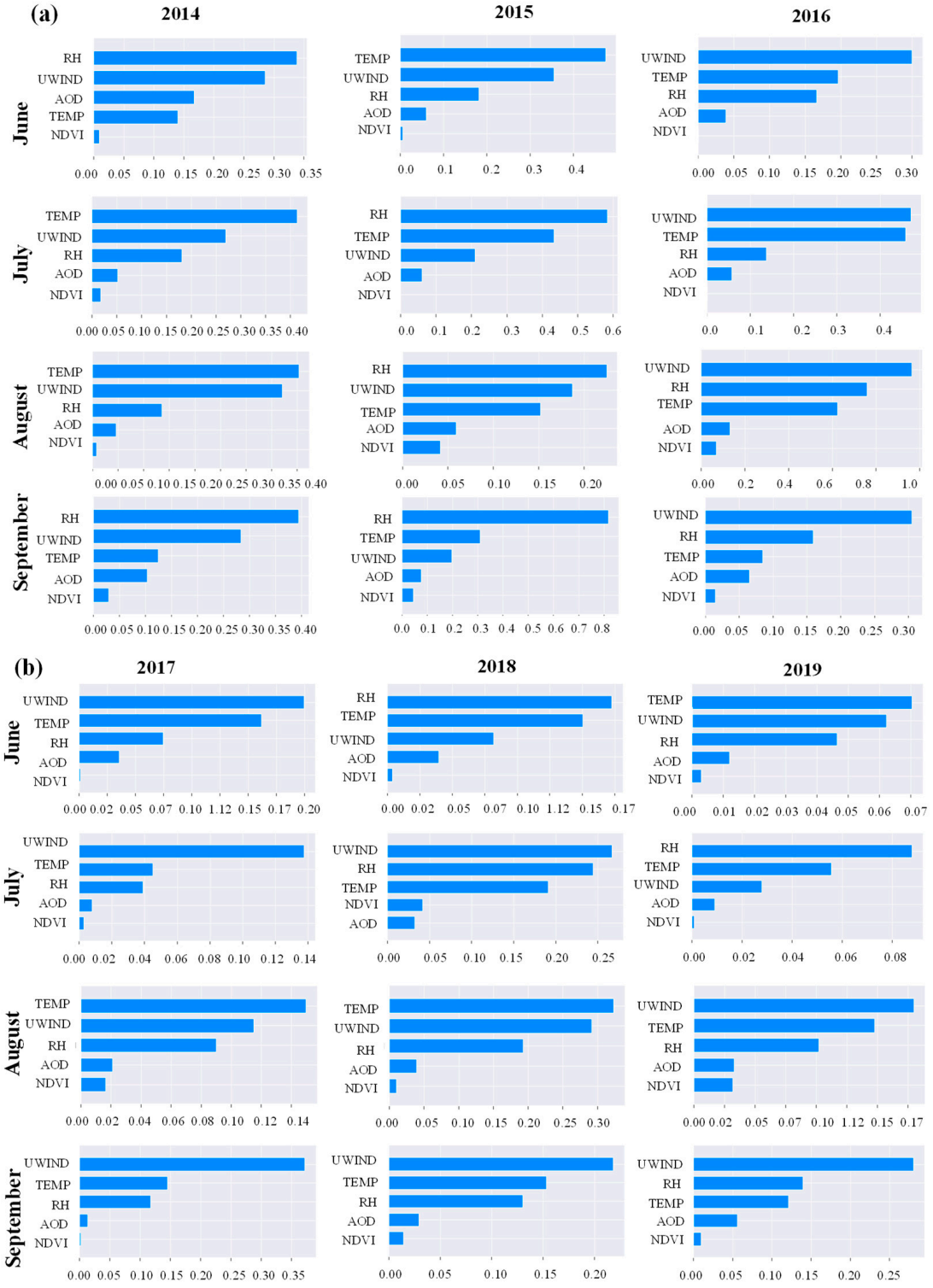

Figure 6. Distribution of SHAP values for RF Tree-based explainer at feature importance levels: (**a**) June–September for the years 2014–2016 and (**b**) June–September for the years 2017–2019. The distribution of feature importance levels ranging from high to low.

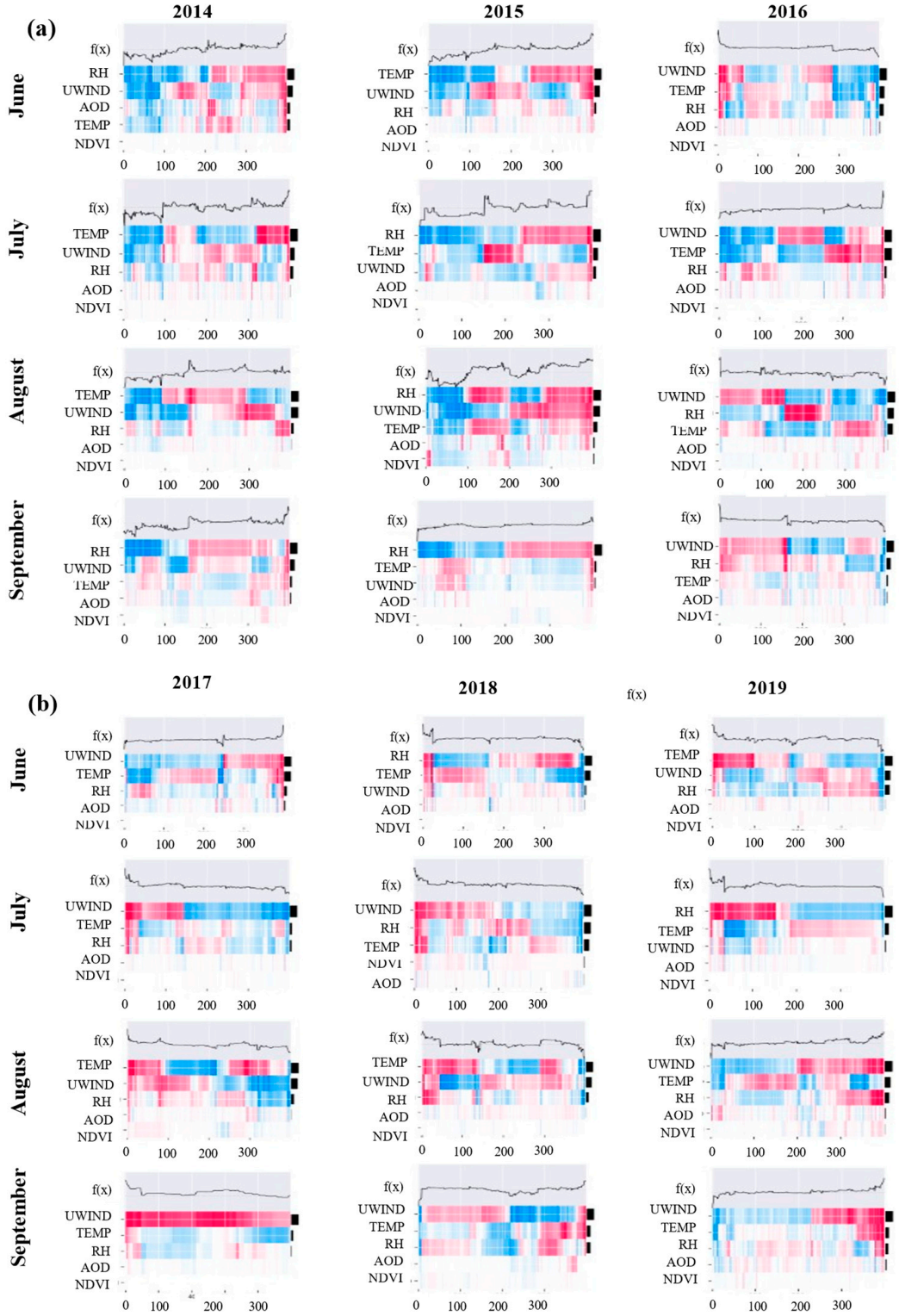

Figure 7. Two heat maps are shown: (**a**) June–September (2014–2016) and (**b**) June–September (2017–2019). Heatmap visualizations of SHAP values for RF Tree-based explainer and feature importance are shown in a color gradient ranging from high to low impact, with the model output displayed on the top *x*-axis in log odds. The *y*-axis shows the order of features by importance, and observations are clustered by function.

4.4. Spatio-Temporal Interpretation Using RF and GF Prediction

The interpretation of $PM_{2.5}$ temporal and spatial variations of both RF and GF algorithms is made from Figures 8 and S17. The high and lower levels in $PM_{2.5}$ concentrations are denoted in μg m^{-3}. The un-optimized algorithms are considered for the spatial and temporal variation based on the highest performance of GPI. August 2016 exhibited higher values when using RF compared to other periods. From RF prediction monthly variation of $PM_{2.5}$, the highest spatial density variation is found in central Singapore. Compared to other months in 2016, August had the highest variation in $PM_{2.5}$ concentrations than in the northern part with the highest and the lowest concentrations in the eastern part of Singapore. According to XAI analysis, UWIND exerts a strong influence among all the parameters even over the entire year of 2016; additionally, RH and TEMP also make significant contributions, although the intensity of variation was not the same.

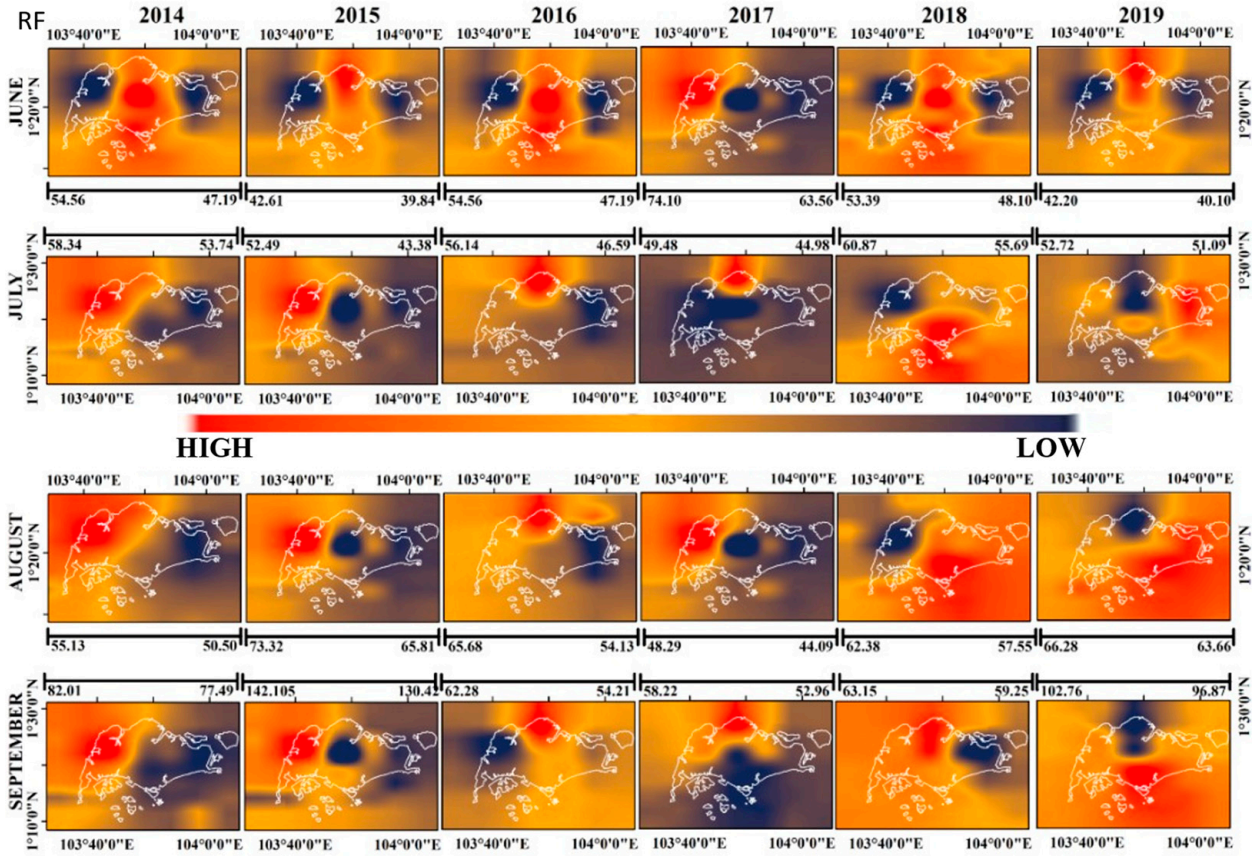

Figure 8. Spatial map using RF regression $PM_{2.5}$ concentration predictions for June–September (2014–2019).

Chen et al. [67] estimated that the overall R^2 values were around 0.88 using the RF model for locations in China, and Hu et al. [68] estimated that the daily $PM_{2.5}$ variation using RF models has achieved R^2 values of 0.80 for the United States. In our study in Singapore, the overall temporal mean R^2 values were around 0.65. The variation in the values is likely due to the dynamic variations in meteorological and topographic structures.

4.5. Insights, Strengths and Limitations

The changes in predicted outcomes of ML models for $PM_{2.5}$ concentrations, specific points are emphasized here for a clear insight:

- The structure of the ML algorithm: Depending on the structure of ML algorithms (RF, GB, and TP), the prediction mechanism and the spatial estimation of the outcome undergo changes. For instance, as an ensemble method, RF combines multiple decision trees to make predictions; GB builds an ensemble of decision trees sequentially, where each tree corrects the errors of the previous one; TP is an optimization algorithm used for hyperparameter tuning and model/feature selections. These predictions can capture complex relationships between the datasets.
- Integrating external factors for prediction: NDVI, temperature, wind speed, and humidity can impact the dispersion, transformation, and accumulation of PM particles in the air. Thus, these factors contribute more to the temporal and spatial dynamics of PM concentration.
- Need for spatial and time series prediction with data-driven analysis: Best ML prediction provides insights regarding how well $PM_{2.5}$ concentrations across different locations in the study area are predicted spatially. Time series and spatial prediction help to understand the yearly-based monthly patterns of $PM_{2.5}$ concentration included with meteorological variables' effects.

In the XAI-interpretable analysis with the RF model, meteorological variable variations like UWIND and RH strongly influenced $PM_{2.5}$ predictions due to dynamic variations in meteorological variables and seasonal influences [69]. The performance of the models considering the yearly variation showed significant variation in magnitudes of GPI values overall, especially in the years 2014, 2015, 2016, 2017, and 2018.

Among all the machine learning explainers, the tree-based explainers were found to predict better than others, which denotes that the tree-based models exhibit good performance in predicting $PM_{2.5}$. Even with the lack of spatial pixel information at some random areas, the tree models performed better in prediction. However, the availability of the data sets in both spatial and temporal resolutions is a limitation, which can affect the accuracy of outcomes. Particularly in the context of policy decisions related to air quality management, this analysis could be highly supportive.

Ji et al. [34] used RF to explore the potential impacts of several air-pollutant concentrations including $PM_{2.5}$ on the incidence of pediatric respiratory diseases in Taizhou, China. RF served as the best-performing model in [34], and this supports our results indeed. Gu et al. [33] used a hybrid model that combines an interpretable model and a deep neural network, achieving an overall RMSE of 15.0835. Bai et al. [70] using LSTM obtained RMSE values ranging from 13 to 14 for $PM_{2.5}$ prediction in China; comparatively, our predictions resulted in ~0.49 for RF and ~0.77 for TP, representing the highest values among the time series predictions.

5. Conclusions

The analysis of $PM_{2.5}$ variability using regression models and XAI provided insights into the important features and their contribution to predicting $PM_{2.5}$ concentrations. SHAP value analysis and global explanation provided a powerful means of achieving the prediction of $PM_{2.5}$ concentrations and ensuring the reliability and transparency of air quality models. Analyzing the feature contribution highlights the need for a comprehensive and dynamic approach to predicting air quality. In this study, different features contributed significantly over different periods for predicting $PM_{2.5}$. Although this prediction includes the lack of data availability and uncertainty (noises, coarse resolution, etc.), this feature contribution analysis was considerably good and reasonable, even when investigated directly. The Tree explainer performed comparatively well with other SHAP explainers on RF, GB, and XGBoost. RF considerably outperformed GB regarding MAE and MAPE values.

While comparing the outcomes of RF and GB without hypertuning the algorithm, RF showed good performance in predicting the $PM_{2.5}$ concentrations, while GB overestimated the prediction. Through hypertuning of the XGB regressor and comparing MSE, RMSE, MAE and MAPE, July 2019 $PM_{2.5}$ concentration values were better predicted compared to other periods, resulting in genetic iterations with the TP model. The study results also suggest that the XGBoost regression model is an effective tool for predicting $PM_{2.5}$ concentrations with lower MAE and MAPE values compared to other models. RH was the important feature for more significant prediction. Global explanation using SHAP values provided insights into the relative importance of the input features for both RF and GB models. UWIND was the most important feature in predicting $PM_{2.5}$ concentrations in August 2016, while RH was the most important in July 2019. This suggests that different features were important in different periods or under other environmental conditions, which significantly ensures the dependency of micro-climate.

Regression models and XAI tools are essential to predict and interpret $PM_{2.5}$ concentrations. Particularly in the context of policy decisions related to air quality management, this analysis could be highly supportive.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/ai4040040/s1, Figure S1: RF and GB local comparison line plots are shown at the top and overall contributions at the bottom for July 2019; Figure S2: RF and GB local comparison line plots are shown at the top and overall contributions at the bottom for August 2019; Figure S3: RF local comparison based on the direction for August 2019; Figure S4: Showing the local interpretation comparison plot for the AOD feature in RF for July 2019; Figure S4: cont. Same as Figure S4 but for the GB method; Figure S5: Showing the local interpretation comparison plot for the UWIND feature in RF and GB methods for July 2019; Figure S5: cont. Same as Figure S5 but for the RH feature; Figure S6: Showing the local interpretation comparison plot for the UWIND feature in RF and GB methods for August 2019; Figure S6: cont. Same as Figure S6 but for the RH feature; Figure S7: Showing the local interpretation comparison plot for the AOD feature in the RF method for August 2019; Figure S7: cont. Same as Figure S7 but for the GB method; Figure S8: RF local comparison based on the direction for July 2019; Figure S9: GB local comparison based on the direction for July 2019; Figure S10: GB local comparison based on the direction for August 2019; Figure S11: Displaying RF performance graph with and without random values on the left and explanation errors, computation time, and overall model output on the right for July 2019; Figure S12: Displaying GB performance graph with and without random values on the left and explanation errors, computation time, and overall model output on the right for July 2019; Figure S13: Displaying XGB performance graph with and without random values on the left and explanation errors, computation time, and overall model output on the right for July 2019; Figure S14: Displaying RF performance graph with and without random values on the left and explanation errors, computation time, and overall model output on the right for August 2019; Figure S15: Displaying GB performance graph with and without random values on the left and explanation errors, computation time, and overall model output on the right for August 2019; Figure S16: Displaying XGB performance graph with and without random values on the left and explanation errors, computation time, and overall model output on the right for August 2019; Figure S17: Spatial map using GB regression PM2.5 concentration predictions for June–September (2014–2019); Table S1: Best pipeline model given by TP algorithm for June–September (2014–2019).

Author Contributions: M.S.S.S.: Conceptualization, writing—original draft preparation, review and editing, and visualization; V.A.T.: conceptualization, writing—review and editing, and supervision; A.K.: conceptualization, writing—review and editing, and supervision; B.T.: conceptualization, writing—review and editing, and supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: MODIS datasets were acquired from https://ladsweb.modaps.eosdis.nasa.gov (accessed on 1 January 2023); ERA5 data available globally in grids with time scales for the fifth generation of the ECMWF from https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5 (accessed on 1 January 2023). Ambient air pollutants $PM_{2.5}$ acquired from the National Environmental Agency (NEA) of Singapore, https://www.nea.gov.sg (accessed on 1 January 2023). Analyzed information of this study may be available upon reasonable request.

Acknowledgments: The authors would like to acknowledge the National Institute of Technology Rourkela for providing lab facilities. The authors are thankful to Yuming Guo, Monash University, for his valuable suggestions and stimulating scientific discussions. We are also thankful to the National Environmental Agency, Singapore, for providing the in situ observations of $PM_{2.5}$ over Singapore.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chae, S.; Shin, J.; Kwon, S.; Lee, S.; Kang, S.; Lee, D. PM_{10} and $PM_{2.5}$ Real-Time Prediction Models Using an Interpolated Convolutional Neural Network. *Sci. Rep.* **2021**, *11*, 11952. [CrossRef] [PubMed]
2. Jat, R.; Gurjar, B.R. Contribution of Different Source Sectors and Source Regions of Indo-Gangetic Plain in India to $PM_{2.5}$ Pollution and Its Short-Term Health Impacts during Peak Polluted Winter. *Atmos. Pollut. Res.* **2021**, *12*, 89–100. [CrossRef]
3. Naghan, D.J.; Neisi, A.; Goudarzi, G.; Dastoorpoor, M.; Fadaei, A.; Angali, K.A. Estimation of the Effects $PM_{2.5}$, NO_2, O_3 Pollutants on the Health of Shahrekord Residents Based on AirQ+ Software during (2012–2018). *Toxicol. Rep.* **2022**, *9*, 842–847. [CrossRef] [PubMed]
4. Bai, H.; Shi, Y.; Seong, M.; Gao, W.; Li, Y. Influence of Spatial Resolution on Satellite-Based $PM_{2.5}$ Estimation: Implications for Health Assessment. *Remote Sens.* **2022**, *14*, 2933. [CrossRef]
5. Balasubramanian, R. Comprehensive Characterization of $PM_{2.5}$ Aerosols in Singapore. *J. Geophys. Res.* **2003**, *108*, 4523. [CrossRef]
6. Li, Y.; Zhu, Y.; Tan, J.Y.K.; Teo, H.C.; Law, A.; Qu, D.; Luo, W. The Impact of COVID-19 on NO_2 and $PM_{2.5}$ Levels and Their Associations with Human Mobility Patterns in Singapore. *Ann. GIS* **2022**, *28*, 515–531. [CrossRef]
7. Fang, G.; Wu, Y.; Huang, S.; Rau, J. Review of Atmospheric Metallic Elements in Asia during 2000–2004. *Atmos. Environ.* **2005**, *39*, 3003–3013. [CrossRef]
8. Lelieveld, J.; Evans, J.S.; Fnais, M.; Giannadaki, D.; Pozzer, A. The Contribution of Outdoor Air Pollution Sources to Premature Mortality on a Global Scale. *Nature* **2015**, *525*, 367–371. [CrossRef]
9. Burnett, R.T.; Arden Pope, C.; Ezzati, M.; Olives, C.; Lim, S.S.; Mehta, S.; Shin, H.H.; Singh, G.; Hubbell, B.; Brauer, M.; et al. An Integrated Risk Function for Estimating the Global Burden of Disease Attributable to Ambient Fine Particulate Matter Exposure. *Environ. Health Perspect.* **2014**, *122*, 397–403. [CrossRef]
10. Guo, H.; Kota, S.H.; Chen, K.; Sahu, S.K.; Hu, J.; Ying, Q.; Wang, Y.; Zhang, H. Source Contributions and Potential Reductions to Health Effects of Particulate Matter in India. *Atmos. Chem. Phys.* **2018**, *18*, 15219–15229. [CrossRef]
11. Zhu, W.; Wang, M.; Zhang, B. The Effects of Urbanization on $PM_{2.5}$ Concentrations in China's Yangtze River Economic Belt: New Evidence from Spatial Econometric Analysis. *J. Clean. Prod.* **2019**, *239*, 118065. [CrossRef]
12. Chen, G.; Li, S.; Knibbs, L.D.; Hamm, N.A.S.; Cao, W.; Li, T.; Guo, J.; Ren, H.; Abramson, M.J.; Guo, Y. A Machine Learning Method to Estimate $PM_{2.5}$ Concentrations across China with Remote Sensing, Meteorological and Land Use Information. *Sci. Total Environ.* **2018**, *636*, 52–60. [CrossRef] [PubMed]
13. Chen, Y.-C.; Li, D.-C. Selection of Key Features for $PM_{2.5}$ Prediction Using a Wavelet Model and RBF-LSTM. *Appl. Intell.* **2021**, *51*, 2534–2555. [CrossRef]
14. Ma, Z.; Dey, S.; Christopher, S.; Liu, R.; Bi, J.; Balyan, P.; Liu, Y. A Review of Statistical Methods Used for Developing Large-Scale and Long-Term $PM_{2.5}$ Models from Satellite Data. *Remote Sens. Environ.* **2022**, *269*, 112827. [CrossRef]
15. Pu, Q.; Yoo, E.H. Ground $PM_{2.5}$ Prediction Using Imputed MAIAC AOD with Uncertainty Quantification. *Environ. Pollut.* **2021**, *274*, 116574. [CrossRef] [PubMed]
16. Mabasa, B.; Lysko, M.D.; Moloi, S.J. Validating Hourly Satellite Based and Reanalysis Based Global Horizontal Irradiance Datasets over South Africa. *Geomatics* **2021**, *1*, 429–449. [CrossRef]

17. Gupta, P.; Levy, R.C.; Mattoo, S.; Remer, L.A.; Munchak, L.A. A Surface Reflectance Scheme for Retrieving Aerosol Optical Depth over Urban Surfaces in MODIS Dark Target Retrieval Algorithm. *Atmos. Meas. Tech.* **2016**, *9*, 3293–3308. [CrossRef]
18. Sekertekin, A.; Inyurt, S.; Yaprak, S. Pre-Seismic Ionospheric Anomalies and Spatio-Temporal Analyses of MODIS Land Surface Temperature and Aerosols Associated with Sep, 24 2013 Pakistan Earthquake. *J. Atmos. Sol.-Terr. Phys.* **2020**, *200*, 105218. [CrossRef]
19. Xiang, Y.; Ye, Y.; Peng, C.; Teng, M.; Zhou, Z. Seasonal Variations for Combined Effects of Landscape Metrics on Land Surface Temperature (LST) and Aerosol Optical Depth (AOD). *Ecol. Indic.* **2022**, *138*, 108810. [CrossRef]
20. Shukla, K.; Kumar, P.; Mann, G.S.; Khare, M. Mapping Spatial Distribution of Particulate Matter Using Kriging and Inverse Distance Weighting at Supersites of Megacity Delhi. *Sustain. Cities Soc.* **2020**, *54*, 101997. [CrossRef]
21. Feng, X.; Li, Q.; Zhu, Y.; Hou, J.; Jin, L.; Wang, J. Artificial Neural Networks Forecasting of PM2.5 Pollution Using Air Mass Trajectory Based Geographic Model and Wavelet Transformation. *Atmos. Environ.* **2015**, *107*, 118–128. [CrossRef]
22. Beucler, T.; Ebert-Uphoff, I.; Rasp, S.; Pritchard, M.; Gentine, P. Machine Learning for Clouds and Climate. In *Clouds and Climate*; Cambridge University Press: Cambridge, UK, 2020.
23. Ghorbanzadeh, O.; Blaschke, T.; Gholamnia, K.; Meena, S.R.; Tiede, D.; Aryal, J. Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional Neural Networks for Landslide Detection. *Remote Sens.* **2019**, *11*, 196. [CrossRef]
24. Ma, J.; Ding, Y.; Gan, V.J.L.; Lin, C.; Wan, Z. Spatiotemporal Prediction of PM2.5 Concentrations at Different Time Granularities Using IDW-BLSTM. *IEEE Access* **2019**, *7*, 107897–107907. [CrossRef]
25. Nazar, M.; Alam, M.M.; Yafi, E.; Su'ud, M.M. A Systematic Review of Human–Computer Interaction and Explainable Artificial Intelligence in Healthcare With Artificial Intelligence Techniques. *IEEE Access* **2021**, *9*, 153316–153348. [CrossRef]
26. Martinez-Seras, A.; Del Ser, J.; Garcia-Bringas, P. Can Post-Hoc Explanations Effectively Detect Out-of-Distribution Samples? In Proceedings of the IEEE International Conference on Fuzzy Systems, Padua, Italy, 18–23 July 2022.
27. Huang, F.; Shangguan, W.; Li, Q.; Li, L.; Zhang, Y. Beyond Prediction: An Integrated Post–Hoc Approach to Interpret Complex Model in Hydrometeorology. *Environ. Model. Softw.* **2022**, *167*, 105762. [CrossRef]
28. Kakogeorgiou, I.; Karantzalos, K. Evaluating Explainable Artificial Intelligence Methods for Multi-Label Deep Learning Classification Tasks in Remote Sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *103*, 102520. [CrossRef]
29. Singh, H.; Roy, A.; Setia, R.K.; Pateriya, B. Estimation of Nitrogen Content in Wheat from Proximal Hyperspectral Data Using Machine Learning and Explainable Artificial Intelligence (XAI) Approach. *Model. Earth Syst. Environ.* **2022**, *8*, 2505–2511. [CrossRef]
30. Stadtler, S.; Betancourt, C.; Roscher, R. Explainable Machine Learning Reveals Capabilities, Redundancy, and Limitations of a Geospatial Air Quality Benchmark Dataset. *Mach. Learn. Knowl. Extr.* **2022**, *4*, 150–171. [CrossRef]
31. Betancourt, C.; Stomberg, T.T.; Edrich, A.K.; Patnala, A.; Schultz, M.G.; Roscher, R.; Kowalski, J.; Stadtler, S. Global, High-Resolution Mapping of Tropospheric Ozone-Explainable Machine Learning and Impact of Uncertainties. *Geosci. Model Dev.* **2022**, *15*, 4331–4354. [CrossRef]
32. Stirnberg, R.; Cermak, J.; Kotthaus, S.; Haeffelin, M.; Andersen, H.; Fuchs, J.; Kim, M.; Petit, J.E.; Favez, O. Meteorology-Driven Variability of Air Pollution (PM1) Revealed with Explainable Machine Learning. *Atmos. Chem. Phys.* **2021**, *21*, 3919–3948. [CrossRef]
33. Gu, Y.; Li, B.; Meng, Q. Hybrid Interpretable Predictive Machine Learning Model for Air Pollution Prediction. *Neurocomputing* **2022**, *468*, 123–136. [CrossRef]
34. Ji, Y.; Zhi, X.; Wu, Y.; Zhang, Y.; Yang, Y.; Peng, T.; Ji, L. Regression Analysis of Air Pollution and Pediatric Respiratory Diseases Based on Interpretable Machine Learning. *Front. Earth Sci.* **2023**, *11*, 1105140. [CrossRef]
35. Tan, S.T.; Mohamed, N.; Ng, L.C.; Aik, J. Air Quality in Underground Metro Station Commuter Platforms in Singapore: A Cross-Sectional Analysis of Factors Influencing Commuter Exposure Levels. *Atmos. Environ.* **2022**, *273*, 118962. [CrossRef]
36. Land Transport Authority Public Transport Ridership. Available online: https://www.lta.gov.sg/content/dam/ltagov/who_we_are/statistics_and_publications/statistics/pdf/PT_Ridership_2015_2019.pdf (accessed on 3 August 2021).
37. Government of Singapore Total Land Area of Singapore. Available online: https://data.gov.sg/dataset/total-land-area-of-singapore (accessed on 3 August 2021).
38. Barudgar, A.; Singh, J.; Tyagi, B. Variability of Fine Particulate Matter ($PM_{2.5}$) and Its Association with Health and Vehicular Emissions Over an Urban Tropical Coastal Station Mumbai, India. *Thalassas* **2022**, *38*, 1067–1080. [CrossRef]
39. Sahu, S.K.; Tyagi, B.; Pradhan, C.; Beig, G. Evaluating the Variability, Transport and Periodicity of Particulate Matter over Smart City Bhubaneswar, a Tropical Coastal Station of Eastern India. *SN Appl. Sci.* **2019**, *1*, 383. [CrossRef]
40. Gogikar, P.; Tyagi, B.; Gorai, A.K. Seasonal Prediction of Particulate Matter over the Steel City of India Using Neural Network Models. *Model. Earth Syst. Environ.* **2019**, *5*, 227–243. [CrossRef]
41. Hari, M.; Tyagi, B. India's Greening Trend Seems to Slow Down. What Does Aerosol Have to Do with It? *Land* **2022**, *11*, 538. [CrossRef]
42. Sahu, R.K.; Hari, M.; Tyagi, B. Forest Fire Induced Air Pollution over Eastern India during March 2021. *Aerosol Air Qual. Res.* **2022**, *22*, 220084. [CrossRef]

43. Gogikar, P.; Tyagi, B.; Padhan, R.R.; Mahaling, M. Particulate Matter Assessment Using In Situ Observations from 2009 to 2014 over an Industrial Region of Eastern India. *Earth Syst. Environ.* **2018**, *2*, 305–322. [CrossRef]
44. Gogikar, P.; Tyagi, B. Assessment of Particulate Matter Variation during 2011–2015 over a Tropical Station Agra, India. *Atmos. Environ.* **2016**, *147*, 11–21. [CrossRef]
45. Peng-in, B.; Sanitluea, P.; Monjatturat, P.; Boonkerd, P.; Phosri, A. Estimating Ground-Level $PM_{2.5}$ over Bangkok Metropolitan Region in Thailand Using Aerosol Optical Depth Retrieved by MODIS. *Air Qual. Atmos. Health* **2022**, *15*, 2091–2102. [CrossRef]
46. Sethi, J.K.; Mittal, M. Monitoring the Impact of Air Quality on the COVID-19 Fatalities in Delhi, India: Using Machine Learning Techniques. *Disaster Med. Public Health Prep.* **2020**, *16*, 604–611. [CrossRef] [PubMed]
47. Mustakim, N.A.; Ul-Saufie, A.Z.; Shaziayani, W.N.; Noor, N.M.; Mutalib, S. Prediction of Daily Air Pollutants Concentration and Air Pollutant Index Using Machine Learning Approach. *Pertanika J. Sci. Technol.* **2023**, *31*, 123–135. [CrossRef]
48. Gautam, S.; Patra, A.K.; Brema, J.; Raj, P.V.; Raimond, K.; Abraham, S.S.; Chudugudu, K.R. Prediction of Various Sizes of Particles in Deep Opencast Copper Mine Using Recurrent Neural Network: A Machine Learning Approach. *J. Inst. Eng. Ser. A* **2022**, *103*, 283–294. [CrossRef]
49. Adong, P.; Bainomugisha, E.; Okure, D.; Sserunjogi, R. Applying Machine Learning for Large Scale Field Calibration of Low-cost $PM_{2.5}$ and PM_{10} Air Pollution Sensors. *Appl. AI Lett.* **2022**, *3*, e76. [CrossRef]
50. Ha, N.T.; Manley-Harris, M.; Pham, T.D.; Hawes, I. The Use of Radar and Optical Satellite Imagery Combined with Advanced Machine Learning and Metaheuristic Optimization Techniques to Detect and Quantify above Ground Biomass of Intertidal Seagrass in a New Zealand Estuary. *Int. J. Remote Sens.* **2021**, *42*, 4712–4738. [CrossRef]
51. Bergstra, J.; Bengio, Y. Random Search for Hyper-Parameter Optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
52. Eiben, A.E. Introduction to Evolutionary Computing. *Assem. Autom.* **2004**, *24*, 324. [CrossRef]
53. Nunnari, G. Modelling Air Pollution Time-Series by Using Wavelet Functions and Genetic Algorithms. *Soft Comput.* **2004**, *8*, 173–178. [CrossRef]
54. Saini, J.; Dutta, M.; Marques, G. A Novel Application of Fuzzy Inference System Optimized with Particle Swarm Optimization and Genetic Algorithm for PM_{10} Prediction. *Soft Comput.* **2022**, *26*, 9573–9586. [CrossRef]
55. Garouani, M.; Ahmad, A.; Bouneffa, M.; Hamlich, M.; Bourguin, G.; Lewandowski, A. Using Meta-Learning for Automated Algorithms Selection and Configuration: An Experimental Framework for Industrial Big Data. *J. Big Data* **2022**, *9*, 57. [CrossRef]
56. Le, T.T.; Fu, W.; Moore, J.H. Scaling Tree-Based Automated Machine Learning to Biomedical Big Data with a Feature Set Selector. *Bioinformatics* **2020**, *36*, 250–256. [CrossRef] [PubMed]
57. Olson, R.S.; Bartley, N.; Urbanowicz, R.J.; Moore, J.H. Evaluation of a Tree-Based Pipeline Optimization Tool for Automating Data Science. In Proceedings of the GECCO 2016—2016 Genetic and Evolutionary Computation Conference, Denver, CO, USA, 20–24 July 2016.
58. Delavar, M.R.; Gholami, A.; Shiran, G.R.; Rashidi, Y.; Nakhaeizadeh, G.R.; Fedra, K.; Afshar, S.H. A Novel Method for Improving Air Pollution Prediction Based on Machine Learning Approaches: A Case Study Applied to the Capital City of Tehran. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 99. [CrossRef]
59. Srivastava, C.; Singh, S.; Singh, A.P. Estimation of Air Pollution in Delhi Using Machine Learning Techniques. In Proceedings of the 2018 International Conference on Computing, Power and Communication Technologies, GUCON 2018, Greater Noida, India, 28–29 September 2018.
60. Khan, M.A.; Kim, H.C.; Park, H. Leveraging Machine Learning for Fault-Tolerant Air Pollutants Monitoring for a Smart City Design. *Electronics* **2022**, *11*, 3122. [CrossRef]
61. Arun, G.; Rathi, S. Real Time Air Quality Evaluation Model Using Machine Learning Approach. *J. Inf. Technol. Digit. World* **2022**, *4*, 23–33. [CrossRef]
62. Zhu, B.; Feng, Y.; Gong, D.; Jiang, S.; Zhao, L.; Cui, N. Hybrid Particle Swarm Optimization with Extreme Learning Machine for Daily Reference Evapotranspiration Prediction from Limited Climatic Data. *Comput. Electron. Agric.* **2020**, *173*, 105430. [CrossRef]
63. Aas, K.; Jullum, M.; Løland, A. Explaining Individual Predictions When Features Are Dependent: More Accurate Approximations to Shapley Values. *Artif. Intell.* **2021**, *298*, 103502. [CrossRef]
64. Lundberg, S.M.; Lee, S.; Lundberg, S.M.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems*; NIPS 2017; The MIT Press: Cambridge, MA, USA, 2017; pp. 4765–4774.
65. Jiang, H.; Senge, E. On Two XAI Cultures: A Case Study of Non-Technical Explanations in Deployed AI System. *arXiv* **2021**, arXiv:2112.01016.
66. Conibear, L.; Reddington, C.L.; Silver, B.J.; Chen, Y.; Knote, C.; Arnold, S.R.; Spracklen, D.V. Statistical Emulation of Winter Ambient Fine Particulate Matter Concentrations From Emission Changes in China. *GeoHealth* **2021**, *5*, e2021GH000391. [CrossRef]
67. Chen, Z.-Y.; Jin, J.-Q.; Zhang, R.; Zhang, T.-H.; Chen, J.-J.; Yang, J.; Ou, C.-Q.; Guo, Y. Comparison of Different Missing-Imputation Methods for MAIAC (Multiangle Implementation of Atmospheric Correction) AOD in Estimating Daily $PM_{2.5}$ Levels. *Remote Sens.* **2020**, *12*, 3008. [CrossRef]
68. Hu, X.; Belle, J.H.; Meng, X.; Wildani, A.; Waller, L.A.; Strickland, M.J.; Liu, Y. Estimating $PM_{2.5}$ Concentrations in the Conterminous United States Using the Random Forest Approach. *Environ. Sci. Technol.* **2017**, *51*, 6936–6944. [CrossRef]

69. Yang, H.; Liu, Z.; Li, G. A New Hybrid Optimization Prediction Model for $PM_{2.5}$ Concentration Considering Other Air Pollutants and Meteorological Conditions. *Chemosphere* **2022**, *307*, 135798. [CrossRef]
70. Bai, Y.; Zeng, B.; Li, C.; Zhang, J. An Ensemble Long Short-Term Memory Neural Network for Hourly $PM_{2.5}$ Concentration Forecasting. *Chemosphere* **2019**, *222*, 286–294. [CrossRef]