

Article

Predicting Fine Particulate Matter (PM2.5) in the Greater London Area: An Ensemble Approach using Machine Learning Methods

Mahdieh Danesh Yazdi ^{1,*}, Zheng Kuang ^{1,2}, Konstantina Dimakopoulou ³, Benjamin Barratt ⁴, Esra Suel ⁵, Heresh Amini ^{1,6}, Alexei Lyapustin ⁷, Klea Katsouyanni ^{3,4} and Joel Schwartz ^{1,8}

- ¹ Department of Environmental Health, Harvard TH Chan School of Public Health, Boston, MA 02115, USA; zhk746@mail.harvard.edu (Z.K.); heresh.amini@sund.ku.dk (H.A.); joel@hsph.harvard.edu (J.S.)
- ² Foundation Medicine, Cambridge, MA 02141, UK
- ³ Department of Hygiene, Epidemiology and Medical Statistics, Medical School, National and Kapodistrian University of Athens, 115 27 Athens, Greece; kdimakop@med.uoa.gr (K.D.); klea.katsouyanni@kcl.ac.uk (K.K.)
- ⁴ MRC Centre for Environment and Health, King's College London, London SE1 9NH, UK; benjamin.barratt@kcl.ac.uk
- ⁵ Department of Global Environmental Health, Imperial College, London SW7 2AZ, UK; esra.suel@imperial.ac.uk
- ⁶ Environmental Epidemiology Group, Section of Environmental Health, Department of Public Health, Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark
- ⁷ Climate and Radiation Laboratory, Goddard Space Flight Center, NASA, Greenbelt, MD 20771, USA; alexei.i.lyapustin@nasa.gov
- ⁸ Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA 02115, USA
- * Correspondence: mahdieh_daneshyazdi@g.harvard.edu

Received: 15 January 2020; Accepted: 10 March 2020; Published: 12 March 2020



Abstract: Estimating air pollution exposure has long been a challenge for environmental health researchers. Technological advances and novel machine learning methods have allowed us to increase the geographic range and accuracy of exposure models, making them a valuable tool in conducting health studies and identifying hotspots of pollution. Here, we have created a prediction model for daily PM_{2.5} levels in the Greater London area from 1st January 2005 to 31st December 2013 using an ensemble machine learning approach incorporating satellite aerosol optical depth (AOD), land use, and meteorological data. The predictions were made on a 1 km × 1 km scale over 3960 grid cells. The ensemble included predictions from three different machine learners: a random forest (RF), a gradient boosting machine (GBM), and a k-nearest neighbor (KNN) approach. Our ensemble model performed very well, with a ten-fold cross-validated R² of 0.828. Of the three machine learners, the random forest outperformed the GBM and KNN. Our model was particularly adept at predicting day-to-day changes in PM_{2.5} levels with an out-of-sample temporal R² of 0.882. However, its ability to predict spatial variability was weaker, with a R² of 0.396. We believe this to be due to the smaller spatial variation in pollutant levels in this area.

Keywords: air pollution; particulate matter; machine learning; exposure modeling

1. Introduction

Environmental research has long dealt with issues in exposure assessment, particularly in studies involving air pollutants. Direct individual measurements using personal monitors are costly, difficult to implement, and inconvenient for participants—effectively limiting the number of individuals that may be recruited and the length of time that exposures are measured. Moreover, some monitors have



large exposure measurement bias [1]. On the other hand, data from centrally placed monitors are limited, both spatially and temporally. Using these would restrict the populations and areas that we could study epidemiologically, increase uncertainty, and potentially introduce bias into health impact assessments. Modeling exposures offers a solution that addresses these deficiencies.

The most widely studied component of air pollution is fine particulate matter, or PM_{2.5}, which refers to particles suspended in the air with an aerodynamic diameter of less than 2.5 micrometers. These particles are of particular concern, as they are small enough to penetrate deep into the respiratory system, cross biological membranes, and cause systemic damage [2]. Exposure to PM_{2.5} has been linked with mortality [3–9], cardiovascular outcomes [8,10–12], cerebrovascular outcomes [8,10,12,13], respiratory outcomes [8,10,12], neurological outcomes [14,15], etc. These associations have been observed even at levels below current regulatory standards, suggesting that no amount of particulate matter in ambient air is safe [6,7,10,16]. Furthermore, it affects at least 95% of individuals [17], making PM_{2.5} an essential component of current and future environmental research.

In recent years, the modeling of PM_{2.5} has centered around the use of aerosol optical depth (AOD) obtained from satellite data as a key predictor of fine particulate matter levels [18–22]. Traditional modeling has usually relied on land-use regression (LUR) models and chemical transport models (CTMs), to describe the relationship between predictors of PM_{2.5} and measured PM_{2.5} levels, with later models incorporating both into generating predictions [23–30]. However, LURs are limited in their ability to capture temporal variation and CTMs often do not correspond well with surface pollution levels on their own.

Novel machine learning techniques allow us to create models with greater accuracy and flexibility that can combine remote sensing, land use, meteorological, and CTM inputs. They are also better at incorporating temporal variation than standard LURs. Machine learning algorithms allow us to non-parametrically examine the relationship between the predictors of pollutant concentrations and measured pollutant concentrations [28,31–36]. By allowing the learner to designate the predictor-outcome relationship, this burden is removed from the researcher and automated, saving time and making the most of both the data and the computational infrastructure now available. Overfitting is avoided by tuning models, based on performance on held out measurements. These methods also provide us with measures of variable importance, so that key consistent predictors may be identified. These predictors can then be used in other models, as well as to inform mitigation policies.

The primary utility of these exposure models is their use in epidemiological studies. They ease exposure assignment to the unit of study, and with greater accuracy, reduce the amount of measurement error of exposure, and thus are preferable for use in statistical analyses [5,6,10,37–39]. The output may also be used to identify areas with high exposures, as a predictor in other prediction models, and in risk assessments.

Here, we incorporate AOD, land-use data, and meteorological data to predict $PM_{2.5}$ levels on a 1 km × 1 km scale, from 1st January 2005 to 31st December 2013 in the Greater London area, using an ensemble model and four machine learning algorithms, which were calibrated using data derived from a wide network of monitors.

2. Materials and Methods

We created a prediction model for daily $PM_{2.5}$ in the Greater London area from 1st January 2005 to 31st December 2013. The predictions were made on a 1 km × 1 km scale. The total map consisted of 3960 grid cells. In order to optimize the predictions, we utilized four different machine learning methods: a gradient boosting machine (GBM), a random forest (RF), a deep neural network (NN), and a k-nearest neighbor (KNN) learner. The predictions from each model were put into a final generalized additive model (GAM) with a smoothed spatial term, to obtain the final average daily $PM_{2.5}$ levels in each grid cell. We expected that the use of an ensemble approach would result in better predictions than those from a single model [40].

2.1. Machine Learning Algorithms

The detailed specifics of how these machine learning algorithms work have been described elsewhere [41–44]. The GBM and RF are both decision tree machine learning algorithms. The GBM is a boosting decision tree method. This means that weak learners are created and the residuals from these models are used to create stronger learners, with previous models essentially "boosting" subsequent models and improving the predictions [41]. In a random forest, a large number of decision trees are constructed and predictions from these individual trees are averaged to obtain the output [42]. The neural network, on the other hand, takes the input variables, much like a neuron responds to stimuli, processes them through various combinations and weights, and generates predictions [43]. The k-nearest neighbor algorithm relies on the assumption that there is a proximal relationship between values. It calculates a weighted average from the closest designated "k" neighbors, in order to generate predictions [44]. Each machine learner was run separately on the data.

The individual machine learners were ensemble-averaged using a GAM, which included a smoothed function of the predictions from each individual learner, plus a smoothed function of latitude and longitude. The predictions from this GAM are the ensemble-averaged predictions. The smoothing terms allow the weights given to each learner to vary with the pollution level, in case one learner performs better in a specific range of $PM_{2.5}$.

2.2. Input Variables

The covariates used in training the models were population density (persons/km²), cloudiness (okta), barometric pressure (mBar/hPa), wind direction (°N), wind speed (m/s), dew point temperature (°C), temperature (°C), aerosol optical depth (AOD), land use type, distance to water (km), distance to Heathrow airport (m), inverse of the height of the planetary boundary layer (m^{-1}) , normalized difference vegetation index (NDVI), traffic counts, sine of day of the year, cosine of day of the year, day of week, number of days from time of origin (1st January 2005), average daily PM_{2.5} across the greater London area (μ g/m³), year, light at night, elevation (m), distance to nearest major road (km), length of major road (km) in grid cell, number of bus stops in grid cell, distance to nearest bus stop (km), average building height (m), and number of buildings in the grid cell. We selected these variables, primarily based on expert knowledge and data availability. Meteorological variables are known to affect the dispersion and transport of fine particulate matter. Land-use variables represent potential sources of PM_{2.5} and areas of higher concern. The time variables account for the seasonal variation in PM_{2.5} levels and the trend over several years. As previously mentioned, AOD is a key predictor of PM_{2.5}, with higher levels of AOD indicating higher PM_{2.5} levels [18–22]. We also used cross-validated R² values in the GBM algorithm to determine whether additional variables would improve model performance in deciding whether to include them or not.

2.3. Data Sources

The meteorological variables were obtained from the UK Meteorological Office. Traffic counts were obtained from the Department of Transport in the United Kingdom. The land use type was derived from the Land Cover Map of Great Britain from 2007. The average building height, number of buildings, distance to nearest major road from grid cell centroid, and length of major road in grid cell were calculated using data provided by the Ordnance Survey. Elevation data was obtained from the CGIAR Consortium for Spatial Information, who used Shuttle Radar Topography Mission (SRTM) data from the United States Geological Survey (USGS) and NASA. Data on bus stops was obtained from Transport for London on the London Datastore website.

We used AOD from the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument on the Aqua and Terra satellites, as provided by the MAIAC algorithm at 1 km² resolution [45]. There were missing AOD values due to cloud cover and snow reflectance. We imputed these missing values using a random forest approach and land use and meteorological predictors. The random forest approach

had an internal cross-validation R^2 score of 0.913 with a root mean square error (RMSE) of 0.037 for AOD-terra and an R^2 of 0.914 and RMSE of 0.036 for AOD-aqua, indicating very good predictive ability for both these variables. NDVI was also obtained from MODIS satellite measurements. It was measured every 16 days. Daily values were imputed spatially and temporally.

Mean annual light at night from 2015 was obtained from the Visible Infrared Imaging Radiometer Suite (VIIRS) satellite, at a horizontal resolution of 750 m.

Population density was obtained on a 1 km × 1 km scale from Columbia University's Center for International Earth Science Information Network (CIESIN) [46].

2.4. PM_{2.5} Data

Predicted $PM_{2.5}$ was compared to measured $PM_{2.5}$ to train the models and assess the accuracy of the various methods used. Measured $PM_{2.5}$ was obtained from 24 monitoring sites for 2005–2008 and 33 sites for 2009–2013 across the Greater London area. Measured PM_{10} information was obtained at 108 monitoring sites for 2005–2008 and 115 monitoring sites from 2009–2010. These values were taken from the London Air Quality Network (londonair.org.uk) [47] and the UK Automatic Urban and Rural Network (uk-air.defra.gov.uk) [48]. In order to predict $PM_{2.5}$ at fixed sites with no $PM_{2.5}$ measurements, but which measured PM_{10} and NO_x , we used two methods: (1) a regression model and (2) a random forest approach. The predictions from the two methods were used as independent variables in a generalized additive model (GAM), in order to improve the fit of the model and therefore provide a greatly enhanced database of $PM_{2.5}$ estimated concentrations. These values were then treated as our measured $PM_{2.5}$, which we used for training and cross-validation. All PM measurements were gravimetric equivalent [49]. Overall, measurements from 124 sites were used. The location of these monitors can be seen in Figure 1.



Figure 1. Location of monitoring sites in the Greater London area.

2.5. Hyper-Parameter Tuning

Although machine learning methods are non-parametric and do not require distributional assumptions, they do require the specification of hyper-parameters, parameters that control the learning process. In order to optimize the hyper-parameters for the algorithms, we used a grid search and looked at the mean square error (MSE) and cross-validated R^2 values. For the gradient boosting machine, we looked at the following hyper-parameters: number of trees, maximum tree depth, column sample rate, and learning rate. For the random forest, we also tuned the number of trees and maximum tree depth. For the neural network we used two hidden layers and tuned the number of neurons, the number of times the data is run through the network, the adaptive learning rate, and two shrinkage parameters. For the k-nearest neighbor, we found the optimal value of "k" using cross-validation.

2.6. Predictions

After hyper-parameter tuning, we estimated average daily $PM_{2.5}$ for 3960 grid cells, which covered the greater London area, from 2005 to 2013, using each of the four methods. We then used predictions from the four models in a generalized additive model (GAM), in different combinations, with a spline term for longitude and latitude to allow the predictions to vary spatially. If any of the final predictions were negative, we set the value of $PM_{2.5}$ to zero. This represented less than 0.00014% of predictions.

We used ten-fold cross-validation to check the robustness of our model. We divided the monitoring stations into ten groups. Each model was trained on data from ninety percent of the monitors and predicted in the held-out ten percent. This process was repeated ten times to fully recreate the measured dataset from the portion of the data in which training did not occur. We then looked at the correlation of the predicted $PM_{2.5}$ with the measured $PM_{2.5}$. In order to look at the model's ability to capture spatial variation, we compared annual average predicted $PM_{2.5}$ to the measured annual average $PM_{2.5}$ at monitoring sites, as seen in the equation below:

Annual Measured PM2.5_{ij} =
$$\beta_0 + \beta_1 Annual Predicted PM2.5_{ij}$$
 (1)

where *i* is the monitoring site and *j* is the year. In order to look at the temporal accuracy, we looked at the difference between predicted and measured $PM_{2.5}$ levels and their annual averages, as seen in the equation below:

Daily Measured
$$PM2.5_{ij} - Annual Measured PM2.5_{ij}$$

= $\beta_0 + \beta_1 (Daily Predicted PM2.5_{ij} - Annual Predicted PM2.5_{ij})$ (2)

We chose our final prediction model based on the overall, spatial, and temporal adjusted R² values. To assess the linearity of the relationship between predicted and measured PM_{2.5}, we regressed the final predictions against the measurements for both the spatial and temporal component, using a penalized spline, which chooses the degree of nonlinearity based on the restricted maximum likelihood. The spatial component was modeled using the following equation:

Annual Measured PM2.5_{*ij*} = s(Annual Predicted PM2.5_{*ij*}) +
$$\varepsilon_{ij}$$
 (3)

In this equation, *i* is the monitoring site and *j* is the year, and *s* is a smoothing function. We modeled the temporal component using the equation below:

Daily Measured
$$PM2.5_{ij} - Annual Measured PM2.5_{ij}$$

= $s(Daily Predicted PM2.5_{ij} - Annual Predicted PM2.5_{ij}) + \varepsilon_{ij}$ (4)

In this equation, *i* is the monitoring site and *j* is the day/year of the day, and *s* is a smoothing function.

All data cleaning and processing operations were done in R Statistical Software Version 3.6.1. The machine learning algorithms and predictions, in particular, were run using the "H2O" and "caret" packages [50,51].

3. Results

The results of the ten-fold cross-validation can be seen in Table 1.

 Table 1. Ten-Fold Cross-Validation Results.

Model	Overall R ²	RMSE	Intercept	Slope	Spatial R ²	RMSE	Intercept	Slope	Temporal R ²	RMSE	Intercept	Slope
RF	0.830	4.278	-0.120	0.989	0.386	2.660	-0.827	1.032	0.886	3.297	0.000	0.988
GBM	0.826	4.331	0.081	0.978	0.393	2.644	-0.328	1.003	0.880	3.381	0.000	0.978
NN	0.793	4.728	0.179	0.956	0.266	3.033	4.92	0.671	0.861	3.642	0.000	0.976
KNN	0.791	4.721	0.107	0.965	0.237	2.985	2.356	0.826	0.863	3.623	0.000	0.972
Final Ensemble Model	0.828	4.231	0.058	0.979	0.396	2.637	-0.216	0.996	0.882	3.556	0.000	0.979

Taking all the results into consideration, the final model chosen was the GAM with the RF, GBM, and KNN. The equation for this model can be seen below:

$$PM_{ij} = s(longitude_i, latitude_i) + s(rf_prediction_{ij}) + s(gbm_prediction_{ij}) + s(knn_prediction_{ij}) + \varepsilon_{ij}$$
(5)

In this equation, *i* refers to the grid cell and *j* refers to the day, and *s* is a smoothing function. The NN showed fairly weak results and was subsequently dropped. The RF performed better than the other machine learning methods, overall and temporally. However, the GBM had the strongest spatial predictability. As such, incorporating both benefited the overall model in terms of its predictive abilities across attributes.

The slope and intercept were obtained from regressing the predicted against the measured in the ten held-out cross-validation samples, which recreated the full measured dataset. This both tests for bias in the estimates and represents a form of regression calibration. The final ensemble model had virtually no additive bias (intercept of 0.058), and little multiplicative bias (slope of 0.979 vs 1.0). a biased slope could induce measurement error bias in epidemiological studies.

This model performed well during different seasons. The cross-validated R^2 values were 0.851 for winter, 0.809 for spring, 0.743 for summer, and 0.837 for fall.

The spatial R² was not strong for any of the individual or ensemble averaged models. We modeled the spatial residuals from various years (Figure 2) to see whether a pattern existed which our learners had failed to capture.



Figure 2. Spatial Residuals of Predictions (1 km²) for Each Year in the Greater London Area.

The residuals showed no discernible pattern for us to address. We further tested this idea by calculating Moran's I and seeing whether there was any spatial autocorrelation between the residuals. The results in Table 2 suggest that the dispersion of the residuals is random and not guided by a spatial trend.

Year	Moran's I	<i>p</i> -Value
2005	-0.0132	0.446
2006	-0.0123	0.603
2007	-0.0125	0.953
2008	-0.0118	0.848
2009	-0.0114	0.552
2010	-0.0119	0.380
2011	-0.0116	0.167
2012	-0.0118	0.338
2013	-0.0122	0.500

Table 2. Moran's I and Statistical Test.

We also attempted to model the spatial variability separately from the temporal variability using the GBM machine and found very poor spatial model performance ($R^2 = 0.09-0.12$), indicating that our ensemble approach would indeed be preferable.

We further modeled the population-weighted standard deviation of the annual predicted $PM_{2.5}$ across the Greater London area to see whether annual $PM_{2.5}$ changes in different parts of London were the same across the years or not. Figure 3 shows that while there is not an obvious pattern across the years, annual $PM_{2.5}$ concentrations were slightly more homogeneous at the end of the study period in 2013 than at the beginning in 2005. This implies that daily changes in $PM_{2.5}$ in a year are less dramatic for later years than earlier years, after adjustment for population density.





Figure 3. Population-Weighted Annual Standard Deviations (µg/m³).

The final model was used to generate daily predictions for 3960 1-km by 1-km grid-cells which cover the Greater London Area from 1st January 2005 to 31st December 2013. The mean daily measured PM_{2.5} at monitoring stations was 16.1 μ g/m³ with a standard deviation of 9.2 μ g/m³. Across all 3960 grid-cells, the mean daily predicted PM_{2.5} level was 14.9 μ g/m³, with a standard deviation of 9.0 μ g/m³, indicating that the monitors were located, on average, in more polluted locations. Annually, the measured PM_{2.5} was 16.1 μ g/m³ with a standard deviation of 0.6 μ g/m³, as compared to a predicted level (across all grid-cells) of 14.9 μ g/m³, with a standard deviation of 0.6 μ g/m³. The measured and predicted values at monitoring sites closely resembled one another (Table 3).

	Minimum	25th Percentile	Median	Mean	75th Percentile	Maximum	Standard Deviation
Daily Measured PM _{2.5}	2.9	10.0	13.1	16.0	19.0	77.5	9.2
Daily Predicted PM _{2.5} (Monitoring Sites)	2.9	10.0	13.2	16.0	19.0	77.4	9.2
Daily Predicted PM _{2.5} (Grid-cells)	2.8	9.0	12.2	14.9	17.9	74.4	9.0
Annual Measured PM _{2.5}	15.3	15.6	16.1	16.1	16.2	17.1	0.6
Annual Predicted PM _{2.5} (Monitoring Sites)	15.3	15.6	16.1	16.1	16.2	17.0	0.6
Annual Predicted PM _{2.5} (Grid-cells)	14.1	14.5	14.7	14.9	15.2	15.9	0.6

Table 3. Distribution of Daily and Annual PM_{2.5} (µg/m³).

Annual predictions for the Greater London area can be seen in Figure 4. The center of London is consistently the location with highest levels of fine particulate matter. The southern and western regions of the area seem to have the lowest levels of pollution. Over the years, a general decrease can be seen in the levels of $PM_{2.5}$. By 2013, the average concentration in London had fallen to 16.1 µg/m³, from a peak of 17.1 µg/m³ in 2011 and an initial concentration of 16.9 µg/m³ in 2005. This is most likely attributable to regulatory policies and technological improvements.



Figure 4. Annual Predictions in the Greater London Area (μ g/m³).

Of the variables used as predictors, six of the ten most important were the same for the RF and the GBM (the KNN did not provide measures of variable importance). These six were average daily city-wide $PM_{2.5}$, height of the planetary boundary layer, average wind speed, wind direction, distance to Heathrow airport, and light at night (Table 4). Of these, the average daily city-wide $PM_{2.5}$ level was by far the most informative predictor. We also looked at SHAP (shapely additive explanations) values. The SHAP values measure the contribution of each predictor to the prediction for each observation. They can be aggregated to obtain a global value of variable importance, as seen in Figure 5 [52]. Of the variables we considered, the average daily city-wide $PM_{2.5}$ level was by far the

most informative predictor. This would be expected given the correlation values seen in Figure 6 with average daily city-wide particulate levels having a correlation coefficient of 0.9 with measured $PM_{2.5}$.

Random I	Forest	Gradient Boosting Machine			
Variable	Relative Contribution (%)	Variable	Relative Contribution (%)		
Average city-wide daily PM _{2.5}	58.31	Average city-wide daily PM _{2.5}	66.75		
Height of the planetary boundary layer (inverse)	10.28	Average wind speed	6.36		
Average wind speed	6.90	Wind direction (categorical)	5.00		
Wind direction (categorical)	4.65	Height of the planetary boundary layer (inverse)	2.69		
AOD (from aqua satellite)	1.17	Time (days from January 1, 2005)	1.26		
Average barometric pressure	1.10	Distance to Heathrow airport	1.22		
Distance to Heathrow airport	0.99	Population density	1.10		
Longitude	0.96	Light at night	1.05		
Light at night	0.94	Average building height in grid cell	1.00		
Average temperature	0.93	Number of buildings in grid cell	0.98		

 Table 4. Variable Importance for Machine Learning Algorithms.



Figure 5. GBM variable importance based on SHAP values.



Figure 6. Correlation Matrix of Continuous Predictors and PM_{2.5}.

When modeling the spatial and temporal predicted vs. measured $PM_{2.5}$ using a smoothing term, we found a virtually perfect linear relationship for both components (Figure 7).



Figure 7. (a) Spatial Measured vs. Smooth Predicted Annual $PM_{2.5}$ (µg/m³). (b) Temporal Measured vs. Smooth Predicted Annual $PM_{2.5}$ (µg/m³) (all values are on a 1 km² scale).

4. Discussion

Our final model incorporated the GBM, RF, and KNN in a GAM, with a smoothing term for longitude and latitude. This model had a very strong performance in terms of the cross-validated overall R^2 and the cross-validated temporal R^2 . The spatial R^2 was not as robust. However, the linear regression looking at measured $PM_{2.5}$ versus predicted $PM_{2.5}$ showed very little bias in the spatial model (i.e., the intercept was approximately zero and the slope was approximately one). The distribution of

the spatial residuals did not demonstrate any particular pattern (Figure 2)—a conclusion supported by Moran's I values (Table 2)—nor did modeling the spatial variability separately improve the spatial R^2 . We hypothesize that the reason behind the lackluster spatial R^2 was that the Greater London area did not have very much spatial variation on this scale to begin with; that is to say that most of the change in PM_{2.5} levels was due to temporal factors. This can be seen by looking at the standard deviation of the measured PM_{2.5}. The standard deviation of the daily measured PM_{2.5} is 9.2 µg/m³, while the standard deviation of the annual averages across the monitoring sites is 0.6 µg/m³, showing that most of the variation in the area can be attributed to day-to-day variation in pollution levels rather than spatial differences. If the predictions were done on a finer scale, the R² for the spatial variation would likely be higher due to greater variability on a smaller scale.

In order to check for bias across the range of particulate matter levels, we modeled the spatial component and temporal component of the measured PM against the predicted one using a smoothing parameter. This would have allowed us to see any non-linearity that might have existed between the predicted and measured at any point. As can be seen in Figure 7, there is a virtually perfect linear relationship between the predicted and measured PM_{2.5} levels at the monitoring sites for both the spatial and temporal component of our predictions. It demonstrates that our model performs well across the range of pollution levels.

Our model joins a number of other PM_{2.5} models that have been created for London, though it covers a more comprehensive time frame of nine years. a 2016 study by researchers at King's College London modeled hourly PM_{2.5} on a 20 m × 20 m scale in 2011. This model was called the Community Multiscale Air Quality (CMAQ)-urban model and it was generated using a combination of dispersion and chemistry models, and meteorological models with a very fine emissions inventory. It had an R² coefficient of 0.59 [26]. a newer hybrid daily model, which incorporated CMAQ-urban, a LUR model, as well as machine learning methods for 2009–2013 at the lower layer output super area (LSOA) level, had a spatial R² 0.22 and a temporal R² of 0.93 for background monitors and a spatial R² of 0.29 and a temporal R² of 0.95 at roadside monitors [53]. a 2014 study by Singh et al., which modeled hourly PM_{2.5} for the year 2008 on an adjustable 10 m to 100 m scale using a dispersion model, had an overall R² of 0.64 when comparing the predicted time-series to the measured time-series. Furthermore, this model had a spatial R² of 0.27, indicating limited spatial predictability [54]. a 2012 land use regression model created for the European Study of Cohorts of Air Pollution Effects (ESCAPE) using measured pollution data, collected three time over 14 days in three seasons and averaged for an annual estimate, found a spatial leave-one-out cross-validated R² of 0.77 in the London/Oxford area [55].

The ensemble model we created compares well to other prediction models for $PM_{2.5}$ using machine learning methods. An ensemble model set to predict daily $PM_{2.5}$ in China from 2013 to 2016 using clustering and an ensemble approach of a an extreme gradient boosting machine, RF, and GAM, found an overall R² of 0.79 [56]. Also in China, a geographically-weighted GBM daily prediction approach across the country for the year 2014 had a cross-validated R² of 0.76 [57]. Another study in China modeling daily $PM_{2.5}$ from 2005 to 2016 and utilizing a random forest approach found a cross-validated R² of 0.83. This machine learning approach outperformed two other non-linear distributed lag models [58]. a study predicting monthly $PM_{2.5}$ in British Columbia based on remote sensing data, and in particular AOD, using eight different approaches found that the random forest was the strongest machine learner as compared to other algorithms, such as extreme gradient boosting and Bayesian regularized neural networks. All machine learners in this study outperformed the multiple linear regression [59].

There are other studies which did not find the same patterns as we did. a study looking specifically at modeling hourly $PM_{2.5}$ in Beijing and Shanghai using multiple machine learning methods found a variant of the artificial neural network to have the highest R^2 , with a value of 0.92, which outperformed the random forest with a value of 0.88 [60], while our study found the neural network to be one of the weakest learners and the random forest to be the strongest. a study predicting average daily $PM_{2.5}$ in the contiguous United States on a 1 km² scale using an ensemble approach and machine

learning methods also showed the neural network to be the strongest learner [31]. This might be due to differences in the input variables, as well as the structure of the algorithm and how it processes the input parameters from different regions. The substantial difference in spatial scale (US vs. London) may also favor different algorithms.

Our study had several limitations. Firstly, we had a limited number of observations and fixed PM₂₅ station monitors. We expanded this number using a model to predict $PM_{2.5}$ from PM_{10} measurements and correcting tapered element oscillating microbalance (TEOM) monitor measurements to gravimetric equivalents for consistency. We then used this dataset to calibrate our predictions. These monitors were located in a range of site types: curbside, roadside and urban background. As our spatial predictions were at 1-km² resolution, validation against monitors with strong local sources would produce an under-prediction. Secondly, as with any machine learning method, our predictions may be subject to overfitting. We accounted for this using out-of-bag ten-fold cross-validation at monitoring sites. Thirdly, we were limited in the number of input variables by data availability. It is possible that the inclusion of an overlooked variable may have improved the spatial and overall predictability. Finally, as with all analyses, our results rely on the quality of our predictors. This is particularly true of measurements made using remote sensing technology such as AOD. Many PM_{2.5} prediction models rely on AOD as the main predictor of PM_{2.5} concentrations. Yet, there is error associated with satellite measurements of AOD as well [61,62]. However, given that we are predicting a continuous outcome, this measurement error should not cause bias and it is accounted for in the residuals of the prediction model.

Despite these limitations, our prediction model also possessed several advantages. It relied on the power of three different machine learning algorithms to generate predictions. As such, it is likely that one method may have captured some of the variation in daily $PM_{2.5}$ that the other two had missed. Moreover, the overall out-of-bag ten-fold cross-validated R^2 of 0.828 suggests a strong ability to predict fine particulate matter and the temporal R^2 suggests that we are capturing the day-to-day trends very well. Furthermore, even though the spatial R^2 is not very strong, the model comparing predicted and measured $PM_{2.5}$ had an intercept close to zero and an intercept close to one. This suggests minimal bias in capturing the spatial variation. Finally, our model predicted particulate matter in 1-km² grid cells which allows for exposure assignment on a fine scale in other studies that may utilize the model.

5. Conclusions

Our $PM_{2.5}$ model, which incorporated several different machine learning algorithms, has shown to be a robust and accurate measure of pollution levels in the Greater London area. It had very strong performance metrics, overall and temporally. The spatial R^2 was fairly weak, however, the model showed very little bias. Future models of Greater London may need to be done on a smaller spatial scale or with additional predictor values in order to capture greater variability. These exposure measurement models can be used study the effect of $PM_{2.5}$ on health outcomes in epidemiological studies, to identify locations with peak concentrations, and to conduct risk assessments.

Author Contributions: The authors' contributions are as follows: conceptualization, M.D.Y., Z.K., K.K. and J.S.; methodology, M.D.Y. and J.S.; validation, M.D.Y.; formal analysis, M.D.Y., J.S.; data curation, M.D.Y., Z.K., K.D., B.B., E.S., H.A., A.L.; writing—original draft preparation, M.D.Y.; writing—review and editing, all authors; supervision, J.S.; project administration, K.K. and J.S.; and funding acquisition, K.K. All authors have read and agreed to the published version of the manuscript.

Funding: Research in this paper was part of the STEAM project, funded under the MRC UK Grant ref.: MR/N014464/1.

Acknowledgments: The authors would like to thank John Evans, Petros Koutrakis, and Brent Coull for their input and feedback on this project.

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

References

- Sousan, S.; Koehler, K.; Hallett, L.; Peters, T.M. Evaluation of consumer monitors to measure particulate matter. J. Aerosol Sci. 2017, 107, 123–133. [CrossRef] [PubMed]
- 2. Xing, Y.-F.; Xu, Y.-H.; Shi, M.-H.; Lian, Y.-X. The impact of PM_{2.5} on the human respiratory system. *J. Thorac. Dis.* **2016**, *8*, E69–E74. [PubMed]
- Dockery, D.W.; Pope, C.A.; Xu, X.; Spengler, J.D.; Ware, J.H.; Fay, M.E.; Ferris, B.G.; Speizer, F.E. An Association between Air Pollution and Mortality in Six U.S. Cities. *N. Engl. J. Med.* 1993, 329, 1753–1759. [CrossRef] [PubMed]
- Pope, C.A.; Thun, M.J.; Namboodiri, M.M.; Dockery, D.W.; Evans, J.S.; Speizer, F.E.; Heath, C.W.J. Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. *Am. J. Respir. Crit. Care Med.* 1995, 151, 669–674. [CrossRef]
- 5. Wang, Y.; Shi, L.; Lee, M.; Liu, P.; Di, Q.; Zanobetti, A.; Schwartz, J.D. Long-term Exposure to PM_{2.5} and Mortality Among Older Adults in the Southeastern US. *Epidemiology* **2017**, *28*, 207–214. [CrossRef]
- 6. Di, Q.; Wang, Y.; Zanobetti, A.; Wang, Y.; Koutrakis, P.; Choirat, C.; Dominici, F.; Schwartz, J.D. Air Pollution and Mortality in the Medicare Population. *N. Engl. J. Med.* **2017**, *376*, 2513–2522. [CrossRef]
- Vodonos, A.; Awad, Y.A.; Schwartz, J. The concentration-response between long-term PM_{2.5} exposure and mortality; a meta-regression approach. *Environ. Res.* 2018, *166*, 677–689. [CrossRef]
- 8. Atkinson, R.W.; Kang, S.; Anderson, H.R.; Mills, I.C.; Walton, H.A. Epidemiological time series studies of PM_{2.5} and daily mortality and hospital admissions: a systematic review and meta-analysis. *Thorax* **2014**, *69*, 660–665. [CrossRef]
- 9. Amini, H.; Trang Nhung, N.T.; Schindler, C.; Yunesian, M.; Hosseini, V.; Shamsipour, M.; Hassanvand, M.S.; Mohammadi, Y.; Farzadfar, F.; Vicedo-Cabrera, A.M.; et al. Short-term associations between daily mortality and ambient particulate matter, nitrogen dioxide, and the air quality index in a Middle Eastern megacity. *Environ. Pollut.* **2019**, *254*, 113121. [CrossRef]
- Danesh Yazdi, M.; Wang, Y.; Di, Q.; Zanobetti, A.; Schwartz, J. Long-term exposure to PM_{2.5} and ozone and hospital admissions of Medicare participants in the Southeast USA. *Environ. Int.* 2019, 130, 104879. [CrossRef]
- Barnett, A.G.; Williams, G.M.; Schwartz, J.; Best, T.L.; Neller, A.H.; Petroeschevsky, A.L.; Simpson, R.W. The effects of air pollution on hospitalizations for cardiovascular disease in elderly people in Australian and New Zealand cities. *Environ. Health Perspect.* 2006, 114, 1018–1023. [CrossRef] [PubMed]
- Pun, V.C.; Kazemiparkouhi, F.; Manjourides, J.; Suh, H.H. Long-Term PM_{2.5} Exposure and Respiratory, Cancer, and Cardiovascular Mortality in Older US Adults. *Am. J. Epidemiol.* 2017, *186*, 961–969. [CrossRef] [PubMed]
- 13. Leiva, G.M.A.; Santibañez, D.A.; Ibarra, E.S.; Matus, C.P.; Seguel, R. a five-year study of particulate matter (PM_{2.5}) and cerebrovascular diseases. *Environ. Pollut.* **2013**, *181*, 1–6. [CrossRef]
- 14. Kioumourtzoglou, M.A.; Schwartz, J.D.; Weisskopf, M.G.; Melly, S.J.; Wang, Y.; Dominici, F.; Zanobetti, A. Long-term PM_{2.5} exposure and neurological hospital admissions in the northeastern United States. *Environ. Health Perspect.* **2016**, *124*, 23–29. [CrossRef]
- 15. Fu, P.; Guo, X.; Cheung, F.M.H.; Yung, K.K.L. The association between PM_{2.5} exposure and neurological disorders: a systematic review and meta-analysis. *Sci. Total Environ.* **2019**, *655*, 1240–1248. [CrossRef]
- Shi, L.; Zanobetti, A.; Kloog, I.; Coull, B.A.; Koutrakis, P.; Melly, S.J.; Schwartz, J.D. Low-concentration PM_{2.5} and mortality: Estimating acute and chronic effects in a population-based study. *Environ. Health Perspect.* 2016, 124, 46–52. [CrossRef] [PubMed]
- Shaddick, G.; Thomas, M.L.; Amini, H.; Broday, D.; Cohen, A.; Frostad, J.; Green, A.; Gumy, S.; Liu, Y.; Martin, R.V.; et al. Data Integration for the Assessment of Population Exposure to Ambient Air Pollution for Global Burden of Disease Assessment. *Environ. Sci. Technol.* 2018, *52*, 9069–9078. [CrossRef]
- Wang, J.; Christopher, S.A. Intercomparison between satellite-derived aerosol optical thickness and PM_{2.5} mass: Implications for air quality studies. *Geophys. Res. Lett.* 2003, *30*, 2–5. [CrossRef]
- Liu, Y.; Park, R.J.; Jacob, D.J.; Li, Q.; Kilaru, V.; Sarnat, J.A. Mapping annual mean ground-level PM_{2.5} concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States. *J. Geophys. Res. Atmos.* 2004, 109, 1–10.
- Van Donkelaar, A.; Martin, R.V.; Park, R.J. Estimating ground-level PM_{2.5} using aerosol optical depth determined from satellite remote sensing. *J. Geophys. Res. Atmos.* 2006, 111, 1–10. [CrossRef]

- 21. Van Donkelaar, A.; Martin, R.V.; Brauer, M.; Kahn, R.; Levy, R.; Verduzco, C.; Villeneuve, P.J. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: Development and application. *Environ. Health Perspect.* **2010**, *118*, 847–855. [CrossRef] [PubMed]
- 22. Gupta, P.; Christopher, S.A.; Wang, J.; Gehrig, R.; Lee, Y.; Kumar, N. Satellite remote sensing of particulate matter and air quality assessment over global cities. *Atmos. Environ.* **2006**, *40*, 5880–5892. [CrossRef]
- Kloog, I.; Koutrakis, P.; Coull, B.A.; Lee, H.J.; Schwartz, J. Assessing temporally and spatially resolved PM_{2.5} exposures for epidemiological studies using satellite aerosol optical depth measurements. *Atmos. Environ.* 2011, 45, 6267–6275. [CrossRef]
- 24. Kloog, I.; Nordio, F.; Coull, B.A.; Schwartz, J. Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal PM_{2.5} exposures in the mid-atlantic states. *Environ. Sci. Technol.* **2012**, *46*, 11913–11921. [CrossRef]
- 25. Moore, D.K.; Jerrett, M.; Mack, W.J.; Künzli, N. a land use regression model for predicting ambient fine particulate matter across Los Angeles, CA. *J. Environ. Monit.* **2007**, *9*, 246–252. [CrossRef]
- Smith, J.D.; Mitsakou, C.; Kitwiroon, N.; Barratt, B.M.; Walton, H.A.; Taylor, J.G.; Anderson, H.R.; Kelly, F.J.; Beevers, S.D. London Hybrid Exposure Model: Improving Human Exposure Estimates to NO₂ and PM_{2.5} in an Urban Setting. *Environ. Sci. Technol.* 2016, *50*, 11760–11768. [CrossRef]
- Geng, G.; Zhang, Q.; Martin, R.V.; van Donkelaar, A.; Huo, H.; Che, H.; Lin, J.; He, K. Estimating long-term PM_{2.5} concentrations in China using satellite-based aerosol optical depth and a chemical transport model. *Remote Sens. Environ.* 2015, 166, 262–270. [CrossRef]
- Di, Q.; Kloog, I.; Koutrakis, P.; Lyapustin, A.; Wang, Y.; Schwartz, J. Assessing PM_{2.5} Exposures with High Spatiotemporal Resolution across the Continental United States. *Environ. Sci. Technol.* 2016, 50, 4712–4721. [CrossRef]
- De Hoogh, K.; Gulliver, J.; van Donkelaar, A.; Martin, R.V.; Marshall, J.D.; Bechle, M.J.; Cesaroni, G.; Pradas, M.C.; Dedele, A.; Eeftens, M.; et al. Development of West-European PM_{2.5} and NO₂ land use regression models incorporating satellite-derived and chemical transport modelling data. *Environ. Res.* 2016, 151, 1–10. [CrossRef]
- Taghavi-Shahri, S.M.; Fassò, A.; Mahaki, B.; Amini, H. Concurrent spatiotemporal daily land use regression modeling and missing data imputation of fine particulate matter using distributed space-time Expectation Maximization. *Atmos. Environ.* 2019, 117202. [CrossRef]
- 31. Di, Q.; Amini, H.; Shi, L.; Kloog, I.; Silvern, R.; Kelly, J.; Sabath, M.B.; Choirat, C.; Koutrakis, P.; Lyapustin, A.; et al. An ensemble-based model of PM_{2.5} concentration across the contiguous United States with high spatiotemporal resolution. *Environ. Int.* **2019**, *130*, 104909. [CrossRef] [PubMed]
- 32. Di, Q.; Rowland, S.; Koutrakis, P.; Schwartz, J. a hybrid model for spatially and temporally resolved ozone exposures in the continental United States a hybrid model for spatially and temporally resolved ozone exposures in the continental a hybrid model. *J. Air Waste Manag. Assoc.* **2017**, *67*, 39–52. [CrossRef] [PubMed]
- Lary, D.J.; Lary, T.; Sattler, B. Using Machine Learning to Estimate Global PM_{2.5} for Environmental Health Studies. *Environ. Health Insights* 2015, 9, 41–52. [CrossRef] [PubMed]
- 34. Weizhen, H.; Zhengqiang, L.; Yuhuan, Z.; Hua, X.; Ying, Z.; Kaitao, L.; Donghui, L.; Peng, W.; Yan, M. Using support vector regression to predict PM₁₀ and PM_{2.5}. *IOP Conf. Ser. Earth Environ. Sci.* **2014**, *17*, 012268. [CrossRef]
- 35. Wei, J.; Huang, W.; Li, Z.; Xue, W.; Peng, Y.; Sun, L.; Cribb, M. Estimating 1-km-resolution PM_{2.5} concentrations across China using the space-time random forest approach. *Remote Sens. Environ.* **2019**, 231, 111221. [CrossRef]
- Di, Q.; Amini, H.; Shi, L.; Kloog, I.; Silvern, R.F.; Kelly, J.T.; Sabath, M.B.; Choirat, C.; Koutrakis, P.; Lyapustin, A.; et al. Assessing NO₂ Concentration and Model Uncertainty with High Spatiotemporal Resolution across the Contiguous United States Using Ensemble Model Averaging. *Environ. Sci. Technol.* 2020, 54, 1372–1384. [CrossRef]
- Wang, Y.; Lee, M.; Liu, P.; Shi, L.; Yu, Z.; Awad, Y.A.; Zanobetti, A.; Schwartz, J.D. Doubly Robust Additive Hazards Models to Estimate Effects of a Continuous Exposure on Survival. *Epidemiology* 2017, 28, 771–779. [CrossRef]
- Chen, G.; Jin, Z.; Li, S.; Jin, X.; Tong, S.; Liu, S.; Yang, Y.; Huang, H.; Guo, Y. Early life exposure to particulate matter air pollution (PM₁, PM_{2.5} and PM₁₀) and autism in Shanghai, China: a case-control study. *Environ. Int.* 2018, 121, 1121–1127. [CrossRef]

- Qiu, X.; Wei, Y.; Wang, Y.; Di, Q.; Sofer, T.; Awad, Y.A.; Schwartz, J. Inverse probability weighted distributed lag effects of short-term exposure to PM_{2.5} and ozone on CVD hospitalizations in New England Medicare participants—Exploring the causal effects. *Environ. Res.* 2020, *182*, 109095. [CrossRef]
- 40. Van Der Laan, M.J.; Polley, E.C.; Hubbard, A.E. Super learner. Stat. Appl. Genet. Mol. Biol. 2007, 6. [CrossRef]
- 41. Friedman, J.H. Greedy function approximation: a gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
- 42. Breiman, L. Random forests. Mach. Learn. 2001, 45, 5–32. [CrossRef]
- Glorot, X.; Bordes, A.; Bengio, Y. Deep Sparse Rectifier Neural Networks. In Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, Ft. Lauderdale, FL, USA, 11–13 April 2011; pp. 315–323.
- 44. Altman, N.S. An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. Am. Stat. 1992, 46, 175–185.
- 45. Lyapustin, A.; Wang, Y.; Korkin, S.; Huang, D. MODIS Collection 6 MAIAC algorithm. *Atmos. Meas. Tech.* **2018**, *11*, 5741–5765. [CrossRef]
- 46. Center for International Earth Science Information Network—CIESIN—Columbia University Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11; NASA Socioeconomic Data and Applications Center (SEDAC): Palisades, NY, USA, 2018. [CrossRef]
- 47. Environmental Research Groupt at King's College London London Air. Available online: http://londonair. org.uk/LondonAir/Default.aspx (accessed on 15 December 2019).
- 48. Department of Environment Food & Rural Affairs UK Automatic Urban and Rural Network. Available online: https://uk-air.defra.gov.uk/ (accessed on 15 December 2019).
- 49. Analitis, A.; Barratt, B.M.; Green, D.; Beddows, A.; Samoli, E.; Schwartz, J.D.; Katsouyanni, K. Enhancement of the PM_{2.5} Database 2004–2013 for London within the STEAM Project Using Generalized Additive Models and Machine Learning Methods. Unpublished work. 2020; 1–20.
- 50. LeDell, E.; Gill, N.; Aiello, S.; Fu, A.; Candel, A.; Click, C.; Kraljevic, T.; Nykodym, T.; Aboyoun, P.; Kurka, M.; et al. h2o: R Interface for "H2O". Available online: https://github.com/h2oai/h2o-3 (accessed on 1 March 2020).
- 51. Kuhn, M. Building predictive models in R using the caret package. J. Stat. Softw. 2008, 28, 1–26. [CrossRef]
- 52. Lundberg, S.M.; Lee, S.I. a unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **2017**, 2017, 4766–4775.
- 53. Samoli, E.; Butland, B.; Rodopoulou, S.; Atkinson, R.W.; Barratt, B.M.; Beevers, S.D.; Dimakopoulou, K.; Danesh Yazdi, M.; Schwartz, J.D.; Katsouyanni, K. The Impact of Measurement Error in Modelled Ambient Particles Exposures on Health Effect Estimates in Multi-level Analysis: a Simulation Study. Unpublished work. 2020.
- 54. Singh, V.; Sokhi, R.S.; Kukkonen, J. PM_{2.5} concentrations in London for 2008-A modeling analysis of contributions from road traffic. *J. Air Waste Manag. Assoc.* **2014**, *64*, 509–518. [CrossRef]
- 55. Eeftens, M.; Beelen, R.; De Hoogh, K.; Bellander, T.; Cesaroni, G.; Cirach, M.; Declercq, C.; Dedele, A.; Dons, E.; De Nazelle, A.; et al. Development of land use regression models for PM_{2.5}, PM_{2.5} absorbance, PM₁₀ and PM_{coarse} in 20 European study areas; Results of the ESCAPE project. *Environ. Sci. Technol.* **2012**, *46*, 11195–11205. [CrossRef]
- Xiao, Q.; Chang, H.H.; Geng, G.; Liu, Y. An Ensemble Machine-Learning Model to Predict Historical PM_{2.5} Concentrations in China from Satellite Data. *Environ. Sci. Technol.* 2018, 52, 13260–13269. [CrossRef] [PubMed]
- Zhan, Y.; Luo, Y.; Deng, X.; Chen, H.; Grieneisen, M.L.; Shen, X.; Zhu, L.; Zhang, M. Spatiotemporal prediction of continuous daily PM_{2.5} concentrations across China using a spatially explicit machine learning algorithm. *Atmos. Environ.* 2017, 155, 129–139. [CrossRef]
- 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]
- Xu, Y.; Ho, H.C.; Wong, M.S.; Deng, C.; Shi, Y.; Chan, T.C.; Knudby, A. Evaluation of machine learning techniques with multiple remote sensing datasets in estimating monthly concentrations of ground-level PM_{2.5}. *Environ. Pollut.* 2018, 242, 1417–1426. [CrossRef]
- 60. Huang, C.J.; Kuo, P.H. a deep cnn-lstm model for particulate matter (PM_{2.5}) forecasting in smart cities. *Sensors* **2018**, *18*, 2220. [CrossRef] [PubMed]

- 61. Just, A.C.; De Carli, M.M.; Shtein, A.; Dorman, M.; Lyapustin, A.; Kloog, I. Correcting measurement error in satellite aerosol optical depth with machine learning for modeling PM_{2.5} in the Northeastern USA. *Remote Sens.* **2018**, *10*, 803. [CrossRef]
- 62. Li, J.; Carlson, B.E.; Lacis, A.A. How well do satellite AOD observations represent the spatial and temporal variability of PM_{2.5} concentration for the United States? *Atmos. Environ.* **2015**, *102*, 260–273. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).